National Association of Broadcasters

ENGINEERING HANDBOOK

EIGHTH EDITION

NAB™
BROADCASTERS

NAB™
BROADCASTERS

ISBN 0-89324-135-0

National Association of Broadcasters

ENGINEERING HANDBOOK

EIGHTH EDITION

National Association of
NAB
BROADCASTERS

ISBN 0-89324-135-0

For more information on other publications available through the National Association of Broadcasters, call (800) 368-5644, or in the D.C. area (202) 429-5373.

FOR THE BROADCAST ENGINEER

This 8th edition of the *NAB Engineering Handbook* records the march of technology and regulatory changes affecting broadcast engineering since the 7th edition was published in 1985. We hope the 8th edition contributes to every engineer's ability and understanding of these changes.

The broadcasting industry is in the midst of an increasingly challenging economic environment. Under these circumstances the industry critically depends on its engineers to maintain high standards of quality and efficiency in the face of difficult economics and ever-changing technologies. One can only admire the individuals who meet these challenges and, every day, work to keep the broadcasting industry on the air.

Edward O. Fritts
President and Chief Executive Officer
National Association of Broadcasters

March 1992

ACKNOWLEDGEMENTS

The *NAB Engineering Handbook* was produced by NAB Science and Technology in cooperation with NAB Services. We gratefully acknowledge the assistance of those who contributed to the completion of this edition.

We also thank the members of NAB's committees, particularly NAB's Engineering Advisory Committee, along with countless industry experts who reviewed manuscripts and graciously offered their technical guidance and time in the preparation of this book.

NAB Science and Technology

Michael C. Rau,
Senior Vice President
Andy Butler
Lynn Claudy
Janet Elliott
Jane Pessall Frock
John Marino
Ken Springer
Kelly Williams

NAB Services

Teri Lepovitz, Senior Vice President
Nancy Tarr
Anne Zaleski-Byers

F. David Harris
E. Benjamin Crutchfield

FOREWORD

So much has changed since the 7th edition was published in 1985. New technology, revised FCC regulations, better understanding of basic broadcast technologies, are all reflected within this new 8th edition *Handbook*. It was not an easy task preparing this edition! NAB is indebted to all the talented individuals who have worked so hard to bring the 8th edition to your hands. I am proud of the thoroughness and attention to detail that is apparent in any chapter of this book.

The *NAB Engineering Handbook* is a reference book on broadcast engineering skills and practices. Its purpose is to collect, in one place, the information needed by broadcast engineers to operate and maintain state-of-the-art radio and TV stations. There are precious few technical standards and even fewer references to help develop the skills needed by a successful broadcast engineer; so far as we know, the *NAB Engineering Handbook* is the only such reference that comprehensively covers all aspects of radio and television broadcast engineering.

We have chosen not to include materials on HDTV or Digital Audio Broadcasting (DAB). These emerging broadcast technologies have not matured to the point where we can publish an authoritative broadcast engineering guide. We're all still learning about these technologies and how they will affect broadcast operations. NAB offers separate publications that are updated annually on these important new technologies.

NAB Science & Technology welcomes your comments and suggestions for improvement on any aspect of our operations. In particular, suggestions for future editions of the *NAB Engineering Handbook* are especially appreciated.

Michael C. Rau
Senior Vice President, Science & Technology
National Association of Broadcasters

March 1992

TABLE OF CONTENTS

SECTION 1. PROCEDURES AND PRACTICES

1.1	FCC Organization and Administrative Process	1
1.2	FCC Field Operations Bureau	11
1.3	Frequency Coordination	23
1.4	Frequency Allocations for Broadcasting and the Broadcast Auxiliary Services	33
1.5	Environmental Concerns	49
1.6	Consultant Services: When and How to Use Them	59
1.7	Source Guide: Broadcast Standards and Information	63

SECTION 2. ANTENNAS AND TOWERS

2.1	Design, Erection, and Maintenance of Antenna Structures	79
2.2	Lightning Protection for Broadcast Facilities	101
2.3	Transmission Lines	115
2.4	Diplexers, Combiners, and Filters	161
2.5	AM Broadcast Antenna Systems	
	Part I: System Design	187
	Part II: Antenna Coupling and Phasing Systems for AM Broadcasting	215
	Part III: Maintenance	229
2.6	FM Broadcast Antennas	237
2.7	Antennas For Television Broadcasting	273
2.8	Radio Wave Propagation	297
2.9	Human Exposure to RF Radiation	311
2.10	AM Field Strength Measurements and Proof-of-Performance	327
2.11	FM and TV Field Strength Measurements	335

SECTION 3. TRANSMITTERS

3.1	AM Transmitters	353
3.2	AM Stereo Broadcasting	405
3.3	FM Broadcast Transmitters	429
3.4	Subcarrier Transmissions and Stereophonic Broadcasting	475
3.5	TV Transmitters	499
3.6	Multichannel Television Sound	553
3.7	Transmitter Control Systems	569
3.8	Frequency and Aural Modulation Monitoring	609

SECTION 4. PROGRAM TRANSMISSION FACILITIES

4.1	TV Signal Transmission Standards	617

4.2 Microwave Engineering for the Broadcaster 653
4.3 Aural Broadcast Auxiliary Links 671
4.4 Satellite Earth Stations 679
4.5 Fiber Optic Transmission Systems 717
4.6 An Overview of Cable Television 731

SECTION 5. PRODUCTION FACILITIES

5.1 Microphones 749
5.2 The Broadcast Television Camera System 775
5.3 Video Signal Switching, Timing, and Distribution 807
5.4 Magnetic Recording Media 825
5.5 Principles and Current State of the Art
Part I: Analog Magnetic Recording 845
Part II: Digital Audio Technology 863
5.6 Videotape Recording 877
5.7 Videotape Editing 913
5.8 Film for Television 933
5.9 Studio Audio 947
5.10 Transmission Audio Processing 967
5.11 Television Camera Robotics 981
5.12 Planning of Broadcast Facilities 989
5.13 Principles of Acoustics for Broadcast Applications 1013
5.14 Lighting For Television 1033

SECTION 6. REMOTE PROGRAM ORIGINATION

6.1 Radio Electronic News Gathering and Field Production 1051
6.2 TV Field Production: An Overview 1073
6.3 Telephone Network Interfacing 1083
6.4 Telco Audio Program Service 1111

SECTION 7. SPECIAL SYSTEMS

7.1 Broadcast Audio Measurement 1121
7.2 Radio Receivers 1139
7.3 Color Television
Part I: The Basic NTSC System 1149
Part II: Worldwide Color Television Standards 1193
7.4 Color Television Receivers 1209
7.5 Closed Captioning Systems 1237
7.6 Weather Radar 1251
7.7 Distance and Bearing Calculations 1259
7.8 The Emergency Broadcast System 1265
7.9 AC Power Systems 1281
7.10 Low Power TV Transmission Systems 1297

Index 1317
Author Listing 1341

1.1 FCC Organization and Administration Process

Staff
Federal Communications Commission, Washington, District of Columbia

STATUTORY AUTHORITY

Congress, through adoption of the Communications Act of 1934 (the Act), created the Federal Communications Commission (FCC) as an independent regulatory agency. Section 1 of the Act specifies that the FCC was created:

> "For the purpose of regulation of interstate and foreign commerce in communication by wire and radio so as to make available, so far as possible, to all the people of the United States a rapid, efficient, nationwide, and worldwide wire and radio communication service with adequate facilities at reasonable charges, for the purpose of the national defense, for the purpose of promoting safety of life and property through the use of wire and radio communication, and for the purpose of securing a more effective execution of this policy by centralizing authority heretofore granted by law to several agencies and by granting additional authority with respect to interstate and foreign commerce in wire and radio communication. . ."

THE "COMMISSION"

The "Commission" consists of five Commissioners appointed by the President, by and with the advice and consent of the Senate. The President designates one Commissioner as Chairman. The Commissioners make their decisions collectively by formal vote although authority to act on routine matters is normally delegated to the staff.

FCC ORGANIZATION

The staff of the FCC performs day-to-day functions of the agency, including: (1) license and application processing, (2) drafting of rulemaking items, (3) enforcing rules and regulations, and (4) formulating policy. The staff is divided along functional lines into various offices and bureaus. Normally, broadcasters deal with the Mass Media Bureau (MMB) or the Field Operations Bureau (FOB); however, actions by other elements of the agency may directly affect broadcasters. The major organizational units are shown in Fig. 1.

Office of the Managing Director (OMD)

The Managing Director serves as the chief operations and executive official, as supervised and directed by the Chairman. OMD recommends, to the Chairman, program priorities, resources and position allocations, management and administrative policies. OMD operates the agency's personnel office and has responsibility for emergency communications policies.

Office of Public Affairs (OPA)

OPA functions as the FCC's primary point of contact with the public in dissemination of information about the Commission.

Office of Legislative Affairs (OLA)

The Office of Legislative Affairs is the primary point of contact with the Congress in dissemination of information about the Commission. Among its functions, the Office facilitates responses to congressional inquiries, drafts Commission legislative proposals and bill comments, and helps the Commission prepare for congressional hearings.

Office of Plans and Policy (OPP)

This Office makes recommendations to the Commission on development and implementation of communications policies.

Office of General Counsel (OGC)

The General Counsel advises the Commission in all matters of law and litigation, and interprets and implements statutes and treaties.

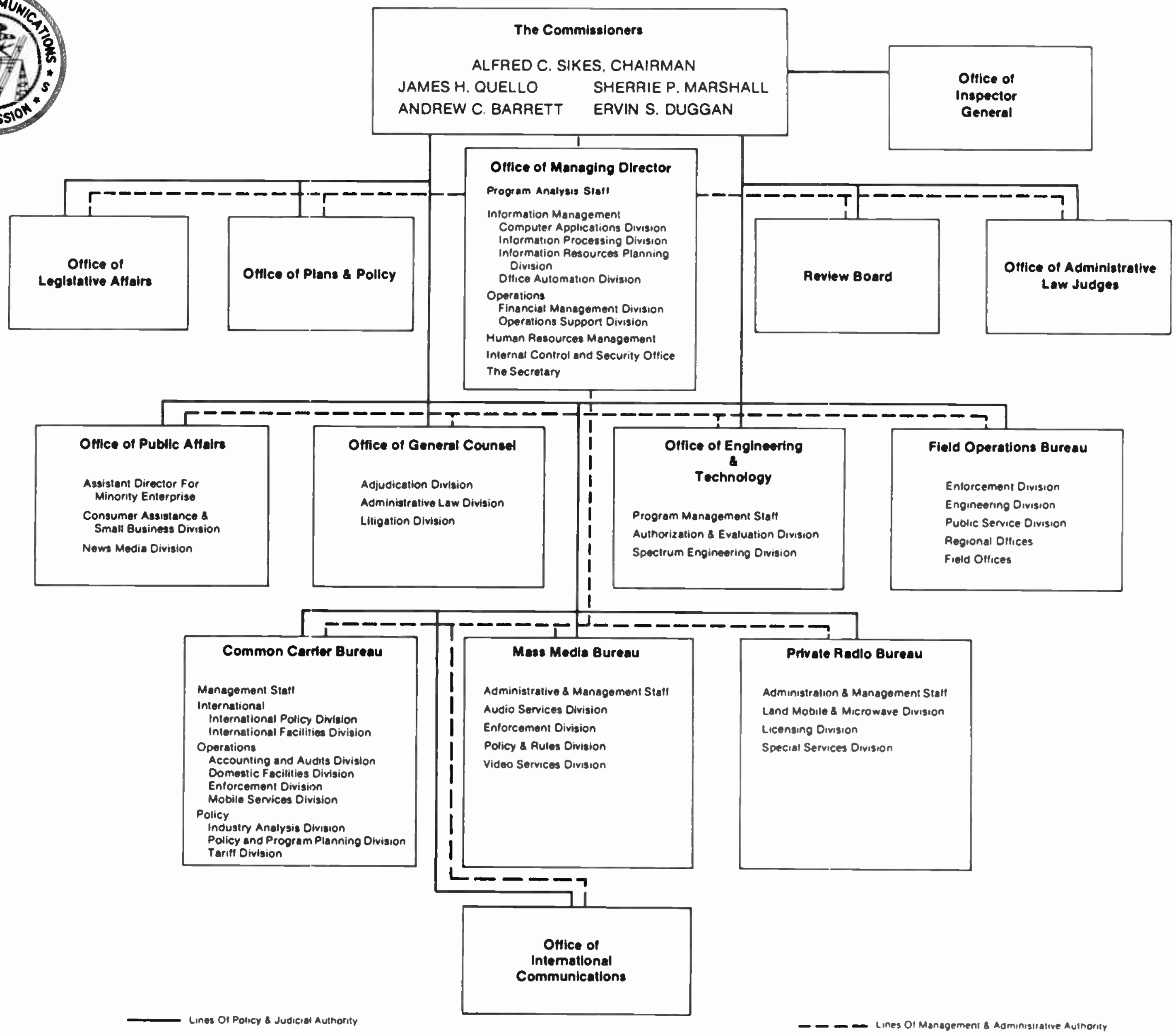


Figure 1. Federal Communications Commission organization chart.

Office of Administrative Law Judges (ALJ)

The Law Judges have responsibility for hearing and conducting all adjudicatory cases designated for any evidentiary adjudicatory hearing other than those to be heard by the Commission en banc.

Office of Science and Technology (OST)

OST conducts scientific and technical studies related to communications and operates the FCC Laboratory in Columbia, Maryland. This Office has primary responsibility for overall spectrum management. The Chief Scientist administers the equipment authorization program (type acceptance, certification, etc.). Additionally, the Office issues authorizations for experimental communications work.

Review Board (RB)

The Review Board (three or more FCC personnel) is set up within the FCC to review decisions of Administrative Law Judges in all adjudicative proceedings, unless the Commission specifies otherwise.

Common Carrier Bureau (CCB)

CCB develops, recommends, and administers policies and programs for the regulation of services, facilities, rates and practices of entities (excluding maritime) which furnish communications services for hire.

Private Radio Bureau (PRB)

PRB develops, recommends, and administers policies and programs for development and regulation of the private radio services, such as those used by persons, businesses, state and local governments, and other organizations licensed to operate their own communications systems for their own use (not for hire).

Field Operations Bureau (FOB)

FOB operates the Commission's field offices, monitoring stations, and mobile monitoring vehicles. It administers the radio operator licensing programs, the field enforcement programs, and the field public service programs. The Bureau has responsibility for the nationwide program of lighting and marking antenna structures.

Mass Media Bureau (MMB)

MMB develops, recommends, and administers policies and programs for development and regulation of the broadcasting and cable television industries. This includes auxiliary services such as translators, low power television, and operational circuits.

MASS MEDIA BUREAU ORGANIZATION

Because the Mass Media Bureau regulates broadcasting and cable television, most broadcast engineers will deal primarily with this Bureau. (Certainly, contact with the Field Operations Bureau can also be expected; see Chapter 1.2: "FCC Field Operations Bureau.") There are four major operational divisions within MMB. Each division has subdivisions, called branches, to handle specific areas of regulation, enforcement, licensing, or policy development, as shown in Fig. 2.

Audio Services Division

Licensing, license modification, and license renewal of AM and FM broadcasting stations fall under the Audio Services Division. The Division also handles the aural auxiliary services, such as Broadcast Remote Pickup, Studio-Transmitter-Links, and FM Translators. Questions concerning pending applications, processing policies, and filing requirements should be directed to the appropriate branch as listed below. Also, requests for waivers of the Rules, requests to operate with parameters at variance (i.e., directional AM antenna out of tolerance), and other requests to operate inconsistently with the station's instrument of authorization (license or construction permit) are handled by this Division. The Audio Services Division branches are:

AM Branch: AM Broadcast radio.

FM Branch: FM Broadcast radio.

Auxiliary Services Branch: Broadcast Remote Pickup, FM Boosters and Translators, and Studio-Transmitter-Links. This division also manages the Mass Media Reference Room.

Video Services Division

Licensing, license modification, and license renewal of television stations are handled by the Video Services Division. In addition to over-the-air television, the Division's responsibilities include Low Power Television (LPTV), Direct Broadcast Satellite (DBS), Cable Television (CATV), Community Antenna Relay Service (CARS), and the Instructional Television Fixed Service (ITFS). Questions concerning pending applications, processing policies, and filing requirements should be addressed to the appropriate branch as listed below. Also, requests for waivers of the Rules or for authority to operate inconsistently with the station's instrument of authorization (license or construction permit) are handled by this Division. The Video Services Division branches are:

Cable Television Branch: CATV and CARS Microwave.

Distribution Service Branch: ITFS and DBS.

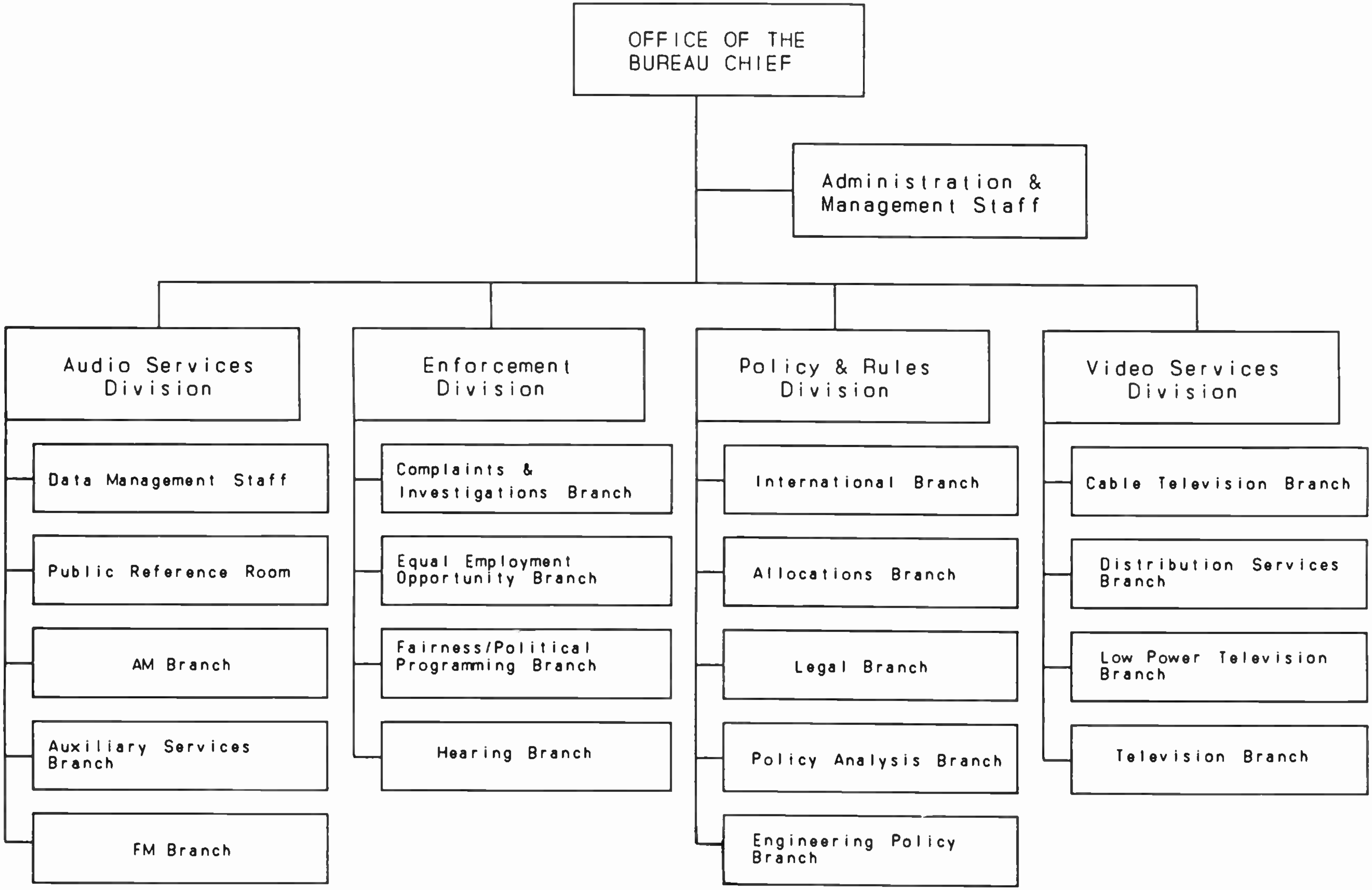
Low Power Television Branch: LPTV and TV Translators.

Television Branch: TV Broadcast.

Enforcement Division

Although the Field Operations Bureau generally conducts station inspections and enforcement monitoring, their work normally concentrates on violations of the technical rules. Enforcement actions in other areas, such as equal employment opportunity or political broadcasting, are handled by the Enforcement Division

MASS MEDIA BUREAU



August 1990

Figure 2. Mass Media Bureau organization chart.

in Washington, DC. This Division also works closely with FOB in enforcement of the technical regulations. Questions regarding enforcement matters should be addressed to the Enforcement Division. The Enforcement Division branches are:

Complaints and Investigations Branch: General.
Equal Employment Opportunity Branch: EEO matters.
Fairness/Political Programming Branch: Fairness Doctrine and Political Broadcasts.
Hearing Branch: Broadcast and CATV hearings.

Policy and Rules Division

New broadcast and CATV rules and regulations generally are written in the Policy and Rules Division. The Division also regularly reviews the rules to eliminate unneeded ones or to modify those that have not kept pace with technology, policies, or legal decision. The Division also interprets broadcasting and CATV rules and regulations. The Division can grant waivers of the Rules to permit experimentation with new technology on broadcast frequencies. The Policy and Rules Division also represents the United States at international forums. Rules and regulations can be classified as: frequency allocations related, legal, policy, or technical. The various branches are:

Allocations Branch: Petitions for Rule Making to amend the FM or TV Table of Allotments.
Legal Branch: Legal rules such as ownership or community ascertainment.
Policy Branch: Rules that are primarily policy decisions such as lottery procedures.
International Branch: International representation, and international coordination of AM and short wave stations.
Engineering Policy Branch: All technical rules.

THE LICENSING PROCESS

Any qualified citizen, firm, or group may apply to the Federal Communications Commission for authority to construct and operate an amplitude modulated (AM), frequency modulated (FM), television (TV), low power television (LPTV) or direct broadcast satellite (DBS) station. The FCC does not license cable television (CATV) systems; however, it does have rules and regulations governing CATV operation and some filings are required. The licensing procedures are outlined in Part 73 of the FCC Rules and Regulations. In general, applicants must satisfy the Commission that they are legally, technically, and financially qualified, and that operation of the proposed station would be in the public interest.

Types of Facilities

Applicants must propose specific station parameters to the Commission. In the case of AM broadcast, applicants generally conduct an interference study to assure that certain field strength contours of the planned station will not overlap existing stations. There are no preplanned AM broadcast assignments to cities or locations. Each application must be uniquely engineered to: (1) serve the community of interest, and (2) prevent interference. Use of a directional antenna will often be required to limit radiation in one or more directions. AM broadcast assignments are made at 10 kilohertz intervals, from 540 kilohertz to 1600 kilohertz, inclusive. Each station has a designation of Class I, II, III, or IV (with additional subdivisions). The classes relate to coverage areas ranging from Class I stations that serve major metropolitan centers and large regions to Class IV stations that serve only one locality. The radiated power of AM broadcast stations varies from 250 watts to 50,000 watts.

Commercial FM and all TV applications differ from AM broadcast in that the Commission has a Table of Allotments in these services. If no allotment exists in the Table, then the applicant must first file a *Petition for Rule Making* to have the proposed city added to the Table. If the proposal meets the Commission's criteria, the proposal will be added to the appropriate Table of Allotments and the applicant can then file for the station. Being successful in amending the Table does not guarantee being granted a station. Once the Table has been amended, other qualified applicants may also file for the station.

Commercial FM stations operate on channels spaced every 200 kilohertz, from 92.1 megahertz to 107.9 megahertz. Stations are classified as A, B1, B, C3, C2, and C, based on location in the country, channel, and power. Class A stations operate between 100 watts and 6,000 watts, while class B stations may operate with up to 50,000 watts and class C stations are permitted 100,000 watts. Television stations have three categories: low VHF, high VHF, and UHF. Channels 2–6 are low VHF and 100,000 watts is the maximum authorized power. High VHF stations operate on channels 7–13 and are permitted 316,000 watts. UHF television stations, which are numbered from 14 to 69, can radiate up to 5 million watts. The coverage areas of all television stations are roughly comparable with the above power levels and with antennas at maximum heights. The power level differences in television compensate for the more difficult propagation conditions at the higher frequencies.

Noncommercial FM (educational) stations have assignments between 88.1 megahertz and 91.9 megahertz. These stations may radiate up to 100,000 watts, but are allocated by actual predictions of service contour overlap rather than by the absolute mileage separations of commercial FM stations. LPTV stations are secondary to full service television stations. LPTV service contours may not overlap full service stations' contours nor cause interference. Successful applicants for direct broadcast satellite transponders receive an exclusive right to use a particular DBS orbital location and channel (or channels) in space. Finally, licensees in the ITFS have exclusive use of a set of frequencies in a given area and may, to a limited extent, merge their facilities with common carrier operations in the Multipoint Distribution Service (MDS).

Applicants may represent themselves before the Commission in licensing matters; however, most applicants employ the services of communications attorneys and consulting engineers. By necessity, to prevent interference and assure license grants to qualified entities, the license process is complex. Professional advice can speed the process and help prevent costly errors.

The First Step to a Broadcasting (AM, FM, or TV) License

After the engineering work has been done to determine that an AM or educational FM station can be designed and built in a community, or that the FM or TV station can be designed and built in a community, or that the FM or TV allotment is now contained in the appropriate Table of Allotments, the next step is to file for a construction permit (CP). Commercial applicants file on FCC Form 301 (Application for Authority to Construct a New Broadcast Station or Make Changes in an Existing Station). Educational station applicants use FCC Form 340 for the same purpose. The application for a CP requests considerable information about the applicant(s) and the proposed facility. Commercial applicants must document the financial ability to operate the station, without revenue, for three months after construction of the station.

If the Commission accepts the application for filing, a *Public Notice* will be released to inform interested parties of the action. After the release of the *Public Notice*, at least 30 days will pass before the Commission again acts on the application. If no objections to the application are received and the Commission otherwise finds the applicants(s) qualified, a construction permit will be issued by the Mass Media Bureau under delegated authority. Again, the Commission issues a *Public Notice* and allows 30 days for interested parties to file for reconsideration of the grant. After grant of the CP, the applicant (now called a "permittee") should file for a call sign. Construction must be completed within 18 months for all classes of station (except TV, which is 24 months) after the grant. Extensions of time may be requested for legitimate reasons.

Contested Applications

Not all applications for construction permits flow smoothly through the FCC. Often other parties will either file applications in competition with the original application or they will contest the application on some grounds. For example, an existing station licensee may disagree with the applicant's basic qualifications or intentions. Any of these protests will undoubtedly delay issuance of the CP, even if the applicant can resolve the conflicts outside the FCC forum. Absent any other method of resolution, the Commission will designate the application for hearing.

Once designated for hearing, the applicant and other parties have 60 days to prepare their respective cases. Generally, the hearing will be conducted by an Administrative Law Judge (ALJ). The ALJ has authority to administer oaths, examine witnesses, and rule on the case based on the evidence presented. After the ALJ closes the record, an initial decision will be issued. The applicant, or any other interested party, may then file exceptions to the finding. The Commission or its Review Board may then hear oral arguments and adopt, modify, or reverse the ALJ's initial decision. If the decision was by the Review Board, the matter may be contested and the case taken to the Commission. Court appeals may be filed after the final decision by the Commissioners.

Station Construction and Equipment Testing

Construction should be exactly in accordance with the terms of the construction permit. If deviations must be made, then an amended FCC Form 301 must be filed. If the station cannot be constructed within the required time period, the permittee must file FCC Form 701 (Application for Additional Time to Construct a Radio Station). It is important to remember that filing of the Form 301 or 701 does not guarantee that an extension of the CP will be granted.

Applying for the Station License

After the station is built and the equipment tested, program testing may begin for nondirectional AM and FM stations and TV stations. At this time, the permittee files for the station license on FCC Form 302. AM and FM stations with directional antennas must apply for and secure authority from the FCC before beginning program testing. Stations operate in the program test mode until a license is received. Processing time varies, but the station license should arrive within a few months, assuming the FCC finds no problems with the Form 302. The FCC may suspend the program test authority because of interference or other appropriate reason.

License Renewal

Broadcast licenses have a life of five years for TV stations and seven years for radio stations. Renewal is not automatic. Licensees file for license renewal on FCC Form 303. If the licensee's record is good and if no one contests the renewal, the renewal process remains relatively simple. On the other hand, the renewal may be placed in hearing status for disposition if the licensee's performance during the previous period is in question, if someone contests the renewal, or if another party files an application which is mutually exclusive with the renewal.

Sales and Transfer

If the holder of a construction permit or license desires to assign it to someone else, application is made on FCC Form 314 (Application for Consent of Assignment of Radio Broadcast Station Construction Permit or License). Should the permittee or licensee wish to transfer corporate control, FCC Form 315 (Application for Consent to Transfer Control of Corporation Holding Broadcast Station Construction Permit or License) is used. FCC Form 316 (Application for Assignment or Transfer—Short Form) may be used when the transfer or assignment involves no substantial change in interest.

Auxiliary and Translator Services

Broadcast station permit and license holders can apply for several types of auxiliary radio stations. For example, Aural Studio-Transmitter-Links are available. Also, Broadcast Remote Pickup (two-way) stations can be used for program production communications or remote broadcasts.

FM licensees and other interested parties may operate FM translators and boosters to fill in areas of the coverage pattern. Likewise, TV licensees and other interested parties may operate TV translators and boosters. Each application varies slightly. Details of these services are contained in Part 74 of the FCC Rules and Regulations.

DBS and LPTV

The application process for these services differs considerably from the usual process for broadcast stations. Unless an application for an LPTV station is not contested (which is very rare), the applicant is selected by weighted lottery. Each applicant receives a weighting in the lottery based on factors such as minority ownership. DBS applicants apply to the Commission for a slot in space in the geostationary orbit. The construction permit can be retained only if a showing of due diligence has been shown in the construction of the facility. Final licensing comes after the station has been placed in operation.

THE FCC RULES AND REGULATIONS

Broadcasters should be familiar with the FCC Rules and Regulations. Every station should have at least one copy of the Rules. The Government Printing Office, not the FCC, sells the Rules through their stores or by mail. New editions are printed yearly. The GPO provides no updates throughout the year when Rules are changed. The following FCC Rule Parts pertain to broadcasting and CATV:

Part 0: Commission Organization.
Part 1: Practice and Procedures.
Part 2: Frequency Allocations and Radio Treaty Matters; General Rules and Regulations.
Part 17: Construction, Marking, and Lighting of Antenna Structures.
Part 73: Radio Broadcast Stations.
Part 74: Experimental, Auxiliary, and Special Broadcast and Other Program Distribution Services.
Part 76: Cable Television Service.
Part 78: Cable Television Relay Service.

All of these Rule Parts can be obtained by purchasing two rule books. The FCC Rules are contained in the Code of Federal Regulations (CFR), Title 47. Parts 0–19 are contained in one book and Parts 70–79 are contained in another. Write the Government Printing Office, Washington, DC, 20402, or telephone (202) 783-3238. Or your local GPO sales office can be contacted directly to order the books.

One or more private publishers offer subscription services to the FCC Rules. Generally, these services are more costly, but regular updates of the Rules are provided to subscribers. Subscriptions can also be obtained to receive copies of daily releases from the FCC.

THE RULE MAKING PROCESS

The FCC, like all other federal government agencies, enacts new rules and regulations through the terms of the Administrative Procedures Act (APA). The APA specifies how rules may be proposed, adopted, and appealed. The APA assures that the public has input into the rule making process. Fig. 3 is a flow chart of the process and may be helpful in understanding the steps. Part 1 of the FCC Rules and Regulations provides detailed information on the FCC's general rules of practice and procedure.

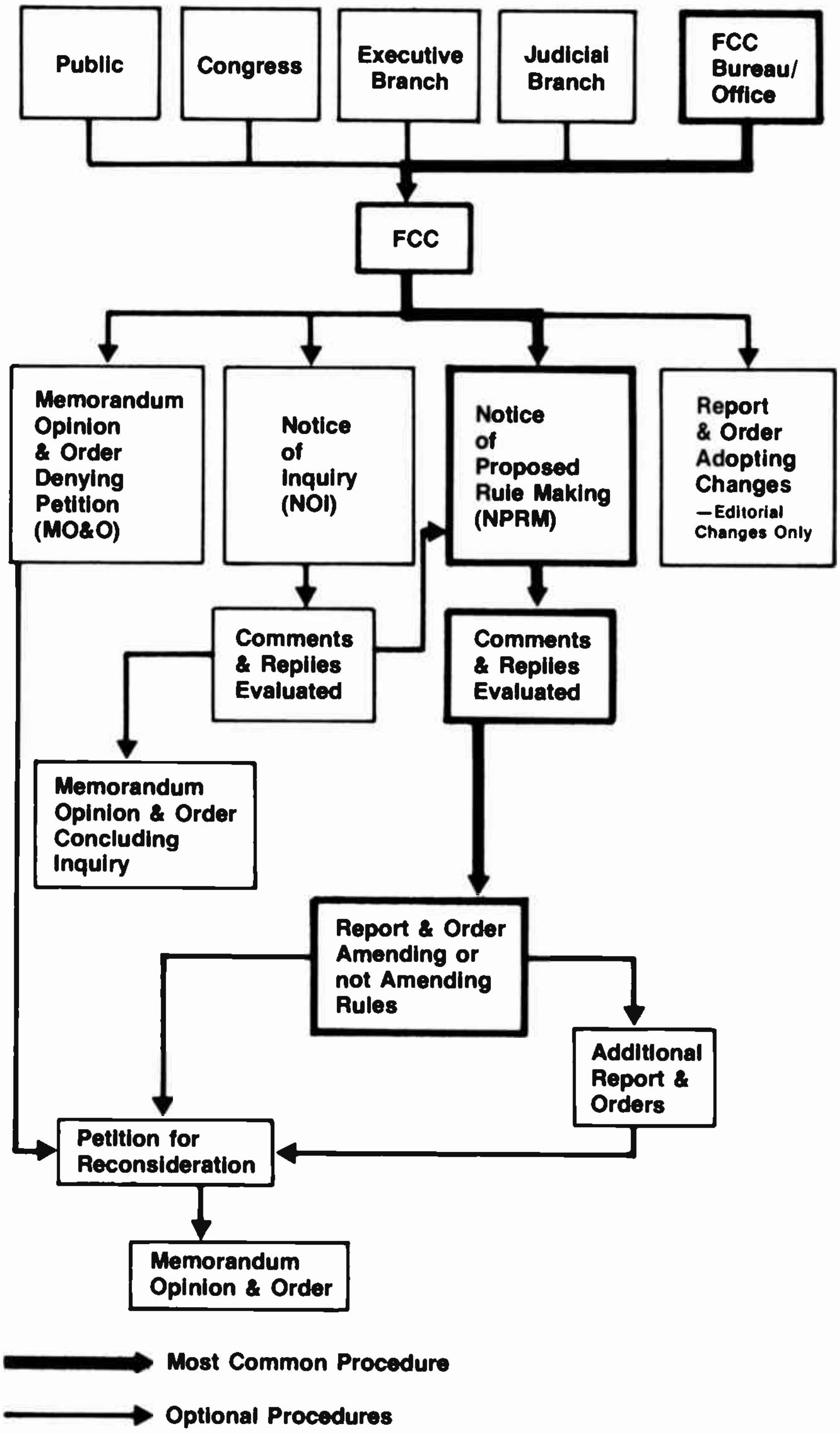
Initiation of Actions

Rule making actions can be proposed by anyone through a *Petition for Rule Making*. Such petitions bring the desires of individuals or groups to the attention of the Commission. The FCC will evaluate the petitions and either dismiss them for one of the reasons in Section 1.401 of the Rules or accept them for action. If accepted, a *Public Notice* will be released giving a brief description of the details of the petition. Others may then submit comments in support of or against the petition.

Other rule makings may begin by direction of Congress, the President, or the courts. The FCC may also initiate rule makings on its own motion.

Early Options

Initiation actions are handled in one of four ways. A petition can be dismissed directly through a *Memorandum Opinion and Order* (MO&O). Minor rule changes of an administrative nature can be made effective directly through a *Report and Order* (R&O). Generally, however, either a *Notice of Inquiry* (NOI) or *Notice of Proposed Rule Making* (NPRM) will be drafted by the staff. If enough information is at hand to draft the proposed rule changes, an NPRM will be issued without first gathering background material in an NOI. After adoption of an NPRM by vote of the Commissioners, the *ex parte* rules become effective (see section below). Because this inhibits direct contact between FCC staff and concerned parties, the NOI may be considered an essential first step in many matters. *Ex parte* does not apply at the NOI stage.



This brief account of how Rules are made at the FCC merely highlights the major components of the process. For details, contact the Dockets Branch.

This article was prepared with the assistance of staff members from several Bureaus and Offices, particularly Sharon Briley, Cable Television Bureau.

Steps:

1. Initiation of Action. Suggestions for changes to the FCC Rules and Regulations can come from sources outside of the Commission either by formal petition, legislation, court decision, or informal suggestion. In addition, a Bureau/Office within the FCC can initiate a Rule Making proceeding on its own.

2. Bureau/Office Evaluation. When a petition for Rule Making is received, it is sent to the appropriate Bureau(s)/Office(s) for evaluation. If a Bureau/Office decides a particular petition is meritorious, it can request that Dockets assign a Rule Making (RM) number to the petition. A similar request is made when a Bureau/Office decides to initiate a Rule Making procedure on its own. A weekly notice is issued listing all accepted petitions for Rule Making; the public has 30 days to submit comments. The Bureau/Office then has the option of generating an agenda item requesting one of four actions by the Commission. If an NOI or NPRM is issued, a Docket is instituted, and a Docket number is assigned.

3. Possible Commission Actions. Major changes to the Rules are presented to the public as either an NOI or NPRM. The Commission will issue an NOI when it is simply asking for information on a broad subject or trying to generate ideas on a given topic; an NPRM is issued when there is a specific change to the Rules being proposed. If an NOI is issued, it must be followed by either an NPRM or an MO&O concluding the inquiry.

4. Comments & Replies Evaluated. When an NOI or NPRM has been issued, the public is given the opportunity to comment initially, and then respond to the comments that are made. When the Commission does not receive sufficient comments to make a decision, a further NOI or NPRM may be issued, again calling for comments & replies. It may be determined that an oral argument before the Commission is needed to provide an opportunity for the public to testify before the Commission, as well as for the Bureau(s)/Office(s) to present diverse opinions concerning the proposed Rule change.

5. Report & Order Issued. A Report & Order is issued by the Commission stating the new or amended Rule, or stating that the Rules will not be changed. The proceeding may be terminated in whole or in part.

6. Additional Reports & Orders Issued. The Commission may issue additional Reports & Orders in the docket.

7. Reconsideration Given. Petitions for reconsideration may be filed by the public within 30 days; they are reviewed by the appropriate Bureau(s)/Office(s) and/or by the Commission.

8. Modifications Possible. As a result of its review of a petition for reconsideration, the Commission may issue a MO&O modifying its initial decision or denying the petition for reconsideration.

Figure 3. How FCC Rules are made.

Comment periods are provided to allow interested parties to express their views. If, based on a review of the comments received, the Commission finds that an NOI should be terminated without further action, it will issue an MO&O stating the reasons for the termination. Otherwise, the matter will be considered for specific rule making in an NPRM. The MO&O is subject to a 30 day period in which parties may file petitions for reconsideration if they have good reason to believe the Commission acted incorrectly.

Notice of Proposed Rule Making

Most rule making items begin as a *Notice of Proposed Rule Making*. The staff prepares the NPRM for consideration by the five Commissioners. An NPRM normally presents the issues and alternatives, may ask specific questions to help finalize the matter, and sets forth proposed rules or rule amendments. An NPRM may address one or multiple topics.

Comments must be filed by the close of the comment period, usually 30 days. Reply comments may then be filed in response to comments, usually for a 15 day period. These two comment periods provide the Commission with a "written debate" of the issues.

The comments and reply comments received are reviewed and enter heavily into the Commission's final actions. However, a comment does not represent a vote for the proposed rules. The Commission must decide on each issue based on the public's interest, convenience, and necessity. Even if the majority of the comments oppose an item, the FCC can nevertheless adopt the proposal. The FCC can also schedule items for oral argument among the interested parties.

Filings should be clear, concise, and address the issues in the NPRM. Emotional responses or uninformed responses are of little value. Copies of the actual texts of the NPRM need to be studied before making comments. All NPRMs are published in the Federal Register and will be available at most larger libraries or law libraries.

The comments are reviewed by Commission staff at the Branch or Division levels. Recommendations are then passed to the Bureau level and ultimately to the Commissioners. The Commissioners will then approve, disapprove, or direct the staff to modify the recommendations.

The staff or the Commissioners may also decide that the record still does not support a decision, and a *Further Notice of Proposed Rule Making* (FNPRM) will be released. Like an NPRM, the comment periods apply. Once the record supports a final decision, the FCC will act by the adoption of a *Report and Order*.

Report and Order

A *Report and Order* may be used to dismiss the proposals in the NPRM, adopt some of the proposals (possibly modified based on the comments), or adopt the rules as proposed. Rules will normally not become effective until 30 days after publication of the R&O in the Federal Register. The R&O constitutes the Commission's final action unless a petition for reconsideration is filed.

Petition for Reconsideration

Any interested person may petition for reconsideration of a final action. The petitions for reconsideration must be filed within 30 days of the date the R&O was published in the Federal Register. The Commission will act on a petition for reconsideration by *Memorandum Opinion and Order*. The initial decision may be modified or the petition may be denied in the MO&O.

Ex Parte Considerations

Most rule making proceedings become *ex parte* upon adoption of an NPRM and remain *ex parte* until 30 days after the Commission's final action (R&O or MO&O for reconsiderations). An *ex parte* contact is any written presentation made to FCC decision making personnel by another person, which is not served on all parties to the proceeding. It is also any oral presentation made to the FCC decision makers, without advance notice to the parties to the proceeding and without opportunity for them to be present. Certain proceedings are "restricted" and no *ex parte* contacts can be made. These would be, for example, contested applications for licenses. Most rule makings are "non-restricted" and do not involve competing claims to a valuable privilege.

Appeal to the Courts

Once the Commission has considered and reconsidered a matter, interested parties may appeal the decision to the federal courts. Ultimately, the Supreme Court could hear the case. Under current law, however, the Federal Communications Commission need not wait for the court decision before enacting the new rules and regulations. As long as the matter has been given full consideration under the Administrative Procedures Act, the FCC may place the new rule in effect until a court rules to the contrary or orders a "stay." The courts typically defer to the expertise of the Commission in technical matters and will often look more to assure that the APA was followed than to try to second-guess the Commission's decision.

1.2 Field Operations Bureau

Staff
Federal Communications Commission, Washington, District of Columbia

THE FIELD OPERATIONS BUREAU

The FCC's Field Operations Bureau (FOB) operates 35 offices throughout the United States, including Alaska, Hawaii, and Puerto Rico. These offices perform liaison between the public and licensees, and carry out enforcement and public service activities.

FOB Headquarters (Washington, DC)

The headquarters of FOB is divided into three Divisions:

1) Engineering Division
2) Enforcement Division
3) Public Service Division

These Divisions formulate policies that guide field work. Responsibilities and functions of the Divisions are:

Engineering Division

The Engineering Division provides engineering expertise for FOB. The Engineering Division develops measurement procedures, designs and builds specialized monitoring vehicles, and develops new tools such as direction-finders. Questions concerning any of these areas may be directed to the Chief, Engineering Division.

Enforcement Division

The Enforcement Division directs all field inspections, investigations, and monitoring (signal analysis) operations. This Division sets priorities for enforcement actions and administers the *Notice of Apparent Liability* (NAL) program. Inquiries concerning enforcement actions may be addressed to the Chief, Enforcement Division. Inquiries may also be made to the Engineer-in-Charge (EIC) of the local field office.

Public Service Division

The Public Service Division administers the Bureau's public service program. This Division assists the public, licensees, and other government agencies with telecommunication-related matters. It also analyzes trends and problems disclosed through public contact. The Division resolves interference problems, supervises the development of publications, and administers the commercial radio operator licensing program.

A subdivision, the Antenna Survey Branch, prescribes painting and lighting requirements for communications towers. Questions in these areas may be addressed to the Chief, Public Service Division.

FOB Facilities

Field offices are grouped into six regions. These offices perform the public service and enforcement functions.

Regional Directors

Each region has a Regional Director whose function is to supervise and administer the operations of the local FOB offices under their jurisdiction. Regional Directors set regional priorities and provide a mid-level supervisory function. They are located in Atlanta, Boston, Chicago, Kansas City, San Francisco, and Seattle.

Local FOB Offices

Most direct contact between broadcasters and the FCC occurs at the local level. Each FOB office is headed by an Engineer-in-Charge, with technical and clerical support staff. These are multi-purpose offices. They provide information, educational activities, radio operator examination and licensing, and conduct enforcement efforts. Off-air and on-site monitoring capabilities are also functions of most FOB offices.

Broadcasters are urged to make contact with their local FOB office and become acquainted with the staff. They can be of invaluable assistance in helping broadcasters understand rules or policies and in resolving interference problems.

Offices are located in: Allegan, MI; Anchorage, AK; Atlanta, GA; Baltimore, MD; Belfast, ME; Boston, MA; Buffalo, NY; Chicago, IL; Dallas, TX; Denver, CO; Detroit, MI; Douglas, AZ; Ferndale, WA; Grand Island, NE; Honolulu, HI; Houston, TX; Kansas City, MO; Kingsville, TX; Laurel, MD; Livermore, CA; Los Angeles, CA; Miami, FL; New Orleans, LA; New York, NY; Norfolk, VA; Philadelphia, PA; Portland, OR; Powder Springs, GA; St. Paul, MN; San Diego, CA; San Francisco, CA; San Juan, PR; Seattle, WA; Tampa, FL; and Vero Beach, FL.

FOB Inspections

Checklists have been prepared by FOB's Public Service Division to assist broadcasters in preparing for an official FCC inspection. It is suggested that the checklists be used on a regular basis to help prevent serious problems from occurring. One checklist covers AM/FM/TV stations, while a second checklist pertains to cable television systems. Copies are available from local FOB offices or from FOB headquarters in Washington, DC.

Elements of an Inspection

Most inspections follow a similar pattern. The inspector will be primarily interested in the technical operation of the station. Although some attention may be paid to required nontechnical records, such as the Public Inspection File, the majority of the time will be devoted to a review of technical parameters. In the case of AM directional stations, the inspection will include visits to the monitoring points.

The following areas are included in inspections, although a particular inspection need not cover all areas, and may be extended to additional areas if needed.

1. Station documents and records, including the Public Inspection File, the tower light inspection log, and other logs as may be required.
2. Transmitter operation.
3. Studio and control point operation.
4. Tower and antenna including painting and lighting.
5. Emergency Broadcast System.
6. Remote control operation.
7. Extension metering.
8. Automatic transmission system.
9. Radio operator licenses.
10. Directional AM antenna parameters.
11. Auxiliary broadcast stations.
12. Signal analysis (frequency, modulation, occupied bandwidth, etc.).

OFF-THE-AIR MONITORING

Technical parameters of station operation may be monitored off-the-air by FCC Field Offices or by FCC enforcement vehicles. Examples of the forms used by the FOB when checking technical parameters are shown in Figs. 1 and 2 at the end of this chapter. Licensees do not usually know about the Commission's monitoring until after the fact. Violations will be brought to a licensee's attention through an *Advisory Notice*, a tier-two *Advisory Notice*, or a tier-one *Violation Notice*, and may also trigger a station inspection.

After the inspection, the inspector will hold a review session with the chief operator and station licensee or general manager. All areas that need improvement will be discussed. Depending on the severity of a particular violation, FCC actions could range from simply a warning to a *Notice of Apparent Liability*, which could ultimately result in a monetary forfeiture.

NOTIFICATION OF VIOLATIONS

If violations were noted during the inspection, the inspector will return to the FCC office and mail an official notice of the violations to the station. One or more of the following kinds of notices may be issued:

Advisory Notice (FCC Form 790)

Near violations will be brought to the attention of licensees using FCC Form 790. Form 790 indicates that an observed condition in some area of the station operation has a high probability of becoming a violation if the problem is not corrected. No reply to an *Advisory Notice* is required. (See Fig. 3 at the end of this chapter.)

Tier-Two Violation (FCC Form 790)

Violations that have little probability of causing interference with other stations will be notified by the tier-two *Advisory Notice*. Again, FCC Form 790 is used. Although no reply is required, licensees should immediately correct the deficiencies. Left unchecked, these violations could lead to more stringent enforcement actions. (See Fig. 3.)

Tier-One Violation (FCC Form 793)

Violations that involve safety of life or property, have a high potential to cause interference, or are of a serious nature for other reasons, will be documented on FCC Form 793. This is an *Official Notice of Violation*. Licensees must reply to such notices within 10 days of receipt. Failure to reply could result in a monetary forfeiture or ultimately in license revocation. Even if the response cannot be fully completed within the 10 days, an initial response should be submitted and an extension of time must be requested. Violations in this category are serious and require immediate attention. Depending on the nature of the violation, whether it is repeated or willful, and the licensee's response, additional sanctions could be imposed. (See Fig. 4.)

MONETARY FORFEITURES

Serious violations can result in a monetary forfeiture. In most cases, the local field office will issue the monetary forfeiture. However, in extreme cases, some forfeitures will be issued by the FCC in Washington, DC. The field-issued forfeitures may be for items related to safety, interference, tower lighting, or important administrative matters (such as not having a licensed operator on duty). More serious violations may have significantly higher monetary forfeitures imposed by the Commission. Failure to pay a forfeiture will result in court proceedings to collect the monetary sanction.

CRIMINAL PENALTIES

Broadcasters normally will not be involved in violations serious enough to be handled as criminal violations. Criminal acts may be committed by persons operating unlicensed stations or marketing illegal radio equipment. The penalties may include fines and/or prison.

APPEALS

As with any governmental enforcement action, licensees have the opportunity throughout the sanction process to appeal the action. If the broadcaster feels that the sanction is inappropriate or that the amount of a monetary forfeiture should be reduced, the office issuing the sanction should be contacted first. Appeals may be pursued through the FCC Regional Director, through the Chief, Enforcement Division, and at higher levels if warranted. It may be useful to obtain professional counsel.

Form FO-794-C
February 1975

FEDERAL COMMUNICATIONS COMMISSION
Field Operations Bureau
Enforcement Division

SAMPLE

FM BROADCAST SIGNAL ANALYSIS REPORT

CALL ____ FREQUENCY ________ LOCATION ____________
MONITORED ________ FROM ________ TO ________
BY ________________

MODULATION MEASUREMENT
_____ % at _____

CARRIER FREQUENCY
_____ Hz at _____

STEREOPHONIC PILOT SUBCARRIER
_____ kHz _____ % at _____

RESIDUAL STEREO SUBCARRIER
_____ % at _____

SCA SUBCARRIER
_____ kHz _____ % _____ kHz modulation at _____

SCA Program Material ________________

SPURIOUS BASEBAND SIGNALS
_____ kHz _____ % at _____

SAMPLE

STATION IDENTIFICATION
Required times _____ Format _____
Dual-city ID _____
Simultaneous AM-FM _____
Discrepancies _____

AUDIO QUALITY ________________

SPURIOUS SIGNALS ________________

REMARKS ________________

STEREOPHONIC PERFORMANCE MEASUREMENTS

LEFT CHANNEL ONLY
_____ Hz _____ dB(M) _____ dB(S)

RIGHT CHANNEL ONLY
_____ Hz _____ dB(M) _____ dB(S) _____ %

MAIN CHANNEL/STEREO SUBCHANNEL CROSSTALK (LEFT PLUS RIGHT)
_____ Hz _____ dB(M) _____ dB(S) Diff. _____ dB - 6 dB = _____ dB crosstalk

SAMPLE

STEREO SUBCHANNEL/MAIN CHANNEL CROSSTALK (LEFT MINUS RIGHT)
_____ Hz _____ dB(S) _____ dB(M) Diff. _____ dB + 6 dB = _____ dB crosstalk

STEREO SUBCARRIER/PILOT SUBCARRIER PHASE RELATIONSHIPS
_____ Hz _____ °phase difference _____ Hz _____ °phase difference

SEPARATION EVALUATION - LEFT INTO RIGHT
_____ Hz _____ (M) _____ (S) _____ ratio = _____ dB separation

SEPARATION EVALUATION - RIGHT INTO LEFT
_____ Hz _____ (M) _____ (S) _____ ratio = _____ dB separation

SAMPLE

EQUIPMENT CONFIGURATION

Signals introduced at _____ through _____
Studio-transmitter link _____
Stereo generator _____ located _____
Exciter _____ Transmitter _____
Modulation monitor _____

REMARKS ________________

Figure 1. Sample FCC FM Broadcast Signal Analysis Report.

FEDERAL COMMUNICATIONS COMMISSION
FIELD OPERATIONS BUREAU

TELEVISION SIGNAL ANALYSIS REPORT

CALL ______ CHANNEL ______ LOCATION ______
MONITORED ON ______ FROM ______ TO ______
BY ______ Monitoring Location ______

VISUAL CARRIER	HORIZONTAL SCAN RATE	VERTICAL SCAN RATE
______ Hz at ______	______ kHz at ______	______ Hz at ______

COLOR SUBCARRIER FREQUENCY	CARRIER SEPARATION
______ MHz at ______	______ Hz at ______

SAMPLE

BLANKING LEVEL*	(72.5-77.5)%	______	______
WHITE LEVEL*	(10-15)%	______	______
SETUP INTERVAL*	(5-10) IRE Units	______	______
HORIZONTAL SYNC PULSE WIDTH	(4.4-5.1) us.	______	______
FRONT PORCH DURATION	(minimum 1.3) us.	______	______
SYNC TO END-OF-BURST DURATION	(maximum 7.9) us.	______	______
SYNC TO START OF VIDEO DURATION	(9.2 min) us.	______	______
TOTAL HORIZONTAL BLANKING INTERVAL	(11.5 max) us.	______	______
COLOR BURST LENGTH	(8-11) cycles	______	______
COLOR BURST AMPLITUDE	(90-110) % of sync	______	______
BREEZEWAY DURATION	(minimum 0.4) us.	______	______
PULSE RISE TIME	(maximum 0.3) us.	______	______
EQUALIZING PULSE WIDTH	us.	______	______
	(45-55% of H sync)%	______	______
SERRATION WIDTH	(3.8-5.1) us.	______	______
SPIKING/OVERSHOOT/TILT		______	______
BLANKING/SYNC TIP VARIATION	(MAXIMUM 5) %	______	______
VERTICAL BLANKING INTERVAL	(18-21) LINES	______	______

*USING IRE ROLLOFF FILTER

(over)

Form FO-794-B
January 1984

VERTICAL INTERVAL TEST SIGNALS

Field 1	Line	Field 2
______	______	______
______	______	______
______	______	______
______	______	______
______	______	______
______	______	______
______	______	______
______	______	______
______	______	______

Discrepancies ______

SPECIAL VIDEO MEASUREMENTS ______ SAMPLE

AURAL MODULATION MEASUREMENT

______ % at ______

SPURIOUS AURAL BASEBAND SIGNALS

______ kHz ______ % at ______

AURAL SUBCARRIER
______ kHz ______ % ______ kHz modulation ______ content at ______

SPURIOUS RF SIGNALS ______

DEMO INPUT ______ db LSB COLOR SUBCARRIER SUPPRESSION ______ db
VISUAL/AURAL RATIO ______ db VISUAL/COLOR BURST RATIO ______ db
VISUAL SIGNAL QUALITY TASO ______ REASON ______
AURAL SIGNAL QUALITY ______ SAMPLE

STATION IDENTIFICATION OK ______ NOT OK ______
TIME ______ FORMAT ______

REMARKS ______

Figure 2. Sample FCC Television Signal Analysis Report.

UNITED STATES OF AMERICA
FEDERAL COMMUNICATIONS COMMISSION
WASHINGTON, D.C. 20554

NOTICE OF
RADIO STATION CONDITIONS

FREQUENCY		CALL SIGN
EMISSION	DATE OBSERVED	TIME OBSERVED
LOCATION OF STATION OR NAME OF CRAFT		RADIO SERVICE OR CLASS OF STATION

NO REPLY
IS NECESSARY

☐ Items are violations of the Commission's Rules detected during monitoring or an inspection. These items must be corrected promptly.

☐ Items identify unsatisfactory or marginal conditions noted during monitoring or an inspection. We suggest you correct these items to avoid violations or operating problems.

PARTICULARS:

SAMPLE

FCC OFFICE ADDRESS	DATE SERVED	ISSUING OFFICER
		SUPERVISOR

CAUTION: Violations, if repeated or willful, may result in imposition of sanctions, including monetary fine, revocation or suspension of operating authority or prosecution. (See Section 501, 502, 503, 312, and 303 (m) of the Communications Act of 1934, as amended and Section 1.89 of the Commission's Rules.) We are retaining a copy of this notice.

(All previous editions of this form are canceled.)

FCC Form 790
November 1978
★ U S GPO 1981-355-475

Figure 3. Sample FCC Advisory Notice.

United States of America
FEDERAL COMMUNICATIONS COMMISSION

OFFICIAL NOTICE OF VIOLATION

1. Name and Address of Licensee

	WARNING: Violations, if repeated or willful, as well as your failure to reply to this notice, may result either in the imposition of monetary forfeitures, the revocation of your station license or suspension of operator license. (See Sections 503, 312, and 303(m) of the Communications Act of 1934, as amended and Section 1.89 of the Commission's Rules.)

(See Instructions and Privacy Act Notice on reverse side)

2. FREQUENCY				
2a. Authorized	2b. Measured	2c. High/Low (Hertz)	3. Emission	
4. Location of Station or Name of Craft	5. Radio Service or Class of Station	6. Hour(s) of Violation (EST-GMT)	7. Date(s) of Violation	8. Call Sign

9. VIOLATION(S)

SAMPLE

ISSUING OFFICER	SUPERVISOR—LOCATION	DATE MAILED/SERVED

The knowing and willful making of any false statement in reply to this NOTICE is punishable by fine or imprisonment under Title 18, United States Code, Section 1001.

(See Reverse Side)

FCC Form 793
January 1985

TO THE LICENSEE:

The facts set forth herein indicate that you have violated the requirements of law or treaty. This Notice is issued in accordance with Section 1.89 of the Commission's Rules.

1. Within 10 days from receipt of this Notice, a written reply shall be addressed to "Federal Communications Commission" and SENT TO THE ADDRESS SHOWN ON THE FACE HEREOF AT THE TOP OF THE PAGE. DO NOT address your reply to an individual.

2. **MAKE CERTAIN THAT YOUR ANSWER:**
   a. Fully explains each violation.
   b. Specifically describes the action taken to correct and to prevent continuation or recurrence of each violation.
   c. Is identified as a reply to this Notice. Include the *call sign* of your station so that your answer may be properly associated with the station file.
   d. Does not refer to a reply to another notice, but is complete in itself.
   e. Is dated and is signed by the licensee or, if appropriate, an officer of the licensee.
   f. Includes all the information requested above in addition to any other information requested in Items 3 and 4, below.

3. **CHANGE OF ADDRESS:**
   If the address appearing in Block 1 on the face of this Notice is not your correct address for receipt of mail, include the correct address in your letter of reply to this Notice.

4. If an "X" appears in the box preceding any of the following instructions, comply with the instruction(s) so indicated and submit the information together with the above described letter.

☐ a. State the name of the person who operated the transmitter at the time of the violation. Does this person hold an operator license or permit issued by the Federal Communications Commission?

☐ b. A second copy of this notice is enclosed for the operator to answer the following questions thereon. That copy must be submitted with the letter of reply described in Item 2 above from the person addressed on the face of this notice. RETAIN THE ORIGINAL OF THIS NOTICE.

I, the undersigned, was the operator on duty at the time of the violation noted hereon, and hereby acknowledge this NOTICE. I hold FCC-issued Radio Operator (Not Station) License or Permit as follows (if none, so state): ____________
Number (If unnumbered, so state): ____________
Issuance date: ________ Birth date: ________
Name (print): ____________
Address: ____________
Signature: ____________

SAMPLE

NOTICE REQUIRED BY THE PRIVACY ACT OF 1974, P.L. 93-579, DECEMBER 31, 1974, 5 U.S.C. 552a(e) (3):

The staff will use all relevant and material information before it, including the information disclosed in your reply, to determine what, if any, enforcement action is required to ensure current and future Rule compliance.

Willful or repeated violation or failure to reply may result in a monetary forfeiture or license revocation.

FCC Form 793
January 1985

Figure 4. Sample FCC Notice of Violation.

INFORMATION BULLETIN

Federal Communications Commission
Field Operations Bureau

Broadcast Service Checklist (AM/FM/TV)
FCC Rule Part 73

This document has been prepared to assist you in conducting an inspection of your station prior to an official FCC inspection. You may wish to use this checklist at regular intervals to ensure that your equipment is properly maintained.

The items in this document are not inclusive of all FCC broadcast rules. Rather, they reflect the most common areas where problems occur and for which the FCC may impose a fine.

Federal Communications Commission
1200 Communications Circle
Virginia Beach, Virginia 23455-3725
Phone: (804) 441-6472

FO Bulletin No. 18
November 1987

I. INTERFERENCE

RULE REFERENCE	SUGGESTED PROCEDURE
[ ] Operation Within Limits Rules: 73.44, 73.62, 73.317, 73.687 73.1545, 73.1560, 73.1570	Develop consistent measurement intervals to maintain operation within authorized limits of your license for: -- operating power -- carrier frequency tolerance -- modulation levels -- spurious or harmonic emission -- directional antenna parameters
[ ] Remote Control, Automatic Transmission System (ATS) Equipment, Extension Metering Rules: 73.57, 73.1410, 73.1500, 73.1550	Provide contingency plan for continuity of operation by direct control if remote or ATS equipment fails. Calibrate indicators and alarms as required to guarantee proper operation and maintenance of technical parameters. Fail-safe capability to turn off the transmitter is essential.
[ ] Directional Antennas - Maintenance Rules: 73.61, 73.62, 73.1745	Perform proper and timely measurements at monitoring point locations specified in license. Change mode of operation as required and review records to verify proper station operation and power levels in appropriate time periods. Keep AM directional antenna parameters within tolerance.

RULE REFERENCE	SUGGESTED PROCEDURE
[ ] Antenna Structures and Lighting Rule: 73.1213	*Develop a verification process for the status of the antenna structure which includes: -- daily observation of lights in accordance with station authorization -- direct an immediate notification to FAA of outage or loss of top beacon -- immediate repair of any defective operation/equipment -- periodic review of condition and color of tower paint.
[ ] Fencing Requirements Rule: 73.49	Prevent access of unauthorized personnel by providing fence and lock for system's antenna area. Check periodically that neither has been tampered with or damaged.
[ ] Emergency Broadcast System Requirements Rules: 73.932, 73.961, 73.1820	Conduct a random test of the EBS system once a week. More often, verify that the EBS monitor receiver and tone generator are operating and set at proper frequencies. Check EBS test entry logs frequently to determine that test has been conducted and received. Familiarize all station personnel with EBS procedures, location of EBS Checklist and Authenticator Word List. Arrange easy accessibility/availability of documents. Ensure that station name is on FCC's mailing list for semi-annual revision of both lists. Be familiar with your State plan.

*Bulletin FO-13, Radio Tower Painting and Lighting, contains additional information. Contact the FCC office nearest you.

III. SERVICE QUALITY

RULE REFERENCE	SUGGESTED PROCEDURE
[ ] Technical Parameters Rules: 73.319, 73.322, 73.682	Monitor equipment as often as necessary to keep parameters within limits. Inspect, adjust or repair equipment frequently to alleviate parameters drifting out of tolerance.
[ ] Modulation Rule: 73.1570	Check equipment frequently to ensure that modulation limits, including stereo and subsidiary subcarrier levels, are maintained. Adjust or repair equipment as required.

IV. ADMINISTRATIVE AND NON-TECHNICAL

[ ] Licensed Operator on Duty Rule: 73.1860	Verify that adequately trained and licensed operating personnel are on duty during all periods of operation.
[ ] Designation of Chief Operator Rule: 73.1870	Ensure that: position of chief operator is a clear-cut designation; documentation is easily accessible; chief operator is well informed on all aspects of technical operation and condition of transmitting system; and that the chief operator keeps management up-to-date with problems detected, status of repairs, adjustments, as well as any unusual operating procedures to be followed.
[ ] Posting of Station and Operator Licenses Rule: 73.1230	Post all required licenses in an easily accessible, highly visible location, visible from the principal control point of transmitter. This helps everyone to be familiar with with the terms of licensing.

IV. ADMINISTRATIVE AND NON-TECHNICAL - CONCLUDED

RULE REFERENCE	SUGGESTED PROCEDURE
[ ] Public Inspection File Rules: 73.3526, 73.3527	Maintain a public inspection file that is complete, current, easily accessible and readily available for inspection by the public, at a permissible location.
[ ] Station Logs Rules: 73.1800, 73.1820	Maintain an accurate, complete and current station log, and ensure that the log is signed.
[ ] Special Technical Logs Rule: 73.1835	Maintain an accurate, complete and current record of any interference situation, malfunction of equipment or significant system adjustments, and operating paramaters as may be directed by the FCC.
[ ] Measurement Records Rule: 73.1225	Update and make available antenna resistance or common point impedance measurements and any required equipment performance measurements.
[ ] Station Identification Rule: 73.1201	Identify station in manner prescribed by rules and as stated in license.
[ ] Specifications for Indicator Instruments Rule: 73.1215	Verify accuracy of instruments and repair, adjust and replace equipment as necessary.

1.3 Frequency Coordination

Robert Van Buhler
KNIX-FM/KCWN-AM, Real Country Network, Tempe, Arizona

Richard Rudman
Radio Station KFWB, Los Angeles, California

WHAT IS FREQUENCY COORDINATION?

The push to leave the studio to originate programs has been a part of broadcasting from the beginning. Now, with over 10,000 radio and 1,400 TV stations on the air, the demand for reliable ways to get news, sports, public affairs, and entertainment programming from the field back to the studio has naturally increased. Part 74 of the FCC's Rules deals with segments of the spectrum that have been assigned to the Broadcast Auxiliary Service to accommodate this demand. Part 74 also covers other services outside the scope of this discussion, such as FM and TV translators, Low Power TV, ITFS, and experimental broadcast stations.

While the number of stations requiring support spectrum has grown tremendously since 1970, the amount of spectrum allocated for this support has stayed relatively constant. Broadcasters using Part 74 spectrum began to experience competition for the limited number of channels allocated not only among broadcasters, but also among the other users who share some of the Part 74 spectrum, such as cable operators. Interference from other broadcasters has gradually become a critical problem in many markets and there are no provisions for exclusive licensing in Part 74. Consequently, frequency coordination is necessary to facilitate the allocation of limited spectrum for a growing number of uses among a growing number of users.

SOME HISTORY OF BROADCAST COORDINATION

Formal coordination has existed for some time for major events such as Olympics, political conventions, and space launches. In the case of the political conventions, the major national television networks have rotated the chairmanship of an "ad hoc" group, known as the National Political Conventions Frequency Coordinating Committee. Washington, D.C. has had unique coordination problems along with a history of cooperative effort for pool feeds to cover events of national interest. Such committee activity has traditionally arisen out of a need, mutually felt by broadcast engineers, to discuss concerns before they become problems. Among the oldest groups formed to deal with mutual operational problems, including coordination, were the Washington Executive Broadcast Engineers (WEBE) and the TV Broadcasters All-Industry Committee in the New York metropolitan area.

The first major region to form a committee to deal exclusively with Part 74 coordination was Southern California. In 1976, the FCC revised Part 74, Subpart D, of its Rules, which deals with the Remote Pickup Service used by radio and television stations for dispatch, on-air remote pickup, cues and orders, communications, and transmitter telemetry. Recognizing the use of narrowband FM technology, the Commission reallocated the UHF (450 and 455 MHz) frequencies from wideband 100 kHz channels to smaller bandwidth 50 kHz, 25 kHz, and 10 kHz channels. Under the FCC's 1976 changes, R and S channels were intended for program material, and cues and orders necessary to implement that programming. N1 and N2 channels were to be used for broadcast program material, cues and orders necessary to that material, and operational communications. Microwave path setup and dispatch functions fall under this last category. P channels may be used only for Operational Communications and telemetry. In 1980 the FCC granted the Southern California Frequency Coordinating Committee (SCFCC) a waiver which further split the S and N1 channel as shown in Table 1.

Note that since the SCFCC plan allows for center channel use of the "split" N1 50 kHz channels, the maximum number of 25 kHz channels can jump from

TABLE 1
1976 FCC 450 Channel Plan Showing SCFCC Waiver Changes

CHANNEL WAIVER DESIGNATION	CHANNEL BAND-WIDTH	PEAK FM DEVIATION	TOTAL # OF CHS.	TOTAL # of SCFCC CHANNELS
N1	50 kHz	10 KhZ	12	0
N2	25	5	24	60
R	50	10	10	14
S	100	35	2	0
P	10	1.5	8	8

24 to 60 since the center channel can be used when geographic separation and or terrain shielding permit.

The FCC will now accept license applications based on channel segments of 5 kHz. These 5 kHz segments can be "stacked" for the Part 74 VHF high band and for the UHF band to form channels of the required bandwidth for a given use. Users should consult either their local coordination group or the FCC's Auxiliary Branch in Washington about a local band plan if questions arise. Regions considering formation of a coordination group can study existing band plans based on segmented channels before implementing new plans. No examples of band plans are given. Plans must be custom tailored for each region.

Coordination activity spans many aspects of broadcasting. Some regions have not experienced congestion and interference with their own licensees, but have organized committees to deal with problems associated with broadcasters who enter their regions for temporary operation. After conducting a survey at the chapter level on how this problem is viewed locally, the Society of Broadcast Engineers (SBE) has identified frequency coordination as a chapter project. The SBE National Frequency Coordinating Committee (NFCC) was formed to assist any chapters interested in this project. In addition, the SBE NFCC publishes a list of volunteer coordinators who are responsible for over 90 regions in 50 states, plus Puerto Rico.

If there is one adjective that can be used to describe the nature of effective frequency broadcast coordination, that word is local. Electronic Newsgathering (ENG) for both radio and television introduces many unknowns. Mobile paths cannot be predicted from day to day, or even hour to hour. Large scale news, public affairs, or sporting events may mean interference for everyone involved unless technical personnel familiar with a number of factors are consulted. Since many of these factors are dependent on local propagation anomalies, terrain considerations, and user patterns, experience has proven that engineers who work with these factors at the local level on a day-to-day basis are the best people to consult with and to ask for advice for successful operations.

PROPAGATION CHARACTERISTICS

A number of factors can modify the expected line-of-sight characteristics of VHF and UHF aural and visual transmission. Many of these can be understood best if the radio waves so affected are compared to light waves. Radio waves sometimes bend around corners in a manner than can be compared to light travelling through the reflective surfaces of a periscope so the viewer can see around corners.

The three most common propagation anomalies are (1) ducting effects, (2) reflection due to thermal inversion layers, and (3) obstacle gain (sometimes referred to as knife edge refraction).

These anomalies can occur for seconds at a time, or may exist in a predictable manner for much longer periods. For instance, obstacle gain can be used in some cases to provide permanent and quite dependable links. In the Part 74 environment, these anomalies have to be recognized and taken into account to assure cochannel and adjacent channel users of adequate protection for both fixed and mobile links. Known anomalies in a region must be factored into any coordinated environment. As in the case of terrain, licensees should seek out engineers who have long-time experience in the region to be coordinated so anomalies that have been identified, and the paths they affect, can be taken into consideration.

THE BROADCAST COORDINATION PROCESS

Successful coordination efforts are all founded on licensee-to-licensee contact. The coordination effort, however it is structured, acts to facilitate this contact. As of this writing, Part 74 of the FCC's Rules places the burden of preventing interference squarely on the licensee. Since broadcasting is highly competitive, no local committee should place itself in a position where it "assigns" spectrum. Eligible radio, TV, cable, and other selected entities have an equal right to use shared frequencies specified under Part 74 or Part 78 of FCC Rules and Regulations, subject only to the service and classifications contained therein. No action taken by a coordination committee should restrain these privileges in any way.

Unless otherwise determined by future FCC actions, committees are voluntary facilitators of the coordination process. The committee has no authority to force adoption of its recommendations. Successful committees never act as policemen. This role is left for the FCC if all other means of settling disputes fails. However, committees can and do use neutral mediators to try to solve the inevitable problems and disputes that arise in a competitive world. The committee exists to facilitate coordination, not to enforce it.

A chaotic situation would already exist in most markets if the engineering community had not developed fundamental coordination techniques and practices. Many cable and network entities, production users, and others who now have access to spectrum originally intended to serve only the needs of local broadcasters would not be able to operate reliably otherwise.

The broadcast or cable engineer does not operate in a purely technical environment with frequency coordination. It is an art and technical discipline that extends beyond the engineering department. There are other key players whose cooperation and participation is essential to efficient frequency coordination.

A news director can create a climate of efficient spectrum management in a radio or TV operation by stressing the importance of adhering to good operating practice and home channel plans.

Both news and engineering managers must realize that frequency (or channel) information must be shared by users for coordination to succeed. The ultimate goal is interference-free operation for all users. This goal must be kept in mind whenever Broadcast Auxiliary Spectrum/Cable TV Relay Service (BAS/CARS) frequencies are employed. Cooperation from all parties, and an active and supportive leadership example from engineering and news managers in charge, is essential. Radio/TV station and cable general managers can make frequency coordination possible by allowing their chief engineers the time and sometimes the financial support needed to get the job done. Many progressive local broadcast associations have seen value in underwriting the costs of frequency coordination by providing computer equipment, postage, stationery, duplication expenses and other support to make frequency coordination a success.

General managers are vital to the process in another way. In many stations, engineering is provided by a contractor whose services are limited to maintenance and repair. By nature of their contract, these providers have very little input on the design of mobile systems or fixed links. In such cases, it is important that the manager insist that his BAS/CARS equipment applications and operation be handled harmoniously with the needs of the local coordinating committee and the broadcast and cable community it serves.

Station or system managers should know if their engineers are involved in frequency coordination. The best approach is to meet with the engineer and discuss the importance of coordination to the well-being of the broadcast and cable community.

Simply requiring chief engineers to attend regular coordination meetings is helpful to the coordination process. By attending the meetings, the engineer is familiarized with the process and becomes acquainted with license changes and coordination problems occurring in the market.

In addition to the reasons mentioned that might lead broadcasters to start coordination activity, one more must be added. Many broadcasters now travel outside their markets to cover news, sports, or public affairs events. In 1983 the FCC revised Section 74.24 of the Rules to make it easier for broadcasters to cover such events. A footnote to this Rules change asked that regional coordinators make themselves known to the Commission. The SBE NFCC was first formed to compile a list of coordinators to go beyond the Commission's request. The SBE NFCC updates this listing several times a year.

COORDINATOR CONSIDERATIONS FOR VARIOUS BROADCAST USES OF PART 74 SPECTRUM

Dispatch and Aural Remote Pickup

Channels for this purpose are currently available in a range from 26 MHz to and including 455 MHz.

In major markets experiencing congestion, and in other markets where older equipment is still in service, 26 MHz operation for dispatch and even some types of remote pickup are still being done. Lower power cuing systems for TV field operations sometimes appear in this band. Since it is near to present Citizen's Band channels, it is subject to interference from this service. Propagation conditions at 26 MHz include "skip," so it is common to experience interference in this band from hundreds, or thousands, of miles away at certain times of the year.

VHF dispatch and aural remote pickup are more common throughout the country. With the advent of excellent commercial narrowband FM mobile tranceivers and remotely operated base stations on hills and mountains, many radio and TV stations have used Part 74 VHF channels reliably for years.

Committees will be confronted with a number of seemingly conflicting uses and users as they sort through the years of accumulated problems in VHF and UHF bands. One guiding principle that has helped many committees is the order of priorities outlined by the FCC in Section 74.403(b). The transmission of remote pickup program material for broadcast, either live or delayed, takes precedence over other permissible uses.

VHF and UHF remote pickup with usable audio frequency response to 15 kHz is available with a variety of equipment. In many regions, interference from adjacent channels licensed to other services makes such operation difficult. A strong adjacent channel signal can "capture" an FM receiver, pulling the discriminator away from the desired signal. As channel use builds in a region, broadband white noise rises, making life difficult for wideband receivers.

Some equipment marketed for broadcast use was not designed for today's levels of channel congestion and sometimes makes the problem appear worse than it really is.

Transmitters should always be connected to antennas through ferrite isolation devices and harmonic filters to minimize generation of mixing products in solid state finals. This applies to repeaters, base stations, and mobile transmitters. This was not as important in the days of tube finals, but is critical in today's non-linear solid state world. Ferrite isolators are used for simplex transmitters. Ferrite circulators are used for repeaters.

Receivers can use combinations of attenuators, helical tuned circuits and cavity filters to protect against first stage overload from out of channel emissions. This is especially important since the 450 MHz Part 74 allocations are within one MHz of Land Mobile dispatch and paging bands. Paging transmitters operating at 250 watts commonly employ two-way sites also used by broadcasters and can greatly overload an unprotected receiver.

Broadband white noise at common sites is now being viewed as much of a culprit to degraded receiver performance as spurious emissions. As the number of transmitters at a site increases, the noise level at the site increases. This noise desensitizes all receivers at the site. If a user is operating a receiver at a common site used for high power paging transmitters that are on the air most of the time, the receiver may be trying to listen in a channel less than 1 MHz away from a 250 watt transmitter with a high gain antenna. The resultant 1000 watts (ERP) can cause problems for even the best equipment. Committees can educate members on this subject as well.

At "quiet" sites, receivers with GAsFET preamplifiers working with preselector/filter network receivers have been used successfully to increase receiver sensitivity or regain insertion losses from various filter elements. Careful design can prevent overloading the receiver with the add-on preamplifier.

Remotely located (satellite) receivers will give users better coverage and better immunity to various types of interference. It is desirable to locate satellite receivers at receive-only locations. Such locations can be selected after spectrum analysis shows that adjacent channel and out-of-band emissions are well within the receiver's ability to function at an optimal performance level. Receive-only locations are also good places to utilize high gain GAsFET preamplifiers. Local coordination groups may be of service in locating such sites, and maintaining their RF integrity. Getting the receiver closer to the transmitter is always preferable to increasing transmitter power as a technique to improve reliability. Leased program circuits can deliver audio back to the studio.

Committees can educate members on the pitfalls of common sites, and work with site managers to let them know of licensee concerns. There is a trend in many markets toward re-engineering common sites using transmitter and receiver combiners, and physical separation of transmit and receive antennas. These changes can improve the transmit and receive performance at most sites. Changes like this are expensive. A committee can educate users on the benefits of such projects, and help generate support to make them happen.

Committees can also make members aware of non-Part 74 licensing options for broadcasters who require two-way communications which are not related to Part 74 priorities such as the business band spectrum or cellular telephones. Cellular telephones have been used by many stations for some years now for coverage of news and even program length material. The FCC makes no distinction between cellular and wired telephones for the purpose of broadcast transmissions as of this writing. Stations should check, however, to see how and if broadcast use of cellular phones conforms with state telephone tariffs. Also, it could be interpreted that extended cellular transmissions from any site that can "lock up" a number of cells may be an abuse of the cellular system.

FCC and Federal Aviation Administration rules prohibit use of cellular phones on airborne aircraft. The FAA is developing rules to govern use of cellular phones while the aircraft is at the departure gate and/or experiencing an extended wait on the ground. Also, many airlines currently have their own rules restricting cellular use on aircraft—airborne or not.

The two blocks of UHF channels available to broadcast licensees are separated by 5 MHz. This is adequate for repeater operations, and many licensees in markets with hilly or mountainous terrain have employed this type of system to allow field units to communicate with the base no matter where they are. The convention of repeater pair frequencies is that the 455 group be used for repeater input and the 450 group should be used for the repeater output. Many markets have taken great pains to correct "reversed" situations for mutual benefit. Simplex channels have to be treated on a case-by-case basis. Repeater operation has the benefit of being truly wireless, since remote base systems usually depend on telephone company local channels for control and audio. Unfortunately, repeater operation requires two channels.

Spectrum efficiency in the UHF bands by splitting channels is an accepted mode of operation in some markets. Since it is possible to derive audio response beyond 5 kHz from a 25 kHz narrowband FM channel, many radio stations which want better audio quality for news reports can get it without using a 50 kHz R channel, provided the receiver used is well into full quieting. Committees can work with their members to make sure everyone is using the least amount of bandwidth consistent with needs.

Similarly, a station engaged in voice dispatch does not need a 50 kHz N1 channel. A 25 kHz N2 channel is more than adequate for this use. Responsible committees can help assure that use is matched to bandwidth.

Narrow band radio systems that exhibit better than 5 kHz audio response with 5 kHz FM deviation made possible use of "splinter" or offset channels where distance separations and relative transmitter power levels would permit. High performance receiver IF filters have been in use for this purpose since the early '80s, usually as a part of audio processing modifications to FCC type accepted transmitters that required an FCC modified type acceptance. Built on a firm base of existing rugged mobile equipment, such systems proved that further channel splitting was practical. Receivers designed to take advantage of the spectrum efficiencies of low FM modulation indexes are now coming to market. These receivers promise even greater quality improvements along with greater spectrum efficiency.

High Band VHF and UHF Coordination

Committees can sometimes arrange sharing of channels between users. This can even involve several licensees operating a community repeater system. This choice not only helps enhance spectrum efficiency, but may mean small stations can become part of a larger system they could not otherwise afford. The following are guidelines for new users of Part 74 communications:

1. The potential user should understand that he must be a Part 73 licensee, or a network (as defined in Part 74) to qualify for Part 74 channels.
2. Know that the committee really exists to help accommodate new users in bands already crowded—not to keep new users out.
3. The potential user should explain his needs during an open committee meeting, and publish what is said in the newsletter so everyone concerned can be aware of what is going on. It may be that someone at the meeting might make a suggestion to the potential user that could improve his system, or even accommodate his needs by some other means. For instance, cellular mobile telephones might be all someone needs to do ENG microwave setup if there are no dispatch channels left in the market.
4. The committee may "recommend" certain frequencies. At this point, it is up to the potential licensee to contact each co-channel and adjacent channel occupant and complete the transaction. It is NOT the responsibility of the committee to do this.
5. In the event the parties reach an impasse regarding an issue, the committee can and should act as a mediating force. Usually a compromise can be arranged.

In a congested RF environment, there is always the possibility of adjacent channel interference. Mobiles operating in the field can find themselves next to other mobiles transmitting on an adjacent channel. If this channel happens to be on the repeater output of the first station, the person in the mobile cannot communicate with its base while the other mobile has its transmitter on. If this is a chronic problem, one station may find it desirable to change frequencies. No amount of coordination will ever solve all of these problems.

There is at least one operating practice a committee can promote that will help: make it a standard in the market for each licensee to use the appropriate station identification and unit designator call sign. Not only is this in compliance with FCC Rules, but call signs make it easier to track down interference problems that will inevitably arise.

THE AURAL STUDIO-TO-TRANSMITTER LINK (STL) BAND

Licensees Operating in the 942–944 Band Segment

Licensees holding authorizations for this band segment dated prior to November 21, 1984, may continue to operate on a co-equal primary basis to other authorized stations and services. Stations in Puerto Rico may continue to be authorized on a primary basis.

950 Band Channel Stacking

A number of coordination groups have suggested standards and band plans. Channel stacking is authorized for this band. The FCC has published a list of authorized center frequencies for channel stacks in Section 74.502(b).

950 Band Recent Developments

In January 1990, the SBE filed a rulemaking petition with the FCC proposing to extend minimum antenna standards and minimum path length requirements to stations in the 950 Aural STL/ICR band. The same category A and B antenna requirements now in place for Private Operational Fixed Microwave Service (POFS) stations in the 952–960 MHz band were proposed by the SBE.

Category A half power beamwidth – ±7°
Category B half power beamwidth – ±10°

Note: Limits apply to parallel polarized antenna performance. No limits apply to cross polarized antenna performance.

A minimum non effective isotropic radiated power (EIRP) restricted path length of 22 kilometers (13.7 miles) has been proposed by the SBE, along with a maximum transmitter output of 10 watts. A 0 dB desired-to-undesired ratio to first adjacent channel stations and a 50 dB ratio to existing co-channel stations was proposed to define a technical standard for 950 aural STL/ICRs. Readers are advised to refer to the latest edition of Part 74 Subpart E, for the most recent regulations.

Wireless Microphones

Wireless microphones are becoming more common and should not be overlooked in the coordination process. Coordinators are often asked to act as a clearinghouse for major events where a large number of such mikes are in use. Wireless microphone operation is a permitted use in the band specified in Part 74.802.

Non-RF Alternatives

There are non-RF alternatives available for STL systems. Laser links are feasible for very short distances where there is little likelihood of any interruption of the line-of-sight path.

Many stations still rely on program channels leased from their local telephone company. For short distances that do not involve very many telephone exchanges, this may still represent the best solution for some stations. This may be true especially in markets that have reached STL saturation.

TV MICROWAVE CONSIDERATIONS

Television has come to rely on either their own or common carrier microwave links to interconnect their studios with their transmitters and to facilitate live remote broadcasts. Rules governing links operated by broadcasters are found in Part 74 of the FCC Regulations and cover frequencies from 2, 2.5, 6.4, 7, 13, 18, 21, 23, and 40 GHz bands.

Microwave transmissions can accommodate one or more subcarriers imbedded in the video signal to carry mono or stereo television sound, aural service for a co-owned AM or FM station, program cuing, telemetry, or a combination of these elements. In congested markets one should use the lowest possible subcarrier frequency combinations to keep occupied bandwidth as low as possible.

In major markets the number of stations engaged in field operations exceeds the number of available channels. While spectrum is available in the 2, 2.5, 6.4, 7, 13, 18, 23, and 40 GHz bands at this time, many stations and coordinating committees in these major markets have opted to make 2 GHz the primary band for ENG. However, this band will have some level of STL and ICR activity for the foreseeable future due to its long haul propagation characteristics.

The use of 2 GHz for ENG with its better propagation characteristics means stations have to install fewer ENG receive locations in their service area. Path lengths up to 50 miles can be operated reliably and under optimum conditions, including low humidity, path lengths up to 100 miles may be possible.

In many markets, there is intense competition for the seven primary 2 GHz channels (1.9 through 2.1 GHz) and for the 2 (+1 grandfather) channels at 2.5 GHz. The 2.5 GHz band is shared with other users such as the industrial, scientific, and medical (ISM), service and the private operational fixed microwave service. This factor has contributed to the rise in frequency coordinating committee formation and activity.

Stations having 2 GHz STLs have been, for the most part, relocating to 7 or 13 GHz for several reasons. First, they may themselves be involved in 2 GHz ENG. Second, a fixed link in this band has a high potential risk of interference from mobile operations. Since modern TV ENG portable equipment is frequency agile, and non technical personnel are more involved in field operations, this is a very real risk.

Coordination for ENG is being done more and more on a "real time" basis. Coverage of breaking news cannot be planned. Coupled with the limited number of 2 GHz channels, coordination will only be more difficult in the future.

SUGGESTED OPERATIONAL GUIDELINES FOR TV ENG OPERATIONS

1. Form a Microwave Subcommittee. Compile a list of persons responsible for ENG technical and operational decisions at each station. This list should contain telephone numbers providing access 24 hours a day, should be circulated to all stations involved in ENG operations, and should be updated as needed.
2. The Microwave Subcommittee should be a forum to devise a Home Channel Plan for the region. The purpose of this plan is to give each user a primary channel for his operation. Alternate polarization is assigned to each channel. Split channel operation should be the long range goal for the market. Every user is furnished with a copy of this plan for reference.
3. As part of this plan, each user should set a goal to become frequency, polarization, and power agile.
4. Once the Home Channel Plan has been developed, a method of coordinating this plan in real time is needed. This can be accomplished by a hard wired ring down network between ENG control points or by using a two-way radio system with base stations located at the control point of each ENG license.

   Each ENG control site should have a direct private phone line to which all ENG users have access as a minimal requirement. This has the disadvantages of requiring several phone calls to track down real time interference, or request use of time on a channel, but will satisfy minimum needs.
5. Each licensee should set a goal to equip the ENG operation at the following level:

   A. Mobile transmitters should be frequency agile.
   B. Each transmitter shall have continuously variable control of output power.
   C. Each mobile truck should ideally be equipped with an antenna with excellent sidelobe suppression of the "Silhouette" type.
   D. Each mobile truck should be capable of changing the polarity of their feed.
   E. Each truck should have a means of identifying its test bars, and ideally be able to identify its transmissions in the vertical interval. If this latter type of equipment is employed, the ID is viewed by unlocking the vertical hold control of a video monitor. (This signal must be stripped by a processing amplifier prior to on-air broadcast!)
   F. Use frequency agile ENG receive sites with remote steerable antennas.
   G. Employ multiple receive sites.
   H. Use steep-skirted channel filters ahead of receivers.

6. Truck crews are instructed never to light up until they receive permission from their control point, even on their home channel.
7. Truck crews should receive careful instruction on polarity settings, using only enough power for the path involved, and in the overall operational agreements with other stations engaged in ENG activity.
8. Stations normally stay on their home channel. They use the ENG control point communication system to request different channels for extra activity, to eliminate inevitable adjacent channel interference during some operations, or for paths that conflict on their home channel with Inter-City

Relay (ICR) links, or STLs. The request is made of the station which is assigned the desired channel as its Home Channel.

9. Remember, no one really "owns" a channel. Anyone might be called upon to share their channel on a real time coordination basis.

The plan must enjoy unequivocal support from all levels of management. It is imperative that this support be articulated and understood throughout the entire ENG operation, both in technical and non-technical areas, verbally and in writing.

All reasonable requests for temporary sharing of channels must be accommodated.

Spectrum use for all major events should be preplanned. When advance notice of a major event is given, a coordinator is assigned. Requests for channel assignment are solicited, operating channel assignments are made, and contingencies considered and addressed.

Examples of this would include community celebrations of a regional or national interest, a papal visit to a particular market, or the opening in New York of the United Nations General Assembly. On a larger scale, examples would include Olympic coverage or political conventions. The foregoing are documented examples of successful coordination.

The 7 GHz Band

The 6.875 to 7.125 GHz band of ten channels is typically used for STLs and ICRs. Some use for remote pickup (RPU) has taken place and it has proven to be very satisfactory. The bulk of all STLs throughout the country are in this band. There is some shared activity possible in this band with common carriers in the 6.425 and 6.525 GHz range.

The 13 GHz Band

The 12.7 to 13.25 GHz band with 44 channels (including offsets) is used also for ICR and STL activity. It is also being used more and more for low power "window" microwave to relay signals from isolated cameras in and out of the studio. (Fig. 1). This band is shared with the cable television relay service (CARS) and with private operational fixed service users on a co-equal basis. The chief liability of this band (and any microwave band at approximately 8 GHz or higher) is degradation due to rain and snow attenuation. Careful engineering is required to insure uninterrupted service, particularly for fixed links.

Fig. 1. Mobile communications gear used on a typical KFWB news vehicle. (Photo courtesy KFWB Radio)

TV STL COORDINATION

Minimum Antenna Standards

TV STL systems licensed as of October 1, 1981 were given a ten year grandfather period to comply with minimum antenna standards established in FCC Docket 21505. Commission action in Gen. Docket 82–334 extended minimum antenna standards to the 2 GHz, 7 GHz, 18 GHz, and 31 GHz TV STL/ICR bands. Although October 1, 1991, would have been the compliance deadline, the FCC extended the deadline to April 1, 1992.

The minimum standards for "Category A" and "Category B" antennas are found in Section 74.641 of the Rules. The FCC had once proposed that Category A antennas be used in "congested areas." However, due to several difficulties, including the problem of defining a "congested area," the FCC abandoned this plan and, in September, 1991, decided not to define any geographic areas as being "congested." Thus, no stations will be required to employ Category A antennas. However, all stations, by April 1, 1992, must employ at least Category B antennas on all fixed microwave links.

Minimum Path Length

A second requirement for minimum path lengths adopted in 1987 in Gen. Docket 82–334 allows short paths at lower microwave frequencies as long as the effective isotropic radiated power (EIRP) be reduced in accordance with formulas stated in Section 74.644(a). After April 1, 1992, grandfathered links will become subject to the same minimum path length requirements that links licensed since 1987 have had to meet. These links must reduce their EIRP for paths shorter than 17 kilometers (10.6 miles) at 2 and 7 GHz.

If the path is shorter than 5 kilometers (3.1 miles) for 13 GHz, EIRP must also be reduced if a new licensee's operation would be hindered. There is no exclusion from this requirement for new licensees.

Power Limitations

Transmitter peak output for fixed stations should not be greater than necessary and should never exceed the power specified in Section 74.636.

Microwave Licensing Considerations

Form 313 is used to apply. Instructions are included with blank forms. A license is always required for an

STL or ICR. The FCC currently issues a "one step" Construction Permit (CP) and license. Contact with a local coordinating committee has not been made mandatory by the Commission.

Temporary Operation for TV Microwave RPU

As in the aural services, temporary operation away from a station's service area without a special license is permitted for up to 720 hours or 30 days per year under Section 74.24 of the Rules. As mentioned earlier in this chapter, the Coordinator Listing compiled and maintained by the SBE can make coordination for this type of operation less painful. The coordination burden is still on the licensee who is visiting not to cause interference to the local stations.

During this type of operation, visitors use their broadcast on-air call sign, not their Part 74 call sign. They should work closely with the local coordination group. Confirm all communications in writing.

A visiting broadcaster should also obtain an accurate database to determine existing cochannel and adjacent channel users. If such a database is not available through a local coordinator, it is incumbent on the temporary user to accurately ascertain who would be affected.

It is important to notify existing co-channel and adjacent channel users of the intended link giving transmitter power output, dates and times of operation, and antenna type and gain information. Exchange names, addresses, and 24-hour telephone numbers.

An engineering study of the path should be performed. This should include detailed path calculations. Several computer programs using accepted path consideration algorithms are available from a number of sources, including the SBE National Frequency Coordination Committee.

A field spectrum analysis is a valuable tool for engineering any link, fixed or temporary. It may uncover licensees who have not been entered in the database, newly authorized construction that may interfere with the path, temporary users who have not notified affected licensees or the local coordination group, Section 74.24 users, or multipath from co-channel links on different paths.

This rule is primarily designed to cover special events and should not be substituted for filing licenses.

Major Event Frequency Coordination for RF Equipment

The FCC has traditionally recognized ad-hoc coordination groups for many years for events such as the national political conventions, major sporting events such as the Olympic Games and the America's Cup races, and news events like NASA Shuttle launches and landings.

Such events require a great deal of planning and hard work. Preparation for the national conventions typically begin 24 to 18 months in advance, although no substantive work can commence until the city sites are announced.

The FCC's normal course of action for such events is to suspend Section 74.24 of the Rules for a period before and during the event. This has the effect of requiring all Part 73 licensees wanting to use RF equipment of any type at the event to contact and coordinate with other users via the ad-hoc coordination group.

The Society of Broadcast Engineers, through its Washington, D.C., attorney will assist local groups with making the proper request and showing to the FCC for such a waiver to be issued for an event.

If an event takes place in an enclosed venue, it may be possible to use channels in use by local broadcasters outside the venue. Using very low power transmitters is an excellent way to allow a number of users to coexist inside such a venue. A large number of handheld transceivers can often be used in adjacent channels by the simple expedient of removing the "rubber duck" antenna and substituting a 50 ohm termination resistor. TV microwave transmitters should be cut back or attenuated to 50-11 milliwatts. If all users cooperate, receivers in the various bands being used will be quite happy, and the needs of all users will be met with little or no interference.

Weekend events may make it possible to use TV and radio ENG channels only used during weekdays. Compatible sharing should take place in the spirit of good cooperation between engineers.

Assuring all equipment used for such events is working properly is an important consideration. While this is the responsibility of the licensee, the coordinating body can sometimes help them deal with this burden. Various techniques have proven to be effective. For example, the FCC usually sends one of its sophisticated trucks to the major political conventions to look for spurious and unauthorized emissions. NASA and the United States Air Force usually check all equipment as it comes into Edwards Air Force Base for Shuttle landings. NASA requires everyone to participate in a "full light up test" two hours before each landing. For smaller events, local groups commonly borrow or rent a spectrum analyzer as a service to guest broadcasters.

If it is not possible to accommodate all users within the existing Part 74 bands for major events, the FCC has granted waivers in the past for one or more unused TV UHF channels, usually for two-way dispatch. This is not possible in all markets. When this spectrum is finally allocated, this option will no longer be available.

Whatever the event, the primary goals of special event coordination are:

1. Protect the existing base of local users.
2. Accommodate itinerant users to the greatest extent possible.
3. Be present at the event to help unsnarl the inevitable "glitches" that will happen.
4. Report to the industry and the FCC after the event so all can learn for the future.

The Benefits of Coordination for Microwave Licensees

As in the case of all Part 74 coordination, committees exist to help new and existing licensees share very limited amounts of spectrum. The move to frequency agility has made real time coordination possible for mobile operations. It mandates cooperation and sharing. In return, it allows licensees to operate their field ENG equipment and STLs with reduced risk of interference.

The Risks and Penalties of Not Coordinating

Failure to coordinate can be an expensive proposition for those who do not participate in this voluntary process. Uncoordinated mobile and fixed links have been taken off the air and re-engineered when they interfered with licensees operating under rules and procedures they agreed to abide by along with other broadcasters in the market. Precedent has been set for FCC support of local agreements where consensus is evident.

Expensive recrystalling, reprogramming, and retuning, antenna and site changes may be the only way to correct interference problems, most of which could have been avoided by communication with the coordinating committee and other users prior to their installation.

Because the broadcasting and cable industries are not monolithically structured, frequency coordination techniques cannot be identical in each market. Only general suggestions can be given as to what each manager or news director can do to participate in and support the coordination process. However, it is clear that coordination cannot succeed without the active support of general managers and news directors at whatever level is practical in a given situation.

INTERFERENCE AND THE PART 74 LICENSEE

Interference to a Part 74 licensee's operations may come from almost any source of RF one cares to mention. Radiation from unshielded computer cables has interfered with wireless microphone operation, for instance. Channel interference from other broadcasters can be thought of as a problem that can be solved by additional attempts at coordination, followed, if necessary, by mediation by a neutral party.

Locating the source of spurious emissions can be a real challenge. The committee can offer some advice and support when a suspected case of interference from an out-of-channel source is reported. Quite often, other stations in the market can make spectrum analyzers available to eliminate the possibility of a mixing product that is actually generated in the front end of the receiver due to an overload condition. This type of problem is common for both voice and video equipment.

Older solid state receivers are sometimes quite prone to front end overload. When an overload-prone receiver is operating at a common site with an appreciable amount of adjacent channel activity, the receiver may be rendered useless. Filtering and attenuator pads can sometimes be suggested as cures. Video microwave receivers operating with preamplifiers can overload easily as a result of proper adjacent channel operation. In such cases, a suggestion to provide a bypass switch is in order.

Committees should encourage licensees to use call letters and unit identifiers for all voice transmissions. All video ENG trucks should be able to transmit bars with ID, and ideally should be able to place ID in the vertical interval as well as for identification during program transmission.

Tracking unidentified interference is not unlike detective work. For video, tuning through the on-air signals may provide a clue. If bars with no ID are seen at an ENG receive location, triangulation may be possible if the interference can be seen by two other locations equipped with steerable dishes. Trucks have been identified by reading street signs that appear in shots. Talent on other stations can sometimes be a good means of identification.

Interference to aural links can be very challenging to trace. If call letters are not being used, anything heard then becomes a clue. If a street address is mentioned, a newsroom "criss-cross" telephone directory can be used to find a telephone number at or near the address. A call to the number identifying yourself, your reason for calling, and if a taxicab or delivery truck has been there (or will be there) often works wonders.

Land mobile transmitters are sometimes designed so the final has power applied to it at all times. Any instability in the final can start it oscillating. These transmitters can sometimes migrate through many megahertz of spectrum, almost at random. When their exciters are keyed, they return to their proper channel. Sometimes only a spectrum analyzer and a lot of patience can trace this type of problem.

Inexpensive direction finding equipment is available for VHF and UHF tracking. Some units work with any hand-held portable. For unmodulated carriers and other types of interference that are present for long periods of time, direction finding can be a very effective (though time consuming and expensive) process.

If the licensee(s) involved, even with committee guidance, are unable to locate the interference, a call to the FCC may be in order. You will be asked for all pertinent information, such as time the interference occurs, its nature, duration, and any other details you might have learned. They generally will not take interference complaints unless you can supply this information. Remember that coordinating committees and their members are never to act as police in such matters.

Curing Interference

Dealing with the broad topic of interference cures is beyond the scope of this chapter. When committees are working to solve such problems, creativity is definitely an asset. For instance, in a case where an aural STL on an adjacent channel on a parallel path was causing interference, the cure that was tried and worked involved installing a larger dish at the transmitter. The dish was aimed slightly off azimuth from its receiver so that the receive site being interfered with appeared in a deep null of the interfering transmit antenna.

MORE INFORMATION

For more information on frequency coordination, including a subscription to the SBE National Frequency Coordinators' Listing, or information on starting a local coordinating committee, contact the Society of Broadcast Engineers, P.O. Box 20450, Indianapolis, Indiana 46220. (317) 842-0836.

1.4 Frequency Allocations for Broadcasting and the Broadcast Auxiliary Services

William R. Meintel
Datel Corporation, Arlington, Virginia

INTRODUCTION

This chapter provides an overview of frequency allocations[1] and a listing of the frequencies available for AM, FM, and TV broadcast stations as well as for the auxiliary broadcast services that support broadcasting operations. Necessarily, such an overview needs to include a description of the decisionmaking process that is involved in allocating a frequency band for a specific purpose.

This complex subject of allocations involves more than just the location of the service in the frequency spectrum; it includes decisions as to the number and width of channels, power and antenna limitations, as well as decisions concerning the technical standards that define how the spectrum will be shared with other users. Because each broadcast service presents unique service and interference objectives, the allocation process has been and will continue to be different for each of these services.

Here it is only possible to provide a brief description of the allocation process before turning to the current situation for each of the services. Where appropriate, the discussion will touch on changes expected in the foreseeable future.

HOW SPECTRUM IS ALLOCATED

It is a fundamental characteristic of radio wave propagation that radio waves follow the laws of physics and thus ignore political or geographic boundaries. As a result, decisions concerning radio frequency allocations cannot be made solely at the local level but must take into account their anticipated impact outside the station's coverage area. Recognizing that coordination in the allocation and use of spectrum is essential, an international mechanism has been established to perform this function, and in the United States, a parallel coordination system has been established at the federal level.

International Allocation

At the international level, frequency allocation decisions are made by the International Telecommunication Union (ITU), a specialized agency of the United Nations, headquartered in Geneva, Switzerland. Like the United Nations, the ITU is a consortium of more than 160 governments, whose purposes are to propose, develop, revise, and administer worldwide frequency allocation plans. Clearly international cooperation will minimize interference and maximize use of the spectrum.

Worldwide allocations for radio services are made through international conferences called World Administrative Radio Conferences, or WARCs. Through a process of give and take, the WARCs make decisions on how to allocate spectrum. General WARCs that encompass the use of the entire radio frequency spectrum are held at roughly 20-year intervals; the last such general WARC took place in 1979, and considered over 15,000 individual proposals dealing with numerous aspects of world telecommunications. Other WARCs are also held as necessary to consider more specific topics of concern on a worldwide basis.

The ITU also convenes Regional Administrative Radio Conferences (RARCs) to consider questions that are unique to a specific ITU Region. Often these RARCs consider implementation of decisions made at an earlier WARC.

Technical Submissions

In addition to allocation proposals and related submissions by ITU member countries, technical input for WARCs and RARCs comes from the International Radio Consultative Committee (CCIR). The CCIR is the arm of the ITU that develops recommendations and provides reports dealing with technical standards and operating procedures. Its work is conducted through 12 study groups,[2] each of which is devoted to a particular radio communication service or specific technical issue.[3]

Overall coordination for the work of the CCIR is provided by its director, elected by the ITU members, who functions at the ITU headquarters in Geneva. The work of two study groups is relevant to broadcasting: Study Group 10 (sound broadcasting) and Study Group 11 (television broadcasting). Under both of these study groups, there are subgroups to deal with specific areas of interest. There also are joint study groups that deal with questions that are of relevance to more than one study group. Each study group has an international chairperson and one or more vice chairpersons who are provided by interested participating administrations (participating governments).

ITU and CCIR work is carried out not only by the members of the ITU, but also by recognized users and standard setting groups such as the European Broadcasting Union and the U.S. broadcasting networks. Within individual administrations, organized structures often are created to provide input for the administration to submit to the CCIR.

In the United States, this activity is chartered by and operates under the Department of State (DOS). There is a U.S. national chairman appointed by DOS, two national vice chairmen from the Federal Communications Commission (FCC), and one from the National Telecommunications and Information Administration (NTIA). DOS also appoints a U.S. national chairman for each of the individual study groups. Each is responsible for the work of the group and heads the U.S. delegation to international study group meetings. In the United States, CCIR activities are open to participation by the public. Contributions to CCIR work in the United States come mainly from the private sector. Elsewhere in the world, CCIR work is performed primarily by government employees.

In addition to providing technical input timed to coincide with scheduled WARCs and RARCs, the CCIR's basic work is carried out in four-year cycles. Three international CCIR meetings are normally held; the first two meetings, usually held at the 18- and 36-month points, are for the individual study groups and are referred to as the *interim* and *final* meetings of the study group. The four-year cycle concludes with a combined meeting of all the study groups, referred to as the plenary meeting. At the plenary meeting, all documents from the study groups are submitted for approval; an outline of work for the next cycle is produced, and the CCIR director and study group chairmen for the next cycle are elected. In addition to these meetings, provision is made for extraordinary meetings to address issues of critical importance that cannot be resolved during the regular study group meetings.

Administrations may suggest any matter of interest for study that is within the purview of the CCIR, but priority is given to issues that are relevant to a scheduled WARC or RARC.

The cumulative recommendations and reports of the CCIR are updated and published at the end of each four-year cycle in books commonly referred to as the Green Books. These contain a great deal of information and recommendations.[4] The material relates not only to spectrum allocation but includes related information concerning measurement procedures and standards for audio and video recording equipment used for the exchange of broadcast programming.

The Allocation Process

The allocation of radio frequency spectrum occurs as a result of a series of interrelated decisions. On the first, most basic level, blocks of frequencies are allocated on a worldwide basis by the ITU WARC process. Exactly which blocks of frequencies are allocated to particular services is determined by evaluating the many specific proposals submitted to the WARC for each frequency band. Technical input is obtained from propagation studies and other engineering analyses undertaken as part of the CCIR process and from the submissions of individual administrations. Frequently, blocks of spectrum are allocated for the same purpose on a worldwide basis, but may also be allocated for different purposes on a regional basis.

Member nations theoretically retain the sovereign right to domestic use of the spectrum so long as such use is not in contravention of the International Radio Regulations or the international agreements to which that administration is a party. However, as a practical matter, the flexibility of administrations to use the spectrum is limited by the worldwide allocation system and the need to avoid harmful interference.

Unlike the FCC, the ITU does not license users of the spectrum. Instead, it operates only as a coordinator, maintaining a Master International Frequency Register (MIFR or Master Register) of radio stations worldwide. The agency within the ITU responsible for that activity is the International Frequency Registration Board (IFRB). Member administrations have agreed to provide notifications to the IFRB of the new stations or modifications in existing stations operating within their respective countries. The IFRB studies these notifications for compliance with the existing world or regional agreements and provides the results of its studies to the member nations. Only those notifications that comply with the existing agreements are placed in the Master Register of stations. Once a station has been placed in the Master Register, the member nations are obliged to provide it with the internationally agreed level of interference protection.

Because the WARC and RARC agreement texts provide only a general framework, many specific matters are left to individual nations to resolve and implement. In the U.S., the FCC and NTIA share the responsibility for implementing agreements to which the U.S. has assented, including the bilateral or multilateral agreements negotiated with our neighbors.

There are differences in the treatment of the various broadcast services. Because AM signals propagate over great distances, international decisions have a much greater impact on AM broadcasting than FM and television, whose VHF or UHF propagation is much more limited. This means that the restrictions on FM and TV allocations imposed by international agreements are usually applied only to areas near the borders. However, international agreement on technical transmission standards often is desired in order to foster the absence of interference and the worldwide free flow of communication.

Domestic Allocation Process

Regulation of spectrum began with the U.S. Department of Commerce in the early 1920s, when the Secretary of Commerce granted the first AM broadcasting licenses. By 1927, the number of AM stations had increased to 733, and over six million radio receivers had been manufactured. However, because of an unfortunate court decision that precluded the Secretary of Commerce from dealing with the specific choice of location, power, and operating frequency, these matters were left largely to the discretion of the broadcaster. This led to a chaotic use of the spectrum, with widespread interference, a situation that led to the creation of the Federal Radio Commission in 1927. Seven years later the Federal Radio Commission was replaced by the Federal Communications Commission (FCC), formed pursuant to the Communications Act of 1934. Ever since, anyone desiring to operate a broadcast station, or most any kind of radio transmitting device, must apply to the FCC and be granted a license before commencing operation. Today, a broadcast license sets forth all essential technical parameters of station operation.

The NTIA performs a similar function and coordinates the spectrum used by government agencies. The FCC works with the NTIA where there is a need for coordination between government and private uses of the spectrum.

Spectrum for domestic use must be allocated by FCC rule making proceedings. The FCC can initiate domestic allocation rule making proceedings on its own motion or in response to requests from the public, but in so doing it must not contravene international agreements to which the United States is a party. The FCC rule making process is a complex subject in its own right, but for present purposes, only a brief description is required. Rule making proceedings are based on a public record developed through responses to the issuance of a Notice of Proposed Rule Making that are filed by interested parties. In addition to filing comments on the FCC's proposal, the public may reply to the comments of other parties. After the Commission evaluates the responses to its Notice, it may decide to either adopt the proposal as originally set forth, modify it, based on the comments received, or possibly reject the proposal.[5]

Many FCC proceedings are controversial in nature. Allocations proceedings may be especially controversial, since a particular communication industry's livelihood may depend in part on how much spectrum is allocated. Thus, FCC allocation decision making is not simply a matter of technical evaluation but must be seen as part of the political process as well, as happens when entire industries compete for a limited amount of spectrum. Where the number of users in a particular frequency band is expected to be relatively small or their use sporadic, the FCC may propose sharing of this spectrum with other users. Such proposals may also be controversial, since sharing spectrum with a dissimilar service invites the possibility of interference and difficulties in coordinating the use of the frequencies. Because of these and other factors, FCC allocations proceedings consume a great deal of its time and energy and can impose burdens on the organizations that participate in them.

From time to time, alternative methods have been proposed for allocating spectrum. These ideas usually envision the removal of the FCC as the arbiter of mutually exclusive requests for spectrum and instead substituting marketplace forces. Under a market allocation system, frequencies would become used by entities who would pay for them; noneconomic, social, or public policy factors would not be considered. The Commission's role would be reduced to that of a technical "traffic cop" of the airwaves. To date, this approach has not been significantly implemented.

AM BROADCASTING FREQUENCY ALLOCATION

In the United States, amplitude modulated (AM) stations operate with carrier frequencies in the center of 10 kHz channel bandwidths. Until recently, most people thought of the AM band as including 107 channels in the band from 535–1605 kHz, as has long been the case in the United States. However, the frequency band allocated to AM broadcasting actually includes a total 118 channels, each 10 kHz wide, in the band 525–1705 kHz, as shown in Fig. 1. As of this writing, the lowest channel (530 kHz) is not used for conventional broadcasting in the United States but is used solely for Travelers Information Stations (TIS). A recent FCC Rule Making has made the 10 channels from 1605 kHz to 1705 kHz available for use in the very near future.

Currently, there are over 4,800 commercial and noncommercial AM stations operating in the United States. These stations operate with various powers levels, up to a maximum of 50 kilowatts. About half of these stations use multitower directional antennas to restrict radiation in certain directions, for the purpose of controlling interference or maximizing radio service in particular directions.

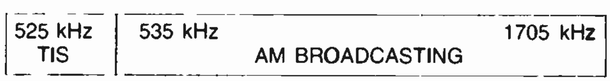


Figure 1. AM broadcasting band, including the Travelers Information Service (TIS).

Allocation decisions for the AM broadcast band are the most complex of the broadcast services. Since propagation varies with time of day, geographic latitude, and frequency, the engineering analyses necessary to establish interference protection for other stations can be quite complicated. Engineers, the FCC, and the IFRB have sophisticated computer programs that analyze the input of a new or modified AM station proposal. Before going into the details of the AM broadcast allotment system currently in place in the United States, it is necessary to provide a brief history of recent AM broadcasting allocations.

From the beginning, the countries in the North American area recognized the need to cooperate in the use of AM frequencies, and in 1937 they reached agreement on how to proceed. Soon, however, this agreement was found to be inadequate. Negotiations began on a new agreement. Although the North American Regional Broadcasting Agreement (NARBA) was signed on November 15, 1950, it did not go into effect until 10 years later, on April 16, 1960. Signatories to NARBA include the United States, Canada, Cuba, the Dominican Republic, and the United Kingdom on behalf of Jamaica and the Bahama Islands. Mexico, an earlier participant, removed itself from these negotiations, and a bilateral agreement between the U.S. and Mexico was reached in 1957. It was superseded by a new U.S.-Mexican agreement in 1968.

These international agreements became necessary principally because nighttime AM propagation has the potential for causing widespread interference to neighboring countries unless mutual allocations criteria and related technical standards could be agreed upon and implemented. To this end, NARBA provided for a partitioning of AM broadcast channels into three basic classes. The first of these are the so-called *clear* channels, whose high-powered stations would have primary access to the frequency. Other stations could use the channel subject to providing full protection to the clear channel's dominant station(s). Clear channel dominant stations were designed to provide service over extensive areas by means of skywave as well as ground wave signals. NARBA set aside 60 of the 107 channels then available for clear channel use. Each NARBA country, except Jamaica, received a priority on one or more clear channels, with the U.S. receiving a major portion of available priorities. NARBA countries without a priority on a given clear channel could still assign stations on that channel, provided that these stations protected the area-wide service of the dominant station in the country with the NARBA priority.

The second class of channels were called *regional* channels, and these occupied an additional 41 channels. Unlike clear channels, these channels were shared on a equal basis by all the NARBA countries. Stations operating on these channels were intended to provide service to a considerable area, but, unlike the clear channel stations that received protection for their skywave as well as groundwave service, only the groundwave service provided by the regional stations was protected.

The remaining six frequencies were the *local* channels that provide an even more limited type of groundwave service. Only limited interference protection was provided to these relatively low power operations.

Recognizing the need for updating these agreements and for developing more efficient coordination throughout ITU Region 2 (North and South America), the 1979 General WARC called for a conference to be held in Region 2 to address AM broadcasting and sharing criteria. That conference was held in two sessions in 1980 and 1981 and resulted in the adoption of an agreement among most of the countries of the hemisphere. Included as part of that agreement, referred to as the 1981 Rio Agreement[6] (for Rio de Janeiro, the location of the second conference session), was a list of all of the operating stations in the hemisphere along with information indicating whether or not the stations were receiving or causing harmful interference according to the technical criteria set forth in the agreement. Stations not causing interference were placed in the IFRB Master Register and accorded protection from interference as defined by the agreement. In situations where interference already existed, the countries involved were asked to meet and work out mutually satisfactory solutions.

Because the general framework of the Rio Agreement did not deal with the particular needs and desires of the United States, Canada and Mexico, separate new bilateral agreements have been negotiated which incorporate the required additional items concerning coordination and technical parameters. Although the Rio Agreement applies throughout most of Region 2, as of this writing, relations with the Bahamas and the Dominican Republic continue to be governed by NARBA, as neither country has taken the necessary steps to replace the NARBA provisions with the Rio 1981 Agreement. Relations with Cuba regarding the 535–1605 kHz band are governed solely by the international Radio Regulations rather than by agreement.

The 1981 Rio Agreement changed the NARBA station classifications. No longer are the channels themselves classified. Stations are now classified without regard to the channel on which they operate. Stations providing wide-area service, both groundwave and skywave, are now designated Class A stations, while stations providing the equivalent of the regional and local services are designated as Class B and C respectively. The new agreement permits any class of station to operate on any channel so long as it provides protection to other stations based on their classification.

Over the years, there have been many changes in the nature of AM broadcasting in the United States. In early years, clear channel stations provided the only service available in many areas of the U.S., but with the end of World War II, demand increased greatly and many AM stations were established in all areas of the country. Because of its early development of AM radio service, the U.S. experience has been used as a model for the regional and bilateral agreements. Figs. 2, 3, and 4 show the classification system that exists today. Until 1990, for domestic purposes, the U.S. continued to use the old method classification of stations based on frequency, rather than the new system adopted in the Rio Agreement. In 1990, the FCC proposed to align the U.S. domestic classification system for AM stations with the 1981 Rio Agreement system. The current status of AM station classification is contained in Section 73.21 of the FCC Rules and Regulations.

The AM broadcasting system in the United States has continued to evolve as the demand for more stations has grown. One result of the demand for facilities is the increased use of directional antennas to provide required interference protection while enhancing coverage in other directions.

Another post-war development was a large increase in the number of daytime-only stations. These are stations authorized to operate only during daylight hours. Since propagation conditions during these daylight hours do not normally support significant skywave transmission, there are many locations where a station can operate during the daytime without causing harmful interference to other stations. Based on this concept, the FCC over the years licensed approximately 2,500 stations for daytime-only operation.

Recognizing that daytime-only stations (and even some full-time stations that operate with restrictive directional antenna patterns during the night) are unable to provide effective service during early morning hours, the FCC originally allowed these stations to operate during this period so long as no interference complaint had been received. Ultimately, this proved to be unworkable, and a more formalized approach was adopted. The FCC began granting Presunrise Authorizations (PSRAs), permitting many of these stations to operate with their daytime facilities with powers up to 500 watts during the presunrise period between 6:00 AM and local sunrise. While some interference occurred, the FCC believed the interference was balanced by the public's need for local informational services during this very important morning time period.

While presunrise operation did provide some relief, it did not end the economic problems many stations faced in effectively competing with full-time stations. The FCC was pressed to provide relief in the form of post sunset operation for daytime-only stations. With the removal of international impediments, the FCC did provide such relief.

International classes of AM stations	Corresponding U. S. classes prior to 1990	Classes of Channels available in U.S. for each class
Class A	I-A	Clear channels
	I-B	"
	I-N	"
Class B	II	"
	II-A	"
	II-B	"
	II-C	"
	II-D	"
	II-S	"
	III	Regional channels
	III-S	"
Class C	IV	Local channels

Figure 2. International and domestic classifications of stations and channels.

Clear Channels

Frequency	Class of Station*	Frequency	Class of Station*	Frequency	Class of Station*
640 kHz	I-A	650 kHz	I-A	660 kHz	I-A
670 kHz	I-A	680 kHz	I-B	700 kHz	I-A
710 kHz	I-B	720 kHz	I-A	750 kHz	I-A
760 kHz	I-A	770 kHz	I-A	780 kHz	I-A
810 kHz	I-B	820 kHz	I-A	830 kHz	I-A
840 kHz	I-A	850 kHz	I-B	870 kHz	I-A
880 kHz	I-A	890 kHz	I-A	940 kHz	II
1000 kHz	I-B	1020 kHz	I-A	1030 kHz	I-A
1040 kHz	I-A	1060 kHz	I-B	1070 kHz	I-B
1080 kHz	I-B	1090 kHz	I-B	1100 kHz	I-A
1110 kHz	I-B	1120 kHz	I-A	1130 kHz	I-B
1140 kHz	I-B	1160 kHz	I-A	1170 kHz	I-B
1180 kHz	I-A	1190 kHz	I-B	1200 kHz	I-A
1210 kHz	I-A	1500 kHz	I-B	1510 kHz	I-B
1520 kHz	I-B	1530 kHz	I-B	1540 kHz	I-B
1550 kHz	II	1560 kHz	I-B		

* In addition to the Class I-A or Class I-B stations that can be assigned to the above channels, various Class II stations may also be assigned.

Regional Channels

(Class III or III-S stations are assigned to these channels)
550, 560, 570, 580, 590, 600, 610, 620, 630, 790, 910, 920, 930, 950, 960, 970, 980, 1150, 1250, 1260, 1270, 1280, 1290, 1300, 1310, 1320, 1330, 1350, 1360, 1370, 1380, 1390, 1410, 1420, 1430, 1440, 1460, 1470, 1480, 1590, and 1600 kHz

Local Channels

(Class IV stations are assigned to these channels)
1230, 1240, 1340, 1400, 1450, and 1490 kHz

Figure 3. Channel utilization in the United States prior to 1990.

Class of Station Prior to 1990	Power in kW Daytime hours Minimum	Maximum	Power in kW Nighttime hours Minimum	Maximum
I-A	10.0	50.0	10.0	50.0
I-B	10.0	50.0	10.0	50.0
I-N	10.0	50.0	10.0	50.0
II	10.0	50.0	10.0	50.0
II-A	10.0	50.0	10.0	50.0
II-B	0.25	50.0	0.25	50.0
II-C	0.25	50.0	0.25	1.0
II-D	0.25	50.0	Daytime only	Daytime only
II-S	0.25	50.0	less than 0.25	
III	0.25	5.0	0.25	5.0
III-S	0.25	5.0	less than 0.25	
IV	0.10	1.0	0.10	1.0

Figure 4. Power limitations by class of station.

Once again, taking into consideration the need for more service of a local nature, and recognizing the changing nature of propagation conditions in which full nighttime conditions do not exist until several hours after sunset, the FCC granted Post Sunset Authorizations (PSSA) for most of the daytime only stations. These authorizations permit operation for periods of up to two hours past sunset with power reduced to prevent interference. The FCC also changed its rules concerning the minimum power at which a station is permitted to operate. This, in turn, led to a subsequent decision allowing many PSSA stations to operate throughout the night, albeit with reduced power.[7]

Still another change that has occurred pertained to the use of the Class I-A clear channels. At one time, only a single station was permitted to operate on these channels at night, but in two FCC decisions, first some and now all of these channels have been broken down to permit the authorization of additional nighttime operations. Clear channel stations that at one time provided service for a major portion of the country during nighttime hours now are protected only out to a distance averaging 750 miles.

Use of the Expanded Band

The FCC has issued a rule making to implement the results of the expanded band Regional Administrative Radio Conference sessions held in 1986 and 1988. At that conference, participating administrations reached agreement on the criteria for expanding the AM band in Region 2 by the adding ten new channels from 1605 kHz to 1705 kHz. Although the Regional agreement established specific technical criteria for the implementation of the new channels, (including the granting of priority usage of certain channels) the United States still retained considerable latitude in its domestic implementation. In a large country like the U.S., the use of the channels is unrestricted except in the relatively few areas near the borders with our neighbors.

The basic criteria set forth in the Regional Agreement are as follows: stations may operate with 1 kW of power with a nondirectional antenna height of 90°; or stations may operate with a power not in excess of 10 kW by employing a directional antenna to provide equivalent protection to stations in other countries. The channels allotted to the United States in the border areas vary from location to location, but, as noted, over a large portion of the U.S. all ten of the channels may be used. Nonallotted channels are not precluded from use, but the allotted channels in the other countries must remain fully protected.

Persons seeking an authorization for a new AM broadcast station may do so by filing an application with the FCC. Such application must provide documentation that the proposed operation will comply with all applicable FCC Rules as well as the appropriate international regulations. The details and methodology for allocating AM radio stations along with basic design specifications for AM directional antennas can be found in FCC Rules at Sections 73.14 to 73.190. A reader interested in further details regarding AM allocations criteria is referred to these sections of the FCC's Rules.

FM BROADCASTING FREQUENCY ALLOCATION

The frequency band 88 MHz to 108 MHz is allocated for FM broadcasting in Region 2 and, with some exceptions, the same is true in Regions 1 and 3 as well. However, unlike AM, FM broadcast allotments are largely a domestic matter, especially in large countries such as the United States, due to the limited nature of signal propagation at these frequencies. Although there are some international regulations regarding FM broadcasting, there is no region-wide FM agreement in Region 2. Instead, there are bilateral agreements between the U.S. and Canada and between the U.S. and Mexico. Both regulate the use of FM Channels in the border areas and specify technical standards in order to insure system compatibility.

The basic plan of the FM broadcast band is shown in Fig. 5. The band is divided into 100 channels, each 200 kHz wide. In the United States, the lower 20 channels, located between 88 MHz and 92 MHz, have been reserved for noncommercial broadcasting, but some of these stations operate on the upper 80 channels as well. In addition, although it is part of television channel 6, the frequency 87.9 MHz can be used for low power noncommercial FM stations, but its use is severely restricted.

As with AM broadcasting, there are different classes of FM stations designed to provide different types of service. In recent years, the number of different classes of stations has grown considerably as the demand for more stations increased. In response to this demand, the FCC has significantly modified the criteria concerning the use of the frequencies. In June 1983, the FCC concluded a lengthy rule making proceeding and modified the domestic allotment criteria for FM broadcasting. Prior to this action, there were three classes of stations on the 80 commercial channels. Twenty of these channels were used for lower power Class A stations having a maximum effective radiated power (ERP) of 3 kW and a maximum antenna height-above-average-terrain (HAAT) of 300 feet. Class A stations have a 1 mV/m service radius of about 15 miles. Higher power Class B or C stations operated on the remaining 60 channels. Whether a station was

82–MHz CHANNEL 6	88 MHz–92 MHz U.S. Noncommercial	92 MHz–108 MHz U.S. Commercial
TELEVISION	FM BROADCASTING	

Figure 5. FM broadcasting band. The FM sound carrier for TV channel 6 is 87.75 MHz or 15 kHz below the first FM broadcast channel (220 at 87.9 MHz).

designated Class B or C depended on where in the U.S. it was located.[8] A Class B station is located in Zone I or I-A. Zone I is the northeast U.S., extending south to the Virginia-North Carolina border and west to the Mississippi River. Zone I-A is all but the northernmost portion of California, plus Puerto Rico and the Virgin Islands. Class C stations operate elsewhere in the country, which is referred to as Zone II.

Class B stations operate with a maximum ERP of 50 kW at 500 feet HAAT and have a service radius of about 33 miles. Class C stations operate with a maximum power of 100 kW at 2000 feet HAAT, for a service radius of approximately 57 miles. FM stations in each of these classes may elect to operate at a HAAT above the maximum, but in such cases they are required to make a compensatory reduction in ERP as noted in FCC Rules and Regulations Section 73.211.

This system was changed drastically in 1983.[9] The FCC:

1. Permitted Class A stations to operate on channels previously reserved for Class B or C stations.
2. Created three new classes of FM stations. Class B1 stations are permitted to operate in Zones I and I-A with a maximum ERP of 25 kW at 100 meters (328 feet HAAT); Class C1 and Class C2 stations are permitted to operate in Zone II. Class C1 stations are permitted a maximum ERP of 100 kW at 300 meters (984 feet) HAAT, and Class C2 stations are permitted a maximum ERP of 50 kW at 150 meters (492 feet) HAAT.
3. Required stations that were previously licensed as a Class B or C, and were not operating at the minimum level specified for their class under the new rules, to upgrade their facilities within three years. Otherwise, the under-minimum facilities would be reclassified to the appropriate lower class based on the facilities they used.
4. Increased the maximum antenna HAAT for Class A stations to 100 meters (328 feet).

In 1989 the FCC further modified the rules to permit Class A stations to operate with an ERP of 6 kW and a HAAT of 100 meters, and added an additional classification (C3) that permits operation in Zone II with a maximum of 25 kW at 100 meters HAAT. In addition, there also are Class D stations that operate as noncommercial educational stations with power not in excess of 10 watts. However, applications for this class of station are no longer accepted.[10] A complete list of the station classes and a summary of FM allotment standards can be found in Figs. 6A and 6B.

Unlike AM broadcasting, where a new station may be applied for at any location where it can meet applicable criteria, the use of commercial FM channels (channels 221–300) is governed by the FM "Table of Allotments" found in Section 73.202 of the FCC's Rules. This table lists all FM channel allotments that have been made available for use. Most are already in use. If the Table does not list a vacant channel in the desired community, the prospective applicant must file a rule making petition with the FCC seeking to add such a channel for the community. The rule making petition proposing such addition must provide a showing that the proposal meets the separation requirements that are applicable to the class of station being proposed. A complete list of the spacing requirements, including those that pertain to stations located in Canada and Mexico, is provided in Figs. 7, 8, 9, and 10. Alternatively, a petitioner can propose to modify the Table by deleting a vacant existing allotment or by changing the frequencies of an existing station and thereby achieve compliance with these spacing requirements.

Station class	Maximum ERP	Maximum HAAT in Meters (ft.)	Expected service radius
A	6 kW	100 (328)	28 km
B1	25 kW	100 (328)	39 km
B	50 kW	150 (492)	52 km
C3	25 kW	100 (328)	39 km
C2	50 kW	150 (328)	52 km
C1	100 kW	299 (981)	72 km
C	100 kW	600 (1968)	92 km

Figure 6A. Standards for FM Allotments for locations other than Puerto Rico and the Virgin Islands.

Station class	Maximum ERP	Maximum HAAT in Meters (ft.)	Expected service radius
A	6 kW	240 (787)	42 km
B1	25 kW	150 (328)	46 km
B	50 kW	472 (1549)	78 km

Figure 6B. Standards for FM Allotments in Puerto Rico and the Virgin Islands.

Once a location has been added to the Table of Allotments, the FCC will announce a period of time called a *window* when it will accept applications for the location. If, as often is the case, there are multiple applicants, the winning applicant will be determined by a comparative hearing. It also should be noted, that in cases where it is not possible to locate a transmitter site that meets the mileage separation requirements, the FCC does permit the use of reduced power and/or antenna height, as well as directional antennas, in order to provide equivalent protection to other stations. However, this only applies to the filing of an application for a location that is already in the Table of Allotments. The FCC will not accept a proposal to modify the Table unless it is shown to meet applicable separation requirements.

Assignment of stations on the noncommercial educational channels (200–220) is accomplished more in the manner that is followed in AM, where an application includes a showing that interference will not be caused to other stations. In addition, for proposals to use Channels 218, 219, and 220, compliance with applicable separation requirements to any allotments on higher, adjacent commercial channels is required.

In addition to the regular FM broadcast stations, there are two other types of stations that are permitted to operate in the FM band on a secondary basis. These

Relation	Co-channel	200 kHz	400/600 kHz	10.6/10.8 MHz
A to A	115(71)	72(45)	31(19)	10(6)
A to B1	143(89)	96(60)	48(30)	12(7)
A to B	178(111)	113(70)	69(43)	15(9)
A to C3	142(88)	89(55)	42(26)	12(7)
A to C2	166(103)	106(66)	55(34)	15(9)
A to C1	200(124)	133(83)	75(47)	22(14)
A to C	226(140)	165(103)	95(59)	29(18)
B1 to B1	175(109)	114(71)	50(31)	14(9)
B1 to B	211(131)	145(90)	71(44)	17(11)
B1 to C3	175(109)	114(71)	50(31)	14(9)
B1 to C2	200(124)	134(83)	56(35)	17(11)
B1 to C1	233(145)	161(100)	77(48)	24(15)
B1 to C	259(161)	193(120)	105(65)	31(19)
B to B	241(150)	169(105)	74(46)	20(12)
B to C3	211(131)	145(90)	71(44)	17(11)
B to C2	241(150)	169(105)	74(46)	20(12)
B to C1	270(168)	195(121)	79(49)	27(17)
B to C	274(170)	217(135)	105(65)	35(22)
C3 to C3	153(95)	99(62)	43(27)	14(9)
C3 to C2	177(110)	117(73)	56(35)	17(11)
C3 to C1	211(131)	144(90)	76(47)	24(15)
C3 to C	237(147)	176(109)	96(60)	31(19)
C2 to C2	190(118)	130(81)	58(36)	20(12)
C2 to C1	224(139)	158(98)	79(49)	27(17)
C2 to C	249(155)	188(117)	105(65)	35(22)
C1 to C1	245(152)	177(110)	82(51)	34(21)
C1 to C	270(168)	209(130)	105(65)	41(25)
C to C	290(180)	241(150)	105(65)	48(30)

Figure 7. Minimum distance separation requirements in kilometers (miles).

Relation	Co-channel	200 kHz	400/600 kHz	10.6/10.8 MHz
A to A	132	85	45	8
A to B1	180	113	62	16
A to B	206	132	76	16
A to C1	239	164	98	32
A to C	242	177	108	32
B1 to B1	197	131	70	24
B1 to B	223	149	84	24
B1 to C1	256	181	106	40
B1 to C	259	195	116	40
B to B	237	164	94	24
B to C1	271	195	115	40
B to C	274	209	125	40
C1 to C1	292	217	134	48
C1 to C	302	230	144	48
C to C	306	241	153	48

Figure 8. Minimum distance separation requirements in kilometers—Canadian Agreement.

Relation	Co-channel	200 kHz	400/600 kHz	10.6/10.8 MHz
A to A	105(65)	65(40)	25(15)	8(5)
A to B	175(110)	105(65)	65(40)	16(10)
A to C	210(130)	170(105)	105(65)	32(20)
A to D	95(60)	50(30)	25(15)	5(5)
B to B	240(150)	170(105)	65(40)	25(15)
B to C	270(170)	215(135)	105(65)	40(25)
B to D	170(105)	95(60)	65(40)	16(10)
C to C	290(180)	240(150)	105(65)	48(30)
C to D	200(125)	155(95)	105(65)	25(15)
D to D	18(11)	10(6)	5(3)	3(2)

Figure 9. Minimum distance separation requirements in kilometers (miles)—Mexican Agreement.

FM Class	TV Zone I	TV Zones II and III
A	17	22
B1	19	23
B	22	26
C3	19	23
C2	22	26
C1	29	33
C	36	41

Figure 10. Minimum distance separation requirements in kilometers to TV channel 6 from FM stations on channel 253 (98.5 MHz).

are FM translator stations and FM booster stations. An FM booster station retransmits the signal of a primary station on the primary station's channel in order to serve areas where the primary station's signal is inadequate. An FM translator station is similar to an FM booster station, except that the signal is not retransmitted on the same channel but instead is translated to a different channel. These stations are authorized in accordance with Part 74, Subpart L of the FCC Rules which, among other things, requires that such stations provide protection from interference to all regular FM Broadcast Stations.

TELEVISION BROADCASTING FREQUENCY ALLOCATION

There are several different frequency bands used for television broadcasting within the United States. The plan as shown in Fig. 11 includes the low-VHF band TV channels 2 to 4 (54 MHz to 72 MHz) as well as channels 5 and 6 (76 MHz to 88 MHz), the high-VHF band channels 7 through 13 (174 MHz to 216 MHz), and the UHF band channels 14 through 69 (470 MHz through 806 MHz). The greater portion of all of these bands is allocated for broadcasting throughout the world; but this allocation is not uniform and in many areas other uses such as land mobile are permitted on a secondary basis. This is also the case in the U.S., where certain UHF television channels are now used for radio services in some major cities. In general, due to the limited extent of radio wave propagation in the television band, TV allocations, like FM, are basically a domestic matter, with few international regulations. Although no regional agreement exists, the United States does have agreements with both Canada and Mexico concerning television allocation.

Like FM broadcasting, TV allocations in the United States are governed by a Table of Frequency Allotments. This Table, which can be found in Section 73.606 of the FCC Rules, contains all commercial as well as noncommercial allotments (the latter are

54–72 MHz	76–88 MHz	174–216 MHz	470–806 MHz
TV Channels 2–4	TV Channels 5–6	TV Channels 7–13	TV Channels 14–69

Figure 11. TV broadcasting bands.

identified by an asterisk). The Table is based on the separation criteria contained in Sections 73.610 and 73.698 of the FCC Rules. As can be seen in Figs. 12 and 13, there are different requirements for different zones or areas of the country. Zone I is the northeast U.S. extending south to the Virginia-North Carolina border and west to the Mississippi River. Zone II consists of that portion of the United States that is not in Zone I or Zone III, along with Puerto Rico, Alaska, the Hawaiian Islands and the Virgin Islands. Zone III is that portion of the southeast United States extending from the east coast of Georgia westward to the Mexican border. An exact description of the zones is contained in Section 73.609 of the FCC Rules.[11] As with FM, the zones reflect the differing population densities in various parts of the country. In addition, differences in propagation conditions were also considered. The closer spacings in Zone I recognize the fact that in the Northeast portion of the United States there are many large population centers needing stations of their own, which are close enough to one another to lessen the need for wide area service. Zone II is characterized by fewer population centers, usually smaller and further apart. For them, wide-area service is a necessity. Finally, because the area of the country along the Gulf Coast (Zone III) is susceptible to high levels of tropospheric propagation, stations in that area need to be spaced further apart so as not to interfere with one another. Although there are differences in the power levels authorized for low-VHF, high-VHF, and UHF stations, unlike AM and FM broadcasting, there are no class designations as such in television. The differing power limitations reflect the differences in signal propagation for low-VHF, high-VHF and UHF.

In addition to the above constraints, there are others on UHF channels due to various types of problems caused by the mixing, in TV receivers, of the signals of stations operating on different channels. A complete table of these restrictions referred to as the "UHF Taboos" is contained in Section 73.698 of the FCC Rules.

To obtain authority to operate a television station, an application is filed with the FCC. The applicant may request authority to operate a station on any of the allotments that are not already in use. Alternatively,

Minimum co-channel television separation distances in kilometers (miles)

ZONE	CHANNELS 2–13	CHANNELS 14–69
I	272.7 (169.5 miles)	248.6 (154.5 miles)
II	304.9 (189.5 miles)	280.8 (174.5 miles)
III	353.2 (219.5 miles)	329.0 (204.5 miles)

Minimum adjacent channel separation distances

Channels 2-13	95.7 kilometers (59.5 miles)
Channels 14-69	87.7 kilometers (54.5 miles)

Figure 12. Minimum co-channel and adjacent channel television separation distances.

FM Class	TV Zone I	TV Zones II and III
A	17	22
B1	19	23
B	22	26
C3	19	23
C2	22	26
C1	29	33
C	36	41

Figure 13. Minimum distance separation requirements in kilometers to TV channel 6 from FM stations on channel 253 (98.5 MHz).

a petition for rule making may be filed with the FCC seeking to modify the Table to include a new allotment. Such petitions must provide a showing that the proposal complies with applicable separation criteria that serve to prevent mutual interference. Generally speaking, few, if any, VHF allotments are available, but there are a number of vacant UHF channels remaining in scattered areas of the U.S.

In addition to the regular television broadcast stations, there are three other types of stations that are permitted to operate in the television bands with low power, on a secondary basis. These are low-power television stations (LPTV), TV booster stations, and TV translator stations. Low-power television stations may retransmit the signals of another station or they may originate programming. A TV booster station retransmits the signal of a primary station on the primary station's channel in order to serve areas where the primary station's signal is inadequate. A TV translator station is similar to a TV booster station, except that the signal is not retransmitted on the same channel but instead is translated to a different channel. All of these stations are authorized in accordance with Part 74, Subpart G of the FCC Rules, which requires full protection from interference to all regular television broadcast stations. In addition, these stations are not accorded any protection from the operation of regular television broadcast stations.

Aside from establishment of the LPTV service, there has been little change in the allocation of television broadcast stations over the past several years. The last major change was the reallocation of channels 70 through 83 (806 MHz to 890 MHz) from TV to the land mobile radio services in the early 1970s. At that time, sharing of UHF channels 14 to 20 (470 MHz to 512 MHz) with land mobile radio services also was permitted in certain major markets. As the demand for radio spectrum continues to grow, land mobile interests may press for further changes of this type, which broadcasters are likely to oppose.

Currently, studies are underway at the FCC and elsewhere looking toward a new type of high definition television system (HDTV) capable of providing greatly enhanced picture and sound quality. As of this writing, it appears the implementation of this new type of service will initially require stations to operate two separate systems. One would continue to provide service for existing television receivers and the second would provide the new service on a second channel. In view of this, it appears that the opportunity for any new stations will be extremely limited because of spectrum requirements to accommodate the existing station expansion into the new system.

AUXILIARY BROADCAST SERVICES FREQUENCY ALLOCATION

Although all auxiliary broadcast services share a common role in support of AM, FM, or TV broadcast operations, there are important differences between them. The nature of the service they provide varies as does the frequency band in which they operate. For these reasons, the allotments available for each auxiliary service needs to be discussed separately. However, before turning to a specific discussion of each service, some general comments are necessary.

The steady growth in the number of AM, FM, and TV stations, as well as their desire to use more advanced technology in every aspect of their operation, has greatly increased demand for spectrum in all auxiliary services. Because of the continuing demand for spectrum from a multitude of other broadcast and nonbroadcast activities, it has been difficult for the FCC to allocate more spectrum to alleviate the congestion faced by the auxiliary services. At the same time, some relief has come through changes in the FCC Rules to facilitate the use of newer, more spectrum-efficient, technology. In using this material, the reader should be aware that additional changes may take place. Accordingly, the careful reader should check the current regulations governing that particular radio service.

Recognizing that available spectrum is limited, it is imperative for broadcasters to use it efficiently. To do this, it is first necessary to understand what spectrum is allocated, how it may be used and what advantages or disadvantages, if any, may be involved in the use of a particular band, for each of the auxiliary services discussed below.

Remote Pickup Broadcast Stations (RPUs)

Remote pickup stations are mobile or portable facilities used to transmit live on-the-air programming from a temporary remote location, such as a shopping center or football game, to the station's studio facilities. This material can be taped for later rebroadcast, or it can be incorporated into actual on-going live broadcasts. Radio stations typically have several RPUs that may be licensed to one or more frequencies.

In November 1984, the FCC significantly revised its radio broadcast auxiliary frequency allocations to permit operational use of narrow band technologies in the Broadcast Remote Pickup Service. The Commission's goal was to foster spectrum efficiency in a flexible manner. Broadcasters and equipment manufacturers who wanted to operate narrow band equipment, the Commission believed, should not be precluded from doing so by rigid FCC Rules. However, these changes have not yet been formally implemented by the FCC due to delays in setting up the mechanism to handle the expected applications, but formal implementation is expected in the near future. In the meantime, the FCC will accept applications based on this revised allocation, which is detailed below.

The following are the frequency allocations for radio broadcast remote pickup stations.

1. 25.67–26.47 MHz: There are a total of 25 frequencies available in this band for use by remote pickup broadcast stations. Bandwidth is limited to 20 kHz except between 25.87 MHz and 26.03 MHz where 40 kHz is permitted. Please note that the use of the frequencies between 25.87 and 26.09 is subject to the condition that no harmful interference is caused to broadcast stations sharing this band. It is also noted that the frequencies between 26.100 and 26.175 have been allocated on a worldwide basis to the Maritime Mobile Services. Although this service has not yet been implemented, at such time as the maritime service is implemented, these frequencies may no longer be available for use by remote pickup stations. It is further noted that the frequencies between 26.175 and 26.47 are allocated on a worldwide basis for use by various types of fixed and mobile operations. Hence, in selecting a frequency in this band for remote pickup use, the implications of the above restrictions should be taken into consideration.
2. 152.875–153.3625 MHz: There are a total of 54 frequencies available in this band for use by remote pickup broadcast stations. Each channel is 5 kHz wide, and channels may be stacked to form a single channel with a maximum bandwidth of 30 kHz. This band is shared with the Private Land Mobile Radio Service and the Maritime Service, and operation of remote pickup stations is subject to the condition that no harmful interference is caused to these other services. Please note that these frequencies are not available to networks entities or for use on board aircraft unless they were previously licensed for such purpose.
3. 160.8625–161.3975 MHz: There are a total of 30 frequencies available in this band for use by remote pickup broadcast stations in Puerto Rico or the Virgin Islands where they are shared with the Public Safety and Land Transportation Radio Services. Each channel is 5 kHz wide and channels may be stacked to form a single channel with a maximum bandwidth of 30 kHz.
4. 161.6275–161.7725 MHz: There are a total of 102 frequencies available in this band for use by remote pickup broadcast stations. Each channel is 5 kHz wide and channels may be stacked to form a single channel with a maximum bandwidth of 30 kHz. These frequencies are not available to network entities and are not available for use in Puerto Rico or the Virgin Island. Also, Public Safety and Land Transportation Radio Service stations may continue to operate on these frequencies on a noninterference basis.

5. 166.25 and 170.15 MHz: These frequencies may be used by remote pickup stations with a maximum bandwidth of 25 kHz; however, the area in which they may be used is restricted. A description of the area in which they may be used is found in FCC Rule Section 74.402.
6. 450.01, 450.02, 450.98, 450.99, 455.01, 455.02, 455.98, and 455.99 MHz: These frequencies may be used by remote pickup stations only for the transmission of operational communications, including tones for signaling and for remote control and automatic transmission system control and telemetry. Bandwidth is limited to a maximum of 10 kHz.
7. 450.0275–455.6225 MHz: There are a total of 240 frequencies available in this band for use by remote pickup broadcast stations. Each channel is 5 kHz wide, and channels may be stacked to form a single channel with a maximum bandwidth of 50 kHz.
8. 450.6375–455.8625 MHz: There are a total of 20 frequencies available in this band for use by remote pickup broadcast stations. Each channel is 25 kHz wide, and channels may be stacked to form a single channel with a maximum bandwidth of 50 kHz. Users committed to 50 kHz bandwidths to transmit program material will have primary use of these channels.
9. 450.900, 450.950, 455.900, and 455.950 MHz: These frequencies are available for use by remote pickup broadcast stations. Each channel is 50 kHz wide, and channels may be stacked to form a single channel with a maximum bandwidth of 100 kHz. Users committed to 100 kHz bandwidths to transmit program material will have primary use of these channels.

Aural Broadcast Auxiliary Stations

Aural broadcast auxiliary stations include studio-transmitter link (STL), intercity relay (ICR), and microwave booster stations used by radio broadcast stations. STL stations are fixed stations used for transmitting program material between the studio and the transmitter of a broadcasting station. ICR stations are fixed stations used for the transmission of program material between broadcasting stations, except international, for simultaneous or delayed broadcast. Microwave booster stations are used to relay the signals of a STL or ICR station over a path that cannot be covered with a single station. They receive and transmit on the same frequency. One or more microwave booster stations may be authorized to licensees of STLs or ICRs. It should be noted that stations in the aural broadcast auxiliary service may be authorized on a secondary noninterference basis to licensees of TV broadcast stations to transmit aural material.

The following frequencies are available for assignment to STL, ICR, and microwave booster stations:

1. 942.5, 943.0, 943.5, and 944 MHz: These frequencies are available for use in Puerto Rico and the Virgin Islands. Also, stations licensed in other parts of the United States prior to November 21, 1984 may continue to operate on a co-equal primary basis in this band.
2. 944–952 MHz: There are 320 channels 25 kHz wide available in this band. The channels may be stacked to form a single channel up to 300 kHz wide. Separately, stations also may be authorized additional 25 kHz channels up to a grand total of 20 channels. The use of these frequencies by ICR stations is subject to the condition that no harmful interference is caused to other classes of stations.
3. 18760–18820 MHz and 19100–19160 MHz: There are 24 channels, each 5 MHz wide, available in these bands. These frequencies are shared on a co-primary basis with other fixed services and their use is subject to the rigorous coordination requirements of FCC Rule Section 21.100(d).

Part 94 Operational Fixed Services

Although not strictly broadcast auxiliary frequency bands, the Operational Fixed Services frequencies, particularly those above 21200 MHz and the 6425–6525 MHz band, in fact are available to broadcasters and may be less congested than the typical auxiliary frequencies. However, some of them are subject to significant limitations. Accordingly, licensees interested in making use of these frequencies should refer to the provisions applicable to the band in question. Applications and instructions are available from the Land Mobile and Microwave Division in the FCC's Private Radio Bureau. Suitable equipment is available from several microwave equipment manufacturers.

Television Broadcast Auxiliary Stations

The demand for the spectrum allocated for television auxiliary services is significantly greater than the spectrum allocated for radio broadcasting services. In addition to the extensive local demand, network remote units travel extensively and compete with local broadcasters for available frequencies, resulting in increased spectrum congestion. Finally, it is also important to recognize that the variety of activities undertaken by television broadcasters usually requires more complex auxiliary systems than is the case in the radio industry.

The following are the types of television broadcast auxiliary stations: TV pickup stations, TV STL stations (studio transmitter link), TV relay stations, TV translator relay stations, and TV microwave booster stations.

TV pickup stations are land mobile stations used for the transmission of TV program material and related communications from the scenes of events to TV broadcast or low power TV (LPTV) stations.

TV STL stations are fixed stations used for the transmission of program material and related communications from the studio to the transmitter of a TV broadcast or LPTV station.

TV relay stations are fixed stations used for transmitting visual program material between TV broadcast or LPTV stations or for the relay of transmissions from a remote pickup station to a single TV station.

TV translator relay stations are fixed stations used for relaying programs and signals of TV broadcast stations to LPTV stations or TV translator stations.

TV microwave booster stations are fixed stations used to receive and amplify signals of TV pickup, TV STL, TV relay, or TV translator relay stations and retransmit them on the same frequency. These stations are used to transmit signals over a path that cannot be covered by a single transmitter.

The following are available for assignment to TV pickup stations.

1. 1990–2110 MHz: Seven channels are available in this band which is also available for assignment to all the other types of television broadcast auxiliary stations.
2. 2450–2483.5 MHz: Two channels are available in this band which is also available for assignment to all the other types of television broadcast auxiliary stations. This band is not available for use by network entities.
3. 6425–6525 MHz: The channels available in this band are co-equally shared with mobile stations licensed under Parts 21, 78, and 94 for the FCC Rules. The available channel bandwidth varies from 1 MHz to 25 MHz and Section 73.602 of the FCC Rules contains further explanation concerning the usage of this band.
4. 6875–7125 MHz: Ten channels, each 25 MHz wide, are available in this band which is also available for assignment to all the other types of television broadcast auxiliary stations.
5. 12700–13250 MHz: This band contains 43 channels; however, the channels overlap and therefore, if use of this band is contemplated, Section 73.602 of the FCC Rules should be consulted for a more complete understanding of their usage. This band is also available for assignment to all the other types of television broadcast auxiliary station.
6. 38.6–40 GHz: This band is available without channel bandwidth limitation on a secondary basis to fixed stations.

The following are available for assignment to TV STL, TV relay, or TV translator relay stations.

1. 1990–2110 MHz: Seven channels are available in this band which is also available for assignment to TV pickup stations.
2. 2450–2483.5 MHz: Two channels are available in this band which is also available for assignment to TV pickup stations.
3. 6875–7125 MHz: Ten channels, 25 MHz wide, are available in this band which is also available for assignment to TV pickup stations.
4. 12700–13250 MHz: This band contains 43 channels; however, the channels overlap and therefore, if use of this band is contemplated, Section 73.602 of the FCC Rules should be consulted for a more complete understanding of their usage. This band is also available for assignment to TV pickup stations. In addition, the channels between 13150 MHz and 13200 MHz are not available within 50 km of the top 100 markets.
5. 17700–19700 MHz: Frequencies in this band are shared on a co-equal basis with stations in the fixed service authorized by Parts 21, 78, and 94 of the FCC Rules. The available channel bandwidth varies from 2 MHz to 80 MHz and Section 73.602 of the FCC Rules contains further explanation concerning the usage of this band.
6. 31.0–31.3 GHz: Frequencies in this band are shared on a co-equal basis with stations in the fixed service authorized by Parts 21, 78, and 94 of the FCC Rules. The available channel bandwidth varies from 25 MHz to 50 MHz and Section 73.602 of the FCC Rules contains further explanation concerning the usage of this band.

In addition to the above frequencies, TV STL and TV relay stations also may be authorized to use UHF-TV channels 14–69 on a secondary basis provided no interference is caused to TV and LPTV stations operating in this band.

Furthermore, the aural portion of television broadcast program material may be transmitted over an aural broadcast STL or ICR station on a secondary, noninterference basis. Likewise, remote pickup stations may be used to transmit the aural portion of television program material.

CONCLUSION

As can be seen from the material contained in this chapter, frequency allocation is a very complex matter and is subject to frequent changes. Therefore, the reader is advised to consult the FCC Rules and Regulations for a complete description of the current allocation situation, including the procedures and policies now being applied by the FCC concerning a particular band.

ENDNOTES

1. Technically the term "allocation" refers to the process by which a frequency band is made available for a specific purpose. However, it is used here in a broader sense that includes the *allotment* of frequencies within a band, and the *assignment* of individual stations.
2. A complete list of the 12 CCIR study groups is contained in Appendix 1.
3. A sister technical committee, the CCITT, performs similar functions for wireline telephone and telegraph communications.
4. Copies of the Green Books may be purchased from:
   International Telecommunication Union
   General Secretariat, Sales Service
   Place des Nations
   CH-1211 Geneva 20
   Switzerland

5. The nature of the Commission's rule making process is described in greater detail in Chapter 1.1 "FCC Organization and Administrative Procedures" of this handbook.
6. Copies of the 1981 Rio Agreement may be purchased from:
   International Telecommunication Union
   General Secretariat, Sales Service
   Place des Nations
   CH-1211 Geneva 20
   Switzerland
7. The rules pertaining to PSSA and PSRA are contained in FCC Rules and Regulations Section 73.99.
8. The concept of allowing different classes of stations in different areas or zones is based on the population density of the areas. The Commission has assumed that there is less need for wide areas of service in areas of dense population. A map depicting Zone I is contained in Appendix 2.
9. See FCC Rules Sections 73.210 and 73.211 for a discussion of station classes.
10. Section 73.512(c) FCC Rules and Regulations.
11. Maps depicting zones I and III are contained in Appendix B.

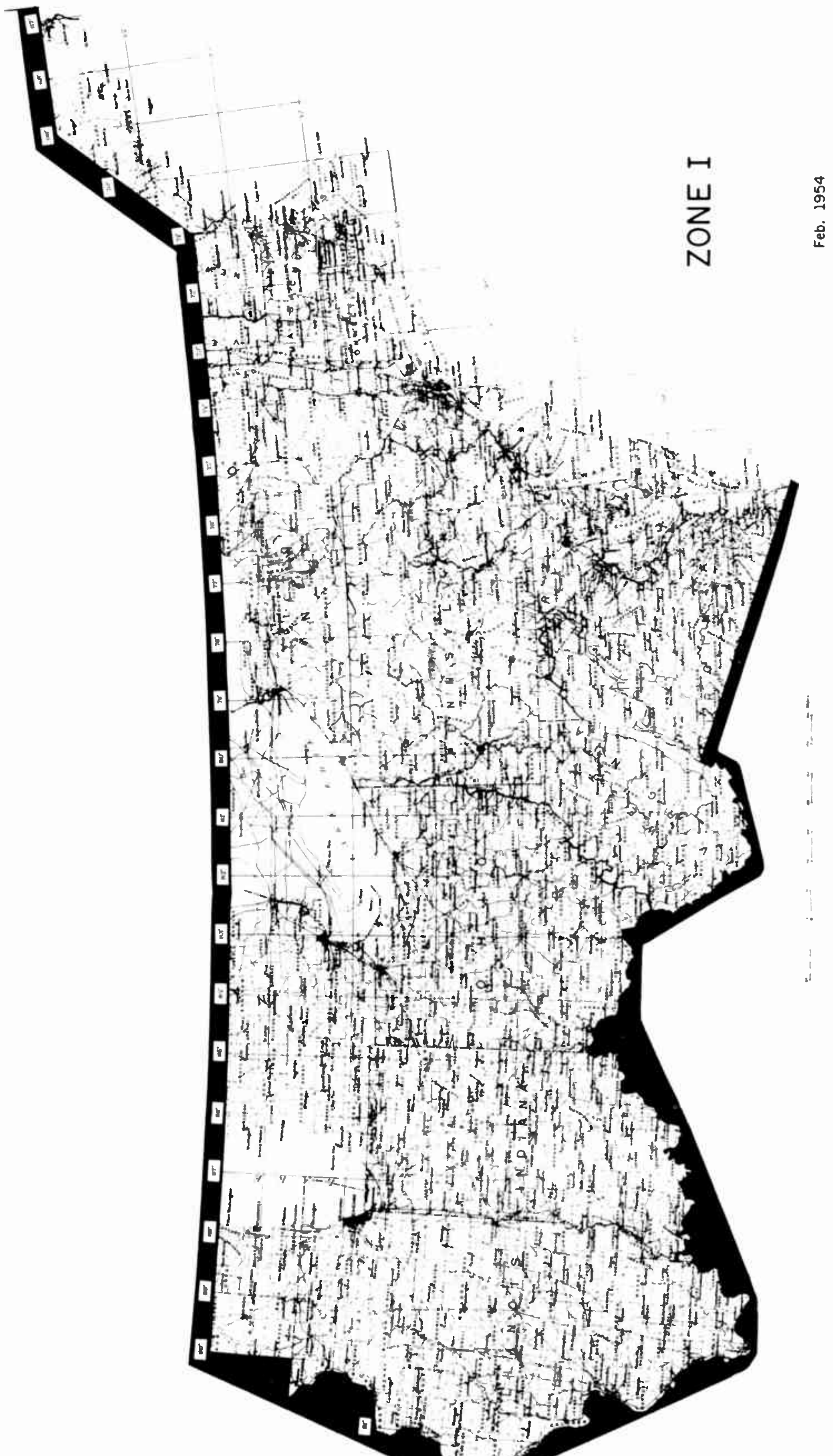
APPENDIX A

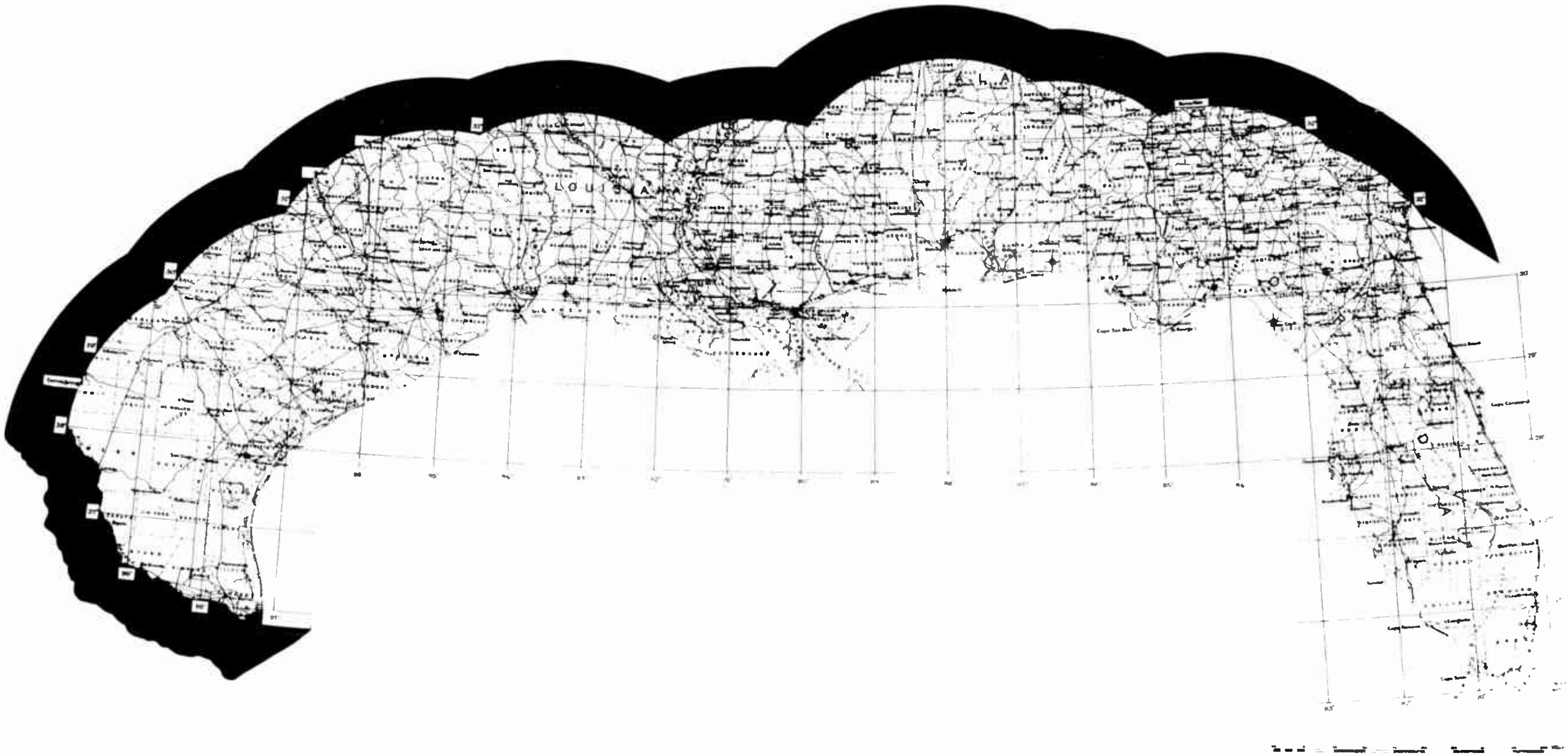
CCIR Study Groups

Study Group Subjects

1. Spectrum utilization and monitoring.
2. Space research and radioastronomy.
3. Fixed service at frequencies below about 30 MHz.
4. Fixed service using communications satellites.
5. Propagation in nonionized media.
6. Ionospheric propagation.
7. Standard frequency and time-signal services.
8. Mobile services.
9. Fixed service using radio relay systems. Coordination and frequency sharing between systems in the fixed satellite service and terrestrial radiorely systems (subjects common to study groups 4 and 9).
10. Broadcasting service (sound) including audio recording and satellite applications.
11. Broadcasting service (television) including video recording and satellite applications.
12. Cross-service spectrum sharing issues. (This group was created in 1990. It does not originate any studies but is given specific tasks by the CCIR Director. Its usefulness will be evaluated at the end of the four year cycle and a decision made as to whether it should continue to exist.)

APPENDIX B





ZONE III

MARCH 1957

APPENDIX B

1.5 Environmental Concerns

Warren P. Happel
Brecksville, Ohio

INTRODUCTION

We work in a world of regulations. Regulations can easily exist today which are not known but which must be followed. As broadcasters we are well aware of the Federal Communications Commission (FCC) and Federal Aviation Administration (FAA) rules which companion our day-to-day business activities. What is not as apparent are the federal Environmental Protection Agency (EPA) and Occupational Safety and Health Administration (OSHA) rules which affect broadcasters as well as other business firms. Most states, many cities, and some counties have parallel rules which often contain additional requirements. Meeting all federal requirements does not necessarily mean all state and local rules have been met. Fines can be imposed for rule violations, even though the rules may not be well publicized and therefore unknown. Penalties can be from federal, state, or local agencies.

This overview will alert you to rules, agencies, dates of compliance, and where to find more information. In comparison to the multitude of pages which have been written about each of the subjects discussed herein, this alert is brief and not exhaustive. The purpose is to guide you to seek more information relative to the regulations that apply to your operation.

The EPA and/or OSHA have rules, regulations and guidelines which affect broadcasters covering:

- Air Contaminants
- Asbestos
- Drinking Water Coolers
- Electromagnetic Radiation
- Hazard Communications Standard
- Lockout/Tagout
- Occupational Noise Exposure
- Polychlorinated Biphenyls (PCBs)
- Protection of Stratospheric Ozone
- Radon
- Underground Storage Tanks

The goal of the OSHA Occupational and Health Standards is to make the workplace safe, and the EPA rules are designed to protect the environment. These rules are not specifically aimed at the Broadcaster, but are for industry/business in general. The EPA and OSHA have not gone out of their way to inform individual business operations. For example, consider PCBs. The EPA rules name power and food industries as PCB users but no mention is made within the rules of the broadcast industry. It wasn't until a PCB transformer was broken open in Oregon, that broadcasters became aware of the PCB problem. Funding, specifically for notification, may be part of the problem, considering there are so many businesses affected and not enough staff available for the purpose of informing and enforcing. To help spread the word, broadcasters and newspapers will continue to receive requests to report OSHA and EPA penalties which have been imposed for rule violations.

INFORMATION SOURCES

Information sources are available that publish the EPA and OSHA Rules, but they are not generally distributed and usually must be bought. Obviously the rules will not even be sought if it is not known that they exist. As an excellent secondary source, the National Association of Broadcasters (NAB) has provided member stations with useful information concerning underground storage tanks (USTs), asbestos, polychlorinated biphenyls (PCBs), and electromagnetic radiation (EMR). Some law firms are now notifying their clients of compliance requirements. Newspapers also serve as a secondary source, but since the articles are usually written to provide general information it is easy to miss what isn't specifically directed toward the broadcast industry. Not knowing that rules exist will not prevent being fined if a penalty is applicable.

The primary source of the OSHA and EPA rules is the Code of Federal Regulations (CFR) which is the general and permanent collection of rules published by the Federal Register (FR). The Code is divided into 50 titles (e.g., title 29 for Labor), each title is divided into volumes (e.g., title 29 has eight volumes), each volume is divided into chapters which usually bear the name of the issuing agency (e.g., Chapter XVI Occupational Safety and Health Administration, Department of Labor) and each chapter is further subdivided into numbered parts (e.g., volume 5 is title 29, contains parts 1900 to 1910 and is all Chapter XVII).

Each volume of the Code is revised at least once each calendar year. In order to stay current on FCC Regulations, one may subscribe to a service that sends FCC updates, learn of major changes from NAB, subscribe to the Federal Register (FR), or some combination of the above.

And so it is with the EPA and OSHA rules. It is not sufficient to buy "Title 29 Parts 1900 to 1910," revised as of July 1, 1991, because just as with the FCC Regulations, changes must necessarily continue with the EPA and OSHA rules.

To be aware of changes made to the Code, one can subscribe to the "List of CFR Sections Affected" (LSA), without subscribing to the FR. (The LSA is included with a subscription to the FR.) Upon learning from the LSA that a regulation change has been made, it would be necessary to pursue the change in the FR. Some local libraries subscribe to the Federal Register.

At the end of each topic discussed herein, references and contacts are listed so that the reader may pursue more sources for information. It is necessary to keep alert and question any hint that a compliance requirement applies to some aspect of the business.

For additional information about the Federal Register and CFRs, contact the Superintendent of Documents, Government Printing Office, Washington, D.C. 20402–9371, (202) 783–3238 or U.S. Government Printing Office Bookstores located in Atlanta, Birmingham, Boston, Chicago, Cleveland, Columbus OH, Dallas, Denver, Detroit, Houston, Jacksonville FL, Kansas City, Los Angles, Milwaukee, New York, Philadelphia, Pittsburgh, Pueblo CO, San Francisco, Seattle, Washington D.C. (3), and Laurel MD.

Additionally, informational booklets are often available through state agencies, such as Department of Transportation, Department of Occupational Safety and Health, and Department of Environmental Regulation, to mention just three departments which are defined in some states.

AIR CONTAMINANTS

Air contaminants are defined by OSHA in a final rule (651 pages), and this rule became effective March 1, 1989 with a compliance date of September 1, 1989. This rule modifies many of the existing permissible exposure limits (PELS) of substances, establishes PELS for substances for which no previous exposure limits exist and includes short-term exposure limits (STEL) to complement the 8-hour time weighted average (TWA) limits. There are 428 substances listed in the rule which include such items as wood dust and carbon tetrachloride. It is estimated that every 20 minutes a new and potentially toxic chemical is introduced into industry. Thus, one can expect this list of 428 substances to be increased as substance studies continue on new and existing substances. For example, the PELS for fibrous glass and mineral wool, two popular insulation materials, were not given in the original 428 substance list. They are named as having their PELS delayed.

This current revised rule on air contaminants should not have a significant impact on the broadcasting industry since there would not normally be exposure to many of the substances listed or at least in amounts to trigger concern. However, the only way to be sure is to review the list of 428 substances and determine if a substance is being used in significant amounts. As an example, wood dust is listed as causing respiratory effects and carries a TWA of five milligrams per cubic meter. This TWA can be met with proper ventilation. More serious are those substances which can cause cancer, kidney or liver effects, and other long term physiological concerns. It is obviously beyond the scope of this review to list the 428 substances listed and their health effects. Under the Hazard Communications Standard, discussed herein later, one should already be aware of hazardous substances being used in the workplace.

In 1987, the EPA was asked to consider indoor air pollution. While there is not now an EPA or OSHA standard for indoor air pollution (other than the PELS established for the 428 substances) the American Society for Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) has recommended standards for indoor ventilation requirements. Additionally, some states have established ventilation requirements. Air contaminants such as mold (which can build-up in ventilation systems), smoke, carbon dioxide, and gasses emanating from building materials can cause some severe reactions in people. Complaints of fatigue, dizziness, eye irritation, headache, nausea, and skin irritation can be caused by ventilation problems rather than a cold, the flu, or allergies.

There are companies which specialize in the measurement of indoor air contamination and make recommendations for the solution to indoor air problems. A building's air should be checked for contaminants if a ventilation problem is suspected.

For additional information see:

"Air Contaminants: Guide and Bibliography to Final Rule," OSHA, 29 CFR 1910, FR, March 28, 1989, pages 12792–12868.

"Air Contaminants," Final Rule, OSHA, 29 CFR Part 1910, FR, January 19, 1989, pages 2332–2983.

Copies of the Air Contaminants document may be obtained from the OSHA Publications Office Rm. N-3101, U.S. Department of Labor, 200 Constitution

Avenue, N.W., Washington, DC 20210, (202) 523-9667, or any OSHA regional or area office.

ASBESTOS

Asbestos removal and disposal are regulated by the EPA, and OSHA regulates the exposure to asbestos in the workplace. The National Emissions Standards for Hazardous Air Pollutants (NESHAPS) requires that asbestos material must be removed when it is friable, that is, crumbled or in any form which could free it into the air. By EPA definition, "Friable asbestos material means any material containing more than one (1) percent asbestos. . . . that, when dry can be crumbled, pulverized, or reduced to powder by hand pressure." Asbestos is one of eight known naturally occurring carcinogens.

Asbestos is not required to be removed if it is enclosed or encapsulated. However, eventually asbestos may need to be removed and some property sales are being conditioned on "no asbestos present." Discussion continues as to whether there is more risk involved in the removal of asbestos which is totally encapsulated rather than leaving it in place.

Asbestos was used commercially as fire-proofing as early as the 1890s, as insulation in early 1900s, and at the height of its popularity in more than 3,000 items, including such widely distributed products as brake linings and hair dryers. Although one can easily identify some material forms of asbestos, by color and consistency, final identification should be done by a lab which can also measure the quantity in a sample. Removal of asbestos should be handled by a licensed contractor, and it is expensive. One should deal with contractors who are experienced, have certified trained workers and supervisors, carry proper insurance, and have submitted a satisfactory work plan. The air needs to be monitored during removal. It is also a good idea to have occurrence type liability insurance for asbestos related work.

The asbestos abatement contractor must certify that the asbestos removed has been disposed of properly in an EPA licensed landfill. Under the EPA published NESHAP revisions of November 20, 1990, the owner or operator of an asbestos removal site is considered the asbestos waste generator. The employer must inform employees of asbestos removal and post signs. Employee notification is required under the OSHA Hazard Communication rules. No smoking is permitted in areas where asbestos removal is taking place. Smoking severely increases the chance of cancer if asbestos lodges in the lungs.

The rules do not end at the federal level. Some states had regulations before the EPA or OSHA did. Not handling asbestos abatement properly can result in a health hazard, the delay of job completion, and fines. The OSHA asbestos standard states that medical surveillance must be provided to any employee exposed to an airborne concentration of asbestos in excess of 0.1 fiber per cubic centimeter of air calculated as an eight-hour time-weighted average. Further, medical records on employees exposed must be maintained by the employer, indefinitely if necessary, even after employment ends.

For additional information see:

"National Emission Standards for Hazardous Air Pollutants; Asbestos NESHAP Revision"; Final Rule, EPA, FR, November 29, 1990, pages 48405–48433.

"Occupational Exposure to Asbestos," Final Rule; partial response to court remand, OSHA, FR, February 5, 1990, pages 3724–3732.

"Occupational Exposure to Asbestos, Tremolite, Anthophyllite and Actinolite" Final Rule; partial response to court demand; 29 CFR Parts 1910 and 1926, OSHA, FR December 20, 1989, pages 52024–52028.

"Dealing with Asbestos in TV and Radio Stations," NAB Info-Pak/December 88. "Occupational Exposure to Asbestos, Tremolite, Anthophyllite, and Actinolite;" Final Rules; Amendment, 29 CFR Parts 1910 and 1926, OSHA, FR, September 14, 1988, pages 35610–29.

"Asbestos, Tremolite, Anthophyllite, and Actinolite," OSHA,
29 CFR 1910.1001, July 1, 1987, pages 682–723.

"Asbestos," EPA, 29 CFR 763.

"Consumers Patching Compounds Containing Respirable Freeform Asbestos Ban," Consumer Product Safety Commission, 16 CFR 1304.

"Occupational Exposure to Asbestos, Anthophyllite, and Actinolite," OSHA, FR, June 20, 1986, 29 CFR Parts 1910 and 1926, pages 22612–790.

Contact:

Office of Toxic Substances, EPA, 401 M St. SW, Washington, D.C., 20460 (202) 554–1404.

DRINKING WATER COOLERS

Drinking water coolers that are not lead free was the subject of an EPA final and proposed list of drinking water coolers which was published in the FR January 18, 1990. The EPA believes that the most serious cooler contamination problems are associated with water coolers which have water reservoir tanks lined with lead materials. On October 31, 1988, the Lead Contamination Control Act of 1988 was enacted. Its major provisions include a mandate for the Consumer Product Safety Commission to order the repair, replacement, or recall and refund of drinking water coolers that EPA has identified as containing lead-lined water tanks and a ban on the manufacture or sale of water coolers which are not lead free.

For additional information:

"Drinking Water Coolers That Are Not Lead Free," EPA, FR January 18, 1990, pages 1772–1776.

Contact: National Defense Library

Office of Drinking Water (WH-550), EPA, 401 M Street N.W., Washington DC 20460, (202) 475-8499.

EPA Safe Drinking Water Hotline (800) 426-4791.

ELECTROMAGNETIC RADIATION

Electromagnetic radiation is covered in Chapter 2.9 of this NAB Engineering Handbook.

For additional information see:

"Evaluating Compliance With FCC-Specified Guidelines for Human Exposure to Radiofrequency Radiation," OST Bulletin No. 65, October 1985, Office of Science and Technology, Federal Communications Commission.

"A Broadcasters Guide To FCC Radiation Regulation Compliance," (includes above), National Association of Broadcasters.

"Request for Declaratory Ruling: Radiofrequency Radiation Compliance" General Docket 88-469:FCC 88-291, Federal Register, October 19, 1988.

"Power Line Fields and Human Health," February 1985, IEEE Spectrum. "Biological Effects of Electromagnetic Fields," May 1984,

HAZARD COMMUNICATIONS STANDARD

The OSHA hazard communications standard is a rule which required compliance as of May 23, 1988. It applies to nonmanufacturing industries and can even apply to white collar offices with a copy machine that uses a chemical which is listed among the more than 400 substances considered to be hazardous. Because of the disaster at Bhopal, India, Congress enacted Title III, the Community Right To Know Law under the Superfund Amendment and Authorization Act (SARA) of 1986. SARA was originated to cover the clean-up of hazardous waste sites under the Comprehensive Environmental Response, Compensation and Liability Act, (CERCLA). The EPA is also involved since it lists the CERCLA-defined hazardous substances and, among other things, also generates reporting forms. If all of this seems confusing, it is. It is not apparent that there is any effort to consolidate or communicate all of the necessary reporting requirements to the millions of businesses which are affected. Most do not know they need to comply, or how. To comply, an employer must develop and maintain a written Hazard Communications Program, which explains and accomplishes the following:

1. Identify and list all of the hazardous chemicals in the workplace.
2. Obtain, (usually from whom one bought the product) the material safety data sheet (MSDS). Any manufacturer who manufactures a product which contains one or more of the chemicals listed by CERCLA and EPA must prepare an MSDS. (The Office of Management and Budget under the paper work reduction act has exempted drugs regulated by the FDA in the nonmanufacturing sector.)
3. Make the MSDS sheets available to the employees.
4. Identify those workers who should be trained and provide training to include the dangers of the substance and the correct and safe use.
5. Know what to do if there is a spill, fire, or personal contact.
6. All substances must be properly labeled and stored.
7. Appoint someone in charge of the program.
8. Any company required to have available MSDSs must submit copies of the MSDSs, or a list of MSDS chemicals to the State Emergency Response Commission, the Community Emergency Planning Commission, and the local fire department. Under some instances, an EPA-designed inventory report form must also be sent to the before mentioned three entities. Those were due March 1, 1988.

There was for a period of time a consumers product exemption which suspended the above reporting requirements if a product was not used differently in a business than in a home. This exemption has, at the time of this writing, been eliminated by the Third Circuit Court.

Most states now have OSHA-approved state hazard communication rules. These state regulations, as do the federal regulations, define the financial penalties which can be imposed for rule infractions.

Effective May 20, 1990, it became no longer permitted to distribute hexavalent chromium water treatment chemicals for use in comfort cooling towers. After May 18, 1990, it could no longer be used by owners and operators. The purpose of the rule is to prevent unreasonable risks associated with human exposure to air emissions of hexavalent chromium.

For additional information see:

"Chemical Hazard Communication," OSHA 3084, 1989.

"Hazard Communication Guidelines for Compliance," OSHA 3111, 1988.

"Hazard Communication A Compliance Kit," OSHA 3104, 1988.

"Prohibition of Hexavalent Chromium Chemicals In Comfort Cooling Towers," 40 CFR Part 749, EPA, FR, January 3, 1990, pages 222–241.

"Emergency Planning and Community Right-to-Know: Availability of Guidance Materials, Brochures, and Electronic Files, Notice of Availability," EPA, FR March 8, 1989, pages 9888–9889.

"List of Hazardous Substances and Reportable Quantities," CERCLA, FR, 3/16/87, pages 8150–71.

"List of Extremely Hazardous Substances and Their Threshold Planning Quantities," EPA, FR 4/22/87, pages 13395–410.

"List of Toxic Chemicals," FR, 6/4/87, pages 21169–77.

"Hazard Communications: Final Rule," 29 CFR Parts 1910, 1915, 1917, 1918, 1926, and 1928, OSHA, FR, 8/24/87, pages 31852–86.

"Emergency and Hazardous Chemical Inventory Forms and Community Right-to-Know Reporting Requirements: Final Rule," 40 CFR 370, EPA, FR, 10/15/87, pages 38344–77.

"Toxic Chemical Release Reporting: Community Right To Know"; Final Rule: 40 CFR Part 372, EPA, FR, 2/16/88, pages 4500–54.

"Extremely Hazardous Substances List," Final Rule, 40 CFR Part 355, EPA, FR, 2/25/88 pages 5574–5 and FR, 12/17/87 pages 48072–4.

"Hazard Communication: Display of Office and Management and Budget Control Numbers Assigned To Collection of Information,"
29 CFR parts 1910, 1915, 1917, 1918, 1926, and 1928, OSHA, FR, 12/4/87, pages 46075–80 and also FR, 4/27/88, pages 15033.

See 40 CFR 702, 704, 710, and 717 for more information.

Contact:

TSCA, Toxic Substances Control Act Assistance Office.

LOCKOUT/TAGOUT

On September 1, 1989, OSHA published the standard for the control of hazardous energy, known as the "Lockout/Tagout" rule, 29 CFR 1910.147. This rule became effective January 2, 1990. The rule describes lockout/tagout procedures to follow in order to disable machinery/equipment and prevent the release of hazardous energy (mechanical) while maintaining/servicing equipment. While lockout/tagout is for the purpose of controlling the electrical source to prevent machine operation and injury, this lockout/tagout rule does not cover exposure to electrical hazards (shock/electrocution). Electrical hazards are addressed under the later rule, "Electrical Safety-Related Work Practices" (ESWP), OSHA 29 CFR 1910.333.

The lockout/tagout rules apply to general industry and requires employers to implement the procedures which are specified in 1910.147, based on the workplace hazards that are encountered. The potential hazards at a broadcast facility under this rule would include, for example, heating/ventilation/air-conditioning (HVAC), moveable satellite antennas, blower motors (especially in transmitters), television commercial spot players, and elevators in buildings and on towers.

Of this group, HVACs and elevators are generally serviced by outside contractors who would be responsible for the lockout/tagout of those items. This is covered under 1910.147(f)(2) Outside Contractors. Under some conditions, the place-of-business may want to place a secondary lock/tag on the equipment being serviced by others if employees are also working or potentially involved by being on/near the locked-out/tagged-out equipment. Employers are required to become familiar with the lockout/tagout procedures used by the contractor.

Under 1910.147, a lockout/tagout program must now be in effect, and those individuals who have the responsibility/authority to lock/tag equipment must be instructed/trained concerning the procedure. One may include the lockout/tagout requirements of 1910.333, (as permitted by OSHA page 32002, FR, August 6, 1990), with those of 1910.147.

The ESWP standard covers electrical safety related work practices of employees who work on, near, or with electric circuits and equipment. Training is required for all qualified persons. A qualified person is someone who is well acquainted with and thoroughly conversant with the electrical/electronic equipment and the electrical hazards involved in the work being performed. Electrical/electronic technicians and engineers are defined as qualified persons by the rules.

As discussed under History, the rule requires the lockout/tagout of electrical/electronic equipment to prevent injury. Lockout/tagout does not apply to equipment which is normally bench maintained and electrically connected with a power cord.

The ESWP rule became effective December 4, 1990 with the training requirement of Section 1910.332, training became effective August 6, 1991. The rules should be available for employees to review.

One must have appropriate OSHA approved tags and locks. The tags and locks are available through a number of sources, catalog and local suppliers.

For additional information see:

"Electrical Safety-Related Work Practices, Final Rule," OSHA, 29 CFR 1910, FR, August 6, 1990, pages 31983–32020.

"Control of Hazardous Energy Source (Lockout/Tagout); Final Rule," OSHA, 29 CFR Part 1910, FR, September 1, 1989, pages 36643–36696.

Contact:

Department of Labor, Room N3649, 200 Constitution Avenue NW., Washington, DC, 20210, (202) 523–8148. For copies of standards, OSHA Office of Publications, Room N3103, (202) 523–8148.

OCCUPATIONAL NOISE EXPOSURE

The occupational noise exposure rules became effective in 1983. A disc jockey probably won't turn down his headphone volume to be in compliance with the rules, even though the sound level exceeds the limit. If someone does listen to sound levels which exceed the OSHA levels, it would be a good idea to limit the power available to the headphones. A broadcaster who is asked to set up an earth-shaking PA system may want to become familiar with CFR 29 Part 1910.95. While normally operating audio systems would not be in violation, it is important to know that the rules exist and what the limits are.

Exposure to noise levels exceeding the permissible limits listed in CFR 29 Part 1910.95, can occur with the operation of some lawn equipment and chain saws. Under such circumstances, proper ear protection should be worn.

For additional information see:

"Occupational Noise Exposure," OSHA, 29 CFR 1910.95

POLYCHLORINATED BIPHENYLS (PCBs)

PCBs are listed by the EPA as toxic and persistent. Based on animal data, the EPA concludes that in addition to chloracne, PCBs may cause reproductive effects, developmental toxicity, and tumors (oncogenicity) in humans. PCBs do not biodegrade rapidly and accumulate in fatty animal tissue.

Should a capacitor explode all over the inside of a transmitter, the spill must be cleaned up, all of the cleanup material disposed of properly, and the site must be certified or sampled depending on the quantity of the spill. A broadcaster could be off the air during cleanup if an uncontaminated standby facility is not available. And if PCBs are in a fire, the use of the building could be lost. In transformers, PCBs are usually mixed with tri-chlorobenzene (TCB) to alter the viscosity of the PCB material. During a fire, dibenzofuran and dioxin which are more toxic that PCBs and TCBs can be generated.

The EPA can impose severe fines for not following the rules under the Toxic Substances Control Act (TCA) which was passed by Congress in 1976. Any authority not granted to the EPA under TCA is given under the Resource Conservation and Recovery Act, (RICRA).

The final EPA transformer rule was published July 19, 1988, in the Federal Register. The owners of PCB transformers which are in or near a commercial building need to take action. As of February 25, 1991, all radial and network PCB transformers and those with secondary voltages below 480 volts, not located in sidewalk vaults, must be equipped with electrical protection to prevent high current faults. By October 1993, all transformers below 480 volts in sidewalk vaults in use near commercial buildings must be removed from service.

The use of PCB transformers and capacitors is still permitted in TV and radio transmitters if the PCBs would be contained, should a leak occur, and there is no reasonable risk of injury to health and the environment. However, as stated earlier, there is risk in using items containing PCBs, and as time goes on it may become more difficult and costly to dispose of PCB items. The revised EPA Penalty Policy became effective April 9, 1990. The revised policy incorporates the enforcement related provisions of all PCB rules and policies to date, including the Notification and Manifesting Rule. The Policy will be used to calculate all penalties in all administrative actions concerning PCBs regardless of the date of the violation. For a copy of the Policy, contact: Environmental Assistance Division (TS-799), EPA, 401 M St. SW, Washington, DC 20460, (202) 554–1404.

On February 5, 1990, the EPA amendments to the disposal and storage regulations for PCBs became effective. Under this Rule, the EPA expanded the meaning of "generator of PCB waste" to include the owner of the PCB material at the time of disposal. Items unregulated for disposal, such as small capacitors (by EPA definition) are not included within this expanded meaning.

A generator of PCB waste must obtain an EPA identification (ID) number before turning the PCBs over to commercial storers, transporters, and disposers of PCBs who must also have an EPA ID number. An ID number can be obtained by filling out the required "Notification of PCB Activity," EPA form number 7710–53. Additionally, a manifest, EPA form 8700–22 must be prepared by the generator to accompany a PCB shipment. The generator must track the PCB waste under the EPA defined time requirements until the PCB waste is finally destroyed. Any deviations from the manifest and time requirements from day of shipment to disposal must be reported by the generator to the EPA Regional Administrator for the region in which the waste originated.

Generators of PCB waste do not relieve their potential for PCB disposal violations by entering into contracts with disposers. Also, any state or local requirements are not preempted by meeting the EPA requirements. Additionally, items exempted from notification by the EPA, may be required for notification under local or state regulations.

Many articles have been published concerning the extensive rules pertaining to PCBs. By this time the identification and labeling of PCB items should have been completed along with the required record keeping and monitoring. Therefore, rather than repeat the extensive PCB information here, please refer to those published articles which are identified below, since they are readily available.

For additional information see:

PCB Alert, NAB Today, September 22, 1986.

PCB Prohibition Deadlines, NAB Info-Pak, November 1987.

"Polychlorinated Biphenyls in Electrical Transformers: Final Rule," 40 CFR Part 761, EPA, FR November 26, 1990, pages 49043–45.

"Polychlorinated Biphenyls: Notification and Manifesting for PCB Waste Activities: Final Rule," 40 CFR Part 761, EPA, FR, December 21, 1989, pages 52716–52756.

"Polychlorinated Biphenyl Spill Cleanup Policy: Amendments and Clarifications," Final Rule: amendment and clarification of policy statement, EPA, 40 CFR Part 761, FR, October 19, 1988, pages 40882–4.

"Polychlorinated Biphenyls in Electrical Transformers:" Final Rule, EPA, 40 CFR Part 761, FR, July 19, 1988, pages 27322–29.

"Polychlorinated Biphenyls (PCBs) Manufacturing, Processing, Distribution in Commerce, and Use Prohibitions, EPA, 40 CFR Part 761, July 1, 1987, pages 194–246

Brad Dick, "Managing The PCB Risk," Broadcast Engineering, October 1988, pages 68–94.

Jack G. Pfrimmer P.E., "Identifying and Managing PCBs in Broadcast Facilities," Proceedings 41st Annual Broadcast Engineering Conference, NAB, 1987.

Contact:

Office of Toxic Substances, 401 M St. NW, Washington, D.C., 20460, (800) 424-9065, and in D.C., (202) 554-1404.

PROTECTION OF STRATOSPHERIC OZONE

The protection of stratospheric ozone has been receiving attention because of the suspected environmental effects. From the Federal Register, April 17, 1989, page 15228, Background, first paragraph:

> "The threat of depletion of the ozone layer from chlorine released chloroflorocarbons (CFCs) was first raised well over a decade ago by the scientific community. (Molina and Rowland 1974). The initial theory suggested that because CFCs are relatively inert, emissions of these chemicals would not break down in the lower atmosphere, but would instead slowly migrate to the stratosphere where they would break apart releasing chlorine. Once freed in the stratosphere, the chlorine would catalytically destroy ozone. From a health and environmental perspective, depletion of the ozone layer would allow more harmful ultraviolet radiation to penetrate the atmosphere and strike the earth's surface. This would increase the incidence of skin cancers and cataracts, suppress the human immune system, damage crops, forests, and aquatic systems, accelerate weathering of certain plastics, and increase the formation of ground level (tropospheric) ozone or smog."

On August 12, 1988, the EPA issued its final rule (53 FR 30566) implementing the Montreal Protocol, an international treaty to limit the worldwide production of CFCs and halons. On April 17, 1989, the EPA in a Notice of Proposed Rule Making considered the possibility of adding carbon tetrachloride to the list of ozone depleting chemicals.

The producers of CFCs and halon are required to decrease the production of these chemicals, and users will pay higher prices and will need to use alternate substances such as HCFC-123. However, the use of HCFC-123 is only an interim solution since its production is scheduled to cease by the year 2015. And the use of HCFC-123 requires equipment modifications and reduces the efficiency of the system. Manufactures are seeking replacements for CFCs, HCFCs, and halon.

For those who are using halon fire suppression systems, it is not too early to begin planning for a substitute for halon fire suppression. It is also likely the use of carbon tetrachloride will be curtailed and substitute solvents will need to be used.

In order to keep presently operating CFC refrigeration systems operating as the supply of CFCs diminishes and becomes much more expensive, it will be necessary to reclaim CFCs from both operating and abandoned refrigeration systems. Present air conditioning systems which release refrigerant into the atmosphere with required moisture purges will need to have CFC recovery units installed. Refrigeration systems which must be emptied of refrigerant for repair will need to have the CFC saved by pumping down the system rather than dumping into the environment. If the company that services one's air conditioning equipment has not approached one about reducing the loss of CFCs, one should initiate the CFC recovery conversation and perhaps if there is no interest, one should discuss it with another service company.

For more information:

"Protection of Stratospheric Ozone; Temporary Final Rule," 40 CFR Part 82, EPA, FR, March 6, 1991, pages 9517–31.

"Protection of Stratospheric Ozone; Notice," Notice Listing Ozone-Depleting Substances, EPA, FR, January 22, 1991, pages 2419–24.

"Ozone Depleting Chemicals; Toxic Chemical Release Reporting; Community Right-To-Know; Addition of Chemicals; Technical Amendment," 40 CFR Part 372, EPA, FR, August 30, 1990, page 35434.

"Ozone Depleting Chemicals; Toxic Chemical Release Reporting; Community Right-To-Know; Addition of Chemicals," EPA, FR, August 3, 1990, pages 31594–98.

"Protection of Stratospheric Ozone; Final Rule," EPA, 40 CFR 82, FR, June 22, 1990, pages 25812–14.

"Protection of Stratospheric Ozone; Final Rule," EPA, 40 CFR 82, FR, June 15, 1990, pages 24490–96.

"Protection of Stratospheric Ozone; Final Rule," 40 CFR Part 82, EPA, FR, February 9, 1989, pages 6376–6379.

Contact:

Daniel Blank, Program Analyst, Global Change Division, Office of Atmospheric and Indoor Air Programs, Office of Air and Radiation (ANR-445), EPA, 401 M Street, SW, Washington DC 20460, (202) 475-8894. David Lee (above address), (202) 475-7497.

RADON

Radon is a radioactive gaseous element which occurs in nature as a result of the natural disintegration of radium. It is a colorless, odorless gas and is found in various levels of concentrations in the ground and ground water. Levels vary in different parts of the United States. Radon can accumulate in basements and buildings by seeping through porous or cracked floors/foundations. Concentrations of radon in structures can expected to be higher if there is insufficient outside air exchange, sometimes referred to as the tight building syndrome. In homes especially, levels of radon can be expected to be higher in the winter.

Concentrations of radon gas are considered to be a health hazard. There does not seem to be agreement regarding the level of radon to be considered dangerous. The EPA presently warns that radon levels exceeding 4 picocuries per liter of air pose a health threat. In Canada, action to reduce levels is not considered necessary until the level exceeds 20 picocuries. The average outdoor level is 0.2 picocuries.

Some states are beginning to pass legislation which requires a buyer or lessee to be notified of the possible threat of radon gas, and some states are passing laws to regulate and license radon testers and mitigators.

Testing for radon gas can be accomplished by placing radon testing canisters in areas where concentrations may be anticipated. The results can vary drastically, depending on the season of the year and the location(s) selected for measurement. If one is placing the canisters oneself rather than having a recognized testing service perform the task, it is suggested that in addition to the test canisters placed within the area of interest, a test control canister be placed outdoors as a calibration check.

To reduce radon concentrations, cracks in foundations and floors must be sealed, and drains must be checked for potential gas leakage. Porous foundation surfaces (such as concrete block) and floors must be sealed.

UNDERGROUND STORAGE TANK REGULATIONS (USTs)

The underground storage tank regulations were published in final rule form September 23, 1988, in 165 pages of the Federal Register. An EPA Form 7530–1 or a similar State Certificate of Notification should already be filed with the environmental agency of the state. Beginning December 22, 1988, the owner/operator of an underground storage tank (UST) needed to begin keeping records which include the date and nature of inspections, repairs, and replacement of USTs in use. Within a ten year period of December 22, 1988, depending on the age of the tank, the USTs must be made corrosion resistant. Since October 24, 1988, any person who sells a tank to be used as an underground storage tank must notify the purchaser of the requirement to report bringing the tank into use.

An underground storage tank is defined as a tank and piping which has at least 10% of its volume underground. There are a few EPA exclusions (not necessarily state exclusions) to the definition of USTs, which at least presently include:

1. Farm or residential tanks of 1,100 gallons or less capacity, storing motor fuel for noncommercial purposes.
2. Tanks storing heating oil for use on the premises where stored.
3. Septic tanks.
4. Storage tanks situated on or above the floor in basements.
5. Tanks that have a capacity of less than 110 gallons (not excluded in the Ohio State regulations for example).

The UST rules define new tank design, construction, and installation. Also defined are the requirements for leak detection, spill prevention, record keeping, reporting, tank closure, and the corrective action for spills. Under the Superfund Amendments and Reauthorization Act of 1986 (SARA), the EPA is given authority to clean up petroleum releases or to require owners or operators to do so. If the EPA cleans it up, do not expect them to seek a low bid and do expect an invoice and a fine.

An EPA study indicated that 42% of the tanks which were 15 to 20 years old and 30% of the 10 to 15 year old tanks were leaking. Release detection at existing UST systems must be added. Older tanks which are unprotected from corrosion must have release detection as of December 22, 1989, with new tanks that are protected from corrosion to have release detection within five years. A protected tank is one with proper cathodic protection or a tank made of an acceptable material, such as fiberglass over steel. Periodic tank tightness testing every five years combined with monthly inventory control is allowed at new tank installations for ten years after installation, but after ten years, monthly release detection is required. Monitor wells are a method of leak detection and are required by some states and may be in addition to whatever other method is used. The above are federal regulations. Many states have requirements which are more stringent.

Tank owners and operators must report actual and suspected releases. Owners and operators of leaking UST systems must follow measures for corrective action. Cleanup levels will be established on a site-by-site basis as approved by the implementing agency. Tanks over ten years of age must be either internally inspected or lined to meet upgrade requirements. This procedure may not be practical for small tanks. It may be more practical to replace an older tank with a new protected one. When a new tank is installed, the installer must certify that the UST system was installed according to the rules. All substandard existing UST systems must be closed, replaced, or retrofitted with corrosion protection within ten years of December 22, 1988 which are also the years of grace before adding overfill protection to all tanks. The required overfill protection leaves much to the discretion of the operator but to wait ten years before installing protection burdens the environment, and could burden the operator financially if there is a spill. Overfill protection can be as simple as automatic supply shut-off when the tank is 95% full. At some locations overspill protection, such as a retaining basin, may be appropriate. If more than 25 gallons are spilled, the spill must be reported. No exceptions here. Failing to report a spill can result in very large fines. In all cases, the spill or overfill must be cleaned up immediately, and if not, even spills under 25 gallons must be reported.

This is only a small sample of the regulations. In addition to the EPA rules, there are state/city regulations which additionally may impose requirements not contained in the EPA rules. Many states had rules which were in effect before the EPA wrote theirs. For example, although tanks associated with emergency generators are currently deferred from release detection by the EPA, they are not deferred in Florida. It really is not wise to defer meeting the requirements of the rules since the consequences of a spill or leak can be severe to both the environment and the bottom line. The cost to replace a 1,000 gallon storage tank in Cleveland, including monitor wells, is around $12,000. If it would be necessary to remove tons of contaminated soil, the costs could be unaffordable.

If a tank is leaking, it must be replaced and any contaminated soil removed. It may be a better choice to replace aged tanks rather than wait until a leak starts. If a tank is buried and not leaking and its removal would be too expensive or not practical, it may be emptied by removing liquid and sludge, and then filled with inert solid material. The choice of inert fill material is not specified. The EPA accepts sand and concrete, some states do not accept both. However, concrete could cause a future construction problem. Many states will dispatch an inspector to review a tank installation before the tank may be filled. The state often requires drilled soil samples in the area of the tank before filing is permitted.

Civil fines for violation of the regulations can be up to $10,000 a day with criminal fines topping $25,000 a day. Corporate officials face fines up to $250,000 and up to 15 years in prison, with corporate fines running to $1 million. Additionally, the clean-up costs of an underground water system could be astronomical.

The EPA addressed financial responsibility in the October 26, 1988, issue of the Federal Register. All UST owners and/or operators, including owners of USTs that are used to store fuel for generators, must be in compliance. This responsibility requires owners, such as broadcasters, to maintain financial assurance of at least $500,000 per occurrence, $1 million aggregate, to insure that damage resulting from a leak will be paid for by the owner.

Some states have established financial assurance funds which require UST owners to file a yearly statement along with an assessment fee.

For the list of agencies designated to receive notifications, refer to the January 1989, NAB UST Info-Pak.

On April 9, 1990, the National Oil and Hazardous Substances Pollution Plan became effective. The 200-page document was printed in the Federal Register March 8, 1990. The purpose of this National Contingency Plan (NCP) is to provide the organizational structure and procedures for preparing for and responding to discharges of oil and releases of hazardous substances, pollutants, and contaminants. The NCP would apply in cases of the discharge of oil into or upon navigable waters of the United States and releases into the environment of substances which may present an imminent and substantial danger to public health and welfare. Thus under certain conditions, an oil spill could fall under the NCP.

For additional information see:

"Underground Storage Tanks: Technical Requirements," Technical Amendment, 40 CFR Part 280, EPA, FR, April 27, 1990, pages 17753–17754.

"National Oil and Hazardous Substances Pollution Contingency Plan: Final Rule," 40 CFR Part 300, EPA, FR, March 8, 1990, pages 8666–8865.

"Underground Storage Tanks Containing Petroleum: Financial Responsibility Requirements," Interim Final Rule, 40 CFR Part 260, EPA, FR, November 9, 1989, pages 47077–47082.

"Administration Penalty Procedures: Interim Final Rule," 40 CFR Part 22, EPA, FR, May 16, 1989, pages 21174–21177.

"New EPA Requirements for Underground Storage Tanks," Counsel from the Legal Department, NAB Info-Pak/January 1989.

"Underground Storage Tanks Containing Petroleum-Financial Responsibility Requirements and State Program Approval Objective," 40 CFR 280 and 281, EPA, FR, October 26, 1988, beginning page 43322.

"Underground Storage Tanks: Technical Requirements and State Program Approval;" Final Rules, 40 CFR Parts 280 and 281, EPA, FR, September 23, 1988, pages 37082–37247.

Contact:

RCRA/Superfund hotline (800) 424-9346 and in D.C., (202) 382-3000.

In summary, this brief collection of information and reference materials has been pulled together to alert the broadcaster to the EPA and OSHA rules which affect the industry and are of an environmental concern. This information is far from complete. Lengthy seminars are often given on just a single subject. There are companies whose main business is the preparation of materials for compliance with the various regulations. Other companies are in the business of meeting environmental clean-up needs. Further rule changes can be expected as well as additions to the rules.

1.6
Consultant Services: When and How to Use Them

Warren M. Powis, P.E.
Cohen, Dippell and Everist, P.C., Washington, District of Columbia
(Chapter based on material prepared by Alan E. Gearing, P.E. and John F.X. Browne, P.E.)
Association of Federal Communications Consulting Engineers (AFCCE)[1]
Washington, District of Columbia

RANGE OF CONSULTING SERVICES

Broadcast applicants and licensees employ the services of consultants to assist with various aspects of the conduct of their business such as engineering, financial, legal, marketing, personnel, and programming areas. The need for these services may arise for many reasons including the unavailability of staff, the need for expertise not found in-house, or the need for additional manpower to ensure timely completion of a project. Broadcasters selecting engineering or technical consultants may be unfamiliar with the variety of services normally offered by consultants or with the procedures for selecting and engaging a consultant.

Services offered by consultants ranges from evaluation and recommendation, to supervision of installation, to complete turnkey responsibilities. The types of consultants range from those handling transmission and general broadcast engineering matters to those specializing in narrowly defined areas. Examples of the latter include experts in broadcast building design and acoustics, tower design and maintenance, and telephone system planning.

Tables 1, 2, and 3 contain a partial listing of the types of services provided. Table 1 includes services applicable to all types of broadcast facilities. Services specific to AM broadcast stations are listed in Table 2, and Table 3 lists services mainly applicable to FM and TV broadcast stations.

Many consultants offer services of the types listed in the tables for other communications activities in which some broadcasters may also be involved, such as:

- Instructional Television Fixed Service Systems
- Multichannel Multipoint Distribution Systems
- Satellite Up Link, Down Link, and Space Facilities
- Microwave Relay Systems
- Paging Systems
- Land Mobile Systems
- Cellular Radio Systems
- International Broadcast (Shortwave) Facilities
- Subcarrier Services
- LPTV and Translator Systems
- Remote Pickup and other Auxiliary Systems
- Fiber Optic Systems

QUALIFICATIONS, REGISTRATION, AND CERTIFICATION

Some of the activities engaged in by consultants may affect the public in matters of health, safety, and welfare, such as radio frequency power density computations and measurements. Therefore, various means have been used to provide some assurance to the public that practitioners have at least threshold qualifications. For example, the states and some professional organizations have established minimum qualifications for persons acting in certain professional capacities including the practice of engineering. All fifty states and the District of Columbia have engineering registration laws.

The practice of broadcast and telecommunications consulting engineering is, with few exceptions, a professional field in which most states require the practitioner to be registered as a *professional engineer* or "PE". In most cases, in fact, registration is required by law in order to represent oneself as a consulting engineer or to provide professional engineering services.

Professional engineers are registered or licensed through a process involving verification of educational credentials (e.g., a degree in engineering or related science), the attainment of specific engineering experience (usually a minimum of four years, post graduation), and passing an examination gauging the applicant's mastery of engineering fundamentals and his or her specific field of engineering practice (usually a two-day, 16-hour written exam).

The services listed in the Tables at the end of this chapter include both those which fall within the realm of the professional engineer and those which may not require such certification. While the guidelines for selecting a consultant, contained in this chapter, refer to the professional engineer or consulting engineer, they can, and should, be used as a general guide in the selection of any type of consultant service.

CODE OF ETHICS

Professional Engineers usually subscribe to a code-of-ethics such as those adopted by the National Society of Professional Engineers (NSPE) and the Association of Federal Communications Consulting Engineers (AFCCE). Many engineering societies representing the various engineering disciplines also support ethical codes for their members.

These codes require a standard of professional conduct which govern the consultant's activites and his relationship with clients in areas such as:

- Advertising
- Solicitation of clients
- Conflicts of interest
- Confidentiality of client activities
- Practice in field of expertise
- Use of the work of other engineers

Copies of these representative codes of ethics may be obtained free-of-charge from the National Society of Professional Engineers, 1420 King Street, Alexandria, Virginia, 22314 and the Association of Federal Communications Consulting Engineers, P.O. Box 19333, 20th Street Station, Washington, D.C., 20036.

SELECTING A CONSULTING ENGINEER

Selection of a consulting engineer does not have to be a confusing formal process. Initial contacts with prospective consulting engineers are often made by referral from communications attorneys or on the recommendations of other broadcasters. Additionally, "professional cards" in trade publications or referrals by engineering societies are sources of listings on professional practitioners.

When selecting a consulting engineer the broadcaster should, as a minimum, discuss the following topics with the candidate consultant:

- Professional qualifications and registrations
- Knowledge of FCC Rules and Regulations and international agreements
- Expertise in the particular area of interest
- Recent experience
- Availability of professional staff, time, and facilities
- Estimated time for completion
- Estimated fees or fee structure

1. Seek names of recommended individuals or firms from people within industry.
2. Review professional cards in magazines.
3. Obtain copy of AFCCE Membership Directory.
4. Select several candidates and discuss your requirements with each. Determine with each discussion if that individual/firm is experienced in providing services required by your project and can work in the desired time frame. If considering a firm, obtain the name of an assistant who can be contacted in the event that the prime contact is unavailable.
5. For portions of work not requiring a consultant, obtain from the consultant recommendations of companies who he believes provide that work.
6. Determine the consultant's hours. Are they compatible with your schedule? Do they have facsimile (fax), answering systems, etc., if the need arises to contact them at irregular hours?

Figure 1. Selecting a consultant.

WORKING WITH A CONSULTING ENGINEER

Once a consulting engineer has been selected for a specific project, an agreement regarding the scope of work, timetables, and fees is normally executed. In its simplest form this may merely be an exchange of informal letters; on complex projects, particularly those involving large scale construction, there may be a need for formalized contractual documents. A retainer fee to be applied against initial work may also be required.

There should be a clear understanding between the parties as to the scope of work and responsibilities of each party, particularly if the client (broadcaster) desires to perform part of the engineering work in-house or plans to employ other consultants for portions of the project. There is a great potential for both duplicative efforts and omissions in these arrangements when areas of responsibility are not clearly defined. While the initial objective may have been to minimize the consulting engineer's time (and fees) the result could very well be the opposite.

To assist the consulting engineer in the performance of his work, the client should provide any and all background material as well as earlier engineering documentation that may have been developed on the project. It is important that the client recognize that the use of such information is at the discretion of the consultant, and that a fee quoted on the basis of the validity and applicability of such data to a new project may have to be adjusted if the data is found to be unsuitable or inaccurate. In some contexts, such as FCC filings, the consultant must attest to the accuracy of the data being presented and, therefore, he must have a high level of confidence in data which he did

1. Establish rapport with the consultant and discuss your goals and requirements.
2. You can provide items which will assist the consultant. Prompt site selection is desirable, and information provided should include the following details:

- Property boundaries of proposed site
- Address(es) of proposed site
- Coverage goals if a broadcast station
- If using an existing tower, pertinent structure information including transmit and receive frequencies, and location of other antennas (including manufacturer and model numbers)
- Availability of power
- Site accessibility

Figure 2. Using your consultant wisely.

not personally generate in order to incorporate it. There may also be legal or ethical problems associated with incorporating the engineering work product of other engineers, especially that prepared by non-registered or unqualified practitioners.

Changes in technical details which may appear to be minor to the client often have major impact on the work completed by the consulting engineer. A transmitter site relocation of only a mile, for example, may involve recomputation of service areas, operating parameters, interference protections or separations, or aeronautical impact. For these reasons the client should attempt to finalize such details—with guidance from the consulting engineer—as early as possible, but especially before authorizing the completion of final reports or applications.

FCC Applications

In many cases FCC applications must be filed before certain dates which are established by processing procedures or mandated in FCC docket actions. Since the preparation of these filings can be a rather lengthy, the client should give the consulting engineer sufficient advance notice of such filing deadlines.

In the absence of formal arrangements to the contrary, consulting engineers do not generally assume the responsibility of monitoring FCC actions on behalf of present or former clients, to note activities (such as competitive filings, notices, or grants) which could have an impact on a client's existing or proposed facility. Such monitoring is often performed by the station's legal counsel and is also available directly from some of the computer database firms serving the broadcast industry. One way in which consulting engineers do provide such monitoring, relative to potential engineering impact, is through retainer agreements in which the client employs the consultant's services under an ongoing arrangement. The agreement usually specifies the level of services, such as FCC monitoring, that will be provided and the method of reporting to be used. In some cases, such agreements also specify that the consultant will make available a specified minimum amount of staff time which may be used over the course of the agreement to respond to the client's needs as they arise.

Data Rights

The client should have an understanding with the consultant regarding ownership of documentation produced by the consultant. It is common practice for the consulting engineer to retain all rights to files, drawings, reports, and other work products. The consultant grants to the client permission to use such data for specified FCC filings and for retention as part of the station's files. Normally, the data cannot be transferred to another party, or copied and used as part of another filing, or incorporated as part of a design for another project or facility without the express consent of all parties.

There are some consultants offering services which are oriented to field projects, such as equipment installation, construction, and measurements. These persons or firms operate as extensions of the broadcaster's staff and usually can bring a wealth of practical experience and additional manpower that the broadcaster may not have in-house. Professional engineering certifications may not be required for these services, but the possession of an appropriate FCC Commercial Operator's license may be required if transmitter related services are performed. Any measurements made by these consultants, which will ultimately be used in connection with FCC filings, should be collected, assembled, and presented (including certifications) in an acceptable form. Coordination of these activities with the station's consulting engineer may be desirable. References concerning the qualifications and capabilities of such consultants may be obtained from the station's consulting engineer or the other referral sources listed earlier.

TABLE 1

Services applicable to all types of broadcast facilities.

Engineering Allocation Studies and Applications

- Allocation studies to determine the feasibility of establishing new stations at specified communities.
- Allocation studies to determine the feasibility of improving existing facilities.
- Determination of appropriate areas for transmitting facilities to be located.
- Consideration of air hazard aspects and preparation of notice to FAA of proposed construction.
- Preparation of engineering portions of applications to the FCC for construction permits and licenses.
- Review of engineering proposals by other stations or applicants
- Engineering in support of petitions to deny or replies to petitions to deny.
- Provision of expert testimony in FCC hearings or court proceedings or local authority hearings.

- Population and area service analyses.
- Determination of existing and proposed services in a particular area
- Environmental and Electromagnetic Compatibility
- Environmental impact studies.
- Electromagnetic interference studies to determine whether a proposed new or improved facility would cause interference to an existing station or to another proposed new or improved facility.
- Electromagnetic radiation power density level computations and measurements.

Transmission Systems

- Preparation of specifications for transmitting systems.
- Transmitter plant design and construction.
- Measurement and adjustment of transmission systems to establish proper operation and compliance with FCC Rules.
- Analysis and design of special-purpose antennas, waveguides, transmission lines, cavity resonators, and other electromagnetic devices.
- Antenna tower planning, including placement of antennas, intermodulation studies, and multicoupler and filter requirements to avoid interference.
- Tower design, erection, and maintenance.
- Structural analysis of towers and antennas.

Studio Systems

- Evaluation and recommendation of equipment for studios, newsrooms, remotes, and automation.
- Studio design and construction including proper acoustics, sound proofing, lighting, heating and air conditioning.
- Design and installation of in-house telephone systems.

Miscellaneous Services and Systems

- Appraisals and on-site inspections of facilities for buyers, sellers, and brokers.
- Mock station inspections to check compliance with FCC Rules and Regulations.
- Field review of station facilities to determine compliance with FCC Rules and to determine the general condition of station facilities.
- Development of operational and monitoring procedures, including specification of test equipment.
- Evaluation of manufacturer's equipment bids.
- Design and installation of equipment for lightning protection.
- Training of technical personnel.
- Preparation of specifications for, installation of, and maintenance of computer systems.

TABLE 2

Services applicable to medium frequency AM broadcast stations (530 to 1700 kHz).

- Measurements of RF field strength relative to allocation matters.
- Design of directional antenna systems.
- Design of phasing and coupling equipment for directional antenna systems.
- Adjustment of directional antenna systems (proof of performance).
- Investigation of effects of reradiating obstacles (buildings, power lines, water towers, utility poles, irregular terrain, nearby antennas and arrays) upon antenna performance.
- Design of matching systems for multiple operations into a single vertical radiator.
- Design of special filters or detuning networks which may be required.
- Investigation of electromagnetic effects of guy wires upon AM broadcast antennas and arrays.

TABLE 3

Services applicable to FM and TV broadcast stations.

- Study of feasibility of adding allotments to FCC Table of Allotments.
- Engineering support for petition for rulemaking.
- Review of engineering aspects of comments and providing engineering support for reply comments.
- Study of and recommendations for optimum antenna height, radiated power, tower location, and vertical and horizontal radiation pattern characteristics.
- Propagation studies, including use of alternate models and accounting for actual terrain features.
- Field strength and signal quality surveys of existing facilities.
- Preparation of specifications for antenna systems including review of manufacturer's range test results.
- Investigation of electromagnetic effects of support towers and guy wires upon TV and FM broadcast antenna.

ENDNOTE

1. Full members of AFCCE are registered professional engineers who subscribe to the Code of Ethics for engineers as adopted by the National Society of Professional Engineers. A directory listing of AFCCE full members and individual firms is available upon request.

1.7 Source Guide: Broadcast Standards and Information

David K. Bialik
WKDM Radio
United Broadcasting Co., New York, New York

Section on Broadcast Acronyms contributed by
John Reiser
Federal Communications Commission, Washington, District of Columbia

INTRODUCTION

The purpose of this chapter is to provide an easily accessible guide to organizations whose activities affect the work of broadcast engineers. This includes government agencies, trade and professional associations, and private companies.

FEDERAL GOVERNMENT

Department of Agriculture

14th and Independence Avenue, Southwest
Washington, D.C. 20250
(202) 447-2791

The Department of Agriculture (USDA) Forest Service oversees the use of 190 million acres of land under the domain of the National Forest System. More often than not, the use of a mountaintop transmitter site is controlled by the Forest Service. Land use permits are obtained from the forest supervisor for the particular forest of interest. The forest supervisor has detailed policy directives for permissible uses and fee structures of forest land.

Department of Commerce

National Institute of Standards and Technology (NIST)

Gaithersburg, MD 20899
(202) 921-1000

The National Institute of Standards and Technology of the Department of Commerce has been authorized by Congress to undertake the following functions: "The custody, maintenance, and development of the national standards of measurements and the provisions of means and methods for making measurements consistent with those standards, including the comparison of standards used in scientific investigations, engineering, manufacturing, commerce, and educational institutions, with the standards adopted or recognized by the Government."

Time and Frequency Division. The Time and Frequency Division, located in Boulder, Colorado, is the part of the NIST that carries out the above functions which relate to time and frequency. This division is responsible for distributing the standards and for finding new and improved methods of dissemination. The dissemination services for time and frequency are presently available from stations WWV and WWVB in Fort Collins, Colorado, and from WWVH in Kauai, Hawaii. In addition, services using network television and satellite signals are also available. Correspondence pertaining to station operations may be addressed to: Engineer-in-Charge, NIST Radio Stations WWV and WWVB, 2000 East County Road 58, Fort Collins, CO, 80524, (303) 484-2372. Or, Engineer-in-Charge, NIST Radio Station WWVH, P.O. Box 417, Kekaha, Kauai, HI, 96752, (808) 335-4361.

National Telecommunications and Information Administration

14th Street and Constitution Avenue N.W.
Washington, D.C. 20230
(202) 337-1551

The National Telecommunications and Information Administration develops policy for the executive branch; manages federal use of radio spectrum; conducts technical research on radio wave transmissions and other aspects of telecommunications; and serves as an information source for federal and state agencies on the efficient use of telecommunications resources.

Department of Interior

The Bureau of Land Management of the Department of Interior oversees the use of 360 million acres of "public domain" lands which are not under the domain of the USDA Forest Service. Broadcaster use of these lands for transmitter sites is coordinated by the State director of the Bureau of Land Management.

Topographic Maps

Topographic maps usually used in the prediction of coverage and other engineering studies, which require accurate information about the position and elevation of terrain features, may be obtained from the U.S. Geological Survey (USGS), a bureau of the U.S. Department of Interior. Standard topographic quadrangles on a scale of 1:24,000 (7.5 minutes) or 1:62,500 (15 minutes) are usually used, but other scale maps are available. To order free map indexes or to purchase maps, contact one of the following offices.

For areas east of the Mississippi River (including Minnesota, Puerto Rico, and The Virgin Islands): Eastern Distribution Branch, U.S. Geological Survey, 1200 South Eads Street, Arlington, VA, 22202.

For areas west of the Mississippi River (including Alaska, Hawaii, Louisiana, Guam, and Samoa): Western Distribution Branch, U.S. Geological Survey, Box 25286, Federal Center, Denver, CO, 80225.

Department of Labor

200 Constitution Ave., N.W.
Washington, DC 20210
(202) 523-6666

Broadcast engineers should be aware that two federal laws, the Federal Labor Standards Act and Occupational Safety and Health Act, empower the Department of Labor (DOL) to regulate work place safety standards and the wages and hours of employment for broadcast employees. Occupational safety standards are enforced by OSHA, the Occupational Safety and Health Administration of DOL, while the minimum wage and hours of employment are enforced by DOL's Wage and Hour Division.

Applicability of OSHA and Wage and Hour Division regulations depend upon many factors; such as the station's geographic location, the number of station employees, and the type of work they perform. Broadcast engineers should consult their station's attorney to determine what regulations apply to them and how to abide by them.

Department of Transportation

Federal Aviation Administration
400 7th Street, S.W.
Washington, D.C. 20590
(202) 426-4000

When construction of a new broadcast tower or a change in height of an existing tower is proposed, the Federal Aviation Administration (FAA) becomes involved by studying the potential impact that the proposed structure may have on navigable airspace. Broadcasters are required to notify the FAA of proposals for new or modified tower structures. The FAA applies its obstruction standards (FAA Rules Part 77) to determine whether or not the tower constitutes a hazard to air navigation. Broadcasters may need to accommodate an adverse finding by the FAA with a modified tower location and or tower height. When the proposal is granted by the FCC and the FAA, the license will contain specifications for marking and lighting the structure, if it is required. The FAA must receive notice immediately should any tower lighting systems fail.

Environmental Protection Agency

401 M Street, S.W.
Washington, D.C. 20460
(202) 382-2090

When broadcasters construct station facilities such as antenna towers or large satellite earth stations they should be aware that these activities may fall under the scope of the National Environmental Protection Act (NEPA). NEPA grants the Environmental Protection Agency (EPA) authority to regulate activities which may affect the "quality of the human environment." The FCC cooperates with EPA in enforcing provisions of NEPA that relate to telecommunication licensees.

The FCC monitors construction of telecommunications facilities and requires information about the operation of transmitters and other broadcast equipment. For example, the FCC, with the assistance of EPA, monitors the level of employee and general public exposure to nonionizing radiation emitted from the equipment. Generally, licensees who propose to construct such facilities must file studies of the potential environmental consequences with the FCC and may be required to take steps to minimize possible environmental hazards.

Federal Communications Commission

1919 M Street N.W.
Washington, D.C. 20554

For information on the Federal Communications Commission, see Chapter 1.1, "FCC Organization and Administrative Procedures," and Chapter 1.2, "FCC Field Operations Bureau."

Every station should have a current copy of the FCC's broadcast rules. There are two ways to obtain them:

1. Order a set of the Code of Federal Regulations (CFR) issued annually by the U.S. Government Printing Office. Call them at (202) 783-3238 and order "47 CFR Parts 0–19" for the Tower Regulations in part 17 and "47 CFR Parts 70–79" for the Broadcast Rules Part 73 and Broadcast Auxiliary Rules Part 74.

The CFRs may also be ordered by writing the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

2. For a continuously updated subscription, call the Rules Service Company at (301) 424-9402. Parts 17,73 and 74 in loose-leaf form (updated quarterly). Or write to them at: 7658 Standish Place, Suite 106, Rockville, MD 20855.

National Labor Relations Board

1717 Pennsylvania Avenue, N.W.
Washington, D.C. 20570
(202) 254-9044

Broadcast engineers should be aware that the National Labor Relations Act (NLRA) protects the right of station employees to bargain collectively with management over the "terms and condition of employment." The National Labor Relations Board (NLRB), an independent federal agency, was established to enforce the right of workers to organize and engage in "concerted activity." As with the Department of Labor regulations, the extent of these rights depends upon many factors such as the type and the number of employees involved.

STATE AND LOCAL GOVERNMENTS

A number of aspects of a broadcast engineer's job are affected or controlled by state or local government agencies. It is important for the engineer to have some familiarity with the laws, codes, and zoning ordinances governing such matters as building construction, electrical wiring, and fire safety. Regulations may vary from one community to another, even within the same state or county.

While there are model national codes, these codes may or may not be adopted by a state or local government. If adopted, there may be some changes from the national model. The only way to determine this is to check with the local agency or agencies having jurisdiction over the matter in question. For more information on state or local regulations see the following:

- The broadcast station's local lawyer
- The county or city business licensing office
- The county or city building inspector, fire marshall
- A licensed local contractor who does the type of work in question

TRADE ASSOCIATIONS

Association of Federal Communications Consulting Engineers (AFCCE)

P.O. Box 19333
20th Street Station
Washington, DC 20036–0333

The AFCCE is a professional association of communications engineers practicing before the Federal Communications Commission. Engineering for broadcast stations in the AM, FM, and TV services, for microwave, cellular radio, paging systems, and for satellite facilities are some of the areas in which AFCCE members offer their professional services. Full members of the AFCCE are registered Professional Engineers in the jurisdictions where they practice. For more information on AFCCE and consultants, in general, see Chapter 1.6, "Consultant Services: When and How to Use Them."

Association for Maximum Service Television (MSTV)

1400 16th Street, N.W.
Suite 610
Washington, DC 20036
(202) 462-4351

The purpose of MSTV is to assure the maintenance and development of an effective nationwide system of free over-the-air television based on local broadcast stations providing service of high technical quality. MSTV seeks to meet both the present and future needs of developing an effective nationwide VHF and UHF system; to provide free over-the-air broadcasting; to develop local stations to reflect the needs and interests of local communities; to safeguard against broadcast of only a few purely national voices; and to protect against interference and degradation of the public's broadcast service. MSTV develops and preserves opportunities for local stations to use new technologies.

Advanced Television Systems Committee

1776 K Street, N.W.
Suite 300
Washington, DC 20006
(202) 828-3130

The Advanced Television Systems Committee (ATSC) was established in 1983 to develop and coordinate voluntary national technical standards for advanced television systems, and to recommend to the U.S. Department of State positions for the United States within international standards organizations such as the International Radio Consultative Committee (CCIR).

The ATSC is composed of fifty member companies, all of which are major factors within the United States television industry. These members include broadcast networks and stations, cable television companies, program producers, professional and consumer equipment manufacturers, and satellite communications companies.

Cable Television Laboratories Inc. (CableLabs)

1050 Walnut Street
Suite 500
Boulder, Colorado 80302
(303) 939-8500

CableLabs is a research and development consortium of cable television system operators representing more than 85% of the cable subscribers in the United States.

Committee for Digital Radio Broadcasting (CDRB)

c/o Society of Broadcast Engineers
P.O. Box 20450
Indianapolis, IN 46220
(317) 253-1640

The role of the CDRB is to support the development of a digital radio broadcast system in the United States. A viable system should provide radio broadcasts with a sound quality comparable to compact discs in a cost effective manner. This technology could be used to augment the current distribution system of radio broadcasting via the FM and AM bands in the United States. The CDRB will:

1. Sponsor demonstrations of digital radio broadcast technology in the U.S.
2. Analyze current developments in digital radio broadcasting.
3. Analyze legal and practical implications of a system on the existing U.S. radio broadcast industry.
4. Determine what elements are needed to identify appropriate spectrum and implement a viable system in the U.S.
5. Promote and support research and development of digital radio broadcasting in the U.S.
6. Advocate spectrum allocation.
7. Serve as industry liaison to the government, consumer electronics industry, and international organizations.
8. Promote industry consensus on digital radio broadcast standards.

Participation in committee activities is open to all interested parties with a direct interest in the development of digital radio broadcasting in the United States, including broadcasters, manufacturers, associations, and laboratories.

Electronic Industries Association (EIA)

2001 Eye Street, N.W.
Washington, DC 20006
(202) 457-4900

The Electronic Industries Association is a trade association representing the manufacturers of telecommunications; industrial and consumer equipment; components (parts, tubes, and solid state products); and equipment and systems for the government. Their engineering activity includes the publication of voluntary standards. Those of interest to broadcasters are standards covering transmitters, towers, microwave transmission systems, and tape cartridges. The standards work is primarily done in the Broadcast Television Systems Committee (BTSC), the National Radio Systems Committee (NRSC), and standing engineering committees. The BTSC and NRSC are joint engineering committees with the National Association of Broadcasters (NAB). The catalog of standards and engineering publications may be obtained from: Standard Sales Department, Electronic Industries Association, 2001 Eye Street, N.W., Washington, D.C., 20006, (202) 457-4966.

A selection of EIA voluntary standards of interest to broadcast engineers is given below along with the date of acceptance or reaffirmation. Interested readers should contact EIA for a current listing and prices.

EIA-189-A Encoded Color Bar Signal. 1976.

EIA-200-A Circular Waveguides. 1975.

EIA-211-D Processed Analog Disc Records and Reproducing Equipment. 1981.

EIA-215 Basic Requirements for Broadcast Microphone Cables. 1958.

EIA-219 Audio Facilities for Radio Broadcasting. 1959.

EIA-221-A Polarity or Phase of Microphones for Broadcasting, Recording and Sound Reinforcement. 1979.

EIA-222-D Structural Standards for Steel Antenna Towers and Antenna Supporting Structures. 1987.

EIA-225 Rigid Coaxial Transmission Lines. 1975.

EIA-232-D Interface Between Data Terminal Equipment and Data Circuit Terminating Equipment Employing Serial Binary Data Interchange. 1986.

EIA-238-B Standards for Stylus Tips Used for Disc Phonograph Record Reproducing. 1981.

EIA-250-C Electrical Performance Standards for Television Relay Facilities. 1977.

EIA-258 Semi-flexible Air Dielectric Coaxial Cables and Connectors. 1962.

EIA-261-B Rectangular Waveguides. 1979.

EIA-264 Magnetic Recording Tape Cartridge Dimensions. 1962.

EIA-288 Audio Magnetic Playback Characteristic at 7½ IPS. 1963.

EIA-295 Disc Recording Characteristic. 1982.

EIA-297-A Cable Connectors for Audio Facilities for Radio Broadcasting. 1970.

EIA-355 Standard dimension for Unrecorded Magnetic Sound Recording Tape. 1974.

EIA-386 Recommended Measurement Method for Phonographic Rumble. 1978.

EIA-455 now has about 100 sections dealing with broadcasting.

EIA-462 Electrical Performance Standards for Television Broadcast Demodulators. 1979.

EIA-508 Electrical Performance Standards for Television Broadcast Transmitters. 1987.

EIA-549 NRSC AM Pre-emphasis/De-emphasis and Broadcast Audio Transmission Bandwidth Specifications. 1990.

EIA-560 Standard Method of Measurement for Compact Disc Players. 1989.

TR-101-A Electrical Performance Standards for Standard Broadcast Transmitters. 1948.

TR-107 Electrical Performance Standards for FM Broadcast Transmitters. 1949.

TR-117 Antennas and Combination of Antennas for FM Broadcasting Stations. 1949.

National Association of Broadcasters (NAB)

1771 N Street, N.W.
Washington, D.C. 20036–2891
(202) 429-5300

The NAB represents the broadcasting industry before Congress, the courts, regulatory agencies, at the White House, and before the general public. They serve a membership of over 5,100 radio station and approximately 1,000 television stations. The NAB department of Science and Technology represents the industry before the FCC and other agencies on issues affecting spectrum management and technical regulations. The department fosters and promotes research and development of new broadcast technologies, in part, by using the resources of NAB's engineering laboratory.

The department provides timely, useful, and accurate technical information for NAB members in Radio & TV TechCheck (a weekly technical news fax), and through other NAB publications. Other engineering publications are: Broadcast Engineering Conference Proceedings and HDTV World Proceedings (published annually at the conferences); Broadcast and Audio System Test CD I and II; NAB Test Record; "Engineering Technical Standards"; "Guide to Advanced Television Systems II"; "A Broadcaster's Guide to FCC RF Radiation Regulation Compliance"; and "Radio and Television Towers: Maintaining, Modifying, and Leasing."

NAB publications may be ordered by calling (800) 368-5644.

National Cable Television Association (NCTA)

1724 Massachusetts Avenue N.W.
Washington, D.C. 20036
(202) 775-3550

The National Cable Television Association (NCTA) is the major trade association for the cable industry. They represent the cable television industry before Congress and Federal Agencies, in the courts, and before state regulatory agencies. Members comprise over 400 cable operating companies and over 600 equipment manufacturers, programmers, and service companies.

Telecommunications Industry Association (TIA)

1722 Eye Street, N.W.
Suite 440
Washington, DC 20006
(202) 457-4936

The Telecommunications Industry Association is the result of a merger between the Electronic Industries Association (EIA) and the United States Telecommunications Suppliers Association (USTSA). The TIA represents U.S. suppliers of telecommunications equipment before Congress, the courts, regulatory agencies, at the White House, and before the general public. TIA is a provider of product-oriented standards. The standards programs presently operate under the accreditation awarded to the EIA Engineering Department by the American National Standards Institute (ANSI).

PROFESSIONAL ASSOCIATIONS

Broadcast Technology Society

Institute of Electrical and Electronic Engineers (IEEE) 345 East 47th Street
New York, New York 10017
(212) 705-7900

Members of the IEEE include engineers and scientists in electrical engineering, electronics, and allied fields as well as over 30,000 students. The IEEE holds numerous meetings and special technical conferences, conducts lecture courses at the local level on topics of current engineering and scientific interest, assists student groups, and awards medals, prizes, and scholarships for outstanding technical achievement. Along with other societies, they support the Engineering Societies' Library in New York City.

Publications of the IEEE include the Proceedings (monthly), Spectrum (monthly), and Directory (annually). The Societies and councils publish journals, magazines, and conference proceedings.

Society of Broadcast Engineers (SBE)

8445 Keystone Crossing, Ste. 140
Indianapolis, Indiana 46240
(317) 253-1640

The SBE is the largest national organization for broadcast engineers. It is devoted to the professional development of its members. The SBE promotes communication between SBE members, and provides national representation for those members before federal and state regulatory agencies, the courts, manufacturers, and the general public. The SBE through its 100 plus local chapters holds periodic meetings and technical conferences.

There are five levels of membership in the society.

- Honorary Member
- Fellow
- Senior Member
- Member
- Associate and Student

The SBE administers a certification program, recognizing four levels of engineering achievement.

- Broadcast Technologist
- Broadcast Engineer
- Senior Broadcast Engineer
- SBE Professional Broadcast Engineer

Frequency Coordination

The SBE National Frequency Coordinating Committee (NFCC) was established in 1982 when the FCC asked the broadcast industry to identify local contacts for Part 74 frequency coordination. The SBE formed the NFCC to develop local data bases of frequencies and users to assist the FCC and Part 74 users. Voluntary frequency coordination is handled on a local basis by SBE coordinators.

For additional information on SBE frequency coordination, see Chapter 1.3, "Frequency Coordination."

Ennes Foundation

The Ennes Educational Foundation was incorporated in 1986 as the educational arm of the SBE.

Each year the Foundation awards scholarships and grants for entry level training as well as continuing education in broadcast technology.

Society of Motion Picture and Television Engineers (SMPTE)

595 West Hartsdale Avenue
White Plains, New York 10607
(914) 761-1100

The membership of SMPTE is comprised of professional engineers and technicians in motion pictures, television, and allied arts and sciences. The Society advances engineering technology, disseminates scientific information, and sponsors lectures, exhibitions, and conferences to advance the theory and practice of motion picture and television engineering. As an accredited standards developer under the American National Standards Institute, they develop national standards for motion pictures, television, and sound associated with motion picture and television images. The Society also develops Recommended Practices and Engineering Guidelines. For copies of SMPTE engineering documents, contact their Standards Department.

SMPTE also makes available picture and sound test films for use as standardized measuring tools, serves as administrator of the Secretariat of ISO Technical Committee 36 on Cinematography, and of the U.S. Technical Advisory Groups for ISO TC36 and IEC SC60B on Video Recording. The Society sponsors technical courses at universities on such subjects as sound techniques, laboratory processing, special effects, and lighting for technicians and students. The Society presents nine annual awards for outstanding contributions to motion picture and television engineering.

SMPTE engineering committees include: Audio Recording and Reproduction Technology; Film Technology; Laboratory Services Technology; Projection Technology; Video Recording and Reproduction Technology; Television Technology; and New Television Technology. Publications of the Society include the monthly Journal, a yearly book of papers presented at the Society's annual Television Conference, and other books on motion picture and television technology.

RELATED ORGANIZATIONS

American National Standards Institute (ANSI)

655 15th Street N.W.
Suite 300
Washington, D.C. 20005
(202) 639-4090

Members include industrial firms, trade associations, technical societies, labor organizations, consumer organizations, and government agencies. ANSI serves as the clearinghouse for nationally-coordinated voluntary safety, engineering, and industrial standards. The institute gives status as American National Standards to standards developed by agreement from all groups concerned, in such areas as: definitions, terminology, symbols and abbreviations; materials; performance characteristics; procedure and methods of rating; methods of testing and analysis; size, weight, volume, and rating; practice; and safety, health, and building construction. They also provide information on foreign standards, and represent United States' interests in international standardization work. The ANSI committee is Certification. The Councils are: Company Member, Consumer, Executive Standards, International Standards, and Organizational Member. Publications of the institute are the Reporter (biweekly), Standards Action (biweekly), and Catalog of Standards (annually).

American Radio Relay League (ARRL)

225 Main Street
Newington, CT 06111
(203) 666-1541

Members are licensed amateur radio operators in the U.S. and Canada, and others interested in amateur radio, communication, and experimentation. The league operates a nationwide message handling network: the National Traffic System. Its members serve as official relay stations, observers, phone stations, emergency coordinators, experimental stations, and bulletin stations. They operate an experimental equipment laboratory and maintain the Museum of Amateur Radio. They also sponsor contests and present awards for operating proficiency. ARRL serves as secretariat for International Amateur Radio Union. Publications include QST (monthly) and Radio Amateur's Handbook (annually), as well as special booklets for beginners, and others on antennas, mobile, and radio fundamentals.

Association for Broadcast Engineering Standards (ABES)

2000 M Street, N.W.
Suite 600
Washington, D.C. 20036
(202) 331-0606

An independent, nongovernmental, voluntary organization of licensees and permitees of U.S. radio broadcast stations "united to assist the appropriate government authorities and the industry in assuring optimum radio service for the people of the United States and to follow the intent of Congress in the Communications Act of 1934, as amended." Objectives include: "rejecting and opposing proposals which would decrease, impair, or destroy the radio service now available to the American people; and supporting the adoption of legislature and Federal Communications Commission policies that encourage sound technical standards upon which optimum radio service may be obtained." The association analyzes proposed broadcast legislation, FCC proposals for rule making, and industry proposals filed with the FCC that will, directly or indirectly, affect the public interest in optimum radio service. The association formulates and conducts short and long range programs of technical research to assist governmental and industrial groups in maintaining sound technical standards.

Audio Engineering Society (AES)

60 East 42nd Street
Room 449
New York, NY 10017
(212) 661-8528

Members include engineers, administrators, and technicians who design or operate recording and reproducing equipment for radio, television, motion picture, and recording studios, or who produce, install, and operate disc, magnetic tape, and sound amplifying equipment; educators who use recording in teaching, or who teach acoustics, electronics, and other sciences basic to the recording and reproducing of sound; and administrators, sales engineers, and technicians in the sound industry and related fields.

Broadcast Education Association (BEA)

1771 N Street, N.W.
Washington, D.C. 20036
(202) 429-5355

Primarily directed by those in the academic community, its orientation is toward exploring new trends, ideas, and opportunities in broadcasting. The BEA publishes the Journal of Broadcasting and Electronic Media and Feedback. It also administers the Patterson Radio Scholarship and the Harold E. Fellows Memorial Scholarship.

Illumination Engineering Society of North America (IES)

345 East 47th Street
New York, NY 10017
(202) 705-7926

This is a technical society whose members include engineers, architects, designers, educators, students, contractors, distributors, utility personnel, scientists, and manufacturers dealing with the art or science of illumination. They provide assistance with technical problems, reference help, and speakers. They maintain liaison with schools and colleges and offer basic and advanced IES Lighting Courses through local sections and in cooperation with other organizations.

Underwriters Laboratories (UL)

333 Pfingsten Road
Northbrook, Illinois 60062
(312) 272-8800

A testing laboratory that maintains additional laboratories in Melville, New York; Santa Clara, California; and Tampa, Florida. UL seeks "by scientific investigation, study, experiments, and tests to determine the relation of various materials, devices, products, equipment, constructions, methods and systems to hazards appurtenant thereto or to the use thereof affecting life and property, and to ascertain, define and publish standards, classifications, and specifications for materials, devices, products, equipment, construction, methods, and systems affecting such hazards, and other information tending to reduce loss of life and property from such hazards."

BROADCAST ENGINEERING AND RELATED PERIODICALS

ABU Review

(Asia-Pacific Broadcasting Union)
P.O. Box 1164
Pejabat Pos Jalan Pantai Baru
59700 Kuala Lumpur, Malaysia
011 603 282 3108

Audio

Diamandis Communications, Inc.
1633 Broadway, 41st Floor
New York, NY 10019
(212) 767-6331

Broadcast Engineering

Intertec Publishing
P.O. Box 12901
Overland Park, Kansas 66282-2901
(913) 888-4664

Broadcasting

1705 DeSales Street, NW
Washington, DC 20036–4480
(202) 659-2340

db Magazine

Sagamore Publishing Co. Inc.
203 Commack Road Suite 1010
Commack, NY 11725
(516) 586-6530

EBU Review—Technical

(European Broadcasting Union)
Case postale 67
CH-1218 Grand-Sacconex
Geneva, Switzerland
011 41 22 717 21 11

Chilton's ECN Electronic Component NEWS

Box 2011
Radnor, PA 19080–9511
(215) 964-4000

EE Evaluation Engineering

2504 North Tamiami Trail
Nokomis, Florida 34275–9987
(813) 966-9521

Electronic Buyers' News

CMP Publications, Inc.
600 Community Drive
Manhasset, NY 11030–3875
(516) 365-4600

Electronic Engineering Times

CMP Publications, Inc.
600 Community Drive
Manhasset, NY 11030–3875
(516) 365-4600

Electronic Media

Crain Communications Inc.
740 North Rush Street
Chicago, Il 60611
(312) 649-5293

Electronics Design

Penton Publishing Co.
600 Sumner St.
Stamford, CT 06904
(212) 696-7000

Home and Studio Recording

Music Maker Publications, Inc.
22024 Lassen Street
Suite 118
Chatsworth, CA 91311
(818) 407-0744

Journal of the Audio Engineering Society

60 East 42nd Street
New York, New York 10165
(212) 661-2355

MIX

Act III Publishing
6400 Hollis Street, #12
Emeryville, CA 94608
(510) 653-3307

Popular Communications

76 North Broadway
Hicksville, NY 11801
(516) 681-2922

PRO Sound News

PSN Publications, Inc.
2 Park Avenue, 4th Floor
New York, NY 10016
(212) 213-3444

QST

225 Main Street
Newington, CT 06111
(203) 666-1541

Radio Electronics

RADIO GUIDE
500-B Bi-Country Blvd.
Farmingdale, NJ 11735
(516) 293-3000

Radio World

P.O. Box 1214
Falls Church, Virginia 22041
(703) 998-7600

Recording Engineer/Producer

Intertec Publishing
8885 Rio San Diego Drive
San Diego, CA 92108
(619) 299-6655

SMPTE Journal

595 West Hartsdale Avenue
White Plains, New York 10607
(914) 761-1100

Studio Sound and Broadcast Engineering

Spotlight Publications, Ltd.
Ludgate House
245 Blackfrars Road
London, SE1 9UR, UK
011 44 71 620 3636

Television Technology

P.O. Box 1214
Falls Church, VA 22041
(703) 998-7600

Video Systems

P.O. Box 12901
Overland Park, Kansas 66282-2901
(913) 888-6664

BROADCAST ACRONYMS

AA Average Audience
AAAA American Association of Advertising Agencies
AAAA Associated Actors and Artists of America
AAF American Advertising Federation
AAPOR American Association for Public Opinion Research
ABES Association for Broadcast Engineering Standards
AC Adult Contemporar
ACA Advertising Council of America
ACBB American Council for Better Broadcasting
ACC Automatic Chrominance (or Contrast) Control
ACE American Cinema Editors
ACSB Amplitude Compandored Sideband
ACT Action for Children's Television
ACTS All-Channel Television Society
ACTS Association of Cable Television Suppliers
ACTVA American Community TV Association
ACU Antenna Coupling Unit
ADC Analog-to-Digital Converter
ADI Area of Dominant Influence
ADPCM Adaptive Differential Pulse Code Modulation
ADS Alpha Delta Sigma (Advertising Fraternity)
AEA Actors Equity Association
AEA American Electronics Association
AER Alpha Epsilon Rho—College Broadcast Fraternity
AES Audio Engineering Society
AF Audio Frequency
AFC Automatic Frequency Control
AFCCE Association of Federal Communications Consulting Engineers
AFI American Film Institute
AFM American Federation of Musicians
AFN Armed Forces Network
AFRTS Armed Forces Radio & Television Service
AFT Automatic Fine Tuning
AFTRA American Federation of Television & Radio Artists
AFTRCC Aerospace Flight Test Radio Coordinating Council
AGAC American Guild of Actors & Composers
AGC Automatic Gain Control
AGMA American Guild of Musical Artists
AGVA American Guild of Variety Artists
AHAAT Antenna Height Above Average Terrain
AIAA Association of International Advertising Agencies
AID Arbitron Information on Demand
AIM Accuracy in Media
AIRS American Inter-Tribal Radio Society
AITS Association of Independent Television Stations
ALC Automatic Level Control
ALF American Legal Foundation
ALJ Administrative Law Judge
ALPTVA American Low Power TV Association
AM Amplitude Modulation
AMOL Automated Measurement of Lineup
AMPAS Academy of Motion Picture Arts & Sciences
AMVB Association of Music Video Broadcasters
ANA Association of National Advertisers
ANPA American Newspaper Publishers Association
ANSI American National Standards Institute
AOR Album Oriented Rock (program format)
AOR All Over the Road (undefined program format)
APA Administrative Procedures Act
APB Associated Press Broadcasters
APL Average Picture Level
APR American Public Radio
ARCH Automatic Remote Cassette Handler
ARF Advertisers Research Foundation
ARRS Association of Radio Reading Services
ASC American Society of Cinematographers
ASCAP American Society of Composers
ASTA Advertisers Syndicated Television Association
ATFP Alliance of Television & Film Producers
ATM Asynchronous Transfer Mode
ATR Audio Tape Recorder
ATRC Advanced Television Research Consortium
ATS Applications Technology Satellite
ATS Automatic Transmission System
ATSC Advanced Television Systems Committee
ATTC Advanced Television Test Center
ATU Antenna Tuning Unit
AWG American Wire Gauge (for electrical wiring)
AWRT American Women in Radio & Television
BAPSA Broadcast Advertising Producers Society of America
BAR Broadcast Advertisers Reports
BBC British Broadcasting Corporation
BBI Broadcasting Foundation, Inc.
BBM Broadcast Bureau of Management
BCA Broadcast Credit Association

BEA Broadcast Education Association
BEDA Broadcast Executive Directors Association
BER Bit Error Rate
BET Black Entertainment Television
BFM Broadcast Financial Management Association
BFO Beat Frequency Oscillator (in receivers)
BIAC Broadcast Interassociation Council
BIB Board for International Broadcasting
BIB Broadcast Information Bureau
BICIAP Broadcasting Industry Council to Improve America Productivity
BM Beautiful Music (easy listening format)
BMI Broadcast Music Incorporated
BNC Bayonet Connector (for small coaxial cable)
BOC Bell Operating Company
BPA Broadcast Promotion Association
BPME Broadcast Promotion and Marketing Executives
BRC Broadcast Rating Council (see EMRC)
BSC British Society of Cinematographers
BSS Broadcast Satellite Service
BTA Best Time Available
BTRE Broadcast Television Recording Engineers
BTSC Broadcast Television Systems Committee
C/CS County Coverage Service
CAB Canadian Association of Broadcasters
CABSC Canadian Advanced Broadcast Systems Committee
CAMS Cable Audience Measurement Study
CAPAC Composers Authors & Publishers Assn. of Canada.
CARS Community Antenna Relay Service
CATA Community Antenna Television Association
CATS Center for Advanced Television Studies
CAVS Component Analog Video Signal
CBA Community Broadcasters Association
CBA Commonwealth Broadcasting Association (United Kingdom)
CBC Canadian Broadcasting Corporation
CBO Congress of Broadcast Organizations
CBO Control Board Operator
CC Closed-Captioned
CCB FCC Common Carrier Bureau
CCC Citizens Communication Center
CCD Charge Coupled Device Camera
CCIR International Consultative Committee for Radio
CCR Central Control Room
CCSB Clear Channel Broadcasting Service
CCTA Canadian Cable Television Association
CCTV Closed Circuit Television
CD Compact Disc (audio recordings)
CD-ROM Compact Disc-Read Only Memory (in computer systems)
CE Chief Engineer (Chief Station Operator)
CEBA Communications Excellence to Black Americans
CEG/EIA Consumer Electronics Group of Electronic Industries Association
CFR Can't Find the Road (undefined program format)
CFR Code of Federal Regulations
CH Critical Hours
CHR Contemporary Hit Radio (format)
CHUT Cable Households Using Television
CIRT Mexican Association of Broadcasters
CLI Cumulative Leakage Index
CMOS Complementary Metal Oxide Semiconductor
COFMD Coded Orthogonal Frequency Division Multiplex
CONTAM Committee on Nationwide Television Audience Measurement
CP Circular Polarization
CP Construction Permit
CPB Corporation for Public Broadcasting
CPCS Common Program Control Station (for EBS)
CPD Committee for Prudent Deregulation
CRT Cathode Ray Tube (Video Monitor)
CRT Copyright Royalty Tribunal
CRTC Canadian Radio Television & Telecommunication Commission
CS Close Shot (in video or film production)
CSSB Compatible Single Sideband
CTAM Cable Television Administration & Marketing Society
CTCM Chrominance Time-Compression Luminance Multiplex
CTS Communications Technology Satellite
CTY Country Music (program format)
CU Close Up (view in video or film production)
CUB Council for UHF Broadcasters
CUME Cumulative Audience
CUT Coordinated Universal Time (see UTC)
CV Composite Video
CVCC Canadian Videotext Consultative Committee
CVSD Continuously Variable Slope Delta Modulation
CWA Communications Workers of America
DA Directional Antenna
DA Distribution Amplifier
DAB Digital Audio Broadcasting
DAC Digital-to-Audio Converter
DAF Demographic Adjustment Factor
DAMA Demand Assigned Multiple Access
DASH Digital Audio Stationary Head Tape Recorder
DAT Digital Audio Tape
DATS Digital Audio Transmission Service
DATV Digitally Assisted Television
DB Delayed Broadcast
DBA Daytime Broadcasters Association (disbanded)
DBS Direct Broadcast Satellite
DBSA Direct Broadcast Satellite Association
DE Director of Engineering
DGA Directors Guild of America
DIP Dual Inline Package (integrated circuit assembly)
DJ Disk Jockey (announcer for programs of recorded music)
DMA Designated Market Area

DMM Digital Multimeter (used for servicing electronic equipment)
DMM Direct Metal Master (for LP record manufacturing)
DNA Data Not Available
DOC Department of Commerce (USA)
DOC Department of Communications (Canadian)
DPCM Differential Pulse Code Modulation
DPSK Differential Phase Shift Keying
DRAW Direct Read After Write (Optical Data Disk)
DSK Downstream Keyer
DSP Digital (audio) Signal Processing
DTMF Dual Tone Multiple Frequency (telephone tone calling system)
DTTR Digital Television Tape Recorder
DVCR Digital Video Cassette Recorder
DVE Digital Video Effects
DW Deutsche Welle (German International Broadcast Service)
DX Long Distance Communications
EAN Emergency Action Notification
EBR Electron Beam Recording
EBS Emergency Broadcast System
EBU European Broadcasting Union
ECU Extreme Close-Up (shot in video or film production)
EDTV Extended Definition Televisio
EEPA Electromagnetic Energy Policy Alliance
EFP Electronic Field Production
EIA Electronic Industries Association
EIAJ Electronic Industries Association of Japan
EIRP Equivalent Isotropic Radiated Power
EJ Electronic Journalism (now ENG)
EMP Electromagnetic Pulse
EMRC Electronic Media Rating Council
ENG Electronic News Gathering
ENP Electronic News Processing
EPA Environmental Protection Agency
EPG Electronic Program Guide
EPM Equipment Performance Measurements
EPROM Erasable Programmable Read Only Memory
ERP Effective Radiated Power
ET Electrical Transcription
ETV Educational Television
FAA Federal Aviation Administration
FCBA Federal Communications Bar Association
FCC Federal Communications Commission
FDM Frequency Division Multiplex
FEC Federal Election Commission
FEC Forward Error Correction (in data transmission system)
FEMA Federal Emergency Management Agency
FET Field Effect Transistor
FIM Fairness in Media (Organization)
FIT Foundation to Improve Television
FM Frequency Modulation
FMX FM Extended Range Transmission System
FNG Film News Gathering
FOB FCC Field Operations Bureau
FOIA Freedom of Information Act
FSK Frequency Shift Keying
FSS Fixed Satellite Service
FTC Federal Trade Commission
FWH Fast Weekly Household Audiences Report
GAO General Accounting Office (federal government)
GMT Greenwich Mean Time (no obsolete—see UTC)
GOE Group of Experts (International Radio Conferences)
GPO Government Printing Office
GRP Gross Rating Points
GSO Geostationary Satellite Orbit
HAAT Height Above Average Terrain
HASL Height above Mean Sea Level
HBI Horizontal Blanking Interval
HDTV High Definition Television
HF High Frequency (shortwave broadcasting)
HFR Hold for Release
HID High Intensity Discharger (for lighting lamps)
HPA High Power Amplifier (used in SNG terminals)
HSD Home Satellite Dish
HTR Household Tracking Report
HUT Household Using Television
IAAB Inter-American Association of Broadcasters
IAB International Association of Broadcasting
IATSE International Alliance of Theatrical Stage Employee
IBA Independent Broadcasting Authority
IBC International Broadcast Convention
IBEW International Brotherhood of Electrical Workers
IBS Intercollegiate Broadcasting System
ICPM Incidental Carrier Phase Modulation
ICR Inter-City Relay Station
ICTV Independent Community Television Alliance
ID Station Identification
IDTV Improved Definition Television
IEC International Electrotechnical Commission
IEEE Institute of Electrical and Electronics Engineers
IFB Interruptable Fold-Back (for SNG circuits)
IFRB International Frequency Registration Board
IHF Institute of High Fidelity (merged into EIA)
IM Intermodulation
IMD Intermodulation Distortion
IMPA Independent Music Producers Association
INTELSAT International Telecommunications Satellite Organization
INTV Association of Independent Television Stations
IPA Intermediate Power Amplifier
IPS Inches per Second (linear recording tape speed)
IRAC Interdepartmental Radio Advisory Committee
IRD Integrated Receiver Decoder (for satellite TV reception)
IRE Institute of Radio Engineers (now IEEE)

IRS Insertion Reference Signals (for auto video processing)
IRTS International Radio and Television Society
IRTV International Radio & Television Foundation
ISDN Integrated Services Digital Network
ISI InterSymbol Interference (in digital transmission systems)
ISL Inter-Satellite Link
ISO International Standards Organization
ITFS Instructional Television Fixed Service
ITNA Independent Television News Association
ITS Inserted Test Signals (in the video signal)
ITS Institute of Telecommunication Sciences (div. of NTIA)
ITS International Transcription Service (FCC Copy Service)
ITU International Telecommunication Union
IX Interference (to radio communications circuits)
JCIC Joint Council on Intersocietal Coordination
JFET Junction Field Effect Transistor
LAN Local Area Network (for computer systems)
LAPLPS North American Presentation Level Protocol Standard (Vidtex)
LCU Line Coupling Unit (in transmission systems)
LED Light Emitting Diode
LFCC Local Frequency Coordinating Committee
LIDIA Local ID Insertion Automatically
LMCC Land Mobile Communications Council
LMRS Land Mobile Radio Service
LNA Low Noise Amplifier
LO Local Origination (channel)
LOP Least Objectionable Program
LP Long Play Phonograph Recording (33⅓ rpm microgroove
LPAS Low Power Auxiliary Service (for broadcast stations)
LPTV Low Power Television
LS Local Sunset (as used in FCC AM station authorizations
LS Long Shot (view in video or film production)
LSI Large Scale Integrated Circuit
LTU Line Tuning Unit (in transmission systems)
M/X Mix Effects
MAC Multiplex Analog Component TV Transmission
MADI Multichannel Audio Digital Interface
MAP Media Access Project
MCR Master Control Room
MCS Multichannel Sound (for TV)
MCTV Multichannel TV
MCU Medium Close-Up (shot in video or film production)
MDC Multistage Depressed Collector (TV transmitter tube type)
MDS Multipoint Distribution Service
MEOV Maximum Expected Operating Value
MERPS Multiple Event Record & Playback System
MIDI Musical Instrument Digital Interface
MIRF Master International Frequency Register (of ITU)
MLB Major League Baseball
MLS Medium Long Shot (in video or film production)
MM Metered Market Service
MMB FCC Mass Media Bureau
MMDS Multichannel-Multipoint Distribution Service
MNA Multi-Network Area
MO&O Memorandum Opinion and Order
MOL Maximum Operating Level
MOP Minute of Program
MOR Middle of the Road (radio programming format)
MOSFET Metal Oxide Semiconductor Field Effect Transistor
MOT Main (Satellite) Operating Terminal
MP Modified Permit
MP Monitoring Point (for AM directional antenna systems)
MPA Multiple Product Advertisement
MPAA Motion Picture Association of America
MSA Metro Survey Area
MSI Market Statistics Inc.
MSO Multiple System Operator (CATV)
MSS Mobile Satellite Service
MST (Now MSTV, see below)
MSTV Association for Maximum Service Televisio
MTBF Mean Time Between Failure(s)
MTBR Mean Time Between Repair(s)
MTS Multichannel Television Sound
MTV Music Television
MUF Maximum Usable Frequency (in HF broadcasting)
MUSE Multiple Sub-Nyquist Sampling Encoding (for TV)
MUX Multiplex (communications transmissions)
MX Mutually Exclusive Condition (conflicting FCC applications)
N/T News-Talk (program format)
NAB National Association of Broadcasters
NABB National Association of Better Broadcasting
NABER National Association of Business & Educational Radi
NABET National Association of Broadcast Employees and Technicians
NABOB National Association of Black-Owned Broadcasters
NABTS North American Broadcast Teletext Standard
NAC National Audience Composition
NACT National Association of Community TV Broadcasters
NAD National Audience Demographics
NAEB National Association of Educational Broadcasters (disbanded)
NAFB National Association of Farm Broadcasters
NAITPD National Association of Independent TV Producers & Distributors
NAL Notice of Apparent Liability (issued by FCC)
NAMBA North American National Broadcasters Association
NAPTE National Association of Television Program Executives

NAPTS National Association of Public Television Stations
NARAS National Academy of Recording Arts and Sciences
NARBA North American Radio Broadcasting Agreement
NARTE National Association of Radio & Television Engineers
NARTSH National Association of Radio Talk Show Hosts
NARUC National Association of Regulatory Utility Commissioners
NASA National Aeronautics and Space Administration
NASB National Association of Spanish Broadcasters
NATAS National Academy of Television Arts & Sciences
NATO National Association of Theater Owners
NBEA National Broadcast Editorial Association
NBI Nielsen Broadcast Index
NBMC National Black Media Coalition
NBN National Black Network
NCCB National Citizens Committee for Broadcasting
NCFM Noncommercial FM
NCIPBP National Coalition of Independent Public Broadcast Producers
NCPAC National Conservative Political Action Committee
NCTA National Cable Television Association
NCTV National Coalition on Television Violence
NEC National Electrical Code
NEH National Endowment for the Humanities
NEMA National Electrical Manufacturers Association
NEMO Not Emanating from Main Office (a remote broadcast)
NET National Educational Televisio
NFCB National Federation of Community Broadcasters
NHI Nielsen Home Video Index
NHK Japanese National Broadcast System
NIAC National Industry Advisory Committee (EBS)
NICAD Nickel Cadmium (alloy used in rechargeable batteries)
NIER Nonionizing Electromagnetic Radiation
NILPTV National Institute for Lower Power TV
NNA National Newspaper Association
NOI Notice of Inquiry
NPACT National Public Affairs Center for Television
NPBS Nielsen Post-Buy Service
NPR National Public Radio
NPRM Notice of Proposed Rule Making
NQRC National Quadraphonic Radio Committee
NRB National Religious Broadcasters
NRBA National Radio Broadcasters Association (merged with NAB)
NSA National Security Agency
NSCA National Satellite Cable Association
NSEP National Security & Emergency Procedures
NSI Nielsen Station Index
NSRC National Stereophonic Radio Committee
NTA National Translator Association
NTI Nielsen Television Index
NTIA National Telecommunications and Information Administration
NTIS National Technical Information Service (Dept. of Commerce)
NTSC National Television System Committee
O&O Owned and Operated
OET FCC Office of Engineering & Technology (formerly OST)
OFS Operational Fixed Service
OIC FCC Office of International Communications
OIC Department of State Office of International Communications
OIRT International Radio and Television Organization
OMB Office of Management and Budget (U.S. Government)
OMD FCC Office of Managing Director
OPA FCC Office of Public Affairs
OPP FCC Office of Plans and Policy
OQPSK Offset Quadrature Phase Shift Keying
OROM Optical Read Only Memory (for data storage)
ORTF French National Broadcasting Organization
OS Off Screen
OSHA Occupational Safety and Health Administration
OST FCC Office of Science & Technology (renamed OET)
OSTP Office of Science & Telecommunications Policy
OTO One Time Only
OTP Office of Telecommunications Polity (now NTIA)
PA Power Amplifier
PA Public Address System
PA Public Affairs
PAC Political Action Committee
PACCT Political Action Committee for Cable Television
PACT Private Access Communications Terminals (Satellite Service)
PAL Phase Alternate Line (color TV system used in Europe and elsewhere)
PBS Public Broadcasting Service
PBX Private Branch (Telephone) Exchange
PC Personal Computer
PCB Polychlorinated Biphenyl (fluid used in electrical equipment; e.g., transformers and capacitors)
PCB Printed Circuit Board
PCM Pulse Code Modulation
PD Program Director—Production Director
PD Public Domain Material
PDC Program Delivery Control (by code within video signal)

PDM Pulse Duration Modulation
PDT Published Data Tapes
PE Professional Engineer (registered or licensed)
PGA Producers Guild of America
PI Per Inquiry
PLL Phase Lock(ed) Loop
PLP Presentation Level Protocol
PLT Private Line Telecommunications Circuit
PM Permanent Magnet
PMRC Parents Music Resource Center
POE Panel of Experts
POM Professional Owners & Managers
POT Plain Old Telephone
POV Point of View (shot in video or film production)
PPI Peak Program (level) Indicator
PPM Peak Program Meter
PPV Pay Per View
PRB FCC Private Radio Bureau
PRS Program Rating Summary Report
PSA Presunrise Service Authorization
PSA Public Service Announcement
PSK Phase Shift Keying
PSM Pulse Step Modulation
PSRA Presunrise Authority
PSSA Post Sunset Authority
PSSC Public Service Satellite Consortium
PSTN Public Switched Telephone Network
PTAR Prime Time Access Rule
PTT Postal Telephone & Telegraph Administration
PTT Push to Talk (microphone)
PTV Pay Television
PTV Public Television
PVC Polyvinyl Chloride (common plastic)
PWM Pulse Width Modulation
QM Quadrature Modulation
QPSK Quadrature Phase Shift Keying
R&F Reach & Frequency
R&O FCC Report & Order (in rule making proceeding)
RAB Radio Advertising Bureau
RADAR Radio All Dimension Audio Research
RADAR Radio Detection and Ranging
RADET Rapid Deployment Earth Terminal (for SNG)
RAM Random Access Memory (in computer systems)
RARC Regional Administrative Radio Conference
RASO Radio Allocations Study Organization
RBOC Regional Bell Operating Company
RCA Radio Club of America
RCC Radio Common Carrier
RCI Radio Canada International
RCL Remote Control Location
RDD Random Digital Dialing (for audience surveys)
RDS Radio Data Service (on FM 57 kHz subcarrier)
RFA Regulatory Flexibility Act
RFE Radio Free Europe
RFI Radio Frequency Interference
RFR Radio Frequency Radiation
RGB Red-Green-Blue (color monitor signal)
RIAA Recording Industry Association of America
RIAS Radio in American Sector (Berlin)
RL Radio Liberty
RMS Root Mean Square
RN Radio Nederland (Dutch International Broadcast Service)
RNA Radio Network Association
RO Receive-Only (earth terminal)
ROM Read Only Memory (in computer systems)
ROS Run of Schedule
ROSP Report on Syndicated Programs
RP Rear Projection
RP Restricted Radiotelephone Operator Permit
RPU Remote Pickup Unit
RSS Root Sum Square
RTCA Radio-Television Correspondents Association
RTES Radio & Television Executives Society
RTNDA Radio & Television News Directors Association
RTRC Radio & Television Research Council
RTTY Radio Teletypewriter Terminal
RX Radio Receiver
S/I Signal to Interference Ratio
S/N Signal to Noise Ratio
SAC Sales Advisory Committee of TBA
SAG Screen Actors Guild
SAP Second Audio Program (on TV subcarrier)
SAR Specific Absorption Rate of Radio Frequency Radiation
SAW Surface Acoustic Wave (type of electronic filter)
SAWA Screen Advertising World Association
SBA Small Business Administration
SBC Standard Broadcast Calendar
SBCA Satellite Broadcast and Communications Association
SBE Society of Broadcast Engineers
SBS Sound Broadcast Satellite
SCA Subsidiary Communications Authority
SCG Screen Cartoonists Guild
SCMS Serial Copy Management System (CD & DAT anti-copy code)
SCPC Single Channel Per Carrier
SCTE Society of Cable Television Engineers
SDX Sigma Delta Chi (Journalism Society)
SE Sound Effects (also see SFX)
SECAM Sequential Coleur Avec Memoire (French & Russian Color TV System)
SECC State Emergency Communications Committee (for EBS)
SEG Screen Extras Guild
SESAC Society of European Songwriters Artists and Composers
SFX Sound Effects
SH Specified Hours (Licensed hours of operation)
SI International System of Units (Metric system units)
SIA Storage Instantaneous Audiometer
SID Source Identification Signal (in TV VBI)

SIR Signal to Interference Ratio
SIS Sound in Sync (sound signal inserted in TV synchronization pulses)
SIU Sets in Use
SMA Special Market Area
SMATV Subscription Master Antenna Television
SMPTE Society of Motion Picture and Television Engineers
SMSA Standard Metropolitan Statistical Area
SNG Satellite News Gathering
SNR Signal to Noise Ratio
SNV Satellite News Vehicle
SOF Sound on Film
SOT Sound on Tape
SPACE Society for Private & Commercial Earth Stations
SPARS Society of Professional Audio Recording Studios
SPD Spectral Power Density
SPL Sound Pressure Level
SPTV Still Picture Television
SQAM Superposed Quadrature Amplitude Modulation
SRA Station Representative Association Inc.
STA Special Temporary Authority
STC Satellite Television Corp.
STL Studio Transmitter Link
STV Subscription Television
STVA Subscription Television Association
SVI Standard Volume Indicator (VU Meter)
SVRC Satellite Viewing Rights Coalition
SWARC Satellite World Radio Conference
T/R Transmit—Receive (earth terminal)
TARPAC Television And Radio Political Action Committee
TARPEC Television And Radio Political Education Committee
TASO Television Allocations Study Organization
TBA To Be Announced
TBC Time Base Corrector
TBD To Be Determined
TBR To Be Recorded
TD Technical Director
TDD Telecommunications Devices for the Deaf
TDMA Time Division Multiple Access (Satellite system)
TF Till Forbid (open ended advertising schedule)
TFN Till Further Notice (scheduled for broadcast)
THD Total Harmonic Distortion
TI Terrestrial Interference to space communications
TIA Telecommunications Industry Association
TIM Transient Intermodulation Distortion
TIO Television Information Office
TMC Time Multiplex Component Video
TMP Test Market Profile
TNC Threaded Connector (for small coaxial cable)
TOC Television Operator Council
TPO Transmitter Power Output
TRF Tuned Radio Frequency (type of receiver)
TRL Telemetry Return Link
TSA Total Survey Area
TSL Transmitter to Studio Link
TTY Teletypewriter Terminal
TVB Television Bureau of Advertising
TVI Television Interference
TVRO TV Receive-Only Satellite (earth terminal)
TVRS Television Viewer Response Service
TWTA Travelling Wave Tube Amplifier
TX Radio Transmitter
UC Urban Contemporary (program format)
UHF Ultra High Frequency
UL Underwriters Laboratory
USIA United States Information Agency
UTC Universal Coordinated Time (formerly GMT)
VAC Voltage—Alternating Current
VAR Video to Audio Carrier Ratio
VBI Vertical Blanking Interval
VBIC Vertical Blanking Interval Code
VCO Voltage Controlled Oscillator
VCR Video Cassette Recorder
VDP Video Disk Player
VDR Video Disk Recorder
VHF Very High Frequency
VHS Video Home System (a video cassette recording format)
VIMCAS Vertical Interval Multi-Channel Audio System
VIRS Vertical Interval Reference Signal
VITC Vertical Interval Time Code
VITS Vertical Interval Television Signal
VJB Video Juke Box
VOA Voice of America
VOD Video on Demand (system of cable program delivery)
VOM Volt-Ohm Meter (used in electronic service work)
VPR Video Production Recording
VPS Viewers Per Set
VPVH Viewer Per Viewing Household
VSAT Very Small [satellite] Antenna Terminal
VSWR Voltage Standing Wave Ratio
VTR Video Tape Recorder
VTVM Vacuum Tube Volt Meter (used in electronic service work)
VU Volume Unit (measurement of program level)
WAN Wide Area Network (for computer systems)
WARC World Administrative Radio Conference
WATCH Washington Association of Television & Children
WATS Wide Area Telephone Service
WDM Wavelength Division Multiplex
WGA Writers Guild of America
WHCA White House Communications Agency
WHCA White House Correspondents Association
WIC Women in Cable
WPFC World Press Freedom Committee
WST World System Teletext
WW Working Women
ZI Zoom In
ZO Zoom Out

2.1 Design, Erection, and Maintenance of Antenna Structures

John Windle, P.E.
Stainless, Inc., North Wales, Pennsylvania

Tower Lighting Section by:
Lewis Wetzel
EG&G Electro-Optics Division, Salem, Massachusetts

INTRODUCTION

The purpose of this section is to provide broadcast engineers and managers information concerning the design, erection, and maintenance of antenna structures. While fundamental principles of the design and behavior of these structures will be discussed, this section is not intended to enable readers to design and build their own tower, but provides instead a basic understanding of these unique structures to facilitate planning, modifying, and maintaining broadcast facilities.

TOWER CHARACTERISTICS

Types

All towers may be classified in one of the two basic groups: guyed or self-supporting. As their names imply, guyed towers depend on cables running to anchors located some distances from the tower base for their structural integrity, while self-supporting towers rely solely on their own construction as a cantilevered space truss.

With only a few exceptions, the cost of the actual tower structure and foundations is considerably less for a guyed tower than for one that is self-supporting. The advantage of the self-supporting tower is the relatively small land area required. Therefore, the choice between guyed or self-supporting depends to a large degree on the availability and cost of real estate.

A self-supporting tower requires a nearly square plot of land with equal sides that are 8% to 20% of the tower's height.

The amount of land required for a guyed tower depends on the distance between the tower base and the guy anchors. This distance is preferably between 70% and 80% of the height, which would require a rectangular plot having sides equal to 125% and 145% of the height.

Because of the great flexibility in guyed tower design, it is possible to reduce the anchor distance to as little as 35% of height, thereby requiring a much smaller land area. However, the cost of the tower increases as the anchor distance decreases. The approximate relationship of cost to anchor distance for a representative 1,200 feet television broadcast tower is shown in Fig. 1.

It is often possible to position a guyed tower on an irregularly shaped plot or to obtain long term lease

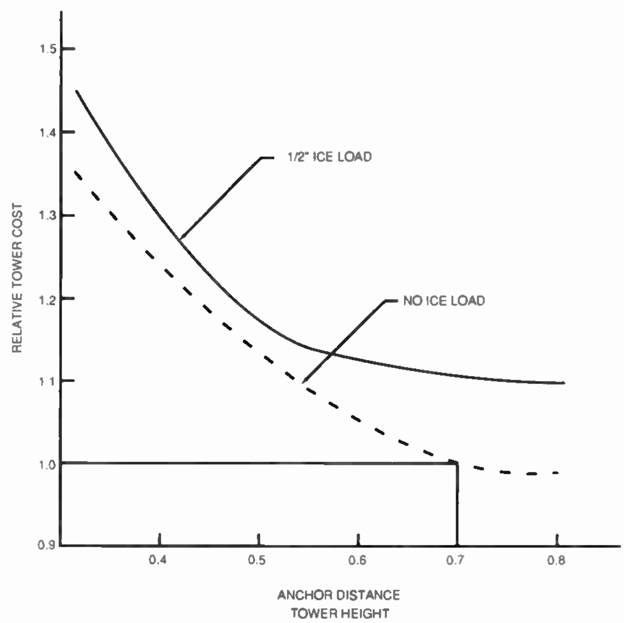


Figure 1. Effects of anchor distance on cost of a 1,200 foot guyed TV broadcast tower using ANSI/EIA/TIA Standard 222-E-1991.

agreements or easements for guy paths and anchor locations in order to minimize the tower cost without obtaining large, rectangular land areas.

Configurations

Self-Supporting Towers

Self-supporting towers may be either square or triangular in cross section. While it is usually more economical to use a triangular cross section, there are situations where a square cross section is a better choice. The principal structural elements are the legs, the web bracing in each face, and if required for stability, horizontal diaphragm bracing. The legs are usually sloped (tapered) to provide adequate strength and stability as the height increases. The degree of slope is an option of the designer to suit the equipment supported, the required rigidity, and the available land area. The slope is sometimes varied within a tower to maintain a desirable balance between the costs of leg members and bracing, or to reduce the foundation loads. Frequently the legs in the top section of the tower will be parallel to simplify the mounting of equipment.

There are several different configurations of bracing members for the individual truss panels. The choice is influenced by the width of the panel, the magnitude of the wind and ice loads imposed, the location of equipment, and required stability. Continuity in transferring the applied loads through the structure without significant eccentricity is essential regardless of the configuration used.

Guyed Towers

Guyed towers are almost always of triangular cross section although there are a few unique conditions for microwave and panel type FM and TV antenna supports where a square cross section is advantageous. The principal structural elements are the legs, the web bracing in each face, and the guy support systems. Except for sections at the tower base and locations where the width changes, the legs are parallel. The width of the tower is usually constant throughout the height of the tower with the exception of sections supporting antennas requiring a specific width of support structure. The base section is often tapered to a single point to provide a pivot support to eliminate large bending and torsional moments.

Theoretically, there are an infinite number of arrangements of guy cables to support a tower. The most common arrangement is three cables spaced at 120° with one attached to each leg, as shown in Fig. 2A. This is the minimum number of cables that can be used. When the tower supports equipment which imposes large twisting moments (torques), it is necessary to provide six cables at a level to maintain torsional stability. If the torque is localized, the guys at the location may be attached to triangular frames as shown in Fig. 2B. If the torques occur throughout the height, it may be desirable to double-guy the tower at every level as shown in Fig. 2C.

The number of levels of guy cables to support the tower is dependent on a number of factors including the height of the tower, width, location of equipment, and the environmental loading conditions. Because the tower is an axially compressed column, its strength is a function of its slenderness. While design codes permit slenderness ratios resulting in triangular towers having a span-to-width ratio as great as 49, it is usually economical to limit the ratio to a maximum of 30. While there is no upper limit to the number of guy levels imposed by any code, a practical limit for economical design is ten.

The position of equipment on the tower is an important factor in determining the location of guy levels. Preferably, guy attachments should not be located

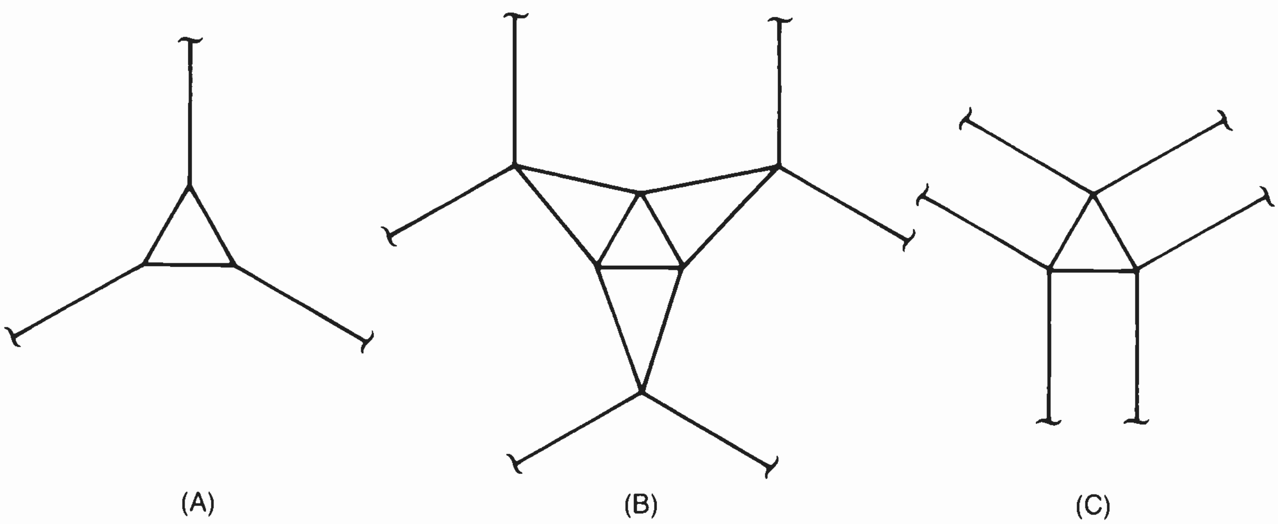


Figure 2. Typical guy arrangements.

within the apertures of side mounted TV and FM broadcast antennas. Equipment producing large localized wind loads, such as microwave antennas or clusters of two-way radio cabinets and antennas, should not be positioned near the center of a span between guys.

If the tower will be subjected to ice loading, it is desirable to reduce the number of guy levels to minimize loads imposed on the tower by ice accumulation on the guy cables.

The number of anchors in each guy direction is dependent on several factors including the number of guy levels, the soil conditions, topography, and obstacles. As a general guideline, it is desirable to limit the number of guy levels attached to a single anchor to five. However, there is nothing absolute about this number, and other conditions may dictate using an anchor for a greater number. There are some soil conditions where it may be economical to provide two or more smaller anchors, while in another instance the use of one large anchor might be desirable. If minimizing the area within which the tower would fall in the event of a collapse is a consideration, a minimum of two anchors should be used in each direction. Where the elevations of the anchors differ from the tower base, it is desirable to vary the distance of the anchors from the tower base to maintain nearly equal initial tensions in the guy cables. Anchors higher than the tower base should be moved toward the tower; and anchors that are lower, away from the tower. The amount of movement should be specified by the designer.

Materials

Tower Structures

Nearly all broadcast towers are made from steel because it provides the most economical structure. The selection of the grade and shape of steel is obviously an important design consideration.

Steels used for towers commonly have low carbon content with yield strengths in the range of 36,000 psi to 60,000 psi. These materials have good ductility and are suitable for welding. Some towers have been built using higher grade materials with yield strengths up to 100,000 psi, but the savings in weight are more than offset by higher base prices and increased fabrication costs. Regardless of the grade of material, the steel's physical and chemical properties should be certified by the producing mill to ensure that it conforms to the design requirements.

The shape of the material as well as its size and strength affects the tower's load carrying capacity. The shape also has a significant effect on the magnitude of loads produced by wind. Design standards permit a reduced wind load on round members as little as 57% of the wind load for flat or angular members of the same width. For this reason, solid round bars, round structural tubes, and pipe are often used. This advantage in wind load is offset somewhat by increased fabrication costs, due to the necessity of welding plates to connect the various members.

There is no one grade or shape of material that is best. The choice depends to a large degree on the preference of the designer and the type of fabricating facilities available.

A factor equally as important as the selection of the grade and shape of the structural steel is the design of the connections. For shop welded connections, the compatibility of the base and filler metals and required preheat temperatures must be considered. The procedures used must be qualified and the welders certified to use them. Inspection procedures should be compatible with the weld design.

Bolts for field assembly may be of various types. Usually those for the main load carrying members are high strength. If positive resistance to slippage of the connections is required, they should be designed as friction connections.

Guys

The most common material for tower guys is galvanized steel strand. This material has excellent strength and durability. Its "structural" elongation due to the seating of the individual wires in the strand is small and can be almost entirely eliminated by prestressing the strand to 50% of its breaking strength at the factory. This should be done for guys on tall towers with factory connected end fittings.

For guys on AM towers, and those close to FM and TV antenna apertures, a nonconductive material is sometimes desirable. Two such materials that have been used are Kevlar rope and fiberglass rods. When using these materials, careful attention must be given to protection against corona effects, fatigue, and deterioration from exposure to ultraviolet light. Also, their elongation characteristics under load must be evaluated. They require delicate handling at all times.

Just as for the tower structure, the connections for the guys are as important as the guy material itself. Some of the more common connections are as follows:

1. *Sockets* of forged or cast steel attached with molten zinc or epoxy resins develop the full strength of the guy. They are normally installed at the factory and proof loaded to 50% of the guy breaking strength. This type of fitting is most common for the larger guys used on tall towers.
2. *Dead end grips* are preformed spiral wire loops in the shape of a large hairpin. The two legs of the hairpin are wrapped around the guy with its closed end forming an eye. These grips are used for guys up to one inch in diameter and usually develop their full strength. They are easily installed in the field, but the ends must be completely snapped into place and a protective device installed to prevent ice from sliding down the guy and loosening the grip.
3. *Clips* used to clamp the ends of guys (when properly applied and tightened) develop 90% of the guys' strength for sizes up to and including ⅞

inch and 80% for larger sizes. To install them it is necessary to bend the strand back on itself to form a loop; thus, the use of clips on large cables is difficult. The saddle of U-bolt type clips must be on the load side and not the dead end which provides another potential error in their installation.

4. *Swaged sleeves* develop between 85% to 100% of a guy's strength depending on the size of the guy and equipment used to squeeze the sleeve. These fittings are usually installed at the factory and can be proof loaded. They are advantageous for connecting closely spaced insulators where the length of dead end grips is unacceptable.
5. *Wedge Type Sockets* are available for guys up to 1¼-inch diameter and develop 100% of their strength. They are most advantageous for guys larger than those for which dead end grips are available.
6. A *serving* is a connection made by rolling the individual wires of a strand back on the strand itself. This method has for the most part been replaced with dead end grips, but it is advantageous for small guys with closely spaced insulators.

Insulators

Insulators in radio frequency applications must withstand mechanical and electrical stress in a varied, changing exterior environment. Selection of insulators should be made with these factors in mind. Insulators primarily designed for 60 Hz applications are unsuitable, particular at high RF powers.

The most common insulating material is a wet process porcelain which has excellent compressive strength and good insulating capabilities for frequencies up to 2 MHz. Synthetic materials are also used.

1. *Base insulators* for AM towers are made from porcelain with appropriate steel end plates or ferrous/nonferrous castings. For guyed towers, a rocking arrangement is provided in the form of a convex plate and pin at the top, or a pivot pin at the bottom of the assembly to hold the tower in place and relieve the porcelain from bending loads which could cause cracking. For self-supporting towers, the insulators are bolted between the tower leg and base pier, and are designed to sustain both uplift and download while keeping the porcelain in compression.
2. *Sectionalizing insulators* are sometimes required to isolate sections of a guyed tower. Where a compression load only is applied, a guyed tower base with minor modifications can be used. If a tension load is anticipated, a push-pull insulator similar to the type used for self-supporting towers is required. Under no circumstances should the porcelain be put in tension.
3. *Guy insulators* are available for primary insulation and for break-up purposes. Primary insulation (insulators next to the tower) should be selected to withstand the full voltage appearing at the guy attachment point. This is to ensure that sufficient insulation remains if all the break-up insulators in the guy line flash over. Break-up insulators, used to reduce reradiation, are selected to withstand the transmitter induced voltage and static voltage. Break-up insulators are usually low voltage types, sometimes protected from flashover and subsequent power arc by a static dissipation device. Guy insulators are available in many styles classified as either compression or tension types.

   A. *Compression* insulators are designed such that the porcelain element is in compression. Simple low voltage types are a single piece of porcelain placed between interlocking loops of the guy. Such insulators are available for mechanical working loads up to 40,000 lbs. For higher loads, oil filled and open types are used. The most common uses of compression insulators in broadcasting are for break-ups and as primary insulators (in groups of three or four) on low-power antennas.

   B. *Tension* insulators come in many forms, including porcelain rods (not permitted in structural applications), fiberglass rods, synthetic ropes, and oil filled safety core types. Tension insulators are used as primary insulators with corona rings to reduce the electrical field stress at the end fittings. One insulator is required at each guy attachment point. Since the voltage level is different at each point on the tower, different voltage ratings may be required for some insulators at certain guy attachment points. Tension insulators are available in a wide range of electrical and mechanical ratings to meet most needs.

Finishes

Corrosion Protection

Steel is susceptible to deterioration from atmospheric corrosion. To prevent deterioration, the tower members and hardware must be given a protective coating. This coating is usually zinc which has excellent resistance to corrosion, and, because it is higher in the electrochemical series of the periodic table of elements, it provides cathodic protection to exposed steel surfaces adjacent to it. Even though the zinc coating may be scraped or otherwise damaged, it continues to inhibit corrosion of these exposed areas, and rust will not develop beneath adjacent zinc coats.

There are several methods for applying the zinc including hot dip galvanizing, flame spraying electroplating, and painting. All must be applied to clean surfaces.

1. *Hot-dip galvanizing* consists of dipping the steel into a bath of molten zinc. A metallurgical bond develops between the steel and the zinc which

adheres to it. When galvanizing tubular members, it is necessary to provide holes in both ends to ensure that the inside surfaces are coated. Careful attention must be given to the type of base and welded metals used, as well as to the welding and forming procedures used in fabrication, to safeguard against possible embrittlement of the steel when galvanized. When properly applied, this process provides the most durable coating.
2. *Flame spraying* consists of spraying molten zinc at high pressure onto the steel surfaces. The bond in this process is mechanical rather than metallurgical. The coating produced is more porous and has less resistance to abrasion than the hot-dip galvanized coating. It cannot be used for the inside of hollow sections or other cavities where access is difficult.
3. *Electroplating*, while suitable for small objects, does not produce a coating thick enough to withstand a hostile exterior environment. This method is not recommended for tower parts or hardware.
4. *Zinc rich paint* consists of extremely finely divided zinc in an inorganic or organic vehicle. It is not a metal coating method, but rather a painting procedure. Its resistance to abrasion and durability are less than hot-dip galvanizing. This procedure is, however, useful for maintenance.

Aircraft Marking

When required by the FCC construction permit, towers must be marked as a hazard to air traffic. One such marking system requires a tower to be painted with contrasting colors of white and international orange in alternate bands. Selection of the paint materials must be compatible with the surfaces to which they will be applied. There are several manufacturers who have one-coat paint systems that can be applied directly to galvanized surfaces. Like all paint procedures, clean, dry surfaces and suitable temperatures are essential to obtaining satisfactory adherence.

Alternate sections of international orange (Federal Standard Number 12197) and aviation white (Federal Standard Number 17875) paint provide maximum visibility of an obstruction by contrast in colors. The chromaticity and luminance standards of aviation orange and white paint should conform to Federal Standard FED-STD-595. These paints meet the specific color requirements when freshly applied to a structure. However, all outdoor paints deteriorate with time. While it is not practical to give a maintenance schedule for all climates, surfaces should be repainted whenever the color changes noticeably, or its effectiveness is reduced by scaling, oxidation, or chipping. An orange color tolerance chart is available upon request from the FAA for determining when repainting is required. The color should be sampled on the upper half of the structure, where the weathering will be greater. This tolerance chart may be obtained from:

Manager, Flight Information and Obstruction Branch, ATO-210
Federal Aviation Administration
800 Independence Avenue, SW
Washington, DC 20591

When high intensity white lights are installed on the tower and are operated during daytime and twilight, the international orange and white paint may be omitted. When medium intensity white lights are operated on structures up to 500 feet in height, the international orange and white paint may also be omitted. See the section on, "Electrical Systems for Antenna Structures," below.

The cost of repainting an antenna structure depends on many variables such as accessibility, size, height, neighboring properties, weather, elevator availability, etc. A rough guideline for estimating the cost for repainting a structure is to multiply the face dimension in feet by the height in feet to get the cost in dollars. Due to the number of variables that have to be factored in, this formula will provide only a rough estimate.

Ice Prevention

Coatings are available to reduce the adherence of water to surfaces and subsequently the formation of ice on them. However, no reliable means exists to completely remove the risk of severe ice accretion.

Access Facilities

A tower must have some access facilities in order to maintain it and the equipment the tower supports. For small towers, the bracing members of the tower itself often serve as steps, or step bolts are attached to one leg or face.

Ladders

For taller broadcast towers, a fixed ladder inside the tower is desirable. The Occupational Safety and Health Administration (OSHA) standards for these ladders require a minimum clear width between side rails of 16 inches and a maximum rung spacing of 12 inches. OSHA also requires that any continuous ladder more than 20 feet in height be equipped with a safety device. This device consists of a continuous rail, either rigid or cable, running up the center of the ladder. A clamping device attached to the climber's safety belt rides along this rail. As long as the climber is in a normal position, the clamp slides freely; if he begins to fall, a cam actuated mechanism freezes the clamp to the rail and prevents his falling.

Elevators

For tall towers supporting multiple antennas, it is often desirable to install an elevator. Most tower elevators are of the power, cable driven type with a capacity of 500 pounds to 750 pounds and a speed between 80 feet and 100 feet per minute. They consist of a drive mechanism, car, guide rails, hoist cable with supporting sheaves, tension weights, electronic controls, and a two-way communications system.

Considerable attention must be given to elevator safety features. These should include limit switches to prevent travel beyond the upper and lower landings on the tower, an automatic brake on the driving mechanism that is activated by an interruption in power, a mechanism to automatically clamp the car to the rails in the event of a broken hoist cable, and interlocks to prevent operation with the car gate open. It is advisable to determine the applicable state or municipal government regulations that may apply, and whether permits, tests, and inspections are required before the tower and elevator system are designed.

The added wind and dead loads from an elevator system are substantial and must be considered in the tower design. Also, careful attention must be given to the positioning of the ladder, RF transmission lines, and electrical conduits in relationship to the elevator. The ladder must be positioned so it is accessible from the elevator car and can be used for an emergency descent. While the elevator hoist cables can be restrained in guides on the return side, they are free to move about under wind load on the lifting side. Therefore, the conduits and transmission lines must be protected from hoist cables striking and damaging them. If a side mounted TV or FM antenna produces a high RF field within the hoistway, protection must be provided to prevent arcing between the hoist cables, the tower structure, and other appurtenances.

Transmission Line Bridges

To allow for thermal expansion and contraction, it is necessary to locate broadcast towers some distance from the transmitter building. Unless the transmission line is placed underground, it is necessary to provide a structural support for it at a height compatible with the transmitter location in the building. The top of the support can be covered with steel grating or plate to protect the line from falling ice. The details of this structure can become quite involved for sites with multiple antennas, uneven terrain and roadways, or obstacles between the tower and building.

Stairways

The lower landing for a tall, guyed broadcast tower with an elevator is often 30 feet or more above ground level. A stairway may be desirable to permit easier access to the landing. This structure can be combined with the transmission line support bridge, or it may be completely separate. It may also be desirable to install a small capacity boom above the lower landing to lift radio cabinets or other equipment onto the landing.

DESIGNING ELECTRICAL SYSTEMS FOR ANTENNA STRUCTURES

Several types of electrical systems are commonly installed on towers to obtain the greatest utility from the structure. These include systems for tower lighting, antenna de-icers, utility circuits, and power for ENG, two way and other communication facilities. To ascertain overall requirements, the criteria and specifications for these systems should be determined, before the general specifications for the tower are completed.

Marking and Lighting Antenna Structures to Meet FAA/FCC Requirements

Since an antenna structure may be a hazard to air navigation, the Federal Aviation Administration (FAA) has established certain marking and lighting requirements for broadcast towers. In general, these requirements are spelled out in an advisory circular published by the FAA entitled "Standards for Marking and Lighting Obstructions to Air Navigation" (AC 70/7460–1). Prior to submission of an application for construction to the FCC, the various options for antenna structure marking and lighting should be carefully considered, and a preference for an option should be requested on the application.

New Certification Procedure For FAA Approved Lighting Equipment

Effective January 1, 1990, ETL Testing Laboratories, Inc., Industrial Park, Cortland, New York 13045 (607-753-6711), will be administering the certification process for airport and obstruction lighting equipment. Lighting equipment approved by the FAA prior to January 1, 1990, will be accepted for the ETL certification program without additional product qualification testing for a period of five years, provided sufficient documentation exists to support a certified status. The equipment covered by this clause will still be subject to ETL's quality control audit, site production testing, and inspection. Should the FAA revise a standard or specification, the applicable equipment must then be recertified and verified by ETL.

The FAA will continue to publish Advisory Circular 150/5345–1, "Approved Airport Equipment," listing all certified lighting equipment that ETL verifies as meeting the applicable FAA specifications.

When building or modifying an antenna structure which involves the purchase of obstruction lighting equipment, it would be wise to require a certificate of compliance from the manufacturer.

FCC Lighting Requirements

Temporary warning lights. During construction of an antenna structure, for which obstruction lighting is required, a temporary light is installed at the uppermost point of the structure. This light should be similar in type and intensity to the permanent light that is required for that level. In addition, as the height of the structure exceeds each level at which permanent obstruction lights will be required, another level of temporary lighting must be used. If practical, the permanent obstruction lights may be installed and operated at each required level as construction progresses.

Inspection of tower lights. The licensee of any radio or television station, which has an antenna structure requiring illumination, is required (FCC Rules Section 17.47 (a)) to do the following:

1. Make an observation of the antenna structure lights at least once each 24 hours, either visually or by observing a properly maintained automatic indicator designed to register any failure, to ensure that all such lights are functioning properly as required.

OR

2. Provide and properly maintain an automatic alarm system designed to detect any failure of such lights, and to provide indication of such failure to the licensee.
3. Inspect, at intervals not to exceed three months, all automatic or mechanical devices, indicators, and alarm systems associated with the antenna structure lighting to ensure that such apparatus is functioning properly.

Notification of extinguishment or improper functioning of lights. The licensee of any radio or television station which has an antenna structure requiring illumination is also required (FCC Rules Section 17.48 (a)) to:

1. Report immediately to the nearest Flight Service Station (FSS) or office of the FAA any observed or otherwise known extinguishment or improper functioning of any top steady burning light or any flashing obstruction light, regardless of its position on the antenna structure, not corrected in 30 minutes. Such reports must set forth the condition of the light or lights, the circumstances which caused the failure, and the probable date for restoration of service. Further notification must be given immediately upon resumption of normal operation of the light or lights.

An extinguishment or improper functioning of a steady burning intermediate sidelight or lights must be corrected as soon as possible, however, notification to the FAA of such extinguishment or improper functioning is not required.

Recording of antenna structure light inspections in the station log. The licensee of any radio or television station, which has an antenna structure requiring illumination, must make (FCC Rules Section 73.1820 (a) (1)) the following entries in the station log in the event of any observed or otherwise known extinguishment or improper functioning of an antenna structure light:

1. The nature of such extinguishment or improper functioning.
2. The date which such extinguishment or improper operation was observed or otherwise noted.
3. The date, time, and nature of adjustment, repairs, or replacements made.

Automatic monitoring for light failures. Although the FCC no longer requires daily logging of antenna structure lighting operation, the logging of a light failure is required along with notification to the FAA of the failure and the proposed schedule for resumption of normal operation. Automatic monitoring is available in accordance with FAA/FCC requirements to alert station personnel of a light failure. Computer compatible monitoring systems are available for use with remote logging systems.

In the electronic power supply that also performs the controller function in the high and medium intensity systems, a failure relay should be included to provide automatic monitoring of the system. From these contacts on the failure relay, a signal can alert the licensee of a failure.

Time when lights should be exhibited.

1. All red obstruction lighting must be exhibited from sunset to sunrise unless otherwise specified.
2. All medium and high intensity white obstruction lighting must be exhibited continuously unless otherwise specified.

Marking With Red Lights

Flashing red beacon (L-864). This is a flashing beacon which produces aviation red light. The peak effective intensity must be 2,000 candelas ±25% in red when measured at any horizontal angle. The flashing rate must not be more than 40 nor less than 20 flashes per minute. The beacon should be lit from one-half to two-thirds of the total cycle. The light intensity during the "off" period must be less than 10% of the peak effective intensity. A change in the latest revision of AC 70/7460–1(H) states, "Intermediate level flashing beacons (L-864) may be installed either within or outside the structure if the transmission cable or the tower legs do not have an effective diameter of more than three inches. If either is more than three inches, two beacons must be installed diametrically opposite each other and on the outside of the structure."

To ensure the proper light output, the operating voltage provided at the lamp socket should not vary by more than 3% from the rated voltage of the lamp. For instance, 3% low voltage results in a 10% drop in light output, and 3% high voltage results in a 30% reduction of the rated lamp life. The measurement of lamp socket voltage must be made while under normal station operating conditions.

Steady burning red lights (L-810). These obstruction lights consist of one or more steady burning lamps ranging from 45 watts to 116 watts per fixture. The intensity must not be less than 32.5 candelas at all horizontal angles. Three fixtures are required per level.

Cost considerations. The equipment cost plus installation for a red obstruction light group (one or two flashing beacons and three steady burning side lights) can vary from $2,000 to $6,000 depending on the size of the antenna structure, its location, and other local conditions affecting installation costs.

To estimate power costs, one must first add the number of lights required. Under the latest revision of AC 70/7460–1, the red obstruction lights are specified to be mounted on the structure in groups (one or two L-864 flashing beacon and one set of L-810 steady burning sidelights). One group consumes 1,175 watts to 2,000 watts of power if the flashing beacon is set to operate for ⅔ of the cycle. This assumes zero power consumption during the off portion of the flashing beacon cycle. Assuming electrical power costs of ten cents per kWh and a yearly average of 12 hours per day of red light operation, the yearly power costs per group is approximately $515 to $875.

Lamp replacement costs depend on the labor costs of riggers, if needed, but the interval can be estimated with the following guidelines. Steady burning sidelight lamps have a rated life of 6,000 hours. These lamps should be replaced after being operated for not more than 75% of their rated life. This allows for a yearly replacement schedule. However, many tower operators replace all lamps at the same time if early replacement of one lamp is necessary due to failure.

The 640 watt flashing beacon lamps have a rated life of 3,000 hours. Seventy-five percent of rated life is reached approximately every 6 months. A 700 watt lamp with a rated life of 6,000 hours is available. This lamp has the same light output as the 640 watt lamp that is normally specified. With the 700 watt lamps installed, the yearly power cost per group increases to approximately $565 to $975.

Marking With White Xenon Lights

The FAA allows the use of white xenon lights for marking antenna structures over 200 feet in height. Lighting with high and medium intensity white obstruction lights is done to provide a high degree of conspicuousness necessary to warn pilots on a potential collision course with a structure during both day and night.

The FAA spent several years looking for the best way to mark obstructions so they would be more visible during daytime. After trying different color painting ideas, and various lighting schemes, the white xenon lights proved to be the best solution. When more tests were conducted in the early 1980s to qualify the use of medium intensity on structures 200 feet to 500 feet tall, again the medium intensity white xenon lights proved to be better than international orange and white paint under all conditions except bright front-lit sunlight.

Effective intensity is probably the least understood term when it comes to the practical use of electronic strobe lighting equipment. The term *candela* is a unit of measurement of light output. *Effective candela* is specified by taking into account the so-called Blondel-Ray relationship.

The eye is able to integrate a light pulse with virtually no reciprocity failure for pulse lengths much shorter than the decay time constant for the eye which is approximately 0.2 seconds. Therefore, the eye is sensitive to the integrated value of the light pulse and not to the peak value, which for short pulses can be stated in impressively high but meaningless figures, as far as the eye is concerned.

The effective intensity can be determined from the Blondel-Ray relationship.

$$I_e = \frac{I_m(t + 1)}{t + 0.2}$$

where:

I_e = Effective intensity in candelas
I_m = Measured intensity in candle-seconds
t = Pulse width of the light pulse (usually between the ⅓ amplitude points) in seconds
0.2 = The decay time constant for the eye in seconds

For short light pulses, the effective intensity can approach five times that of a steady light source with the same candle-second intensity.

Factors Affecting Installation, Reliability, and Life

Power conversion. The power transformers that supply white xenon lights today are typically of the constant current/voltage regulating type. This type of transformer is very efficient when used to charge capacitors and is capable of supplying a near constant current from zero volts to full flash operating voltage on the capacitor. This type of transformer is not harmed if a flash capacitor develops a short, and is less likely to destroy the capacitor. Another feature of this type of transformer is its inherent immunity to voltage transients and spikes on the power line. Since the flash energy is proportional to the square of the flash voltage, capacitor safety margins are maintained and light output can be kept within the FAA specifications over the normal input power voltage fluctuations. To guard against lightning-induced voltages on the structure, the breakdown voltage from the primary to case and to other windings should be designed for at least 5.0 kV with insulation of high enough heat rating to hold up for the desired useful life of the light.

Energy storage. The energy storage capacitors should be properly derated for the operational voltage to provide long life. Operational deratings for capacitors in common use during the past 15 to 25 years have shown deratings to 60% as being safe and long-lived. New technology capacitors using state of the art dielectric materials have shown deratings to 80% as being safe and long-lived. Because of the significant energy levels used in these lights, energy storage capacitors using minimal deratings should have built-in protection devices to minimize the possibility of overheating, rupturing, and causing damage not only to themselves but to fixture interiors as well.

Flashlamp characteristics. In the nonionized state, a flashlamp has high impedance (tens of mega-ohms); therefore, all current from the power supply initially flows into the capacitor. As the voltage across the capacitor is increased, a point is reached, called the breakdown voltage, where xenon atoms are ionized and the impedance of the flashlamp starts to drop. In a short period of time, enough xenon atoms are ionized so that a low-impedance path is formed from anode to cathode, and current flows from the capacitor through the flashlamp. As this occurs, more xenon atoms are ionized; the arc impedance drops to the milliohm region; and the arc expands outward, eventually filling the bore of the flashlamp. Most of the energy stored in the capacitor is expended in a matter of microseconds so that, eventually, the current through the flashlamp drops to such a low level that the tube de-ionizes and stops conducting. At this point, the capacitor starts recharging. An inductor is placed in series with the flashlamp and the capacitor. The values of the inductance, the capacitance, and the charging voltage are chosen carefully so that the energy is transferred to the flashlamp in a critically damped pulse. To ensure long life, the flashlamp should be designed so that at maximum ratings its loading is not more than 6% of explosion energy.

Optical design. For environmental considerations, the optical design is very important. These lights are specifically designed to alert pilots to the presence of the structure. Stray light reaching inhabited areas on the ground may be minimized by using louvers to block the direct radiation from the flash tube and polyfocal parabolic reflectors to provide a very sharp cutoff on the lower side of the light beam. Because of the critical beam shaping, care must be taken in the initial adjustment of the lights, to give the local residents the protection available from the optical design.

Physical installation requirements. The size of the lighting fixture must be considered because of the wind loading that is added to the structure. The size and weight can also become a factor during maintenance procedures. Lighting fixtures must be manufactured with noncorrosive enclosures such as stainless steel or fiberglass material.

Medium Intensity Lights for 200 Foot to 500 Foot Towers

When to use medium intensity lights. Medium intensity white lights can be used to mark antenna structures with an overall height not exceeding 500 feet and eliminate the need for painting. The red light for nighttime marking is usually eliminated also; however, in some locations neighboring landowners have insisted on the use of red lights for nighttime. A change in the latest revision of AC 70/7460–1(H) states, "Intermediate level medium intensity beacons (L-865) may be installed either within or outside the structure if the transmission cable or the tower legs do not have an effective diameter of more than three inches. If either is more than three inches, two beacons must be installed diametrically opposite each other and on the outside of the structure."

Cost considerations. The equipment cost for medium intensity lighting is about $2,400 to $3,000 per beacon, and includes all installation material. Because of variables such as location, weather, other work being done on structure etc., the installation cost can vary from about $500 per beacon up.

Each beacon system operating 24 hours per day will consume about 920 kWh per year. Assuming a rate of ten cents per kWh, the yearly operational cost is approximately $92 per year.

Medium intensity systems warranties vary from one to two years on parts. One system manufactured is a self contained unit with the electronic power supply built into the beacon flashhead. All others manufactured have the electronic power supply in a separate cabinet which is usually installed at the base of the structure. Unless the beacon fails to operate properly, a scheduled yearly routine maintenance check of the electronic power supply is all that is necessary.

Optics designed to shield ground. Two types of medium intensity beacons are presently available. One type uses a Fresnel lens to form a beam and thus obtains a gain in the peak light output. The Fresnel lens redirects the light that would normally fall on the ground near the structure and up to the sky towards the horizontal. The other type uses three linear flash lamps with parabolic reflectors to control objectionable light radiated toward inhabitants near the structure.

High Intensity Lights for 200 Foot to 2,000 Foot Towers

When to use high intensity lights. The FAA will require high intensity lights on an antenna structure when its location and height are in the airspace normally used by visual flight rules (VFR) pilots. This includes four miles on either side of a vector airway, along natural flyways such as rivers, railroads, interstate highways, and areas near airports.

The licensee may choose to install high intensity light on an antenna structure over 500 feet because of the long term economic gains. The FAA must approve the request to use high intensity lights.

Cost considerations. A high intensity lighting system requires a level of lights to be installed every 350 feet or less. At least three lights are used at each level. The average cost per level for the lighting equipment is about $10,000. The cost of installation can vary widely due to structure location, height, and other variables.

Each level using three high intensity lights operating 24 hours per day consume about 3,942 kWh per year. Assuming a rate of ten cents per kWh, the yearly power cost is approximately $395.

High intensity system warranties vary from one to two years for parts and labor. In order to continue to meet the light output specification, the flashtubes should be changed approximately every two to three years. Unless there is a light failure, an annual routine maintenance check to perform preventative maintenance should keep the system in good working order.

Shielding the ground with louvers. A set of louvers can shield the ground from the light radiating from the flashtube directly toward the ground near the tower. The design of the reflector and other light controlling devices can also direct the light output so that a very small percentage of the light is radiated toward the ground.

Other Electrical Circuits to Consider in Antenna Structure Planning

Antenna Deicing Circuits

Antenna deicing circuits should be designed to be installed using the main electrical conduit. If a choice of deicer operating voltage is available, consider choosing the highest voltage under 600 volts in order to minimize wire size. In many cases a 208/240 volt deicer system is fed with 480 volts using a step down transformer to reduce the I^2 R losses incurred in the long runs feeding systems on tall towers. Deicing systems should be designed to provide no more than a 5% to 10% voltage drop to ensure the proper operation of the deicer since the available heating power is proportional to the square of the voltage. Most automatic deicer control circuits require a shielded triplet which may also be installed in the main electrical conduit.

AC Utility Circuits
An AC utility circuit to provide access to 120 volt power at several elevations on the structure will often save time during maintenance and repair work on the structure and associated systems.

Elevator Control Circuits
Wiring for the elevator control circuits, if needed, may also be installed in the main electrical conduit.

Sound-Powered Telephone
A sound-powered telephone system in a separate ½-inch conduit to provide communications with ground personnel from the various working platforms can likewise save time during maintenance and repairs.

Power at Platforms for Communications Systems
With the recent increase in two-way communications and cellular telephone activity, many tall tower installations have found it advantageous to provide several platforms with up to 150 amp, 120 volt service which will accommodate a large number of communication systems.

A separate conduit with 20 to 100 pair shielded cable is normally used to provide audio input for the two-way communications systems.

Design Criteria For Electrical Systems

The following criteria should be regarded as guidelines for tower and system design. Final designs must meet all applicable state and local regulations and codes.

Conduit Sizing
The size of the conduit is determined by the number and size of conductors that must go into the conduit. If all conductors are the same size, a table in the National Electrical Code (NEC) spells out the size of the conduit needed. However, in most cases the various electrical circuits on a tower require a large variety of conductor sizes. The following table from the NEC Handbook shows the maximum fill to be used when dealing with many sizes of conductors.

Percent of Cross Section of Conduit for Conductors

Number of Conductors	1	2	3 or more
All Conductor Types (Except lead covered)	53	31	40

In practice, most tower installations are specified for 30% or less conduit fill.

Junction and Pull Box Sizing
In straight pulls the length of the box should not be less than eight times the trade diameter of the largest conduit. Where angle or U pulls are made, the distance between each raceway entry inside the box and the opposite wall of the box should not be less than six times the trade diameter of the largest conduit.

Selecting Conductor Sizes
Each electrical circuit must be evaluated to determine the voltage drop that can be tolerated. With the voltage drop determined, the size of the conductors needed can be calculated. The following table is a guide for acceptable voltage drops for various antenna structure circuits:

Type of Circuit	Percent Voltage Drop
Red Lights	3
Strobe Lights	10
Deicer Circuits	10
Utility Circuits	10

Using boost transformers. Using boost transformers to make up for voltage drop in a circuit is a practice commonly used in the design of incandescent lighting for towers to allow for the use of smaller wire sizes. Fluctuations of the voltages in these circuits can shorten the life of the lamps.

Circuits for Isolated Antenna Structures
To connect a 60 Hz supply to an insulated antenna structure, a device to isolate the RF from the 60 Hz supply is required. The two methods commonly used are chokes with a very high RF impedance or isolation transformers. The lighting load and RF voltage at the antenna structure base must be considered during the selection process.

Chokes are relatively low cost and are suitable for low power applications. For high power and most directional arrays, where a high degree of tuning stability is required, ring isolation transformers offer a low fixed capacitance and high flashover capability.

DESIGN STANDARDS

Industry

The vast majority of towers in the United States have been designed in accordance with the Electronic Industries Association EIA-222, Structural Standards for Steel Antenna Towers and Antenna Supporting Structures. This standard has been used since 1959 when it replaced the Radio-Electronic-Television Manufacturers Association (RETMA) Standard TR-116. The current revision "E" of EIA-222 was issued in 1991. It is an approved American National Standard and carries the designation ANSI/EIA/TIA-222-E-1991.

This standard is intended to provide *minimum* criteria for specifying and designing steel antenna towers and antenna supporting structures. Unlike general specifications and building codes, it is applicable only to antenna towers and supporting structures. As such it contains criteria specific to these structures that are not readily available elsewhere. Therefore, it is always advisable to specify that your tower must conform to this standard.

"Appendix A: Purchase Checklist" of this standard is provided to alert the purchaser to the most common areas where site-specific data may be required to supplement the minimum criteria of the Standard.

Statutory

Most municipal and state governments have statutory codes regulating the design of structures. Many of these are patterned after or include one of several model codes. The most common of these are:

- Building Officials and Code Administrators International (BOCA) *Basic Building Code*
- International Conference of Building Officials (ICBO) *Uniform Building Code*
- Southern Building Code Congress (SBC) *Standard Building Code*.

These codes cover all types of structures and are directed primarily toward conventional types of buildings. As such, they do not contain all the criteria necessary to design broadcast towers. For example, none of them includes a recommended safety factor for guy cables.

The industry standard ANSI/EIA/TIA-222-E-1991 is compatible with these codes. In fact, its use for calculating and applying wind loads is required by the BOCA *Basic Building Code* and, as an approved American National Standard, is permitted by the ICBO *Uniform Building Code*.

Since it is necessary to comply with the applicable statutory requirements, it is important to determine what these requirements are and include them in the purchase specification for the tower.

LOADS, ANALYSIS, AND SAFETY FACTORS

Loads

In addition to a tower's own dead weight and the dead weight of the appurtenances and equipment it supports, the tower must withstand the forces of nature: wind, ice, temperature changes, and earthquakes.

Wind Load

Wind produces a principal load on tower structures. For design purposes it is represented as a horizontal static force.

Wind load is specified in terms of a basic wind speed at 10 meters (33 ft.) above ground level. The ANSI/EIA/TIA 222-E-91 Standard provides a tabulation of recommended minimum values for this speed for each county in the United States. This standard also gives specific procedures and factors for calculating wind loads considers the following:

1. Wind pressure is proportional to the square of wind speed.
2. Wind speed and, consequently, wind pressure vary with respect to the height above ground.
3. The effects of gusts of brief duration which exceed the fastest mile basic wind speed.
4. The effects of the configuration, size, proportions, shape and orientation with respect to the wind direction of the structural components of the tower and its appurtenances.

An example of wind load calculations for a typical broadcast tower is given in Fig. 3.

Since the wind may act from any direction it is necessary to apply the calculated wind loads in any horizontal direction to determine the maximum stresses produced in the structure. For a triangular tower, three directions must be considered; while for a square tower, two are sufficient. These are shown in Fig. 4.

In addition to this direct load in the direction of the wind (drag), there may also be a component of load perpendicular to the wind direction (lift). These lift components are calculated in a manner similar to that for drag forces using different shape coefficients that vary with respect to the angle of attack between the member's geometric axis and the wind direction. They are most significant for wind acting on guy cables, microwave antennas, and rectangular waveguides.

Ice Load

Ice accumulations have two effects on a tower. The weight of the ice acts directly on the structure in the same manner as the dead weight. The ice accumulation also increases the area exposed to the wind and consequently the load produced by the wind. This increase is substantial on small components such as guy cables, tension rods, ladders, small diameter transmission lines, and reflector screens for antennas. It is also possible for the ice accumulation to alter the aerodynamic shape of members, thereby requiring the use of a different coefficient in calculating the wind load. An example is a set of closely spaced parallel coaxial lines. Without ice each would be considered a round cylindrical member. With accumulated ice, they would present a large flat area to the wind requiring a different coefficient.

Ice produces an entirely different stress distribution in a tower than wind, so it is not reasonable to merely increase the design wind load to provide for ice accumulations. It is also a misconception that ice will break up and blow off the tower, and therefore, ice and wind need not be considered simultaneously. The

PROCEDURE

1. Determine basic wind speed V (mph). Minimum recommended values in Sect. 16
2. Calculate exposure coefficient $K_Z = \left(\frac{Z}{33}\right)^{2/7}$ where Z is the height (ft) above ground level to mid point of section. $1.0 \leq K_Z \leq 2.58$.
3. Calculate velocity pressure. $q_Z = 0.00256\, K_Z V^2$ (lbs/sq. ft).
4. Calculate gust response factor. $G_\mu = 0.65 + 0.60/(h/33)^{1/7}$ $1.0 \leq G_\mu \leq 1.25$ where h is the total height (ft) of the tower structure.
5. Calculate the projected areas A_F and A_R (sq. ft) of the flat and round structural components in one face including any linear appurtenances attached to a face and considered as structural components. Include the thickness of any ice load.
6. Calculate the solidity ratio $e = \frac{A_F + A_R}{A_G}$ where A_G is the gross area (sq. ft) of one tower face as if the face were solid.
7. Calculate the structure force coefficient
   $C_F = 3.4e^2 - 4.7e + 3.4$ (Triangular cross section)
   $C_F = 4.0e^2 - 5.9e + 4.0$ (Square cross section)
8. Calculate reduction factor for round structural components $R_R = 0.51e^2 + 0.57 \leq 1.0$
9. Determine direction factors D_F and D_R from Table 2 for the wind direction considered.
10. Calculate the effective projected area A_E (sq. ft) of the structural components.
    $A_E = D_F \cdot A_F + D_R \cdot R_Z \cdot A_R$
11. Calculate the projected area A_A (sq. ft) of each individual appurtenance not considered as a structural component.
12. Determine the appropriate force coefficient C_A from Table 3 for each individual appurtenance and sum the products $C_A \cdot A_R$ (sq. ft).
13. Calculate total wind force.
    $F = q_Z\, C_{TH}\, [C_F A_E + \Sigma\, C_A A_A]$ (lies)

EXAMPLE

1. V = 80 mph.
2. $K_Z = \left(\frac{600}{32}\right)^{2/7} = 2.29$
3. $q_Z = 0.00256(2.29)(50)^2 = 37.56$ lbs/sq. ft
4. $G_\mu = 0.65 + 0.60/(1200/33)^{1/7} = 1.009$
5.

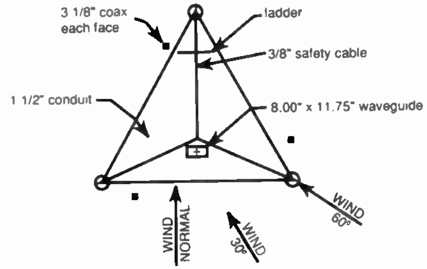
	Structural Component	A_F	A_R
Legs	$2 \times \frac{4}{12} \times 7.5$	—	5.00
Gusset P	$4 \times \frac{6 \times 8}{144}$	1.33	—
Horizontal	$\frac{2.5}{12} \times 8.67$	1.81	—
-Diagonals	$2 \times \frac{0.75}{12} \times 10.84$	—	1.36
3½" coaxes	$\frac{3.125}{12}\left(7.5 - \frac{2.5}{12} - \frac{2 \times 93}{12}\right)$		1.36
	Σ	3.14	8.22

6. $e = \frac{3.14 + 8.27}{10.33 - 7.5} = 0.147$
7. $C_F = 3.4(0.147)^2 - 4.7(0.147) + 3.4 = 2.78$
8. $R_R = 0.51(0.147)^2 + 0.57 = 0.58$
9. For wind normal $D_e = 1.0$, $D_R = 1.0$
10. $A_e = (1)(3.14) + (1)(0.58)(8.22) = 7.91$
11.

	Appurtenance	A_A	C_A	$C_A \cdot A_A$
Ladder rails	$\frac{2 \times 0.75}{12} \times 7.5$	0.74	1.2	1.13
Ladder rungs	$\frac{7 \times 0.75}{12} \times 1.33$	0.58	1.2	0.70
Safety cable	$\frac{0.375}{12} \times 7.5$	0.23	1.2	0.28
1½" Conouit	$\frac{1.9}{12} \times 7.5$	1.19	1.2	1.43
Waveguide	$\frac{11.75}{12} \times 7.5$	7.34	2.0	14.68

12. $\Sigma\, C_A \cdot A_A = 18.22$ Sq. ft
13. $F = (37.56)(1.009)[(2.73)(-.91) + 15.22] = 1524$ lbs.

$F = \frac{15.24}{7.5} = 203.1$ lbs/ft.



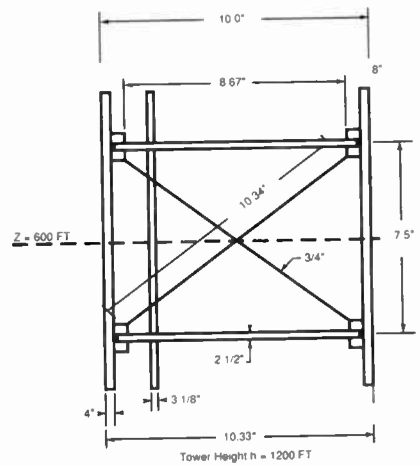


Figure 3. Wind force calculations for latticed towers using ANSI/EIA/TIA Standard 222-E-1991.

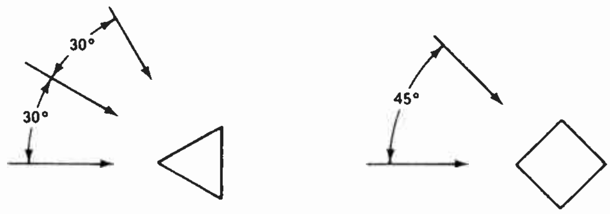


Figure 4. Wind directions to be considered.

ANSI/EIA/TIA-222-E-91 Standard states that unless otherwise specified by the purchaser, 75% of the windload shall be applied in combination with the ice load. This is equivalent to reducing the basic wind speed by a factor of 0.866. In areas where ice is likely to form, it is advisable to specify a basic wind speed to act concurrently with the ice load as well as the basic wind speed without ice, i.e., 70 mph with ice and 80 mph without ice.

While the 222-E Standard emphasizes the need to consider ice loads, it does not provide specific recommendations for the magnitude of the accumulation. This responsibility is left with the purchaser.

Temperature Changes

Changes in temperature have no significant load producing effects on self-supporting towers, but they can on guyed towers. Because of their differences in length, the guy cables expand and contract different amounts than the tower itself and thereby require elastic deformations from stress changes. The effects are greatest for those cables having the flattest angle with the ground. While the stresses produced are considerably less than those produced by wind and ice loads, they should be considered in the design of guyed towers.

Seismic Loads

Loads due to earthquakes are considered to act horizontally and are dependent on the mass and stiffness of the tower. They are usually less than those produced by wind but are distributed in a different manner. Procedures for calculating these loads are given in some of the design standards including ANSI A58.1, but not in ANSI/EIA/TIA-222-E-91. While a tower properly designed for wind loads is usually adequate for seismic forces, they cannot be neglected in areas with frequent and intense earthquake occurrences.

Structural Models and Analysis

A self-supporting tower may be described structurally as a cantilevered space frame or truss. Although it may have many different members, it is a relatively simple structure, and the determination of the forces in the individual members due to the applied static loads is easily done using fundamental principles of structural mechanics. The potential modes of failure are buckling of individual leg or bracing members under compressive loads, and shear or tension failures of the connections.

A guyed tower is a much more complex structure than a self-supporting tower. Whereas there is only one basic path through a self-supporting tower for the loads to be transferred to the ground, there are several for a guyed tower. The distribution of the loads among these paths is dependent upon the relative stiffnesses of guy systems and the tower shaft.

Each span of the tower has a stiffness with respect to relative deflections from axial and shear forces and bending and torsional moments. These stiffnesses are a function of several variables, including the geometric configuration, the mechanical properties, and the sizes of the individual members.

Each guy cable also has a stiffness with respect to movement of its attachment point to the tower that is a function of the amount of initial tension, the magnitude of ice load, and the magnitude and direction of wind load on the cables. By evaluating all of these, it is possible to simulate all the guys at a given level as a spring having a specific stiffness. Because of the nonlinearity of some of the relationships involved, the spring constant derived is only valid for a specific set of conditions and for a finite range of translation. Similarly, a torsional spring constant can be derived. It is interdependent with the translation stiffness and is also valid for only a finite range of translation.

Another difference between a guyed and self-supporting tower is the magnitude and significance of the axial load. For a self-supporting tower this is composed only of the gravity loads from the tower, its appurtenances, and any ice load. It is independent of wind load, and its effects on individual member loads are relatively small. The axial load for a guyed tower includes in addition to the gravity loads, the vertical components of the tensions in the various guys. Since these tensions are directly affected by the wind loads, the axial load is now dependent upon wind load, and its effects on the individual leg members are relatively large. Tension in the guy wires also produces an additional bending moment on the tower equal to the product of the axial load and the deflection of the tower.

Despite the complexity of the relationships involved, the availability and widespread use of computer systems permits accurate structural analysis of guyed towers. There are several different structural models that may be used.

One of the most commonly used idealizes the tower shaft as a continuous beam-column on nonlinear elastic supports (the guys) subjected to simultaneous transverse (wind and/or seismic) and axial (dead, ice, and vertical components of guy tensions) loads.

The modes of failure are buckling of individual leg or bracing members under compressive loads; rupture of bracing members, guys, or guy anchor arms under tensile loading; and shear or tension failures of the connections.

Dynamic Considerations

As previously mentioned, even though wind and earthquakes involve kinetic energy, their effects are simulated by equivalent static loads determined in accordance with the design standards. In recent years there have been more sophisticated efforts to investigate the actual response of tower structures to the dynamic aspects of wind gusts. A conclusion drawn from these studies is that the bending moments in the upper portions of tall guyed towers are considerably higher than those determined by the usual static analysis. Consequently, the loads imposed on the vertical legs and their splice connections would be amplified beyond safe limits indicating a potential failure condition. Considering the usual fundamental periods of tall guyed towers, it appears that towers taller than 1,200 feet to 1,300 feet should be investigated dynamically as well as statically.

There are two other phenomena related to the dynamics of wind that are important in guyed tower design. These are aeolian vibrations and "galloping," both of which involve periodic loading.

Aeolian vibrations are low amplitude, high frequency movements which occur in the tower guy cables due to a phenomenon known as vortex shedding. If they are not suppressed through the use of dampers, they can result in destruction of the filaments in the tower lights at the least, or fatigue failure of a guy cable and collapse of the tower at the worst. Dampers attached at one or both ends of the guy cables have proven effective in controlling these vibrations and should be considered for all tall guyed towers.

Galloping is a condition of instability involving large amplitude, low frequency movements. It is caused by the perpetual amplifications of periodic loads due to the motion of the body itself. The most dramatic and well-known example of galloping is the collapse of the Tacoma Narrows suspension bridge in 1940.

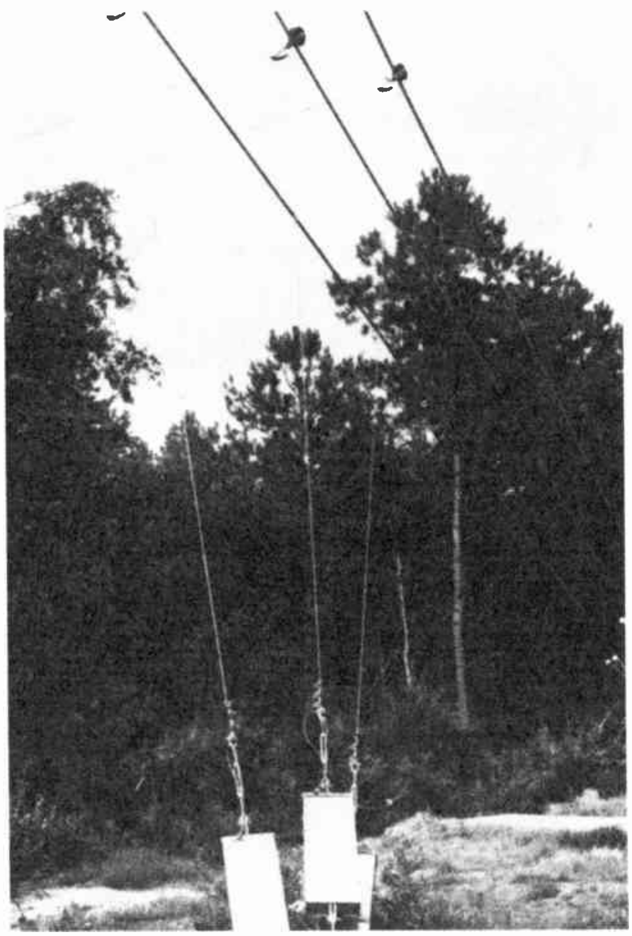


Figure 6. Anchor connection with snubber system to prevent galloping.

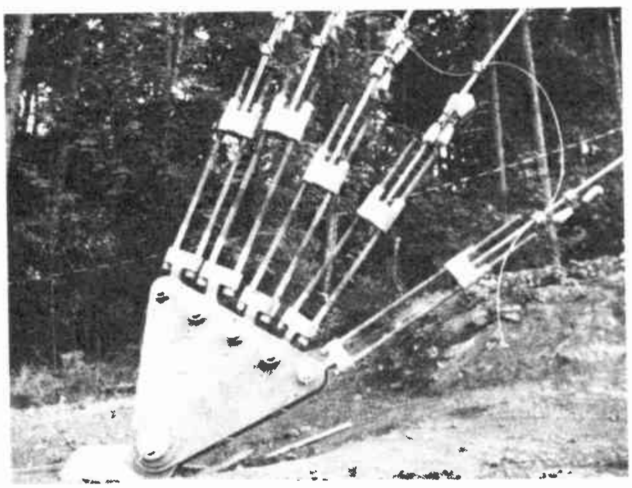


Figure 5. Typical anchor connection with dampers to prevent aeolian vibrations.

For tower structures, galloping is usually associated with the guy cables on tall towers, but in at least one instance it was related to a large rectangular waveguide. There have been several different methods involving detuning and energy dissipation used for preventing galloping in guy cables that appear to be successful. In the case of the rectangular waveguide, galloping was controlled by moving the waveguide inside the tower along the centroidal axis from its original position on the outside of one face. This reduced the torsional rotation of the structure which was the source of the perpetuating force. Based on this experience, it would appear prudent to always install this type of waveguide inside the tower unless adequate torsional rigidity is provided throughout the height of the tower.

Allowable Stresses and Safety Factors

Towers, like all other structures, are designed so that the maximum anticipated stresses are less than those which would cause failure. This ratio of failure stress to maximum allowable stress is known as the safety factor. It is intended to provide for several variations from the ideal conditions assumed for design including loads greater than anticipated, imperfections in materials, and tolerances in fabrication and construction.

The ANSI/EIA/TIA-222-E-91 Standard refers to the American Institute of Steel Construction (AISC) "Specification for the Design, Fabrication and Erection of Structural Steel for Buildings" for the design of the structure's members and to the American Concrete Institute (ACI) "Building Code Requirement for Reinforced Concrete Structures" for the design of the reinforced concrete foundations and guy anchors.

For towers under 700 feet in height the allowable stresses given in the AISC specification may be increased by one third. For towers 1,200 feet or taller, no increase is permitted. For towers between 700 feet and 1,200 feet, the amount of increase permitted is determined by linear interpolation.

In a similar manner, the required reinforced concrete strength for towers under 700 feet in height must equal 1.3 times the calculated reactions and for towers 1,200 feet and taller, 1.7 times the calculated reactions. Linear interpolation between these two values is used for towers between 700 feet and 1,200 feet to determine the required strength.

The minimum safety factor for guy cables is two for towers under 700 feet in height and 2.5 for towers 1,200 feet and taller with linear interpolation applied between these two heights.

EFFECTS OF ANTENNAS AND TRANSMISSION LINES

Except for AM radiators, the tower is the "necessary evil" to support the broadcast antennas and transmission lines at a suitable height above ground. Thus the effects of this equipment are of paramount importance.

Loads

Every antenna imposes a wind load and a dead load on the tower. If the antenna is mounted atop the tower, it also imposes an overturning moment. If it is mounted on a side of the tower, the antenna imposes a torsional moment. For TV and FM broadcast and microwave antennas, these loads are relatively large, and their location has a significant effect on the placement of guy cables.

Transmission lines feeding the various antennas also impose wind and dead loads on the tower. These loads are distributed uniformly between the antenna and their entry point near the base of the tower. The total load produced by a coaxial line or waveguide is frequently greater than that produced by the antenna itself. The shape of the transmission line influences the magnitude of the wind load, with circular or elliptical lines having loads that are 60% of those for rectangular lines with the same projected area.

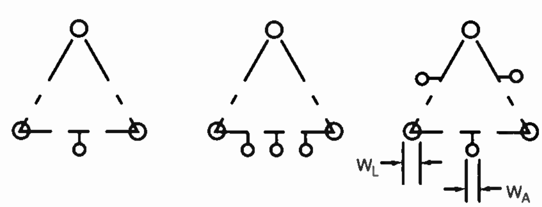
It is important not to overlook the support system required for transmission lines. Some large waveguides have support systems that require nearly continuous vertical structural members that add substantial wind and dead loads. Small, flexible lines require supports at a maximum interval of three to four feet, which is often less than the vertical spacing of horizontal members in the tower. Thus, it may be necessary to provide an additional support structure for these lines, again adding to the total load.

An important consideration when locating transmission lines is that the 222-E Standard permits linear appurtenances that are attached to the tower face, and do not extend in width beyond the normal projected area of the face, to be treated as structural components rather than individual appurtenances for purposes of calculating wind load. The effect of this is to substantially reduce the magnitude of the calculated wind load on these appurtenances. An example is shown in Fig. 7. When this procedure is used, it is necessary that the transmission lines be installed around the tower faces as assumed for the design. Because this may not be easily controlled for lines possibly installed in the future, it may be prudent to prohibit the use of this procedure.

Width Restrictions

Some antennas impose restrictions on the width of the supporting tower. One common example is a side mounted FM antenna requiring a maximum width of 18 inches to 24 inches. For antennas with more than eight bays, this results in a very slender structure. When placed at the top of a tall guyed tower, the design of the guy system for this structure becomes

COMPARATIVE WIND LOADS ON
LINEAR APPURTENANCES



RS–222–C	1.00 (Ref.)	3.00	3.00	CYLINDRICAL $W_A \leq W_L$
RS–222–D	1.53	4.59	1.78	
RS–222–E	1.53	4.59	1.78	
RS–222–C	1.00 (Ref.)	3.00	3.00	FLAT $W_A > W_L$
RS–222–D	1.70	5.10	5.10	
RS–222–E	1.70	5.10	2.07	

Figure 7. Effects of linear appurtenance locations on calculated wind loads.

extremely critical. Use of a cantilevered pole structure above the top of the main tower should be considered for these cases.

Another example of width restriction is a panel-type TV or FM antenna mounted on the faces of the tower. Here, too, it is often better to support these antennas on a cantilevered structure above the main tower rather than placing guys within the aperture of the antenna.

As previously mentioned, it is desirable to place large waveguides inside the tower near the vertical centroidal axis to prevent large torsional loads. This requires a tower having a minimum face width in the range of seven feet to eight feet to accommodate the waveguide and its supports.

Initial and Future Considerations

Because the antennas and transmission lines have such a significant effect on the tower design, it is important to consider all possible uses for a tower before it is designed. It is better to have unused capacity than to undergo expensive modifications or replacement in several years to obtain additional height or accommodate another antenna. This has become apparent in recent years with the proliferation of microwave, two-way communications, and cellular radio systems.

When providing for multiple antennas, it is important to determine not only the number and type of antennas and lines, but also their location on the tower. The distribution of load is equally important as magnitude.

Triangular top platforms ("candelabras") to support broadcast antennas on each corner have been successfully used for many years. They have the advantage of placing all antennas at the same height. A variation of this platform to support only two antennas ("tee-bars") has also been used. Both of these systems require multiple guy cables at the top platform to provide adequate torsional stability. It is possible to design the tower for a multiple antenna support platform without installing all antennas at the same time.

Another arrangement of multiple antennas is stacking, i.e., installing one antenna atop the tower and arranging others along the tower, one below the other. This arrangement can also be combined with a multiple antenna support platform.

If capacity for microwave antennas is required, it should be provided near guy levels and preferably above to minimize interference with the guy cables. The guy system and web bracing at these levels must be designed to provide adequate torsional rigidity.

Capacity for small antennas may be provided at various locations throughout the height of the tower. One arrangement for a large number of antennas is to provide a platform around the outside of the tower that is large enough to support the radio equipment for these antennas. The antennas can be mounted on the outside railing of the platform, thereby requiring only a short run of coax. Electrical power must be provided to the platform. This arrangement imposes a large concentrated load at the platform location with a relatively small uniform load between the tower base and the platform. If the same number of antennas were mounted along the tower and each fed by an individual coax line from the base, there would be only small concentrated loads at the antenna locations, but a relatively large uniform load due to the lines. This is an entirely different distribution of load, and would have a pronounced effect on the design.

Another important consideration for future antennas or height extensions is the electrical system. If an extension in height is planned, the wiring for the aircraft warning light system should be designed so that any additional lights can be connected to the system without adding or replacing wires in the existing conduit. The same holds true for any circuits required for future antennas. If the necessary wiring cannot be provided during the initial installations, capacity should be provided for additional conduits to hold the future circuits.

Replacement, Relocation, or Additions to Existing Towers

Since every tower has been designed for a specified arrangement of equipment, changes should not be made without considering their effects on the structural adequacy of the tower.

Two common misconceptions related to changes in equipment are "lower is better," and "smaller is better." Neither is necessarily correct, especially for guyed towers. Decisions based on these premises can have serious consequences.

It is much better to have a structural analysis of the tower made by a structural engineer experienced in tower design. Because of the significant changes that have been made in the methods of specifying loads in the various revisions of the design standard, the analysis should be made using the same criteria for wind and ice loads used for the original design and also for the current revision of the standard. An example of the differences in calculated loads is given in Fig. 8. This analysis will determine if any overstresses would occur in the tower or its foundations, and what modifications and reinforcing would be required to retain the structural integrity. To make this analysis, it is necessary for the engineer to have complete data on the tower and its foundations including configuration, member sizes, and material strengths. The use of presumptive values can result in an analysis with little value.

FOUNDATIONS AND ANCHORS

It is most difficult to predict the cost of the foundation system of a tower installation. This is due to the nonhomogeneous nature of soils, and the uncertainty of the conditions that may exist below grade. Therefore, it is necessary to have an investigation made of the subsurface soil conditions.

It is important to note that the soil design parameters given in the ANSI/EIA/TIA-222-E Standard are in-

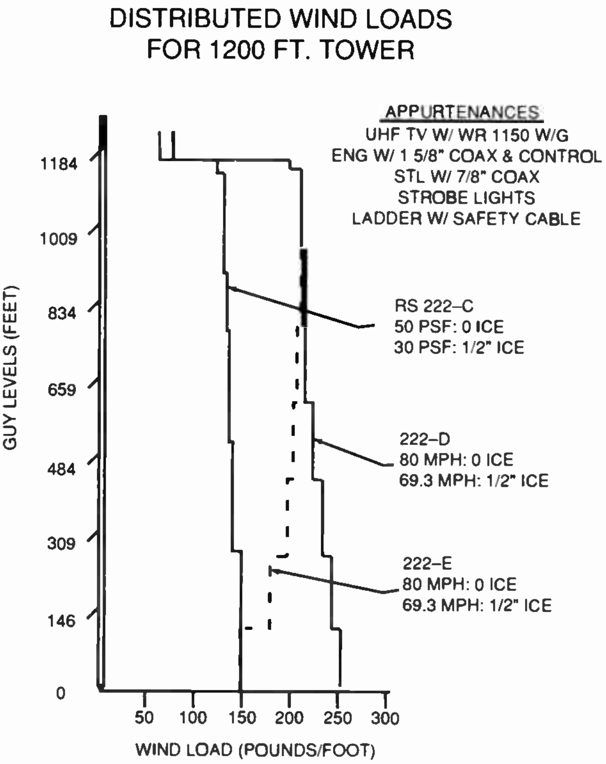


Figure 8. Comparison of calculated wind loads for different revisions of the design standard.

tended to serve only as a basis for preliminary design and estimating of foundation costs prior to obtaining specific soil data. They should not be used for the final design without verification by geotechnical investigation.

Soil Investigation

The soil investigation should be made by an engineering firm which specializes in soil investigations and evaluations, and is familiar with the general area of the tower site. It should consist of making a test boring at each foundation and guy anchor location, analysis of soil samples taken from the borings, determination of the structural properties of the soils, determination of ground water levels, recommendations of parameters for designing the foundations, identification of any special construction procedures required, and recommended backfill specifications. If piles or rock anchors are necessary, recommendations related to these should be provided. It should also address requirements for frost protection and buoyancy effects.

Because the loads imposed on foundations for towers are unique from those for conventional buildings (tower foundations have large uplift and horizontal components), it is important to provide the soil engineer with the loading conditions before they make their investigation. This will enable the engineer to plan their work in a manner suitable for obtaining and reporting the characteristics relevant to designing for the projected foundation loads.

Self-Supporting Tower Foundations

Except for relatively small towers with narrow base spreads, isolated foundations at each leg are usually more economical than a single mat for all legs. These foundations may be spread footings, drilled caissons or driven piles. If sound rock is present at shallow depths it is often economical to anchor the footing to the rock. These anchors should be proof-loaded to ensure their holding capacity in uplift.

Since these foundations are subjected to large uplift forces, it is important to consider buoyancy effects if ground water is present. Also, if driven or cast-in-place piles are used, they must be adequately anchored to the reinforced concrete cap.

Guyed Tower Base Foundations

These foundations may be spread footings, drilled caissons, or driven piles. Since they are subject only to downloads with relatively small horizontal forces, they require no special anchorage details for uplift, unless they are placed above expansive soils. Buoyancy is usually not a problem.

Guy Anchors

Deadmen (buried reinforced concrete blocks), drilled caissons, or driven piles may be used for these foundations. If sound rock is present at shallow depths it is often economical to anchor the foundation to the rock.

These foundations are subject to large horizontal forces as well as vertical uplift. Therefore, deadmen must have a large enough frontal area bearing against the soil to resist sliding; drilled caissons must have sufficient diameter and depth to prevent excessive lateral deflection as well as pull out from uplift; and driven piles must be sloped to prevent large lateral loads being imposed on them. Rock anchors may be installed along the slope of the resultant of the horizontal and vertical loads, or they may be installed vertically and post-tensioned to clamp the concrete cap to the rock to prevent sliding. Because of the uplift forces, it is important to consider buoyancy due to ground water and to provide adequate anchorage for driven or cast-in-place piles.

Construction

Since nearly the entire foundation system will be below finished grade and not subject to later inspection, it is important to carefully monitor its construction. The following items should be verified:

- Location and alignment of anchors in plan and elevation
- Condition of excavation surfaces on which concrete will be placed
- Position, size and grade of reinforcement steel
- Placement of concrete to prevent voids and air pockets

- Strength of concrete using test cylinders for 7- and 28-day break tests
- Protection of concrete against freezing during the curing period
- Placement and compaction of backfill
- Driving records and/or load tests of piles
- Proof loading and post-tensioning of rock anchors

For towers with extensive foundation systems, it is advisable to retain an independent inspections service for this work. Often the firm making the subsurface soil investigation can also provide this service.

ERECTION

The erection of towers is a highly specialized field and should be done only by firms having the proper equipment and experienced rigging personnel. It is also important that the firm have adequate insurance coverage including worker's compensation, general and automobile liability, and builder's all-risk for direct damage to the tower and antennas being erected.

Owner's Preparation

Prior to the arrival on site of the erection crew, the site should be made ready for work to begin. These preparations include:

Access
Suitable access from public roads for delivery of the tower materials and erection equipment is required. While a paved roadway is not necessary, the access must be able to handle heavy trucks and construction equipment.

Permits
All necessary building and construction permits should be obtained and posted as required. Any inspections required during construction should be noted.

Clearing
A work area must be cleared to permit unloading, sorting, and assembling the tower. Paths from the tower base to the guy anchors must be cleared for a width adequate to permit hauling the guy cables to the anchors and pulling them to the tower. Paths must also be cleared for the hoist line from the tower base to the hoist location, and for the tag line used to stabilize the loads as they are uplifted. The sizes and locations of these cleared areas should be agreed upon beforehand with the erector. A typical layout is shown in Fig. 9.

Electrical Power
Power for operating the temporary aircraft warning lights must be available before erection begins.

Assembly

The usual procedure for erecting a guyed tower is to assemble the individual sections on the ground and then lift them one at a time as an assembled unit. For a self-supporting tower, the wider sections near the bottom of the tower are often assembled in the air as the tower is constructed.

Assembly of the tower sections should be done on a level bed to ensure that they will be straight and not racked or twisted. Bolts must be properly tightened and have a locking device. For high strength galvanized bolts, tightening by the "turn-of-the-nut" method is preferable to using a calibrated torque wrench.

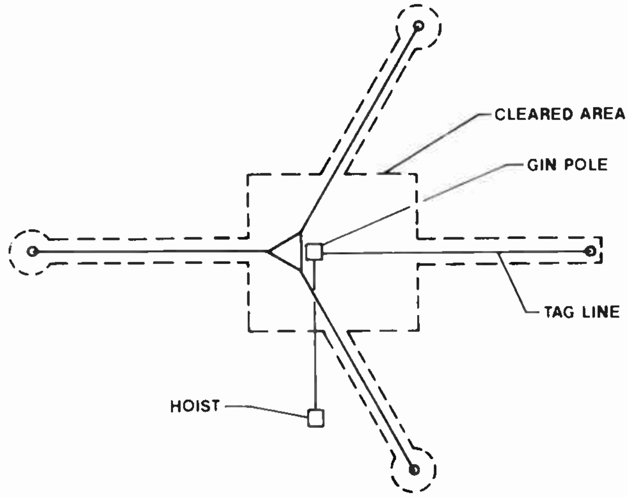


Figure 9. Typical layout for guyed tower.

Stacking

For a guyed tower, the first group of three to six sections are often joined together on the ground and then lifted into place using a crane. This portion of the tower is then guyed with temporary cables, and the remaining sections are erected one at a time using a vertical boom or "gin pole." This boom is moved or "jumped," up the tower as each section is installed. This arrangement is shown in Fig. 10. Temporary guys to stabilize the tower should be used when instructed by the designer.

For a self-supporting tower, a crane is often used to lift as many of the tower sections as possible, after

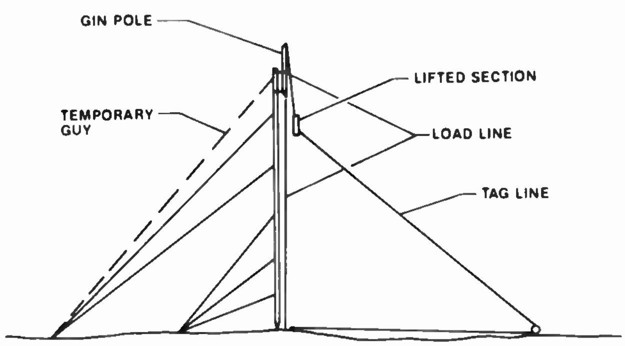


Figure 10. Typical erection setup for guyed tower.

which a gin pole is installed and used for the upper sections beyond the crane's reach.

Temporary aircraft warning lights must be installed at the top of the construction at the end of each day.

The tower should be grounded as soon as the first section is in place.

Guy Installation

When the tower reaches a guy attachment level, the cables at that level are installed. The guys in all three directions should be pulled out simultaneously to prevent any large unbalanced loads on the tower.

The tower should be checked for plumbness as each set of guys is installed and tensioned. Maintaining a plumb tower during erection eliminates the need for time-consuming adjustments later. Final tensioning of the guy cables and a plumbness check are done after the entire tower is erected.

REINFORCEMENT AND MODIFICATION OF EXISTING TOWERS

When equipment is replaced, relocated, or added to a tower it is often necessary to reinforce or replace existing structural components. The details and specifications for this work must be developed by the structural engineer who analyzes the tower.

Leg members may be strengthened by installing additional bracing or by field-welding additional material to them. Specific procedures in accordance with The American Welding Society *Structural Welding Code* must be provided and followed if field welding is required.

Bracing members and guy cables may be replaced with stronger components. Careful attention must be given to the connections for the new components to ensure their compatibility with the existing tower, as well as providing the required strength for the new components. When replacing components it is essential that temporary bracing or guy cables be installed before removing any existing component, and that it remain in place until the new component has been installed.

It is necessary that damage to the protective finish on existing members due to field welding or reaming be repaired. If required, the affected areas must be painted for aviation obstruction marking.

Foundation and guy anchors are the most difficult components to strengthen, and they may prove to be the limiting factor in determining a tower's capacity. The nature and feasibility of strengthening these components depend on the specific soil conditions.

INSPECTION AND MAINTENANCE PROCEDURES

To ensure trouble-free performance of a tower and its appurtenances, it is desirable to have a regular inspection and maintenance program. Portions of the program can be done by station personnel while others require experienced tower personnel.

Safety precautions should be observed at all times when working on or around the tower. If the tower itself is energized or if a high intensity RF field exists from antennas mounted on it, no work should be done on the tower without clearing it with the station engineer. When climbing the tower, safety belts and climbing devices should always be used. Automatic safety features on elevators should never be by-passed to save time. It is a good idea to never work alone. Failure to observe proper safety measures can result in serious injury or death.

Tower Structure

Damaged or Deformed Member

A visual inspection should be made of the entire tower structure to determine if any of the members have been deformed or damaged. Any bowed or kinked member should be noted as to type, location in tower, and nature and magnitude of deformation or damage. This information should be reported to the tower designer for evaluation and recommended action.

Condition of Paint

A visual inspection should be made of the entire tower structure to determine the condition of the paint. If the painting of the tower is for aircraft observation marking only, and not for corrosion protection, it is necessary only to note any general deterioration rather than small blemishes and scratches. If repainting is necessary, it is important to properly prepare all surfaces and select paints that are compatible with the existing finish.

Corrosion

Small scratches in the galvanized surface are not detrimental as the exposed surfaces will be protected by cathodic action of the adjacent zinc. If corrosion is observed, the source should be determined and noted. The affected areas should be wire-brushed clean to bare metal then painted with a zinc rich prime coat and, if necessary, a finish enamel coat of the appropriate color.

Connections

All bolts should be checked for tightness. Any loose bolts should be tightened in accordance with the original installation instructions.

Alignment

The tower structure should be checked for alignment using an engineer's transit. This check should be done only on a calm day, i.e., with wind velocity less than 10 mph, and in conjunction with measuring the guy tensions (described later).

Both plumbness and twist of a tower can be calculated from the measured horizontal deviations of each tower leg member from true vertical. Thus three transit set-ups (one on each leg azimuth) are required for a triangular tower, and four for a square tower. When the transit has been properly leveled, set the vertical cross-hair on the edge of the vertical leg at the tower base and lock the instrument in this position. By moving the telescope upward, it is then possible to observe the straightness of the leg over its entire height. The magnitude of misalignment can be accurately estimated by comparison with the tower leg diameter. A record should be made of the observations of each leg at each guy level.

Tolerances for plumbness and straightness should be as provided by the designer. EIA-222-E gives a plumbness tolerance that limits the horizontal distance between the vertical centerlines at any two elevations to 0.25% of the vertical distance between the two elevations. This should never be exceeded. A good rule of thumb in the absence of other data is to keep the tower plumb and straight within the diameter of the leg members. EIA-222-E gives a twist tolerance of 0.5° in any ten feet and a total twist limit of 5°.

If straightening of the tower is required, it should be done by adjusting the guy wires as described later.

When checking the plumbness of top mounted poles and pylon antennas, the effects of direct sunlight on them must be considered. It is best to make these checks early in the morning or on a cloudy day.

Guys and Guy Insulators

Inspection of the guys can be done visually only for those portions adjacent to the anchors and tower. The range of this visual inspection can be extended by using binoculars, but its reliability is limited. If experienced riggers are available, it is possible to ride down the guy on a bosun's chair, but this method should be used only under the supervision of qualified personnel.

Damaged components. A visual inspection should be made of the guy cables, insulators, and hardware. Cables and dead end grips should be checked for nicks or cuts in the individual strands. All porcelain insulators should be checked for chips, cracks, and oil leaks where appropriate. Fiberglass rods should be checked for surface tracking (black carbon track marks on surface of the rod), breakdown of the epoxy surface, and exposure of the individual glass strands. The manufacturers should be consulted with regard to corrective action.

Corrosion. If the guy cables show signs of corrosion, consideration should be given to coating or replacing them. The cost of cleaning and coating the cables should be considered along with the life expectancy of the coating when comparing it to the cost of replacement. All guy hardware should be checked using the same procedures for inspection and corrective actions as previously described for the tower structure.

Connections. All pins should be checked for tightness and the condition of the cotter keys. Dead end grips should be checked to ensure that their ends are completely snapped close, preventing any ice from forming inside. The surface appearance of the guy strand immediately next to the connections should be noted for evidence of slippage. Threads should be given a light petroleum coating.

Tensions. Guy tensions should be checked in conjunction with the tower alignment. These tensions should be measured at the anchor end and compared to the specified values. It is important to remember that they are dependent upon the ambient temperature.

For the usual guy arrangement with cables in three directions, it is necessary to measure the tensions in only one direction while keeping the tower plumb in all directions. For guy arrangements with cables in four or more directions, it is necessary to measure the tensions in only one of the two guys in the same vertical plane while keeping the tower plumb in that plane.

There are several methods of measuring guy tension with varying degrees of accuracy. For small guys up to ¾ inch, a shunt dynamometer calibrated for the size and type of strand is often used.

For larger guys, a series dynamometer may be placed in a temporary line between the anchor and a clamp on the cable. This line is then tightened until the permanent connection is relieved, and the tension is indicated on the dynamometer. Hydraulic jacks with a calibrated pressure gauge or load cells can be used in place of the temporary line and dynamometer. These are particularly effective for large guys attached with bridge sockets.

There are two indirect methods of measuring tensions in guys that do not have any large insulators or other loads in them. The intercept method consists of sighting along a straight bar attached at the bottom of the guy and measuring the vertical distance between the point where the line-of-sight intercepts the tower and the point where the guy is attached. This distance can be accurately estimated by counting the number of bracing panels. The tension in the guy is directly related to this intercept distance, the weight of the guy, and its length and slope.

The tension in a guy cable is also directly related to its length, weight, and natural frequency of free vibration. The natural frequency can be determined by putting the guy in motion with your hand and measuring the fundamental period with a stop watch. It should be noted that because a guy slopes, the tension on it varies along its length, and this method will only provide the average tension and not the tension at the anchor point. For long cables, this difference can be significant.

All tension measurements should be recorded along with temperature and wind speed and direction. If any substantial changes are noted from the values previously measured, careful checks for slippage of all connections should be made.

Tolerances for guy tensions should be as provided by the designer. In the absence of any other tolerance, tensions should be within plus or minus five percent of the specified values.

Any necessary adjustments in tensions can be made by adjusting the turnbuckle or bridge socket at the anchor. Make such adjustments slowly and carefully.

Never leave less than three threads sticking through the turnbuckle body or nut on the socket U-bolt. Remember that the tower must be kept plumb.

Base Insulator

The porcelain surface should be wiped clean with a soft cloth to remove accumulated dirt. A check should be made for cracks or chips on the porcelain surface. Scratches are often mistaken for cracks. Oil-filled insulators will display a wet surface or leak if cracked. If an oil stain or leak appears at the bottom of the porcelain on an oil-filled insulator and a crack cannot be found, incorrect loading possibly due to settlement of the pier should be suspected. A cracked base insulator should be replaced as soon as practical. Any sign of corrosion in the upper and lower bearing plates, rain shield, or lightning gap should be noted and corrected in a manner similar to that described for the tower structure. The lightning gap should be adjusted in accordance with instructions from the station engineer.

Tower Base and Guy Anchors

The tower base and guy anchors above grade should be visually inspected for spalling and cracking of the concrete. The soil surrounding the tower base foundation should be inspected for evidence of settlement. The anchor arms and surrounding soil should be examined for evidence of movement of the anchor. Any such settlement or movement should be noted.

Steel anchor shafts exposed directly to the soil should be inspected below grade for evidence of galvanic or electrolytic corrosion, especially in areas of high ground conductivity. Extreme caution should be exercised when excavating and backfilling during this inspection to ensure that the anchor's effectiveness is maintained.

Appurtenances

Ladder and safety device. The ladder and its connections should be checked for corrosion and tightness along with the tower. The sleeve and belt of the safety device should be visually examined and tested near the ground level before each use.

Elevator System. Inspection and maintenance of the elevator system should be in accordance with the manufacturer's instructions. It is a good practice to operate the elevator at least once a month.

Lighting System. Inspection and maintenance of the lighting system should be in accordance with the manufacturer's instructions. Checks for corrosion in the conduit, junction boxes, and light fixtures should be made along with the tower inspection. Any obstructions in the breather or drain in the conduit should be removed. Broken or cracked glass and any leaking gaskets should be replaced.

If the tower is equipped with an isolation transformer, its surface should be inspected for cracking and splitting. The surface should be painted with glyptal or a good quality enamel paint. Badly cracked surfaces should be recovered with cotton tape prior to painting.

Frequency of Inspection and Maintenance

A suggested schedule for inspections and maintenance performance is shown in Fig. 11.

Reports

A written report of each maintenance and inspection procedure performed should be made and filed with the station engineer.

SUGGESTED INSPECTION AND MAINTENANCE SCHEDULE

ITEM	Daily	Monthly	Before Each Use	Annually	After a major wind or ice storm	Manufacturer's Recommendation
Tower Structure:					•	
Damaged or deformed members				•		
Condition of paint				•		
Corrosion				•		
Connections				•		
Alignment				•	•	
Guys and Insulators:						
Damaged components				•	•	
Corrosion				•		
Connections				•	•	
Tensions				•	•	
Base Insulator				•		
Tower Base and Guy Anchors				•	•	
Ladder Safety Device			•			
Elevator System						•
Operate		•				
Lighting System						•
Lamp Failure	•					
Conduit Systems, fixtures				•		

Figure 11. Suggested inspection and maintenance schedule.

2.2 Lightning Protection for Broadcast Facilities

Frank C. Breeze,
Tilden, Lobnitz & Cooper, Inc., Orlando, Florida

INTRODUCTION

While there is still debate among atmospheric physicists regarding the exact mechanism of charge separation in thunderstorms, much is known about lightning itself. Lightning currents and radiated fields have been carefully studied for many years, leading to a better understanding of how to successfully protect against damage and injury. Manufacturers have used this understanding to develop a wide variety of lightning protection and surge suppression products. Codes and practices are also continually being revised to reflect increasing levels of understanding.

This chapter takes a very practical look at measures which will help the broadcaster prevent damage to equipment and disruption of service in the presence of lightning. The focus of this chapter has changed significantly from the last edition. Discussion on lightning physics, structural protection and designing surge suppression devices from a component level has been minimized, shifting the focus to practical measures which will help equipment and systems remain alive and well in a moderate to high lightning environment.

THE NATURE OF LIGHTNING

Thunderstorm cells are formed when warm moist air rises into cooler regions of the atmosphere, driven either by convective or frontal activity. Vertically circulating air currents are established which, through freezing and thawing of moisture particles, are believed to cause a separation of charge between upper and lower regions of the cell. (See Fig. 1.) A typical thunderstorm is comprised of many such cells which form and die away, each with an average life span of about twenty minutes.

The abundance of electrons accumulating in the lower regions of a cell creates a difference in potential with the less negative regions in the upper cell area. These electrons also repel free electrons in the surface of the earth, again resulting in a potential difference between cloud and earth. This charging process continues until the dielectric property of the air between charged bodies is exceeded.

A lightning discharge may occur within a thunderstorm cell, between cells, or from a cell to the earth below. Intense ionization of the air occurs between the charged bodies, followed by a faintly visible stepped leader which forks and branches from the negative body toward the positive. In the case of a cloud-to-ground discharge, the approach of this leader triggers an upward streamer, usually emanating from a tall grounded object. A complete circuit is formed when the streamer connects with the stepped leader, resulting in an avalanche of current known as the first return

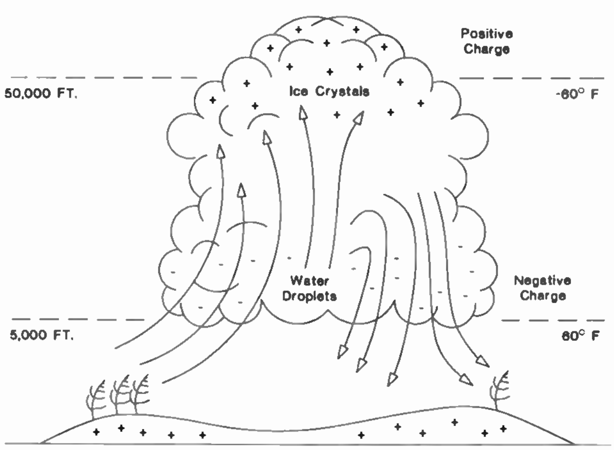


Figure 1. This thundercloud has reached its mature stage and a downdraft has developed on the cloud's lee side, producing a sudden drop in temperature, wind gusts, rain and perhaps hail, and lightning flashes.

stroke. Additional or subsequent return strokes may occur in rapid succession as depleted charge is replenished from surrounding areas of the cloud and earth.

Fig. 2 describes the annual frequency of thunderstorm days in various regions of the United States. A thunderstorm day is simply a day on which thunder is heard. Within the broad contours shown on the map there are often microclimates created by local terrain which may exceed or fall short of the average for the region.

Coarse methods of risk calculation[2] utilize isoceraunic data with the latitude and attractive area of a structure to establish the approximate frequency of strikes. Numerous lightning position and tracking systems are also in operation throughout the U.S. which record both the time and position of flashes to ground.

The strength of a lightning discharge is a function of the amount of stored charge, the quality of the channel through which the discharge occurs and the ability of the charged media to efficiently deliver stored charge to the discharge channel. Fig. 3 indicates the relative intensity of lightning discharges as a function of their frequency of occurrence. Approximately 50% of all lightning strikes deliver first return stroke current of 20,000 amperes or below. Only 2% reach levels of 140,000 amperes and larger strokes are even more rare.

Subsequent return stroke currents are always lower than the initial stroke as there is less charge remaining between the two bodies after the initial discharge. The importance of understanding the intensity of lightning current quickly becomes apparent when you consider that, by Ohms law, even an average 20,000 ampere lightning current flowing into a ten-ohm grounding system will cause the system to rise to a potential of about 200,000 volts above normal.

The radiated field from a lightning channel can pose a significant hazard to equipment and systems, particularly those which are interconnected by long lengths of cable. Fig. 4 is a composite of lightning electromagnetic field measurements made by a number

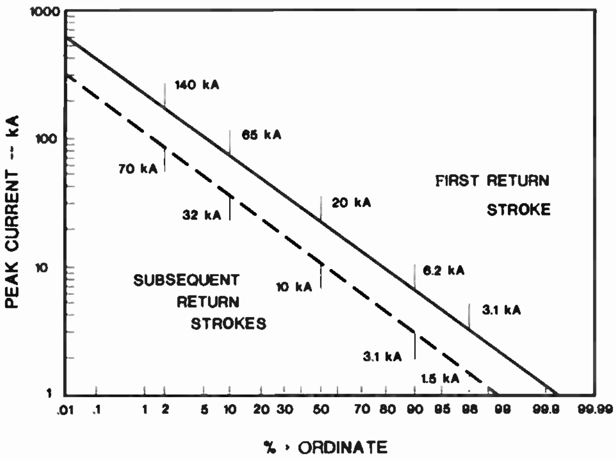


Figure 3. Lightning stroke intensity. Courtesy SRI International (Stanford Research Institute), Menlo Park, California 94025.

of researchers which have been normalized to a distance of ten kilometers. The figure shows a frequency domain distribution which peaks at about ten kilohertz at an intensity of slightly more than one volt per meter. It is important to realize that nearer strikes can create field strengths many orders of magnitude higher than those shown. The predominant low frequency component is also very effective in coupling energy into systems of wiring, even when buried in the ground.

The time domain current waveform associated with a typical lightning strike is characterized by a very fast leading edge or risetime, followed by a more gradual decay. Technically, risetime is the period of time required for the wave to increase from 10 to 90% of its crest value. Decay time is normally expressed as the time measured between the wave crest and 50% of the crest value. A description of a waveform such as 1.5 × 50 microseconds would indicate a single

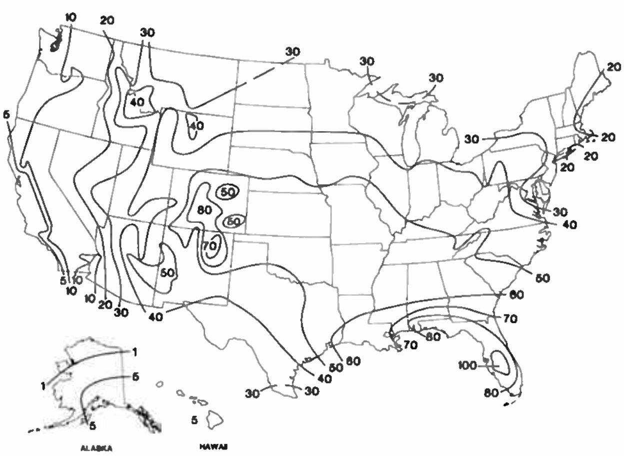


Figure 2. Average annual frequency of thunderstorm days in the United States.

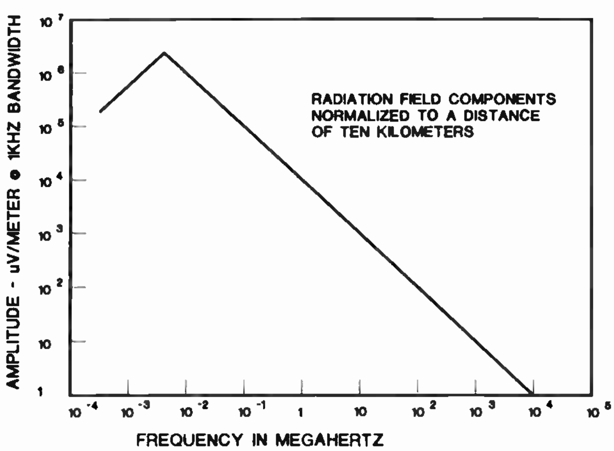


Figure 4. Lightning signal amplitude vs frequency.

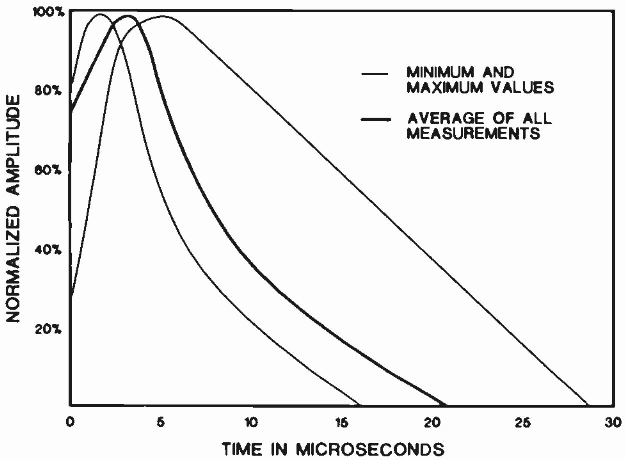


Figure 5. Lightning waveforms.

impulse waveform with a risetime of 1.5 microseconds and a decay time of 50 microseconds.

There are many variations on actual lightning waveforms seen in real world circuits. Waveforms such as those shown in Fig. 5 may be found with risetimes measured in a fraction of a microsecond near the point of lightning entry to a circuit. As a wave propagates through a wiring system, the risetime and the decay time will lengthen. The polarity of the impulse may be either positive or negative. Inductive and capacitive properties of the wiring system may cause the circuit to act as a resonant tuned circuit, producing a ringing wave which alternates in polarity. References which categorize the waveforms and current levels for several types of circuits are listed at the end of this chapter.[5,6]

PROTECTION OBJECTIVES

Lightning protection objectives may be grouped into three basic categories:

Structural Protection

Providing a means to intercept a lightning discharge, conduct the lightning current safely through or around the structure and dissipate the current into the earth. Lightning current passing uncontrolled through a structure may result in ignition of combustible materials, generation of explosive forces in masonry and other moisture bearing materials and burning or tearing of roofing systems. Secondary flashing between the primary current path and nearby grounded objects may also pose a threat to persons or elements of the structure.

Personnel Protection

Protecting personnel from the threat of a direct lightning strike and secondary flashing (sideflash), and controlling differences in potential between different parts of their bodies during a lightning event. Step potentials are voltage gradients seen along the surface of the earth as a lightning current radiates hemispherically from its point of entry into the soil. Touch potentials are voltage differences developed in a structure, natural object, or system during the passage of lightning current. Both step and touch potentials can be hazardous.

Equipment Protection

Controlling voltage potential differences between the various metallic circuitry leaving an item or system and between these circuits and the system's ground reference maybe "is key to protecting equipment." Controlling the potential differences to a value below the equipment damage threshold will insure the equipment survives. Providing tighter control to a value below the equipment upset threshold will help to insure the system rides through the lightning event without any noticeable effect.

STRUCTURAL PROTECTION STRATEGY

A traditional structural lightning protection system utilizes a system of air terminals to control the point of contact between a lightning discharge and the structure. Cables, or in the case of a tower or other steel framed structure, the structure itself is utilized to safely conduct the lightning current to a system of grounding electrodes. These electrodes, which may range from driven rods to radial wires, are utilized to pass the lightning current safely into the earth.

The basic premise behind traditional lightning protection design is based on the zone of protection created by any tall grounded metallic body. On conventional structures, a system of air terminals is used to create a network of overlapping protective zones. These zones insure that a lightning discharge contacts an air terminal rather than the structure. Towers, masts, and other conductive objects of sufficient mass and conductivity may be used as air terminals if properly bonded to the lightning protection system.

Fig. 6 shows traditional lightning protection for a studio facility with an adjacent microwave relay tower. Protective zones are based on an imaginary "rolling ball" 150 feet in radius being passed over the structure. The ball is also rolled around the structure tangent to earth. Air terminals are placed in such a way that the ball never contacts the structure or other objects requiring protection. Properly protected masts, light poles, and adjacent buildings may be taken into account when establishing zones of protection.

The protective zones created by the 150 foot radius ball have proven statistically adequate for most facilities and are the basis behind most codes and standards.[7,8,9] An interesting way to think about structural lightning protection design is to imagine rolling an inked ball of 150 foot radius around and over a structure. The ball should roll only on the earth, air terminals and other suitable metallic components which have been connected to the lightning protection system. Any other area which receives ink should be considered unprotected.

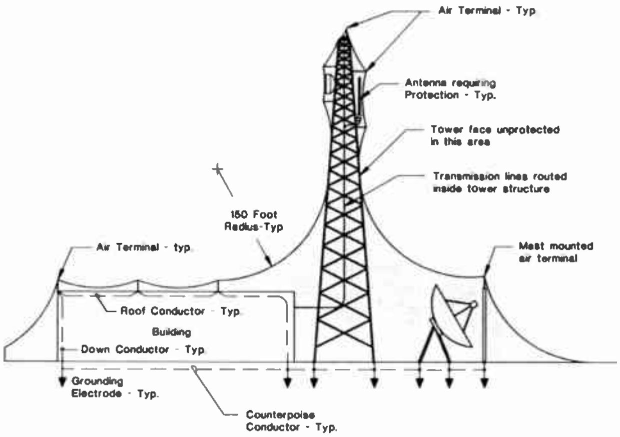


Figure 6. Traditional lightning protection for a studio facility with an adjacent microwave relay tower.

Most broadcast towers, and much of the equipment they support, have sufficient mass and conductivity to resist damage from direct lightning contact. Localized burning and pitting may occur in the immediate area of contact with the lightning channel, however, the brief duration of a lightning stroke causes little heating in areas away from the point of lightning contact. Fig. 6 shows how air terminals, properly bonded to the tower, may be used to protect more sensitive objects.

Whenever practical, transmission lines should be extended up the inside face of the tower to minimize the possibility of direct lightning contact. It is common practice to bond transmission lines to the tower at their highest and lowest points, forcing them to track the voltage gradient created in the tower by lightning. With the cable and tower at the same relative potential, there can be no secondary flashing between the tower and cable.

As a last note on the performance of towers, a passive grounded tower with continuous metallic guys will always stand out as the best performer when struck by lightning. The guys serve as additional lightning protection down conductors, significantly reducing the voltage gradient seen from top to bottom on the tower. Voltage gradients in the earth near the base of the tower (and around the station building) are also reduced as a large percentage of the lightning current is delivered to ground at the remote guy anchor points.

Self-supporting towers and AM broadcast towers with nonconductive guy segments will deliver the full lightning current to ground at their base, resulting in a larger voltage gradient in the earth near the base. AM towers normally use a discharge gap across a static discharge choke or tuned stub at the base insulator to provide a controlled point of flashover between the tower and its grounding system.

A comprehensive description of structural protection fundamentals is beyond the scope of this chapter. Readers desiring more information on structural protection are encouraged to obtain copies of the standards for structural protection listed at the end of this chapter.

PERSONNEL PROTECTION STRATEGY

The threat to personnel during a lightning strike ranges from the obvious danger of direct contact with a lightning stroke to the more obscure effects of step and touch voltages. Protection from a direct strike is accomplished with traditional structural protection methods, insuring that an adequate protective zone is provided in areas frequented by personnel.

Step and touch potentials are created as a lightning current passes through resistive soil and other available paths as it dissipates into the earth. A person in contact with only one point of the gradient will simply rise and fall in potential with the gradient without injury. A person in contact with multiple points on the earth or objects at different potentials along the gradient will become part of the current path and may sustain injury or death.

Fig. 7 indicates a number of methods for protecting personnel from the direct and secondary effects of lightning. A typical tower/transmitter site is used as an example. A technician responding to a service problem during a thunderstorm would likely exit his vehicle outside the gate, unlock and open the gate, and move his vehicle into the inside yard. The technician would then leave the vehicle, and enter the building.

The threat of a direct lightning strike to the technician has been minimized by establishing a protective zone over the areas to be traversed. This zone is created by the tower, air terminals on the roof of the transmitter building, and air terminals mounted atop light poles.

Step potentials are minimized through the use of a ground mat buried just below the surface of the area where the technician is expected to be outside the vehicle. Ground mats are commercially available, fabricated in a two-foot square pattern using #6AWG

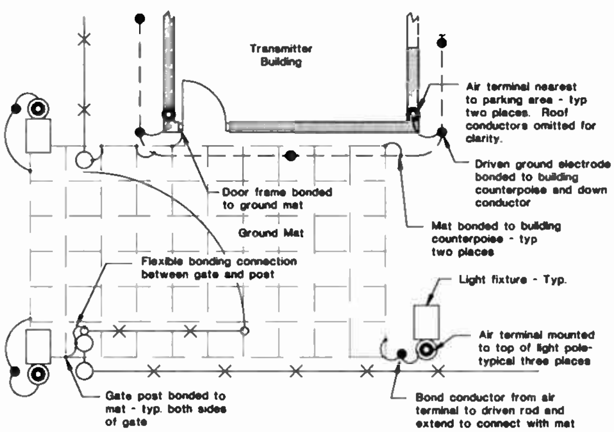


Figure 7. Personnel protection methods.

bare copper wire. Each intersection is welded creating, for all practical purposes, an equipotential plane which short-circuits the step potential gradient in the area above the mat. The mat as a whole will rise and fall in potential, however, there will be little difference in potential between the technician's feet. Mats should be covered with six inches of crushed stone or pavement.

The threat of dangerous touch potentials is minimized by bonding the ground mat to the fence at each side of the gate opening, bonding to the door frame of the transmitter building door, and providing a flexible bonding connection between the swing gate and its terminal post. Such bonding insures that the object being touched by the technician is at or near the same potential as his or her feet.

Bonding both sides of the gate opening to the mat helps to insure that the technician and both sides of the gate are at approximately the same potential while the gate is being handled. The flexible bond between the gate and its support post may be accomplished using a commercially available kit or by exothermically welding a short length of flexible 2/0 AWG welding cable between the two elements.

EQUIPMENT DAMAGE MECHANISMS

Most lightning damage to equipment occurs as the result of potential differences which exceed the tolerance level of the equipment. These potential differences may be presented to the equipment or system through external metallic circuits as a conducted or induced transient voltage surge. They may also occur as the result of differences in ground potential at various items of equipment which are connected together to form a system. Fig. 8 describes these mechanisms in greater detail.

Common mode surges, which are also referred to as "longitudinal mode" in many documents, arrive at the equipment with approximately equal potential on both sides of a balanced pair or on a number of circuits simultaneously. These surges may be induced into a wiring system by nearby lightning, directly coupled into the circuit or even created by the action of an upstream suppression device as it clamps two or more conductors together. Common mode surges may enter equipment on power phases, signal, and other circuits. Damage is normally sustained due to potential differences between the affected circuit(s) and equipment chassis or other uninvolved circuits.

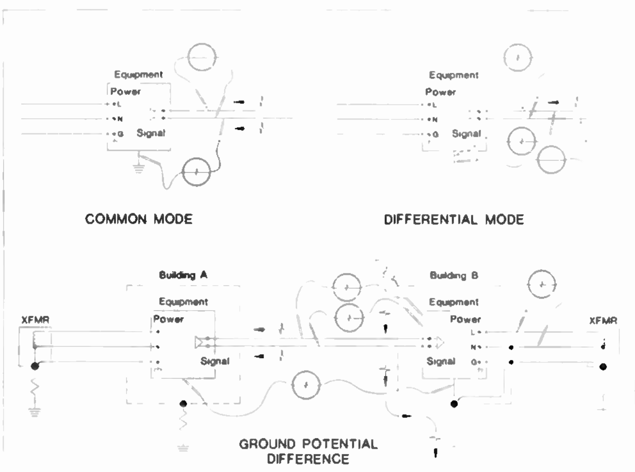


Figure 8. Equipment damage mechanisms.

Differential or normal mode surges are often more damaging than their common mode counterpart as most equipment is designed to operate in a differential fashion. In power circuits, a differential mode surge may appear on one or more phases relative to the neutral and ground conductors. In signal circuits, especially those operating on a balanced differential basis, the tolerance to differential mode surges is lower than for common mode. Common mode surges on electrical systems are routinely converted to differential mode at electrical services where one side of the service is referenced to ground. The same conversion process can occur on a balanced circuit when an upstream transient voltage surge suppressor (TVSS) device clamps one side of a pair before the other.

Ground differential damage is a bit more obscure than either of the other mechanisms, however, it is responsible for a great deal of damage to systems with equipment in multiple locations. Equipment in different buildings, or even equipment within different areas of the same building can be damaged through ground differentials.

In the simple example in Fig. 8, assume that building B receives a lightning strike of 20,000 amperes. Also assume the grounding system resistance is two ohms. As the lightning current flows into the earth through the two ohm grounding resistance, a 40,000 volt potential rise will be produced in the building B grounding system. Since the equipment in building B references the local building through its power cord and bonding conductor, its chassis will rise to about the same potential.

Circuitry within the building B equipment will attempt to track the building ground potential rise, except for the components which attach to wiring from building A. These components see a large difference in potential between the balance of their circuitry and the wiring to building A. Building A has not been involved in the ground potential rise and these circuits are still near ground potential.

Component breakdown occurs within the equipment in building B and a small fraction of the total lightning current attempts to find a path to ground through the wiring leaving for building A. Upon reaching the equipment at the remote building A, this current presents itself as a common mode surge causing damage to the equipment.

EQUIPMENT PROTECTION STRATEGY

Protecting equipment from the effects of lightning involves a combination of grounding, bonding, and surge suppression. Grounding provides a path to introduce lightning currents into the earth. Bonding serves to equalize lightning potential differences between various elements of equipment. Surge suppression limits differences in potential on active circuits which cannot be directly bonded.

The External Ground System

The effectiveness of a grounding system is a function of the type and extent of electrode system used, and resistivity of the surrounding soil. Soil resistivity is dependent on the quantity of free ions (chemical salts) in the soil, temperature, and moisture content. The character of the soil below a particular site may also vary significantly with depth and location due to layering of different types of soil, the presence of hardpan layers, and subsurface rock.

Temperature is a major concern in shallow grounding systems as it has a major effect on soil resistivity. During winter months the grounding system resistance may rise to unacceptable levels due to freezing of liquid water in the soil. The same shallow grounding system may also suffer from high resistance in the summer as moisture is evaporated from soil. It is wise to determine the natural frost line and moisture profile for an area before attempting design of a grounding system.

Fig. 9 describes a four-point method for in-place measurement of soil resistivity. Four uniformly spaced probes are placed in a linear arrangement and connected to a ground resistance test meter. An alternating current (at a frequency other than 60 hertz) is passed between the two most distant probes resulting in a potential difference between the center potential probes. The meter display in ohms of resistance may then be applied to the formula to determine the average soil resistivity in ohm-centimeters for the hemispherical area between the C1 and P2 probes.

Soil resistivity measurements should be repeated at a number of locations to establish a resistivity profile

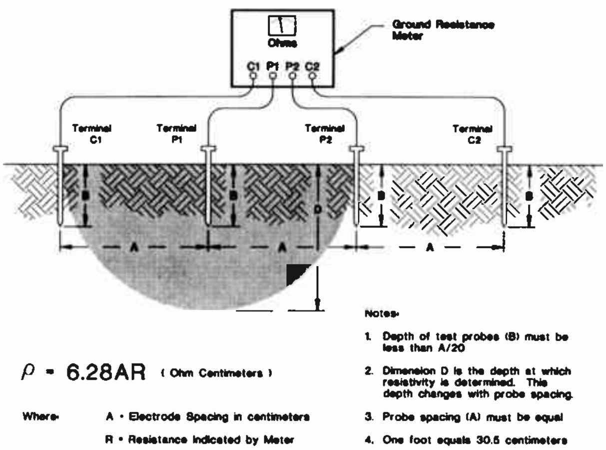


Figure 9. Four-point method for soil resistivity measurement.

for the site. The depth of measurement may be controlled by varying the spacing between the probes. In no case should the probe length exceed 20% of the spacing between probes.

Once the soil resistivity for a site is known, calculations can approximate the effectiveness of a variety of grounding system configurations. Fig. 10 includes equations for several driven rod and radial configurations which may be used for the purpose of estimating system resistance. Generally, driven rod systems are appropriate where soil resistivity continues to improve with depth or where temperature extremes indicate seasonal frozen or dry soil conditions.

Radials are also quite effective if placed below the frost line. They are often the only practical solution in areas with shallow subsurface rock. There have been instances at bald rock mountain top sites where radials were either grouted into saw cuts in the rock or simply pinned against the face of the rock.

The performance of a grounding electrode in high resistivity soil can often be improved through the addition of chemical salts. These salts leach into the soil, increasing the number of free ions with a proportionate decrease of soil resistivity in the area of the rod. Magnesium sulphate (epsom salts), copper sulphate (blue virol), calcium chloride, sodium chloride (table salt), and potassium nitrate have all been used for this purpose.

Fig. 11 describes the trench and well methods for applying chemical treatment. A typical precharged chemical ground rod installation is also shown. In a precharged rod, moisture from the air enters the rod through breather holes at the top of the rod and leaches through the chemicals inside, gradually exiting the rod through weep holes. As one might expect, chemically enriched grounds require recharging after a number of years to maintain their effectiveness. It is also wise to check with governing environmental agencies before introducing any foreign chemical into the soil.

The bentonite ground shown in Fig. 11 is a popular way of decreasing grounding system resistance of a rod electrode in uniform soil by up to 30%. Instead of driving the rod, it is placed in the center of a six to 12 inch augured hole. A slurry consisting of bentonite clay and water (well drillers mud) is then poured around the rod. As the water settles out, the resulting clay remains moist through absorption of moisture from the surrounding soil. Popular additives to the slurry include up to 75% powdered gypsum (calcium sulphate) and up to 5% sodium sulphate (galvanic anode backfill).

Ground Electrode Testing

Testing of grounding electrodes before they are connected to form a complex network is a fairly simple process. Fig. 12 describes the use of a ground resistance meter with the P1 potential electrode and C1 current electrode attached to the electrode under test. Most meters utilize a removable strap between these terminals so that only one test lead is required for the C1/P1 connection except during soil resistivity testing.

Hemisphere radius a	$R = \frac{\rho}{2\pi a}$
One ground rod length L, radius a	$R = \frac{\rho}{2\pi L}\left(\ln\frac{4L}{a} - 1\right)$
Two ground rods s > L; spacing s	$R = \frac{\rho}{4\pi L}\left(\ln\frac{4L}{a} - 1\right) + \frac{\rho}{4\pi s}\left(1 - \frac{L^2}{3s^2} + \frac{2L^4}{5s^4} \cdots\right)$
Two ground rods s < L; spacing s	$R = \frac{\rho}{4\pi L}\left(\ln\frac{4L}{a} + \ln\frac{4L}{s} - 2 + \frac{s}{2L} - \frac{s^2}{16L^2} + \frac{s^4}{512L^4} \cdots\right)$
Buried horizontal wire length 2L, depth s/2	$R = \frac{\rho}{4\pi L}\left(\ln\frac{4L}{a} + \ln\frac{4L}{s} - 2 + \frac{s}{2L} - \frac{s^2}{16L^2} + \frac{s^4}{512L^4} \cdots\right)$
Right-angle turn of wire length of arm L, depth s/2	$R = \frac{\rho}{4\pi L}\left(\ln\frac{2L}{a} + \ln\frac{2L}{s} - 0.2373 + 0.2146\frac{s}{L} + 0.1035\frac{s^2}{L^2} - 0.0424\frac{s^4}{L^4} \cdots\right)$
Three-point star length of arm L, depth s/2	$R = \frac{\rho}{6\pi L}\left(\ln\frac{2L}{a} + \ln\frac{2L}{s} + 1.071 - 0.209\frac{s}{L} + 0.238\frac{s^2}{L^2} - 0.054\frac{s^4}{L^4} \cdots\right)$
Four-point star length of arm L, depth s/2	$R = \frac{\rho}{8\pi L}\left(\ln\frac{2L}{a} + \ln\frac{2L}{s} + 2.912 - 1.071\frac{s}{L} + 0.645\frac{s^2}{L^2} - 0.145\frac{s^4}{L^4} \cdots\right)$
Six-point star length of arm L, depth s/2	$R = \frac{\rho}{12\pi L}\left(\ln\frac{2L}{a} + \ln\frac{2L}{s} + 6.851 - 3.128\frac{s}{L} + 1.758\frac{s^2}{L^2} - 0.490\frac{s^4}{L^4} \cdots\right)$
Eight-point star length of arm L, depth s/2	$R = \frac{\rho}{16\pi L}\left(\ln\frac{2L}{a} + \ln\frac{2L}{s} + 10.98 - 5.51\frac{s}{L} + 3.26\frac{s^2}{L^2} - 1.17\frac{s^4}{L^4} \cdots\right)$
Ring of Wire – diameter of ring D, diameter of wire d, depth s/2	$R = \frac{\rho}{2\pi^2 D}\left(\ln\frac{8D}{d} + \ln\frac{4D}{s}\right)$
Buried horizontal strip length 2/L, section α by b, depth s/2, b < a/8	$R = \frac{\rho}{4\pi L}\left(\ln\frac{4L}{a} + \frac{a^2 - \pi ab}{2(a+b)^2} + \ln\frac{4L}{s} - 1 + \frac{s}{2L} - \frac{s^2}{16L^2} + \frac{s^4}{512L^4} \cdots\right)$
Buried horizontal round plate, radius a, depth s/2	$R = \frac{\rho}{8a} + \frac{\rho}{4\pi s}\left(1 - \frac{7}{12}\frac{a^2}{s^2} + \frac{33}{40}\frac{a^4}{s^4} \cdots\right)$
Buried vertical round plate, radius a, depth s/2	$R = \frac{\rho}{8a} + \frac{\rho}{4\pi s}\left(1 + \frac{7}{24}\frac{a^2}{s^2} + \frac{99}{320}\frac{a^4}{s^4} \cdots\right)$

Notes:

1. Approximate formulas, including the effect of images
2. Dimensions must be in centimeters to return result in Ohms
3. ρ =resistivity of earth in Ohm–Centimeters
4. For 10 ft (3m) rods of 1/2" (12.7mm), 5/8" (15.88mm) and 3/4" (19.05mm) diameters, the grounding resistance may be quickly determined by dividing the soil resistivity ρ in ohm–centimeters by 292, 302 and 311 respectively.
5. Data source IEEE Green Book (Std 142–1982)

Figure 10. Formulas for calculation of resistance to ground.

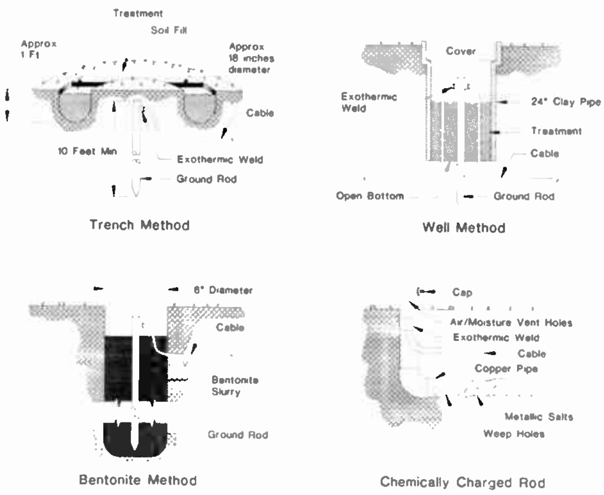


Figure 11. Trench and well methods for chemical treatment.

Probe C2 is inserted into the ground away from the electrode under test according to the instructions provided with the meter. The P2 probe is normally inserted along an imaginary line between the electrode and C2 probe at 62% of the distance to probe C2.

Testing is accomplished by passing an alternating current through the soil between the test electrode and C2 probe while reading the potential at the P2 probe. The ground resistance meter will evaluate these parameters, displaying the resistance of the electrode under test directly in ohms. If the C2 probe is placed too close to the electrode under test an invalid reading will result due to a shallow uniform voltage gradient between the C1 electrode and probe.

Testing of a completed grounding system is performed in essentially the same manner as an individual electrode. The logistics of performing the test can, however, become quite complicated due to the distance required between the probes.

The first step in the test process is to identify the spacing required between the nearest point of the grounding system and the C2 probe. See Fig. 13. A recommended starting point is five times the longest diagonal dimension of the grounding system. In a long linear site, the diagonal distance is usually apparent. For other sites, it is often easier to use the diameter of the smallest circle which completely encloses all of the components of the grounding system. The C2 probe should be placed in a direction away from all underground metallic utility lines and power lines as they may short-circuit the current path and invalidate the readings.

Once the C2 electrode is placed, resistance readings should be taken along the path between the ground system connection and C2 and plotted in terms of resistance versus distance. If there is adequate distance between the system under test and the C2 electrode, a flat will be present in the resistance curve. The correct resistance may then be read from the curve at 62% of the distance to probe C2. If no flat exists, the C2 probe is too close to the system under test.

On a new site it is often possible to perform ground system tests before the power company ground/neutral conductor is attached to the system. It is worthwhile to conduct a before and after test with probes in the same position to determine the influence of the power company attachment during future tests. It is also worthwhile to install permanent electrodes and marker monuments at the original P2 and C2 probe positions to insure the repeatability of future tests.

Grounding Resistance

The bottom line question in designing a grounding system is usually; "What value of grounding resistance do I need"? At electrical services the answer is dictated by the National Electric Code. If the first ground rod doesn't provide 25 ohms or less, drive a second rod. For towers, transmitter manufacturers will often have

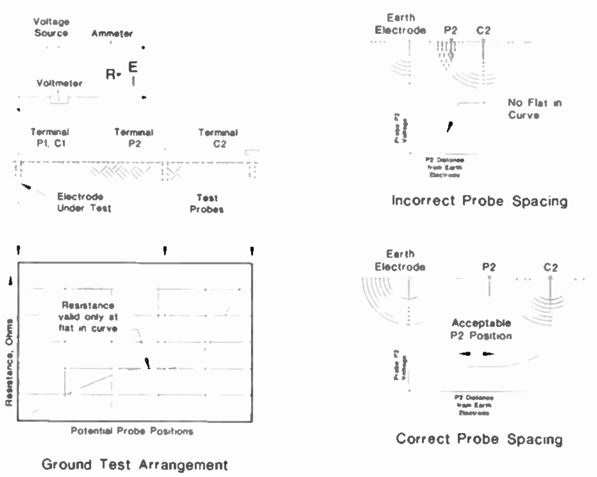


Figure 12. Ground electrode testing.

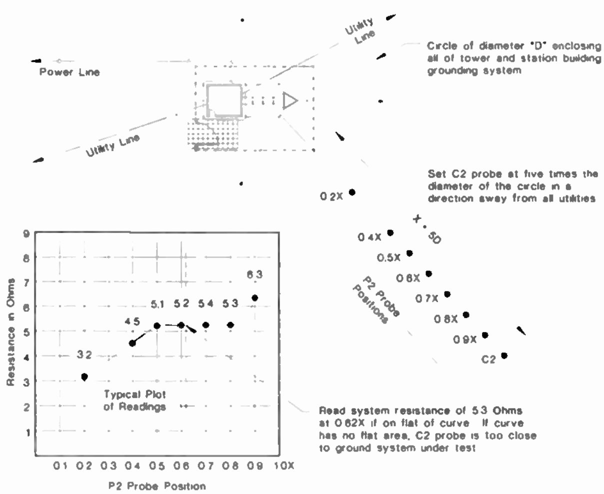


Figure 13. Ground system testing.

their own practices which dictate a specific value of grounding resistance. Lacking other guidance, the answer is to provide as good a ground as you can reasonably afford, focusing first on those locations subject to the greatest lightning currents, and to a lesser degree on areas of smaller exposure.

Tower-Building System

In a typical tower-transmitter building arrangement, the tower is normally subject to more frequent and larger lightning currents than the station building. It is therefore reasonable to place emphasis on the tower grounding system with less emphasis on that for the station building. Improved grounding at the tower will result in less current flowing between the tower and station building grounding systems, reducing potential differences between the two systems.

Fig. 14 describes a typical grounding configuration for a guyed tower and associated transmitter building. Tower grounding may be accomplished either by a system of driven electrodes between the tower base and guy points or by radial counterpoise conductors. In the driven electrode configuration, one radial counterpoise conductor is extended from the base of the tower to each guy point. This conductor serves as an attachment point for driven electrodes near the tower base, at guy locations and at intermediate points.

The rings shown feeding the system of radial conductors should stop within a few feet of the tower base. The complex mesh created by multiple bonds between the rings and radials will help to feed lightning current efficiently from the tower legs into each of the radial conductors. Apart from providing more copper in contact with the earth, there is no advantage in adding additional rings in the area between the tower base and guy points. Current flowing from the tower base out on the radials will produce approximately equal potentials between adjacent radials. With nearly equal potentials at both ends, additional bonding conductors

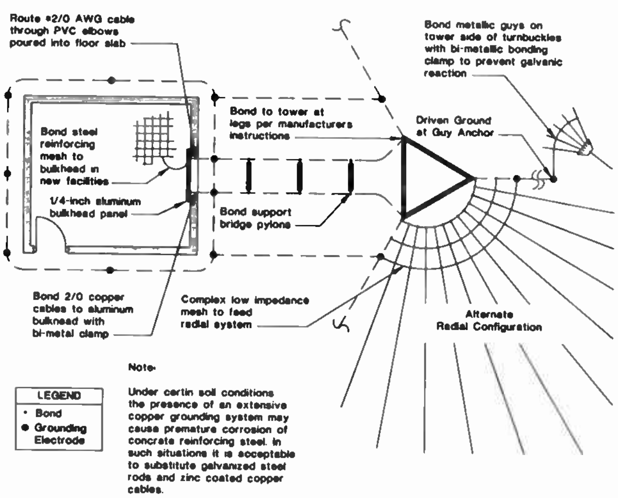


Figure 14. Typical tower ground system.

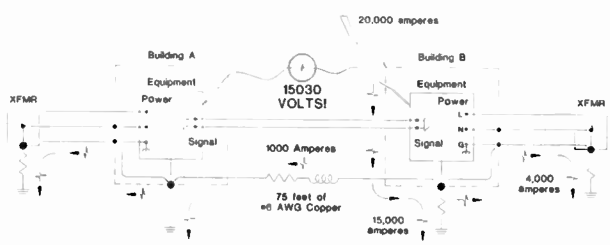
between the radials will carry little or no current. There is only a slight advantage in bonding between radials for currents introduced into the grounding system at guy points.

A ground ring is shown which encircles the station building helping to equalize potential differences within the building. The station building ring also serves as a connection point for driven electrodes, fencing and other objects which must be bonded (see Fig. 7). Four parallel bonding conductors are shown between the station building ground ring and the tower grounding system. These conductors reduce the level of current carried between the tower and station building by the transmission lines.

A $\frac{1}{4}$-inch aluminum (or copper) bulkhead plate is shown on the side wall of the station building which is bonded to the station grounding ring. This plate serves as a single point ground for all equipment within the station building. In new construction, the steel reinforcing mesh in the station building floor should be bonded together to the bulkhead panel to minimize potential differences between the equipment and floor during a lightning strike.

In the previous discussion, a great deal of attention was devoted to the subject of bonding as a means of equalizing potentials during a lightning strike. Fig. 15 repeats the earlier example of damage due to ground potential differences. This time, however, a bonding conductor is provided between the two grounding systems in an attempt to keep both at the same potential.

Kirchoff's law tells us that current will divide itself among all of the available parallel paths through a circuit in proportion to the impedance of each path. Lightning currents behave in the same way, flowing through all available paths to ground. In direct current circuits, the voltage produced across any circuit component is the product of current and resistance. When dealing with rapidly changing lightning current, inductance of the circuit plays a far larger role than simple resistance. Recall that an inductor tends to oppose any change in current until it has stabilized its magnetic field.



I = 1000 amperes
R = .03 ohms (from wire tables)
L = 22.5 microhenries (0.3uh/ft × 75 feet)
di = 1000 amperes (same as I)
dt = 1.5 microseconds risetime

$$E = IR + L\frac{di}{dt}$$

$$E = 1\text{x}10^{3} \times 3\text{x}10^{-2} + 22.5\text{x}10^{-6} \times \frac{1\text{x}10^{2}}{1.5\text{x}10^{-6}}$$

$$E = 30 + 15000$$

$$E = 15030 \text{ volts}$$

Figure 15. Example of ground potential difference.

The bonding conductor shown in Fig. 15, by virtue of its connection to ground at remote building A, serves as one of several paths for lightning current to follow on its way to ground. In this case, only 5% or 1,000 amperes of the lightning current flows through the 75 foot conductor with the remainder of the 20,000 amperes flowing into the grounding system of building B and its electrical service. Assuming a 1.5 microsecond risetime for the lightning current, the peak end-to-end voltage on the conductor is 15,030 volts. 15,000 volts of the total is the result of inductance in the conductor. The remaining 30 volts is the result of the conductor's resistance.

The normal reaction to lowering the potential difference in the previous example is to suggest a larger cable. After all, larger cables have less resistance and lower voltage drop. Changing to a larger cable, however, has little effect on the circuit inductance, affecting primarily the 30 volt portion of the total.

Fig. 16 provides a comparison of inductance values for a one-foot length of various sizes of round conductors and strip materials. Strip materials are considerably more effective for the same cross sectional area as a round conductor, making them more attractive as a bonding medium for lightning protection purposes. Another interesting property of strip material is that once you reach a width of about four inches, the major reason for thickness is mechanical strength and mounting convenience.

As a final note on the inductance of bonding materials, *never expect an insulated conductor in steel conduit to carry lightning current effectively*. Ferrite beads make a reasonably effective low pass filter when placed around an insulated conductor. Steel conduit around an insulated conductor creates the same effect, increasing the inductance of the cable within the conduit at least an order of magnitude. Where this condition exists (and there are many locations) a marginal compromise is to bond both ends of the cable to the conduit, permitting the conduit to serve as part of the circuit.

Single Point Grounding

If, after the exercise in bonding, you conclude there is no way of preventing potential differences in a conductor carrying lightning current, you are correct. There is, however, a method of preventing lightning current flow through a bonding circuit and with no current flow there can be no potential difference between the bonded items. This method is called single point grounding.

In Fig. 17 the equipment chassis and all metallic circuits leaving the equipment for the outside world have been bonded together and to a ground conductor at a single point. There is no possibility of a difference in potential between the circuits entering the equipment or between these circuits and the chassis as they are all bonded together. There is also no possibility of current flow from the single point ground into the equipment through any of the circuits as the equipment is isolated from the structure.

A surge entering on the power or signal lines cannot present itself to the equipment in differential mode as the lines are all connected together. A common mode surge arriving at the single point ground will pass harmlessly to ground through the grounding conductor and ground electrode resistance. There will be potential rise at the single point ground due to inductance and resistance in the grounding circuit, however, no current can flow through the equipment as it remains isolated from other points of ground reference. The equipment will simply rise and fall in potential, tracking the potential of the single point ground.

A lightning strike to the building or other structures connected to its grounding system will also cause the single point ground to rise and fall in potential. The equipment, however, sees no potential difference as its chassis and all external metallic circuits are tied "together at the single point ground. The only difficulty with the example in Fig. 17 is that nothing works.

Fig. 18 provides a more realistic approach to single point grounding. The only difference between Fig. 18 and the preceding example is that Transient Voltage Surge Suppression (TVSS) devices are used on the

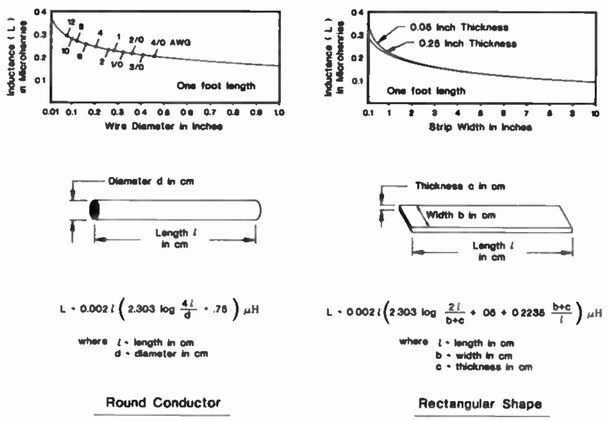


Figure 16. Inductance values for round conductors and strip materials.

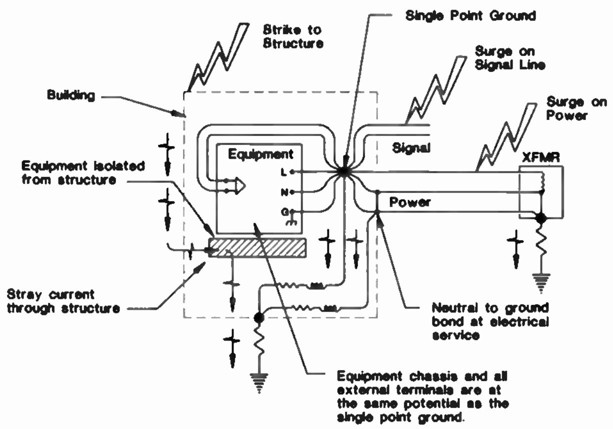


Figure 17. Single point grounding.

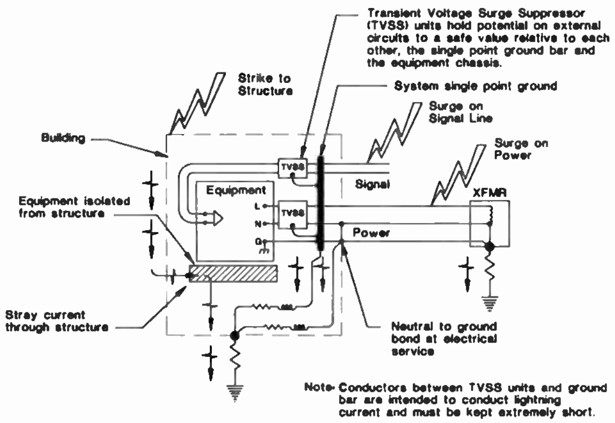


Figure 18. Single point grounding with transient voltage surge suppression units.

active circuits which, for obvious reasons, cannot be directly bonded to the single point ground.

For the purpose of this discussion, it is helpful to think of surge suppressors as a conditional bond, clamping or limiting the excursion of voltage on active circuits to a safe level relative to each other and to the single point ground. The single point ground may rise and fall in potential as the suppressors discharge current into it or during a strike to the building, but the difference in potential presented to the protected equipment is always held within safe limits.

It is worthwhile noting that TVSS devices will clamp in response to a rise in potential on their ground terminal as well as for legitimate transients on their active conductors. A strike to the building or nearby structure will cause a significant elevation in ground potential. The single point ground will rise in potential by virtue of its connection(s) to the building grounding system. The TVSS units, seeing their ground terminals rise in potential above their remotely connected active circuits, will clamp, forcing the active circuits to track the potential of the single point ground and the chassis of the protected equipment. Again, the voltage excursion seen by the equipment is held to a safe level and no damage is sustained.

Applications of Single Point Grounding

The application of single point grounding is normally limited to equipment within a room or a group of rooms. While it is possible to design larger configurations, the need to bring circuits in at different locations soon dictates the need for multiple locations, each treated as an island of equipment with its own suppression devices and single point of ground reference. Larger single point grounding systems are also more susceptible to induced voltages from nearby lightning by virtue of their increased cable lengths.

As an example of two extremes in scale, a computer room which serves terminals throughout a station complex may be engineered with a single point grounding system and proper surge suppression on its external circuits. The terminals and their printers, however, are scattered throughout the building, referencing ground at each location through their power cords. It is possible to designate the ground pin on the receptacle for each terminal/printer combination as the single point ground for the equipment at that location. A combination power and data suppressor may be provided for each location which insures that these conductors are held within safe limits of the receptacle ground pin and chassis of the equipment. The equipment is isolated from stray grounds by placement on a desk top.

Fig. 19 is an example of how single point grounding applies to a typical broadcast transmitter building. A ¼-inch bulkhead panel in the wall of the building serves as a single point ground reference for all equipment within the small facility. All coaxial cables, waveguides, and raceways from the tower are bonded to the bulkhead as they pass into the building.

The physical size of the electrical equipment dictates that it cannot be located directly at the bulkhead panel. To minimize the effect of bonding system inductance, a six-inch wide bonding bus is extended to each side of the bulkhead. Width of the strip provides the necessary low inductance. Its ¼-inch thickness, while not necessary for electrical reasons, provides the installer with a bus which may be drilled and tapped to accept short bonding pigtails to the equipment.

Support hangers for cables and raceways serving the protected equipment are isolated from the roof structure to prevent inadvertent current flow through the raceways. Isolation is provided between the equipment feet and floor slab by a polycarbonate plastic pad. Such isolation may not be necessary if adequate isolation is provided by the equipment feet. Nylon bolts may be used with conventional expansion anchors to secure equipment to the floor without violating the integrity of the single point grounding system.

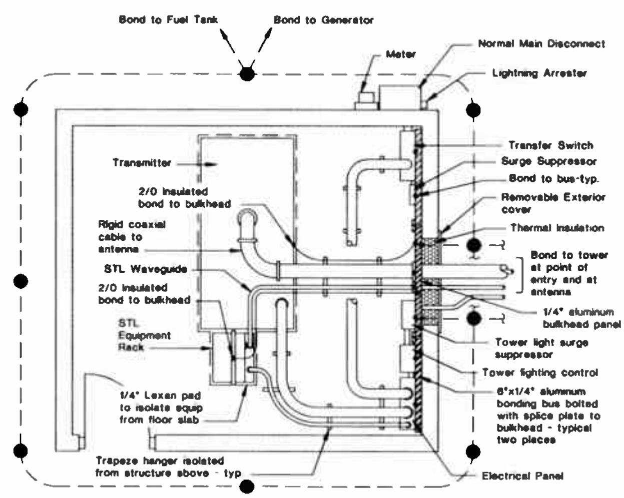


Figure 19. Single point grounding to a typical broadcast transmitter building.

Surge Suppression

Surge suppressors are shown at the transfer switch to protect the incoming power circuits and on the circuitry from tower lights. Both suppressors should be bonded to the grounding bus, keeping the length of their bonding lead as short as possible. Suppression should also be provided on all metallic communications lines and circuits serving lighting, winches and other electrical items outside the building. Again, locate these suppressors at the point of entry for the circuits and bond their ground leads to the bus with the shortest possible lead length.

The lightning arrester shown at the main disconnect outside the building is an inexpensive device, however, it serves an important purpose. Under normal operation, the main disconnect is closed and the transfer switch is connected to utility power. In this configuration, the lightning arrester and surge suppressor at the output of the transfer switch are in parallel. It is unlikely that the surge suppressor will allow the arrester to operate due to the difference in clamping level. During operation on emergency power, however, the scenario is quite different.

A lightning strike to the utility line will propagate along the line as a traveling wave in both directions from the point of lightning contact. Upon reaching the open circuit input of the transfer switch, the wave will reflect back on itself, effectively doubling its initial crest value. The same condition can occur at the main utility company disconnect if it should be open during servicing.

Flashover within electrical equipment is serious in itself, however, the problem is compounded when operating voltage is present. The flashover arc provides a low impedance path for 60 Hz fault current and significant damage to the equipment may occur. The arrester will help prevent this condition from occurring by limiting the traveling wave voltage.

Fig. 20 indicates a method of connecting a three-phase surge suppressor to an electrical panel. Many manufacturers are recommending connection of their products using a thermal circuit breaker rather than drilling and tapping the buses of the panel or attaching the suppressor to the main lugs. It is of great importance to minimize the length of the suppressor leads and that of the grounding conductor as their inductive voltage drop is additive with the initial clamping voltage of the suppressor.

The surge suppressor example in Fig. 20 indicates the use of a shunt wired device. There are, however, large surge suppression devices on the market which permit up to several thousand amperes of load current to flow through them. On these devices, only the length of the grounding lead between the suppressor and ground bus is critical.

Signal Line Suppressors

Fig. 21 shows a common surge suppressor configuration for telephone and signal line applications. Most suppressors of this type are constructed as multi-stage hybrid devices utilizing a high-energy first stage, a fast acting second stage, and an impedance in series between the two stages to coordinate their clamping behavior. Because of the multi-stage design, these devices must be installed in series with the protected circuits.

The treatment of shields is often an issue when dealing with signal line surge suppressors. Fig. 21 shows shields being bonded to the suppressor ground bus to force them to track the single point ground. If ground loop or other technical restrictions prevent direct bonding of the shields, they should be protected as any other active circuit.

The bonding lead distance for signal line suppressors is often more critical than for power devices. Tolerance levels of signal circuits are normally lower than for power supply inputs, and the relatively small voltages developed in the suppressor ground leads can become significant.

Signal line surge suppressors are sold in a variety of shapes and sizes for different applications. The

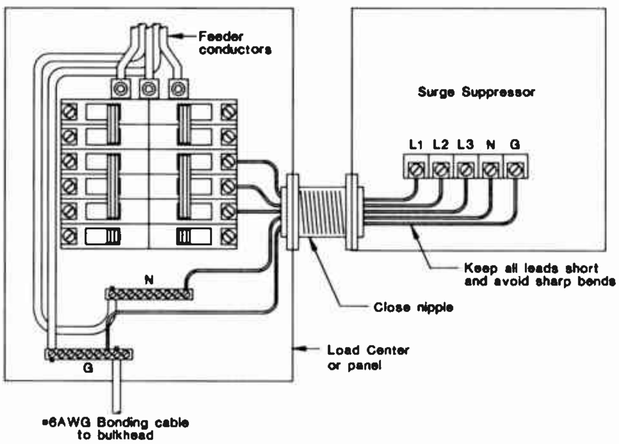


Figure 20. Three-phase surge suppressor.

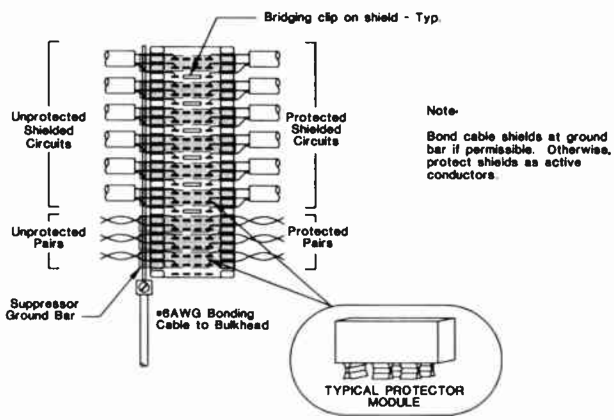


Figure 21. Signal line suppressor.

suppressors should generally clamp transient voltage on a circuit to within 150% of normal peak operating voltages and even lower in some applications. Since most signal line suppressors are inserted in series with the circuit being protected, it is wise to evaluate the effect of their series impedance on the operation of a circuit.

The effect of suppressor capacitance can be important in many high speed data, RF, and video applications. One simple way of evaluating the effect of this capacitance is to equate it to equivalent cable feet. For example, if the desired suppressor exhibits capacitance of 100 picofarads and the cable used in the circuit is rated at ten picofarads per foot. Will the circuit tolerate an additional ten feet of cable? If so, the suppressor capacitance should produce no noticeable effect on the circuit.

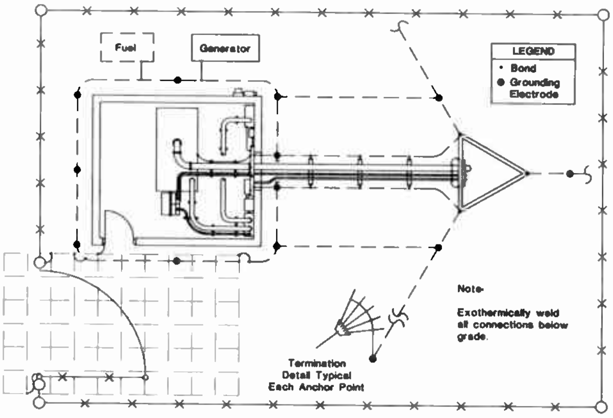


Figure 23. Typical bonding and grounding recommendations.

Isolated Ground Receptacle

Fig. 22 shows an isolated ground receptacle circuit commonly used in computer room grounding applications. The receptacles used in this type of circuit differ from the norm in that their ground sockets are electrically isolated from their mounting tabs. They are therefore isolated from their outlet box and structural ground at each receptacle location. A dual system of grounding conductors insures that equipment plugged into an isolated ground (IG) receptacle references ground first at the single point ground.

The use of isolated ground receptacles helps to insure that plug-in terminals, printers, diagnostic and other ancillary equipment is properly referenced to the single point ground and not the local structure. It only takes one item of equipment connected between the protected equipment and a remotely grounded receptacle to compromise the integrity of the grounding system.

Fig. 23 is a composite of the bonding and grounding recommendations for the typical transmitter site. While complicated in appearance, each component has its purpose as part of a simple to understand subsystem.

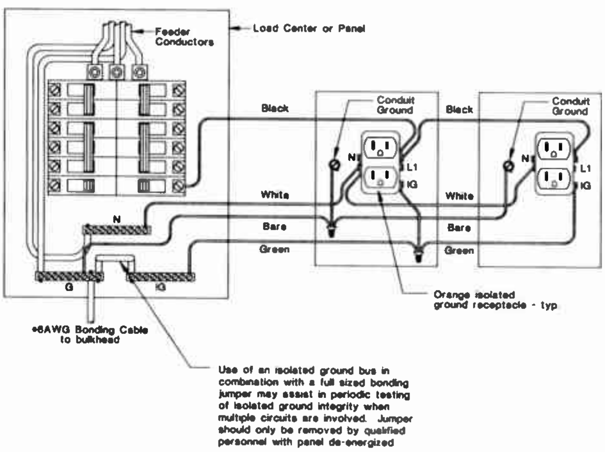


Figure 22. Isolated ground receptacle.

CONCLUSION

This chapter has been written in tutorial form as every site is different and no single set of recommendations will apply to every situation. The principles set forth, while tailored to a broadcast environment, apply equally to other systems. There are still a few mysteries to be solved in completely understanding lightning, but once it enters a wiring system it becomes an electrical current which is both predicable and understandable.

Many of the elements covered in this chapter have been touched on lightly to keep the chapter to a reasonable size. A list of recommended publications is included at the end of this chapter which deal with the subject material in greater depth.

REFERENCES

1. United States Isoceraunic contours from National Oceanographic and Atmospheric Administration maps.
2. Military Handbook 419A, "Grounding, Bonding & Shielding for Electronic Systems," December, 1987. (Available from the US Government Printing Office, Philadelphia, PA.)
3. "A Ground Lightning Environment," Cianos & Pierce, Courtesy SRI International (Stanford Research Institute), Menlo Park, CA.
4. "Lightning Protection Using Transzorbs," General Semiconductor Corp.
5. ANSI/IEEE C64.41–1980 (IEEE 587), "Guide for Surge Voltages in Low Voltage AC Power Circuits."
6. FCC Rules & Regulations, Chapter 68, section 3.02—"Lightning Protection for Telecommunications Equipment."

7. NFPA 78—"Lightning Protection Code," National Fire Protection Association, Quincy, MA.
8. UL–96A—"Standard for Lightning Protection Systems," Underwriters Laboratories, Northbrook, IL.
9. LPI–175—*Standard of Practice*, Lightning Protection Institute, Woodstock, IL.
10. IEEE 142 (Green Book) "Grounding Practices for Electrical Systems."
11. Portions adapted from Military Handbook 419A listed in reference 2.
12. Stromberg Carlson Practices, Section 53-001-03, "DCO-CS Grounding Engineering Considerations," 1989.

2.3 Transmission Lines

Harry D. McKee
Bruce J. Carlson
Robert D. Leonard
John A. Midkiff
Andrew Corporation, Orland Park, Illinois

INTRODUCTION

Transmission lines transmit or guide energy from one point to another. Usually, this needs to be done with maximum efficiency, keeping losses due to heat or radiation as small as possible.

This chapter discusses the three major types of transmission lines in common broadcast use.

1. Semi-flexible transmission lines with two concentric conductors (coaxial) transmitting in transverse electromagnetic (TEM) mode. In a TEM mode, both the electric (E) and magnetic (H) fields are entirely transverse to the direction of propagation (Fig. 1A).
2. Rigid transmission lines with concentric conductors (coaxial) which also are transmitting in TEM mode.
3. Transmission lines incorporating a single hollow conductor (waveguide) in which the E or the H (or both) fields always have a component in the direction of propagation. In Fig. 1B, for example, the E field in the rectangular waveguide shown is transverse to the direction of propagation, hence the "TE" designation for *transverse electric*. Note that the magnetic field has a component in the direction of propagation. All waveguides used at UHF television frequencies transmit in a transverse electric configuration.

General Selection Criteria

The three major criteria for selecting transmission line are:

1. Frequency band of operation
2. Power handling capability
3. Transmission efficiency (attenuation)

Table 1 shows the general characteristics of the most commonly used broadcast transmission line types together with an approximate rating of associated installation complexity, accessories and overall cost. Readers are cautioned that these are very general ratings and individual circumstances may cause wide variations.

Frequency

Frequency of operation in a transmission line is usually a function of the diameter. In semi-flexible

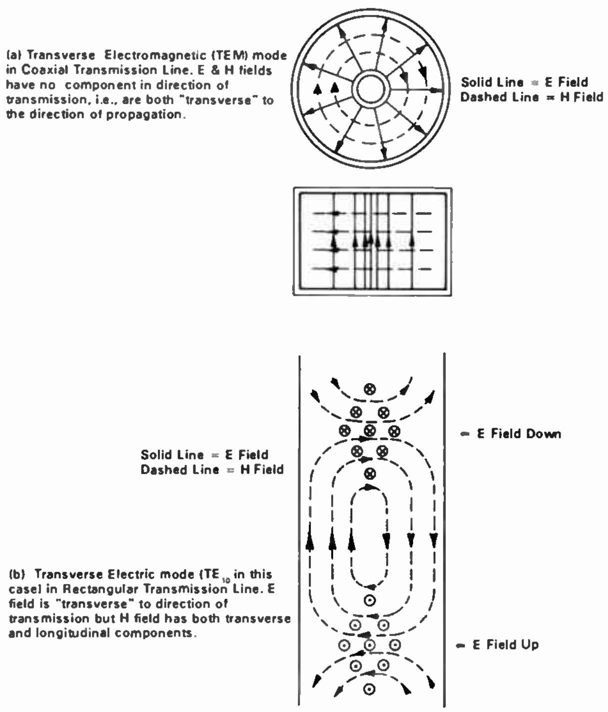


Figure 1. E & H Field configurations in coaxial line (A) and waveguide (B).

TABLE 1
Comparison chart for broadcast transmission lines.

Characteristic	Semi-Flexible Coaxial Cable	Rigid Coaxial Line	Waveguide
Typical Application	AM, HF, FM, LPTV	AM, HF, FM, LPTV, UHF-TV,	UHF-TV
Precise Layout Required	No	Yes	Yes
Channelized Lengths	No	Yes	Yes
Power Handling	Good	Better	Best
Attenuation	Good	Good/Lower	Lowest
Power Loading	Low/Med	Low/Med	Higher
Typical Hanger Type	SST Band	Spring Hgr.	Spring Hgr.
Typical Hanger Spacing	3–5 ft.	10 ft.	10–12 ft.
Pressurization Required	Foam—No Air—Yes	Yes	Yes
Recommended Maintenance	Annual Insp.	Annual Insp.	Annual Insp.
Initial Material Cost	Most Ecomomical	More Expensive	More Expensive
Initial Installation Cost	Lower	Higher	Higher

coaxial lines (Figs. 2A and 2B), outer conductor sizes range from ¼" to 9" diameter. In rigid line (Fig. 2C) typical diameters range from ⅞" to 9³⁄₁₆". In the AM, FM, and TV operating bands, the most commonly used sizes of semi-flexible or rigid coax range from 1⅝" to 8³⁄₁₆" diameter.

For rectangular waveguides used at UHF-TV frequencies (Fig. 2D), sizes range from 5.25" × 11.5" (WR1150) to 9.0" × 18.0" (WR1800). Circular and truncated waveguides (Fig. 2E and 2F) generally range from 13" to 18" nominal diameter.

Power Handling

Power handling in a transmission line is usually limited by the amount of heat generated as current flows on the skin of the conductors, caused by the resistivity of the conductor material. Heat generated in the inner conductor must be transferred by convection, radiation, and conduction to the outer conductor, then passed to the surrounding environment by convection and radiation.

In broadcast applications, the limiting factor is usually the amount of heat which the dielectric support material between the conductors can absorb before it begins to soften. If that happens, the inner conductor can change position, which causes a change in the characteristic impedance of the transmission line. This creates a voltage standing wave ratio (VSWR) mismatch within the transmission line, causing degraded performance due to higher reflected power.

On a size for size basis, corrugated semi-flexible cables, having a greater length of copper per unit of line length than rigid line, generate slightly higher I^2R (heat) losses per unit of length than rigid line.

Waveguide, having no inner conductor and no dielectric support material, has very high power handling capability. Consequently, it is more commonly used in very high power UHF applications.

Attenuation

Size for size, semi-flexible cable will have higher attenuation than rigid line since the corrugated construction technique contains more copper per unit length. In addition, more dielectric is needed to maintain concentricity of the inner conductor of semi-flexible transmission line. Besides absence of corrugations, the copper tubes used for rigid lines tend to support themselves, so less dielectric support material is required.

Waveguide, which eliminates the center conductor and associated supporting dielectric, has the lowest attenuation of the three types. Due to the low attenuation characteristics, it may become the medium of choice in areas where high operating costs (due to higher utility rates) must be minimized. However, the price for low attenuation is a significantly larger size, which adds substantial wind loading to the tower.

SEMI-FLEXIBLE CABLE

General

Semi-flexible transmission line has seen increasing use in the broadcast industry over the years because it is ideal for a wide variety of low and medium power applications. This line achieves its flexibility through the use of corrugated copper conductors. The term semi-flexible is used because in larger sizes, 1⅝" and

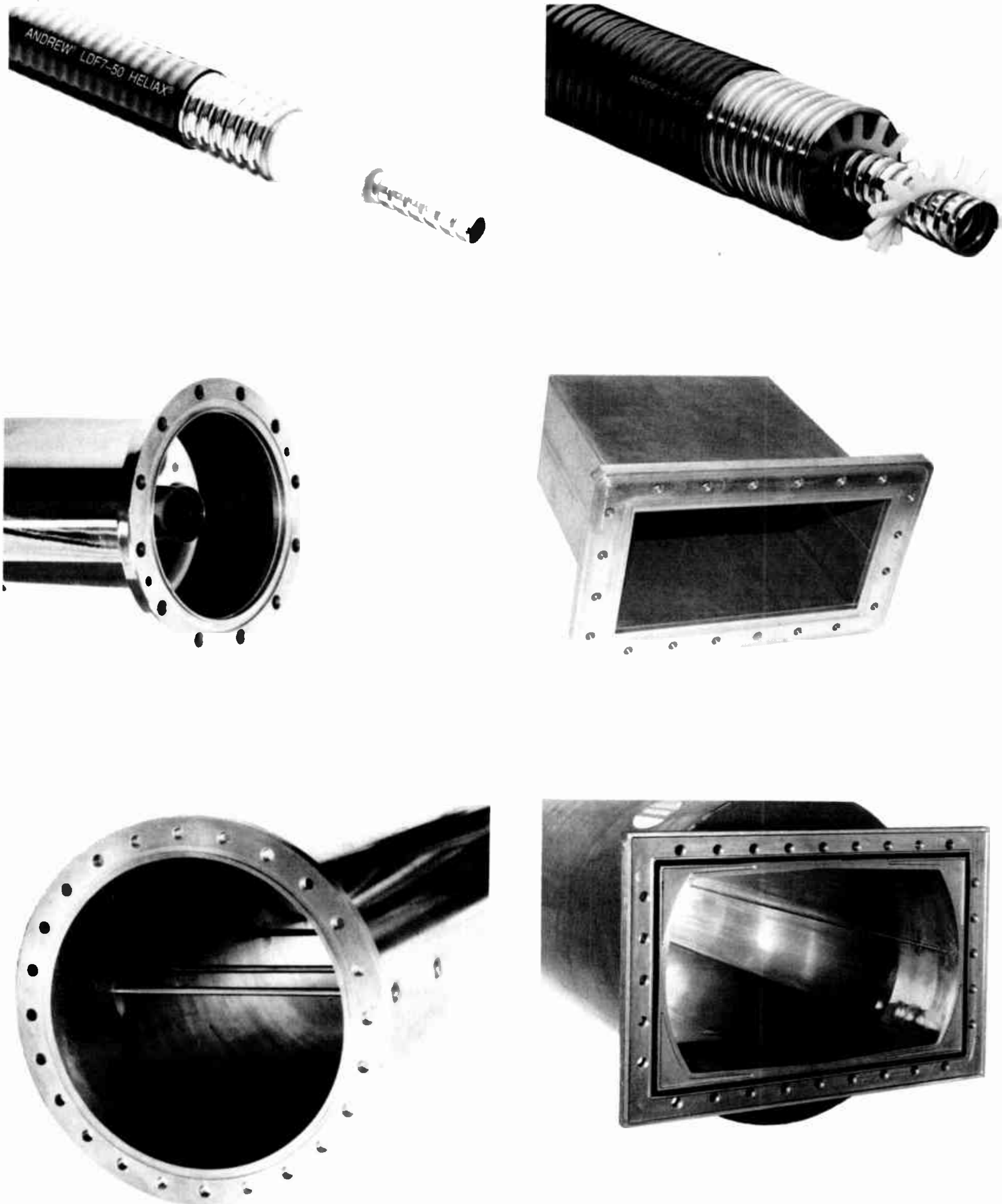


Figure 2. Various broadcast transmission lines.

TABLE 2

Typical minimum bend radius for semi-flexible cable (minimum of 10 bends).

Line Size	Typical Minimum Bend Radius
1⅝"	20"
2¼"	22"
3"	30"
3½"	30"
4", 4⅛"	40"
5"	50"
6⅛"	79"
8"	98"
9"	118"

up, the minimum bending radius becomes quite large (Table 2).

The advantages of semi-flexible cable are:

1. One-piece installation without the multiple connections of rigid line, since it is supplied in continuous lengths.
2. Easier to install than rigid line, since exact length is not critical.
3. High reliability through accommodation of thermal expansion and contraction, due to its corrugated construction.
4. Long life in corrosive environments or if buried, when provided with a protective polyethylene jacket.

Types of Semi-Flexible Cable

Semi-flexible cable is supplied in two types, foam dielectric and air dielectric.

Foam-Dielectric Cable

Foam-dielectric coaxial cables (Fig. 3) are designed for most antenna feeder systems which do not require a pressure path to the antenna. Typical applications include AM and FM radio and low-power television (LPTV). Sizes range from ¼" to 1⅝", with the smaller sizes being available in two versions: standard and superflexible. Superflexible cables, intended for installation in confined spaces, have deeper corrugations than standard foam cables, which permit them to be bent on a smaller radius and with the capability of more repeat bends. They are often used with earth station antenna systems and as jumper cables. Some smaller diameter cables are available with either 50 ohm or 75 ohm (nominal) impedance; 1¼" diameter and larger cables are furnished with 50 ohm impedance only. Foam-dielectric cables are constructed with a closed-cell, low density foam dielectric which prevents water penetration while providing low attenuation and high relative velocity of propagation.

Air-Dielectric Cable

Air-dielectric cables (Fig. 4) utilize a spiral polyethylene (or other polyolefin) spacer to separate the conductors. Compared to foam-dielectric cables, air-dielectric coaxial cables have slightly lower attenuation and higher average power rating, due to the superiority of air as a dielectric. They are pressurized in operation, and so are used to provide dry gas to pressurized antenna systems. When equipped with a pressurization alarm, damage to a pressurized cable can be detected immediately. Air dielectric cable is manufactured in sizes required for most broadcast applications.

A recent variant of air-dielectric cable is sectionalized semi-flexible coaxial cable in large diameter (up to 9"), supplied in 38-foot sections for transportation in standard shipping containers. This eliminates the delivery and installation problems associated with large diameter continuous coaxial cable, which must be shipped on immense reels. Sectionalized coaxial cable is available for HF, MF, and LF frequencies only. Large flanges connect the sections and provide good electrical contact and pressure tightness.

Coaxial Cable Design Considerations

Coaxial transmission line theory is presented in detail in Chipman[5] and Ramo[10]. The following discussion focuses on the information that is of the most practical value to broadcast engineers.

The basic principle in the design of a coaxial transmission line is to maximize the power handling capabilities while minimizing the attenuation for a given size line. This involves adjusting the inner conductor diameter to achieve an optimum value for each of the performance characteristics. The resulting ratio will determine the characteristic impedance of the line.

Impedance

The impedance is related to the dimensions of the inner and outer conductors and the dielectric constant

Figure 3. Typical foam-dielectric cable.

Figure 4. Typical air-dielectric coaxial cable.

of the dielectric material between them. It can be expressed by the following equation:

$$Z_C = \frac{60}{\sqrt{\epsilon'}} \ln \frac{D}{d}$$

where: Z_C = Characteristic impedance, ohms
ϵ' = Dielectric constant = relative permittivity of dielectric (ϵ' = 1.0 for air dielectric)
D = Inner electrical diameter of outer conductor, in
d = Outer electrical diameter of inner conductor, in

Power Handling

Power handling is limited by either of two factors: the maximum peak power, determined by electric field strength, or the maximum average power, determined by the allowed temperature rise of the inner conductor.

For a coaxial construction the electric field is at a maximum near the outer surface of the inner conductor. Maximum electric field strength can be calculated as follows:

$$E_{MAX} = \frac{.278}{d} \sqrt{\frac{P}{\ln \frac{D}{d}}}$$

where: E_{MAX} = Maximum electric field strength, volts/in
P = Power level of signal, watts

E_{MAX} is at a minimum when the ratio D/d is equal to 1.65, resulting in an impedance of 30 ohms for an air dielectric line.

Determining the average power limitation for a coaxial line involves complex thermal models. An approximate optimization is that the optimum ratio of D/d should equal 2.72, resulting in an impedance of 60 ohms.

Attenuation

Attenuation is due to dielectric losses and conductor losses. The loss of the dielectric material is directly proportional to frequency and also to the "loss factor" of the material. For the most commonly used dielectrics in RF transmission lines (i.e., Teflon®, polyethylene, air etc.), the loss factor is fairly small (of the order of 0.0002) and dielectric losses are small compared to conductor losses. Conductor loss is related to dimensions, permeability, and conductivity of the material. It varies with the square root of the frequency and is defined by the following equation for copper conductors:

$$\alpha = \frac{.433}{Z_C}\left(\frac{1}{D} + \frac{1}{d}\right)\sqrt{f}$$

where: α = Attenuation, db/100 ft
f = Frequency, MHz

As can be derived from this equation, attenuation is minimized when D/d is equal to 3.59, which results in an impedance of 77 ohms.

It is seen from this discussion that there is a trade-off in design between optimizing for line peak power, average power, and attenuation. In general, most lines are manufactured with an impedance of either 50 ohms or 75 ohms. The 50 ohm case provides a good balance between the 30 ohm optimum level for peak power and the 60 ohm optimum level for average power. If average power rating is the primary factor, a 50 ohm line should be used; but if attenuation is the main consideration, then a 75 ohm line should be selected. A 50 ohm line will have roughly 8.0% higher attenuation than a 75 ohm line with identical outer conductor diameters.

Cut-Off Frequency

The cut-off frequency of a coaxial line is the frequency above which undesirable modes of propagation are generated. A coaxial cable, therefore, must not be used above this frequency. Cut-off frequency, f_C, is inversely related to conductor dimensions and dielectric constant, as shown in the following equation:

$$f_C(GHz) = \frac{7.52}{\sqrt{\epsilon'}\,(D+d)}$$

Consideration of maximum frequency is important in high power TV applications where the cut-off frequency falls in the UHF bands for larger size coaxial cables (6⅛" and above). A maximum operating frequency is always specified for semi-flexible cable, which is somewhat lower than the calculated cut-off frequency.

Cable Selection

General

Semi-flexible coaxial cable for broadcast applications is selected on the basis of cut-off frequency, power being transmitted, loss (attenuation), size, installation, and cost. The following discussion describes the typical components required in a system and the electrical performance evaluation in terms of power ratings, attenuation, efficiency, connector loss, VSWR, and derating considerations for the various sizes of cable available. Each installation should be considered on an individual basis.

Coaxial Components

A schematic drawing of a typical semi-flexible transmission line broadcast installation is shown in Fig. 5. Table 3 lists some typical components for the various coaxial cable sizes available. This is only a partial list and individual manufacturers' catalogs should be consulted when planning a semi-flexible coaxial cable system.

Power Ratings

There are two power ratings for a coaxial line. One, the peak power rating, is based on voltage breakdown considerations; the other, the average power rating, is based on the maximum heating the cable construction can safely withstand. At VHF, FM, and UHF frequencies, coaxial cables are average power limited, while the peak power rating is usually the limiting factor in amplitude modulation applications at medium freqency (MF). For amplitude modulation high frequency (HF)

TABLE 3

Typical components for semi-flexible coaxial cable installation.

Item No.	Description
2	Semi-Flexible Coaxial Cable
3	Antenna End Connector
4	Transmitter End Connector w/Gas Barrier
5	Hanger Kits
	Hardware Kits
	Angle Adaptor Kits
	Round Member Adaptor
	Tower Standoff Kit
	Threaded Rod Support Kit
	Insulated Hanger
	Angle Adaptor for Insuflated Hanger
	Round Member Adaptor for Insulated Hanger
6	Grounding Kit
7	Wall/Roof Feed Thru
8	90° Elbow
9	Hoisting Grip
10	Dehydrator

applications, the line may be either average or peak power limited depending on the conditions.

Peak power is the maximum RF power which can be reached in a short interval (a few RF cycles). (This is not the instantaneous power when the RF voltage is at a maximum, but is the power averaged over an RF cycle; "peak" here refers to peak amplitude of modulation.) In a continuous wave (CW) carrier (including FM), peak power equals average power. In 100% AM, the power rises to four times the carrier power at the peaks of the modulation envelope, so in this case peak power is four times the carrier power.

The peak power rating of a cable is dependent on voltage breakdown considerations which are essentially not frequency-sensitive; thus, this rating is considered constant with frequency. It is determined by the maximum voltage withstanding capability between the inner and outer conductors. This voltage capability varies with line size, line pressure, and type of pressurization gas.

A conservative method of determining peak power ratings is to use the equation:

$$P_{PK} = \frac{\left(\dfrac{E_p \times 0.707 \times 0.7}{2}\right)^2}{Z_C}$$

where: P_{PK} = Cable peak power rating, standard conditions, kW
E_P = DC production test voltage, volts
0.707 = RMS factor
0.7 = DC to RF factor (empirically verified)
2 = Safety factor on voltage
Z_C = Characteristic impedance, ohms

An adequate safety factor on peak power is necessary to safeguard against voltage breakdown, which can result in permanent damage to the cable. Manufacturers of coaxial cables may use different safety factors for published peak power ratings, so this should be taken into account when comparing catalog values. Cables are available which have been high-voltage tested to the equivalent of 400% of their rated peak power, resulting in a safety factor of 2 on voltage and 4 on peak power.

Peak power ratings can be determined using the average power chart in Table 4 for various sizes of semi-flexible coaxial cable. The peak power ratings correspond to the average power rating numbers at 1 MHz (AM). The ratings are for CW power, terminating VSWR of 1.0 and one atmosphere absolute of dry air pressure. Peak power must be derated for modulation technique and VSWR.

Average power is the power in the signal capable of creating heat. The average power rating of semi-flexible coaxial cable is based on the maximum inner conductor temperature which will permit safe, long-term performance for the particular dielectric material used. Since the average power rating is limited by heating, which is created by line losses, it decreases with increasing frequency. Table 4 is a listing of manufacturer's semi-

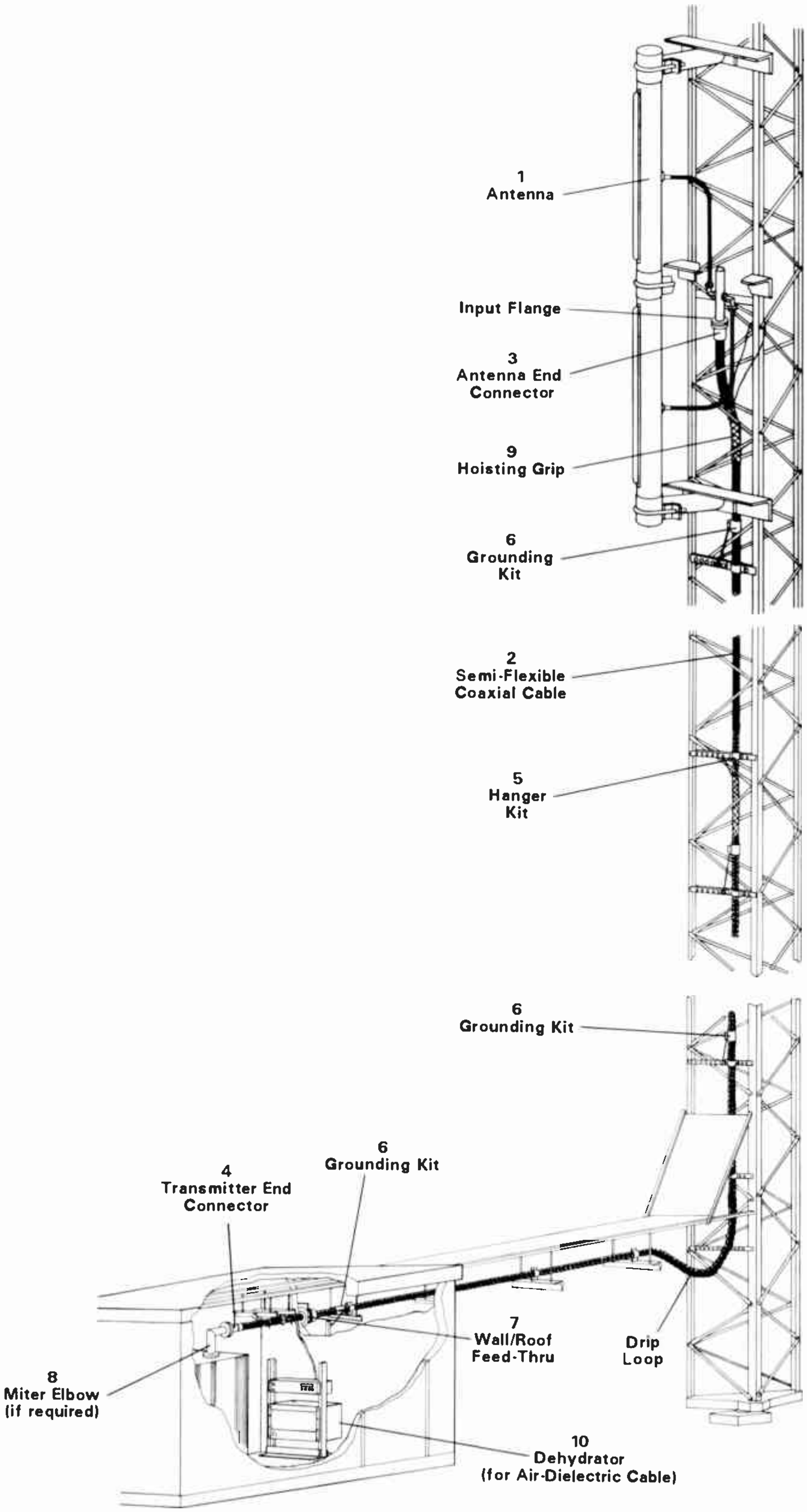


Figure 5. Typical semi-flexible cable broadcast installation.

TABLE 4

Average power ratings of semi-flexible coaxial cable assuming VSWR = 1.0, ambient temperature = 104°F(40°C) and atmospheric pressure, dry air.

SEMI-FLEXIBLE COAXIAL CABLE AVERAGE POWER RATINGS														
		FOAM-DIELECTRIC CABLE			AIR DIELECTRIC CABLES									
AVERAGE POWER, KW*														
	Frequency	7/8"	1-1/4"	1-5/8"	7/8"	1-5/8"	2-1/4"	3"	3-1/2"	4"	5"	6-1/8"	8-7/8"
Channel	(MHz)	LDF5-50A	LDF6-50A	LDF7-50A	HJ5-50A	HJ7-50	HJ12-50	HJ8-50B	HCC-312	HJ11-50	HJ9-50	HF6-1/8"	HF-8
AM*	1.00	44.00	50.00	143.00	44.00	145.00	209.55	320.00	521.24	493.00	765.00	1400.41	2370.68
2	55.25	7.34	12.71	18.31	8.73	19.65	27.42	51.14	67.61	77.62	101.27	183.51	310.65
3	61.25	6.96	12.04	17.34	9.27	18.61	26.10	48.34	64.07	73.34	95.67	174.01	294.57
4	67.25	6.63	11.46	16.51	7.86	17.72	24.95	45.94	61.02	69.66	90.86	165.22	280.70
5	77.25	6.16	10.65	15.35	7.33	16.49	23.34	42.59	56.74	64.55	84.17	154.35	261.26
6	83.25	5.92	10.24	14.76	7.04	15.85	22.52	40.99	54.56	61.94	80.77	148.48	251.35
FM	88.00	5.75	9.95	14.33	6.84	15.40	21.93	39.67	52.39	60.08	78.33	144.27	244.22
FM	98.00	5.44	9.40	13.54	6.47	14.55	20.82	37.41	50.07	56.63	73.82	136.43	230.95
FM	108.00	5.17	8.93	12.86	6.13	13.84	19.70	35.44	47.57	53.72	70.11	129.70	219.56
7	175.25	3.99	6.90	9.93	4.71	10.75	14.92	26.97	36.77	41.46	54.53	100.67	170.42
8	181.25	3.92	6.78	9.73	4.63	10.55	14.66	26.46	35.11	40.75	53.61	98.91	167.43
9	187.25	3.85	6.66	9.56	4.55	10.37	14.41	25.97	35.46	40.06	52.73	97.22	164.52
10	193.25	3.79	6.55	9.40	4.48	10.19	14.17	25.50	34.89	39.40	51.89	95.62	161.57
11	199.25	3.73	6.44	9.24	4.41	10.02	13.94	25.05	34.32	38.78	51.10	94.09	159.28
12	205.25	3.67	6.34	9.09	4.34	9.87	13.72	24.63	33.78	38.11	50.30	92.63	156.80
13	211.25	3.61	6.24	8.94	4.27	9.72	13.51	24.23	33.26	37.47	49.52	91.22	154.42
14	471.25	2.33	4.03	5.71	2.78	6.41	8.75	15.14	21.44	23.37	31.82	59.41	100.56
15	477.25	2.31	4.00	5.66	2.76	6.37	8.69	15.02	21.29	23.20	31.57	59.00	99.88
16	483.25	2.29	3.97	5.62	2.74	6.32	8.63	14.90	21.14	23.03	31.33	58.60	99.20
17	489.25	2.28	3.94	5.58	2.72	6.28	8.57	14.78	21.00	22.86	31.10	58.21	98.54
18	495.25	2.26	3.92	5.55	2.71	6.24	8.51	14.67	20.85	22.70	30.87	57.83	97.89
19	501.25	2.25	3.89	5.51	2.69	6.20	8.46	14.56	20.72	22.54	30.64	57.45	97.26
20	507.25	2.23	3.86	5.47	2.67	6.16	8.40	14.45	20.58	22.39	30.42	57.08	96.63
21	513.25	2.22	3.84	5.43	2.65	6.12	8.35	14.34	20.44	22.23	30.21	56.72	96.02
22	519.25	2.20	3.81	5.40	2.64	6.08	8.29	14.24	20.31	22.08	29.99	56.36	95.41
23	525.25	2.19	3.79	5.36	2.62	6.04	8.24	14.13	20.18	21.93	29.79	56.01	94.82
24	531.25	2.16	3.77	5.33	2.61	6.01	8.19	14.03	20.05	21.79	29.58	55.67	94.23
25	537.25	2.16	3.74	5.29	2.59	5.97	8.14	13.93	19.93	21.65	29.38	55.33	93.66
26	543.25	2.15	3.72	5.26	2.57	5.93	8.09	13.83	19.80	21.51	29.18	54.99	93.09
27	549.25	2.14	3.70	5.23	2.56	5.90	8.04	13.74	19.68	21.37	28.99	54.67	92.54
28	555.25	2.12	3.67	5.19	2.54	5.86	7.99	13.64	19.56	21.23	28.80	54.34	91.99
29	561.25	2.11	3.65	5.16	2.53	5.83	7.95	13.55	19.45	21.10	28.61	54.03	91.46
30	567.25	2.10	3.63	5.13	2.52	5.80	7.90	13.46	19.33	20.97	28.42	53.71	90.93
31	573.25	2.08	3.61	5.10	2.50	5.76	7.85	13.37	19.22	20.84	28.24	53.41	90.41
32	579.25	2.07	3.59	5.07	2.49	5.73	7.81	13.28	19.10	20.71	28.06	53.10	89.90
33	585.25	2.06	3.57	5.04	2.47	5.70	7.76	13.20	18.99	20.59	27.89	52.81	89.39
34	591.25	2.05	3.55	5.01	2.46	5.67	7.72	13.11	18.89	20.47	27.72	52.51	88.90
35	597.25	2.04	3.53	4.98	2.45	5.64	7.68	13.03	18.78	20.35	27.55	52.22	88.41
36	603.25	2.03	3.51	4.95	2.45	5.60	7.64	12.95	18.67	20.23	27.38	51.94	87.93
37	609.25	2.02	3.49	4.92	2.42	5.57	7.59	12.86	18.57	20.11	27.21	51.66	87.45
38	615.25	2.00	3.47	4.90	2.41	5.54	7.55	12.78	18.47	20.00	27.05	51.38	86.98
39	621.25	1.99	3.45	4.87	2.40	5.52	7.51	12.71	18.37	19.88	26.89	51.11	86.52
40	627.25	1.98	3.43	4.84	2.38	5.49	7.47	12.63	18.27	19.77	26.74	50.84	86.07
41	633.25	1.97	3.42	4.82	2.37	5.46	7.43	12.55	18.17	19.66	26.58	50.58	85.62
42	639.25	1.96	3.40	4.79	2.36	5.43	7.40	12.48	18.07	19.55	26.43	50.32	85.18
43	645.25	1.95	3.38	4.76	2.35	5.40	7.36	12.40	17.98	19.45	26.28	50.06	84.75
44	651.25	1.94	3.36	4.74	2.33	5.38	7.32	12.33	17.89	19.34	26.13	49.81	
45	657.25	1.93	3.35	4.71	2.32	5.35	7.28	12.26	17.79	19.24	25.99	49.56	
46	663.25	1.92	3.33	4.69	2.31	5.32	7.25	12.19	17.70	19.14	25.84	49.31	
47	669.25	1.91	3.31	4.67	2.30	5.30	7.21	12.12	17.61	19.03	25.70	49.07	
48	675.25	1.90	3.30	4.64	2.29	5.27	7.18	12.05	17.52	18.94	25.56	48.83	
49	681.25	1.89	3.28	4.62	2.28	5.25	7.14	11.98	17.44	18.84	25.43	48.59	
50	687.25	1.88	3.26	4.60	2.27	5.22	7.11	11.92	17.35	18.74	25.29	48.36	
51	693.25	1.87	3.25	4.57	2.26	5.20	7.07	11.85	17.26	18.65	25.16	48.13	
52	699.25	1.87	3.23	4.55	2.25	5.17	7.04	11.79	17.18	18.55	25.03	47.90	
53	705.25	1.86	3.22	4.53	2.24	5.15	7.01	11.72	17.10	18.46	24.89	47.68	
54	711.25	1.85	3.20	4.51	2.23	5.13	6.97	11.66	17.01	18.37	24.76	47.46	
55	717.25	1.84	3.19	4.48	2.22	5.10	6.94	11.59	16.93	18.27	24.63	47.24	
56	723.25	1.83	3.17	4.46	2.21	5.08	6.91	11.53	16.85	18.18	24.50	47.02	
57	729.25	1.82	3.16	4.44	2.20	5.06	6.87	11.46	16.78	18.10	24.38	46.81	
58	735.25	1.81	3.14	4.42	2.19	5.04	6.84	11.40	16.70	18.01	24.25	46.60	
59	741.25	1.80	3.13	4.40	2.18	5.02	6.81	11.34	16.62	17.92	24.13	46.39	
60	747.25	1.80	3.11	4.38	2.17	5.00	6.78	11.28	16.54	17.84	24.01	46.18	
61	753.25	1.79	3.10	4.36	2.16	4.98	6.75	11.22	16.47	17.75	23.89	45.98	
62	759.25	1.78	3.08	4.34	2.15	4.96	6.72	11.16	16.40	17.67	23.77	45.78	
63	765.25	1.77	3.07	4.32	2.14	4.94	6.69	11.10	16.32	17.58	23.65	45.58	
64	771.25	1.76	3.06	4.30	2.13	4.92	6.66	11.05	16.25	17.50	23.54	45.38	
65	777.25	1.76	3.04	4.28	2.12	4.90	6.63	10.99	16.18	17.42	23.42	45.19	
66	783.25	1.75	3.03	4.26	2.11	4.85	6.60	10.93	16.11	17.34	23.31	45.00	
67	789.25	1.74	3.02	4.24	2.10	4.86	6.58	10.88	16.04	17.26	23.20	44.81	
68	795.25	1.73	3.01	4.22	2.10	4.84	6.55	10.82	15.97	17.19	23.09	44.62	
69	801.25	1.73	2.99	4.20	2.09	4.82	6.52	10.77	15.90	17.11	22.98	44.44	

*Peak power rating

TABLE 5
Average transmitter power calculations.

Modulation	Average Power Calculation
HF, AM	$P_{AVG} = P_C(1 + \frac{M^2}{2})$
FM	$P_{AVG} = P_T$
TV	$P_{AVG} = .8P_{TV}$

P_{AVG} = Average transmitter power (visual & aural)
P_C = Carrier power
M = Amplitude Modulation index
P_T = FM transmitted power
P_{TV} = TV peak sync. power

Note: Commercial AM (530 to 1610 kHz) is usually peak power limited.

flexible coaxial cable average power ratings for a VSWR of 1.0 and ambient temperature of 104°F (40°C). Average power must be derated for ambient temperature and VSWR. A discussion of average power ratings of coaxial cable is given in Chipman[5] and Martin[7].

The average power that is applied to a transmission line is dependent on the nominal CW power of the transmitter and the type of modulation, and can be calculated for different modulation schemes using Table 5. The 0.80 factor used for the TV average power calculation is based on a totally black picture (60% peak TV power) plus aural signal (20% peak TV power). This can be considered a maximum average power level.

Attenuation

The attenuation of a transmission line is an expression of the ratio of input power to output power in decibels. It is determined by the material properties and construction of the line.
A listing of semi-flexible coaxial cable attenuation for various sizes and construction, set up by broadcast channels and frequencies, is presented in Table 6. The constructions listed are 50 ohm foam and air dielectric cables. Sizes correspond to the nominal diameter over the jacket. This data is based on a VSWR of 1.0 and an ambient temperature of 75°F (24°C).

Efficiency

Efficiency is a convenient way of expressing the power loss of a transmission line and is useful in selecting the appropriate line for the application. It is defined as the ratio of power delivered to the antenna to input power into the transmission line. It is easily calculated from the attenuation and length of the transmission line by applying the following equation:

$$\text{Efficiency} = \frac{\text{Power (out)}}{\text{Power (in)}} \times 100\%$$

$$\text{Efficiency} = \frac{100\%}{10^{\alpha/10}}$$

where: α = Total line attenuation in dB

Example: 300 feet of HJ9–50 5″ air dielectric coaxial cable for Channel 6.
Carrier frequency = 83.25 MHz
Attenuation = 0.071 db/100 ft (Obtained from Table 6)
L = 300 ft
α = 0.213 dB (.071 × 3)

$$\text{Efficiency} = \frac{100\%}{10^{.213/10}}$$

$$= \frac{100\%}{10^{.0213}}$$

$$= 95.21\%$$

The remaining power is dissipated in the transmission line as heat.

Connector Loss

The effect of connectors on transmission loss is negligible, except for small connectors (SMA, TNC) at frequencies of several GHz and higher. Therefore, for broadcast use, connector loss can be ignored.

Voltage Standing Wave Ratio (VSWR)

VSWR is an important factor in selecting a transmission line for a particular broadcast application. Typical VSWR values for standard semi-flexible cable and low VSWR semi-flexible cable are listed in Table 7.

The narrow bandwidths utilized in broadcast applications usually allow low VSWR cable to be selected. Therefore, it is important to let the manufacturer know what channel or bandwidth will be used.

Derating

The attenuation and power performance of transmission lines are affected by various external factors which must be taken into consideration. Cable attenuation and average power ratings are stated for particular operating conditions, which most likely will not be representative of a typical application. Performance will be affected by temperature, load VSWR, line pressure, line VSWR, and type of modulation.

Line attenuation will vary with ambient temperature as shown in the curve in Fig. 6. For example, if ambient temperature is increased by 101°F (56°C) above 75°F (24°C), to 176°F (80°C), attenuation will increase by a factor of 1.115 or 11.5%. Therefore, in the previous example for 300 ft of 5″ air dielectric cable, at 176°F total line attenuation would increase by a factor of 1.115, from 0.213 dB to 0.237 dB. This would decrease the efficiency to 94.69%.

The VSWR of the attached antenna will increase the total transmission loss of the system. Typically this is very small if there is a good match between the line and the antenna. Fig. 7 shows the minimum increase in loss with load VSWR, assuming a VSWR of 1.0 at the input of the transmission line.

The average power rating should be adjusted according to the actual ambient temperature, as shown in Fig. 8. In the example above, the average power

TABLE 6
Attenuation of semi-flexible coaxial cable assuming VSWR = 1.0 and ambient temperature = 75°F(24°C).

SEMI-FLEXIBLE COAXIAL CABLE ATTENUATION RATINGS

		FOAM-DIELECTRIC CABLE						AIR DIELECTRIC CABLES					
	ATTENUATION* dB/100 FT												
	Frequency	7/8"	1-1/4"	1-5/8"	7/8"	1-5/8"	2-1/4"	3"	3-1/2"	4"	5"	6-1/8"	8"
Channel	(MHz)	LDF5-50A	LDF6-50A	LDF7-50A	HJ5-50	HJ7-50A	HJ12-50	HJ8-50B	HCC-312	HJ11-50	HJ9-50	HJS-1/8"	HP-8"
AM	1.00	0.035	0.026	0.021	0.035	0.020	0.016	0.013	0.010	0.010	0.007	0.005	0.004
2	55.25	0.271	0.201	0.164	0.274	0.153	0.124	0.103	0.080	0.082	0.057	0.038	0.029
3	61.25	0.286	0.213	0.174	0.288	0.161	0.131	0.109	0.084	0.087	0.060	0.040	0.031
4	67.25	0.300	0.223	0.182	0.302	0.169	0.137	0.114	0.089	0.091	0.064	0.042	0.033
5	77.25	0.322	0.240	0.196	0.324	0.181	0.148	0.123	0.095	0.098	0.069	0.046	0.035
6	83.25	0.335	0.250	0.204	0.337	0.188	0.154	0.127	0.099	0.102	0.071	0.048	0.037
FM	89.00	0.345	0.257	0.210	0.347	0.194	0.158	0.131	0.102	0.106	0.074	0.049	0.038
FM	99.00	0.365	0.272	0.223	0.366	0.205	0.167	0.139	0.108	0.112	0.078	0.052	0.040
FM	108.00	0.384	0.286	0.235	0.386	0.215	0.176	0.147	0.114	0.118	0.082	0.055	0.042
7	175.25	0.498	0.370	0.305	0.502	0.272	0.227	0.194	0.147	0.154	0.107	0.071	0.056
8	181.25	0.507	0.376	0.311	0.511	0.275	0.231	0.197	0.150	0.157	0.109	0.073	0.057
9	187.25	0.516	0.383	0.316	0.521	0.281	0.235	0.201	0.152	0.160	0.111	0.074	0.058
10	193.25	0.525	0.390	0.322	0.530	0.285	0.239	0.204	0.155	0.163	0.113	0.075	0.059
11	199.25	0.534	0.396	0.327	0.539	0.289	0.243	0.208	0.157	0.166	0.115	0.077	0.060
12	205.25	0.543	0.403	0.333	0.547	0.294	0.247	0.211	0.160	0.169	0.117	0.078	0.061
13	211.25	0.551	0.409	0.338	0.555	0.299	0.251	0.215	0.162	0.172	0.118	0.079	0.062
14	471.25	0.856	0.633	0.530	0.843	0.462	0.387	0.351	0.252	0.277	0.184	0.125	0.099
15	477.25	0.862	0.638	0.534	0.848	0.466	0.390	0.354	0.254	0.279	0.186	0.126	0.100
16	483.25	0.868	0.642	0.538	0.854	0.469	0.393	0.357	0.255	0.282	0.187	0.127	0.101
17	489.25	0.874	0.647	0.541	0.860	0.472	0.396	0.359	0.257	0.284	0.188	0.128	0.102
18	495.25	0.880	0.651	0.545	0.865	0.475	0.398	0.362	0.259	0.286	0.190	0.129	0.102
19	501.25	0.886	0.655	0.549	0.871	0.478	0.401	0.365	0.261	0.288	0.191	0.130	0.103
20	507.25	0.892	0.660	0.553	0.877	0.482	0.404	0.368	0.262	0.290	0.192	0.131	0.104
21	513.25	0.898	0.664	0.556	0.882	0.485	0.406	0.370	0.264	0.292	0.194	0.132	0.105
22	519.25	0.903	0.668	0.560	0.888	0.488	0.409	0.373	0.266	0.294	0.195	0.133	0.105
23	525.25	0.909	0.673	0.564	0.893	0.491	0.411	0.376	0.268	0.296	0.196	0.134	0.106
24	531.25	0.915	0.677	0.568	0.899	0.494	0.414	0.378	0.269	0.298	0.198	0.135	0.107
25	537.25	0.921	0.681	0.571	0.904	0.497	0.417	0.381	0.271	0.300	0.199	0.135	0.108
26	543.25	0.926	0.685	0.575	0.909	0.500	0.419	0.384	0.273	0.302	0.200	0.136	0.108
27	549.25	0.932	0.689	0.578	0.915	0.503	0.422	0.386	0.274	0.304	0.201	0.137	0.109
28	555.25	0.938	0.694	0.582	0.920	0.506	0.424	0.389	0.276	0.306	0.203	0.138	0.110
29	561.25	0.943	0.698	0.586	0.925	0.509	0.427	0.392	0.278	0.308	0.204	0.139	0.111
30	567.25	0.949	0.702	0.589	0.931	0.512	0.429	0.394	0.279	0.310	0.205	0.140	0.111
31	573.25	0.955	0.706	0.593	0.936	0.515	0.432	0.397	0.281	0.312	0.206	0.141	0.112
32	579.25	0.960	0.710	0.596	0.941	0.518	0.434	0.400	0.283	0.314	0.208	0.142	0.113
33	585.25	0.966	0.714	0.600	0.946	0.521	0.437	0.402	0.284	0.316	0.209	0.142	0.113
34	591.25	0.971	0.718	0.603	0.952	0.524	0.439	0.405	0.286	0.317	0.210	0.143	0.114
35	597.25	0.977	0.722	0.607	0.957	0.527	0.442	0.407	0.288	0.319	0.211	0.144	0.115
36	603.25	0.982	0.726	0.610	0.962	0.530	0.444	0.410	0.289	0.321	0.212	0.145	0.116
37	609.25	0.988	0.730	0.614	0.967	0.533	0.446	0.413	0.291	0.323	0.214	0.146	0.116
38	615.25	0.993	0.734	0.617	0.972	0.536	0.449	0.415	0.292	0.325	0.215	0.147	0.117
39	621.25	0.998	0.738	0.621	0.977	0.539	0.451	0.418	0.294	0.327	0.216	0.148	0.118
40	627.25	1.004	0.742	0.624	0.982	0.542	0.454	0.420	0.296	0.329	0.217	0.148	0.118
41	633.25	1.009	0.746	0.628	0.987	0.544	0.456	0.423	0.297	0.331	0.218	0.149	0.119
42	639.25	1.014	0.750	0.631	0.992	0.547	0.458	0.425	0.299	0.333	0.220	0.150	0.120
43	645.25	1.020	0.754	0.634	0.997	0.550	0.461	0.428	0.300	0.334	0.221	0.151	0.120
44	651.25	1.025	0.758	0.638	1.002	0.553	0.463	0.430	0.302	0.336	0.222	0.152	
45	657.25	1.030	0.762	0.641	1.007	0.556	0.466	0.433	0.303	0.338	0.223	0.153	
46	663.25	1.035	0.766	0.644	1.012	0.559	0.468	0.435	0.305	0.340	0.224	0.153	
47	669.25	1.041	0.770	0.648	1.017	0.561	0.470	0.438	0.307	0.342	0.225	0.154	
48	675.25	1.046	0.773	0.651	1.022	0.564	0.472	0.440	0.308	0.344	0.227	0.155	
49	681.25	1.051	0.777	0.654	1.027	0.567	0.475	0.443	0.310	0.345	0.228	0.156	
50	687.25	1.056	0.781	0.658	1.032	0.570	0.477	0.445	0.311	0.347	0.229	0.157	
51	693.25	1.061	0.785	0.661	1.036	0.572	0.479	0.447	0.313	0.349	0.230	0.158	
52	699.25	1.066	0.789	0.664	1.041	0.575	0.482	0.450	0.314	0.351	0.231	0.158	
53	705.25	1.072	0.792	0.668	1.046	0.578	0.484	0.452	0.316	0.353	0.232	0.159	
54	711.25	1.077	0.796	0.671	1.051	0.581	0.486	0.455	0.317	0.354	0.234	0.160	
55	717.25	1.082	0.800	0.674	1.056	0.583	0.489	0.457	0.319	0.356	0.235	0.161	
56	723.25	1.087	0.804	0.678	1.061	0.586	0.491	0.459	0.320	0.358	0.236	0.162	
57	729.25	1.093	0.808	0.681	1.066	0.589	0.493	0.462	0.322	0.359	0.237	0.162	
58	735.25	1.098	0.811	0.684	1.071	0.591	0.495	0.464	0.323	0.361	0.238	0.163	
59	741.25	1.103	0.815	0.687	1.075	0.594	0.498	0.466	0.325	0.363	0.239	0.164	
60	747.25	1.108	0.819	0.691	1.080	0.597	0.500	0.469	0.326	0.364	0.240	0.165	
61	753.25	1.113	0.823	0.694	1.085	0.599	0.502	0.471	0.328	0.366	0.241	0.166	
62	759.25	1.118	0.826	0.697	1.090	0.602	0.505	0.473	0.329	0.368	0.243	0.166	
63	765.25	1.123	0.830	0.700	1.095	0.604	0.507	0.475	0.331	0.369	0.244	0.167	
64	771.25	1.128	0.834	0.704	1.099	0.607	0.509	0.478	0.332	0.371	0.245	0.168	
65	777.25	1.133	0.838	0.707	1.104	0.610	0.511	0.480	0.334	0.373	0.246	0.169	
66	783.25	1.138	0.841	0.710	1.109	0.612	0.513	0.482	0.335	0.374	0.247	0.170	
67	789.25	1.143	0.845	0.713	1.114	0.615	0.516	0.484	0.337	0.376	0.248	0.170	
68	795.25	1.148	0.849	0.716	1.118	0.617	0.518	0.487	0.338	0.377	0.249	0.171	
69	801.25	1.153	0.852	0.719	1.123	0.620	0.520	0.489	0.340	0.379	0.250	0.172	

TABLE 7

Typical VSWR and return loss values for standard and low VSWR semi-flexible cable.

Frequency MHz	Band	Standard Semi-flexible Coaxial Cable Typical VSWR (Return Loss)	Low VSWR Semi-flexible Coaxial Cable Maximum VSWR (Return Loss)
0.53-1.61	AM	1.02 (40.1 dB)	N.A.
3-30	HF	1.02 (40.1 dB)	N.A.
54-216	VHF,FM	1.10 (26.4 dB)	1.05 (32.2 dB)
470-740	UHF	1.12 (24.9 dB)	1.08 (28.3 dB)
740-806	UHF	1.15 (23.1 dB)	1.10 (26.4 dB)

rating of the 5″ coaxial cable at 83.25 MHz would be 80.8 kW for 104°F (40°C) ambient temperature. If the long term average ambient temperature is 63°F (17°C), then the average power rating can be increased by a factor of 1.5, resulting in a 121.2 kW average power rating.

Average power should also be derated for VSWR. The derating factor (D.F.) is calculated from the formula given below where F^1 is a factor that varies with frequency and line size. Select the F^1 factor from Fig. 9, calculate the D.F. and divide into the average power.

$$DF = \frac{(VSWR^2 + 1)}{2\,(VSWR)} + \frac{F^1\,(VSWR^2 - 1)}{2\,(VSWR)}$$

The following example is a calculation for the Channel 6 case with 5″ coaxial cable:

VSWR = 1.10 (Typical system VSWR with semi-flexible cable)

F^1 = .39 (From Fig. 9)

$$D.F. = \frac{(1.10^2 + 1)}{2\,(1.10)} + \frac{.39\,(1.10^2 - 1)}{2\,(1.10)} = 1.042$$

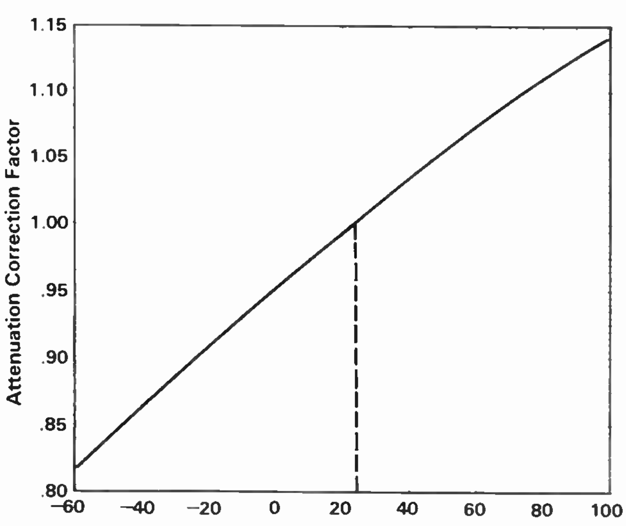


Figure 6. Variation of attenuation with ambient temperature.

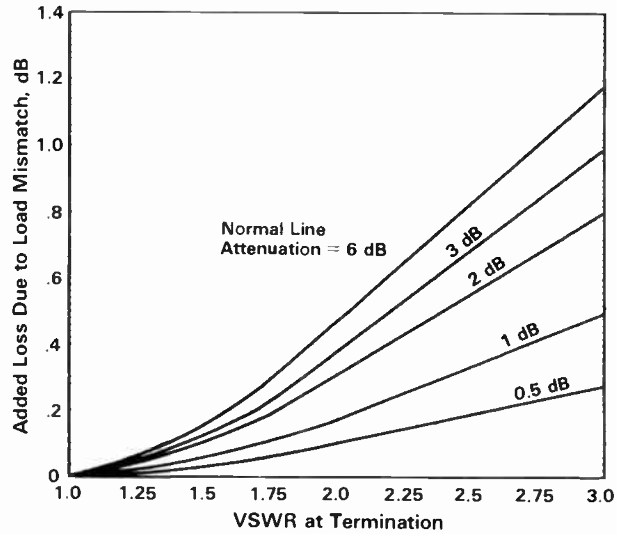


Figure 7. Effect of load VSWR on transmission loss.

Therefore, the average power rating of 80.8 kW (Table 4) will be decreased by a factor of 1.042, resulting in a 77.5 kW rating.

The peak power rating can be increased by pressurization, the use of high-density gases with high dielectric strength, or both. These effects are shown in Fig. 10.

Peak power ratings for semi-flexible cable must be derated for VSWR and modulation technique, as shown in Table 8:

Rated transmitter power must be less than the calculated derated peak power of the semi-flexible coaxial cable for safe operation.

Other Cable Characteristics

After selecting a cable on the basis of operating frequency, power handling, and attenuation, its other

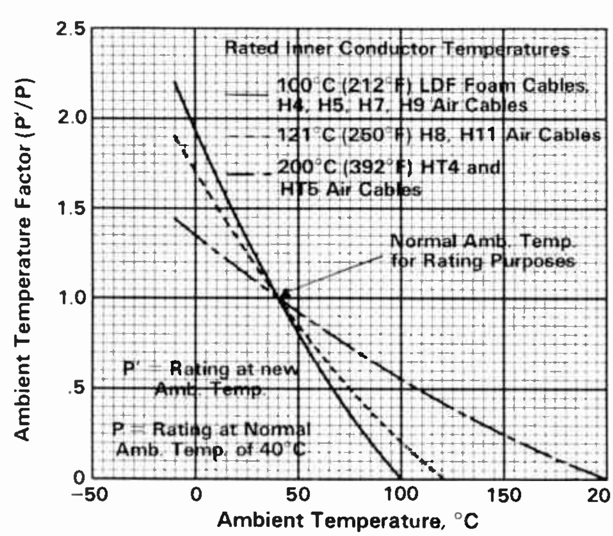


Figure 8. Variation of average power rating with ambient temperature.

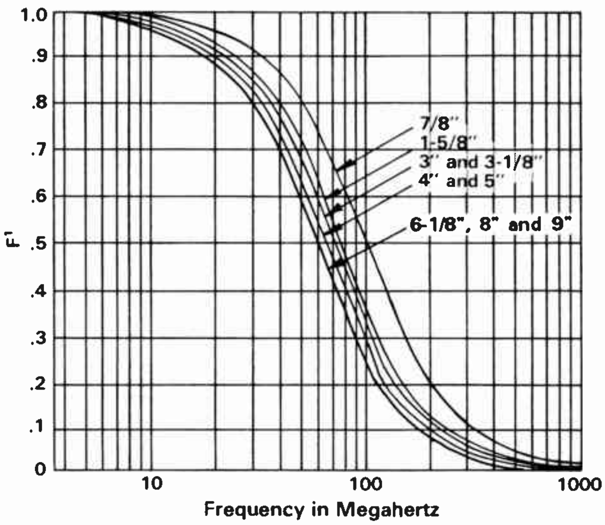


Figure 9. Derating factor for average power due to VSWR.

TABLE 8

Peak power derating for modulation and VSWR.

Modulation	Peak Power Derating Calculation
AM	$P_T < \dfrac{P_{PK}}{(1+M)^2 VSWR}$
FM	$P_T < \dfrac{P_{PK}}{VSWR}$
TV	$P_T < \dfrac{P_{PK}}{(1+AU+2\sqrt{AU})\,VSWR}$ $< \dfrac{P_{PK}}{(2.09)\,VSWR}$

P_T = Rated transmitter power
P_{PK} = Peak power rating of Table 4 or Table 17
M = Amplitude Modulation index
AU = Aural to visual ratio (20% Aural: AU = .2)
2.09 = Modulation derating factor for TV

characteristics should be considered. Some of these characteristics for typical semi-flexible coaxial cables are presented in Table 9. They include relative velocity, nominal inside transverse dimension in centimeters (required for FCC Form 302), diameter over jacket, minimum bend radius, and weight.

Connectors, Splices, and Adaptors

Usually the connector type is dictated by the antenna and transmitter input types and the performance required from the system. However, equally as important as connector selection is the care with which it is attached.

The instructions included with each connector must be followed with close attention to cable trimming dimensions. All surfaces must be as clean as possible to insure good electrical contact for the entire 360° around the cable edge. This is especially critical for higher power applications. All components must be tightened properly, and there must be no gaps on the electrical contact surface.

It is recommended practice to make all connector and splice attachments at ground level so that no chips from the cutting operation fall inside the cable. When a connector or splice must be attached to a cable already installed on the tower, tin snips or an equivalent tool should be used to do the preliminary cutting on the cable so as not to generate metallic chips. Rags should then be stuffed into the open cable ends to prevent chips from the final trimming and flaring operations from falling inside. Once the last chip-producing operations are completed, the rags can be carefully removed to avoid dropping any of the trapped particles.

Proper attachment of the finished connector to the mating antenna or radio equipment prevents loosening of the connection from vibration and other environmental stresses. Connectors with screw-on interfaces should be taped securely, using a two-part taping system. First, a layer of electrical tape is applied over both connectors (Fig. 11). Then a layer of butyl rubber tape is used to fill in any voids and make the outer surface relatively even (Figs. 12 and 13). Finally, additional electrical tape is added to compress the rubber tape and enclose the finished connection (Fig. 14). The resultant shape of this method has led to naming it the "Football Wrap."

Mated EIA style connectors must be carefully tightened. To avoid distorting the flange, tighten bolts that are opposite (not adjacent) to each other. Overtightening the hardware can also cause flange distortion. Care must be taken when seating O-rings to avoid pinching them between the flanges, which could result in poor electrical contact and air pressure leakage. See Table 10 for flange hardware requirements and recommended torque values.

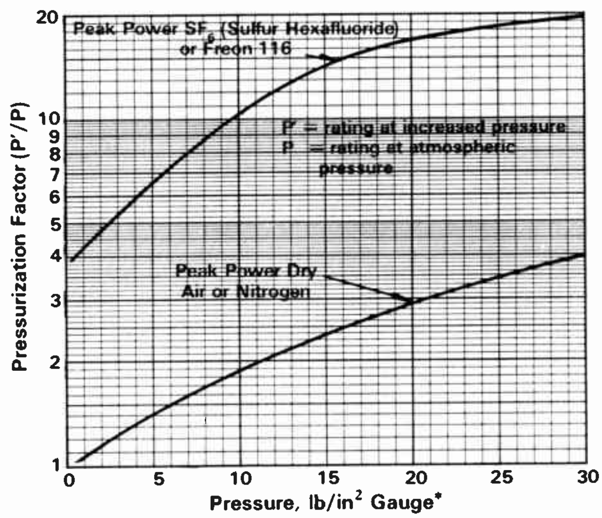


Figure 10. Pressurization factors.

TABLE 9
Characteristics of typical 50 ohm coaxial cables.

Size	Max. Freq GHz	Velocity Percent	Nominal Inside Transverse Dimension cm	Dia. Over Jacket Inches	Min. Bend Radius Inches	Weight lb/ft
Foam-Dielectric Cable						
7/8"	5.00	89.0	2.11	1.09	10	0.33
1 1/4"	3.30	89.0	3.11	1.55	15	0.66
1 5/8"	2.50	88.0	4.05	1.98	20	0.92
Air-Dielectric Cable						
7/8"	5.20	91.6	2.02	1.11	10	0.54
1 5/8"	2.70	92.1	3.99	1.98	20	1.04
1 5/8"	2.70	92.1	3.99	1.98	20	1.04
2 1/4"	2.30	93.1	4.96	2.38	22	1.16
3"	1.64	93.3	6.35	3.02	30	1.78
3 1/2"	1.43	96.0	7.52	3.50	30	1.98
4"	1.22	92.0	8.55	4.00	40	2.50
5"	0.96	93.1	11.30	5.20	50	3.30
6 1/8"	0.86	97.0	14.70	6.73	79	7.33
9"	0.65	97.0	19.50	8.90	98	12.50

Figure 11. Wrap both connectors with layer of 3/4" plastic tape. Overlap tape to half its width and extend wrapping two inches beyond connection.

Figure 12. Cut butyl rubber tape into six 12" lengths. Form a tapered surface by starting with two tapes folded to half their widths. Finish with one full-width tape.

Figure 13. Lay three rubber tapes along the connection so they overlap. Pull tape as necessary for overlap. Press tapes together along overlaps.

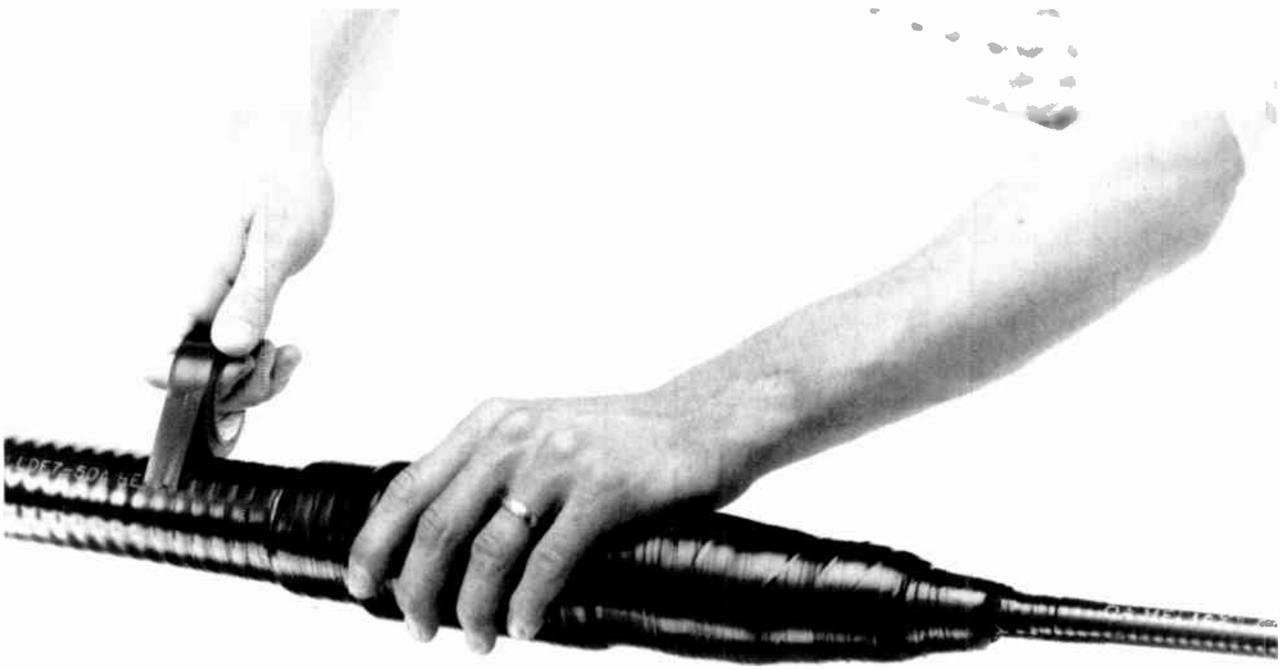


Figure 14. Wrap connection with layer of 2" tape, then layer of ¾" plastic tape. Overlap each winding to half the tape width. Extend wrapping two inches beyond previous tape. DO NOT PULL tape to tear it during last few turns; instead, cut it. (Pulled tape eventually unravels.)

Installation

General

Care is required in installing semi-flexible transmission line to assure a lasting installation and prevent damage. Procedures are outlined in Fig. 15. These tips distill the experience of many years of successful semi-flexible transmission line installation.

TABLE 10

EIA flange hardware requirements and recommended torque values.

EIA Flange Size	Bolt Size	No. of Bolts	Recommended Torque Value
7/8"	1/4"	3	80 lb-in (9.0 N-m)
1 5/8"	5/16"	4	140 lb-in (15.8 N-m)
3 1/8"	3/8"	6	20 lb-ft (27.1 N-m)
6 1/8"	3/8"	12	20 lb-ft (27.1 N-m)
8 3/16"	3/8"	18	20 lb-ft (27.1 N-m)
9" (SCL950)	3/8"	20	25 lb-ft (34.0 N-m)
9 3/16"	3/8"	20	20 lb-ft (27.1 N-m)

Hoisting grips are used to attach the cable to the hoisting line during lifting. Cable hangers (for ½" and larger cables) or nylon cable ties (for ⅜" and smaller cables) are used to affix the cable securely to the tower, and angle and round member adapters eliminate any need to drill holes.

Preparation for Installation

Inspect the cable for possible shipping damage and pressure loss. Most air dielectric cables are pressurized to 10 psig (70 kPa) for shipment. Allowable pressure drop is 1 psig (7 kPa) per 24 hours. Do not install cable which exceeds the standard leak rate without first checking to determine the cause.

Coaxial cables are usually shipped on reels or in cartons. These containers should be carefully handled to avoid damaging the cable. Reels should always be stored and moved on their flange edges and should never be laid flat or dropped during handling.

If a cable has been fitted with factory-attached connectors, leave their pressure plates or protective covers in place until ready to attach them to the mating

Installation "Do's"

1. Do keep reels on their flange edges during transportation and storage.
2. Do keep connectors and accessories stored where they will not be damaged by rain or dirt.
3. Do keep air dielectric cables pressurized and foam dielectric cables capped while in storage.
4. Do use hoisting grips to lift and secure each cable in place.
5. Do protect antenna end connectors from damage while lifting the cable up the tower.
6. Do securely install hangers on cable at the manufacturer's recommended spacing.
7. Do secure the cable coming out of the antenna to a tower member to prevent stress to the antenna input connection.
8. Do tighten EIA style connections to their recommended torque levels sequentially, to avoid flange distortion.
9. Do tighten the antenna end connection firmly to the antenna input.
10. Do tape connections to avoid loosening and water penetration.
11. Do pressurize air dielectric cable to 3–8 psig with dry air or nitrogen.
12. Do ground each cable run, as a minimum, at the top and bottom of the vertical run and at the equipment shelter entrance.

Installation "Don'ts"

1. Don't lay reels on their sides.
2. Don't install any cable that shows signs of physical damage.
3. Don't install foam dielectric cable that has been left uncapped. The exposed end should be cut back several feet and the cable given a TDR test (see Testing section) to be sure that no moisture penetration has occurred.
4. Don't install air dielectric cable that does not have positive air pressure. First, purge out any moisture that may be inside.
5. Don't release pressure from air cables unless making a connection, cutting cable from a bulk length, or purging the line.
6. Don't pay off cable from the reel too quickly, as it could cause tangling.
7. Don't attach cable to the tower where it could chafe against tower components.
8. Don't bend cable tighter than its minimum recommended bend radius.
9. Don't overtighten Type "N" connectors with pliers. These connections are designed for hand tightening only.
10. Don't leave connector interfaces unprotected from the environment.
11. Don't exceed the maximum air pressure level of the component with the lowest maximum pressure rating (such as an antenna or pressure window) when pressurizing a complete transmission line/antenna system.

Figure 15. Coaxial cable installation "do's" and "don'ts."

pieces of equipment. This will protect the connector interfaces from handling damage.

Installation on Towers

Hoisting grips should be attached to the cable and hoist line as shown in Fig. 16. When installing lengths of more than 200 ft (60 m), additional hoisting grips at 200 ft (60 m) maximum intervals are required.

Reels can be used to unroll the cable slowly during hoisting. A pipe inserted through the reel hub and supported by stands will allow controlled unreeling. Pay off the cable from the bottom as shown in Fig. 16. Cable shipped in cartons can be uncoiled on the ground and carefully pulled up.

After the cable has been lifted, the hoisting grips can be permanently attached to the tower. The bottom end of each grip should be securely taped to the cable to prevent loosening and eventual slippage.

Hanger Installation

Proper hanger installation techniques are vital to trouble free operation. The purpose of hangers is to prevent cable movement from wind, rain, ice, and other forces. Excessive movement could damage the cable from abrasion and work hardening of the outer conductor. Each site needs to be evaluated thoroughly to determine what type of hanger (and standoff, if any) should be used and how far apart they should be spaced. As a starting point, the manufacturer's recommended hanger spacing for the size cable being installed should be considered.

The following are examples of the site variables that should be addressed:

- Cable size
- Tower height
- Proximity of tower members and joints
- Wind velocity extremes
- Average wind velocity
- Ice loading
- Temperature extremes
- Accessibility for future maintenance

As a general rule, hanger spacing should be halved (i.e., use double the number of hangers) when the wind velocity is consistently over 40 mph (65 km/h), where

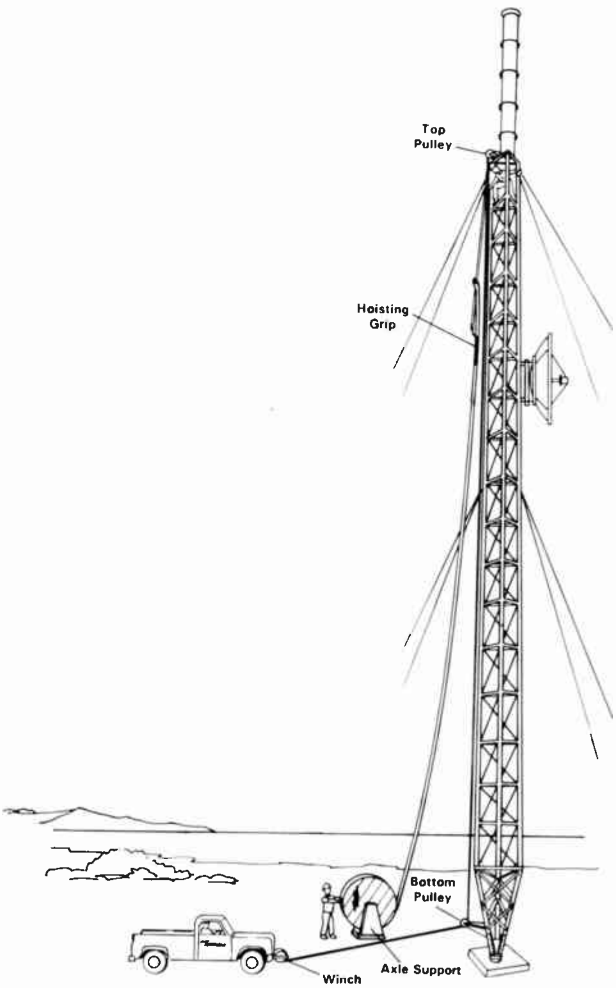


Figure 16. Installation of coaxial cable.

peak wind velocity may exceed 125 mph (200 km/h) and/or when heavy radial ice [½" (13 mm) to 2" (50 mm) thick] is expected to form.

Horizontal Runs

Horizontal coaxial cable runs need to be adequately supported. Typically they are either supported by an ice shield or buried.

Horizontal runs supported by an ice shield. Horizontal cable runs supported by ice shields or other structures normally require the same hanger spacing as the vertical run. A broad drip loop (Fig. 5) will help prevent moisture accumulation at the building entry feed-through. If water emerges from the building entry connection, indicating that the jacket has been damaged somewhere along its length, allowing water entry, a notch in the jacket at the bottom of the drip loop will allow for drainage.

Buried horizontal runs. Jacketed corrugated copper cables are very corrosion resistant and can be safely buried. Cable should be located below the frost line and placed in the middle of a 12" layer of sand to protect the jacket from stones or other sharp objects.

Conduit is also very desirable for burial installations, especially when the cable is to be placed under a service road. Care should be taken so as not to exceed the tensile strength of the cable when pulling it through. See Table 11 for typical tensile values. Conduit pulling lubricants should be applied generously to the cable. Bends should be kept to a minimum and should be long and sweeping to prevent overstressing.

Since any buried cable will likely become wet or occasionally submerged, all buried connectors or splices should be protected with a good weatherproofing kit. See "Connectors, Splices, and Adapters" section above for a step-by-step description of weatherproofing connections. Fig. 17 shows a photograph of a well-sealed splice connection.

Pressurization

Air dielectric transmission lines are pressurized with dry air or nitrogen for two reasons. First, the inside of the line must be kept dry to prevent possible corrosion of the copper inner and outer conductors. Second, electrical performance will be severely impaired if water enters the line, as VSWR and attenuation can be degraded.

Table 12 summarizes the advantages and disadvantages of four pressurization systems: nitrogen, manual regenerative dehydrator, automatic regenerative dehydrator and automatic continuous membrane dehydrator. Automatic dehydrators are recommended for unattended sites or those with a large number of runs.

After all connections have been made, purge and repressurize air dielectric cable to remove any residual moisture. Follow the guidelines set forth in the "Moisture Purging" section.

A gauge pressure of 8 psig is adequate for most installations. A pressure exceeding 10 psig (70 kPa) is unnecessary and could lead to damage to antennas, gas barriers, or other pressure sensitive devices. Usually the upper pressure limit is determined by the pressure rating of some other system component, such as the antenna.

After installation, check all connections for leaks. Use commercially available leak detectors or liquid detergent. An unbroken soap film applied over the entire joint will show even very small leaks.

The desiccant in a dehydrator may become saturated by ordinary atmospheric humidity if the dehydrator

TABLE 11

Typical tensile strength of coaxial cables commonly used for burial in conduit.

Cable Size	Dielectric Type	Tensile Strength
½"	low density foam	250 lbs (1112 N)
½"	air	700 lbs (3113 N)
⅞"	low density foam	325 lbs (1445 N)
⅞"	air	800 lbs (3560 N)
1¼"	low density foam	1500 lbs (6670 N)

Figure 17. A well-sealed splice connection protected with a "football wrap."

does not run for long periods (several weeks). It is good practice to allow some pressure (1 psig to 3 psig) to be released from the cable periodically. This enables the dehydrator to regenerate its desiccant.

Moisture Purging

Removing moisture from air dielectric semi-flexible cable may be necessary due to loss of pressure during storage, installation, or operation. It is also a good preventive maintenance practice to purge the cable periodically, since condensation can form from normal capillary action, especially on larger size cable.

On cables that are known to have been without pressure for some time, make sure there is no water near the bottom connector. This can be done by removing the connector body and lowering that end of the cable to see if any water drips out.

To purge the cable, attach an air line to the transmitter end connector gas port plug and remove the plug from the antenna end connector. Purge the line with dry nitrogen or air from a dehydrator. Set the regulator to 5 psi to 10 psi and purge at the rate of one hour for every 50 ft of cable.

If access to the antenna end connector is limited, an alternate method is to pressurize to 8 psig and let the air escape at the transmitter end of the cable after one hour. Repeat this procedure several times, allowing one hour each time for the air to mix.

After purging, replace the gas port plug and pressurize the cable to 8 psig. Then test the cable for insulation resistance as described in the "Testing" section following. Repeat the purging procedure if necessary.

TABLE 12

Comparison of broadcast pressurization systems.

Advantages	Disadvantages
Dry Nitrogen Tank System	
Good for small tight systems.	Low purge and system volumes.
Low dew point, −90°F (−68°C).	Complex regulator and alarm installations.
No power required.	Cylinder transportation and handling.
No moving parts and noise free.	
Manual Regenerative Dehydrator	
Good for small tight systems.	Requires periodic inspection.
Low power consumption.	Saturated silica gel must be baked and/or replaced.
10–15 year life expectancy.	Requires power.
Oil and lubrication free.	Low purge and system volume.
Dew point in excess of −36°F (−38°C).	Generates low level noise.
Automatic Regenerative Dehydrator	
Provides oil free air.	Requires power.
Dew point in excess of −40°F (−40°C).	Moving parts require maintenance every 5,000 hours.
System easy to install.	Generates periodic noise.
High purge and system volumes.	
No lubrication required.	
Automatic Continuous Membrane Dehydrator	
Provides oil free air.	Requires power.
Dew point in excess of −50°F (−46°C).	Moving parts require maintenance every 6,000 hours.
System easy to install.	Generates low level noise periodically.
No lubrication required.	
Meets UL®, CSA® requirements.	

Grounding

Grounding of semi-flexible transmission line is very important in order to protect against the extremely large currents created during lightning strikes, since the outer conductor forms a direct connection to the transmitter equipment.

The tower presents a lower impedance to ground for lightning than either the outer conductor of the coaxial cable or the antenna input connector; consequently, the outer conductor of the cable must be grounded close to the antenna, allowing the currents to travel down the tower to earth ground. On long vertical runs, grounding the cable at 200 ft (60 m) intervals is considered prudent. It should also be grounded at the bottom of the vertical run and again at the point where it enters the equipment building (to the external building grounding system), as shown in Fig. 5, as this is the last point at which lightning currents can be shunted to ground before they enter the transmitter room. Inside the building, grounding is covered by local requirements; for example, grounding inside the building as well as outside may be required by Article 820–33 of the National Electrical Code.®

Ground connections must be made with high quality copper wire or straps. Military standards and some local codes prohibit the use of braided copper straps or fine stranded wire. These materials will corrode in time, increasing the impedance of the connection and leading to possible equipment damage when lightning strikes occur.

After installation, all ground connections at the cable should be weatherproofed to prevent electrical contact degradation and to keep moisture from collecting between the cable jacket and outer conductor. Local building codes vary, so grounding plans should be checked for compliance to these codes.

Considerations for AM Tower Installations

If the supporting tower is also used as an AM broadcast radiator, it is necessary to prevent the AM energy from being grounded by the transmission line that feeds the TV or FM antenna. One way to accomplish this is to isolate the line up the tower for a distance of a quarter wavelength (at the AM broadcast frequency) from the base, using insulated cable hangers (Fig. 18). Because of the quarter wave isolation from the base of the tower, a very high impedance between the tower and the line is presented to the AM energy. Common practice is to make the isolated section approximately 0.22 wavelength long and to use a variable capacitor at the base of the tower to tune to quarter wave resonance. An RF isolation transformer is also an acceptable technique to avoid grounding.

Caution: Since a high RF potential exists between the line and the tower, the line should be mounted where it will not be accidentally touched by anyone on the tower climbing ladder.

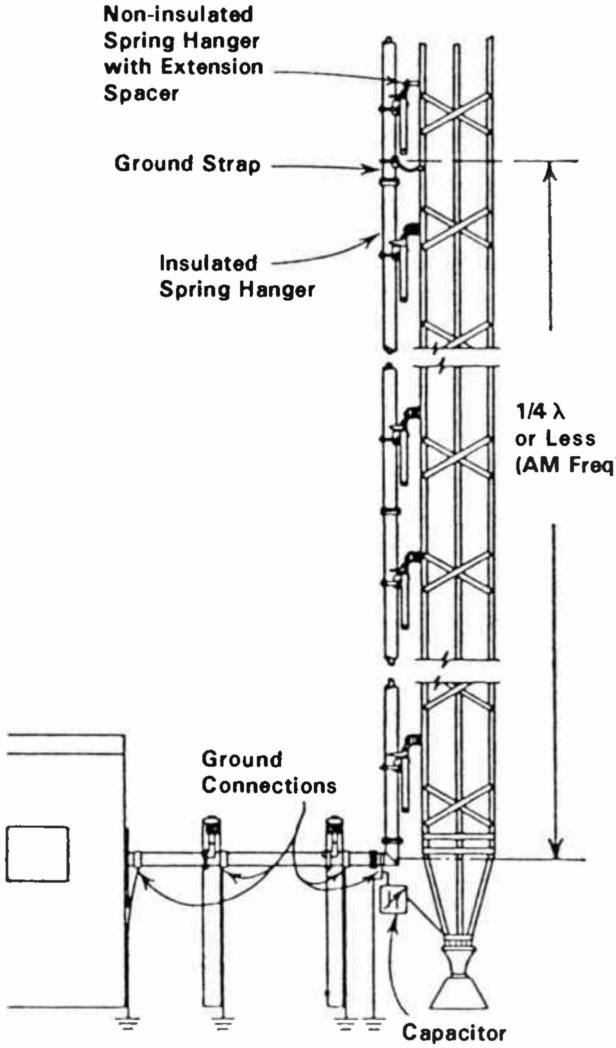


Figure 18. Installation on AM tower.

Sampling Lines

In order to obtain the desired pattern in an AM directional antenna, the relative phase and magnitude of the current in each radiator must be controlled. Monitoring of this current is done by installing sampling loops on one leg of each tower, with a length of coaxial cable terminated via a strap or end terminal. This coaxial cable is called a sampling line.

The purpose of a sampling line is to carry the current that results from the voltage induced across the sampling loop (which should be proportional to the tower current) to a phase monitor. Whether the sampling line current is a true representation of the phase of the radiator current depends on the phase stability of the sampling line.

Ordinary coaxial cable should not be used for sampling lines, since it is subject to significant phase variation with temperature change when first installed. Instead, special phase stabilized foam polyethylene dielectric or air dielectric coaxial cable should be specified. This cable has been repeatedly cycled by the manufacturer through a wide temperature range to eliminate phase instability caused by hysteresis. Foam dielectric cables are generally preferred for sampling lines, these cables being ordered from the manufacturer precut to a precise electrical length.

Testing

General

Electrical testing of semi-flexible coaxial cable should be performed after installation and prior to applying power. Testing should also be used for troubleshooting after VSWR trips have occurred, to locate the cause of the trip. It is advantageous to test while installation personnel are still on site.

Electrical testing is usually done to verify the performance of the transmission system. Common tests that are performed are the return loss (VSWR) sweep test and RF pulse test. These tests operate over the bandwidth of the system and give a good indication of line and antenna performance. Network analyzers with time domain capabilities and lightweight line analyzers can be used for overall system performance checks and are also excellent for locating discontinuities not detectable by RF pulse methods. Traditional time-domain reflectometry (TDR) devices can also be used to locate these discontinuities, but will not give an overall indication of performance over the channel bandwidth. This section reviews each technique and presents actual data taken on transmission line systems.

Sweep Test

The sweep test is a frequency domain test that measures the return loss of the line swept over the channel bandwidth. This test should be performed before power is applied to the line and must be done with the end of the line terminated in either the antenna or a load. The effects of the terminating device will be included in the measurement.

A return loss sweep test curve of a UHF-TV transmission system is shown in Fig. 19. The system consists of 800 ft of HJ9–50 air dielectric semi-flexible coaxial cable and a top mounted antenna. The maximum return loss level is about −26 dB (1.106 VSWR). The rippling effect in the trace arises from the continuous phasing of the reflection from the ends of the system.

RF Pulse Test

The RF pulse test shows the performance of the line in the time domain. This test displays impedance changes versus time and allows the user to correlate relative electrical (return loss) performance to general location in the line.

The test is done by applying a CW burst to the line and observing the reflected signal in time. The frequency of the CW is equal to the center frequency of the channel bandwidth, and the duration of the burst is equal to the inverse of the channel bandwidth times two. This technique allows the performance of the transmission system over the channel bandwidth to be displayed in time (length). The disadvantage of the pulse test is that information over only a small bandwidth is used to derive the time domain characteristics. This means the display will show only gross defects in the system.

In most time domain tests, such as RF pulse, the response curve expresses return loss in decibels versus time on the horizontal axis. A rough conversion of time to distance for air dielectric cables is 2.0 nanoseconds per foot. (This allows for the round trip time of the pulse.) The calculation can be refined by including the velocity of propagation of the particular cable.

Network Analyzers

Network analyzers with time domain capabilities can be used for both sweep testing and simulation of RF pulse and TDR testing. The advantages of this equipment are that one unit contains all the necessary transmission testing functions and the time domain functions allow excellent control over the resolution. The time domain can be used to simulate an RF pulse by looking over only a small frequency range (i.e., channel bandwidth), or over the entire range of frequencies up to line cut-off. Using the equipment up to cut-off gives excellent resolution and allows small impedance variations to be displayed in time (length). Therefore, the analyzer can be used to locate problems that are not seen with pulse testing but could lead to eventual failures. The disadvantages of a network analyzer are size and cost.

An impedance plot of a transmission system using a network analyzer is shown in Fig. 20. The time domain information was calculated from data taken in the frequency domain up to 900 MHz. Note that this sensitive test revealed a slight discontinuity (−43 dB) in the vertical run, arising from a splice in the coaxial cable.

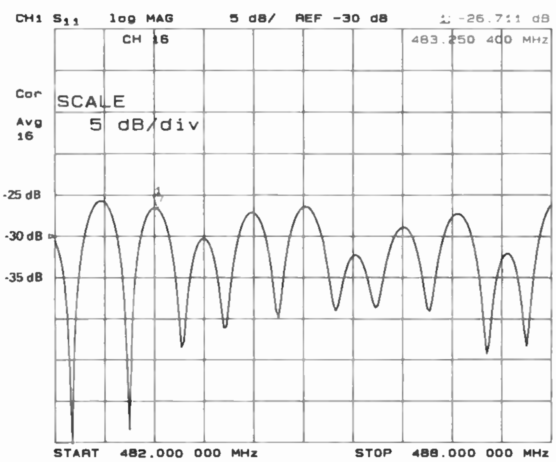


Figure 19. Swept return loss of system including antenna and 800 ft. of HJ9-50, 5″ air dielectric semi-flexible coaxial cable.

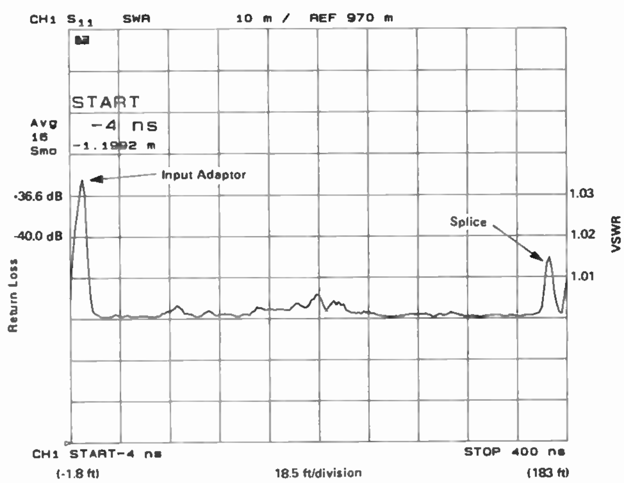


Figure 20. Time domain impedence plot of system including antenna and 800 ft. of HJ9-50, 5″ air dielectric semi-flexible coaxial cable (first 183 ft. shown).

(Note: All of these tests were performed using an HP 8753C Network Analyzer with time domain analysis. This unit is excellent for performing complete line verification and debugging.)

Time Domain Reflectometry (TDR) Test
A TDR test can be used to examine a transmission line for faults. The TDR test set transmits a fast rise time voltage step and measures the reflected signal. This signal will have significant frequency components up to 35% of the inverse of the pulse rise time. Therefore, it is important to adjust the pulse rise time to the cut-off frequency of the line.

Since TDR is a broadband test, it cannot be used to determine the performance of the line over the channel bandwidth.

Transmission Line Analyzers
Commercial transmission line analyzers are available that are made for field use so they are easily hand-carried to the site. The majority are used as fault finders and simulate a TDR test, but it is beneficial to have a unit that will also do a sweep test and simulate an RF pulse test. One such unit is the Systron Donner model 5220 Transline Analyzer. The 5220 will provide all of the characteristics discussed in the network analyzer section, but is easy to use, lightweight and provides a hard copy printout.

Insulation Resistance Test
An insulation resistance test can be performed on a semi-flexible coaxial cable to determine if it is sufficiently dry after purging with dry air before applying power. This test should be performed after any line has suffered from moisture ingress. The insulation resistance between the inner and outer conductor should be greater than 100,000 Megohms. There are many inexpensive units available that can measure insulation resistance.

Maintenance
Semi-flexible coaxial cable should be maintained according to the following guidelines:

1. Maintain at least 3 lb/in^2 of dry air or nitrogen pressure at all times.
2. If a leak occurs, keep positive pressure on the line until the leak is repaired. This will reduce moisture ingress.
3. Locate and repair pressure leaks immediately.
4. If moisture problems are suspected, shake the line near the bottom of the vertical run to see if significant performance changes occur. If so, open line and let water drain.
5. Purge system after pressure leak repair.
6. Do not continually override the transmitter reverse power trip. When a trip occurs, inspect line carefully to determine problem before reapplying power.
7. Visually inspect system on a yearly basis, checking hangers for tightness, grounding kits for good contact, and connectors for pressure leaks.

RIGID LINE

General
This section reviews the common sizes of rigid coaxial transmission lines used in broadcast applications, ranging from ⅞" to 9³⁄₁₆" in diameter. Fig. 21 shows the construction of a typical coaxial rigid line. Rigid lines have inherently low attenuation and VSWR which make them ideal for high power broadcast applications. Inner conductors are fabricated using high conductivity oxygen-free hard copper tubing. Outer conductors are also made of copper, but can be constructed with aluminum to reduce cost and weight. The lengths are made in 20 ft. (nominal) flanged sections. The copper inner conductor is supported in the outer conductor by peg or disc insulators with a low dielectric constant, low dissipation factor, high voltage breakdown, and good stability at high operating temperatures. Teflon, which has these characteristics, is usually used for the support insulator in rigid lines. The outer and inner conductor diameters are chosen to obtain a desired characteristic impedance of the line, normally 50 or 75 ohms.

The inner conductors of adjacent line sections are joined together by inner connector "bullets" which have fitted Teflon insulators that are captured in the outer flanges. Fig. 22 shows a connector for a 6⅛" 75 ohm line. Tension spring fingers on each end of the inner connector fit inside the inner conductors and connect two adjacent inner conductors together mechanically and electrically. This ensures efficient transfer of power. The insulator serves to anchor the inner conductor within the outer conductor in a vertical installation. Silver plating is commonly used on inner connectors to reduce the resistive losses of the joint.

The outer conductors are attached by bolting the flanges together; RF electrical contact is made between flanges by a raised contact surface. A pressure seal is maintained in the line by a flange O-ring. EIA standards RS-225 and RS-259 cover some 50 ohm and 75 ohm rigid coaxial transmission lines.

Propagation Factors

TEM Mode of Propagation
In a coaxial line the usual mode of energy wave propagation is TEM (Transverse Electromagnetic). In this mode the electric and magnetic fields are both perpendicular to the direction of energy propagation (Fig. 1A). This mode determines the impedance, attenuation, phase shift, and velocity of propagation characteristics.

Velocity of Propagation
The velocity of propagation (V_P) is equal to $\frac{1}{\sqrt{\epsilon'}}$ and is expressed as a fraction of the speed of light in a vacuum (approximately 300,000,000 m/sec). For example, TFE Teflon has a dielectric constant, ϵ', of 2.1, so its V_P equals 0.69. Therefore, the TEM wave

Figure 21. 6⅛", 75-ohm rigid coaxial line.

will travel at 207,000,000 m/sec through a solid piece of Teflon.

Determination of Cut-Off Frequency

The cut-off frequency of a coaxial line is the frequency above which undesirable modes of propagation are generated. Therefore a coaxial line must be used below this frequency Cut-off frequency, f_c, is inversely related to conductor dimensions and the relative dielectric constant of the insulation material as shown in the following equation:

$$f_c(GHz) = \frac{7.52}{\sqrt{\epsilon'}(D + d)}$$

The coaxial line should operate below this frequency.

Consideration of maximum frequency is important in high power TV applications where the cut-off frequencies start to fall in the UHF bands for the larger size coax lines (6⅛" and above). The specified maximum operating frequency for rigid coax includes a safety factor and is somewhat lower than the calculated cut-off frequency.

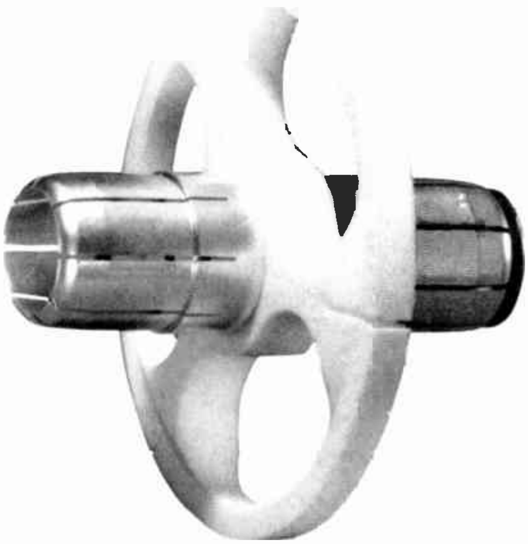


Figure 22. Inner connector for a 6⅛", 75-ohm rigid line.

VSWR, Return Loss, Reflection Coefficient

The voltage standing wave ratio (VSWR), return loss, and reflection coefficient are expressions relating the impedance of a transmission system to the reference impedance (usually 50 or 75 ohms). A change in impedance, from reference, will produce a reflected signal from the incident wave. The reverse and forward traveling signals will produce a standing wave in the line. Some useful conversion equations are given in Table 13.

Narrow band VSWR spikes will occur in a transmission line system when there are inherent periodic impedance changes. The frequency of the VSWR

TABLE 13
Conversion formulas.

$RL = -20\log(\lvert\Gamma\rvert)$	$VSWR = \frac{1+\lvert\Gamma\rvert}{1-\lvert\Gamma\rvert}$	$\lvert\Gamma\rvert = \frac{VSWR-1}{VSWR+1}$
$\lvert\Gamma\rvert = 10^{-\frac{RL}{20}}$		$\lvert\Gamma\rvert^2 = \frac{Pr}{Pn}$

RL: Return Loss, dB
Γ: Reflection Coefficient
VSWR: Voltage Standing Wave Ratio
Pn: Incident Power, Watts
Pr: Reflected Power, Watts

spikes is inversely related to the distance between discontinuities according to the following equation:

$$F_{SPIKE}(MHz) = \frac{492.15 \times V_P \times N}{L}$$

where: F_{SPIKE} = Frequency of VSWR spikes, MHz
N = Any integer
V_P = Velocity of propagation (assuming an approximate speed of 300,000,000 m/sec)
L = Equal distance between impedance discontinuities, ft

The level of the VSWR spike depends on the number of evenly spaced discontinuities and magnitude of impedance change.

High Power/Thermal Issues

Thermal Analysis

In a high power broadcast application, the inner conductor of a transmission line will run substantially hotter than the outer conductor. For example, if 60 kW average power is applied to a 6⅛" 75-ohm rigid line at a carrier frequency equivalent to channel 27 (549 MHz), the inner conductor will typically run 74°F hotter than the outer conductor.

With the thermal expansion of copper being 9.8×10^{-6} in/in/°F, the inner will expand by 0.174" over a 20 ft section of line with respect to the outer. In semi-flexible line this growth is absorbed in the corrugations of the inner and outer conductor. For conventional rigid line, the only place for this movement to occur is on the interface of the inner conductor and inner connector. Since the inner connector at each flange is captured by the Teflon insulator, the movement is contained in each section.

To allow for this movement the inner conductor of the line is made slightly shorter so that it will not bottom out on the shoulder of the inner connector when it expands. When the power level in the line is varied, the inner conductor will move on the inner connector. This movement, which is most severe when the transmitter is turned on for operation, will cause an electrical and mechanical degradation in the inner joint and eventually produce metallic shavings. These shavings, over time, can accumulate on the flange insulators and create a path allowing a voltage flashover to occur. The degradation will also cause an increase in I^2R losses at the joint resulting in additional power losses and possible overheating of the dielectric insulator supporting the inner connector.

Expansion/Contraction Compensation Devices

To overcome this problem in the inner joints of rigid lines, expansion/contraction compensating devices are used. Two basic designs are available for high power rigid coaxial lines: the watchband spring and the bellows (Fig. 23). The watchband spring allows the movement to occur on a spring which is built into one end of the inner connector. This effectively reduces wear between the inner connector and inner conductor. The bellows device is built into the inner conductor and absorbs the differential growth/shrinkage of the inner. This eliminates any mechanical movement of the inner conductor on the inner connector, similar to the corrugation principle used in the construction of semi-flexible transmission lines.

The other major area of concern at the junction of the inner connector and inner conductor is overheating due to poor contact pressure. Localized hot spots, created by years of resistive losses at points of higher resistance around the joint circumference, can ultimately cascade to failure. The watchband spring must provide the proper amount of uniform contact pressure throughout the life of the joint while being subjected to the abrasion of the expansion/contraction process. In the bellows approach, all the surface area of the spring fingers is in contact with the inner conductor and provides the maximum contact pressure desired to assure efficient power transfer. Since no movement occurs at this joint in the bellows line, the amount of bullet insertion force can be maximized and will remain constant.

The different compensation methods must be evaluated by the user to ensure that the possibility of flashover due to metallic shavings, or heating due to varying contact pressure, is reduced as much as possible over the life of the line.

Applications

Generally, rigid lines are used for the main feeder run from the transmitter diplexer output to the antenna input in high power FM, VHF, and UHF television broadcast applications and for interior runs between transmitter and diplexer.

Line Selection Issues

Rigid line is selected on the basis of frequency of operation, power to be transmitted, acceptable loss of power (attenuation), installation, and cost. During the selection process the attenuation and power ratings of the line must be derated to take into consideration the particular characteristics of the installation site. This section discusses what factors must be considered when selecting a rigid line for a particular site. The section concludes with a case study for a given TV application that walks through all of the calculations that are necessary in the selection process.

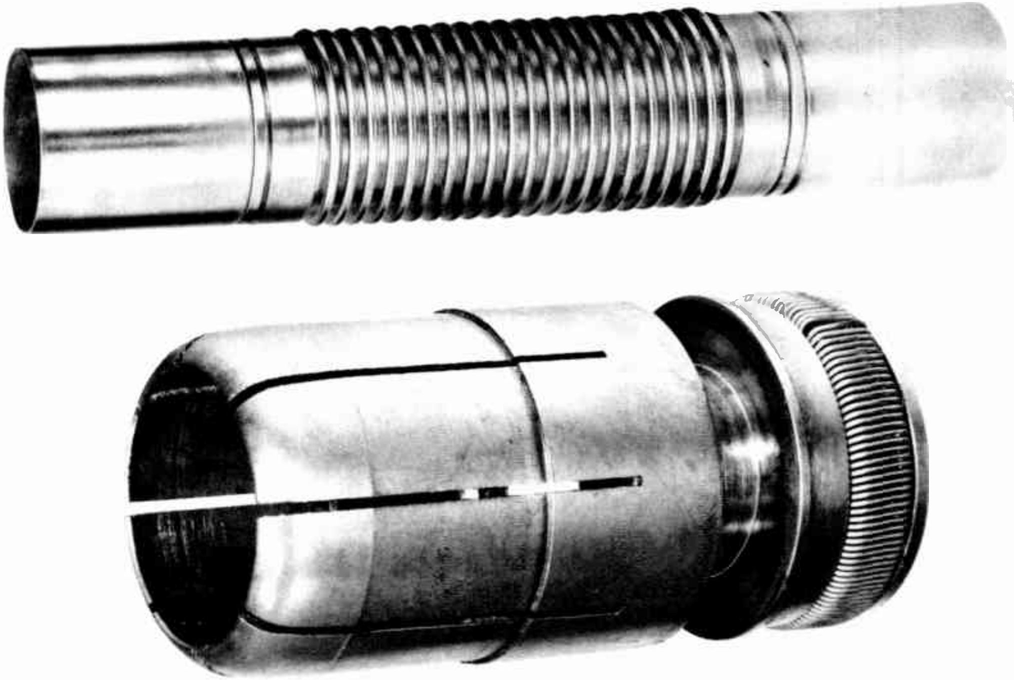


Figure 23. Compensation devices.

Characteristics

Table 14 is a partial listing of rigid line characteristics for the standard sizes and impedances.

Note that for the same size line, 75-ohm versions have a higher cut-off frequency than 50-ohm versions. Also, note that the velocity of propagation is expressed as a percentage of the speed of light in a vacuum.

Section Lengths

When selecting a rigid line for an application, the section length must be taken into consideration. The length must be chosen such that the frequency of operation does not fall into a VSWR reject band. The section length must not correspond to the half-wavelength or any multiple thereof. Table 15 is a listing of recommended lengths versus TV channel number or FM frequency.

While these are not the only choices possible, they do provide the most operating margin from the VSWR reject bands. Other choices can be used as long as the desired operating frequency is at least 2.0 MHz away from the calculated VSWR spike location.

Peak Power

The peak power rating of a coaxial transmission line is limited by the voltage breakdown (flashover) between the inner and outer conductors and is the maximum RMS power which can be reached in any

TABLE 14

Characteristics for commonly used rigid lines.

Nom. OD of Outer Condr.	Z_C ohms	Maximum Freq. MHz	Vel. of Prop. Percent	Nom. OD of Inner inches	Nom. ID of Outer inches	Net Weight lb/ft
⅞″	50	6000	99.8	.341	.785	0.6
1⅝″	50	3000	99.8	.664	1.527	1.3
3⅛″	50	1588	99.8	1.315	3.027	3.0
4¹⁄₁₆″	50	1197	99.8	1.711	3.935	5.6
6⅛″	50	788	99.8	2.600	5.981	7.3
6⅛″	75	900	99.7	1.711	5.981	6.75
8³⁄₁₆″	75	709	99.7	2.293	8.000	9.0
9³⁄₁₆″	50	530	99.7	3.910	9.000	11.45
9³⁄₁₆″	75	600	99.7	2.580	9.000	11.45

TABLE 15
Recommended rigid line section lengths.

20' Sect.	19' 9" Sect.	19' 6" Sect.	19' Sect.
TV Channels			
2,3,5,6,7,8, 11,12,14,15, 18,19,23,27, 31,35,39,40, 43,44,47,48, 52,56,60,64,68	16,20,24,28, 32,33,36,37, 41,45,49,53, 57,58,61,62, 65,66,69	4,9,10,13,17 21,22,25,26, 29,30,34,38, 42,46,50,51, 54,55,59,63,67	
FM Radio			
88.1–95.9 MHz 100.3–107.9 MHz		96.1–98.3 MHz	98.5–100.1 MHz

modulation interval. At one manufacturer, peak power ratings are determined according to the following equation:

$$P_{PK} = \frac{\left(\frac{.7E_P}{2\sqrt{2}}\right)^2}{Z_C}$$

where: P_{PK} = Peak power rating (standard conditions), watts
E_P = DC production test voltage, volts
0.7 = DC to RF factor
2 = Safety factor on voltage
$\sqrt{2}$ = RMS factor
Z_C = Characteristic impedance, ohms

In this case the production test voltage, E_P, is 35% of the theoretical breakdown voltage for a given line size. Peak power ratings vary between manufacturers depending on the voltage safety factor and test voltage values used.

One manufacturer's peak power ratings for common rigid line sizes are given in Table 16 and correspond to the power rating at 1.0 MHz (AM). These values must be derated for VSWR, modulation, and pressurization.

Peak Power Derating

Peak power ratings can be increased by pressurization and/or use of high density gases with high dielectric strength (i.e., sulfur hexafluoride and Freon 116). The effects are shown in Fig. 10. The disadvantage, however, is the inherent risk of voltage breakdown in the event of failure of the pressure system.

Peak power derating calculations for modulation techniques and VSWR are the same as for semi-flexible cable, and are given in Table 8.

Average Power

Average power ratings are based on a maximum inner conductor operating temperature generated from losses in the conductors. Table 16 lists the average power ratings of rigid lines. These are based on an inner conductor temperature of 216°F (102°C), an ambient temperature of 104°F (40°C), a terminating VSWR of 1.0 and dry air atmospheric pressure. When engineering for a specific application, average power must be derated for higher ambient temperature, expected termination VSWR and type of modulation.

Average Power Derating

The average power rating must be adjusted according to the actual ambient temperature since the ambient will influence the maximum temperature of the inner conductors. Fig. 24 shows the variation of average power with ambient temperature.

The average power rating must also be derated to account for VSWR. The derating factor is calculated from the following formula:

$$D.F. = \frac{(VSWR^2 + 1)}{2(VSWR)} + \frac{F^1(VSWR^2 - 1}{2(VSWR)}$$

where F^1 is a factor that varies with frequency and line size and is selected from the appropriate curve of Fig. 25.

The average power of the transmitter must be less than the calculated derated average power rating of the transmission line. As for semi-flexible cable, average power can be calculated for different modulation schemes using Table 5.

Attenuation

The attenuation of a transmission line is an expression of the ratio of input power to output power in decibels (dB). Table 17 is a listing of attenuation values versus carrier frequency based on an ambient temperature of 75°F (24°C) and unity VSWR. The line attenuation must be derated for ambient temperature and load VSWR.

Attenuation Derating

Fig. 6 shows the relationship between ambient temperature and the attenuation correction factor. The VSWR of the attached antenna (load) also contributes to the total transmission loss of the system. Typically this is very small due to the good match that usually exists between the line and antenna. Fig. 7 shows the increase in loss with load VSWR, assuming a VSWR of 1.0 at the input of the transmission line.

Efficiency

Efficiency is an expression of the power loss for a transmission line and is useful in selecting the appropriate line for an application. It is defined as the ratio of power delivered to the antenna to the input power into the transmission line. It is easily calculated by knowing the corrected attenuation and length of the rigid line and applying the following equation:

$$\text{Efficiency} = \frac{100\%}{10^{\frac{\alpha}{10}}}$$

where: α = Total line attenuation, dB

TABLE 16
Rigid line power ratings assuming unity VSWR, 104°F ambient temperature and atmospheric pressure, dry air.

RIGID LINE POWER RATINGS

POWER, kW

Channel	Frequency (MHz)	7/8" 50 OHM	1-5/8" 50 OHM	3-1/8" 50 OHM	4-1/16" 50 OHM	6-1/8" 50 OHM	6-1/8" 75 OHM	8-3/16" 75 OHM	9-3/16" 50 OHM	9-3/16" 75 OHM
AM*	1.00	52.12	145.00	500.00	750.00	1500.00	1000.00	2606.93	3935.09	3211.68
2	55.25	6.97	20.01	78.72	129.15	271.00	237.09	349.90	528.13	430.97
3	61.25	6.62	19.03	74.69	121.89	257.07	224.45	332.28	501.32	409.26
4	67.25	6.31	18.19	71.22	115.67	245.06	213.58	317.06	478.56	390.52
5	77.25	5.89	17.00	66.36	107.02	228.26	198.42	295.77	446.42	364.28
6	83.25	5.67	16.40	63.88	102.62	219.68	190.69	284.88	429.97	350.86
FM	88.00	5.51	15.96	62.10	99.48	213.53	185.15	277.06	418.17	341.22
FM	98.00	5.22	15.15	58.78	93.65	202.08	174.87	262.49	396.16	323.28
FM	108.00	4.97	14.38	55.94	88.68	191.36	165.76	250.00	377.33	307.89
7	175.25	3.89	11.11	43.71	67.60	144.74	126.26	196.05	295.89	241.43
8	181.25	3.83	10.92	42.97	66.33	141.92	123.83	192.76	290.93	237.37
9	187.25	3.77	10.75	42.26	65.13	139.25	121.53	189.64	286.21	233.52
10	193.25	3.71	10.58	41.59	63.99	136.71	119.34	186.65	287.71	229.85
11	199.25	3.65	10.42	40.95	62.90	134.29	117.25	183.81	277.41	226.34
12	205.25	3.60	10.26	40.33	61.86	131.91	115.13	181.09	273.31	222.99
13	211.25	3.54	10.11	39.75	60.87	129.62	113.08	178.48	269.38	219.78
14	471.25	2.36	6.63	26.41	38.81	79.42	68.76	119.19	179.87	146.73
15	477.25	2.34	6.58	26.24	38.54	78.79	68.23	118.43	178.73	145.80
16	483.25	2.33	6.54	26.07	38.27	78.18	67.71	117.69	177.61	144.85
17	489.25	2.31	6.50	25.91	38.00	77.58	67.20	116.96	176.51	143.98
18	495.25	2.30	6.46	25.75	37.75	76.99	66.71	116.25	175.43	143.10
19	501.25	2.28	6.41	25.59	37.49	76.41	66.22	115.54	174.36	142.24
20	507.25	2.27	6.37	25.44	37.24	75.84	65.74	114.85	173.32	141.38
21	513.25	2.26	6.33	25.29	37.00	75.29	65.27	114.17	172.30	140.55
22	519.25	2.24	6.30	25.14	36.76	74.74	64.81	113.51	171.29	139.73
23	525.25	2.23	6.26	24.99	36.52	74.21	64.35	112.85	170.30	138.92
24	531.25	2.22	6.22	24.85	36.29	73.68	63.91	112.21		138.12
25	537.25	2.20	6.18	24.70	36.06	73.17	63.47	111.57		137.34
26	543.25	2.19	6.15	24.57	35.84	72.66	63.04	110.95		136.58
27	549.25	2.18	6.11	24.43	35.62	72.16	62.62	110.34		135.82
28	555.25	2.17	6.07	24.29	35.40	71.67	62.21	109.73		135.08
29	561.25	2.16	6.04	24.16	35.19	71.19	61.80	109.14		134.35
30	567.25	2.14	6.01	24.03	34.98	70.72	61.40	108.56		133.63
31	573.25	2.13	5.97	23.90	34.77	70.26	61.01	107.98		132.92
32	579.25	2.12	5.94	23.78	34.57	69.80	60.62	107.42		132.22
33	585.25	2.11	5.91	23.65	34.37	69.35	60.24	106.86		131.54
34	591.25	2.10	5.87	23.53	34.17	68.91	59.87	106.31		130.86
35	597.25	2.09	5.84	23.41	33.98	68.48	59.50	105.77		130.20
36	603.25	2.08	5.81	23.29	33.79	68.05	59.13	105.24		
37	609.25	2.07	5.78	23.17	33.60	67.63	58.78	104.72		
38	615.25	2.06	5.75	23.06	33.42	67.21	58.43	104.20		
39	621.25	2.05	5.72	22.94	33.24	66.81	58.08	103.69		
40	627.25	2.04	5.69	22.83	33.06	66.41	57.74	103.19		
41	633.25	2.03	5.66	22.72	32.88	66.01	57.41	102.69		
42	639.25	2.02	5.64	22.61	32.71	65.62	57.08	102.21		
43	645.25	2.01	5.61	22.50	32.54	65.24	56.75	101.73		
44	651.25	2.00	5.58	22.40	32.37	64.86	56.43	101.25		
45	657.25	1.99	5.55	22.29	32.21	64.49	56.12	100.78		
46	663.25	1.98	5.53	22.19	32.04	64.13	55.81	100.32		
47	669.25	1.97	5.50	22.09	31.88	63.77	55.50	99.87		
48	675.25	1.96	5.47	21.99	31.72	63.41	55.20	99.42		
49	681.25	1.95	5.45	21.89	31.56	63.06	54.90	98.98		
50	687.25	1.94	5.42	21.79	31.41	62.72	54.61	98.54		
51	693.25	1.94	5.40	21.70	31.26	62.38	54.32	98.11		
52	699.25	1.93	5.37	21.60	31.11	62.04	54.04	97.68		
53	705.25	1.92	5.35	21.51	30.96	61.80	53.74	97.26		
54	711.25	1.91	5.32	21.42	30.81	61.58	53.45	96.85		
55	717.25	1.90	5.30	21.32	30.67	61.36	53.17			
56	723.25	1.89	5.27	21.23	30.52	61.14	52.89			
57	729.25	1.89	5.25	21.14	30.38	60.92	52.61			
58	735.25	1.88	5.22	21.06	30.24	60.71	52.33			
59	741.25	1.87	5.20	20.97	30.10	60.50	52.06			
60	747.25	1.86	5.18	20.88	29.97	60.29	51.80			
61	753.25	1.86	5.15	20.80	29.83	60.08	51.53			
62	759.25	1.85	5.13	20.71	29.70	59.88	51.27			
63	765.25	1.84	5.11	20.63	29.57	59.68	51.02			
64	771.25	1.83	5.09	20.55	29.44	59.48	50.76			
65	777.25	1.83	5.06	20.47	29.31	59.28	50.51			
66	783.25	1.82	5.04	20.39	29.19	59.08	50.26			
67	789.25	1.81	5.02	20.31	29.06	58.89	50.02			
68	795.25	1.80	5.00	20.23	28.94	58.70	49.78			
69	801.25	1.80	4.98	20.15	28.82	58.51	49.54			

* Peak Power Rating Value

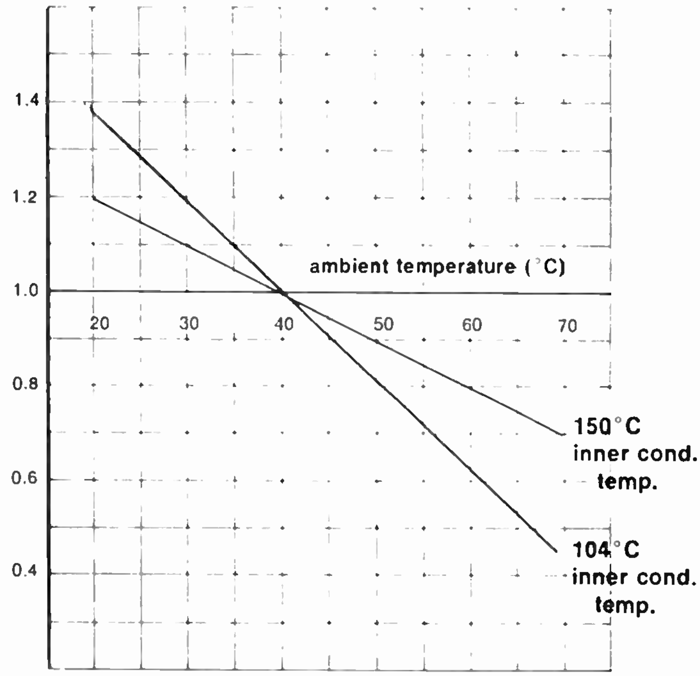


Figure 24. Variation of average power with ambient temperature.

VSWR

Rigid lines offer generally lower VSWR specifications than semi-flexible lines, over a single channel bandwidth. A good rigid line will have a maximum VSWR of 1.02 per component and 1.07 for a complete run (not including antenna). Fine matchers used at the input and output of the line can further improve the VSWR at particular locations in the operating band.

Selection Example

Following is an example of rigid line selection for a UHF-TV broadcast application. For the station characteristics given, a group of acceptable rigid lines are selected.

Station Characteristics.

Channel	56
Frequency Band	722 to 728 MHz
Video Carrier Frequency	723.25 MHz
Transmitter Power (Peak Sync.)	60 kW
Aural Power, 20%	12 kW
Modulation	TV
Line Length	600 ft
Load/Antenna VSWR (Inflated)	1.25
Average Ambient Temperature	68°F (20°C)

Cut-Off Frequency (Table 14). Cut-off frequency (f_C) of line must be greater than 728 MHz (top of channel band). Line sizes ⅞" to 6⅛" meet this requirement.

Peak Power Derating for VSWR and Modulation (Tables 8 & 16).

$$P_{PK} > P_{TRANSMITTER}(1 + AU + 2\sqrt{AU})\ VSWR$$
$$AU = .2\ (\textit{aural is 20\% peak sync.})$$
$$P_{PK} > 60\ kW\,(1 + .2 + 2\sqrt{.2})\ 1.25$$
$$P_{PK} > 60\ kW\,(2.09)\ 1.25$$
$$P_{PK} > 157\ kW$$

Peak power rating (P_{PK}) of line must be greater than 157 kW.

Line sizes 3⅛" to 9³⁄₁₆" meet this requirement.

This leaves a preliminary acceptable range of 3⅛" to 6⅛" rigid line sizes.

Transmitter Average Power Calculation (Table 5).

$$P_{AVG} = 0.8\ (P_{TV}) = 0.8\ (60\ kW) = 48\ kW$$

Rigid Line Average Power Rating Adjustments for Actual Ambient Temperature and Antenna VSWR.

Line Size Line Impedance	3⅛" 50Ω	4¹⁄₁₆" 50Ω	6⅛" 50Ω	6⅛" 75Ω
Rigid Line Average Power Rating (Table 17)	21.23 kW	30.52 kW	61.14 kW	52.89 kW
Adjustment for 68°F (20°C) Ambient (Fig. 25) x 1.37	29.09 kW	41.81 kW	83.76 kW	72.46 kW
Derating for Antenna VSWR F' Factor (Fig. 26)	.01	.01	.01	.01
Derating Factor Calculation:				
Adjusted Average Power ÷ 1.027	28.32 kW	40.71 kW	81.56 kW	70.55 kW

Derating Factor Calculation:

$$D.F. = \frac{(VSWR^2 + 1)}{2\,(VSWR)} + \frac{F'(VSWR^2 - 1)}{2\,(VSWR)}$$
$$= \frac{(1.25^2 + 1)}{2\,(1.25)} + \frac{.01\,(1.25^2 - 1)}{2\,(1.25)}$$
$$= 1.027$$

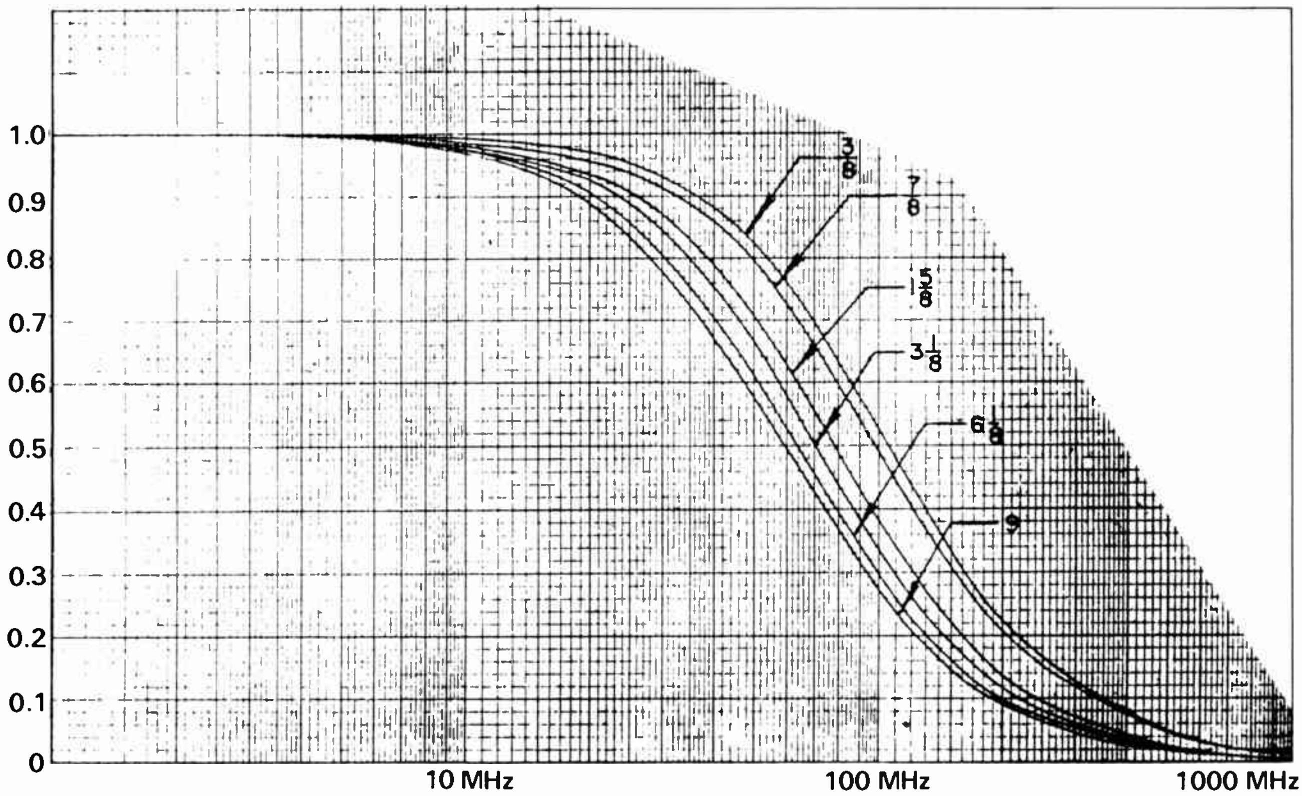


Figure 25. Derating factor for average power due to load VSWR.

Only the 6⅛" 50-ohm and 75-ohm rigid lines meet *all* of the following requirements:

1. Adjusted average power ratings greater than the effective transmitter average power level of 48 kW.
2. Peak power ratings greater than 157 kW.
3. Cut-off frequencies greater than 728 MHz.

Rigid Line Attenuation Adjustments for Actual Ambient Temperature and Load VSWR Effects and Efficiency Calculations.

Line Size Line Impedance	6⅛" 50Ω	6⅛" 75Ω
Total Line Attenuation (Table 18), x 6	.160 dB x 6 = .960 dB	.146 dB x 6 = .876 dB
Adjustment for 68°F (20°C) Ambient (Fig. 6), x .99	.960 dB x .99 = .950 dB	.876 dB x .99 = .867 dB
Derating for Load VSWR (Fig. 7), + .02	.950 dB + .02 = .970 dB	.867 dB + .02 = .887 dB
Efficiency Calc. $Eff = \frac{100\%}{10^{\frac{\alpha}{10}}}$	$\frac{100\%}{10^{\frac{.970}{10}}}$ = 79.98%	$\frac{100\%}{10^{\frac{.887}{10}}}$ = 81.52%

Therefore, we can conclude that either 50-ohm or 75-ohm 6⅛" rigid coaxial line will work for this application. The performance of a system for both types of line is as follows:

	System Req.	50Ω Line	75Ω Line
Cut-Off Frequency, MHz	>728	788	900
Peak Power, Derated, kW	>157	1500	1000
Average Power, Derated, kW	>48	81.56	70.55
Total Attenuation, Derated, dB		.970	.887
Line Efficiency, %		79.98	81.52

Both lines satisfy the peak power, average power, and frequency cut-off requirements, but the 75-ohm line will deliver about 2% more power to the antenna with an average weight reduction over the 50-ohm line of approximately 7.5%.

Rigid Line Systems

General

A rigid coaxial transmission line system is composed of an interior run and exterior horizontal/vertical runs. The exterior runs should be pressurized to prevent

TABLE 17
Rigid line attenuation assuming unity VSWR and 75°F ambient temperature.

RIGID LINE ATTENUATION RATINGS

ATTENUATION dB/100 FT.

Channel	Frequency (MHz)	7/8" 50 OHM	1-5/8" 50 OHM	3-1/8" 50 OHM	4-1/16" 50 OHM	6-1/8" 50 OHM	6-1/8" 75 OHM	8-3/16" 75 OHM	9-3/16" 50 OHM	9-3/16" 75 OHM
AM	1.00	0.037	0.017	0.010	0.007	0.005	0.005	0.003	0.003	0.003
2	55.25	0.277	0.142	0.076	0.053	0.037	0.034	0.025	0.024	0.023
3	61.25	0.292	0.150	0.080	0.056	0.039	0.036	0.027	0.026	0.024
4	67.25	0.306	0.158	0.084	0.058	0.041	0.038	0.028	0.027	0.025
5	77.25	0.328	0.170	0.090	0.063	0.044	0.041	0.030	0.029	0.027
6	83.25	0.340	0.177	0.094	0.065	0.045	0.042	0.031	0.030	0.028
FM	88.00	0.350	0.182	0.097	0.067	0.047	0.043	0.032	0.031	0.029
FM	98.00	0.370	0.193	0.102	0.070	0.049	0.046	0.034	0.033	0.031
FM	108.00	0.388	0.203	0.107	0.074	0.052	0.048	0.036	0.035	0.032
7	175.25	0.496	0.261	0.137	0.094	0.068	0.063	0.046	0.044	0.042
8	181.25	0.504	0.266	0.139	0.096	0.069	0.064	0.047	0.045	0.043
9	187.25	0.512	0.270	0.142	0.098	0.070	0.066	0.047	0.046	0.043
10	193.25	0.521	0.275	0.144	0.099	0.072	0.067	0.048	0.047	0.044
11	199.25	0.529	0.279	0.146	0.101	0.073	0.068	0.049	0.048	0.045
12	205.25	0.537	0.284	0.148	0.102	0.074	0.069	0.050	0.048	0.045
13	211.25	0.545	0.288	0.151	0.104	0.075	0.070	0.050	0.049	0.046
14	471.25	0.819	0.442	0.226	0.156	0.120	0.112	0.077	0.075	0.071
15	477.25	0.824	0.445	0.228	0.157	0.121	0.113	0.078	0.076	0.071
16	483.25	0.829	0.448	0.229	0.158	0.122	0.114	0.078	0.076	0.072
17	489.25	0.835	0.451	0.231	0.159	0.123	0.115	0.079	0.077	0.072
18	495.25	0.840	0.454	0.232	0.160	0.124	0.116	0.079	0.077	0.073
19	501.25	0.845	0.457	0.234	0.161	0.125	0.117	0.080	0.078	0.073
20	507.25	0.850	0.460	0.235	0.162	0.126	0.118	0.080	0.078	0.074
21	513.25	0.855	0.463	0.236	0.162	0.127	0.119	0.081	0.079	0.074
22	519.25	0.860	0.466	0.238	0.163	0.128	0.119	0.081	0.079	0.075
23	525.25	0.865	0.469	0.239	0.164	0.129	0.120	0.082	0.080	0.075
24	531.25	0.870	0.472	0.241	0.165	0.130	0.121	0.082		0.076
25	537.25	0.875	0.475	0.242	0.166	0.131	0.122	0.083		0.076
26	543.25	0.880	0.477	0.243	0.167	0.132	60.123	0.083		0.076
27	549.25	0.885	0.480	0.245	0.168	0.133	0.124	0.084		0.077
28	555.25	0.890	0.483	0.246	0.169	0.134	0.125	0.084		0.077
29	561.25	0.895	0.486	0.247	0.170	0.135	0.125	0.085		0.078
30	567.25	0.900	0.489	0.249	0.171	0.135	0.126	0.085		0.078
31	573.25	0.905	0.492	0.250	0.172	0.136	0.127	0.086		0.079
32	579.25	0.910	0.494	0.251	0.173	0.137	0.128	0.086		0.079
33	585.25	0.914	0.497	0.253	0.174	0.138	0.129	0.087		0.080
34	591.25	0.919	0.500	0.254	0.174	0.139	0.130	0.087		0.080
35	597.25	0.924	0.503	0.255	0.175	0.140	0.130	0.088		0.080
36	603.25	0.929	0.506	0.257	0.176	0.141	0.131	0.088		
37	609.25	0.933	0.508	0.258	0.177	0.142	0.132	0.089		
38	615.25	0.938	0.511	0.259	0.178	0.143	0.133	0.089		
39	621.25	0.943	0.514	0.260	0.179	0.144	0.134	0.090		
40	627.25	0.947	0.516	0.262	0.180	0.144	0.134	0.090		
41	633.25	0.952	0.519	0.263	0.181	0.145	0.135	0.091		
42	639.25	0.957	0.522	0.264	0.182	0.146	0.136	0.091		
43	645.25	0.961	0.524	0.266	0.182	0.147	0.137	0.092		
44	651.25	0.966	0.527	0.267	0.183	0.148	0.138	0.092		
45	657.25	0.970	0.530	0.268	0.184	0.149	0.138	0.092		
46	663.25	0.975	0.532	0.269	0.185	0.150	0.139	0.093		
47	669.25	0.979	0.535	0.270	0.186	0.150	0.140	0.093		
48	675.25	0.984	0.538	0.272	0.187	0.151	0.141	0.094		
49	681.25	0.988	0.540	0.273	0.187	0.152	0.142	0.094		
50	687.25	0.993	0.543	0.274	0.188	0.153	0.142	0.095		
51	693.25	0.997	0.545	0.275	0.189	0.154	0.143	0.095		
52	699.25	1.001	0.548	0.277	0.190	0.155	0.144	0.096		
53	705.25	1.006	0.551	0.278	0.191	0.156	0.145	0.096		
54	711.25	1.010	0.553	0.279	0.192	0.157	0.145	0.097		
55	717.25	1.015	0.556	0.280	0.192	0.159	0.146			
56	723.25	1.019	0.558	0.281	0.193	0.160	0.146			
57	729.25	1.023	0.561	0.283	0.194	0.161	0.147			
58	735.25	1.028	0.564	0.284	0.195	0.163	0.148			
59	741.25	1.032	0.566	0.285	0.196	0.164	0.148			
60	747.25	1.036	0.569	0.286	0.196	0.165	0.149			
61	753.25	1.040	0.571	0.287	0.197	0.167	0.150			
62	759.25	1.045	0.574	0.288	0.198	0.168	0.150			
63	765.25	1.049	0.576	0.290	0.199	0.170	0.151			
64	771.25	1.053	0.579	0.291	0.200	0.171	0.151			
65	777.25	1.057	0.581	0.292	0.200	0.172	0.152			
66	783.25	1.061	0.584	0.293	0.201	0.174	0.153			
67	789.25	1.066	0.587	0.294	0.202	0.175	0.153			
68	795.25	1.070	0.589	0.295	0.203	0.176	0.154			
69	801.25	1.074	0.592	0.296	0.203	0.178	0.154			

ingress of moisture. The fundamental consideration in a rigid line system is differential thermal expansion between the steel tower and copper rigid line. Over a temperature range of −25°F (−32°C) to +125°F (+52°C), expansion of copper is about 1½"/100 ft and steel is about 1"/100 ft. This results in a differential expansion of ½" per 100 ft. Provisions must be made for the line expansion to prevent severe buckling strain on both the transmission line system and tower.

Various hangers are available from rigid line manufacturers that will fasten the line to the tower and horizontal bridge and allow for differential thermal expansion.

Mechanical Interfaces

Interface standards have been established by the Electronic Industries Association (EIA): RS-225, for 50-ohm rigid coaxial lines, and RS-259, for 75-ohm rigid coaxial lines. These standards specify line and flange dimensions and tolerances for most of the available line sizes.

Components are available from line manufacturers that permit the correct interface connection of different line sizes and impedances.

Coaxial Components

Table 18 is a list of typical components used in a rigid line system as depicted in Fig. 26.

TABLE 18

Typical rigid line components and uses.

It.	Component	Application
1	19–20 ft. Line Section	Construct horizontal/vertical run.
2	90° Elbow	Change direction of run.
3	Gas Barrier	Access port to connect pressure system. Isolate pressurized sections of line (e.g. exterior/interior line interface).
4	Adaptor	Connect two components together of different size but same impedance.
	Impedance Transformer	Connect components together of different impedances (e.g. 50-ohm patch panel to 75-ohm line).
	Fine Matcher	Fine tune diplexer output to line and line to antenna.
	Field Flange Kits	Used with field cut line section to obtain a special length line to complete a run.
8	Rigid Hanger	Support line rigidly at top of tower.
9	Vertical Spring Hanger	Support weight of line along tower and accommodate expansion.
10	Lateral Brace	Restrict lateral movement of line at the bottom of tower due to forces created by expansion of vertical and horizontal runs.
11	3-Point Suspension Hanger	Allow vertical movement of rigid line in horizontal run as vertical run varies in length due to thermal effects.
12	Wall Feed-Thru	Provide weather tight access thru building wall.
13	Dehydrator	Maintain positive pressure in line.

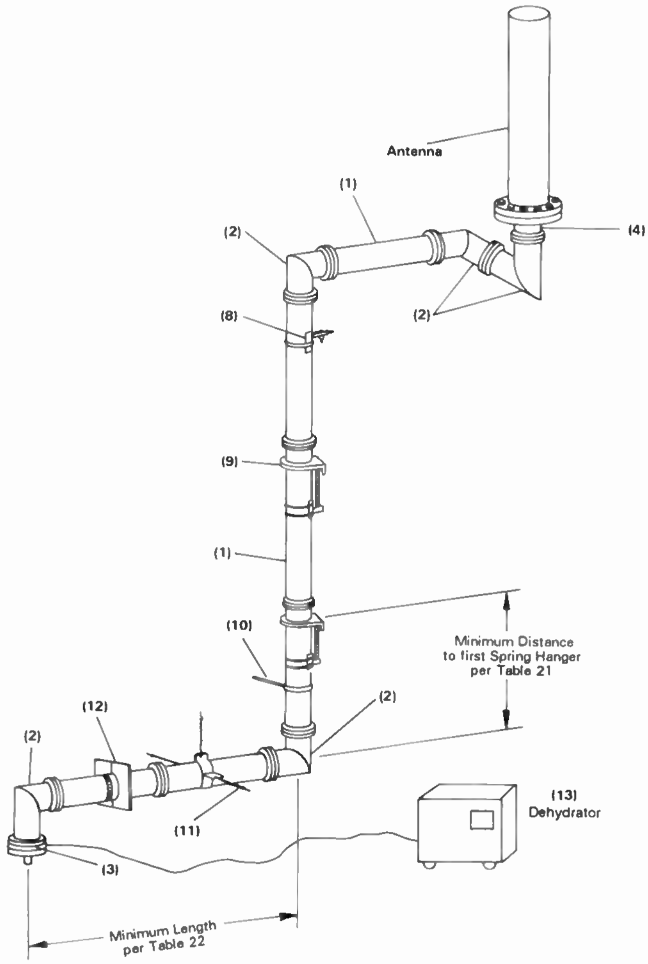


Figure 26. Typical broadcast rigid line system.

Coaxial line components should be supplied with 1 inner connector, 1 hardware kit, 1 "O" ring gasket, and 2 protective end caps. Hangers should be supplied with mounting hardware unless otherwise stated.

Installation

Rigid transmission line installations may begin at either the top or bottom of the tower. However done, they require proper positioning of the bottom elbow to allow for both expansion and contraction of the rigid line. In addition, a special length section is often required, and an "elbow complex" may be needed to allow alignment with the antenna.

Vertical Run

For installations originating at the top of the tower, the following general steps should be followed:

1. Use appropriate adaptor and/or transformer to adapt the antenna feed (size and impedance) to the line.
2. Field cut lengths and multiple elbows can be used to obtain the required orientation of the line to the tower and antenna.

3. Use reinforced elbows for larger line sizes (4¹⁄₁₆" and greater).
4. Mount the first vertical section to the tower using rigid hangers. The number will correspond to the size of the line and length of vertical run.
5. Use vertical spring hangers every ten feet to mount the vertical run to the tower. Tower steel must provide support for a straight vertical run and maintain clearance for movement. Adjust spring hangers to their proper setting.
6. Use a field cut section at the bottom of the tower to bring the vertical run down to the desired height for attachment to the horizontal run.
7. Set the distance from the horizontal run to the lowest vertical spring hanger according to Table 19.
8. Use a lateral brace near the bottom to restrict lateral motion while permitting vertical and horizontal movement.
9. Use 90° reinforced elbows to change direction from vertical to horizontal, or to obtain a nonhorizontal angle of elevation for the run to the transmitter building.

For installations originating at the bottom of the tower the above steps should be followed, except temporary rigid hangers should be used on the first vertical line section installed. The temporary rigid hangers must be removed after the permanent rigid hangers have been placed on the top vertical section to prevent damage due to differential expansion. The position of the bottom elbows must be carefully determined based on current ambient conditions to provide clearance for the expected expansion of the vertical run without interference from tower members. Field cut sections and elbows will be needed at the top of the vertical run to obtain the desired height and orientation for connection to the antenna.

Horizontal Run

Three point suspension hangers are used every ten feet to support the rigid line to the horizontal support structure. It is recommended that the minimum length of horizontal run shown in Table 20 be used to avoid overstress from thermal expansion. The horizontal run should enter the transmitter building through a wall feed-thru with a gas barrier used inside the building to allow pressurization of the exterior run.

TABLE 19

Minimum distance to first vertical spring hanger versus length of horizontal run (refer to Fig. 26).

Horizontal Run (feet)	Minimum Distance to Lowest Hanger (feet) 3⅛"	4¹⁄₁₆"	6⅛"	8³⁄₁₆"	9³⁄₁₆"
20	5	6	9	12	13
40	6	7	11	15	17
60	7	8	13	17	20
80	8	9	14	19	22
100	9	10	16	21	23

TABLE 20

Recommended length of horizontal run.

Vertical Run (feet)	Recommended Lengtyh of Horizontal Run (feet) 3⅛"	4¹⁄₁₆"	6⅛"	8³⁄₁₆"	9³⁄₁₆"
100	15	15	15	20	20
500	25	30	35	40	40
1000	35	40	50	60	60
1500	40	50	60	70	70
2000	45	60	70	80	80

Do's & Don'ts

Installation "do's" and "don'ts" are provided in Fig. 27.

Pressurization

General

External rigid coaxial transmission lines must be maintained under dry gas, nitrogen, or dry air, to prevent electrical performance degradation and possible transmission system failure. If a positive dry air pressure is not maintained inside the lines, "breathing" can occur—allowing moist air and dust to enter the line through the joints. This moisture will condense and accumulate on the internal surfaces of the transmission line system causing oxidation and corrosion, resulting in:

- Increase in line attenuation from oxidation and moisture
- Increase in VSWR from moisture accumulation
- Corrosion in mechanical joints causing localized I^2R losses and hot spots
- Voltage flashover caused by dust and other contaminates, left on the insulators

Transmission lines pressurized with dry gas maintain the high initial performance of the transmission system, but also reduce the risk of damage and costly interruption of service. Dry gas may be obtained from either nitrogen bottle or a dehydrated air system. Pressurization values must not exceed the lowest component pressure rating in the transmission system (usually the antenna).

Selection

Table 12 summarizes the advantages and disadvantages of four pressurization systems: nitrogen, manual regenerative dehydrator, automatic regenerative dehydrator, and automatic continuous membrane dehydrator.

When selecting a pressurization system, the transmission system volume, normal leakage rate, ambient temperature swing, solar load changes and system advantages/disadvantages must be considered to ensure proper operation.

Installation "Do's"

1. Do lift components with connector end up unless marked otherwise.
2. Do check the spring hangers for free travel and proper alignment.
3. Do consult manufacturers' spring loading charts.
4. Do tighten flange bolts alternately, one side, then the other, before torquing.
5. Do check to make sure O-rings are fully seated in groove.
6. Do pressurize the line immediately following installation and maintain a positive pressure at all times.
7. Do use reinforced elbows for exterior installations of larger size coax (4 1/16" and larger).
8. Do perform electrical testing before applying power.
9. Do purge and pressurize line before applying power.
10. Do remove all burrs from field cut conductors.

Installation "Don'ts"

1. Don't hoist coupled sections of line.
2. Don't force components together.
3. Don't assemble line components that contain water or condensation.
4. Don't install line that exhibits any evidence of damage.
5. Don't dismiss rigger until line is completely installed, holds pressure, and electrical testing has been performed.
6. Don't apply power to line until it has been purged and pressurized.
7. Don't use damaged O-rings.
8. Don't assemble a horizontal run without support.
9. Don't exceed pressure rating of any component (antenna, gas barrier, etc.).
10. Don't exceed torque specifications on flange hardware.

Figure 27. Rigid line installation "do's" and "don'ts".

Purging

System purging must be done after any of the following conditions have occurred:

1. When a line has been installed and before power is applied.
2. When a pressure seal has been broken on the outside run of the transmission line for maintenance work.
3. When positive pressure has not been maintained in the line.

The line must be purged with a minimum of three volumes of air to ensure dryness to the desired dew point. This is accomplished by opening a joint or using an available purge valve near the antenna to allow the gas to flow through the entire length of line, drawing moisture out with it.

Testing

General

Electrical testing of rigid line systems should be performed for any of the following conditions:

1. After initial installation (prior to applying power).
2. After VSWR/reverse power trips to determine mode and location of failure.
3. Periodically, to verify performance of system.
4. After persistent signal fluctuations of "in-line" waveform monitor and signal analyzer.

It is advantageous to test when installation personnel are still on site. Electrical testing is usually done to verify the performance of the transmission system. Thorough testing done at this stage will both verify initial system performance and also detect and locate most "bugs" in a system that could lead to eventual failures. These minor defects usually occur during installation.

The most useful tests are the VSWR/return loss sweep test and the RF pulse test. These tests operate over the bandwidth of the system and give a good indication of line and antenna performance. Network analyzers with time domain capabilities and lightweight line analyzers can be used for overall system performance checks, and are also excellent at locating discontinuities (not detectable by RF pulse) which may lead to future line failure. Traditional time-domain reflectometry devices (TDR) can be used to locate these discontinuities, but will not give an indication of performance over the channel bandwidth. This section reviews each technique and presents actual data taken on transmission line systems.

Sweep Test

The sweep test is a frequency domain test that measures the return loss of the line in the channel bandwidth. This test should be performed before power is applied to the line, and must be done with the end of the line terminated in either an antenna or load. The effects of the terminating device will be included in the measurement.

A return loss sweep test curve of a UHF transmission system is shown in Fig. 28. The system consists of 380 ft of 6 1/8" 75-ohm rigid line, various components (elbows, gas barriers, adapters, etc.), and a top mounted antenna. The maximum return loss level is about 20 dB (1.22 VSWR) at the top of the band. The rippling effect in the trace arises from the continuous phasing of the reflection from the ends of the system.

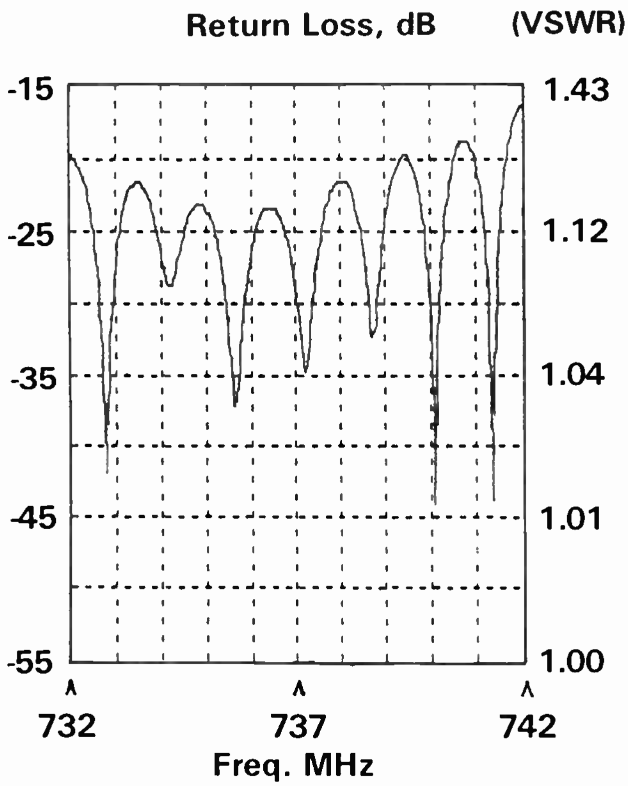


Figure 28. Swept return loss of a 380 ft, 6⅛″ 75-ohm rigid line system including antenna. (VSWR of this system is higher than normally expected.)

The return loss level is high and we should be able to determine the location of the problem using one of the time domain techniques.

RF Pulse Test

The RF pulse test displays the performance of the line in the time domain. This test shows impedance changes versus time and allows the user to correlate relative electrical performance to general location in the line.

The procedure is to apply a CW burst to the line and observe the reflected signal in time. The frequency of the CW signal is equal to the center frequency of the channel bandwidth and the duration of the burst is of the order of inverse of the channel bandwidth. This technique allows the performance of the transmission system over the channel bandwidth to be displayed in time. The disadvantage of this approach is that information over a only a small bandwidth is used to derive the time domain characteristics. This means that the display will show only gross defects in the system. An RF pulse test curve for 1,000 ft of 6⅛″ 75-ohm rigid line and antenna for channel 14 is shown in Fig. 29. An accurate conversion of 2.1 nanoseconds per foot of line can be applied showing that 1 division on the horizontal axis is equal to 238 ft of line. This conversion takes into consideration the round trip time of the pulse. The return loss over the length of the line is very good with the worst spike of −43 dB at approximately 666 ft.

Network Analyzers

Network analyzers with time domain capabilities can be used for both sweep testing and to simulate RF pulse and TDR testing. The advantage of this equipment is that one unit contains all the necessary transmission testing functions, and the time domain functions allow excellent control over the resolution. The time domain can be used to simulate an RF pulse by looking over only a small frequency range (channel bandwidth) or over the entire range of frequencies up to line cut-off. Using the equipment up to cut-off gives excellent resolution and allows smooth impedance variations to be displayed in time/length. Therefore, the analyzer can be used to locate problems that are not seen by an RF pulse but which could lead to eventual failures. The disadvantage of a network analyzer is cost.

An impedance plot of a transmission system using a network analyzer is shown in Fig. 30. This plot was taken for troubleshooting after multiple VSWR trips had occurred. The time domain information was automatically calculated from data taken in the frequency domain up to 900 MHz. In this mode the individual

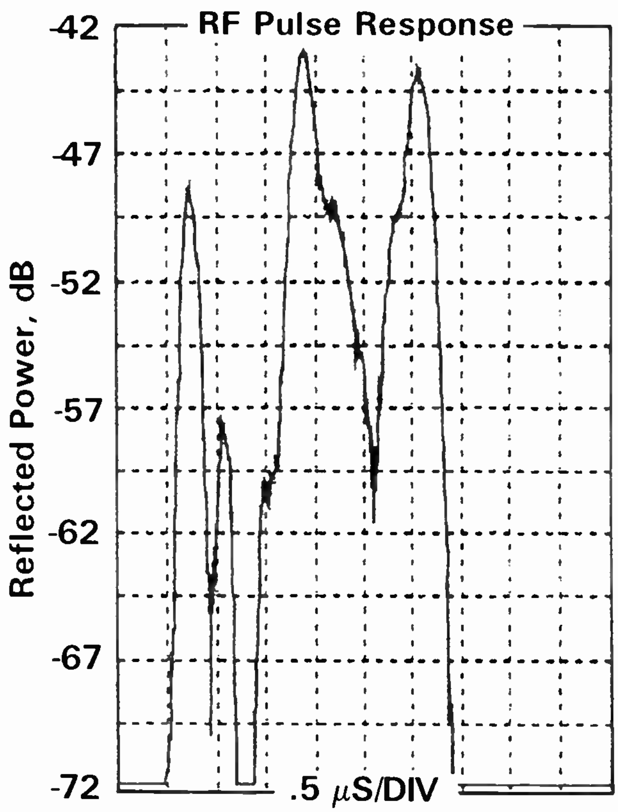


Figure 29. RF pulse test response of a 1,000 ft. 6⅛″ 75-ohm rigid line system and antenna.

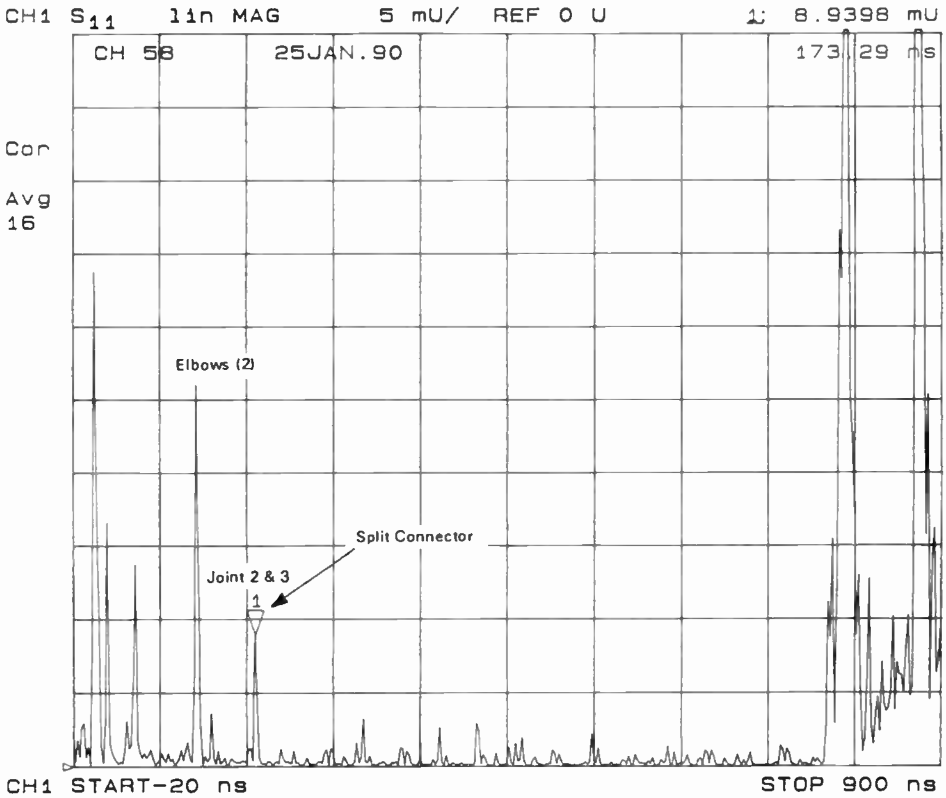


Figure 30. Time domain impedance plot of rigid line.

line components can be observed (adapters, elbows, flanges, insulators, and antenna). In Fig. 30 an impedance spike is seen in the vertical run of rigid line (this turned out to be a split inner connector). Fig. 31 is an impedance plot using only a small amount of frequency information (734 MHz to 740 MHz) to simulate an RF pulse. This test indicated a high impedance near the input to the antenna and around the horizontal-to-vertical junction at the tower base. The former turned out to be a bent inner conductor at the antenna input; the latter was an improperly tuned elbow used to connect the vertical to horizontal run. After replacing these items, the trace in Fig. 32 shows an improvement of approximately 11 dB at both locations.

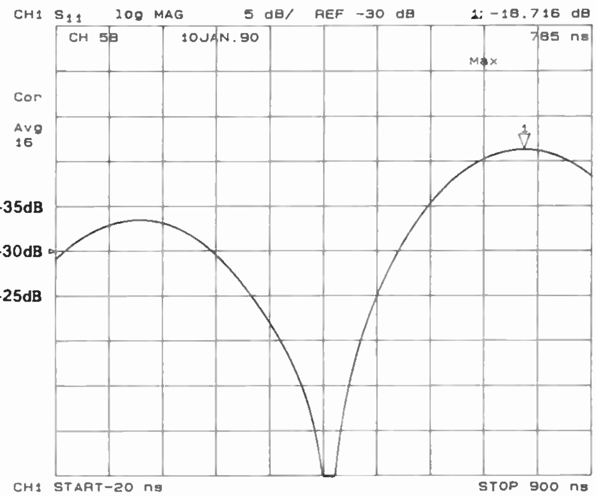


Figure 31. Time domain impedance plot before fix.

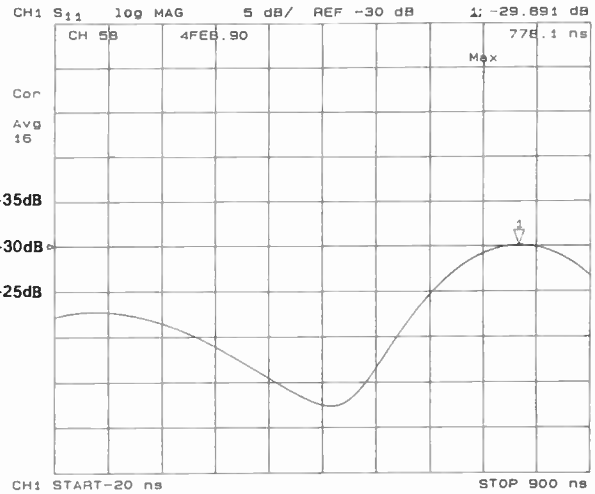


Figure 32. Time domain impedance plot after fix.

All of these tests were performed using an HP 8753C Network Analyzer with time domain analysis. This unit is excellent for performing complete system verification and debugging. A Systron Donner Model 5220 Transline Analyzer, which is made for field use, could also have been used.

Time-Domain Reflectometry (TDR) Test

A TDR (time-domain reflectometry) test can be used to examine a transmission line for faults. The TDR test set uses a fast rise time voltage step transmitted down the line, then measures the reflected signal. Since the signal will have significant frequency components up to 0.35 divided by the rise time of the pulse, the important thing to remember is to adjust the pulse rise time on the TDR test set to the cut-off frequency of the line. Because the TDR test set is a broadband device it cannot be used to determine the performance of the line over the channel bandwidth.

Insulation Resistance Test

An insulation resistance test can be performed on a rigid transmission line after purging with dry air to determine if it is sufficiently dry before applying power. This test should be performed after any line has suffered from moisture ingress. The insulation resistance between the inner and outer conductor should be greater than 100,000 Megohms. There are many inexpensive units available that can measure insulation resistance.

Operation and Maintenance

The following guidelines should be followed to ensure proper operation of a rigid coaxial transmission line system:

1. Establish a swept VSWR and time domain response of the transmission system before applying power.
2. Set and maintain the reverse power VSWR trip to an acceptable level for the method of broadcasting being used.
3. If possible, use an in-line power probe and meter to monitor actual forward and reverse power in the line and to compare with the transmitter meters.
4. Determine the cause of VSWR trips before overriding the system.
5. Always maintain a positive pressure in the line.
6. Do not exceed the pressure rating of the antenna or any other component in the transmission system.
7. Establish and record the dehydrator cycle pattern.
8. Record dehydrator cycle times and compare to the established pattern. Deviation from the established pattern may indicate a problem in the system.
9. Perform periodic maintenance checks on the pressurization system.
10. Always purge the line after an outside seal has been broken or when a positive pressure has not been maintained in the line.
11. Flanges and fittings can be examined for leaks using a soapy water solution.
12. Periodically feel for hot spots along the rigid line. The outer temperature is proportional to the inner temperature and excess outer temperatures may indicate an inner joint heating up.
13. Check the vertical and horizontal line sections for wear caused by thermal expansion and contraction.
14. Monitor the on-air signal and waveform monitor. Fluctuations in the signal could be the result of intermittent contact in the line which could lead to line failure.
15. Remeasure system VSWR and impedance periodically and compare to previous results to detect signs of deterioration.

WAVEGUIDE

General

Waveguide, as the name implies, "guides" the electric and magnetic fields of the travelling wave within its enclosed space. For UHF broadcast applications, the common forms of waveguide to be discussed in this section include (Fig. 33):

- Rectangular; generally aluminum with a width to height ratio of 2 to 1, ranging in size from 11.5" x 5.75" (WR1150) to 18" x 9" (WR1800).
- Circular; also aluminum with diameters ranging from approximately 13" to 18". A recent improvement provides single polarized operation by eliminating the cross polarized TE_{11} field (Fig. 33C).
- Truncated; an aluminum hybrid of the above consisting of a special interior shape (between circular and rectangular) surrounded by a cylindrical cover to reduce wind loading on the tower. Final diameters are comparable to those of a circular waveguide.

Because of the larger sizes which would be necessary at lower frequencies, waveguide is rarely used for broadcast applications other than UHF television.

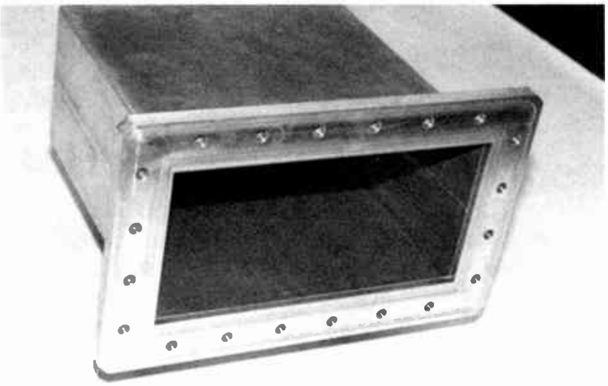
Basic Considerations

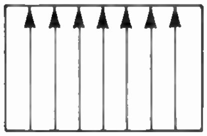
There are a number of important elements to be considered when evaluating the use of waveguide in a UHF broadcast application.
Following are descriptions of the most important.

Lower Attenuation

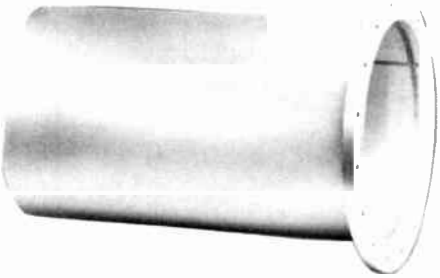
The absence of dielectric support materials together with propagation at waveguide modes reduces attenuation to values well below that of large rigid coaxial lines (Figs. 34 and 35). The lower attenuation of waveguide translates to high efficiency according to the formula:

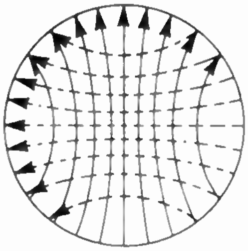
$$Eff = \frac{100\%}{10^{\left(\frac{dB}{10}\right)}}$$



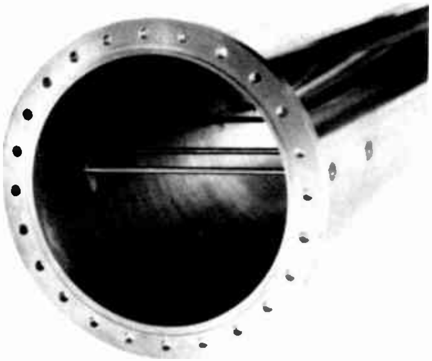


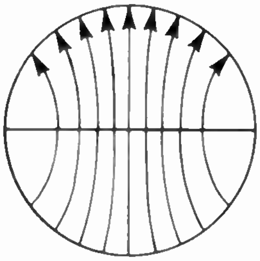
A) Rectangular



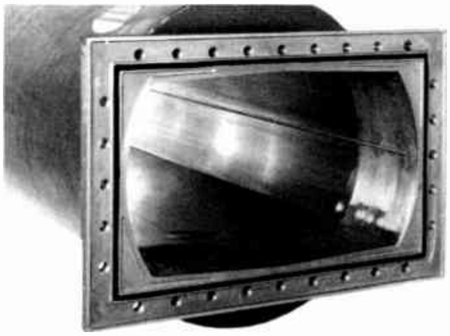


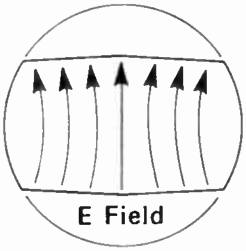
B) Circular





C) Circular with Cross-Polarization
Cancelling Rods





D) Truncated

Figure 33. Common broadcast waveguide types.

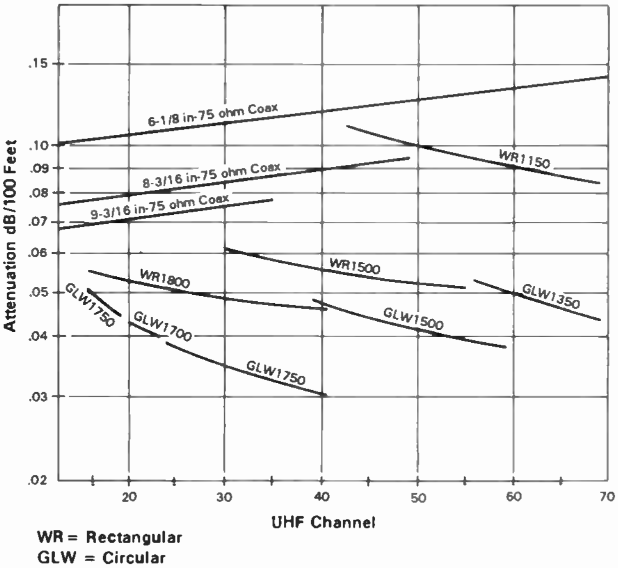


Figure 34. Relative attenuation of waveguide compared to large rigid line.

where dB is the total attenuation of the transmission line at the frequency of interest. The power lost in the line is dissipated as heat.

Table 21 lists the efficiencies of various transmission lines 2,000 ft long at channel 48. Manufacturers should be consulted to confirm attenuation/efficiency figures for actual system configurations.

Lower Operating Costs

Because of the increased efficiency compared to rigid line, desired ERP levels can be achieved with less transmit power; thereby reducing electricity bills. Table 21 shows the savings compared to 8³⁄₁₆″ rigid line assuming 20 hr/day transmission using a 50% efficient 120 kW transmitter with an electric rate of $0.10/kWH. Obviously, even higher savings would result if transmitter efficiency is less than 50%.

As a general rule, waveguide is more efficient than coaxial rigid line, with circular waveguide being the most efficient.

Possibility of Both Horizontal and Vertical Transmission

With increased efficiency of transmission, enough power may be available at the antenna to allow both full ERP in the horizontal polarization plane and additional ERP in the vertical polarization plane. In urban areas 5 to 20 miles from the tower, this vertically polarized signal may serve to improve reception for viewers using indoor antennas.

High Power Handling Capability

The single tube construction of waveguide provides high power handling capacity. There is no inner conductor to heat up, no dielectric support material to soften, and a large surface area over which the I^2R losses can be dissipated. Applications of 240 kW are common. Waveguide is the only practical way to deliver such power levels to the antenna.

Pressurization Requirements Must Be Carefully Evaluated

The large internal volume of a waveguide system should be pressurized. If left unpressurized, moist air may "breathe" in between the flanges under conditions of rapid barometric pressure or temperature change. If the pressure outside increases suddenly (i.e., quickly-moving warm air front), the interior volume of the waveguide acts like a vacuum and immediately tries to achieve pressure equal to the surrounding environment. Once stabilized with ambient air, the moisture now contained within the waveguide can, under dropping temperature conditions, condense and form water droplets. Gravity causes the droplets to accumulate at the bottom of the line, usually at the location of the bend between the vertical and horizontal runs. This moisture can, in some cases, act as a load and absorb a significant amount of transmitted power, thereby reducing efficiency. Alternately, it can raise VSWR high enough to cause transmitter shutdown.

Higher Wind Load

The larger size of waveguide causes more wind load on the tower. This limits the use of waveguide on many existing towers originally designed to support rigid lines only half as large in diameter.

Larger Elbow Complexes

Due to the physical size and electric field (E-field) orientation of waveguide, it is mechanically more complex to route and change direction of the transmission line system since there are no swivel flanges to facilitate compound bends. Consequently, routing waveguide in and around tower members may involve numerous components. Circular pin twist sections

TABLE 21

Annual savings on operating costs of waveguide compared to large rigid line for 2000 ft. system on channel 48.

Line Type	Circular GLW1500	Truncated DTW1500	Rectangular WR1500	8³⁄₁₆″ Rigid
Efficiency (%)	82.1	78.4	74.4	64.7
Proposed XMIT Power for 5 MW ERP (kW)	95	99	104	120
XMTR Power Consumption @ 50% Efficiency (kW)	190	198	208	240
Annual Savings @ $0.10/kWH & 20 hr/day Operation	$36,500	$30,660	$23,360	—

		Rectangular			Truncated			Circular		
Channel	VIS Carrier	WR1150	WR1500	WR1800	DTW1350	DTW1500	DTW1750	GLW1350	GLW1500	GLW1750
14	471.25			0.0567			0.0519			0.0521
15	477.25			0.0559			0.0509			0.0502
16	483.25			0.0552			0.0501			0.0484
17	489.25			0.0546			0.0492			0.0469
18	495.25			0.0539			0.0484			0.0454
19	501.25			0.0534		0.0775	0.0477			0.0441
20	507.25			0.0529		0.0754	0.0471			0.0430*
21	513.25			0.0524		0.0735	0.0465			0.0419*
22	519.25			0.0519		0.0718	0.0459			0.0409*
23	525.25			0.0515		0.0702	0.0454			0.0399*
24	531.25			0.0511		0.0688	0.0449			0.0391
25	537.25			0.0507		0.0675	0.0445			0.0383
26	543.25			0.0504		0.0663	0.0441			0.0375
27	549.25			0.0500		0.0655	0.0436			0.0368
28	555.25			0.0497		0.0644	0.0432			0.0362
29	561.25			0.0495	0.0892	0.0635	0.0429			0.0356
30	567.25		0.0620	0.0492	0.0871	0.0625	0.0425			0.0350
31	573.25		0.0613	0.0489	0.0852	0.0617	0.0422			0.0344
32	579.25		0.0607	0.0487	0.0835	0.0609	0.0419			0.0339
33	585.25		0.0601	0.0485	0.0819	0.0601	0.0416			0.0334
34	591.25		0.0595	0.0483	0.0804	0.0594	0.0414			0.0330
35	597.25		0.0590	0.0481	0.0790	0.0588	0.0411			0.0325
36	603.25		0.0585	0.0479	0.0778	0.0581	0.0409			0.0321
37	609.25		0.0580	0.0477	0.0766	0.0576	0.0407			0.0317
38	615.25		0.0575	0.0475	0.0755	0.0570	0.0404			0.0314
39	621.25		0.0571	0.0474	0.0744	0.0565	0.0402		0.0490	0.0310
40	627.27		0.0567	0.0473	0.0734	0.0559	0.0400		0.0482	0.0307
41	633.25		0.0564	0.0471	0.0725	0.0555	0.0398		0.0474	0.0303
42	639.25		0.0560	0.0470	0.0719	0.0551	0.0397		0.0466	
43	645.25	0.1130	0.0557		0.0711	0.0547	0.0395		0.0459	
44	651.25	0.1110	0.0553		0.0703	0.0542	0.0394		0.0452	
45	657.25	0.1091	0.0547		0.0696	0.0539	0.0392		0.0445	
46	663.25	0.1074	0.0547		0.0689	0.0535	0.0391		0.0439	
47	669.25	0.1057	0.0545		0.0682	0.0532	0.0390		0.0433	
48	675.25	0.1042	0.0542		0.0676	0.0528			0.0428	
49	681.25	0.1028	0.0540		0.0670	0.0525			0.0423	
50	687.25	0.1014	0.0537		0.0664	0.0522			0.0418	
51	693.25	0.1001	0.0535		0.0659	0.0521			0.0413	
52	699.25	0.0989	0.0533		0.0654	0.0517			0.0408	
53	705.25	0.0978	0.0531		0.0649	0.0515			0.0404	
54	711.25	0.0967	0.0529		0.0644	0.0512			0.0399	
55	717.25	0.0957	0.0527		0.0640	0.0511			0.0395	
56	723.25	0.0948	0.0526		0.0636	0.0507		0.0530	0.0392	
57	729.25	0.0939			0.0632	0.0505		0.0523	0.0388	
58	735.25	0.0930			0.0628	0.0503		0.0516	0.0384	
59	741.25	0.0922			0.0624	0.0502		0.0510	0.0381	
60	747.25	0.0914			0.0621	0.0500		0.0504		
61	753.25	0.0907			0.0617	0.0498		0.0498		
62	759.25	0.0900			0.0614	0.0497		0.0493		
63	765.25	0.0893			0.0611	0.0495		0.0488		
64	771.25	0.0887			0.0608	0.0494		0.0483		
65	777.25	0.0880			0.0605	0.0492		0.0478		
66	783.25	0.0875			0.0603	0.0491		0.0473		
67	789.25	0.0869			0.0601			0.0469		
68	795.25	0.0864			0.0598			0.0464		
69	801.25	0.0859			0.0596			0.0460		

* GLW 1700

Figure 35. Attenuation tables for broadcast waveguide types.

(described later) are used in single polarized circular systems to facilitate alignment of the polarization field with the components feeding the antenna and diplexer.

Shape Changes May Affect Performance

In a rectangular waveguide system, too high a pressure (typically anything above 0.5 psig) may cause the waveguide to "bulge" between flanges. This may cause the system VSWR to increase, thereby increasing the amount of reflected power and effectively reducing overall system efficiency. If the reflected signal becomes high enough (usually about 26 dB below the level of the main signal) a "ghost" signal may be evident at receiving locations. This occurs because the reflected signal has been delayed in time by its multiple trips through the system before being transmitted.

Truncated waveguide avoids this problem by enclosing the interior waveguide section in a circular jacket and pressurizing both equally. In this configuration, as in circular waveguide, the existence of higher pressures (2 to 3 psig) does not cause bulging.

However, if winds of high enough velocity distort the shape of any waveguide, overmoding problems can occur. Once additional higher order modes are generated by changes in shape, they reduce the efficiency of the waveguide, since some amount of power is not being delivered to the antenna in the desired mode. Also, this higher order mode energy may, as the shape "normalizes," reconvert to the desired mode. As in the case of a strong reflected signal, this reconverted higher order mode energy is transmitted on a delayed basis and, if strong enough, may also appear as a "ghost" at television receivers.

System Layout Considerations

General Information

Typical system configurations for different types of waveguide are shown in Fig. 36.

Because of the large size of bends, transitions, etc., and the differential expansion between waveguide and tower, careful coordination with the tower engineer and waveguide manufacturer is essential to prevent problems during installation and operation.

In addition, the horizontal run must be long enough to withstand the flexing caused by expansion in the much longer vertical run. Fig. 37 provides a guide to determine this length for circular waveguide systems. As can be seen, the horizontal run needs to be about 10% of the vertical for shorter systems and 5% for long systems. Always consult the manufacturer for confirmation prior to finalizing locations of the tower and transmitter building.

Horizontal and Vertical Straight Sections

Straight sections range from 11.5 ft to 12 ft depending on waveguide type and channel of operation. Follow the manufacturer's catalog recommendations.

The most common starting point for developing the list of required straight sections is the waveguide elbow located at the base of the tower. Working toward the building from the bend, the standard lengths recommended by the manufacturer are used back through the building opening. If the dimension of the final mating pieces cannot be determined accurately, specify a "length to be advised" piece. This will alert the manufacturer that a section must be made available based on last minute measurements.

The vertical run is laid out by stacking lengths end-to-end until the appropriate height is reached for positioning the top anchor plate. The same rule for a "to be advised" length should be applied as in the horizontal run.

Additional pieces will be needed between this point and the input to the antenna. As indicated in the figures, these will be a combination of waveguide-to-coax transitions, waveguide bends and special waveguide lengths. Since the waveguide is fastened rigidly at the top of the run, and the antenna is fastened rigidly at its mounting plate, the waveguide in between must also be allowed to grow and shrink due to thermal expansion and contraction. U-link bends or 90° elbows are the most common method for allowing the necessary "flexing" movement.

Tower Top

The vertical run is secured at the top end with an *anchor plate* or *milkstool* which must be carefully located on the tower to line up with the hangers for the vertical run and with the antenna interconnecting components. The milkstool hanger concept (Fig. 38A) provides adjustment in all directions to allow for minor field variations.

Hangers

Thermal expansion in both vertical and horizontal runs, depicted in Fig. 39, causes both to change positions. At the lower portion of the vertical run, lateral braces (Fig. 38B) are used to provide lateral support while allowing this shifting of position to occur. The supporting hangers for the horizontal run must also allow for an increased amount of movement at this area.

A spring hanger (Fig. 38C) is required on all other sections of waveguide in both the horizontal and vertical runs to provide the expansion/contraction capability necessary. For spring hanger and lateral brace mounting details, the manufacturer's installation bulletins should be thoroughly reviewed. With careful engineering and attention to all the necessary details, the hangers should attach to both the horizontal bridge and the vertical support structure with a minimum of difficulty.

Keep in mind that the coefficients of thermal expansion are 13×10^{-6} per degree F for the aluminum waveguide, and approximately 6.5×10^{-6} per degree F for the tower steel. This results in a relative differential movement of about 3/4" per 100 ft. for a 100°F temperature swing (2.5 cm per 40 m for a 56°C temperature swing). Also, the ambient temperature of the guide and the tower are different due to the additional heating effect of the RF power being transmitted. The resulting

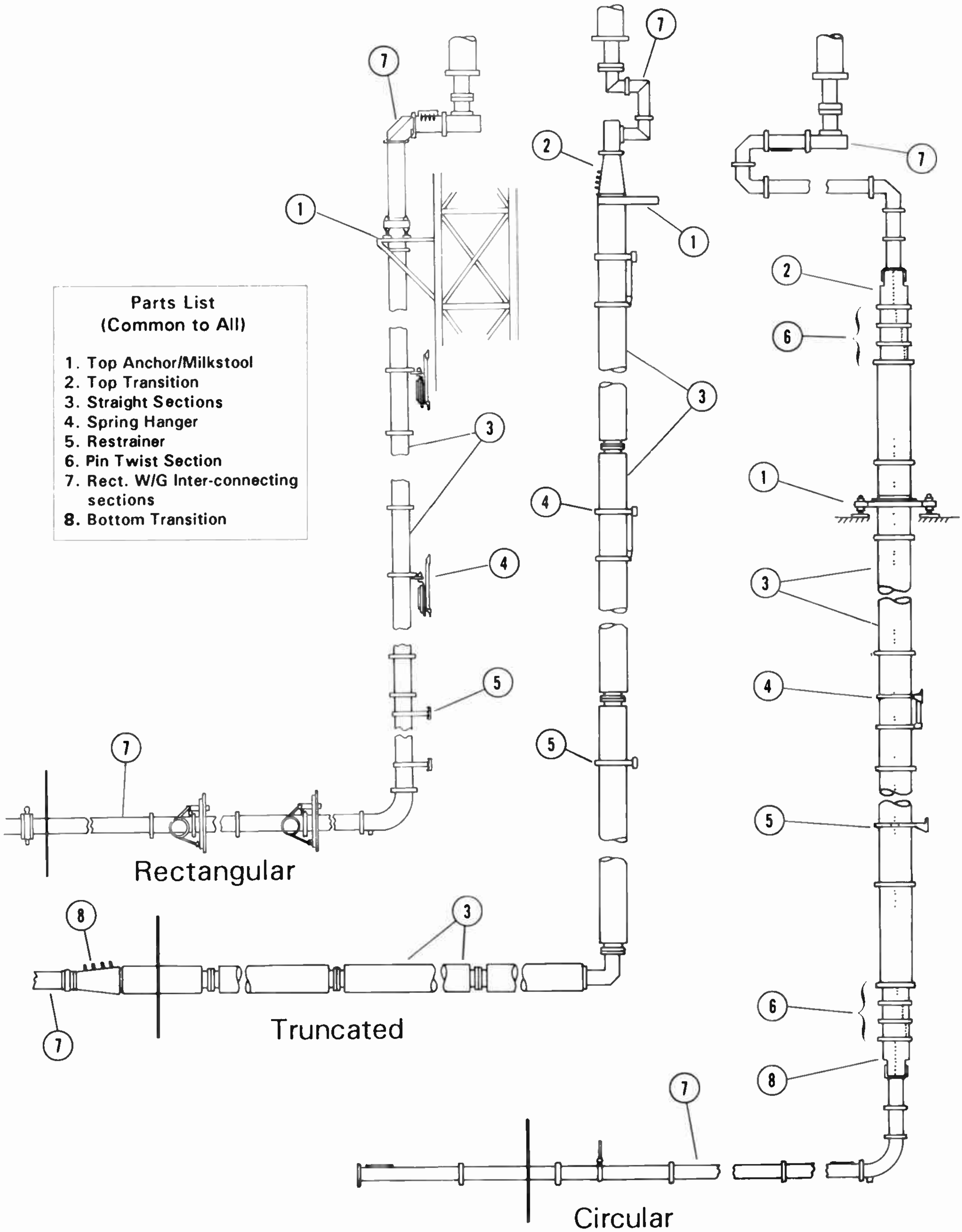


Figure 36. Typical waveguide system layouts.

Vertical Run Height (feet)	Minimum Horizontal Run Length (feet)					
	GLW			DTW		
	1350	1500	1750	1350	1500	1750
2000	75	90	100	72	84	120
1500	66	77	87	60	84	108
1000	54	61	70	60	72	96
500	38	42	49	48	60	84

Figure 37. Typical manufacturer's recommendations for determining minimum length of horizontal run to provide for expansion of vertical run.

expansion/contraction must be accommodated by providing sufficient clearance between all waveguide components and the tower or bridge (Fig. 39).

Transitions and Associated Components

Rectangular waveguide systems may include transitions as shown in the typical layouts of Fig. 36. Rectangular to coaxial transitions are most commonly used to interface the main waveguide run to the antenna and, if equipped with a coax input, the diplexer. Single polarized circular and doubly truncated waveguide systems will also use transitions from their shape to standard rectangular waveguide at both ends (Fig. 38D).

Single polarized circular systems incorporate multisection pin twists (Fig. 38E) at the transitions to facilitate alignment of the electric field (E-field) with the mating rectangular waveguide feeding the antenna and transmitter. This simplifies both the layout of the waveguide system on the tower and the installation. In rectangular-only systems, special twist sections are needed to rotate the run as required for proper alignment.

Tuners

For final system VSWR adjustment, the manufacturer may recommend sections of waveguide with tuners to be located periodically in both the vertical and horizontal runs, especially in rectangular waveguide systems. A tuning section may also be located between the antenna and the anchor plate to facilitate matching the antenna to the waveguide run during final system testing.

Installation

General Considerations

Pages could be written about the handling and installation practices appropriate for waveguide sections. A summary of the major considerations is presented here.

The crew contracted to do the installation will play a major role in the success of the project. Select one with experience, credentials, and recent references. This is not a good place to try to save a few dollars.

If the waveguide will be stored at the site prior to installation, keep the interior surfaces dry and clean by covering the sections with plastic sheeting. Tape the sides and ends securely to keep out wind-driven rain and dirt.

Hanger Installation

Before uncovering any waveguide sections, install all the hangers on the tower and the horizontal bridge. Follow the detail drawings developed by the tower manufacturer to insure the proper interface between the hanger and the tower. Due to the significant windload which the waveguide run will place on the tower, it is important that these instructions be followed carefully.

As shown in the system layout drawings of Fig. 36, several types of hangers will be utilized—a fixed hanger at the top of the tower, spring hangers down the tower, lateral restrainers near the base of the tower and horizontal spring hangers beneath the bridge. Spring hangers with additional travel are provided with certain types of installations to accommodate the larger displacement near the base of the tower in both the vertical and horizontal runs. Be sure that the manufacturer's instructions are followed when installing hangers.

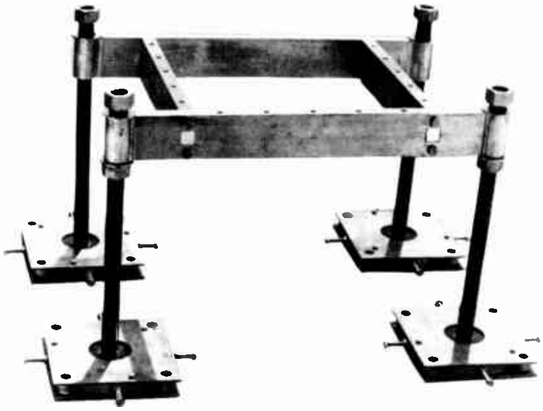
After installation, all hangers must provide the necessary movement to allow waveguide and flange displacement of the type shown in Fig. 39 with no interference at any tower member.

Waveguide Installation

Installation of straight sections usually starts at the base of the tower after properly locating the E-plane bend for clearance at both anticipated temperature extremes. However, single polarized circular waveguide is installed from the top down to eliminate the need for special field-cut circular sections in the vertical run.

Straight sections are normally provided with an O-ring groove in one flange only. This allows the sections to be installed with the groove facing up to provide a cavity for centering the O-ring and prevent "pinching" when bolting mating flanges together. Check for a uniform gap around the flange perimeter after torquing all flange hardware. A pinched O-ring will ultimately need fixing since it will cause VSWR changes, arcing, pressure leaks—or all three.

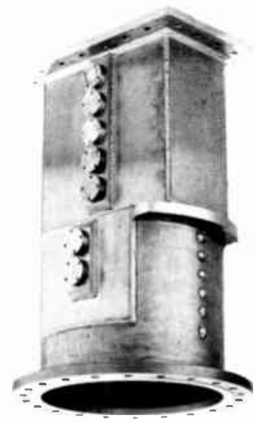
As mentioned above, be sure the bend at the base of the tower is placed in a location which provides freedom of movement for both the vertical and horizontal runs when differential thermal expansion occurs. As indicated, this differential amount can be approximated using the "¾ in/100 ft" (2.5 cm/40 m) guideline. Note that this will be true for both the vertical and horizontal runs, assuming steel is used for the tower and the horizontal bridge. This general rule is based on a 100°F (56°C) change in temperature. Thus, if the temperature is expected to increase 50°F (28°C) from the ambient at time of installation, the



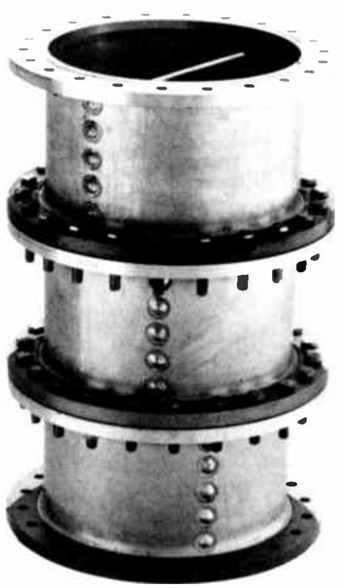
A) "Milkstool" Type Top Anchor

B) Lateral Restrainer for Circular Waveguide

C) Spring Hanger for Circular Waveguide



D) Circular to Rectangular Transition



E) Circular Waveguide Pin Twist Section

Figure 38. Typical waveguide system components.

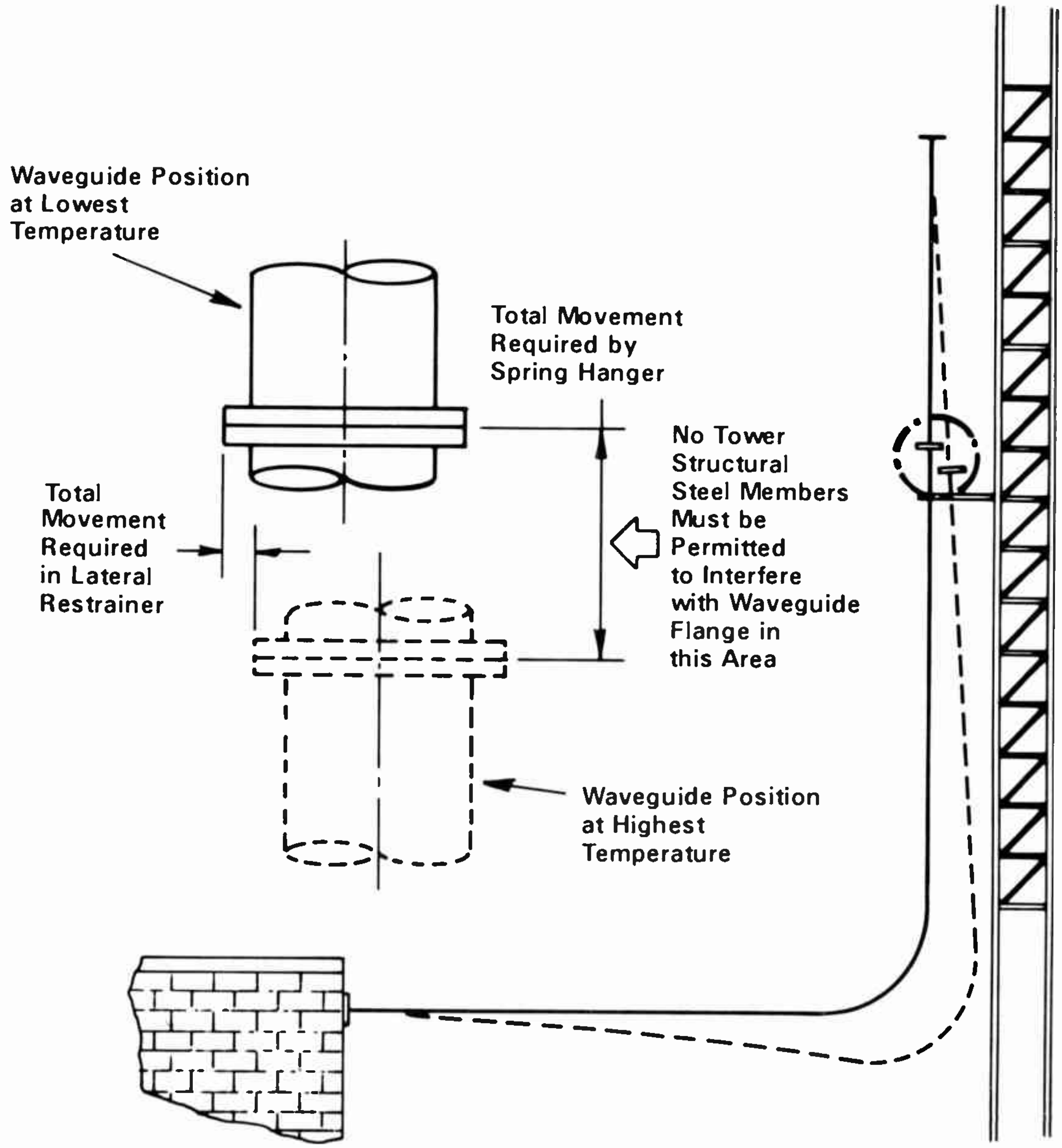


Figure 39. Effects of differential thermal expansion between transmission line and tower.

amount of differential *expansion* will be one-half that indicated above. However, remember that *shrinkage* will also occur as the temperature drops 50°F (28°C) from the ambient, so the same amount of "upward" movement must also be accommodated.

Waveguide flanges in mating sections are normally aligned by pins prior to tightening hardware. Make sure the installers attach these pins to their tool pouches using a long, strong cord. This will prevent the pins from inadvertently being dropped and, subsequently, not used. Misaligned flanges will cause power to be reflected at the point, adding to system VSWR and reducing transmission efficiency.

Pressurization

As mentioned earlier, positive dry air or nitrogen pressure should be maintained inside transmission lines to prevent "breathing." Ambient air can enter the line through the joints due to changes in atmospheric pressure. As the temperature drops, the moisture will condense on the internal surfaces of the waveguide. The accumulation of enough moisture in the waveguide can cause large changes in performance by increasing attenuation, system return loss, or both.

Rectangular waveguide, due to its shape, is usually not pressurized to levels greater than 0.25 psig to prevent distorting or bulging between the flanges.

Circular and truncated waveguides, on the other hand, can be pressurized at levels from 0.5 to 2 psig thereby reducing the likelihood of moisture penetration from the breathing process.

Due to the large volume of enclosed air and large surface area of the guide, sudden changes in ambient temperature and/or solar heating can cause fast changes in internal pressure. A pressure relief value is needed to bleed off rapid increases in pressure which can occur as the rising sun warms up the waveguide. When cooling occurs (due to clouds, sunset, etc.), the pressurization system must have the capacity to provide the required dry air volume quickly.

Due to the cyclic nature of waveguide pressurization systems (pressure relief followed by replacement of the purged volume), automatic dry air systems are more functional than nitrogen tank pressure systems since the latter would need regular replacement.

Testing

Electrical testing of waveguide systems should be performed after any of the following have occurred:

- Initial installation prior to applying power
- VSWR/reverse power trip
- Signal fluctuations of "in-line" waveform monitors or signal analyzers

The purpose of such tests is to verify that the overall performance of the system is as good as expected. Large amounts of reflected power reduce system efficiency, create the potential for "ghosting" and may even cause transmitter shutdown. It is advantageous to test new systems while installation personnel are still on site. Common tests performed are VSWR/return loss sweep and RF pulse test. Network analyzers with time domain capabilities can be used for overall system performance checks and are excellent devices for locating discontinuities not detectable by RF pulse testing. This section discusses such tests and presents actual data taken on waveguide systems.

Sweep Test

The sweep test is a frequency domain test that measures the return loss of the transmission line over the operating channel bandwidth. This will determine how efficiently energy is being transferred from the transmitter system into the transmission line. The measurement is typically performed at the gas barrier inside the transmitter building by applying a 6 MHz swept signal to the transmission line through a directional coupler or return loss bridge. A sample of the forward signal is compared to the reflected signal and the return loss or VSWR can be calculated. If a network analyzer is used, the calculation is done automatically, and the return loss/VSWR versus frequency is displayed on a screen. This test must be done with the end of the line terminated in either an antenna or load. Note that the effects of the terminating device will be included in the measurement.

Since power levels are highest at the visual carrier frequency, most waveguide systems are "optimized" for return loss at that frequency. The goal is at least 32 dB return loss (1.05 VSWR) which means less than a tenth of one percent of the transmitted power will be reflected back toward the transmitter. For a 120 kW UHF application, this translates to less than 120 watts of reflected power and a power transfer efficiency of greater than 99.9%.

A return loss sweep test of a circular waveguide system for a 240 kW application at channel 54 is shown in Fig. 40. This system included 900 ft of 15 in. circular waveguide in the vertical run, 100 ft of WR 1150 rectangular waveguide in the horizontal run, numerous bends and field cut sections, plus gas barrier and transmitting antenna. The maximum return loss level is about 23 dB to 24 dB (1.22 VSWR) at the edges of the band. At visual carrier, return loss is about 34 dB (1.04 VSWR).

The effects of discontinuities having repetitive spacing in a waveguide system, such as the flange joints, are shown clearly in Fig. 41A. The high VSWR "spikes" in this rectangular waveguide system occur at specific "reject" frequencies in the 200 MHz test bandwidth. A system utilizing different section lengths would cause these reject frequencies to change. For this reason, waveguide section lengths vary from about 11 to 12 ft depending on UHF operating channel in the same manner that rigid line lengths vary from 19 to 20 ft Selecting the correct length insures that no VSWR spikes will occur near the 6 MHz operating band.

Rectangular waveguide systems capable of broadband operation are now being provided for special applications. Fig. 41B shows the same waveguide system after adjusting flange tuners to reduce the effects of flange mismatch.

RF Pulse Test

An RF pulse test will display the performance of the line in the time domain. This test displays impedance changes versus time and allows the correlation of relative electrical performance of components to general location in the system when the propagation velocity is known. The test is performed by applying

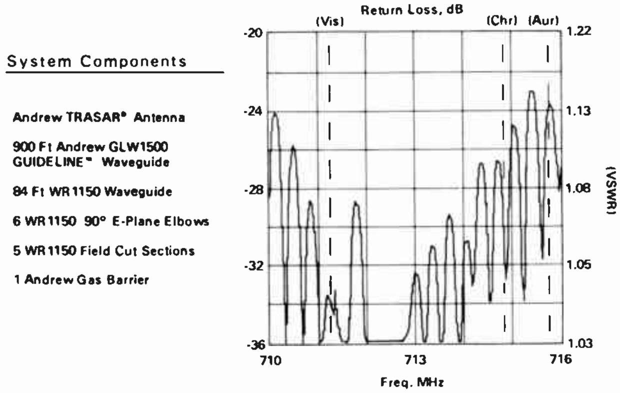


Figure 40. Sweep VSWR test results at channel 54 on 1,000 ft. circular waveguide system.

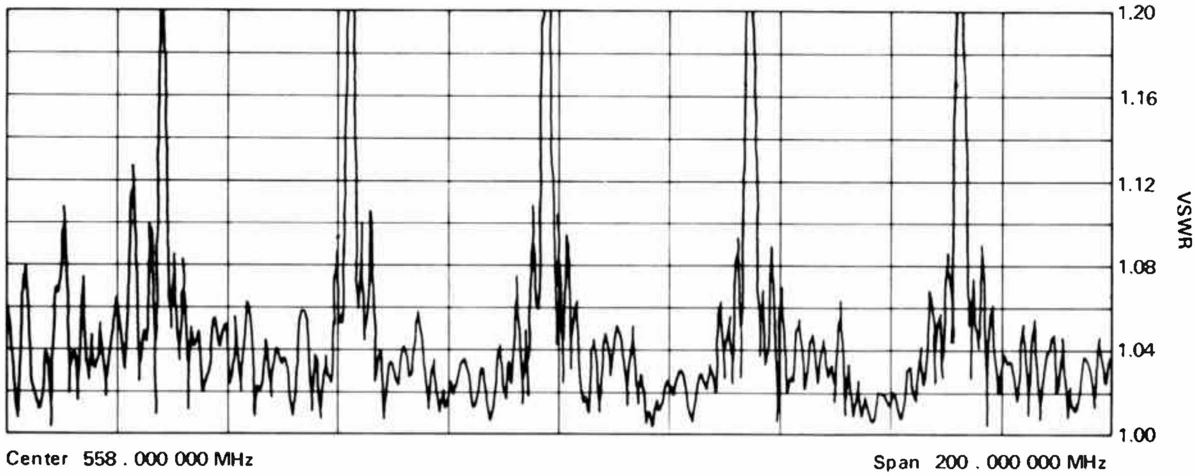


Figure 41A. Broadband (200 MHz) test of rectangular waveguide system to show high VSWR "spikes" at frequencies corresponding to flange joint spacing.

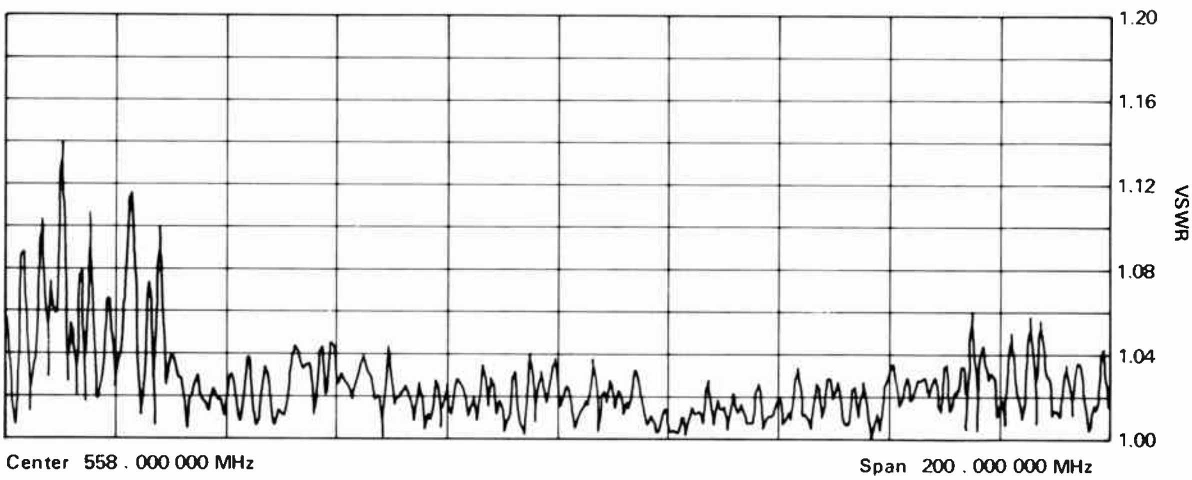


Figure 41B. Same waveguide system optimized for broadband application using tuners at flange joints to counteract flange mismatch.

CW signal bursts of various widths to the system and detecting the reflected signals versus time with a receiver having characteristics simulating an ideal video response. A 0.25 microsecond (2T) pulse has the spectrum of a normal video signal and is used to optimize the system for average picture content. The far end reflections are noted carefully to assure meeting the −32 dB return loss level in order to prevent noticeable ghosting. Because of the pulse width required to simulate the video information, this method does not have the positional resolution of a TDR. However, it provides excellent information regarding the system's transmission performance and can also detect gross problems such as misaligned flanges and pinched O-rings.

The main disadvantage of this test is that there is no standard "off-the-shelf" equipment package available. Therefore, correct interpretation of the test results is somewhat dependent on the skill and experience of the operator.

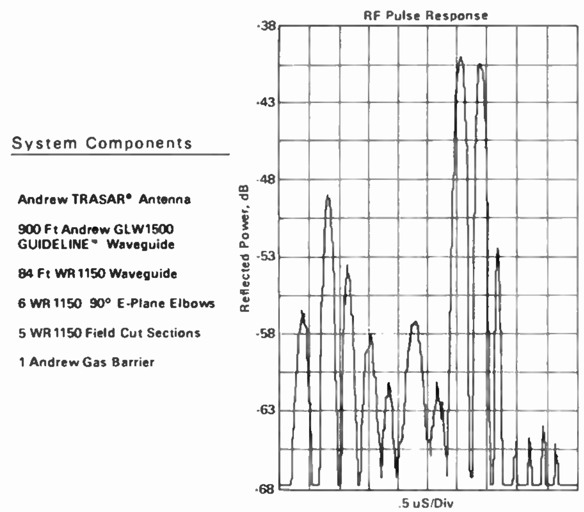


Figure 42. RF pulse test results on system shown in Figure 40.

An RF pulse test curve for the 15″ circular waveguide system described earlier is shown in Fig. 42. As can be seen, there is nothing between the bottom gas barrier up to and including the antenna with a return loss worse than about 40 dB (1.02 VSWR) at the visual carrier frequency of 711.25 MHz.

REFERENCES

1. Andrew Corp., Andrew Catalog 35, Orland Park, Illinois, 1991.
2. Andrew Corp., Pressurization System Planner for Broadcast Antenna Systems, Bulletin SP43–02, Orland Park, Illinois, 1990.
3. Andrew Corp., Pressurization System Planner for Microwave, Cellular and Radio Broadcast Antenna Systems, Bulletin SP43–01, Orland Park, Illinois, 1990.
4. Cablewave Systems, Rigid Coaxial Transmission Lines, Catalog 720A, North Haven, Connecticut, 1991.
5. Chipman, R.A., Theory and Problems of Transmission Lines, McGraw-Hill, Inc., New York, 1968.
6. Kraus, J.D., Electromagnetics, McGraw-Hill, Inc., New York, 1953.
7. Martin, H., "High-Power Radio-Frequency Coaxial Cables: Their Design and Rating," International Wire and Cable Symposium, 1973. 1973.
8. Myat, Inc., Rigid Coaxial Lines and Components, Catalog 85, Bergenfield, N.J.
9. NAB Engineering Handbook, Sixth Edition.
10. Ramo, S., Whinnery, J.R., and Van Duzer, T., Fields and Waves in Communication Electronics, Wiley, New York, 1965.
11. Reference Data for Radio Engineers, 6th ed., Howard W. Sams and Co., Inc., Indianapolis, Indiana, 1975.
12. Smith, P.H., Electronic Applications of the Smith Chart, Kreiger Publications, 1983.
13. Strickland, J., Time-Domain Reflectometry Measurements, Tektronix, Inc., Beaverton, Oregon, 1970.

2.4 Diplexers, Combiners, and Filters[1]

William DeCormier
Dielectric Communications, Raymond, Maine

INTRODUCTION

In this chapter, several methods of connecting the output of the transmitter(s) to the external transmission line system, using either coaxial transmission line or waveguide, will be presented. Various components, ranging from the familiar (simple patch panels) to the recently-developed (circulators and switchless combiner-routing systems), their operation and integration into the transmitter system will be discussed. Filters and filter networks, such as diplexers and multistage multiplexers for FM, TV, and low-power TV, will also be covered.

SWITCHING

Every transmitter installation requires some type of switching in the transmission line system. The complexity of this switching will depend on the individual needs of the station. It can be as simple as a three-port patch panel to connect the output of the transmitter to the station load, or as complex as a multiple transmitter-antenna installation.

Patch Panels

Patch panels provide a convenient method of rerouting the interconnecting transmission lines between various inputs and outputs in the transmitter plant. Since they are manual devices and cannot be changed very rapidly, their use is limited to maintenance functions or as a secondary means of switching.

Coaxial Patch Panels

Coaxial patch panels can have any number of ports, but the more common have three, four, and seven-ports. They consist of the appropriate number of quick disconnect connectors mounted on a panel, and interconnecting transmission lines, usually in the form of U-lines.[2] The connectors are spaced so that the U-links may be used to interconnect any two ports (Fig. 1). For example: the connectors of a three-port patch panel would form an equilateral triangle.

Interlock switches are used to prevent transmitter power from being applied until the U-links are in the proper positions and properly seated. Power handling

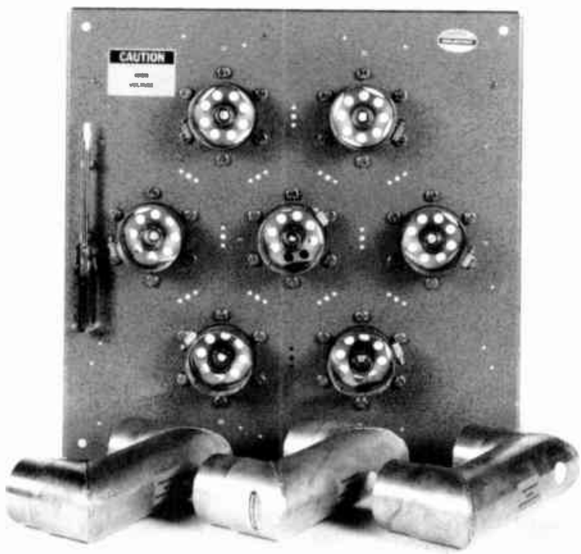


Figure 1. Seven-port patch panel.

[1] Portions of this chapter were retained from the 7th Edition.

[2] Short lengths of transmission line with connectors on both ends.

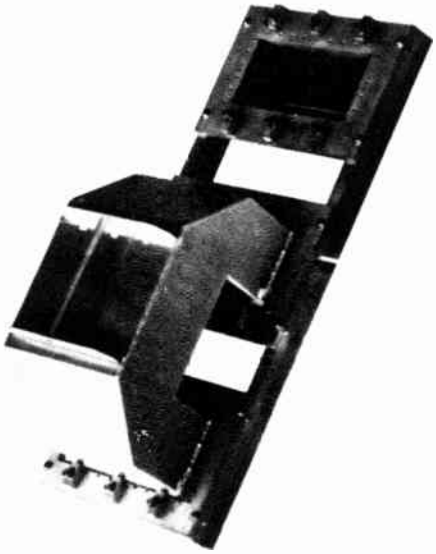


Figure 2. Three-port E-plane waveguide patch panel.

capabilities of the patch panels are essentially the same as the mating transmission lines.

Waveguide Patch Panels

Waveguide patch panels are used in much the same way as their coaxial counterparts. However, since waveguide is larger and inflexible, and most transmitter installations use rectangular waveguide inside the transmitter building, the patch panels are not quite as versatile. Generally, waveguide patch panels are only available as a three-port unit (Fig. 2).

Since the waveguide is rectangular, the ports must be in a straight line in either the broad wall or narrow wall plane.

Manual Coaxial Switches

Manual coaxial switches are generally available in either single-pole-double-throw (SPDT) or four-port transfer-type configurations. They have two distinct advantages over the manual patch panels: ease of operation and speed of switching.

Most manual coaxial switches have either a lever or knob that is turned to change positions of the switch (Fig. 3). This can be accomplished in a few seconds, compared to minutes to change a manual patch panel. Like patch panels, manual coaxial switches have interlock switches to turn off the transmitters during switching.

Power ratings of coaxial switches are approximately 80% of the equivalent coaxial transmission line. Since they are more complicated to build, they cost more than a patch panel. In order to accomplish the same functions performed by a seven-port patch panel, several switches will be required, because they are only available in three-port SPDT and four-port transfer configurations.

Motorized Coaxial Switches

Generally, motorized coaxial switches are very similar to manual switches. To reduce production costs, many of the parts are the same.

Usually, the RF portion of the switch is the same for either manual or motorized switches. The knob or lever is replaced with a motor drive assembly. The motor drive system requires some type of control system to start and stop the motor in the switching sequence.

Most motorized switches are available with a choice of motor voltages, with 115 vac being the more common, and various control circuit voltages. A small control relay isolates the switch from the control circuits of the transmitter system. Generally, both sides of the control relays are available in the electrical connector. Therefore, the user must supply a power source to energize the control relay.

Rotary motorized coaxial switches are available for 1⅝", 3⅛", 4$\frac{1}{16}$" and 6⅛" cable systems. These switches will change positions in approximately two seconds. Their frequency range is good up through the TV frequencies. Some of the larger switches are limited in their use in the UHF range because of possible moding problems.

For use at higher power levels and high frequencies, another type of switch was developed. This is a motorized U-link type switch (Fig. 4). Its power ratings are essentially the same as the comparable transmission line and its problems at the higher frequencies have, therefore, been greatly reduced. Sizes for 3⅛", 4$\frac{1}{16}$", 6⅛", 8$\frac{3}{16}$", and 9$\frac{3}{16}$" cable systems are available. Because of the mass of the moving parts of the larger switches, switching time is increased to approximately 10 seconds.

Figure 3. Coaxial rotary switch.

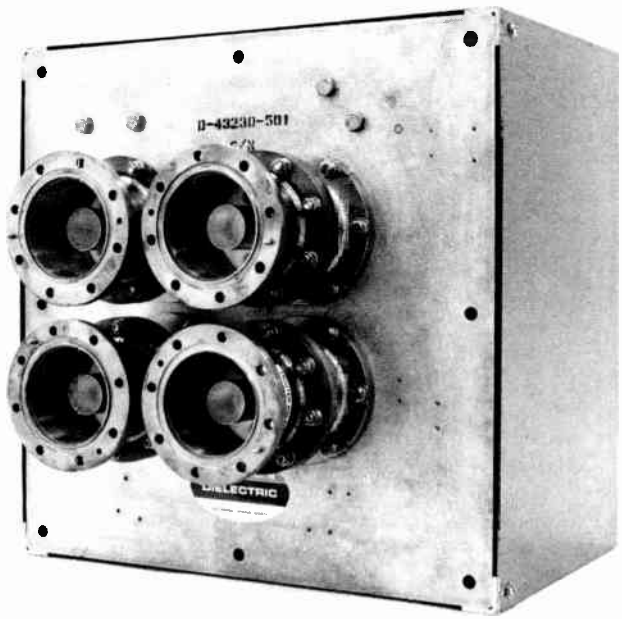


Figure 4. High power coaxial switch.

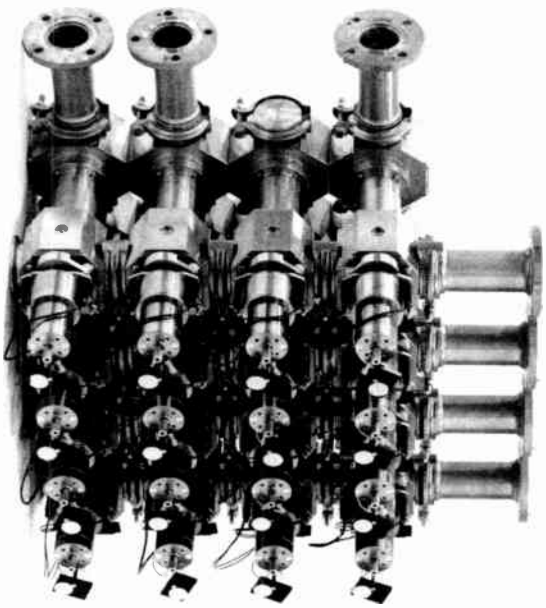


Figure 5. Coaxial crosspoint matrix.

Switches for Shortwave and AM Applications

Open wire switches can be used in lower frequency applications, such as the AM and shortwave bands. These switches often resemble high-power relays or contactors. The major differences are the type of dielectric material and the spacing between the contacts. Various types and sizes are available for use with powers from a few watts to several hundred kilowatts.

In addition, coaxial crosspoint switches are commonly used for AM and shortwave applications. A crosspoint switch consists of two SPDT switches in a single case. When used in a matrix, the transmitters are each connected to the end of a column and the antennas or load are connected to the end of the rows. In the inactive position, each switch passes the energy from one transmitter along its column while simultaneously passing another channel, highly isolated along a perpendicular row. When a switch is activated, its signal along the column is diverted to the row either to the right or the left. In this way, output from any transmitter can be routed to any antenna.

Coaxial sizes in excess of 10 inches in diameter may be switched in this manner, handling average power levels in excess of 1 megawatt. Fig. 5 depicts 1⅝" coaxial cross point matrix (4 rows × 4 columns). Newer designs, such as the one shown, feature >80 dB isolation, modular construction and auxiliary interlocks.

Waveguide Switches

Waveguide switches are available in three-, four-, and five-port versions. The RF sections of both manual and motorized switches are usually the same. Manual units will have a knob or handle to change the positions of the switch. Motorized units have the manual drive replaced with a motor drive assembly. Some type of motor control circuits will also be required.

The four-port transfer switch shown in Fig. 6 is an E-plane type of switch. Both E-plane and H-plane units are available. The four-port units usually take the form of crossed waveguide, with the ports 90° apart. Within the switch is a metal back plate that has

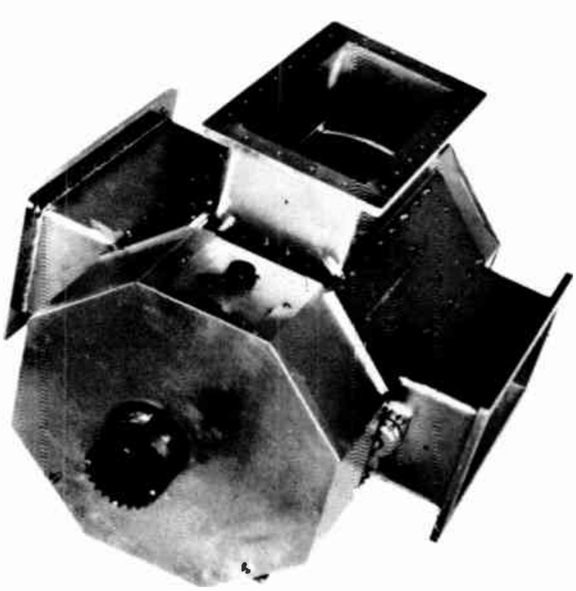


Figure 6. Port E-plane waveguide switch.

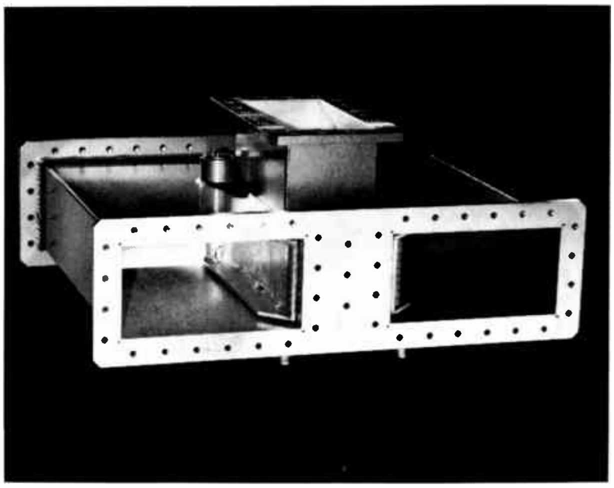


Figure 7. Five-port waveguide switch. (Courtesy Micro Communications, Inc.)

fingerstock around the four sides which contact the waveguide case. The back plate is positioned at a 45° angle to the ports of the switch. Thus, two ports of the switch will be connected together in the form of an elbow. In one position, the top port is connected to the right hand port, while the bottom port is connected to the left hand port. In the other position the opposite is true.

The control circuits of the waveguide switch are very similar to those in the coaxial switches. Various motor and control voltages are available. The waveguide switch is equipped with some type of interlock switches to turn off the transmitters during the switching process.

FILTERS

Filters are used in broadcasting to limit the undesirable emissions from transmitters. More specifically, it is necessary to limit the harmonic content of transmitters to prevent interference at higher frequencies with other services. In addition, it is necessary to limit the intrusion of other RF signals into the final stage of transmitters as a result of antenna coupling.

The FCC addresses these requirements for the FM broadcaster in Volume III, Subpart B, Section 73.317, Paragraph 14 of the Regulations by stating that "any emissions appearing on a frequency removed from the carrier by more than 600 kHz shall be attenuated by 43 dB + 10 log (power) dB below the level of the unmodulated carrier or 80 dB, whichever is the lesser attenuation." If a broadcaster is operating above 5 kW, the 80 dB requirement applies. For comparison, a broadcaster operating at 100 watts would be required to meet a 63 dB requirement under these rules.

The TV requirement is addressed in Volume III, Subpart D, Section 73.687, Paragraph (i)(1). In part, it states ". . .all emissions removed in frequency in excess of 3 MHz above or below the respective channel edge shall be attenuated no less than 60 dB." It goes on to say that this requirement should be considered temporary and that the state of the art might be the more appropriate limit. Broadcasters are encouraged to seek this limit in order to meet future rule making requirements.

Harmonic Filters

Harmonic filters are commonly used on the output of all transmitters used for broadcast applications. They can be built for use with either coaxial or waveguide. The decision to use one form over the other is a matter of convenience (i.e., size), performance, and cost.

The coaxial form is used in the lower portions of the broadcast spectrum, namely VHF, FM, and portions of UHF bands. For low band VHF and FM, where the maximum ERP is limited to 100 kW, coaxial harmonic filters are used since the common EIA line sizes from 3⅛" to 6⅛" coaxial can handle these power levels without degradation due to higher order modes. Waveguide would have to be greater than 60" to operate in the fundamental mode at these frequencies.

Harmonic filters will pass the fundamental frequency with efficiencies of about 98% or −0.1 dB insertion loss. They will reject the second through the fifth harmonic with attenuation of −40 to −50 dB. By virtue of their function of attenuating harmonics they are, by necessity, designed to handle a limited segment of the band (less than one octave). Table 1 lists the typical way the band is divided.

The skirt of the attenuation curve must have a slope sufficiently large to pass the highest fundamental frequency and reject the lowest frequency in the second harmonic. A filter with nine to eleven stages will normally provide 40 to 50 dB rejection at the low end of the second harmonic.

A waveguide harmonic filter is normally used at frequencies on the high end of the UHF band. This is necessary for several reasons. Larger coaxial sizes will support the generation of higher order modes near the high end of the frequency spectrum for UHF-TV. Higher order modes will be sustained at a frequency where

$$\text{wavelength} = \pi\,(a+b)$$

TABLE 1

Channel	Fundamental	2nd Harmonic	Typical Construction
2–3	54–66 MHz	108–132 MHz	Coaxial
4–6	66–88 MHz	132–176 MHz	Coaxial
FM	88–108 MHz	176–216 MHz	Coaxial
7–13	174–216 MHz	348–432 MHz	Coaxial
14–43	470–650 MHz	940–1300 MHz	Coaxial
44–52	650–698 MHz	1300–1396 MHz	Coaxial or Waveguide
52–69	698–806 MHz	1396–1612 MHz	Waveguide

where a and b are the radii of inner and outer conductors. Larger coaxial sizes are also needed to handle the power levels authorized in the UHF band. 8³⁄₁₆", 75 ohm line will support higher order modes at frequencies just above channel 56; and 9³⁄₁₆", 75 ohm cable at frequencies just above channel 40. But the construction of a coaxial filter is a cascade of larger and smaller diameter inner conductors. The larger inner conductors essentially cause moding to occur at longer wavelengths or lower frequencies. This phenomenon lowers the effective frequency at which coaxial filters can be used.

Waveguide filters must therefore be used above channel 40 when transmission line power levels exceed the rating of 6⅛" coaxial. One type is commonly called a waffle iron filter because the broad walls of the waveguide resemble the top and bottom plate in a waffle iron. A waffle iron harmonic filter in WR 1150 waveguide can be used between channels 40 and 69.

Harmonic filters are usually supplied with transmitters since the transmitters cannot meet FCC Rules with regard to harmonic content without a filter. The broadcaster will seldom need to buy a harmonic filter unless he has experienced a severe transmission line failure or acquired a used transmitter.

Band-Pass and Band-Stop Filters

Band-pass filters are used sometimes in combination with band-stop filters to control another class of spurious emission problems. The transmitter (with its harmonic filter) is capable, in the absence of other RF signals, of producing transmissions which are free of any emissions. Potential problems arise, however, when the transmission facilities of two or more broadcast channels are located very close to one another.

Assume, for example, that Channel A and Channel B use antennas on the same tower. This is common today. Channel A's antenna, in addition to transmitting its primary signal, also receives some of Channel B's signal. The magnitude of this received signal depends on the gain and bandwidth of Antenna A at the frequency of Channel B as well as the distance between the two antennas. Several spurious signals can be generated in Transmitter A and transmitted on the air as a result of the presence of RF from Channel B. The transmitter will usually, because of its limited bandwidth, provide several dB turnaround losses for this spur. The most troublesome spur will occur at a frequency which is

$$F = 2A - B$$

A comparable problem could occur in Transmitter B where

$$F = 2B - A$$

The magnitude of the spur will be equal to the power level of the coupled signal minus the turn-around loss. So, if Channel B is present in Channel A's transmission line at a level of −40 dB down from Channel A's power level, and Transmitter A provides −10 dB turn-around loss, then a spur will likely exist at (F = 2A-B) with an amplitude of −50 dB from Channel A's amplitude. In order to comply with FCC Rules for FM, therefore, a filter would have to be installed in Channel A's transmission line which would pass Channel A with minimal insertion loss (usually −0.15 dB) and provide −30 dB rejection at Channel B. This would lower the spur level to −80 dB below Channel A's transmission line level.

Another example of spurious emissions occurs when a channel 14 or channel 69 system is located in close proximity to land mobile communications. Two types of interference have been noted in this situation:

1. An intermodulation product of visual, aural, or color signals mixing to produce a spur at a frequency used by the land mobile system.
2. The power generated by the TV station simply overloads the front end of land mobile receivers although the carrier signals of each are at different frequencies.

The first situation must be addressed by providing filters in the TV station's system to reduce the resulting intermodulation. Fig. 8 illustrates how the intermodulation was suppressed at Home Shopping Network's Channel 69 in Hollywood, Florida.

The second situation must be addressed by installing a filter in the land mobile receiver system. This can generally be achieved with much smaller hardware when required.

The decision to use band-pass versus band-stop filters should be based on the nature and extent of the problem. Each type has virtues and limitations which make them suitable for certain problems.

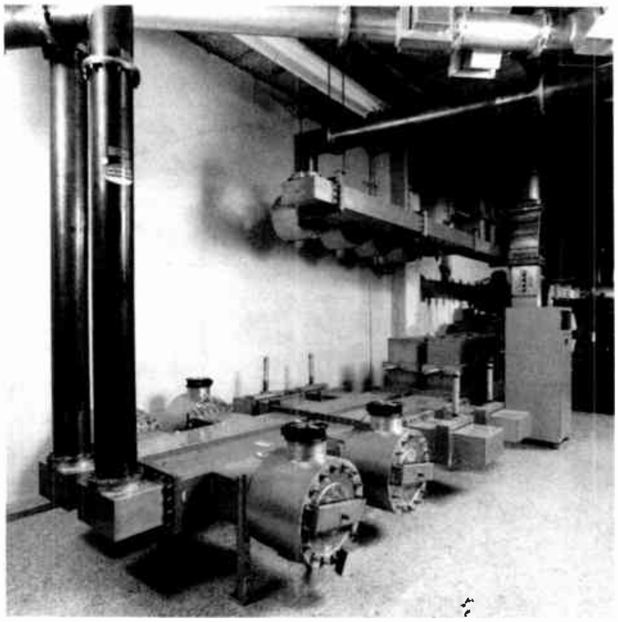


Figure 8. Notch filters at visual and aural inputs in Hollywood, Florida. (Courtesy Home Shopping Network.)

The typical response curves for several combinations are presented in Figs. 9 through 12. The band-stop (as depicted in Figs. 9 and 10) is characterized by rapidly rising skirts which make these filters particularly suited to rejecting frequencies that are extremely close to the desired frequency (i.e., frequencies displaced by as little as 0.8%). However, due to their sharp response, they are more prone to drifting with temperature than the band-pass configuration. They can be built so that drifting does not affect their desired performance, if manufacturers design them to minimize this tendency. Band-stop cavities are commonly used in the FM band. They are also used in VHF or UHF diplexers, but are not used to protect one TV channel from another since their band-stop widths are generally too narrow to reject an entire TV channel.

The notch responses shown were plotted in the FM band near 100 MHz. The center frequency can be scaled to anywhere in the VHF or UHF band, and similar percentage bandwidth would be obtainable. Fig. 9 depicts the response of a single-cavity notch. Fig. 10 shows a dual-cavity response.

Band-pass filter responses for a four-cavity and five-cavity filter are shown in Figs. 11 and 12. Again, these were measured near 100 MHz and their percentage bandwidth would be obtainable at any other frequency in VHF or UHF spectrum. The skirt selectivity is not as sharp for band-pass cavities as for band stop cavities. This suggests that more stages will be needed in band-pass cavities to obtain the same attenuation as a band-stop cavity. The band-pass cavity, however, has some advantages; since the pass band is broad, it is not affected by drifting due to temperature changes. The reject curves are also located symmetrically about the pass band.

The formula for locating the spur,

$$F = 2A - B$$

will always place the spur at exactly the same distance, frequency-wise, from the pass-frequency, but on the opposite side of the pass band. So the band-pass filter attenuates the incoming RF on one side and the resulting spur on the other side of the pass band. It also has the advantage of attenuating all frequencies sufficiently removed from the pass band. This is especially useful when multiple interferences are detected.

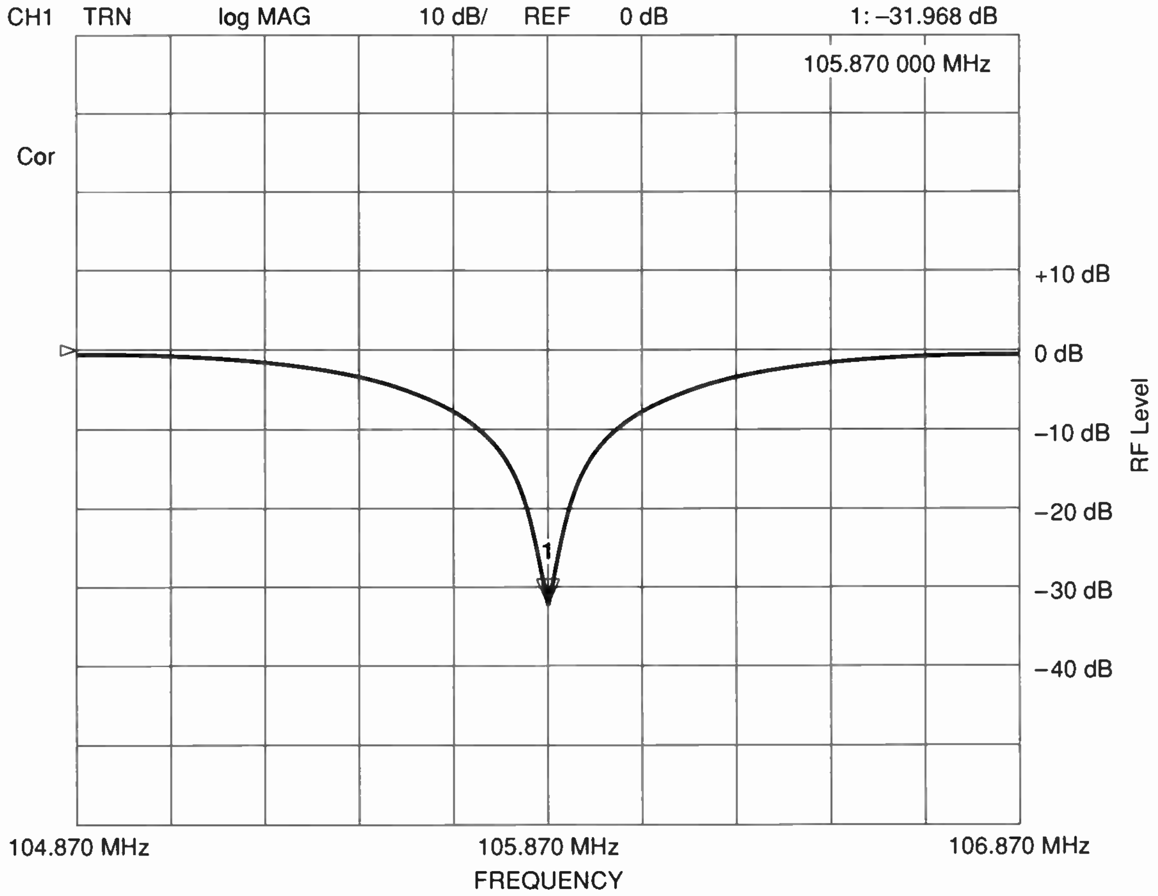


Figure 9. Single-cavity notch filter response 31.9 dB.

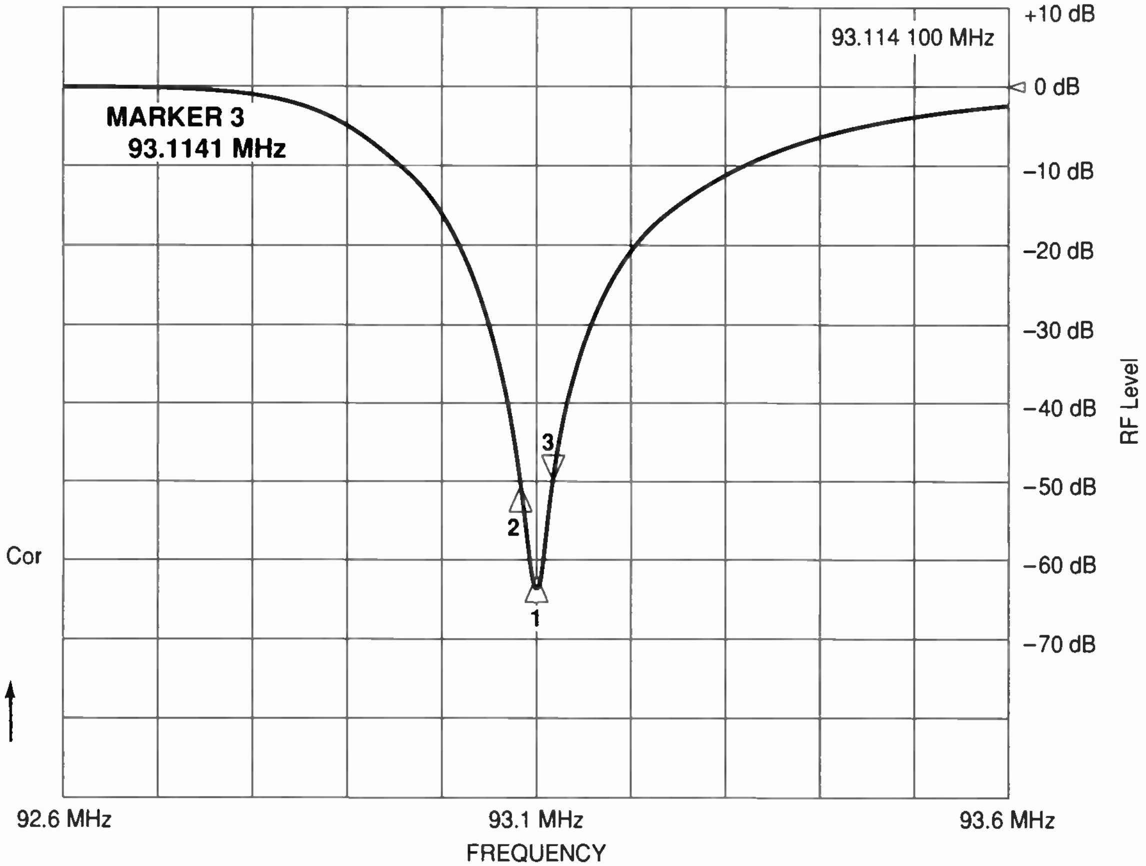


Figure 10. Dual-cavity notch filter response 64 dB.

Combinations of band-pass and band-stop are also available in a single filter. They can be visualized by superimposing a band-stop and a band-pass filter response with the notch located at the edge of the pass band. These filters are useful when interference is caused by a combination of one frequency close to the desired frequency and multiples further removed from it.

All of these responses in the VHF and FM bands can be achieved with coaxial cavities approximately ⅜-wavelength long and 12″ to 24″ in cross section. (See Fig. 13.) In the UHF band they can be produced by using coaxial cavities or waveguide cavities. These filter systems also serve as building blocks for the diplexing and multiplexing systems discussed in the next section.

Once a decision is made to use a band-pass, band-stop, or combination of both filters for a particular application, it is important to specify a few additional parameters to insure that the filter does not degrade the audio or visual content of the broadcast signal.

For example, if a filter is used to prevent Channel B from entering the transmitter of Channel A, the filter will pass Channel A and reject Channel B. But the reject curve for Channel B must not infringe upon the bandwidth of Channel A. Therefore, an insertion loss variation must be specified across the operating bandwidth and perhaps beyond. At FM, where these filters are commonly used, insertion loss variation can be kept within

$$< 0.1\ \text{dB},\ \pm 200\ \text{kHz}$$

when either the notch is sufficiently removed in frequency or the pass band is broad enough. The edge of the reject skirt in either configuration is also characterized by a large deviation in group delay. Group delay is defined as a change in phase divided by a change in frequency.

$$\text{Group delay} = \frac{\text{Change in phase}}{\text{Change in frequency}}$$

The standard for group delay at FM seems to be

$$\pm 25\ \text{ns}\ \pm 150\ \text{kHz}$$

This specification approaches state of the art limits when two FM channels are separated by only 800 kHz. Since group delay can exceed either of the above specifications at frequencies well inside the 3 dB points of a band-pass filter if improperly tuned, specification of group delay is desirable in order to assure proper audio performance.

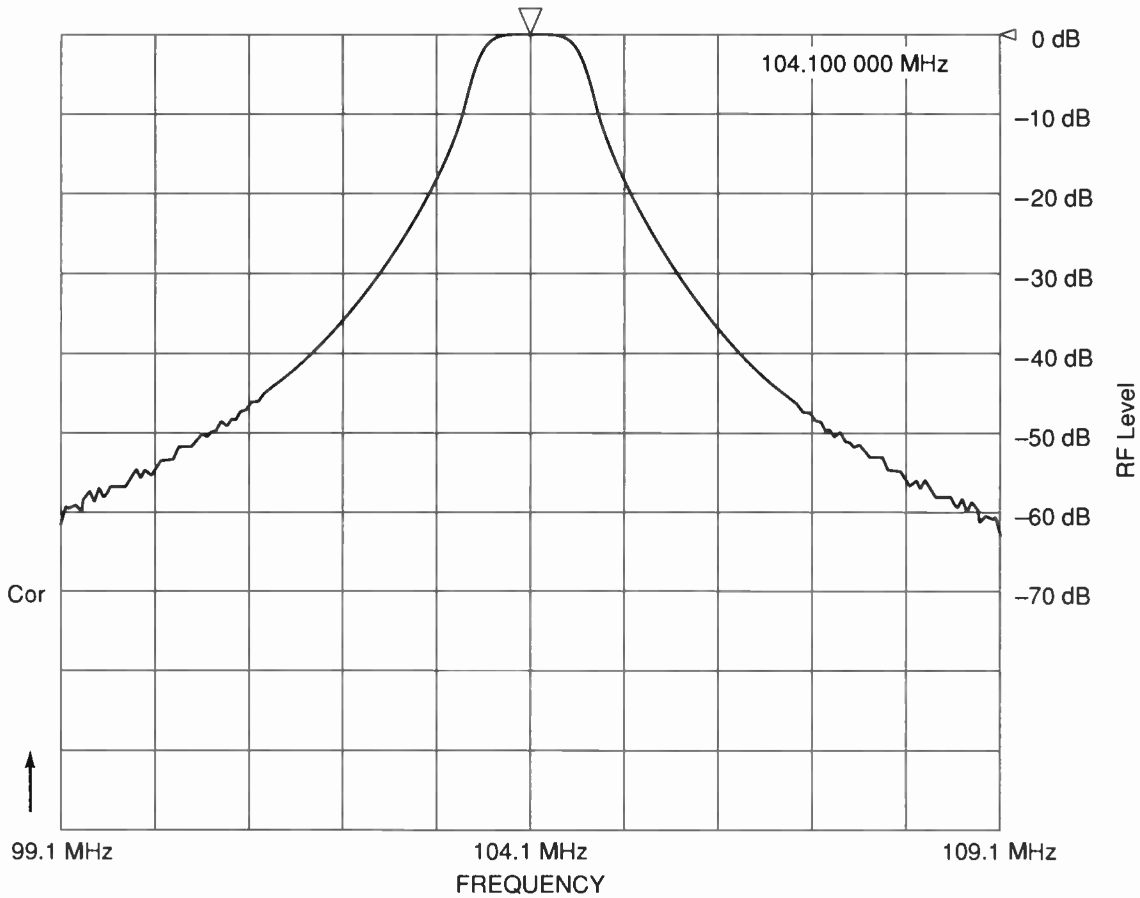


Figure 11. Four-cavity band-pass filter response.

Group Delay Correctors

The group delay response of a band-pass filter is U-shaped. A device which uses a hybrid in conjunction with two notch cavities produces a group delay response shaped like an inverted U. Since group delay is additive, the positive group delay of the bandpass added to the negative group delay of this corrector causes cancellation at the band edges relative to center frequency. The group delay correctors thereby produce a relatively flat group delay. These were first developed as high power devices used in multiplexers. They have since been developed for low power applications and are now used in FM applications and in the aural input of TV diplexers.

Diplexers and Multiplexers

Diplexers and multiplexers are devices which allow broadcasters to combine two or more frequencies in a common transmission line while providing the isolation needed to prevent either transmitter from generating spurious emissions. Figs. 14, 15, 16, and 17 provide schematic examples of configurations which can provide, in varying degree, the necessary response.

In the tee diplexer (Fig. 14) each input leg contains either a band-pass or a band stop filter. Each of these filters is characterized by a good voltage standing wave ratio (VSWR) and low insertion loss within the pass band. The slope of the reject curve away from center frequency is completely dependent on the number of cavities in each leg. Therefore, F1 and F2 must be separated sufficiently in frequency to allow the reject skirt to reach a rejection sufficient to obtain the desired isolation. All of this isolation (other than the 3 dB split of the tee) must be provided by the filter cavities. Figs. 11 and 12 show rejection to be expected with deviations from center frequency. In general, F1 and F2 must be widely separated for band-pass legs, and each operating band must be narrow for band-stop legs. In addition, this configuration has limitation when multiple frequencies must be combined.

Fig. 15 illustrates the branch combiner for three or more channels. This device is not generally used for inputs greater than 10 kW since the VSWR's and bandwidths are not as good as those in the constant impedance units shown in Figs. 16 and 17. Broadcasters with tight budgets may still consider these units if they are willing to accept some performance degradation.

Figs. 16 and 17 respectively are band stop and band-pass versions of a constant impedance diplexing configuration in common usage today at VHF, UHF,

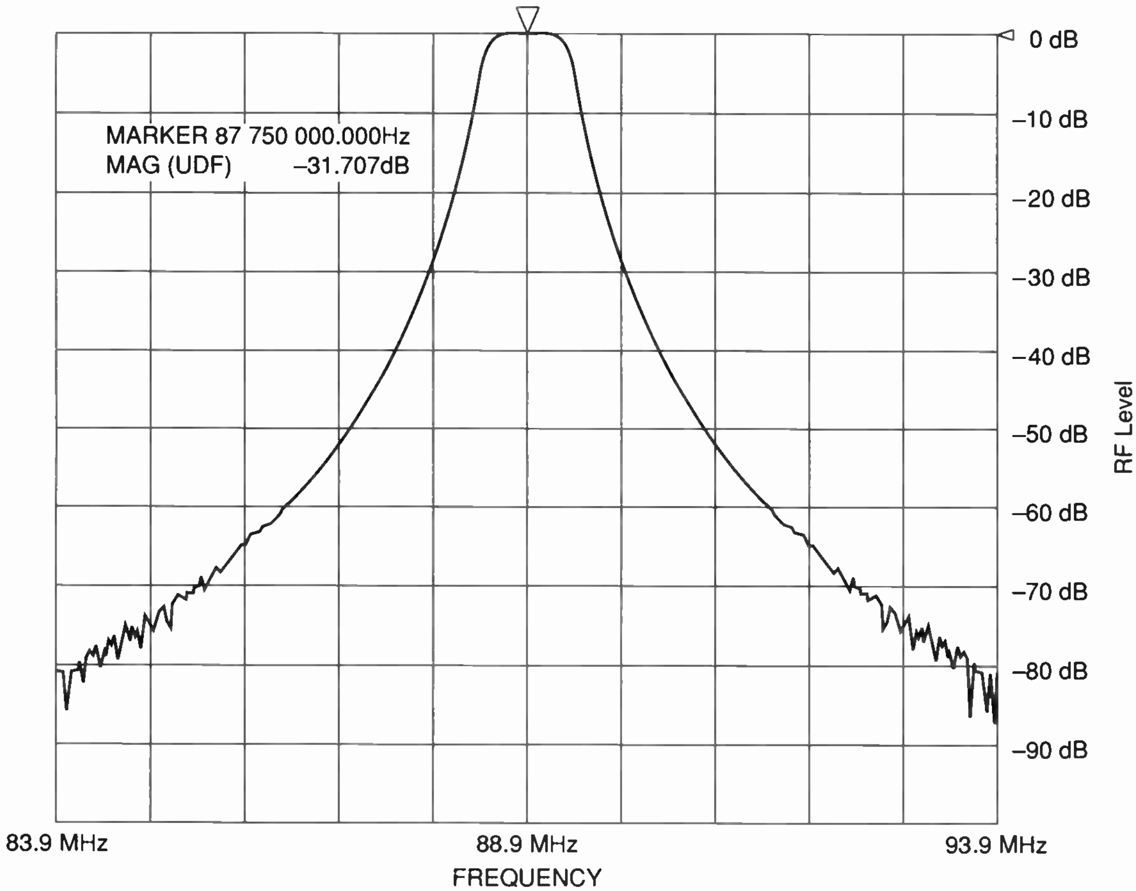


Figure 12. Five-cavity band-pass filter response.

and FM frequencies. In both configurations 3 dB hybrids are used on both ends of the system and there will be an equal number of cavities in each leg between the hybrids. With the exception of the load on the isolated port and possible differences in line size, each system is electrically symmetrical about lines running through the center, both vertically and horizontally. These systems have similar response in both coaxial and waveguide configurations.

The band-stop system of Fig. 16 is typically used for visual to aural diplexers in television applications. F1, in that case, would be aural and F2 would be visual. Aural would enter at F1, be split into both legs of the hybrid, and then be reflected by the notch cavities. The reflected signal would pass through the

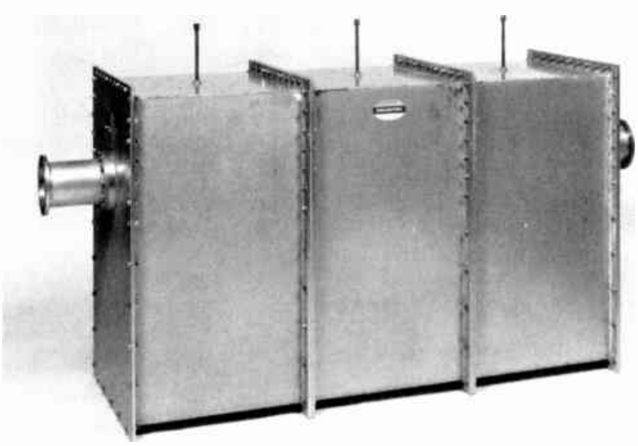


Figure 13. Three-cavity band-pass filter.

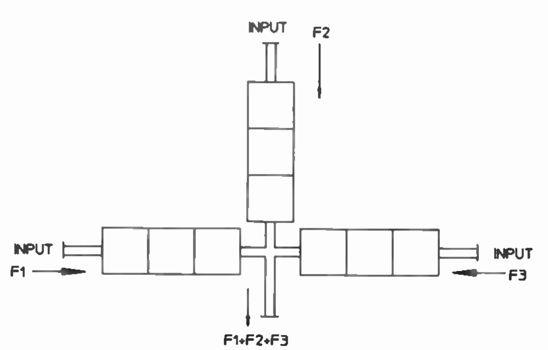


Figure 14. Schematic tee diplexer.

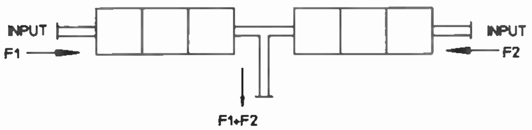


Figure 15. Schematic branch diplexer.

same hybrid as before but recombine in the output port.

The visual signal (F2) would enter the rightmost hybrid, split into the two lines with cavities, pass the cavities unattenuated, and recombine into the output. In most TV diplexers, the second cavity in each leg would be tuned to 3.58 MHz below the visual carrier. This is the lower sideband of the visual color signal. The signal enters F2 with the visual signal and is reflected, in a manner similar to the aural signal, except that this signal is reflected into the load and absorbed.

An analysis would show the isolation of visual (F2) to aural (F1) would be due solely to the isolation inherent in the leftmost hybrid. Over a 6 MHz bandwidth, this is typically 35 dB, but can be as high as 45 dB with special care. The isolation from aural (F1) to visual (F2) is due to two components. The notches produce 27 dB typically and the rightmost hybrid about 35 dB, for a total of 62 dB.

In previous years, this type of diplexer has been used at FM frequencies, but the poor isolation in one direction required the addition of supplementary cavities on the F1 input. In addition, the group delay for this approach was four times higher than achievable with a band-pass system and the lack of symmetry and inverted shape made group delay compensation more difficult. The bandwidth, while adequate for the aural component of a TV system, has proven to be too narrow for FM applications resulting in incidental amplitude modulation.

The notch type constant-impedance diplexer has been widely used for TV applications with excellent results.

The band-pass constant-impedance diplexer, shown in Fig. 17, is widely used for high-power diplexing or multiplexing applications in FM. It is occassionally used for diplexing complete TV signals, each containing visual plus aural. Several units can be cascaded together, each module providing a separate station's

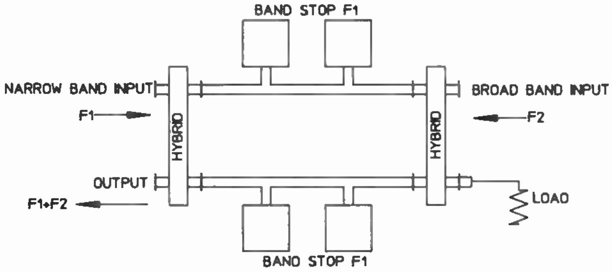


Figure 16. Schematic band-stop constant impedance diplexer.

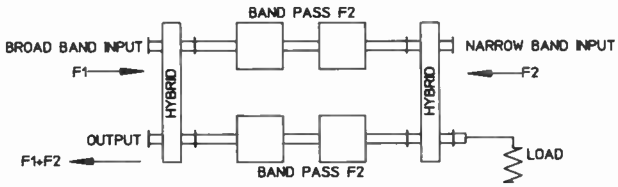


Figure 17. Schematic band-pass constant impedance diplexer.

input. Since the cavities are band-pass units, only one channel passes through them and all others are reflected. For this case F2 becomes an individual-station input. The signal is split by the hybrid and passes with minimal attenuation through the cavities to recombine in the output. Any channel (other than F2) sufficiently spaced to be rejected by the cavities will be injected at F1, be split by the hybrid and then reflected by the cavities. The reflected signal recombines in the hybrid to exit through the output port.

An analysis of isolation in this case shows a deficiency in isolation from F2 to F1. The isolation is entirely that inherent in the hybrid. Across a 20 MHz bandwidth at FM that isolation is typically 35 dB. The isolation from F1 to F2 is very good since it is due to the combined effects of the rejection of the cavities (which is >25 dB) and the hybrid (which is 35 dB) for a combined total of 60 dB.

If a module is used for each channel, then the adjacent module provides the additional rejection to increase the deficient isolation from F2 to F1 to 60 dB.

Fig. 18 shows a four station multiplexer that is using band-pass modules to combine three FM stations with a channel 6. The taller module passes channel 6. Two supplementary notches are shown to the left of the channel 6 module to supplement the isolation from channel 6 (aural) to a closely spaced FM channel.

Fig. 19 shows a VHF notch diplexer used to combine visual and aural. It functions as the schematic in Fig. 16 shows.

Fig. 20 shows a UHF waveguide notch diplexer. It also functions as shown in Fig. 15.

In the UHF portion of the band, where the visual transmitter is actually capable of amplifying the audio if diplexed low level, special motor-driven devices may be used in waveguides to detune the notch so that the entire content of the TV channel can be amplified by the visual transmitter, fed into the broadband port and recombined into the output without being attenuated by the notches. This is an emergency configuration for use when an audio transmitter fails.

TRANSMITTER SYSTEMS

In the following section, several methods of connecting one or more transmitters to the antenna system will be discussed. It would be impossible to cover all of the possible combinations. Therefore, only some of the basic configurations will be discussed. These can

Figure 18. Four-channel multiplexer: Ch. 6, plus three FM's. Corpus Christi, Texas.

be modified or expanded to suit the individual station's requirements. Most broadcast equipment manufacturers will be happy to assist in designing custom systems.

A Caution About Interlocks

When designing an RF output system with motorized coaxial switches, it is essential to make sure that the transmitter cannot produce RF power when the switch contacts open. Most coaxial switches are constructed so that the interlock kills the transmitter before the RF contacts open. However, this timing will vary between different types of switches. Also, the time it takes a transmitter to stop producing RF in response to the interlock signal will vary. Therefore, it is a good idea to check this timing. It may be necessary to turn off the transmitter a short time before commanding the coaxial switch to change positions. If the transmitter is still producing RF power when the RF contacts open, the contacts switch will be burned.

Single and Alternate Main Transmitter Systems

Fig. 23 illustrates a single TV transmitter system utilizing a seven-port patch panel. The patch panel would allow the visual or aural component or the diplexer output to be terminated in station load. If the transmitter has multiplex capability (amplifying both the visual and aural signals in the visual amplifier), the visual amplifier could be connected directly to the antenna, bypassing the diplexer.

An alternate main transmitter with a coaxial switch is shown in Fig. 24. A single ended transmitter (such as an AM or FM unit) would connect as shown. A TV transmitter would require two switches.

Parallel Transmitters

Parallel transmitters are two complete transmitters whose output is combined to double the available output power. In addition to the increased output

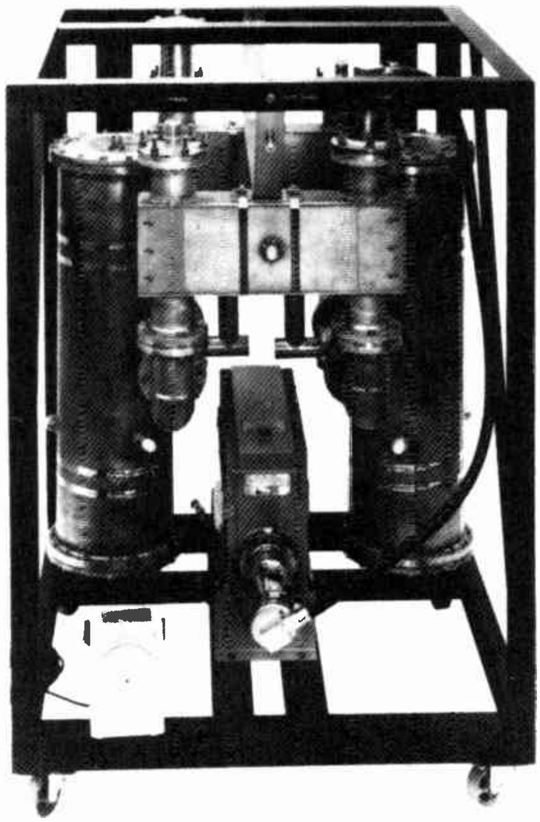


Figure 19. High band VHF notch diplexer.

power, the parallel transmitter has additional advantages, such as redundancy and reduction of ghosting in TV applications.

When one of the transmitters fails, the output power will drop to quarter power. If switching is provided in the output system, the output power can be increased to half power by bypassing the output combiner. This switching can take place at any convenient time after the failure.

In its simplest form, a parallel transmitter consists of an exciter-modulator, input power divider, two amplifier sections, output power combiner, and reject load. (Fig. 25.) In order to properly combine the input signals, they must be of the proper phase and amplitude. These relationships vary with the different types of combiners.

AM and Short Wave Parallel Transmitters

Parallel transmitters are not as common in the AM, and shortwave broadcast bands are not as common as in the higher frequency bands. They are generally used only to double the output power available.

Combiners used at these lower frequencies are usually a bridge circuit. The bridge is usually made of four circuits with a characteristic impedance of 70.7 ohms and constructed with lumped elements (capacitors and inductors). The bridge circuit is shown in Fig. 26A. The phase shifts produced by each leg can either be positive or negative 90°, depending on the elements used. If the shunt elements are capacitors, a negative phase shift will be produced. Using inductors will produce a positive phase shift for the shunts. If the

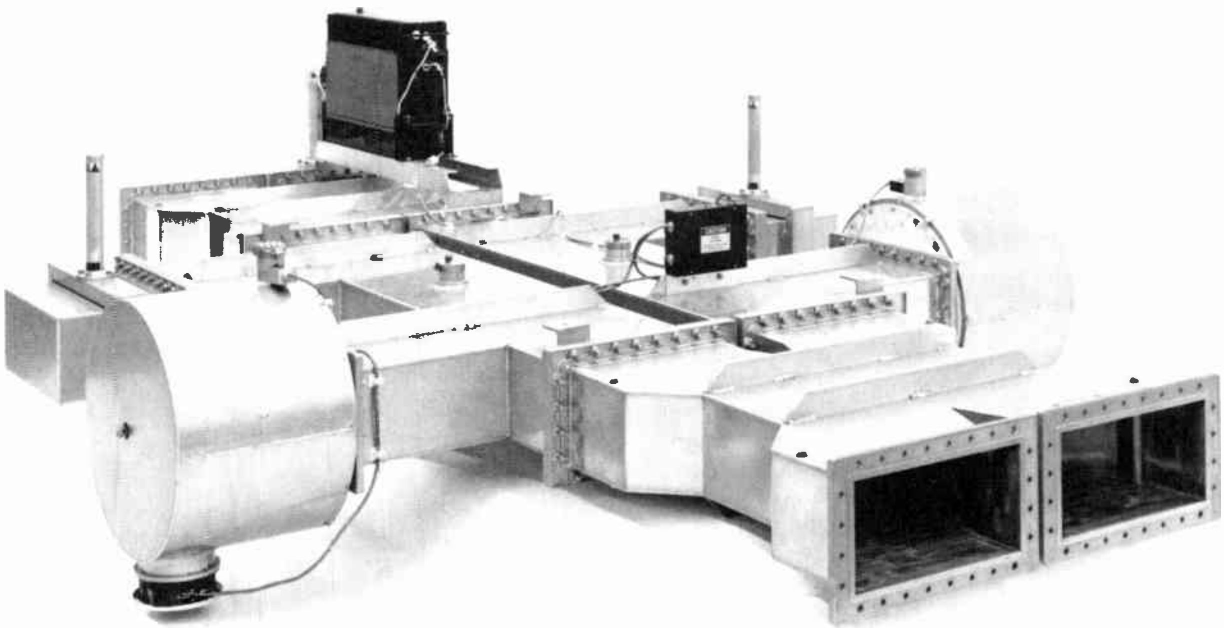


Figure 20. WR1500 waveguide notch diplexer.

Figure 21. UHF low-power TV notch diplexer. (Courtesy Micro Communications, Inc.)

two transmitters are fed to the two inputs in phase, the two signals will be in phase at the antenna output, they will combine. The two input signals will be 180° out of phase at the reject load output, therefore, there will be no power dissipated in the load. When one of the transmitters stops producing power, the power from the remaining transmitter will be split equally to the reject load and the antenna.

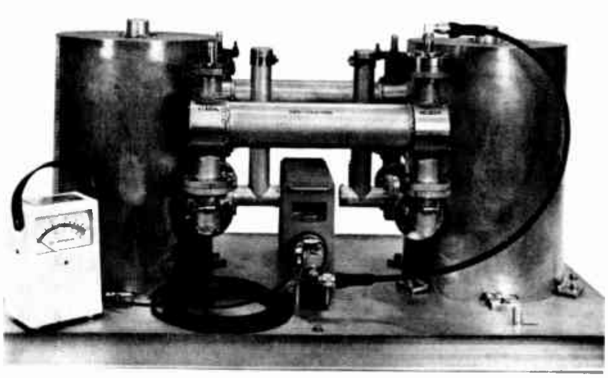


Figure 22. High band low-power TV notch diplexer.

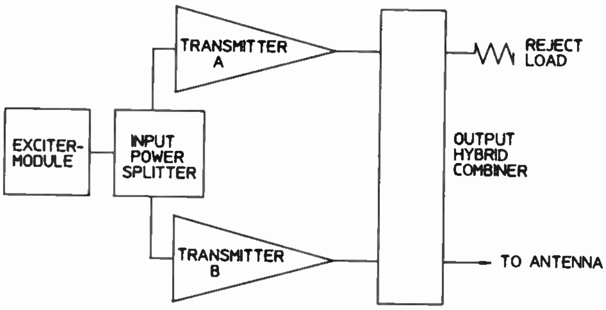


Figure 23. RF flow diagram TV transmitter with seven-port patch panel.

By making three of the legs of the bridge a positive phase shift, the number of circuit components can be reduced. This simplified circuit is shown in Fig. 26B. The two parallel capacitors across the antenna output and Transmitter 1's input can be combined so that only one capacitor is required at each point. The capacitor and inductor that are in parallel across the reject load and Transmitter 2's input will have equal reactance and thus cancel, so neither are required. The bridge combiner can be built with only three capacitors and three inductors, thereby reducing costs. An added advantage of high current carrying devices is that, of the series elements, three are inductors.

Ferrite core transformers may be used as combiners in lower power applications. One parallel transmitter using such a device is shown in Fig. 27. This unit is used to combine two 5 kilowatt transmitters.

The combiner can be thought of as a center-tapped autotransformer that operates at radio frequencies. If equal-amplitude signals that are in-phase are applied to the ferrite combiner, with a common ground, the sum of the two will appear at the center tap signals. Since the two signals are in-phase, there will be no voltage differential across the reject load resistor.

If each input power is 200 watts, for example, there will be 100 volts at 2 amperes across the 50 ohm input of the ferrite transformer. When the two signals are added, current will total 4 amperes. Since the center tap of the transformer has an impedance of 25 ohms, the 4 amperes will produce 400 watts of power. An "L" network is used to match the 25-ohm center tap of the transformer to the 50-ohm output.

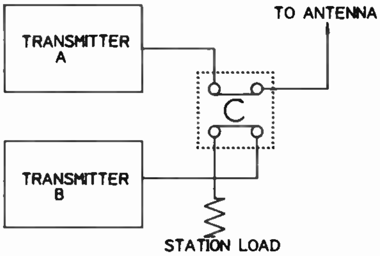


Figure 24. RF flow diagram with alternate main transmitter.

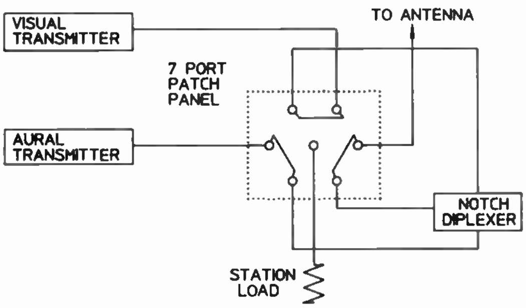


Figure 25. Basic parallel transmitter.

When only one transmitter is operating, the power it produces will be split equally between the antenna and the reject load. The operating transmitter will produce 200 watts, or 100 volts at 2 amperes in the example above. Therefore, there will be only 2 amperes at the center tap of the transformer, which will produce 100 watts of power. Since the other transmitter is not operating, there will be 100 volts across the 100 ohm reject load, which is the remaining 100 watts produced by the operating transmitter.

VHF Parallel Transmitters (Coaxial)

In VHF parallel transmitters, the output combiner is usually a 3 dB, 90° hybrid. In order to properly combine the two transmitter signals, the hybrid requires that the signals be of equal amplitude and phased

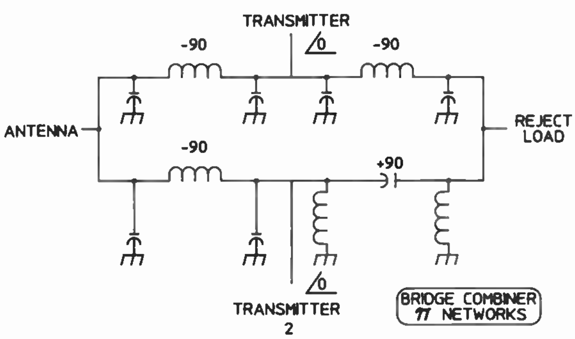


Figure 26A. Bridge combiner PI network.

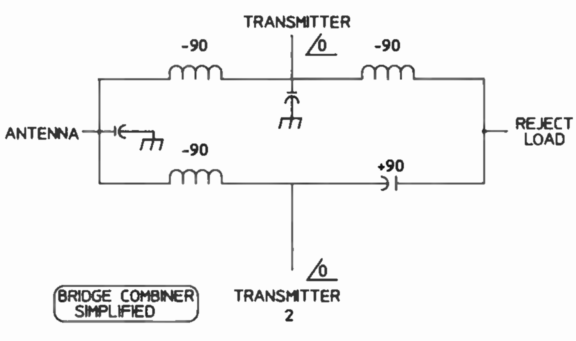


Figure 26B. Bridge combiner simplified.

in quadrature (90°). If we assume that the amplifier sections of the transmitters are identically tuned and that the electrical path lengths and gain are the same, then the input power divider must provide two signals that are of equal amplitude and phased in quadrature. A 3 dB hybrid will provide this type of power division, but other types could be used. An in-phase power divider with a 90° delay in one output would work just as well.

Fig. 28 shows the relationship of phase error to output power of the parallel transmitter system. With no phase error (input signals to the output hybrid in quadrature) 100% of the available transmitter power will be delivered to the antenna. If there is a 90° phase error, both signals in-phase, the power will be divided equally between the antenna and reject load. All available transmitter power will be dissipated in the reject load if the phase error is 180°. A phase error of 20° will only result in an output power reduction of approximately three percent. This would indicate that the phasing is not critical for output power considerations.

The relationship between relative amplitude of the two transmitters and output power is shown in Fig. 29. The graph assumes that one of the transmitters is operating at full power and that output power of the second transmitter is varied from zero to full power. If only one transmitter is operating, the output power will be only 25% of the normal combined output power. With only one input signal, the output hybrid acts as a power divider, applying half the power to the antenna and the other half to the reject load. The power being fed to the antenna is 25% of the normal combined transmitter power. If one of the transmitters is operated at half its normal output power while the other transmitter is operated at full power, the combined output power will be approximately 73% of the normal combined power. Since the two transmitters are only generating 75% of the normal combined power, only about 2% of the power is being dissipated in the reject load. Therefore, not much power is wasted in the reject load. The maximum power that the reject load should be required to dissipate is half of one transmitter's power.

Thus far only the basic parallel transmitter system has been discussed. It was assumed earlier that the amplifier sections of the transmitters were identically tuned, having the same electrical path lengths and gain. From a practical viewpoint this could be done, but with difficulty. Therefore, most parallel transmitter systems provide a means of controlling the phase and gain of the transmitters that is independent of the tuning. This usually means an attenuator for gain and a phase shifter for the phase. Some transmitters use a gain control within the amplifier section and one of the input matching controls of the power amplifier to accomplish these adjustments. This practice is probably more common in FM than TV since the FM bandwidth is smaller.

One of the major advantages of parallel transmitters in television is the reduction of ghosts or reflections

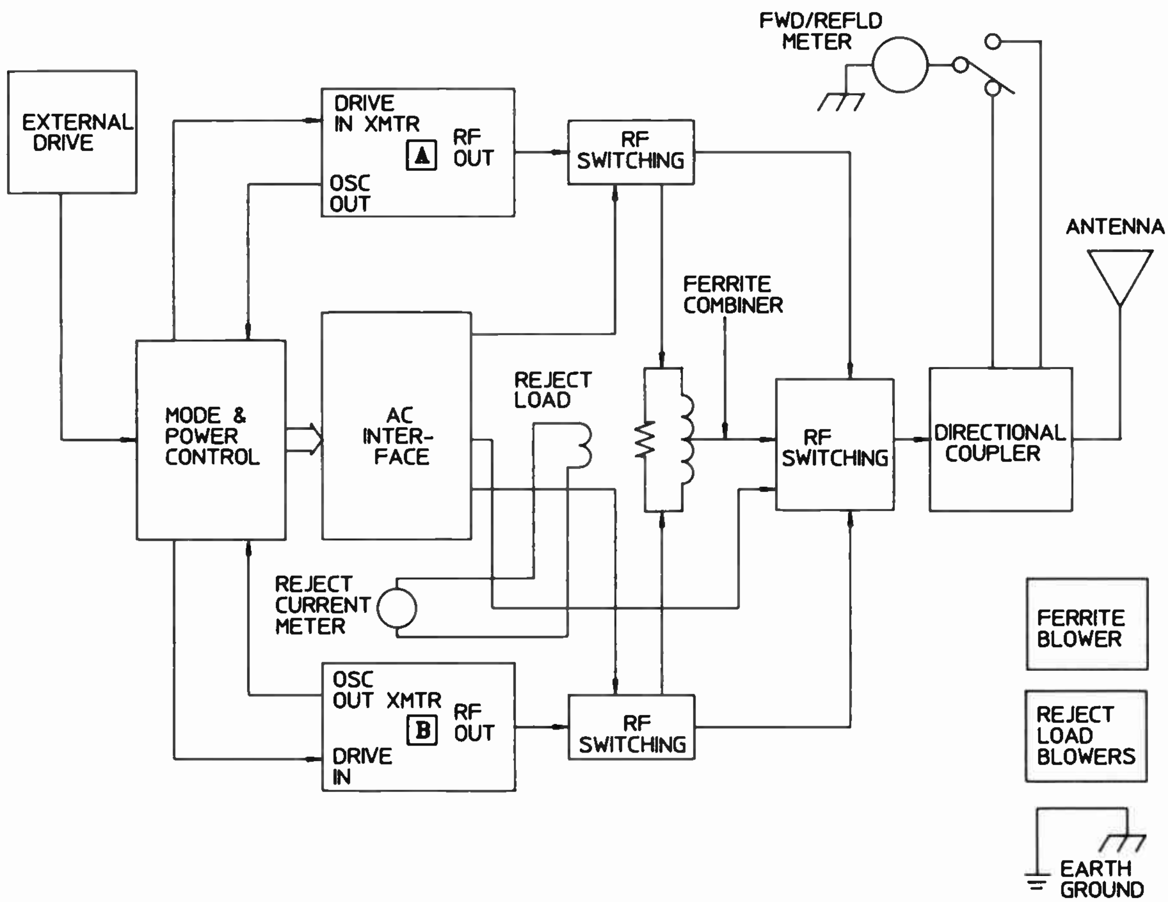


Figure 27. Simplified block diagram of combiner and RF switching using a ferrite core transformer. (Courtesy Harris Corporation.)

from the antenna. Ghosts occur when a portion of the power is reflected from the antenna back to the transmitter, re-reflected, and finally radiated. The distance between the original image and the ghost as shown on a TV receiver can be used to determine the approximate location of the reflection in the transmitting antenna system. Since the horizontal frequency of the TV is 15,734 kHz, the full horizontal line would be equal to 63.6 microseconds. The length of the visible portion of the horizontal line is 53.1 microseconds. Therefore, the time between the image and the ghost can be measured. This time, when compared to the speed of light, will yield the distance the reflected signal had to travel to produce the ghost.

It must be kept in mind that the reflected signal had to travel from the point of origination down the

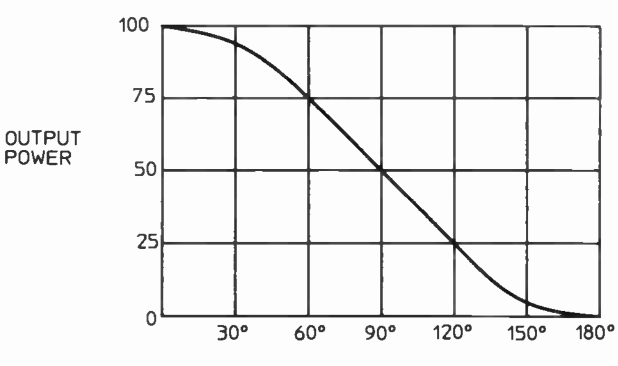


Figure 28. Relative output power versus phase error.

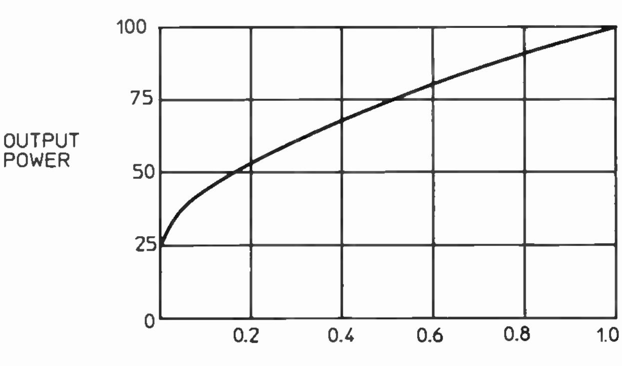


Figure 29. Relative output power versus amplitude ratio.

transmission line to the transmitter and back up the transmission line to the antenna to be radiated. The velocity of propagation of the transmission line must also be taken into account.

The use of the 3 dB hybrid in the parallel transmitter will reduce the reflected signals that produce the ghosts. When the reflected energy from the antenna is applied to the output of the 3 dB hybrid, it is split into two signals at the two transmitter inputs. These signals will be phased 90° apart and will continue until they are re-reflected by the output circuitry of the transmitters. The signals will then be applied to the inputs of the 3 dB hybrid. However, their phases are such that instead of combining in the antenna output, they will combine in the reject load of the parallel transmitter. For optimum ghost reduction the electrical path lengths between the hybrid and transmitter inputs must be the same. The use of slugs or other tuning devices can upset the phase balance or electrical length of the system.

One method to measure the effectiveness of a system for ghost reduction is to measure the reverse VSWR of the system. This is accomplished by placing open or short circuits on the transmission lines that would connect to the transmitter outputs and measuring the VSWR at the output of the combining system. Thus, the path of the ghost signal is being measured. Ideally, this path should be as good as the forward VSWR of the system. However, from a practical standpoint, a VSWR of 1.1:1 or better will reduce the ghosting. It should be noted that equal electrical line lengths are needed for quadrature type combining networks. For systems using in-phase type combining networks, there must be a 90° delay in the proper input. To offset this delay, a 90° delay can be inserted in the input circuitry of the opposite transmitter.

Fig. 30 shows a complete output switcher for a parallel television transmitter. For single-ended transmitters such as used in FM, only half the system would be required. This figure shows the switching options of the output hybrid combiner. There are normally four modes of operation:

1. A and B combined to the antenna.
2. A and B combined to the station load.
3. A to the antenna and B to the station load.
4. B to the antenna and A to the station load.

As drawn, the diagram is shown in the A&B to the antenna mode. By rotating S3 and S6, the system is changed to the "A and B combined to the station load" mode. If S1, S2, S4, and S5 are changed, Transmitter A will be connected to the antenna while Transmitter B will be connected to the station load. If S3 and S6 are rotated, the transmitters will switch Transmitter A to the station load while Transmitter B is connected to the antenna.

Since a parallel transmitter is really two complete transmitters, there should ideally be two exciter-modulators. Some manufacturers may offer the second exciter-modulator as an option. By adding a switch on

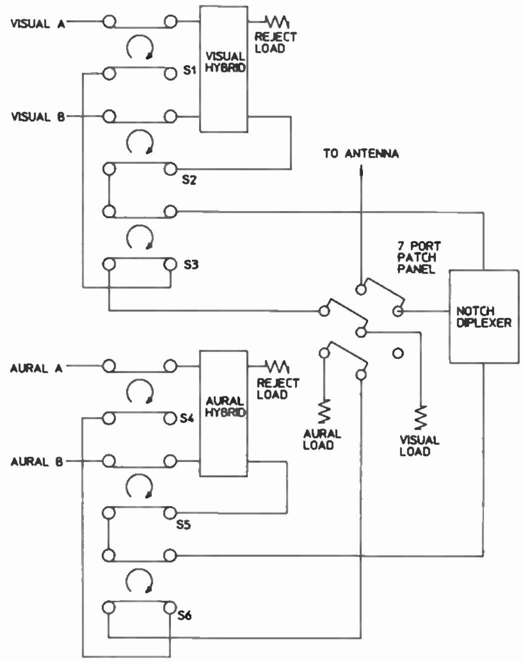


Figure 30. RF flow diagram for output switcher.

the input of the input power divider, either exciter-modulator could be selected. This would provide redundancy, should the active exciter-modulator fail. Since the power level of the exciter-modulator is usually fairly low, the switching could be done under power, which would allow the switching to be automatic. By using relatively fast switches, the transfer could be done with only a small carrier interruption.

Automatic operation of the output switching system is not usually done. Stations would rather choose when the carrier break occurs, since it will be noticeable to the audience. In lower power installations, the carrier break will be two seconds or less. With higher power switches the break could be up to ten seconds.

Again referring to Fig. 30, the combined output of the visual hybrid must pass through S2 and S3. Therefore, these switches must be sized to carry the combined power. Some systems will add another coaxial switch (S7) on the output of the visual hybrid as shown in Fig. 31. This higher power switch will allow the other three switches to be of a smaller size, since the combined power will only be applied to S7. It would then be necessary to add another station load for the combined signal at S7. Generally, this will reduce both the physical size and the cost of the parallel transmitter system. A television system requires a diplexer of some type in order to combine the aural and visual signals. The output switching system usually contains a patch panel which allows the output of the diplexer to be routed either to the antenna or the station load. The output switching system must also contain the necessary monitoring points for combined power and reject power.

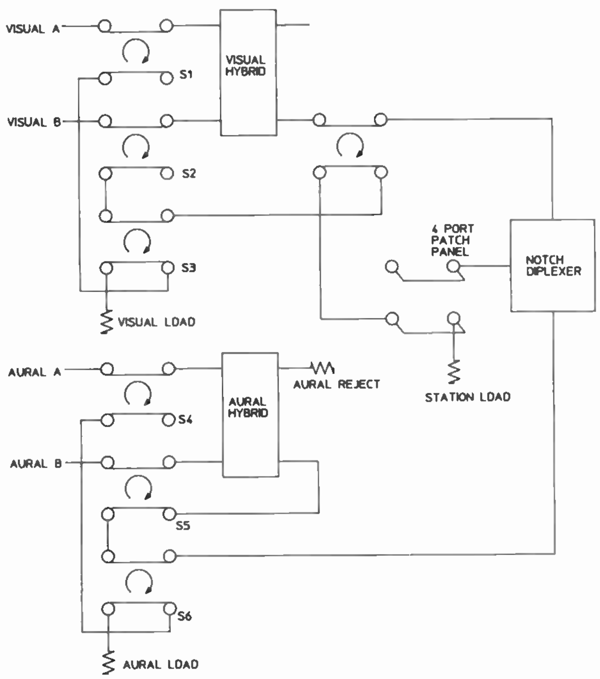
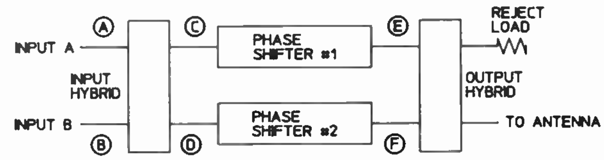


Figure 31. RF flow diagram for high power output switcher.



NOTES:

1. INPUTS A+B ARE IN PHASE
2. PHASE SHIFTER #2 WILL USUALLY CONTAIN A FIXED 90° SECTION AND A VARIABLE 90° SECTION.

MODE	PHASE SHIFTER #1	PHASE SHIFTER #2
A+B ANTENNA	0	90
A-ANTENNA B-LOAD	90	90
B-ANTENNA A-LOAD	0	180

Figure 32. Schematic of basic switchless combining system.

VHF Switchless System (Coaxial)

The switchless system also combines the outputs of two transmitters operating in parallel.

As with the switch method, it provides four modes of operation:

1. A and B combined to the antenna.
2. A and B combined to the station load.
3. A to the antenna and B to the station load.
4. B to the antenna and A to the station load.

There are several methods to accomplish these mode changes without the switches. A basic system is shown in Fig. 32 and consists of two 90° hybrids, a reject load, and some type of phase shifting device. The method used to shift the phases defines the various options. Therefore, the basic system will be presented here, and then the different methods of accomplishing the phase shift will be discussed.

Since the switchless system contains two hybrids, one must understand how the hybrid operates. The 90° hybrid can be used as either a power divider or a power combiner. In the switchless system both are used.

When used as a power combiner, the hybrid will combine two signals that are equal in amplitude and phased in quadrature. Again referring to Fig. 32, if two equal amplitude signals are applied to points E and F, with the signal at F lagging in phase by 90°, the signals will be combined into the antenna. Conversely, if the signal at point E is lagging by 90°, the signals will be combined into the load.

If a signal is applied to the A input of the hybrid (now being used as a power divider), it will be split into two signals of equal amplitude that are phased 90° apart. The signal appearing at Point C will be in phase with the input signal (Point A), while the signal at Point D will lag the input signal by 90°.

The converse is true for a signal being fed into the B input. The signal at Point D will be in-phase, while the signal at Point C will lag 90°. If two signals that are in-phase are applied to inputs A and B, then each of the outputs (Points C and D) will have two signals, one in-phase with the inputs and one lagging by 90°.

If the switchless system is set in the A and B to the antenna mode, phase shifter #2 must have 90° more phase shift than phase shifter #1. The two signals at Point F will lag the two signals at Point E. In this situation, the signals will combine in the antenna output.

If either of the phase shifters are set for an additional 90° of phase shift, the system will be in one of the single transmitter modes of operation. If phase shifter #1 is changed, then Transmitter A will be routed to the antenna, while Transmitter B will be terminated in the station load. Should phase shifter #2 be changed, Transmitter B will be routed to the antenna, while Transmitter A is terminated in the station load.

There are three methods of sending output from the combined transmitters to the station load:

1. A coaxial switch or patch panel could be used at the antenna output to route this output to a separate station load.
2. The combined transmitters could be routed to the reject load by moving the additional 90° phase shift in phase shifter #2 to the #1 phase shifter.
3. The transmitter outputs could also be combined into the reject load by changing the input phases to the switchless system. By adding 180° delay to

the A input, the combined transmitter output will be routed to the reject load.

The latter two methods will require that the power rating of the reject load be increased to the combined power level instead of that of a single transmitter. They also require that the antenna receive any power that is not absorbed in the reject load.

There are several methods of changing the phase of the RF signals in the switchless system.

1. Probably the most familiar method of changing the phase at higher power levels is the line stretcher. It is a piece of transmission line whose length can be changed. For convenience, it often takes the form of a U-link, or trombone, so that the connectors can be mounted and the U-link moved to change the length.
2. Another method uses a 90° hybrid with movable short circuits on two of the arms (Fig 33). If a signal is applied to the input (Point A), the hybrid will divide it into two equal signals that are phased 90° apart. The signal at Point D will lag by 90°. The short circuits attached to Points C and D will reflect the two signals back to the hybrid. The phase of the two signals will be delayed by twice the electrical length of the short circuits. If the two short circuits are the same length, the relative phases of the two signals will still be 90°. Since the signal at Point D is lagging by 90°, the two signals will combine into the output (Point D). By changing the length of the short circuits, the delay or phase shift through the circuit will change. If the short circuits are moved 90°, the phase shift through the circuit will change. If the short circuits are moved 90°, the phase shift through the circuit will be 180°. The signal must travel from the hybrid to the short circuit and then return to the hybrid, twice the distance of the short circuit. The two short circuits must be moved together in order to make the hybrid combine the reflected signals properly.
3. Fig. 34 shows several methods of making a short circuit needed in the above circuit. It could be the traditional short (i.e., a piece of movable metal contacting the inner and outer conductors of the transmission line).
4. A *deep short* could be used which moves the ringer contacts one quarter wavelength away from the short. This greatly reduces the amount of current that the fingers are required to carry. A *noncontacting short* may be used. It is a pair of cylinders that are a quarter wavelength long and shorted at one end. The sizes of the cylinders are such that they fit between the inner and outer conductors of the transmission line. The shorted cylinders are insulated from the transmission line, thus forming a capacitor. The capacitor is large enough to have very little impedance at the operating frequency, therefore, it appears as a short circuit.
5. Another method of creating phase shift is to use an open circuit. A short section of the center conductor is removed from a shorted piece of transmission line. The transmission line appears as an open circuit, since there is very little capacitance between the two pieces of cut center conductor. By moving an insulated metal probe across the gap in the cut center conductor, a large amount of capacity is created. This causes the transmission

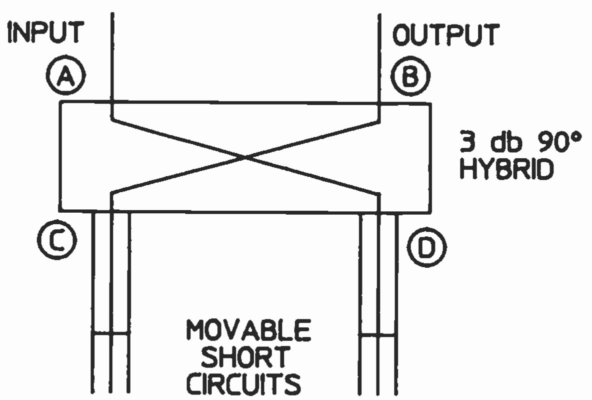


Figure 33. Variable phase shifter assembly.

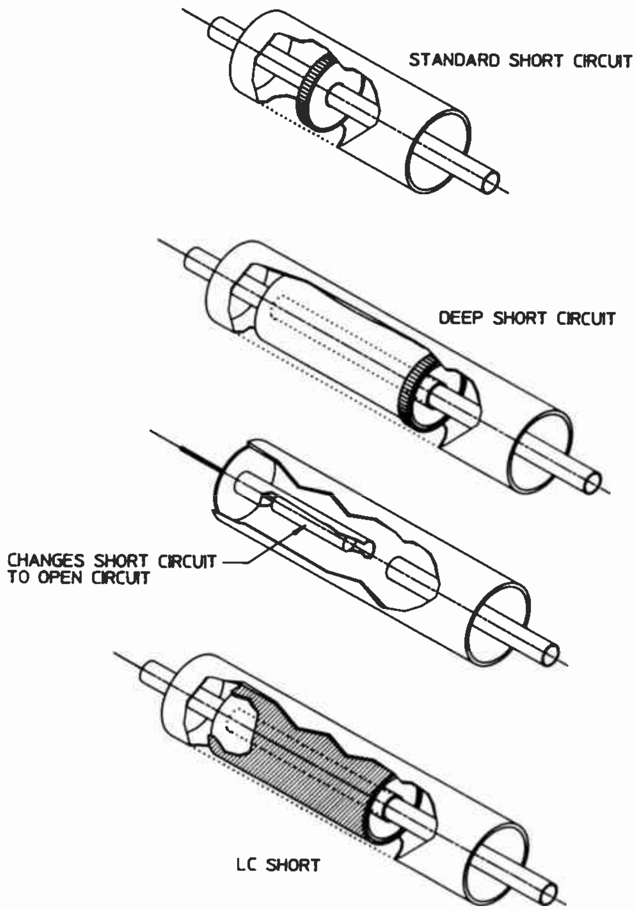


Figure 34. Variable phase shifter options.

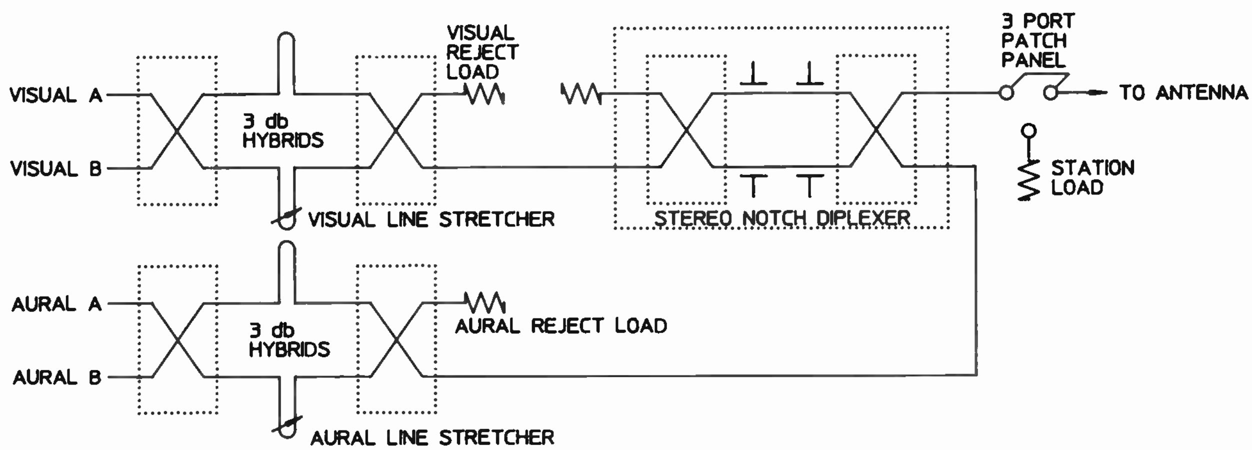


Figure 35. RF flow diagram VHF switchless system with diplexer.

line to appear as a short circuit, thus changing the phase.

All of the above methods will change the phase and allow the switchless system to operate. Each method has advantages and disadvantages that should be taken into consideration upon purchasing a switchless system.

A typical flow diagram for a VHF switchless system is in Fig. 35, and a diagram for a coaxial Opto-SX is shown in Fig. 36.

The switchless system does not require that the transmitters be turned off during the switching process. Since there is no carrier break, the length of switching is not important. Switching may be be done with no regard to program content.

The switchless system will offer the same ghost reduction capability as other types of parallel transmitters. Since the two signals from the transmitters are applied to the switchless system in-phase, there must be an external phase shift to take advantage of the ghost reduction feature. A 90° phase shift must be added between the output of one of the transmitters and the switchless system. This will delay the reflected signal in that path 90° as it passes from the output system to the transmitter, and another 90° as it passes from the transmitter back to the output system. The two reflected signals will now be 180° out-of-phase, and thus be combined in the reject load. Since a 90° delay was added to one of the transmitter signals, it will be necessary to add an equal phase shift to the other transmitter so that the two signals will combine in the antenna output. This delay can be added to the input circuits, allowing the ghost reduction circuit to operate properly.

Safety Considerations

There are some safety aspects of the switchless system that were not present in the switch type

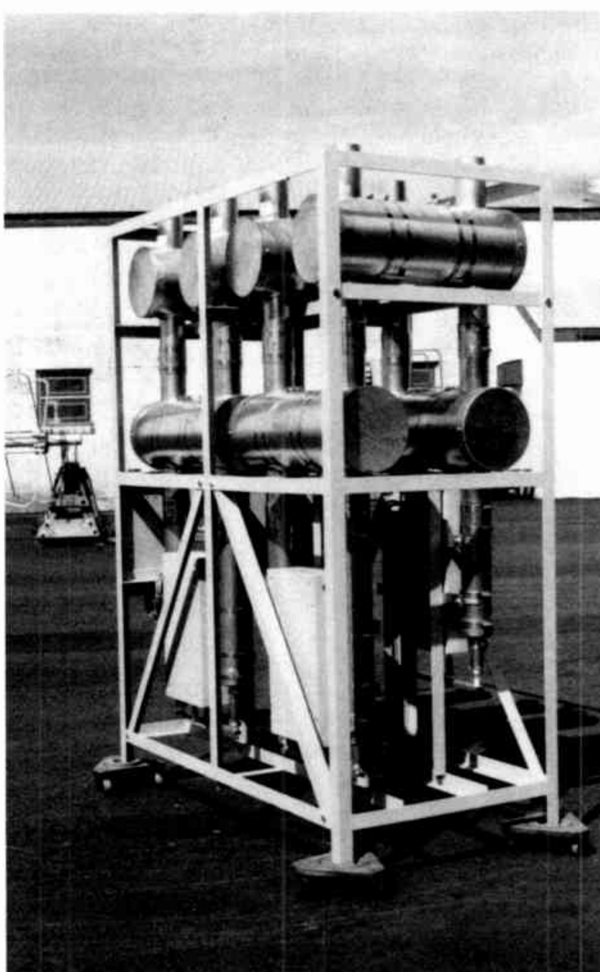


Figure 36. VHF switchless "Opto-SX" switcher. Includes aural and visual inputs.

systems. In a switch type system, the transmitters were isolated by the mechanical switch when operating in the single transmitter mode. The switchless system does not have that isolation. Its isolation is provided only by the two hybrids. Therefore, the 60 dB or better isolation provided by the coaxial switches, is not present in the switchless system. It is possible that voltages could be present on an input, even through the transmitter driving that input is turned off. It is necessary to make sure that safety devices are used to protect technicians. It would be a good idea to delay maintenance on the system until both transmitters could be turned off.

UHF Parallel Transmitters (Waveguide)

UHF parallel transmitters are different from their VHF counterparts. Because of the power rating of the klystron (30 kW to 60 kW), purely parallel transmitters are not as common as in VHF. While there are quite a few 110 to 120 kW transmitters operating, these are not true parallel transmitters, as only the visual klystrons are operating in parallel. A 220 to 240 kW transmitter is the more common parallel UHF transmitter. The UHF transmitters also have aural multiplex capabilities, which reduce the need for parallel aural amplifiers.

A typical RF flow diagram for a waveguide output switching system is shown in Fig. 37. It contains four waveguide transfer switches, a hybrid combiner, and a reject load. S1 and S2 switch the input of the system around the hybrid for single transmitter operation. S3 determines whether the transmitter signals will be routed through the diplexer, while S4 routes the output of the diplexer to the antenna or station load. The aural amplifier is routed directly to the diplexer. This system would have the four basic modes of operation:

1. Visual A and B combined to the antenna.
2. Visual A and B combined to the station load.
3. Visual A to the antenna and visual B to the station load.
4. Visual B to the antenna and visual A to the station load.

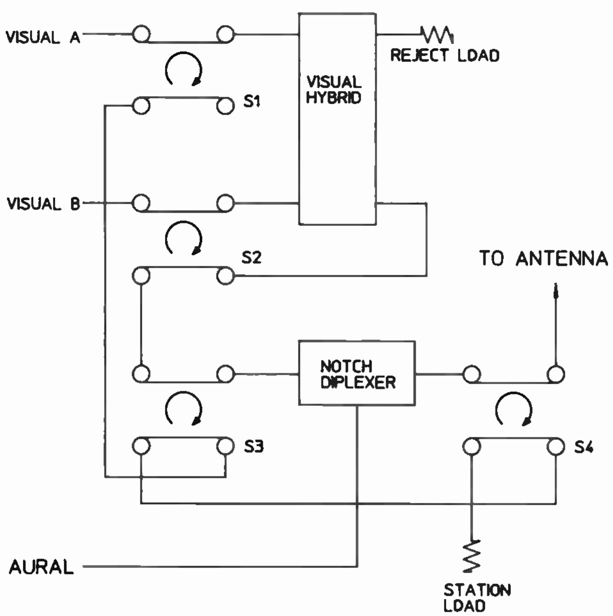


Figure 37. RF flow diagram of output switcher with parallel visual amplifiers.

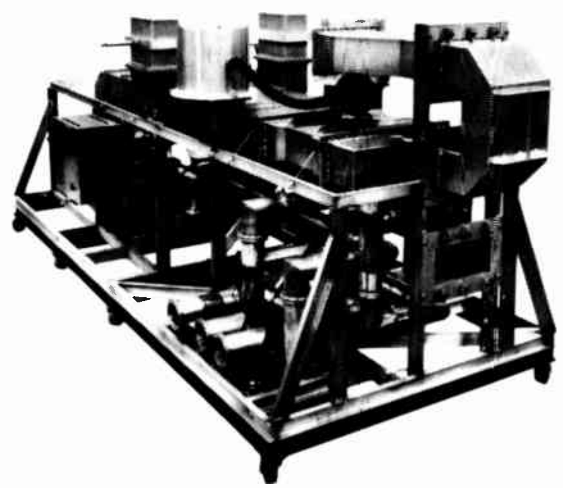


Figure 38. Output switcher 100 kW UHF transmitting system.

In addition, each of the modes could be operated either multiplexed or normal.

Recent developments in diplexers have simplified the switching systems. Aural notch detuners allow the multiplexed signals to pass through the notch diplexer rather than being switched around it. The detuners will raise the aural notch cavity's frequency so that it is above the aural carrier. This allows the multiplexed visual carrier to pass through the diplexer, just as the visual signal does normally.

A station's RF output switching system is usually unique. Either the RF layout of the system will be slightly different, or the mechanical layout of the system must be adapted to the transmitter building. Since this is usually the case rather than the exception, most manufacturers are equipped to handle these situations. Fig. 36 shows a system with coaxial patch panels and switches on the input and a waveguide patch panel on the output of the diplexer.

UHF Switchless System (Waveguide)

One of the more recent developments in RF switching is the switchless phase-shifter systems. These combine the outputs of higher power UHF amplifiers and can be used for either aural or visual service. This system consists of a 3 dB, 90° hybrid, a dual phase-shifter, a reject load, and a magic tee (180° hybrid) and is shown in Fig. 39. To better understand how the system operates, it is essential to examine the individual components.

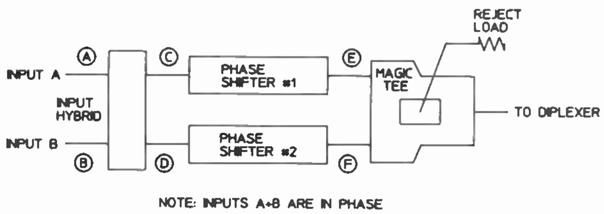


Figure 39. Block diagram UHF switchless switcher.

The magic tee (180° hybrid) works much like a 90° hybrid. The difference is the amount of phase difference which it provides. When it is used as a splitter, the outputs will either be in phase or 180° out-of-phase, depending on which input is being driven. When used as a combiner, the magic tee will combine signals that are in phase into the main output; if the signals are 180° out-of-phase it will route them to the coupled output.

Each phase shifter consists of a movable piece of dielectric material inside a section of waveguide. When the dielectric material is against the side wall of the waveguide, the phase shift will be at the minimum. As the dielectric material is moved toward the center of the waveguide, the phase shift of the signal going through the waveguide will increase. By choosing the type and amount of dielectric material, the device can be adjusted to produce 90° of phase shift. The amount of phase shift with the dielectric material against the side wall is not important in this instance. The critical element is that the phase shifts through the two units be equal.

Assume that there is no phase shift through the unit when the dielectric material is against the side wall and 90° when it is in the center of the waveguide. If a signal is applied to the A input of the hybrid, it will be split into two signals of equal amplitude that are phased 90° apart. The signal appearing at Point C will be in-phase with the input signal (Point A), while the signal at Point D will lag the input signal by 90°. The relationship of the signals at C and D will be reversed for inputs at B. If the result of each of these scenarios is superimposed on one another to simulate the normal mode of operation with signals at both A and B, the outputs C and D each contain two signals at equal amplitude and 90° phase difference. The hybrid effectively causes these signals to be vectorially summed with an equivalent output of amplitude = 2, and phase = −45°. Although these signals have been delayed 45° with respect to the input, they are now in phase with one another at ports C and D.

With the assumption that the phase shifters have no phase shift when set to the minimum position, the signals at the output of the input hybrid (Points C and D) would be applied to the input of the magic tee (Points E and F). There will be signals present at both Points E and F that are in phase with the input signal. The signal at Point E was produced by amplifier A, while the signal at Point F was produced by amplifier B. Since these signals are equal amplitude and in phase

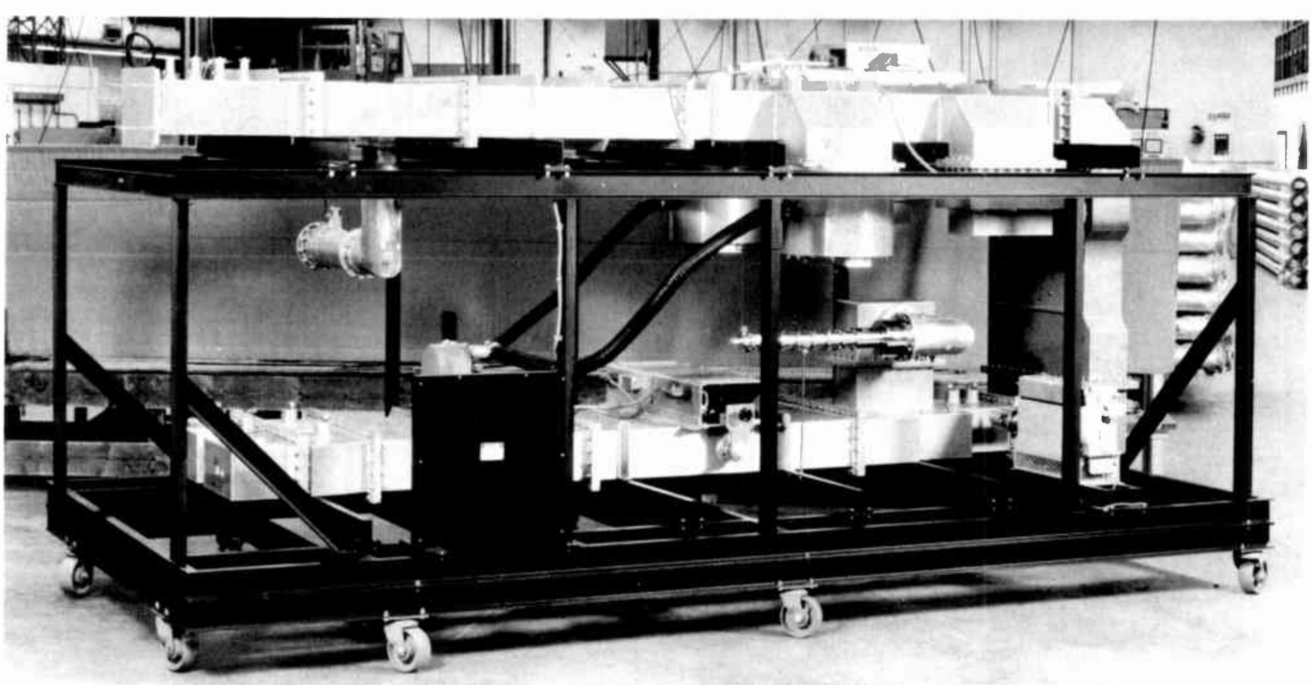


Figure 40. UHF 120 kW switchless diplexer.

they will combine into the main output of the magic tee. At the same time, the other two signals at Points E and F are equal in amplitude and in phase they will combine into the main output of the magic tee. At the same time, the other two signals at Point E and F are equal in amplitude and in phase, therefore, they will combine just as the other signals did. Since these signals are not in phase at the output of the magic tee, the resultant signal will be the vector sum.

By inserting 90° of shift in phase shifter #1, a new mode is created. The phases at the output of the hybrid (Points C and D) will be the same as before. Since 90° is added between Points C and E, the phases at E will now be 90° for the signal from input A and 180° for the one from input B. Therefore, on the inputs of the magic tee, the signals from input A are in phase, while the signal from the B input are 180° out-of-phase. The in phase signals will add at the main output of the magic tee while the signals that are 180° out-of-phase will add in the coupled output.

If phase shifter #1 is returned to the minimum position and phase shifter #2 is set for 90° of phase shift, the signals from the B input will add in the main output, while the signals from the A input will add in the coupled output. By adding the phase shift in the #2 phase shifter, the signals at Point F will be 90° for the one from input B and 180° for the one from input A. Thus the signals from input B will be in phase at the inputs of the magic tee, while the ones from input A are 180° out-of-phase.

Thus far three modes of operation have been accomplished within the switchless system. The fourth mode, both transmitters combined to the station load, can be done two ways. If the phase of the input signals are changed from in phase to 180°, the combining that takes place will be applied to the coupled output (reject load) rather than to the main output. The second method is to add a switch to the output of the system to allow the output to be connected to the antenna or the station load. The latter seems to be the more common method at the present time.

The switchless phase shifter can be changed under power, since there are no contacts to break and make. The only part that is moving is the dielectric material

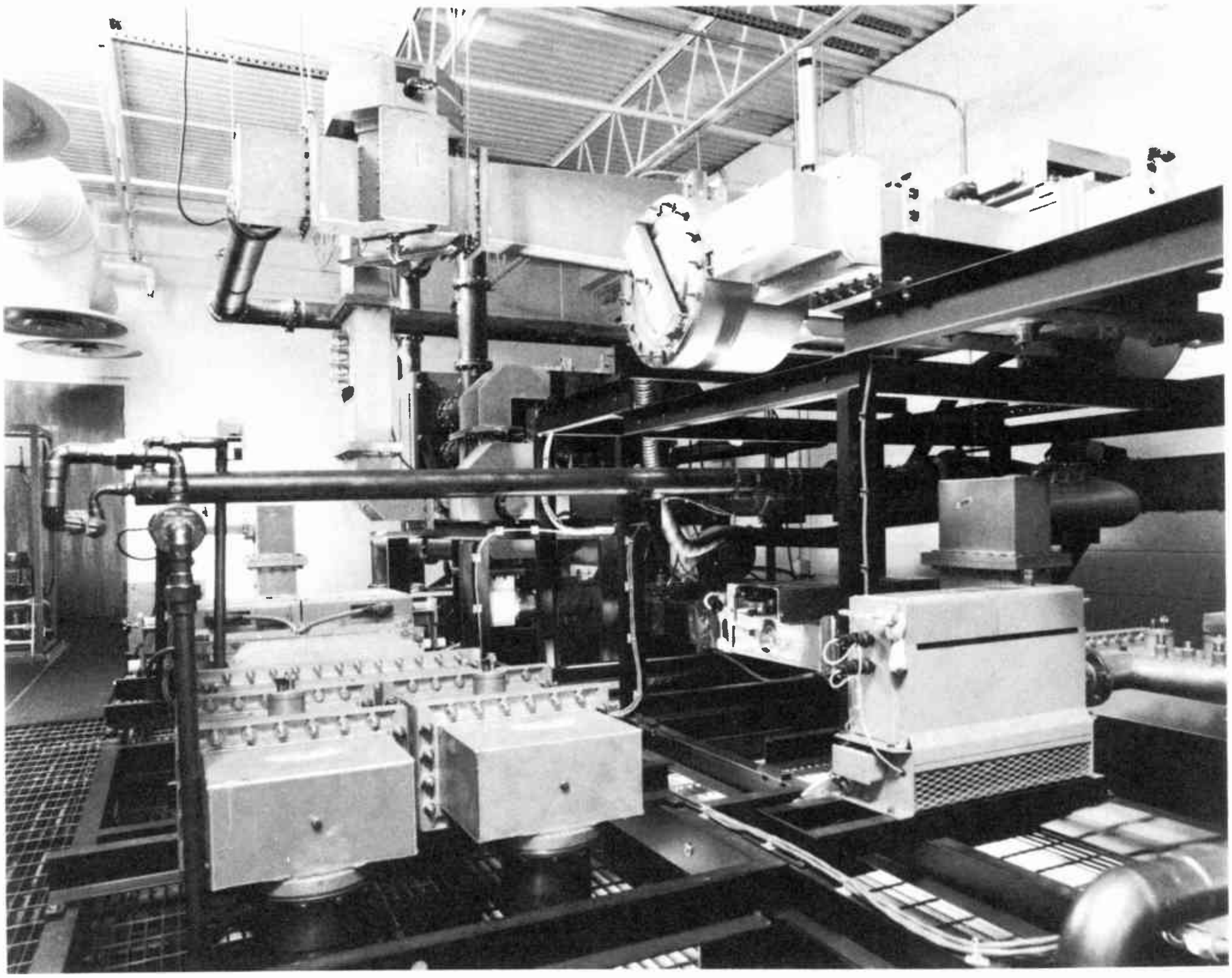


Figure 41. 180 kW three-tube switchless at channel 50 in Tampa, Florida. (Courtesy of Home Shopping Network.)

in the phase shifter section. Since the full power of one transmitter can be applied to the reject load, the load must be capable of handling that power.

In the past, switching systems changed the RF path by mechanically moving the parts of a switch. A typical switch will have around 60 dB of isolation between the paths of the switch. In the switchless system there is no mechanical isolation. Isolation is provided by the the hybrid and by the magic tee within the system. Typically, this isolation will be on the order of 35 to 45 dB, which is sufficient for good performance. However, the operator must be aware that the isolation that he provided in a switch-type system is not there in the switchless system. Therefore, when maintenance is being performed on the system of one transmitter, there may be a small amount of power present from the other transmitter.

Three-Tube Switchless Combiner

A recent addition to UHF switchless systems is the three-tube switchless combiner, introduced by Dielectric in 1988 and covered by U.S. Patent #4,951,013. Until this time, it was not recognized that widely different amplitude of visual inputs could be combined with a switchless system. Once two tubes of a 180 kW system have been combined, the resulting 120 kW must be combined with the remaining 60 kW. The three-tube switchless enables other phase choices in addition to 0°, 90°, and 180° from the phase shifters as shown in Fig. 39. By allowing for 19°, for example, the high efficiencies of a two-tube system may be obtained with the three-tube system. (See Fig. 41.)

This approach is also usable at VHF should the need arise.

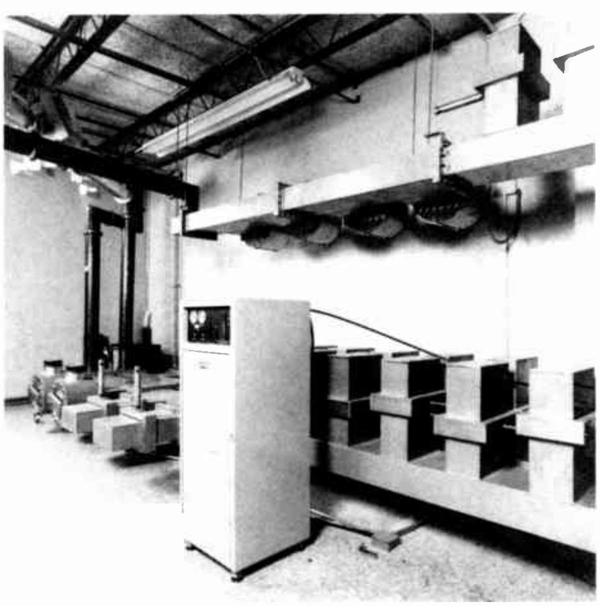


Figure 42. Aural and visual filtering at channel 69 in Hollywood, Florida. (Courtesy of Home Shopping Network.)

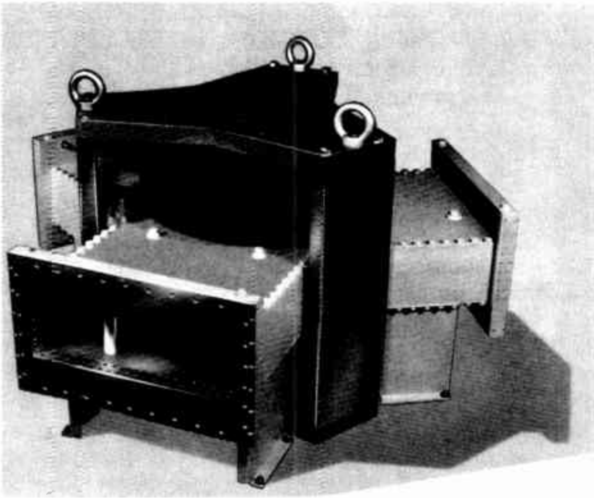


Figure 43. High-power UHF isolator. (Courtesy of Micro Communications, Inc.)

Special Filtering for UHF

The industry is aware of potential problems channels 14 or 69 are installed in localities where land mobile services must coexist. Recent success in planning and installing a channel 69 system in Hollywood, Florida, may allay the fears of those contemplating such an installation. An extensive review was provided by Beifus, Harbaugh, and DeCormier in *Channel 69 Filtering for Land Mobile Compatibility* in the 1990 44th Annual Proceedings of the NAB.

High power filters were provided on the visual and aural inputs to supress intermodulations. (See Fig. 42.)

The super power isolator has also been introduced recently to the broadcast community. (See Fig. 43.) The isolator is a nonreciprocal three-port device. A wave entering at port 1 appears only at port 2 and a wave returning from port 2 appears only at port 3. This circulating action occurs because of the deflection of the wave by a nearby magnetized ferrite cylinder. It is promoted to offer impedance stabilization, ghost elimination, hot switching, and harmonic rejection. The performance is temperature-dependent, requiring a closed loop cooling/heating system.

Multi-Transmitter Different Frequencies Coaxial System

Within the FM band there is a growing demand for multiplexing systems that provide more than the capability of simply combining several stations. They also provide a spare broadband port for emergency use should any module in the system fail and either dual outputs with switching should either half of the antenna fail and extended monitoring capabilities or both.

The basic functions of a multiplexer module have been discussed earlier in this chapter as a category of filters. Here the peripheral features will be explored in

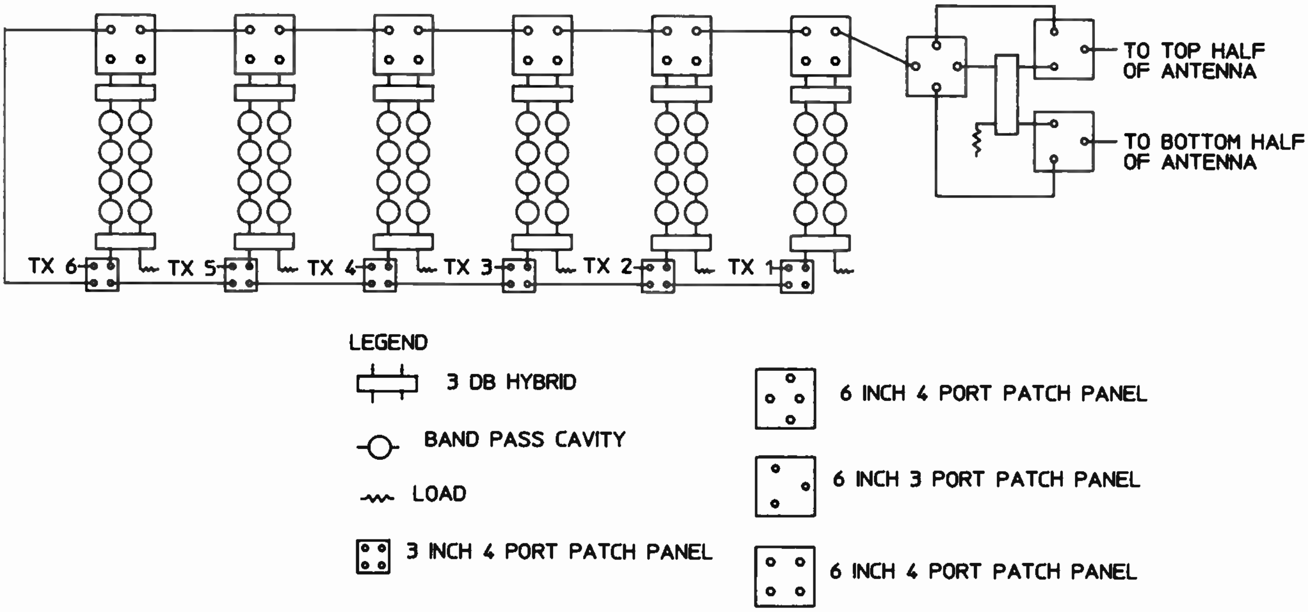


Figure 44. Schematic for six-station FM band-pass constant impedance multiplexer.

Figure 45. Six-station FM band-pass constant impedance multiplexer at St. Louis. (Courtesy EZ Communications, Inc.)

an effort to examine additional benefits of a multiplexer system.

Whether the diplexer modules in the system are made of band-pass or band stop cavities, the system can be built with N modules or N-1 modules where N is the number of channels to be combined. Building the system with N-1 modules is the most economical way while N modules provide a spare broadband input. The spare input can be used as an emergency input if any of the modules fail for any reason. A switch or patch panel is needed both on the input side and output side of each module. See Fig. 44 for a typical schematic of this system.

The band-pass system for FM applications has demonstrated superior performance; in particular, with regard to amplitude response, group delay, and intermodulation rejection. The following are nominal performance characteristics:

- Group Delay +25 ns +150 kHz
- Amplitude Response < 0.1 dB +200 kHz
- Intermodulation Rejection > 80 dB

Fig. 45 depicts a six-station band-pass FM combiner in St. Louis.

The sharing of costs for a system among all participants often makes it possible to consider monitoring systems that one station might have considered too costly. At a minimum, each station must have individual fail-safe power to trip the transmitter. Some broadcasters may consider a computer monitor system superimposed on the fail safe. The computer can monitor forward and reflected power; status of interlocks and heat sensors; and can be programmed to transmit status hourly, daily, or upon an unusual event through a modem to a responsible individual.

All of the features mentioned need to be addressed with the philosophy of the participants in mind. These systems are all custom-made using components that have been standardized over years of use in the industry. Extensive discussions with manufacturers or suppliers are necessary to specify sizes, power capacity, VSWR, insertion loss, bandwidth response, group delay, patch panels, switches, monitoring, interlocks, heat sensors, cooling, and layout to meet individual requirements. When seeking bids, it is prudent to specify all of the above parameters so that all bidders are quoting the same system. These systems can be built with all the bells and whistles or stripped to a bare minimum.

Figure 46. RF monitor and control system in St. Louis. (Courtesy of EZ Communications, Inc.)

Fig. 46 depicts a six-station monitor and control system. The system monitors forward and reflected power for each station, and the combined forward and reflected power on each of the dual outputs to the antenna. The system provides local and remote warnings when VSWR rises to 1.25:1 and trips at 1.50:1 VSWR. It also generates reports at prescribed time intervals and on the occurrence of an event. This makes analysis easier in the event of a failure.

2.5
AM Broadcast Antenna Systems
Part I: System Design

Carl E. Smith
Carl E. Smith Electronics, Cleveland, Ohio

INTRODUCTION

Standard broadcast (AM) antenna systems can reach a high degree of sophistication, much of which is based on advanced mathematics. The purpose of this section is to provide the station engineer with an understanding of some of the basic concepts of antenna design and an appreciation for the complexities of this specialty. An extensive bibliography is included for those who wish to pursue further study.

The chief purpose of a broadcasting antenna system is to radiate efficiently the power supplied to it by the transmitter. A simple antenna can do this job quite well. This is often a single vertical tower that radiates its signal equally in all directions along the ground in a so-called nondirectional or omnidirectional pattern. A second purpose of an AM antenna system is often to concentrate the power in desired directions to cover populated areas, and to suppress it in other directions to protect the coverage of other stations sharing the same or closely-adjacent channels. This directionality may require a very complicated antenna system with several towers if the requirements are stringent.

The antenna is the last point in the system under the control of the broadcaster. The signals radiated from the antenna are propagated through space to each receiving antenna. The factors affecting the strength of the received signal include the strength of the signal radiated by the broadcasting station in a particular direction, the distance to the receiving site, losses incurred by the less-than-perfect conductivity of the ground along the propagation path, terrain obstructions (large hills cast shadows even at AM frequencies), and in the case of skywave transmission, the ionospheric conditions that determine how much of the radiated signal will be reflected back to each distant receiving location. Signal strength in a particular direction can also be affected by the presence of structures such as buildings or towers near the radiation system.

The polarization of the transmitted waves is also a factor; for standard broadcast stations vertical polarization is used because of its superior groundwave propagation and the simplicity of antenna design. The FCC has established maximum transmitter power limits for each of the three classes of AM channels (clear, regional, and local) so the only variables available to the design engineer attempting to maximize the coverage of a radio station involve the antenna location, the pattern design, and a limited choice of power levels. These factors go hand in hand when designing a directional antenna system. Severe constraints are usually imposed on transmitter site selection because of aeronautical, zoning, environmental, and coverage requirements. The constraints encountered in the pattern design relate to the size and shape of the transmitter site; the extent to which the necessary signal suppression can be achieved at the desired transmitter power level; and the cost of design, construction, adjustment, and maintenance of multitower systems. The pattern design can also seriously affect the stability, efficiency, and bandwidth of the completed system. These factors will be discussed later.

RADIATION VERSUS FIELD STRENGTH

Two independent factors determine the signal strength at any given point within a station's service area. First is the strength of the signal radiated in that direction; and second is the path attenuation between the transmitting and receiving antennas. Attenuation is determined by both distance and ground conductivity. It is customary to express the radiation in units of millivolts-per-meter at one kilometer (km) unattenuated. This is the field that would exist at one kilometer over perfectly conducting earth. In this case the field strength would be inversely proportional to the distance from the transmitting antenna; hence, the radiation is also described as the "inverse distance field." The unattenuated radiation cannot be measured directly but can be inferred with great accuracy if sufficient field strength measurements are made to determine the ground conductivity. Field strength measurements are always dependent on radiation, distance, and ground conductivity.

THE SINGLE TOWER NONDIRECTIONAL ANTENNA

Current and Voltage Distribution

The majority of single-tower antennas are neither top-loaded nor sectionalized, and most of them are insulated from ground. For such simple towers, the current is a maximum 90 electrical degrees down from the top (or at the base if the tower is shorter than 90 degrees in height). A typical guyed tower that is 90 degrees high physically is about 95 degrees high electrically, because the velocity of propagation is less in the tower than in the air and is a function of the tower cross-section, slowing down as the cross-section is increased. The approximate shape of the current distribution on a thin tower of uniform cross-section is given by

$$i_a = I_a \sin(G - y)$$

where: i_a = Current in amperes at height y
I_a = The maximum current in amperes
G = The tower height in degrees
y = The height in degrees of the current element i_a

As an example, the general shape of the current and voltage distribution on a thin tower 210 electrical degrees high is shown in Fig. 1. For shorter towers, the distribution would approximate that shown, but with the lower portions cut off; there always being a current node and a voltage maximum at the top of any such tower that does not employ top loading. It is important to visualize the shape of the voltage distribution along the tower because of the need of good insulators at the high-voltage points. Otherwise corona or arc-overs may result to disrupt broadcasting service.

The tower current and voltage are not zero at the nodes shown along the tower. Rather, they reach minimum values and shift rapidly, approximately 180 degrees in-phase in traversing the node region. When towers considerably taller than 180 degrees in height are considered, the current near the base is in the opposite direction from that in the upper portion of the tower. Under these conditions, when viewed in the horizontal plane, the radiation from the lowest part of the tower is canceling a portion of the radiation from the part above the current minimum. Any increase in tower height above the optimum would actually reduce horizontal plane radiation.

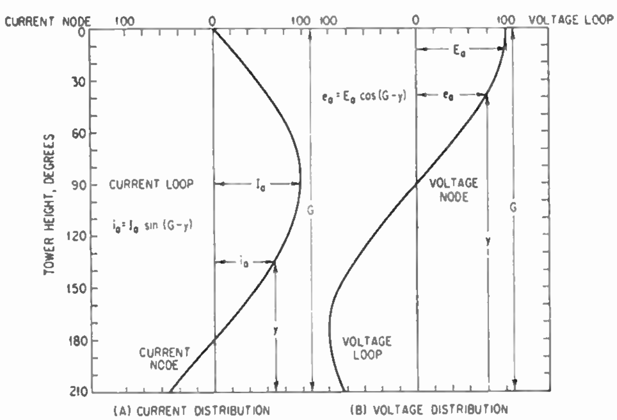


Figure 1. Theoretical current and voltage distribution on a vertical radiator.

VERTICAL RADIATION CHARACTERISTICS

Maximum groundwave radiation for a given transmitter input power occurs when a tower is 225 electrical degrees high (five-eighths wavelength). The variations in tower current distribution with increasing tower height defines the shape of the radiation characteristic in the vertical plane. Fig. 2 shows the size and shape of the vertical plane radiation patterns for a single tower of various heights atop a perfect ground system, fed with one kilowatt of power.

Insulated Tower Base Impedance

The base impedance of a single nondirectional tower is determined principally by its electrical height, its cross-section, the extent of the ground system, and the elevation of the feed point above ground. For typical guyed towers of uniform cross-section, which are base insulated and fed four or five feet above ground level, the resistive and reactive components of the base impedance approximate the values shown in Fig. 3. The base impedance of self-supporting towers departs radically from the values shown, not only because of their large and tapering cross-section, but also because of the capacitance of the base insulators necessary to support each leg of the tower.

Electrically short towers are inefficient radiators, not only because of the shape of their vertical radiation characteristics as shown in Fig. 2, but also because of

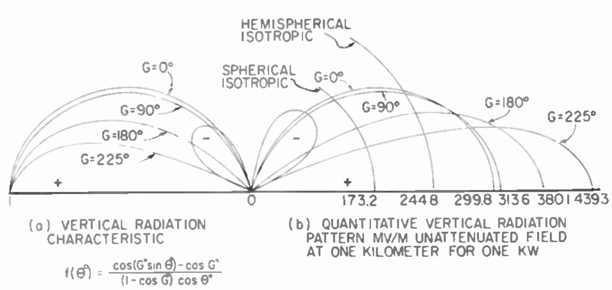


Figure 2. Radiation characteristics in vertical plane.

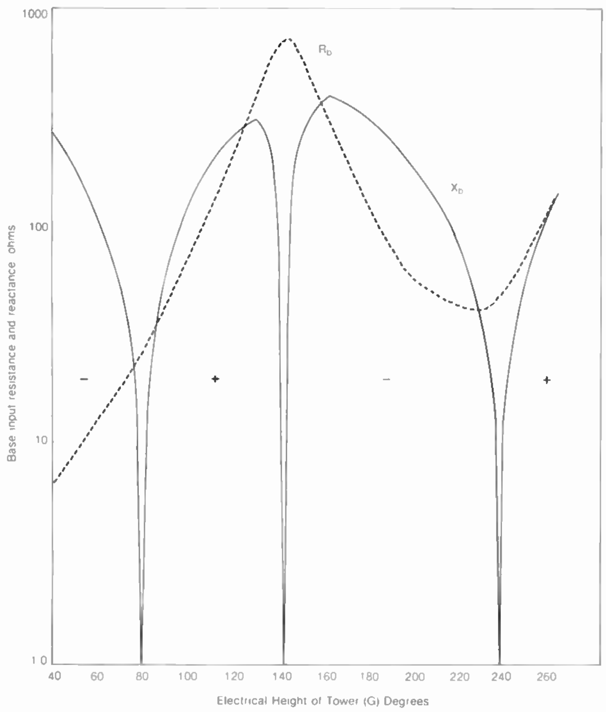


Figure 3. Typical base input resistance and reactance of a uniform cross-section base insulated guyed tower.

proportionately higher ground losses. For example, a tower 48 degrees high with a base resistance of only 9 ohms will have approximately 10% of the available power wasted in ground system resistance losses (typically one ohm).

Grounded Towers, Shunt Fed and Folded Monopoles

Occasionally towers without insulated bases must be used as AM radiators. Such structures include FM or TV towers, water tanks, and ornamental flag poles. Although the impedance at the base of such a tower is necessarily essentially zero, the impedance rises with increasing height of the feed point. It is a simple matter to determine experimentally the height at which a shunt fed tower must be driven to provide a desirable input impedance. A common technique is a *slant-wire* feed in which a wire is attached to the tower at a selected height above ground, and brought down to near ground level at an angle approximating 45 degrees, to serve as the antenna input terminal. A slant-wire feed distorts the otherwise omnidirectional pattern of a single tower and tends to suppress radiation over the sector on the side where the slant-wire is attached. This effect can be avoided if, instead of the slant-wire, the feed conductors are insulated at the base, brought up outside of the tower, and bonded to the tower; for example, 90 degrees above ground to form a folded monopole. The conductors, in this concentric arrangement, in effect form the outer conductor of a coaxial transmission line with a short to the tower at the 90-degree point and an open at the base insulators. This quarter-wave open circuit transmission line in effect puts an insulator at the tower base. The current up on the outer conductors and down on the tower essentially cancel so far as radiation is concerned. The tower with this insulated skirt performs like a base insulated tower. The concentric arrangement of conductors, usually six, are tied together above the conductor base insulators and fed like a base insulated tower. There is a small amount of power loss in the 90-degree concentric transmission line shorted at the top and used to produce the open circuit at the bottom. The radiation current is up on the outer conductors, to where they are connected to the tower at the 90-degree point, and then on up to the top of the tower where the current is zero, as on a base fed insulated tower.

Folded Conical Monopole

The folded conical monopole shown in Fig. 4 broadbands the input impedance. For a 90-degree tower, the vertical radiation characteristic is that of a 75-degree tower, but with an increased value of input resistance and reactance of about 50%. These results were obtained by applying the method of moments to compute the base impedance, current distribution, and vertical pattern.

Top Loading

The performance of an electrically short tower (significantly less than 90 degrees) can be improved, both as to radiation efficiency and bandwidth, by means of top loading. This consists of increasing the capacitance

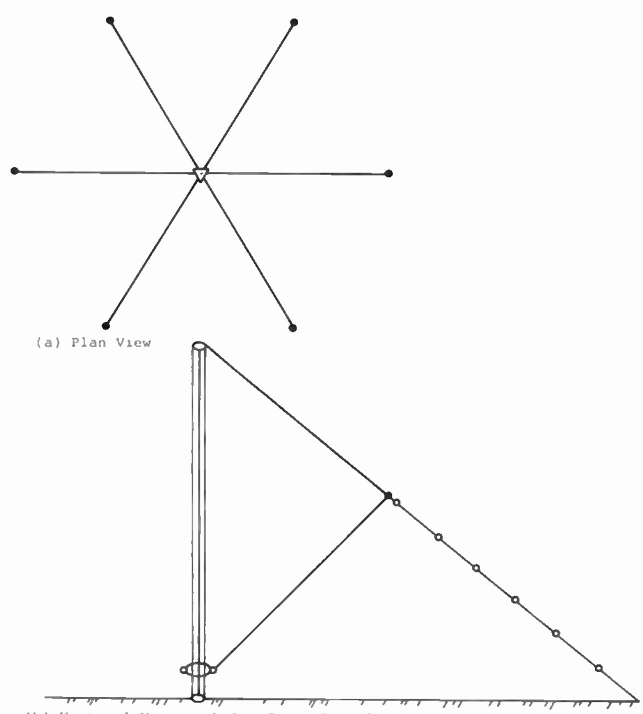


Figure 4. Sketch of folded conical monopole antenna.

to ground from the top of the tower. This loading can take the form of either a flat, more-or-less circular horizontal disk attached to the top of the tower (called a top hat), or as sections of guy wires bonded to the top of the tower and extending down a useful distance before encountering the first of the guy wire insulators. Many variations are possible. Some installations use 3, 6, or even 12 nonstructural guys for top loading that are very effective. By interconnecting the lower ends of the top loading cables, the capacitive loading is increased some, but both this method and spider web connections between the top loading cables increases the construction and maintenance problems. These problems can be eliminated by simply increasing the top loading cables a small amount to give the same increase in capacity effect. Top loading is electrically less desirable than increased tower height, but is useful where towers must be electrically short due to either extremely low carrier frequencies or to aeronautical limitations. Top loading increases the base resistance and lowers the capacitive base reactance, thus reducing the Q and improving the bandwidth on towers less than 90 degrees high. When the tower height is of the order of 130 degrees, top loading can be used to increase the tower's electrical height to give maximum groundwave radiation and minimum skywave radiation.

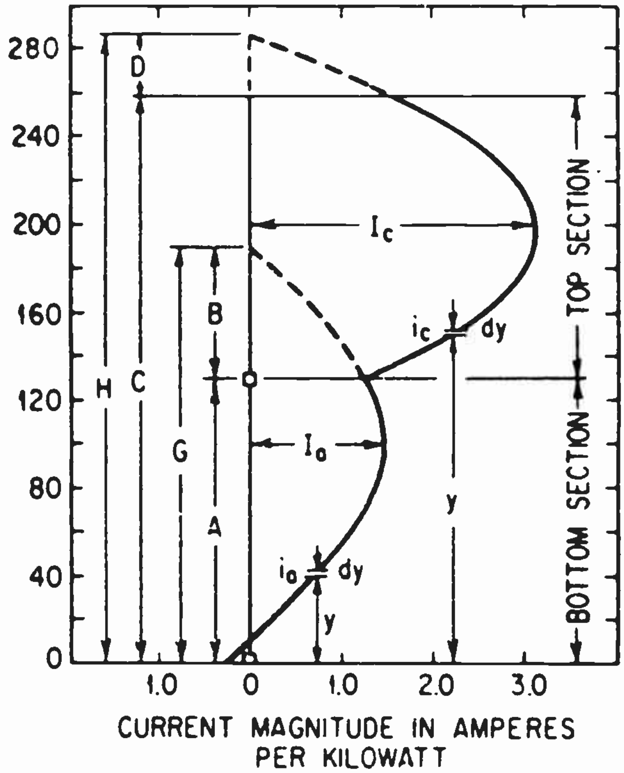


Figure 5. Theoretical current distribution on top-loaded sectionalized tower.

Sectionalized Towers

A utopian vertical radiator would have a constant current throughout its height, but in real life the current must ultimately reduce to zero at the tower top or at the end of the top loading cables. The current can be made to diminish less rapidly by inserting an inductance in series with the tower at a point part way up its height. This is the same technique as using the familiar "loading coil" near the center of the vertical whips often used in mobile radio systems.

Top-Loaded Sectionalized Tower

For a simple vertical radiator, the radiation characteristic can be improved either by increasing the tower height up to 225 degrees for maximum groundwave or by top loading. This in effect raises the position of the current loop with respect to the ground. This principle can also be applied to the top section of a sectionalized tower.

The purpose of top loading a sectionalized tower is to provide a means of further controlling the current distribution on the lower section only. Considering efficiency and stability, it is usually possible to achieve a more favorable radiation characteristic of the whole tower by employing top loading and sectionalization. See Fig. 5. In the case of tall towers used to support FM or TV antennas, it may not be practical to employ top loading.

Depending on the height of tower in wavelengths, the tower can be sectionalized at one or more points to accomplish the highest efficiency consistent with good operating stability. (See Fig. 6.)

Short Low-Loss Antenna with Insulated Counterpoise

Short low-loss antennas are useful for standby use and regular use where height is limited. By using optimum top loading and a tuned counterpoise, the field strength is maximized by adjusting the counterpoise inductor in Fig. 7 to minimize the top hat field through the counterpoise to the lossy ground.

Ground Systems

The current on a tower does not simply "disappear", rather it returns to earth through the capacitance between the earth and each incremental element of the tower or the top loading. For towers not exceeding 90 degrees in height, the tower current is greatest at the base. For such towers the radial ground current is greatest near the tower and decreases with increasing distance from the tower. For single towers the ground currents are radial from the tower base. The ground losses are greatly reduced if the tower has a radial copper ground system, so the ground current will be in the low-loss copper ground system rather than in the earth which has a much higher resistance. A solid copper sheet of infinite radius would be the ultimate ground system, but experiments and experience have defined the dimensions of an adequate ground system. A system of 120 radial ground wires, each 90 degrees

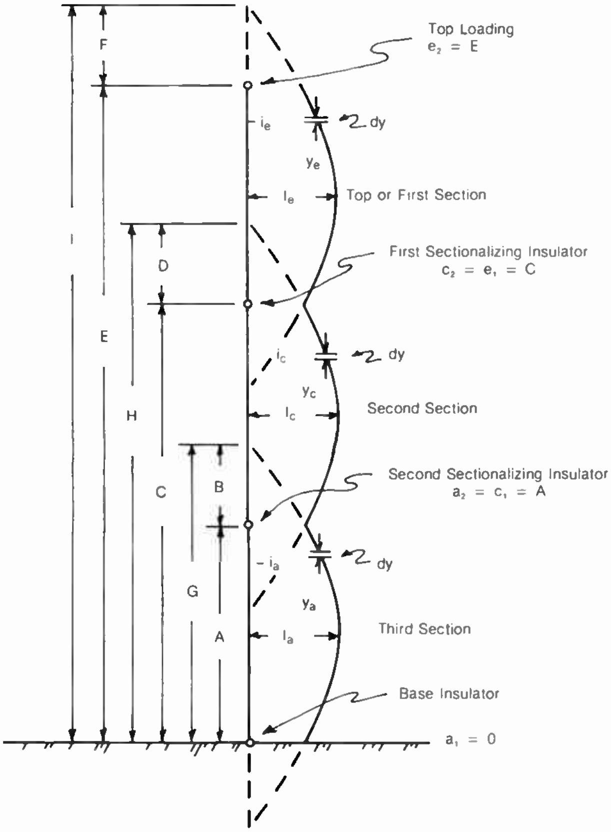


Figure 6. Theoretical current distribution on three section top loaded tower.

long (140 degrees is considered optimum), and equally spaced out from the tower base, constitutes a "standard" ground system. This is often augmented with an additional 120 interspersed radials 50 feet long, or an expanded copper mesh ground screen 25 to 50 feet square centered at the tower. A superior ground screen material is the copperweld, mesh ground mat often utilized by power companies for lightning protection under electrical substations.

Where the antenna site is too small to accommodate all the ground radials at full length, a compromise often used, if easement can not be obtained beyond the property line, is to increase the number of radials by

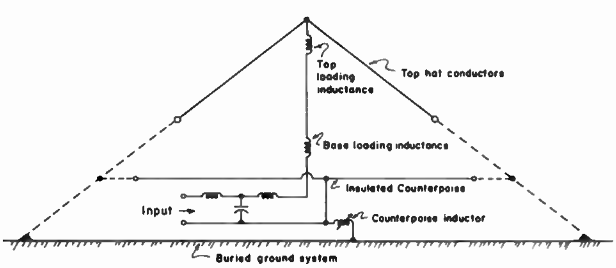


Figure 7. Short, low-loss antenna with counterpoise.

placing them 1 or 2 degrees apart rather than the standard 3 degree separation.

There is no magic in a standard ground system for nondirectional towers; it simply represents a reasonable balance between cost and radiation efficiency.

The antenna system loss including the tower and ground system is normally assumed to be one ohm and is added to the tower base resistance.

Most ground systems under directional antenna arrays consist of the usual 120 radials per tower truncated and bonded to traverse copper straps where the radials from the several towers would otherwise intersect. Stability considerations may dictate larger than standards ground systems under critical directional antenna arrays; changes in soil conditions beyond the ground system can result in small changes in tower base impedance.

Ground system losses are minimized if the radial wires are placed above ground, thus the E-field voltage from the tower and top loading cables terminate on these radial conductors so the H-field current can return to the tower base without penetrating the lossy earth. Ground radials are usually buried 6 to 8 inches for mechanical protection. Burial up to 24 inches is feasible where necessary to permit deep plowing for agricultural crops. However, the ground system should be very near the earth surface in the immediate vicinity of the tower. The earth losses are greater for the buried ground system. Ground systems laid on the earth surface, or tundra in the far north, has the highest loss and least stability of the base impedance. Changes in weather conditions change the dielectric constant and conductivity of any unshielded earth to the detriment of base current stability.

TWO-TOWER DIRECTIONAL ANTENNA

Radiation Pattern Shape

When a nondirectional antenna, with a given power, does not radiate enough field strength to serve the community of interest and/or fails to protect other radio stations, then it is logical to resort to a directional antenna system to achieve these objectives. FCC Rules spell out the protection requirements to be provided to the various classes of stations, both daytime and nighttime on the same and adjacent channels. These limits, which must be met in the directional antenna design, tend to define the shape and size of the most desirable antenna pattern. Since the distances and directions to the other stations requiring protection are rarely the same, most directional antenna patterns are tailored to meet the specific requirements. A directional antenna functions by carefully controlling the amplitude and phase of the radio frequency currents fed to each tower. The resulting field in any direction is the vector sum of the individual tower radiation components. To visualize the resulting pattern in the horizontal plane, one must consider the individual tower radiation components when viewed from distant points in different directions. The relative amplitudes from the individual towers remain unchanged, but the relative phases shift with azimuth because the signal from the closest tower arrives first. In a directional antenna system, one tower is usually defined as the reference tower, and the amplitude and phase of each other tower is measured relative to this reference. The reference tower usually has the greatest current, thus the ratio of the current in each other tower relative to the reference tower current is a fractional number often expressed as a percent of the reference tower current. The relative amplitude and phase of the tower currents is measured by means of an antenna monitor.

The phase of the field, radiated by each tower relative to the reference tower, has two components when viewed from any distant point of observation. The relative electrical magnitude of the current fed to the tower is one component, and is adjustable. The second component is the phase which appears to lead or lag the reference tower by virtue of being more distant or closer than the reference tower to the point of observation. This is termed the space phase component and varies continuously for each tower in a sinusoidal manner as the observation point is moved in azimuth along a distant circle around the array.

Fig. 8 shows three simple directional antennas and their resulting patterns which are easy to visualize. Fig. 8A shows two towers arranged along a north-south line separated by 180 degrees and fed with equal currents in phase. When viewed from the east or west, the fields from the two towers are in phase and the maximum field strength results. When viewed from the north or south, the field from the more distant tower is delayed by the 180 degrees of additional distance, thus canceling the field of the closer tower so as to result in a minimum or null. The deepest minimum or null occurs only when the fields are exactly equal in amplitude and opposite in phase.

Fig. 8A is termed a broadside array because the maximum radiation is broadside to a line through the towers. Fig. 8B shows a similar arrangement, but with the phase of the current in the north tower shifted by 180 degrees. The fields from the two towers cancel each other when viewed from the east or west, but would produce maximum radiation from the north or south. This would be termed an end-fire array, because maximum radiation coincides with a line through the ends of the array. Fig. 8C alters the spacing to 90 degrees and phasing to 90 degrees so as to produce a cardioid pattern. Other combinations of tower spacing and phasing can produce a great variety of pattern shapes. (See Appendix B.)

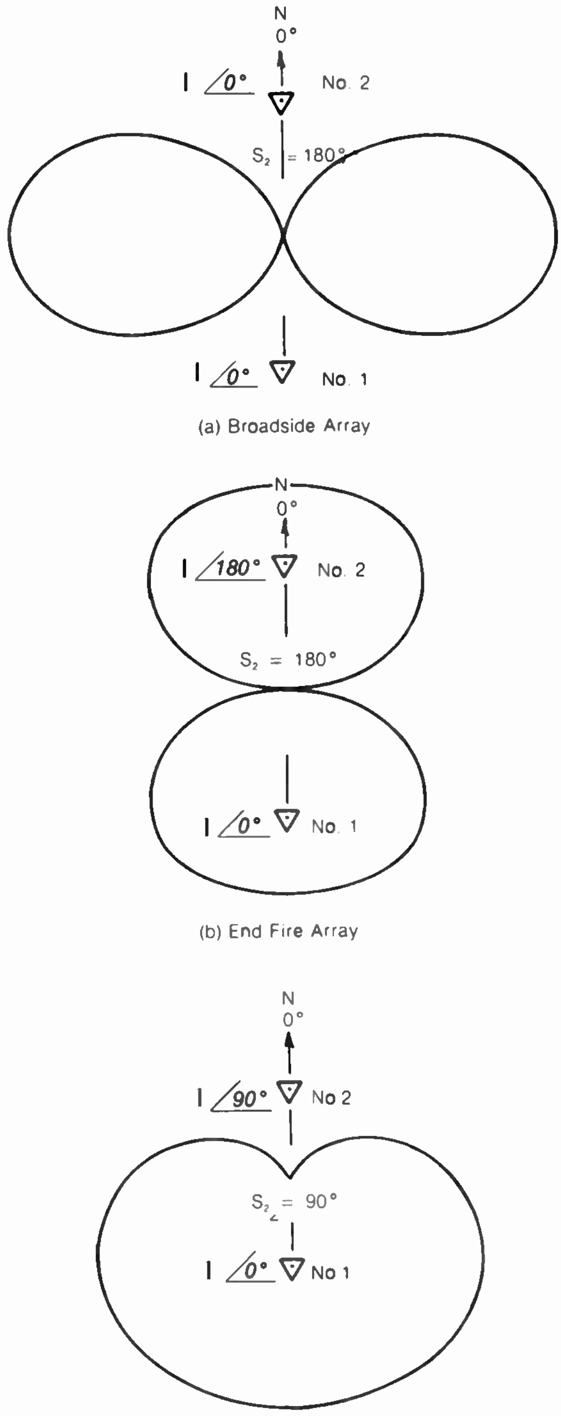


Figure 8. Three simple directional antenna patterns.

Multiplication of Two-Tower Patterns

Perhaps the most widely used method of controlling pattern shape involves the multiplication of two-tower patterns. This is illustrated in Fig. 9. When a two-tower pattern such as pattern No. 1 with nulls at $\pm\theta_{n1}$ is multiplied by pattern No. 2 with nulls at $\pm\theta_{n2}$ the result is pattern No. 3 in a three-tower array. The directions of all of the two-tower array nulls are maintained in the three-tower array. This is a very powerful design technique for protecting other stations and still serving a desired service area. In this special case, the spacings S_2 and S_3 are equal, resulting in an in-line array with fields of towers No. 2 and No. 3 being added in the center tower, and the end tower of

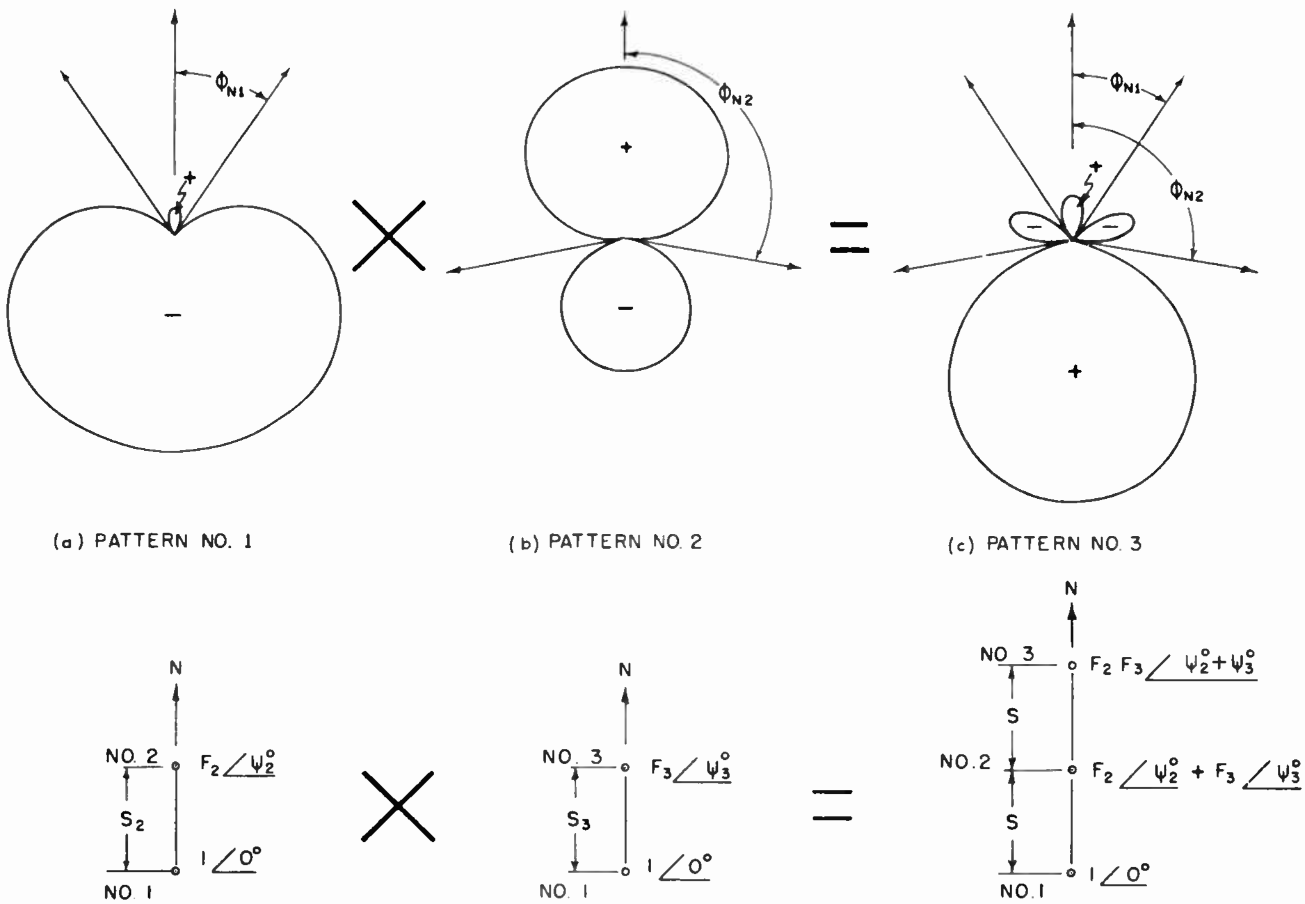


Figure 9. Multiplications of patterns to produce a three tower in line array.

the three-tower array is the multiplication of these fields as shown in pattern No. 3 of Fig. 9.

In the event that the protection directions are not symmetrically located, the two-tower arrays can be placed on different azimuth angles (as shown in Fig. 10) to produce a four-tower parallelogram array. The nulls of the No. 1 pattern are maintained and the nulls of No. 2 pattern are maintained in the four-tower parallelogram array. Furthermore, the spacing from No. 1 tower to No. 2 and No. 3 towers does not have to be the same. By this approach of using one or more parallelograms, a wide variety of asymmetrical patterns are possible with relative simplicity of pattern calculations. However, modern computer techniques can optimize individual tower locations, currents and phases so as to produce an efficient pattern, frequently using fewer towers than required with the parallelogram approach.

Systematization of Patterns

The pattern possibilities resulting from variations in spacing and phasing have been systematized and a sample of two-tower patterns is shown in Appendix B. See Ref. 1 for three-tower patterns.

Radiation Pattern Size

The pattern size is usually determined by integrating the energy flow outward through an imaginary hemispherical surface surrounding the directional antenna array. This method does not give information regarding the distribution of power radiated from the various towers of the directional antenna array, however, it is very useful for making comparisons of pattern size. This computation method is available in digital computer programs and is used by the FCC.

There are other methods of determining pattern size; such as the *mutual resistance method* which employs Bessel functions, and the *driving point impedance method* which uses mesh circuit equations with self and mutual impedance information.

The "method of moments" is now available in large computer programs to determine current distribution on towers and top loading cables, base driving point impedances, and the vertical pattern of directional antenna arrays.

Driving Point Impedance

The input impedance of each tower in an array, called the driving point impedance, is not that of the

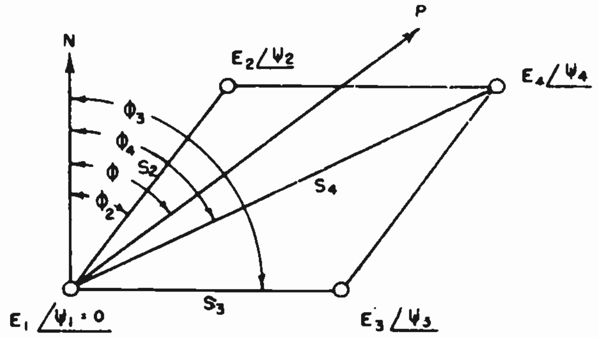


Figure 10. Multiplication of two patterns to produce a four tower parallelogram array.

nondirectional tower. The driving point impedance contains the self impedance plus the mutual impedance multiplied by the current ratios that exist in the array as driven to produce the desired pattern. The driving point impedance will modify the self impedance, depending on the array parameters, and can even make the base resistance negative so that the tower draws power from the other towers and dissipates the power into a load resistor or delivers it back to the phasing system. Because the driving point impedance is affected by the currents in the other towers, it can only be measured by an operating bridge inserted in the tower feed point, while the other towers are operating with their correct current magnitude and phase.

Base Currents Versus Radiated Fields

In a directional array, the tower base current ratios will usually depart substantially from the calculated radiated field ratios when the pattern is correctly adjusted. This is caused by the mutual coupling between towers which distorts the sinusoidal current distribution otherwise assumed for each tower. Thus, the correct pattern is initially proved by means of a series of field strength measurements in significant radial directions from the station rather than by assuming that measurement of tower currents and phases can establish the correct pattern.

Near-Field Versus Far-Field Conditions

Theoretically a directional antenna pattern is not fully formed except at an infinite distance, where the separate towers can be considered as point sources. As a practical matter, near-field effects can persist as far as 32 km (20 miles) from an antenna before far-field conditions prevail. This is especially true in the deep minimums of wide-spaced arrays; however, misleading measurement results can often occur under apparently innocent circumstances. Near-field calculations involve consideration of the actual inverse distance attenuation and the actual phase delay from each antenna element to a series of observation points along a radial.

Fig. 11 shows the results of such calculations on a minimum radial and the resulting analysis of field strength measurements. Line A is the inverse distance line for the theoretical unattenuated radiation at one kilometer. Line B is the result of the near-field calculations assuming only inverse distance attenuation, that is, no soil losses. It converges with the inverse distance line with increasing distance. Line C represents a soil conductivity of 10 mmhos/m as drawn in the conventional manner from analysis of nondirectional measurements on the radial. Line D is a composite of lines B and C. It includes the near-field calculations and is attenuated with distance in accordance with the soil conductivity previously established. This composite line converges with the near-field calculations at short distances where soil attenuation is negligible and converges with the soil conductivity line at great distances where near-field effects disappear. Since curve D accounts for both near-field effects and soil losses, it is the proper curve against which the directional field strength measurement data should be fitted. Note the good fit to the measurement data, both close to the array and at distant points even though the first 19 measurement points fall considerably above the inverse distance line A.

Pattern Size Versus Pattern Shape

The shape of a directional antenna pattern is determined by the adjustment of the phase and ratio parameters, whereas the pattern size is a measure of the power radiated and is affected by the transmitter power output and the losses within the phasing system and the ground system. Since pattern size and shape are essentially independent, it is most expeditious to adjust an array to get the correct shape before expending much concern on the size.

Field strength measurements on a previously-licensed directional antenna may appear to indicate a change in pattern shape or size when the change was in fact due to changes in soil conductivity. Such changes affect distant measurements more than close-in measurements. In some areas of the United States, the conductivity is typically higher during winter and spring months, when the soil is more moist, than in summer and fall months.

Seasonal conductivity variations are not observable in some portions of the country, yet are extreme in other areas. One well-documented case showed a seasonal doubling of signal strength at 32 km (20 miles) in the main lobe of a correctly adjusted system operating on 1380 kHz. To avoid the misleading effects of seasonal conductivity changes that might appear to distort measured directional antenna patterns in size or shape, the FCC requires that all the field strength measurements in a directional antenna proof of performance be made under "similar environmental conditions."

Standard Patterns

Theoretical (also called calculated) patterns can have nulls wherein the radiation at specific azimuths goes completely to zero. In practice, it is not possible to prove by field strength measurements that a null exists.

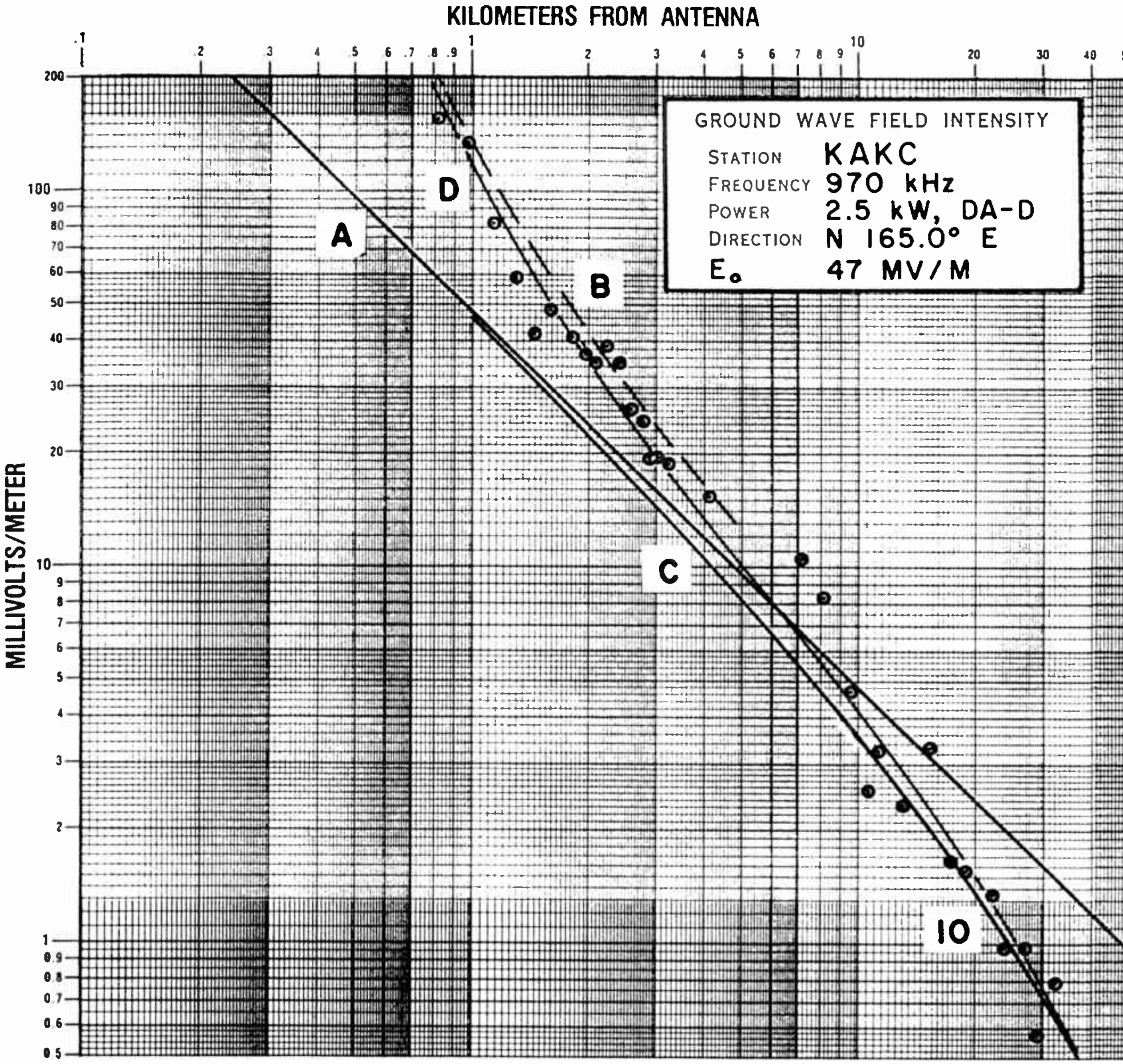


Figure 11. Near-field effects.

Reradiation and scatter from objects external to the array limit the depth to which a pattern minimum can be proven. Additionally, operational variations in phase and ratio parameters will increase radiation in any direction where the deepest possible minimum has been previously established. To accommodate these limitations, the FCC authorizes a "standard" pattern for each directional antenna station. Standard patterns exceed the theoretical pattern at all azimuths by specified and easily calculated amounts. It is required that the radiation from a directional station not exceed its standard pattern. All U.S. stations employing directional antennas have FCC specified standard patterns. These supersede all earlier patterns based on theoretical calculations or on field strength measurements. The standard pattern radiation values are now used exclusively in all calculations of coverage and interference.

Augmented Patterns

Augmentation is applied to the standard pattern when the measured field strength is exceeded in discrete directions but does not cause interference to other stations. When augmentation is desired, it is achieved by applying Eq. (C-9) in Appendix C.

REFERENCES

The first twelve references listed are published by Carl E. Smith Electronics, 8200 Snowville Road, Cleveland, OH 44141, 216/526-4386.

1. "Directional Antenna Patterns"
2. "Theory and Design of Directional Antennas"
3. "Standard Broadcast Antenna Systems"
4. "Design and Operation of Directional Antennas"
5. "Directional Antenna Pattern Shapes"
6. "Radiation Characteristics of Transmitting Antennae (An Introduction to Directional Antenna Pattern Design)"
7. "Directional Antenna Design Example"
8. "Parasitic Reradiation"
9. "Introduction to Directional Antenna Systems"
10. "Instructions for Installation of Radio Broadcast Stations Ground Systems"
11. "Log Periodic Antenna Design Handbook"
12. "Radio Broadcast Ground Systems"
13. Stuart Ballantine, "On the Optimum Transmitting Wavelength for a Vertical Antenna Over Perfect Earth", Proc. I.R.E. Vol. 12, pp. 833–839; December, 1924.
14. Carl E. Smith and Earl M. Johnson, "Performance of Short Antennas", Proc. I.R.E., Vol. 35, pp. 1026–1038; October 1947.
15. Carl E. Smith, John R. Hall and James O. Weldon, "Very High-Power Long-Wave Broadcasting Antennas", Proc. I.R.E., Vol. 42, No. 8, pp. 1222–1235; August, 1954.
16. Ralph N. Harmon, "Some Comments on Broadcast Antennas", Proc. I.R.E., Vol. 24, pp. 36–47; January, 1936.
17. George H. Brown, "A Critical Study of the Characteristics of Broadcast Antennas as Affected by Current Distribution", Proc. I.R.E., Vol. 24, pp. 48–81; January, 1936.
18. Carl E. Smith, "A Critical Study of Several Antennas Designed to Increase the Primary Coverage of a Radio Broadcasting Transmitter", Professional Thesis at Ohio State University, Columbus, Ohio.
19. Carl E. Smith, "A Critical Study of Two Broadcast Antennas", Proc. I.R.E., Vol. 24, pp. 1329–1341; October, 1936.
20. Morrison, J. F., and P. E. Smith, "The Shunt Excited Tower", Proc. IRE, vol. 25, pp. 673–696; June, 1937.
21. Jeffers, C. L., "An Antenna for Controlling the Nonfading Range of Broadcasting Stations", Proc. IRE, Vol. 36, pp. 1426–1431; November, 1948.
22. Smith, Carl E., D. B. Hutton, and W. G. Hutton, "Performance of Sectionalized Broadcasting Towers", IRE Trans., pp. 22–34; December, 1955.
23. Herbert P. Jacobson, Jr., P.E., Carl E. Smith, P.E. and R. Richard Riggs, "High Power Steerable Short-Wave Antennas", IEEE Transactions on Broadcasting, Vol. 34, No. 2, pp. 186–192, June 1988.
24. Smith, Carl E., "Short Low Loss Antennas", IEEE Transactions on Broadcasting, Vol. 35, No. 2, pp. 237–249, June, 1989.

APPENDIX A (CHAPTER 2.5, PART I)

DIRECTIONAL ANTENNAS FOR PATTERN SHAPE

Space Configuration

The plan configuration of the k^{th} tower in an array is shown in Fig. A-1. A space view of the k^{th} tower and observation point *P* is shown in Fig. A-2.

Vector Diagram

The field strength at the point *P* in space for the k^{th} tower is shown in Fig. A-3. The space phasing in the horizontal plane is shown in Fig. A-4 and in the elevation plane the space phasing is reduced further as shown in Fig. A-5.

Generalized Equation

The vector equation to express the vectors in Fig. A-6 is the generalized equation that can be used to express the pattern shape for a directional antenna array of n towers. The equation in condensed form is,

$$E = \sum_{k=1}^{k=n} E_k f_k(\Theta) \quad \beta_k \qquad \text{[A-1]}$$

where:

E = the total effective field strength vector at unit distance (P) for the antenna array with respect to the voltage vector reference axis. This vector makes the angle ß with respect to this axis as shown in Fig. A-6.

k = the kth tower in the directional antenna system

n = the total number of towers in the directional antenna array

E_k = the magnitude of the field strength at unit distance in the horizontal plane produced by the kth tower acting alone

$f_k(\Theta)$ = vertical radiation characteristic of the kth antenna as given in Eq. A-3

Θ = elevation angle of the observation point *P* measured up from the horizon in degrees

$$\beta_k = S_k \cos\Theta \cos(\phi_k - \phi) + \psi_k \qquad \text{[A-2]}$$

= phase relation of the field strength at the observation point *P* for the k^{th} tower taken with respect to the voltage vector reference axis.

$$S_k \cos(\phi_k - \phi) \cos\Theta$$

is the space phasing portion of β_k due to the location of the kth tower and ψ_k is the phasing portion of β_k.

S_k = electrical length of spacing of the kth tower in the horizontal plane from the space reference point

ϕ_k = true horizontal azimuth orientation of kth tower with respect to the space reference axis

ϕ = true horizontal azimuth angle of the direction to the observation point *P* (measured clockwise from true north)

ψ_k = time phasing portion of β_k due to the electrical phase angle of the voltage (or current) in the k^{th} tower taken with respect to the voltage vector reference axis

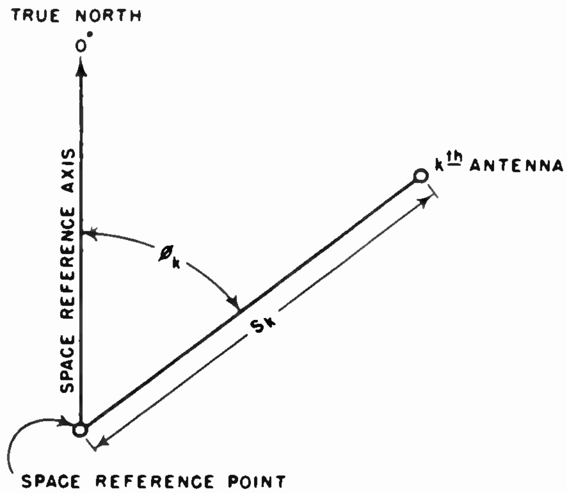


Figure A-1. Plan view of space configuration of k^{th} antenna.

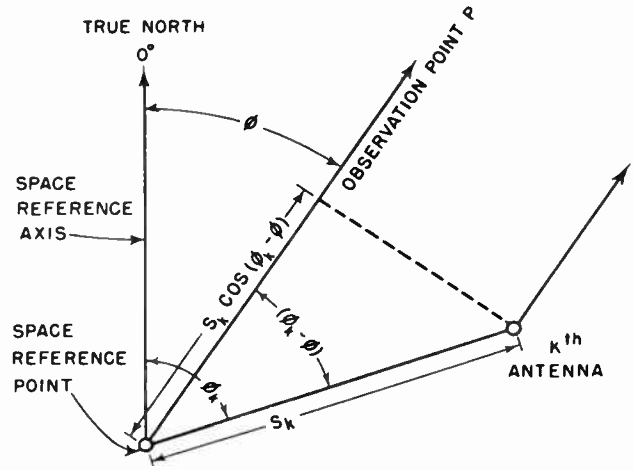


Figure A-4. Plan view of k^{th} antenna showing space phasing in the horizontal plane.

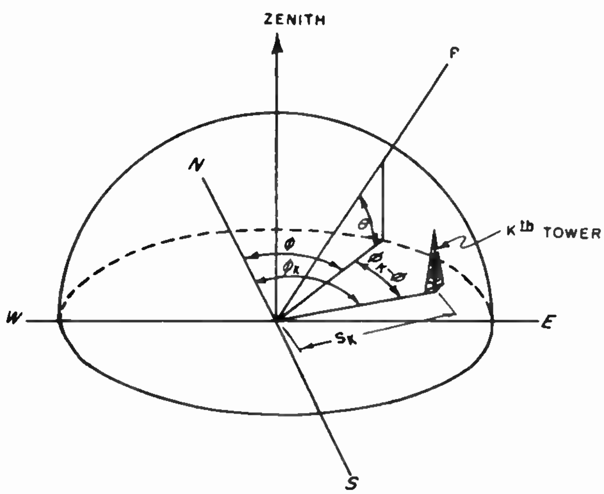


Figure A-2. Space view of observation point P and the k^{th} tower.

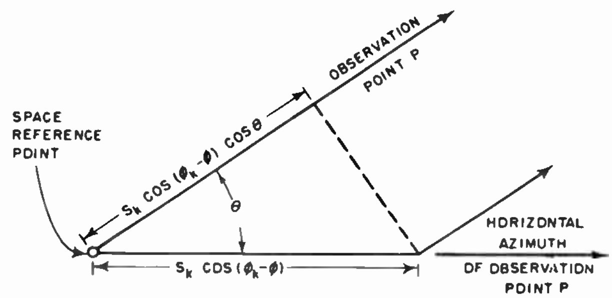


Figure A-5. Elevation angle θ shortens the spacing S_k to the value of $S_k \cos\theta$.

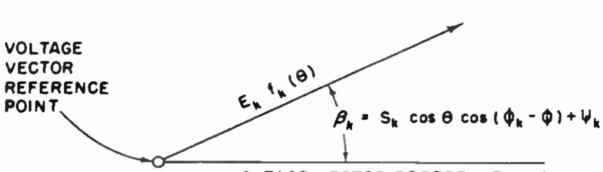


Figure A-3. Voltage vector diagram for the k^{th} antenna.

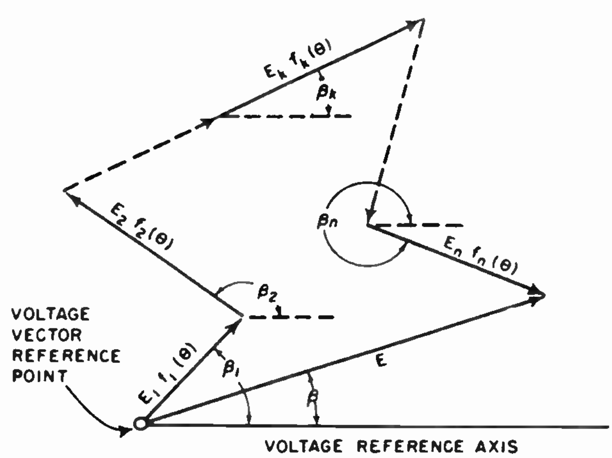


Figure A-6. Summation of field strength vectors for n antennas in the directional antenna array.

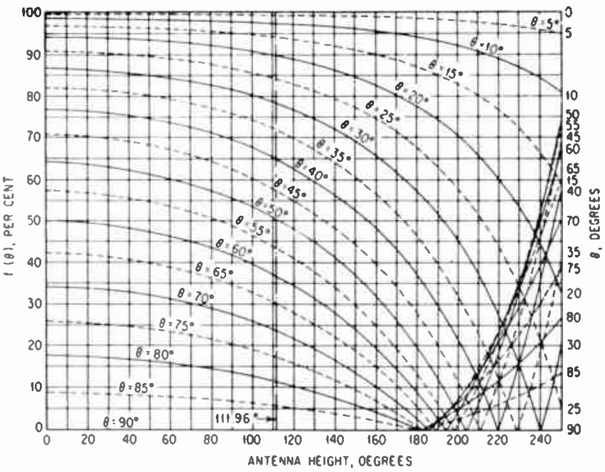


Figure A-7. Vertical-radiation characteristic as a function of electrical tower height for various values of elevation angle.

The shape of any directional antenna pattern can be computed by applying the above equations, however, many directional antenna arrays can be designed by simplified versions of this equation.

For a vertical antenna having a sinusoidal current distribution with a current node at the top, the vertical radiation characteristic takes on the form

$$f(\Theta) = \frac{\cos(G \sin \Theta) - \cos G}{(1 - \cos G)\cos \Theta} \qquad [A\text{-}3]$$

where:

$f(\Theta)$ = vertical radiation characteristic

G = electrical height of the antenna in electrical degrees

Θ = elevation angle of the observation point measured up from the horizon in degrees

The vertical radiation characteristics in Eq. A-3 are graphed in Fig. A-7.

For a top-loaded tower the formula is:

$$f(\Theta) = \frac{\cos B \cos(A \sin \Theta) - \cos G - \sin B \sin \Theta \sin(A \sin \Theta)}{\cos \Theta (\cos B - \cos G)}$$

This is the vertical radiation characteristic for a top-loaded tower of height A and top-loaded to a height of G = A + B.

For a two section top-loaded tower as shown in Fig. 5, the formula is:

$$f(\Theta) = \frac{\cos B \cos(A \sin \Theta) - \cos G + \dfrac{\sin B \cos(H - C)\cos(C \sin \Theta)}{\sin(H - A)} - \dfrac{\sin B \sin \Theta \sin (H - C \sin(C \sin \Theta))}{\sin(H - A)} - \dfrac{\sin B \cos (H - A)\cos(A - \Theta)}{\sin(H - A)}}{\cos\{\cos B - \cos G + [\sin B/\sin(H - A)](\cos \overline{H - C} \cos \overline{H - A})\}} \qquad [A\text{-}4]$$

This is the vertical radiation characteristic equation for a two-section sectionalized tower. The same procedure can be applied if more than two sections are involved. (See Fig. 6.)

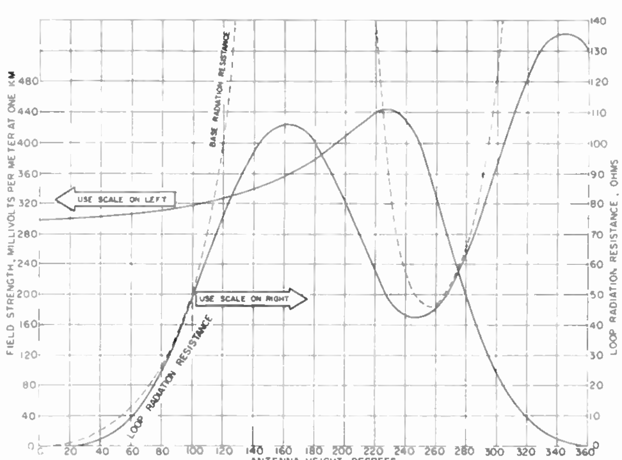


Figure A-8. Inverse field strength at 1 kilometer for 1 kw, loop and base radiation resistance as a function of tower height over a perfectly conducting earth.

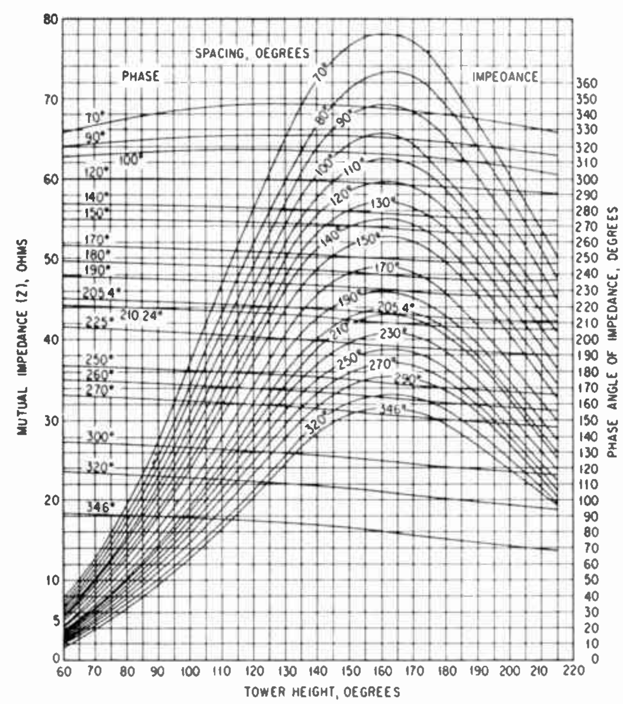


Figure A-9. Loop mutual impedance and phase angle between two towers of equal height.

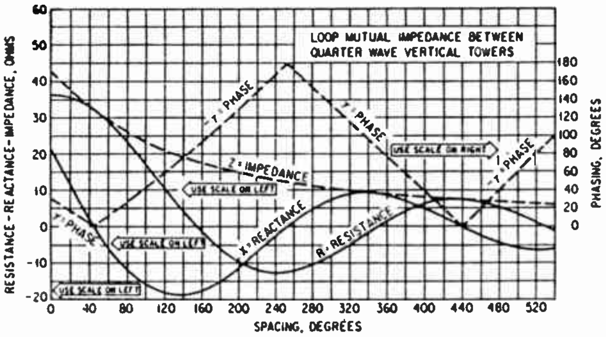


Figure A-10. Loop mutual impedance between quarter-wave vertical towers.

Theoretical Self-Loop and Base-Radiation Resistance

It is useful to know the theoretical loop and base resistance of a vertical radiator. This information is presented graphically in Fig. A-8 along with the theoretical inverse field strength at one kilometer.

Mutual Impedance Curves

The value of mutual impedance for most tower heights and spacing is given in Fig. A-9. The loop mutual impedance between quarter-wave towers is shown in Fig. A-10.

Horizontal RMS Field Strength

The field strength gain or loss of a two-tower array for various values of phasing and spacing is shown in Fig. A-11.

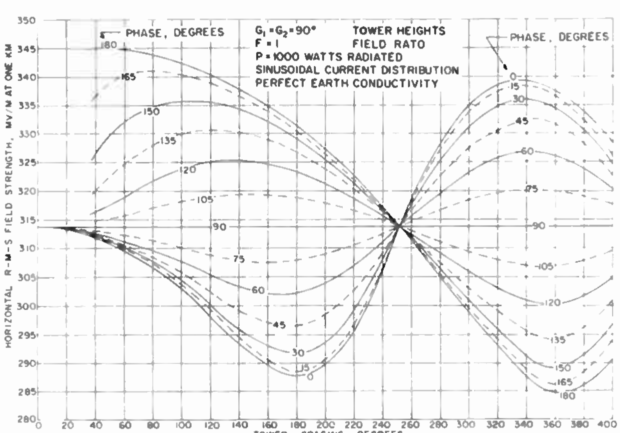


Figure A-11. Horizontal RMS field strength of two-tower directional antenna.

APPENDIX B (CHAPTER 2.5, PART I)

SYSTEMATIZATION OF TWO-TOWER PATTERNS

The pattern numbering system has been devised to furnish the antenna parameters in an orderly fashion as described at the bottom of each page. To simplify the numbering system, the spacing and phasing of each antenna is shifted in steps of 45 degrees. The spacing varies from 45 degrees on page B-1 to 1,440 degrees on page B-4 while the phasing varies from 0 degrees to 315 degrees on each page. The field strength of each antenna, E1 and E2, is expressed in percent of the maximum lobe field strength (50%) of the horizontal pattern. Finally, the last two digits specify, in percent, the RMS field strength in the horizontal plane of the two tower directional antenna array.

The detailed patterns which follow are for spacings in steps of 15 degrees up to 360 degrees. The phasings are only presented from 0 degrees to 180 degrees since the same patterns, turned 180 degrees, give the results for phasings from 180 degrees to 360 degrees.

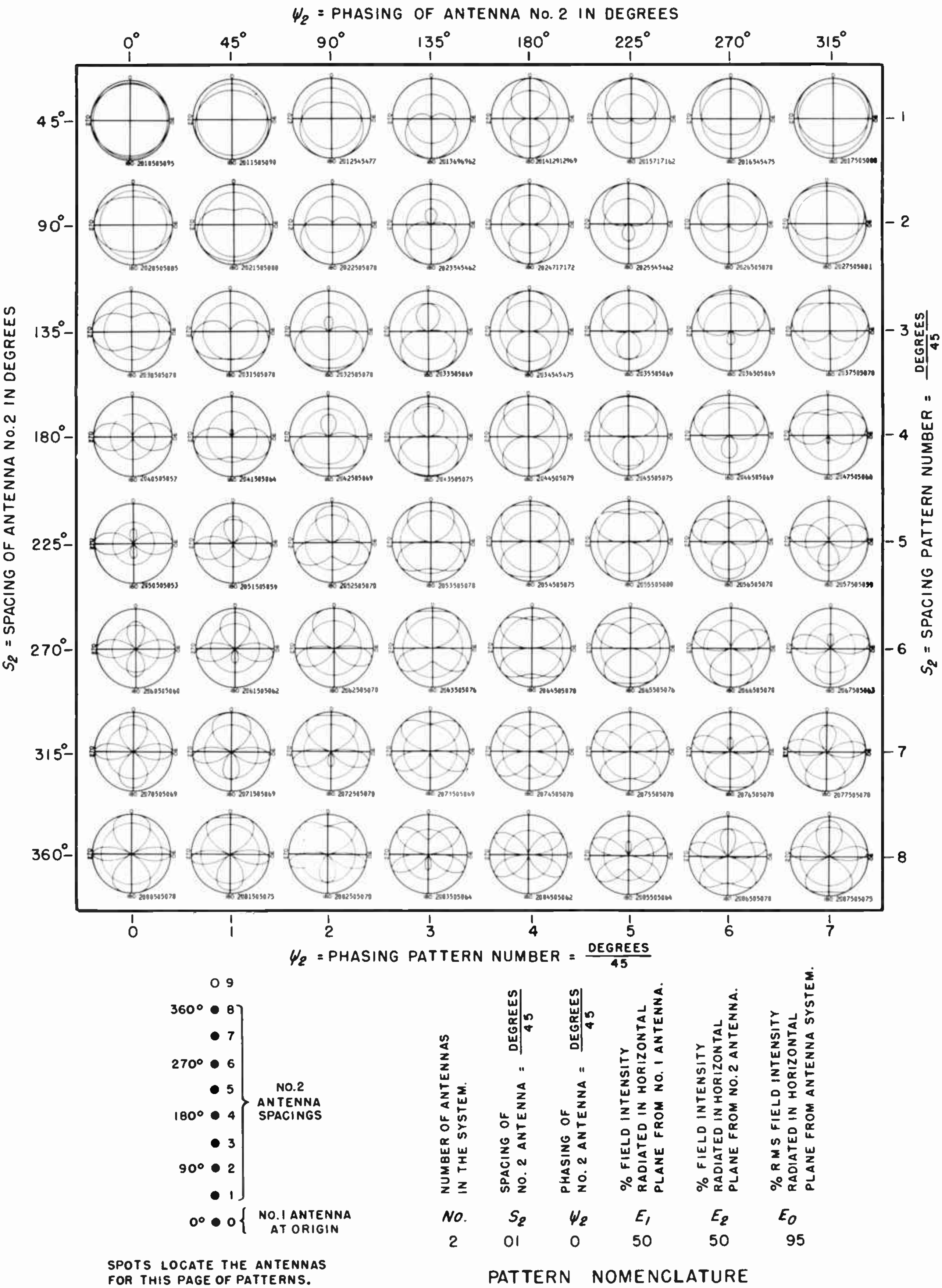
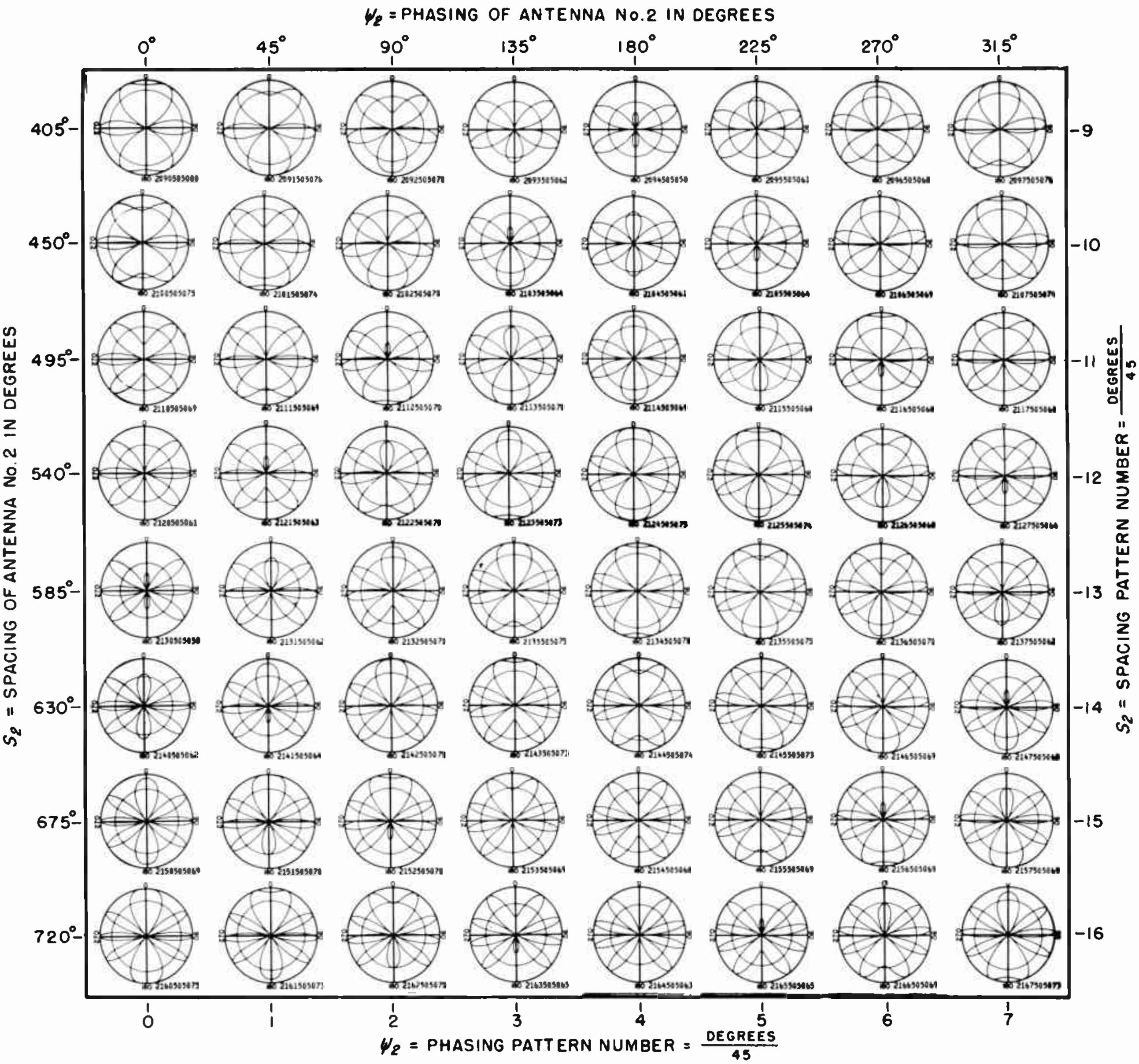
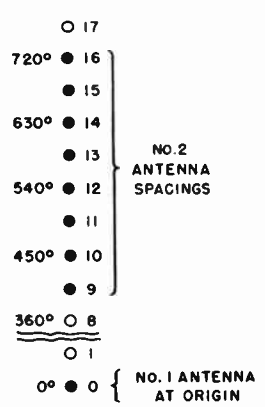

ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
0° 45° 90° 135° 180° 225° 270° 315°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
45° 90° 135° 180° 225° 270° 315° 360°
S_2 = SPACING PATTERN NUMBER = DEGREES / 45
1 2 3 4 5 6 7 8
0 1 2 3 4 5 6 7
ψ_2 = PHASING PATTERN NUMBER = DEGREES / 45
0 9
360° 8
7
270° 6
5
180° 4
3
90° 2
1
0° 0
NO. 2 ANTENNA SPACINGS
NO. 1 ANTENNA AT ORIGIN
SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS.
NUMBER OF ANTENNAS IN THE SYSTEM.
SPACING OF NO. 2 ANTENNA = DEGREES / 45
PHASING OF NO. 2 ANTENNA = DEGREES / 45
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA.
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA.
% R M S FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM.
No. S_2 ψ_2 E_1 E_2 E_0
2 01 0 50 50 95
PATTERN NOMENCLATURE

Figure B-1.



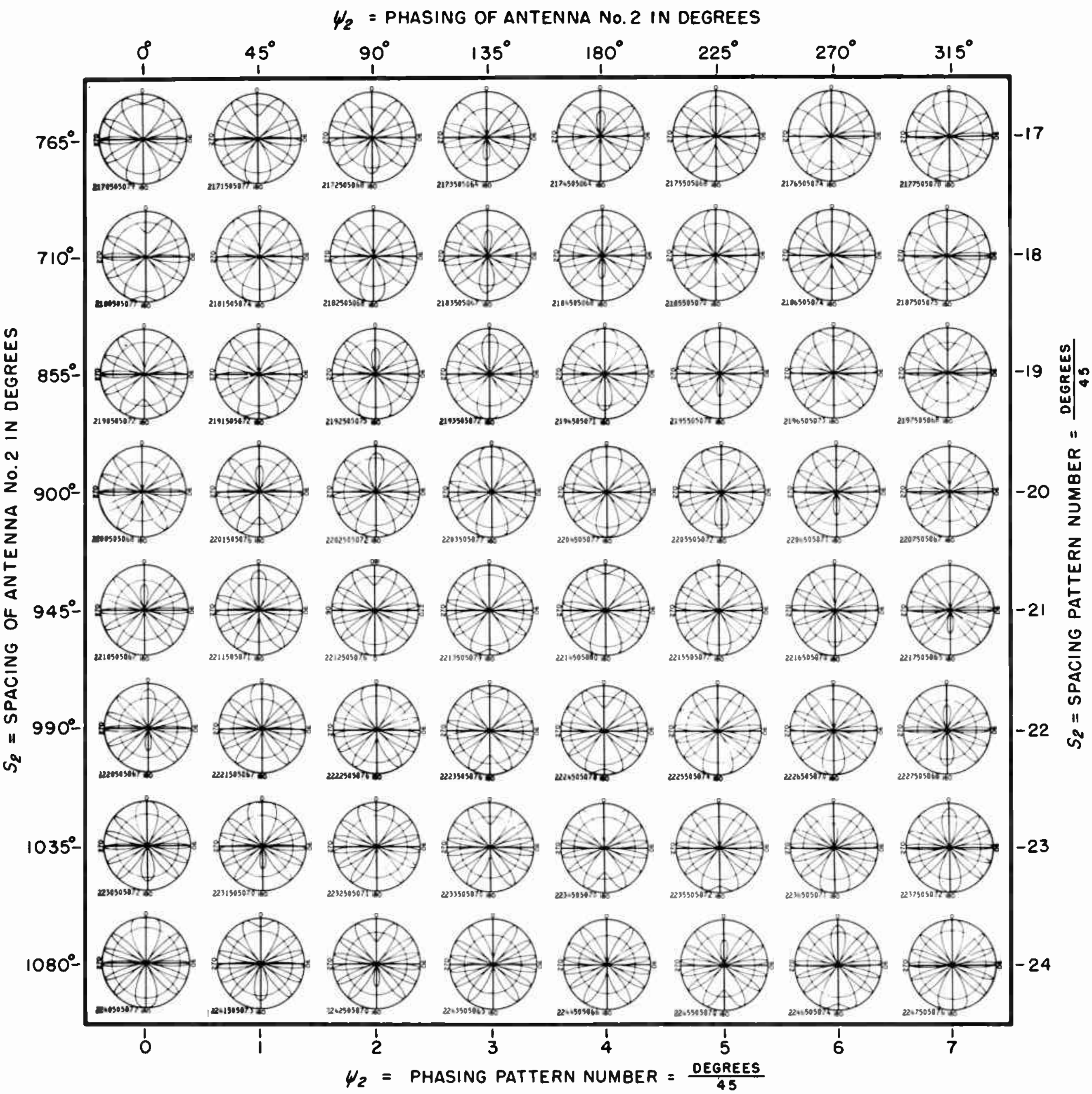


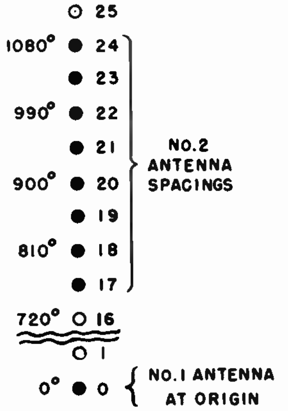
SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS.

NUMBER OF ANTENNAS IN THE SYSTEM.	SPACING OF NO. 2 ANTENNA = DEGREES / 45	PHASING OF NO. 2 ANTENNA = DEGREES / 45	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA.	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA.	% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM.
NO.	S_2	ψ_2	E_1	E_2	E_0
2	09	0	50	50	80

PATTERN NOMENCLATURE

Figure B-2.



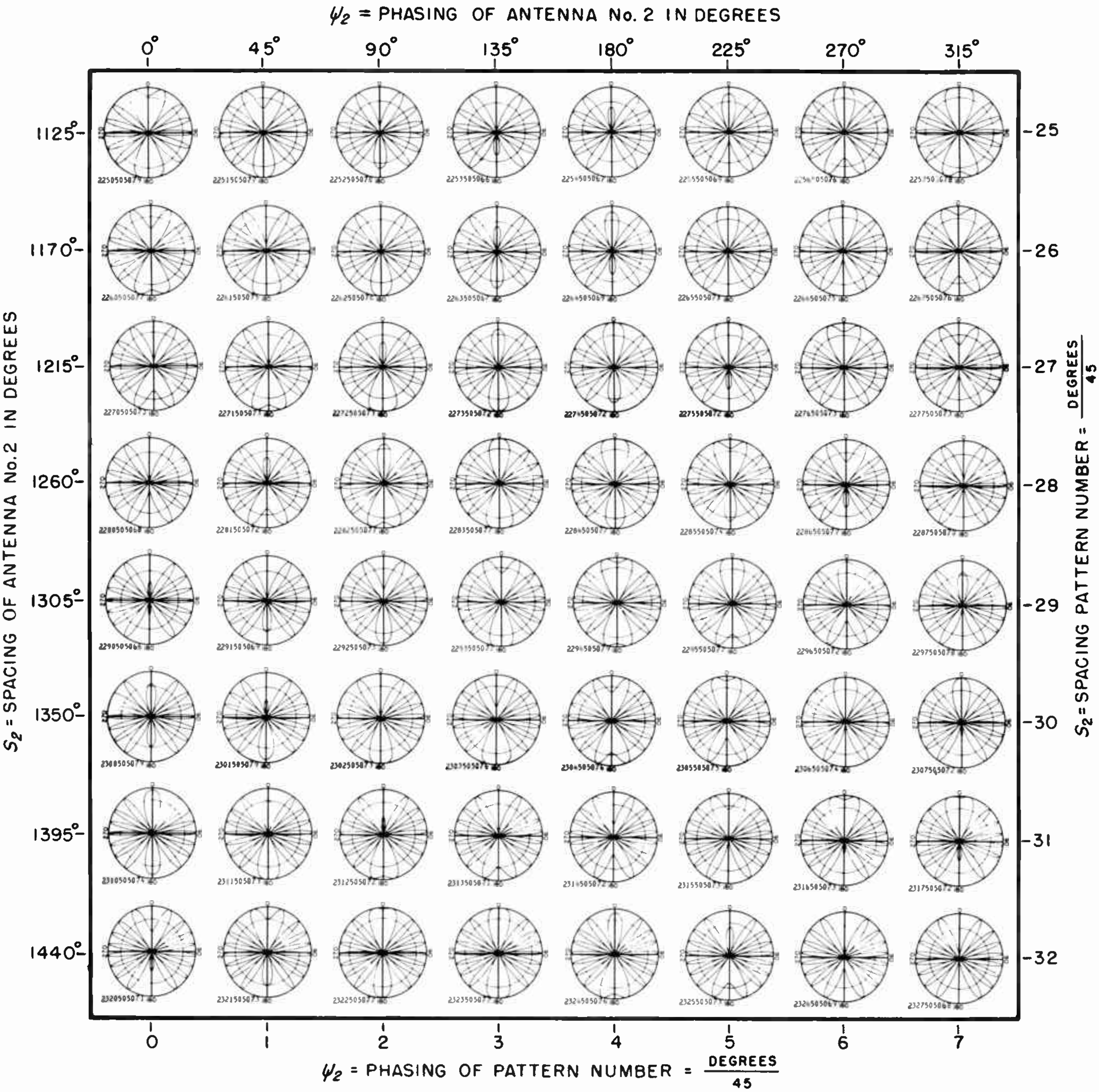


SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS

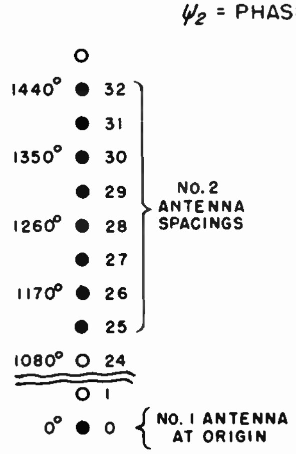
NUMBER OF ANTENNAS IN THE SYSTEM.	SPACING OF NO. 2 ANTENNA = $\frac{\text{DEGREES}}{45}$	PHASING OF NO. 2 ANTENNA = $\frac{\text{DEGREES}}{45}$	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA.	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA.	% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM.
No.	S_2	ψ_2	E_1	E_2	E_0
2	17	0	50	50	79

PATTERN NOMENCLATURE

Figure B-3.



CLEVELAND INSTITUTE OF RADIO ELECTRONICS
TERMINAL TOWER, CLEVELAND 13, OHIO.

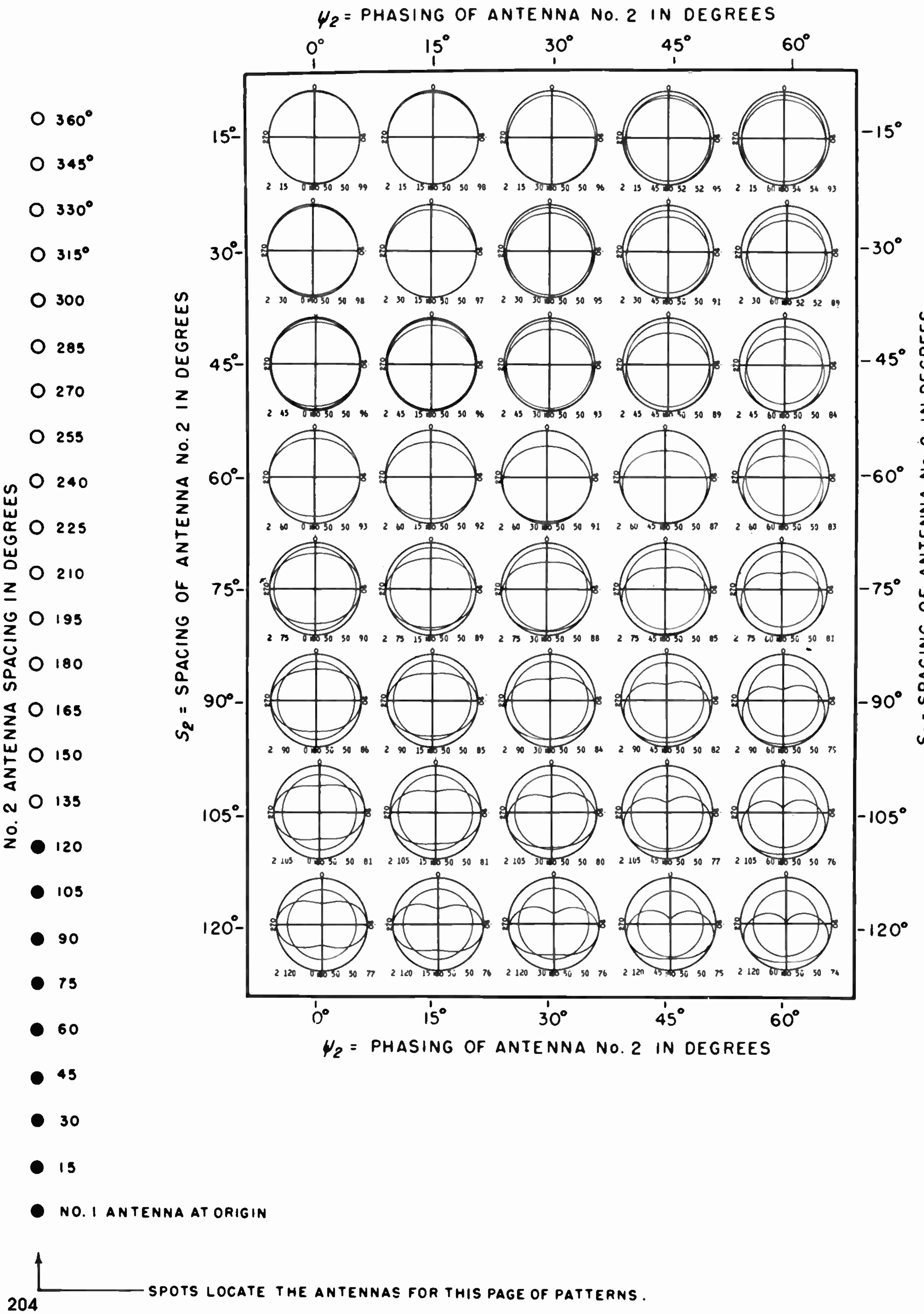


SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS.

NUMBER OF ANTENNAS IN THE SYSTEM.	SPACING OF NO. 2 ANTENNA = DEGREES / 45	PHASING OF NO. 2 ANTENNA = DEGREES / 45	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA.	% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA.	% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM.
No.	S_2	ψ_2	E_1	E_2	E_0
2	25	0	50	50	79

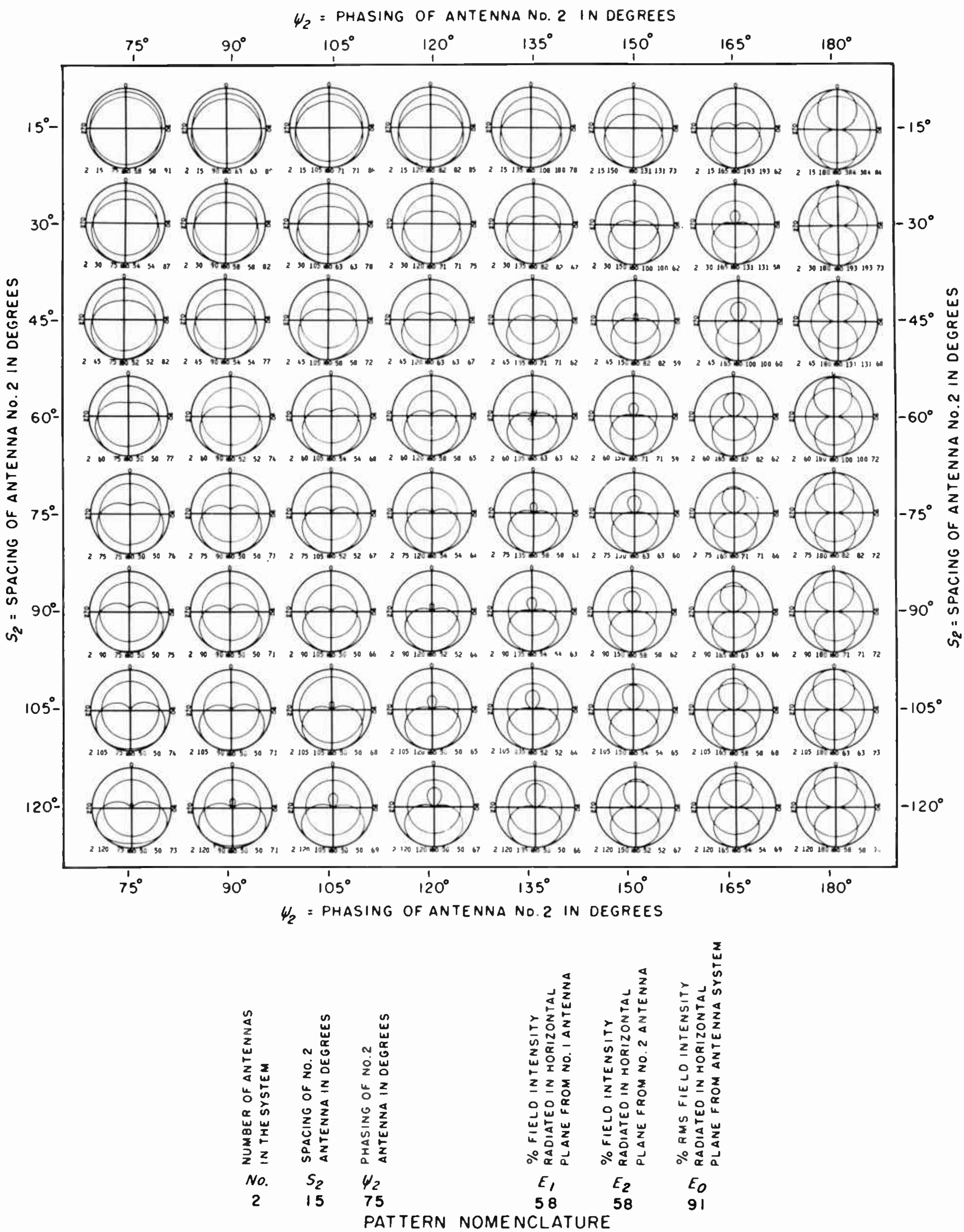
PATTERN NOMENCLATURE

Figure B-4.



ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
0°
15°
30°
45°
60°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
15°
30°
45°
60°
75°
90°
105°
120°
-15°
-30°
-45°
-60°
-75°
-90°
-105°
-120°
No. 2 ANTENNA SPACING IN DEGREES
360°
345°
330°
315°
300
285
270
255
240
225
210
195
180
165
150
135
120
105
90
75
60
45
30
15
NO. 1 ANTENNA AT ORIGIN
SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS.

Figure B-5.



ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
75° 90° 105° 120° 135° 150° 165° 180°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
15° 30° 45° 60° 75° 90° 105° 120°
NUMBER OF ANTENNAS IN THE SYSTEM
SPACING OF NO. 2 ANTENNA IN DEGREES
PHASING OF NO. 2 ANTENNA IN DEGREES
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA
% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM
No. S_2 ψ_2 E_1 E_2 E_0
2 15 75 58 58 91
PATTERN NOMENCLATURE

Figure B-6.

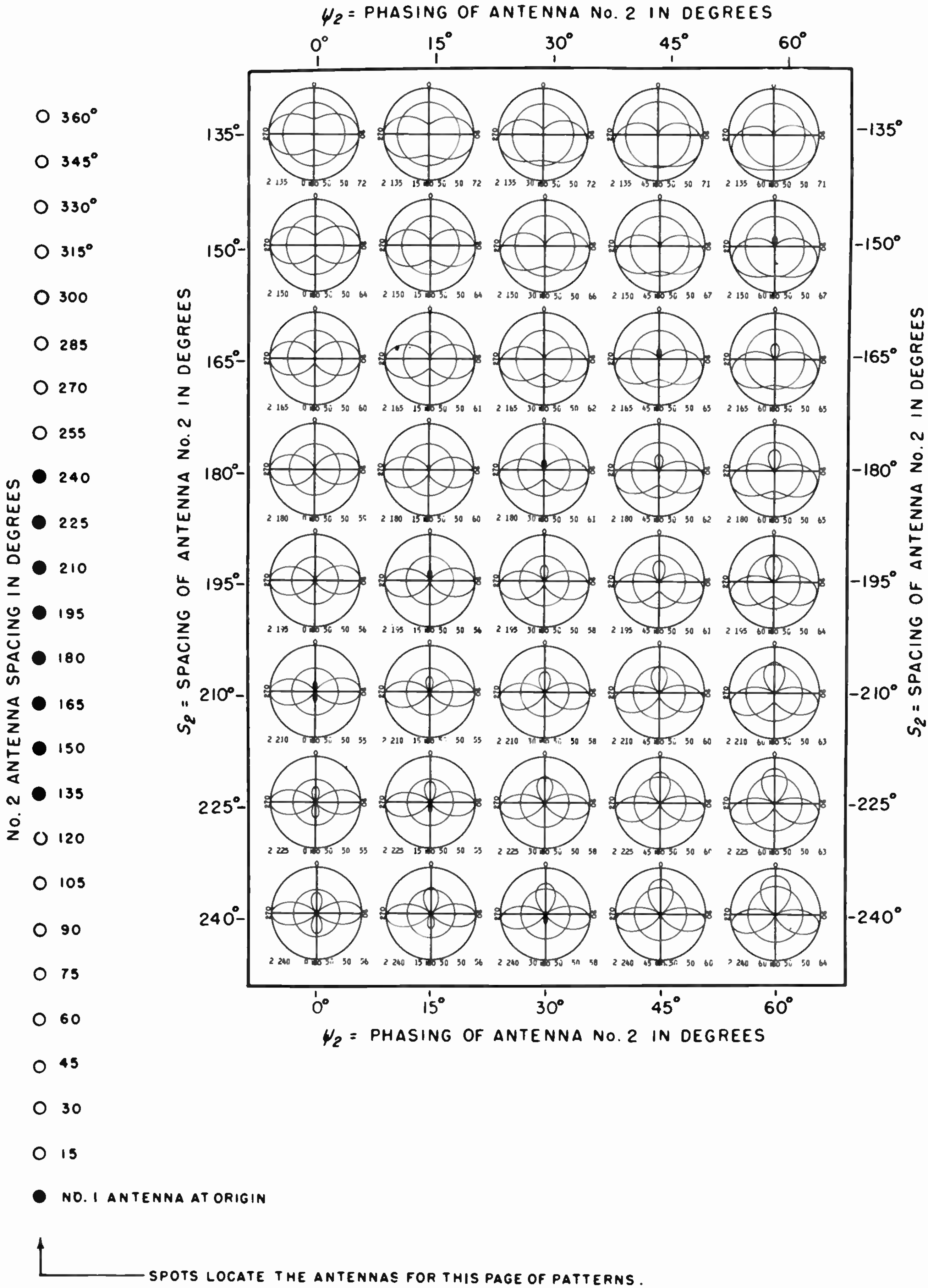

ψ2 = PHASING OF ANTENNA No. 2 IN DEGREES
0° 15° 30° 45° 60°
135° 150° 165° 180° 195° 210° 225° 240°
-135° -150° -165° -180° -195° -210° -225° -240°
S2 = SPACING OF ANTENNA No. 2 IN DEGREES
No. 2 ANTENNA SPACING IN DEGREES
360° 345° 330° 315° 300 285 270 255 240 225 210 195 180 165 150 135 120 105 90 75 60 45 30 15
NO. 1 ANTENNA AT ORIGIN
SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS.

Figure B-7.

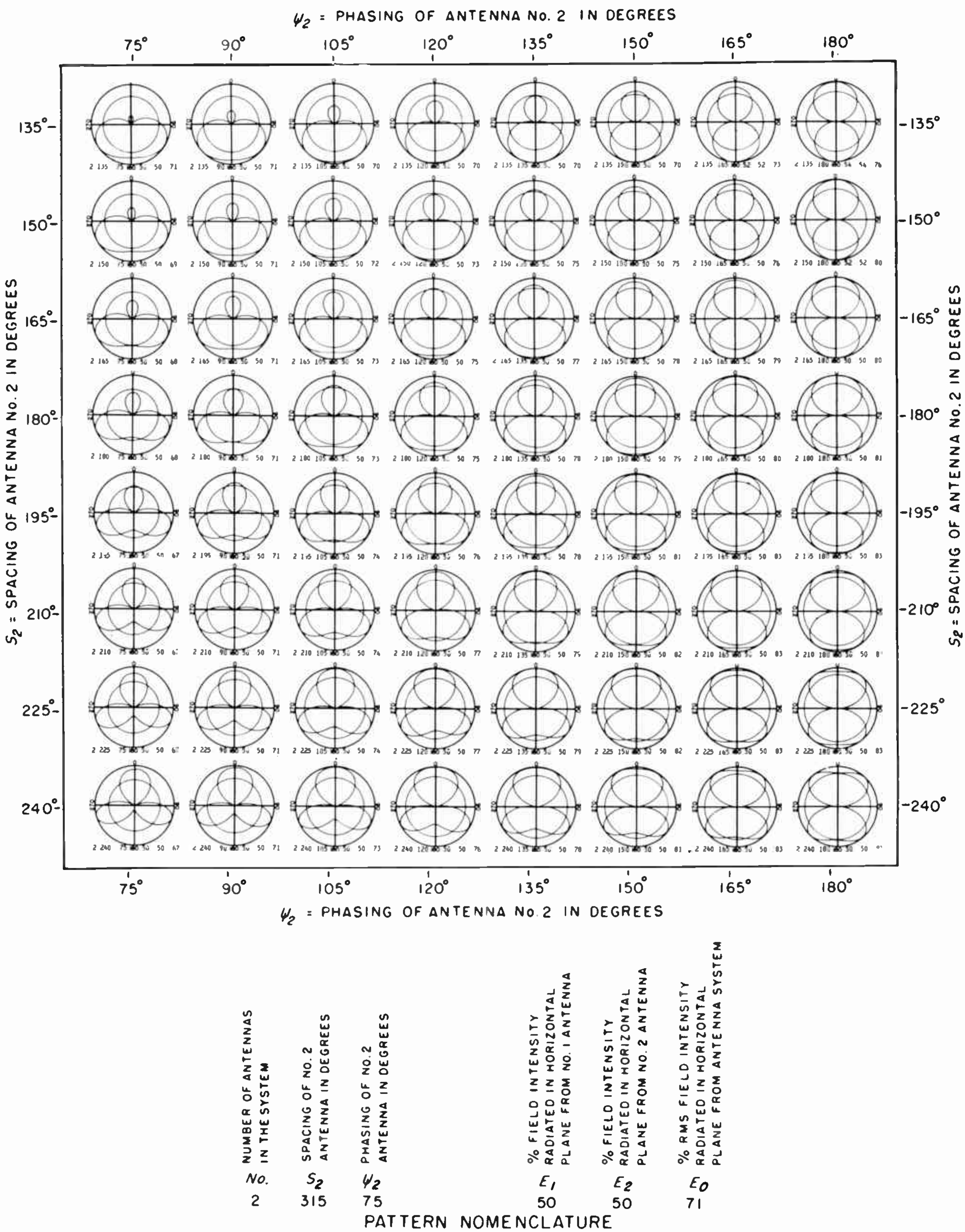

ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
75°
90°
105°
120°
135°
150°
165°
180°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
135°
150°
165°
180°
195°
210°
225°
240°
NUMBER OF ANTENNAS IN THE SYSTEM
SPACING OF NO. 2 ANTENNA IN DEGREES
PHASING OF NO. 2 ANTENNA IN DEGREES
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA
% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM
No.
S_2
ψ_2
E_1
E_2
E_0
2
315
75
50
50
71
PATTERN NOMENCLATURE

Figure B-8.

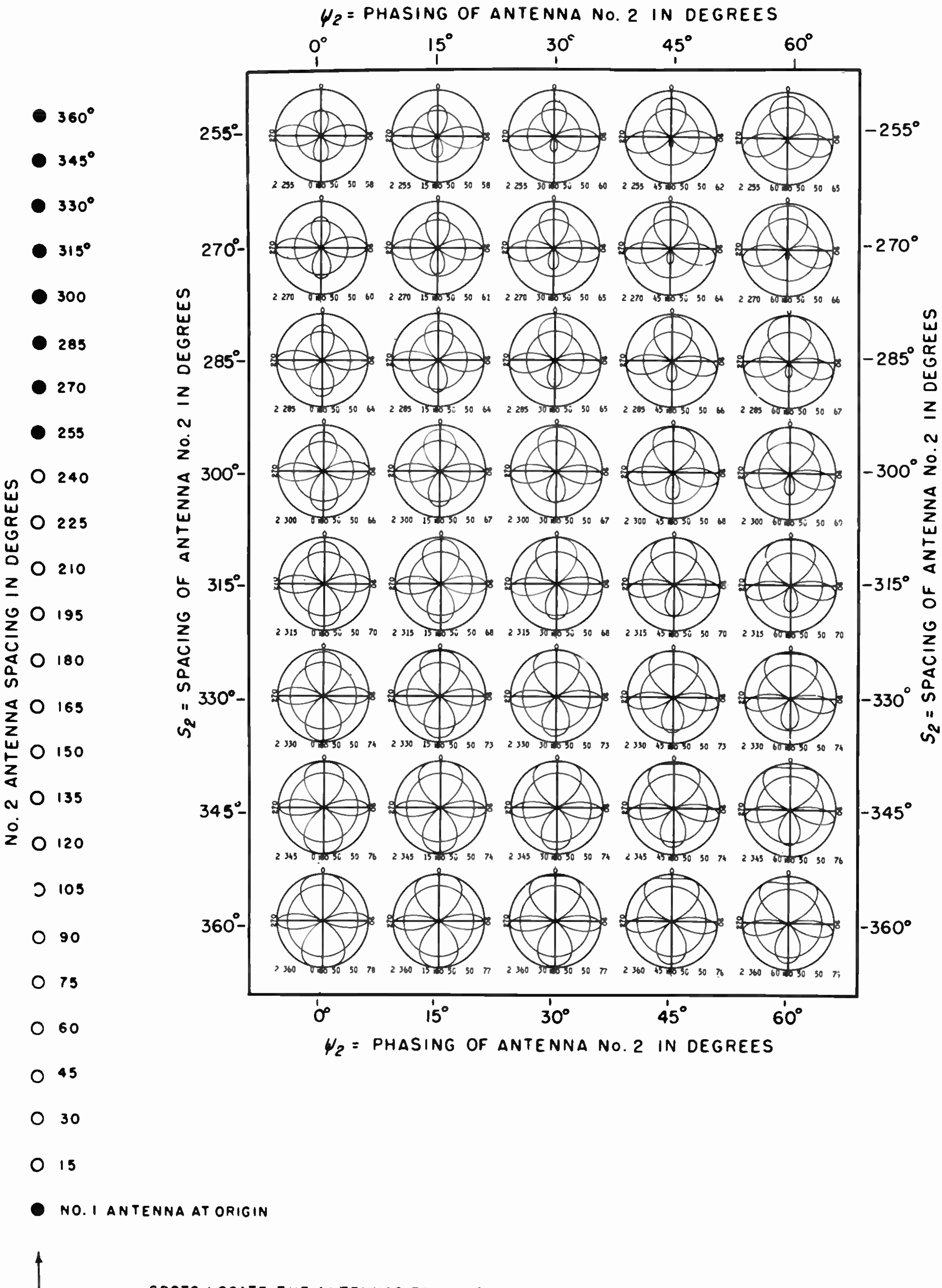
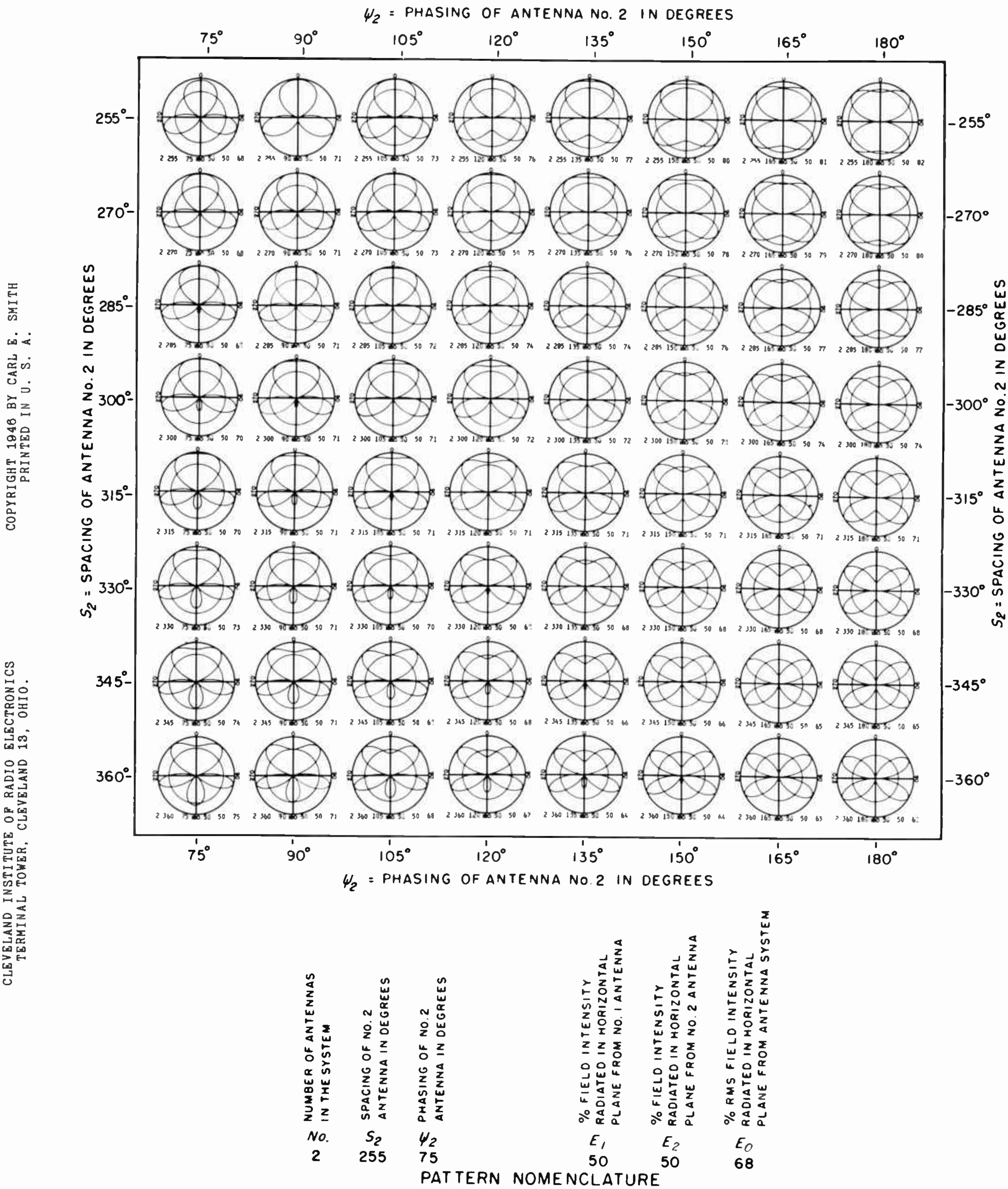

ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
0° 15° 30° 45° 60°
255° 270° 285° 300° 315° 330° 345° 360°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
No. 2 ANTENNA SPACING IN DEGREES
360° 345° 330° 315° 300 285 270 255 240 225 210 195 180 165 150 135 120 105 90 75 60 45 30 15
NO. 1 ANTENNA AT ORIGIN
SPOTS LOCATE THE ANTENNAS FOR THIS PAGE OF PATTERNS

Figure 8-9.



ψ_2 = PHASING OF ANTENNA No. 2 IN DEGREES
75° 90° 105° 120° 135° 150° 165° 180°
S_2 = SPACING OF ANTENNA No. 2 IN DEGREES
255° 270° 285° 300° 315° 330° 345° 360°
COPYRIGHT 1946 BY CARL E. SMITH
PRINTED IN U. S. A.
CLEVELAND INSTITUTE OF RADIO ELECTRONICS
TERMINAL TOWER, CLEVELAND 13, OHIO.
NUMBER OF ANTENNAS IN THE SYSTEM
SPACING OF NO. 2 ANTENNA IN DEGREES
PHASING OF NO. 2 ANTENNA IN DEGREES
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 1 ANTENNA
% FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM NO. 2 ANTENNA
% RMS FIELD INTENSITY RADIATED IN HORIZONTAL PLANE FROM ANTENNA SYSTEM
No. S_2 ψ_2 E_1 E_2 E_0
2 255 75 50 50 68
PATTERN NOMENCLATURE

Figure B-10.

APPENDIX C (CHAPTER 2.5, PART I)

PATTERN DEVELOPMENT OF DIRECTIONAL ANTENNAS

Theoretical Pattern Equation

The theoretical pattern equation of Appendix A can be written as follows by changing the k^{th} tower to the i^{th} tower to conform with the FCC practice, thus:

$$E(\phi, \Theta)_{th} = \left| k \sum_{i=1}^{n} F_i(\Theta) \; \underline{/ S_i \cos\Theta \cos(\phi_i - \phi) + \psi_i} \right| \qquad [C\text{-}1]$$

where:

k = multiplying constant which determines pattern size

Standard Pattern Equation

The standard pattern equation is obtained from Eq. C-1 by adding the quadrature term Q to fill minimums and increase the size by five percent, thus:

$$E(\phi, \Theta)_{std} = 1.05\sqrt{[E(\phi, \Theta)_{th}]^2 + Q^2} \qquad [C\text{-}2]$$

where Q is the greater of the quantities:

$$0.025\, g(\Theta)\, E_{rss} \qquad [C\text{-}3]$$

OR

$$10.0\, g(\Theta) \sqrt{P_{kw}} \qquad [C\text{-}4]$$

where g(Θ) is the vertical plane distribution factor, $f(\Theta)$, for the shortest element in the array (see Eq. C-2 above; also see FCC Rules Section 73.190, Figure 5).

If the shortest element has an electrical height in excess of 0.5 wavelength, g(Θ) shall be computed as follows:

$$g(\Theta) = \frac{\sqrt{\{f(\Theta)\}^2 + 0.0625}}{1.030776} \qquad [C\text{-}5]$$

$$E_{rss} = \sqrt{\sum_{i=1}^{n} E_i^2} \qquad [C\text{-}6]$$

As an example, consider a two-tower array. The theoretical pattern equation (Eq. C-1) becomes:

$$E = E_1 f_1(\Theta) \; \underline{/0^\circ} + E_2 f_2(\Theta) \; \underline{/ S_2 \cos\Theta \cos(\phi_2 - \phi) + \psi_2} \qquad [C\text{-}7]$$

Now, for 5 kilowatts, 90 degree towers, and the following parameters:

Tower No.	Height G°	Field Ratio	Spacing S°	True Bearing φ°	Phase ψ°
1	90	1.0	0	0	0
2	90	1.0	90	0	−90

E_{rss} (theoretical pattern) = 691.92 mV/m
Q (quadrature term) = 21.60 mV/m
E_{rms} (standard pattern) = 726.87 mV/m

A plot of the theoretical and standard patterns are shown in figure C-1.

The minimum horizontal field strength (at one kilometer) when the theoretical field strength goes to zero is given by Eq. C-2 for a standard pattern along the ground using Eq. C-4 with g(Θ) = 1.0. For one kilowatt and under, Q is 6 according to FCC Rules. For various FCC licensed values of power, the minimum field strength values are as follows:

P_{kW}	Q	$E_{mV/m}$
0.25	6.00	10.14
0.50	6.00	10.14
1.00	6.00	10.14
2.50	9.49	16.03
5.00	13.42	22.67
10.00	18.97	32.06
25.00	30.00	50.69
50.00	42.43	71.69

The minimum field strength (at one kilometer) for any elevation of a standard pattern, by Eq. C-5, is:

$$g(\Theta) = \frac{\sqrt{0 + 0.0625}}{1.030776} = 0.2425$$

Augmented Pattern Equation

The augmented pattern equation is obtained by adding an augmentation quadrature term to the standard pattern as given here:

$$E_{(\phi,\Theta)aug} = \sqrt{\{E_{(\phi,\Theta)std}\}^2 + A\left\{g(\Theta)\cos\left(180\frac{D_A}{S}\right)\right\}^2}$$

where: $E_{(\phi,\Theta)aug}$ = augmented pattern radiation value at azimuth, elevation

$E_{(\phi,\Theta)std}$ = standard pattern radiation value at azimuth, elevation

A = augmentation constant (insert equation C-10)

$g(\Theta) = \dfrac{\sqrt{\{f(\Theta)\}^2 + 0.0625}}{1.030776}$

S = span of augmentation in degrees

D_A = angular distance from center of span

The principle of augmentation is illustrated in the cardioid pattern of Fig. C-2.

The FCC has converted all augmented directional patterns to a table for each station as shown in the example of Fig. C-3. In this case there were six augmentations as tabulated in Fig. C-4 and shown on the polar chart of Fig. C-5.

It should be noted that where the spans overlap, Eq. C-9 is applied repeatedly, once for each augmentation, proceeding clockwise from true north.

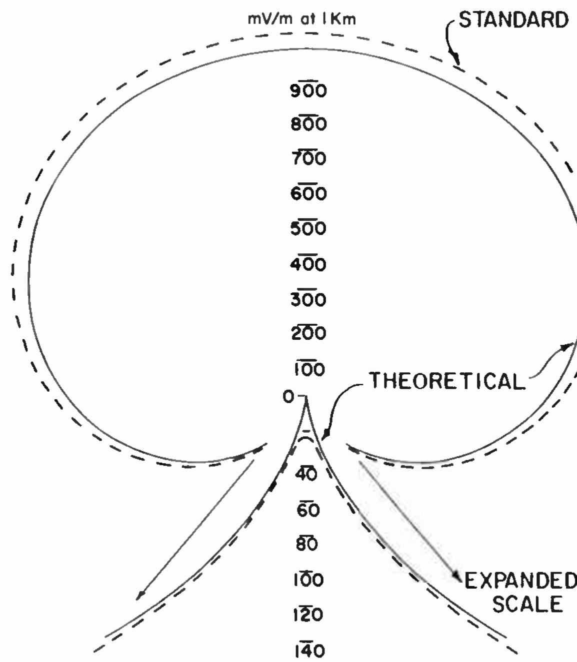


Figure C-1. Theoretical and standard pattern.

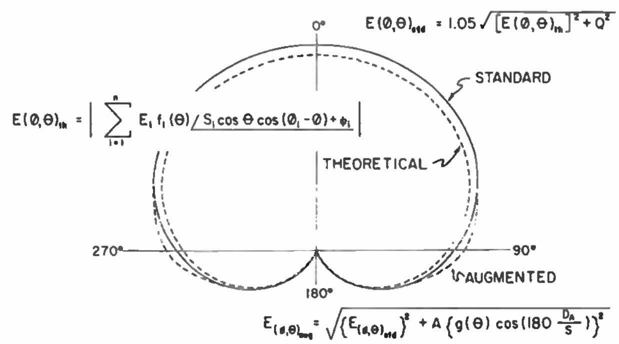


Figure C-2. Theoretical, standard and augmented pattern.

Technical Parameters Resulting From Conversion of AM Broadcast Stations To Standard Patterns

STANDARD PATTERN CONVERSION NO.: 1280–22

FREQ. KHZ	CALL LETTER	CITY	STATE	HRS.	PATTERN STATUS	CLASS
1280	WHVR	HANOVER	PA	N	LIC.	3

POWER KW	LATITUDE	LONGITUDE	PAT-MULT MV/M	TH-RMS MV/M	STD/AUG RMS-MV/M	PAT-RSS MV/M	Q-FACTOR
.500	39-49-11	77-00-25	131.27	143.00	150.54	185.65	6.0000

TOWER NO.	PHYS-HT (A)-DEG	TL-HT (B)-DEG	TOT-HT (C)-DEG	TL-HT (D)-DEG	FIELD RATIO	PHASE DEG.	SPACING DEG.	ORIENT DEG-TR	REF FLG
1	91.0	.0	.0	.0	1.000	149.5	.0	.0	
2	91.0	.0	.0	.0	1.000	.0	90.0	178.0	

AUGMENTATION DATA

CENTRAL AZIM. DEGREES TRUE	SPAN DEGREES	FIELD AT AZIM. MV/M
64.0	12.0	17.0
260.5	55.0	103.0
288.0	14.0	7.5
288.0	10.0	21.2
295.0	14.0	30.0
295.0	10.0	43.3

HORIZONTAL PLANE STANDARD/AUGMENTED RADIATION VALUES

AZ. DEG	FIELD MV/M	AZ. DEG	FIELD MV/M	AZ. DEG	FIELD MV/M	AZ. DEG	FIELD MV/M	AZ. DEG.	FIELD MV/M	AZ. DEG	FIELD MV/M
0	136.8	60	29.1	120	174.0	180	239.4	240	163.9	300	42.2
10	132.8	70	9.0	130	196.3	190	237.0	250	136.2	310	71.0
20	123.1	80	43.5	140	213.5	200	231.5	260	104.7	320	95.4
30	107.5	90	80.0	150	225.9	210	221.4	270	69.3	330	114.4
40	86.3	100	115.0	160	234.0	220	207.2	280	31.3	340	127.7
50	59.8	110	146.7	170	238.4	230	188.0	290	20.4	350	135.1

CONSTRUCTION PERMIT LIMITS

AZIMUTH DEG. TRUE	PRESENT MV/M	NEW MV/M
64.0	17.0	17.0
231.0	179.0	185.8
288.0	33.0	21.2
352.0	131.0	135.9

PATTERN MINIMA

AZIMUTH DEG. TRUE	FIELD MV/M
68.5	6.8
284.5	17.3
290.0	20.4
299.0	40.9

PATTERN MAXIMA

AZIMUTH DEG. TRUE	FIELD MV/M
178.0	239.4
288.0	21.2
296.0	44.2
358.0	136.9

Figure C-3. FCC method of specifying augmentation.

CENTER AZIMUTH OF AUGMENTATION	SPAN DEGREES	EXTENT OF SPAN	FIELD AT CENTER SPAN AT 1 KM
64°	12°(*6°)	(58°–70°)	27.4
260.5°	55°(*27.5°)	(233°–288°)	165.8
288°	14°(*7°)	281°–295°)	12.1
288°	10°(*5°)	(283°–293°)	34.1
295°	14°(*7°)	(288°–302°)	48.3
295°	10°(*5°)	(290°–300°)	69.7

Figure C-4. Table of augmentation data.

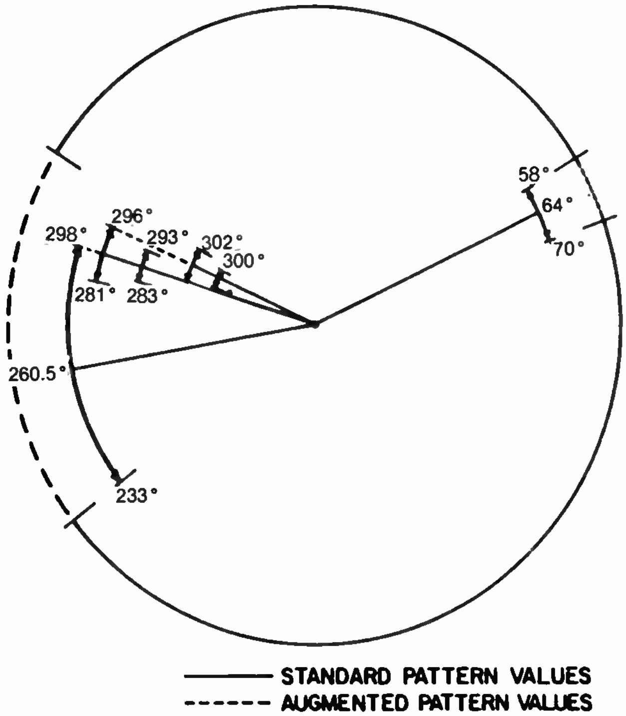


Figure C-5. Augmented pattern flow chart showing overlapping spans.

2.5 AM Broadcast Antenna Systems

Part II: Antenna Coupling and Phasing Systems for AM Broadcasting

Edward Edison, P.E.
Hammett & Edison, Inc., San Francisco, California

AM broadcast antenna coupling and phasing systems consist principally of passive networks of inductors, capacitors, transmission lines, and auxiliary components such as lightning gaps, meters, jacks, and relays. This equipment can range from simple (such as a nondirectional station having the transmitter building located adjacent to the tower base) to complex (such as a multi-tower directional array requiring different patterns day and night). To understand antenna systems and their environment, we must consider the function of such systems and their performance objectives, as well as the characteristics of basic networks, power dividers, transmission lines, sampling systems, detuning systems, and transmitter load optimization.

It is not the intent of this chapter to make directional antenna "experts" out of its readers. However, in providing a broad perspective of the design of such systems, we hope to encourage a rational approach to the everyday problems of directional antenna operation and maintenance. The design and adjustment of the more complex directional arrays require considerable knowledge, experience, and computer power that is beyond the ken of almost all chief engineers and contract engineers. When unusual problems arise that do not yield to reasonable efforts at repair and adjustment, it is time to call in an expert; preferably the consulting engineer or firm that designed or previously adjusted the system.

THE FUNCTION OF A DIRECTIONAL ANTENNA PHASING SYSTEM

The function of a directional antenna phasing system is to distribute current to each tower with controllable phase and amplitude so as to generate the desired directional pattern. This function is accomplished typically by means of the following:

1. A network at each tower matches each tower load to its transmission line. (See Fig. 1.) This network is typically in a weatherproof box, termed an *antenna coupling unit* (ACU), or within a small building called a "doghouse."
2. The remaining equipment, usually housed in one or more indoor cabinets, is termed a *phasor*.
   A. Within the phasor cabinet (usually located in the transmitter building), networks control the phase of the current into each transmission line.
   B. Power divider circuits control the current amplitude in each transmission line.
   C. A common-point matching network adjusts the phasor input impedance to a desirable resistive value (typically 50 ohms) without disturbing the phase or amplitude of the tower currents.

Fig. 1 shows a system with only two towers. Each additional tower requires its own antenna coupling unit, transmission line, and phasing network, as well as additional components in the power divider.

PHASOR PERFORMANCE OBJECTIVES

The essential performance objective is to have both the phasor input impedance and the directional antenna radiation pattern remain essentially constant at all frequencies within ±10 kHz of the assigned channel. The performance may be disappointing if the design process proceeds by a piecemeal approach, where subsystems are independently designed for each of the functions illustrated in Fig. 1 and are then connected together. It is better to use a systems approach that considers the entire collection of components between the transmitter and the towers as a whole rather than by the "building block" approach implicit in Fig. 1. No

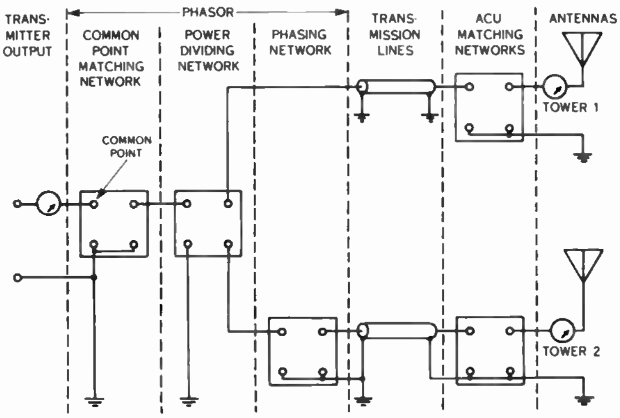


Figure 1. Block diagram of a directional antenna.

one subsystem in an array can be treated in isolation. While it is a simple matter to design a series of subsystems to provide the desired phase and amplitude parameters, these subsystems must properly complement each other if optimum bandwidth is to be achieved.

Phasor Input Impedance

An ideal phasor would present a load to the transmitter that is the same pure resistance at all sideband frequencies. However, phasors consist of coils and capacitors, whose reactance varies with frequency; therefore, this ideal cannot be achieved. The term "impedance bandwidth" is used to describe the degree of constancy of the phasor input impedance across the entire range of sideband frequencies. A useful figure-of-merit to express the impedance bandwidth as a single number is the worst-case voltage standing wave ratio (VSWR) existing at either of the two 10 kHz sideband frequencies when the system is perfectly matched at the carrier frequency. This approach yields VSWR numbers that are related to antenna performance just as television and FM transmitting antenna performance is described by VSWR limits at various sideband frequencies. The sideband VSWR can be determined by analysis of common-point impedance measurements. The resistance and reactance at each sideband frequency of interest are expressed as a percentage of the measured carrier frequency resistance and are then plotted on a Smith Chart. A Smith Chart depicts the relationship between the load impedances at sideband frequencies and the ideal matched load. The distance from the center of the chart for each sideband frequency is a measure of the VSWR at that frequency.

Bandwidth

Bandwidth requirements for each channel change greatly across the broadcast band. At 1600 kHz, a total bandwidth of 20 kHz corresponds to about one percent of the center frequency whereas, at 540 kHz, 20 kHz corresponds to nearly four percent of the center frequency. Comparatively, an AM station at the lower frequencies has a more stringent bandwidth requirement than is needed for acceptable visual transmitter performance on VHF TV channels 7 through 13 or on any of the UHF channels.

The common-point VSWR at 10 kHz above and below carrier often is better than 1.1:1 in good phasor designs. Poor designs can be worse than 3:1. Poor phasor designs manifest themselves by poor audio performance. In the main lobe of a pattern, poor bandwidth will limit frequency response; in the null sectors of a pattern, distortion may be exacerbated by limited bandwidth which causes the carrier frequency to be suppressed more than the sideband frequencies. The audible effect of the resulting distortion is similar to over-modulation.

Radiation Pattern

An ideal phasor would produce an antenna radiation pattern that remains unchanged across the channel. Variations in the phase and ratio of tower currents at the sideband frequencies can result in changes in the location and depth of the intended pattern minimums. The frequencies for which the pattern remains useful describe the *pattern bandwidth*. Pattern bandwidth cannot be quantified by simple numbers, but the concept is useful in comparing alternative designs.

The example of poor pattern bandwidth shown in Fig. 2 is for the three-tower daytime pattern of an actual station at its carrier frequency and at the 10 kHz sideband frequencies. The sideband patterns were calculated from the actual measured ratios and phases of tower currents at both sideband frequencies. This directional antenna system did not adequately protect an adjacent channel station, even though conventional field strength measurements indicated proper operation. A spectrum analyzer displaying the received signal confirmed the extreme sideband asymmetry in critical directions.

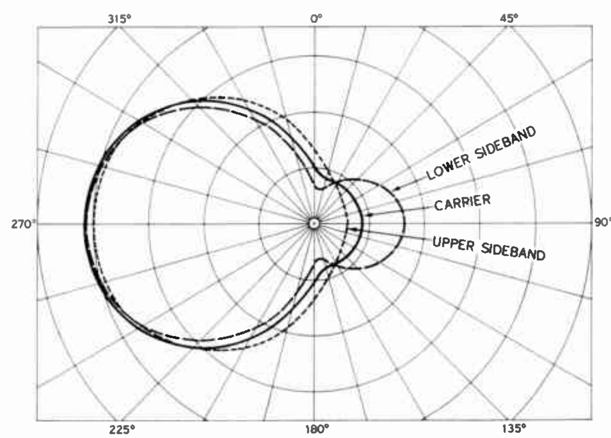


Figure 2. Example of poor pattern bandwidth.

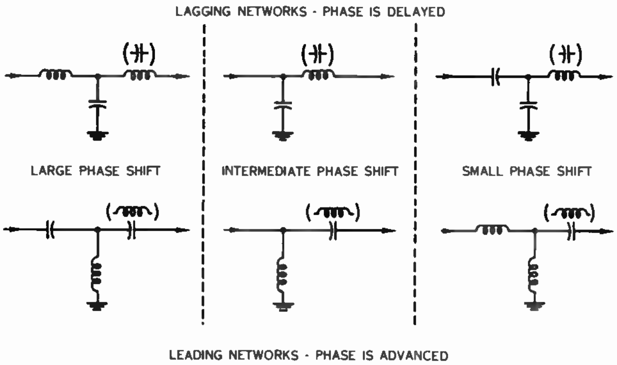
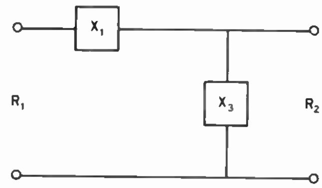


Figure 3. Conventional "T" and "L" networks.



$$X_1 = \pm\sqrt{R_1R_2 - R_1^2}$$

$$X_3 = \mp\frac{R_1R_2}{X_1}$$

$$\beta = \cos^{-1}\sqrt{\frac{R_1}{R_2}}$$

R_2 MUST BE GREATER THAN R_1

X_1 IS INDUCTIVE (+J) FOR LAGGING NETWORK

X_1 IS CAPACITIVE (-J) FOR LEADING NETWORK

X_3 MUST HAVE SIGN OPPOSITE TO X_1

Figure 5. "L" network formulas.

BASIC NETWORKS

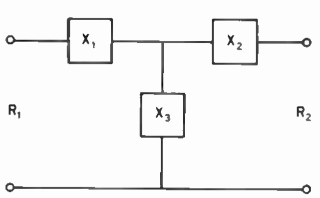
The networks used for matching impedances or for shifting phase are typically "T" or "L" configurations. "Pi" networks having shunt elements at the input and output and a central series element can be made electrically identical to any T network, but T networks are easier to adjust.

The networks at the left in Fig. 3 offer large phase shifts, while the configurations on the right offer small phase shifts. Intermediate between these two conditions are L networks which can be considered as T networks with one zero-reactance arm. L networks do not permit independent adjustment of impedance match and phase shift; however, the resulting phase delay or advance can be calculated easily, with the result often being a value that is compatible with the overall phasor design. The formulas for these networks are all based on matching into resistive loads.

The loads presented by antenna towers usually have a reactive component; therefore, the output arm of the matching network is modified to cancel the tower reactance. When this modification is made, the output arm will occasionally assume the opposite reactance sign, as shown in the parentheses in Fig. 3.

T Networks

The reactances of the input, output, and shunt arms of a T network can be calculated quickly, once the desired input and output resistances and the phase shift (b) are specified, by using the formulas shown in Fig. 4. The formulas are equally applicable for leading (+b) and lagging (−b) networks.



$$X_3 = \frac{\sqrt{R_1R_2}}{\sin\beta}$$

$$X_2 = \frac{R_2}{\tan\beta} - X_3$$

$$X_1 = \frac{R_1}{\tan\beta} - X_3$$

R_1 = INPUT RESISTANCE

R_2 = OUTPUT (LOAD) RESISTANCE

β = PHASE SHIFT (OUTPUT-INPUT)

Figure 4. "T" network formulas.

L Networks

Conventional L networks can provide a match between any two resistance values. The formula for such networks is shown in Fig. 5, which presumes a nonreactive load. If an L network is used to match a tower to a transmission line, the reactive component of the tower impedance must be considered. If the tower resistance is lower than the line impedance, the series arm of the L network is connected to the tower. This arm can then be modified from the value calculated by the formula so as to also cancel the tower reactance.

An L network can also be used to match a tower having a resistance higher than the transmission line impedance, but in this case the formula in Fig. 5 does not apply unless the tower impedance contains no reactance. In the usual case, where the tower impedance is reactive, an L network cannot be so easily calculated. The shunt arm (which is in parallel with the reactive tower load) must be one which makes the resistive component of the parallel combination equal to the transmission line impedance. Then the resulting reactance of the parallel combination is cancelled by the series input arm of the L network. Such networks are easy to adjust; the shunt arm of the L network is adjusted to make the input resistance match the transmission line characteristic impedance, and the series arm is then adjusted to cancel the resulting reactive component. The two adjustments are substantially independent of each other. If the tower is one element of a directional array, an operating impedance bridge (hot bridge) at the network input must be used and the phase and ratio parameters must be approximately correct in order for the tower operating impedance to be that which will exist when the directional antenna array is operating normally.

When the tower resistance is higher than the line impedance, there are two different values of shunt arm reactance (one capacitive and one inductive) that can satisfy the conditions necessary for a match with an L network. If the shunt arm is of the same sign as the tower reactance, the required series arm will be of opposite sign and the phase shift will be small. If the shunt arm is opposite in sign to the tower reactance, the required reactance in the series arm will have the same sign as the tower reactance and the phase shift will be large.

An alternative way to design such L networks (rather than by calculating the shunt arm reactance necessary to produce the desired resistive component for the parallel combination) is to think of them as T networks in which the output arm is represented by the reactive component of the tower load. As such, these networks can be described as *Phantom T Networks*.

Phantom T Networks

A phantom T network uses the tower reactance itself as the output arm (X_2) of the network. It is electrically equivalent to a conventional T network but requires one less circuit element. Because it has only two adjustable elements, the phase shift of a phantom T cannot be independently selected but two choices are possible, corresponding to T networks with small and large phase shifts. When the input resistance and reactance are properly adjusted, the resultant phase shift is then defined. Phantom T networks can be calculated quickly by the T network formulas while using cut-and-try variations of the phase shift until X_2 exactly equals the tower reactance. Phantom T networks with small phase shifts exhibit excellent bandwidth. Those with large phase shifts are somewhat poorer.

POWER DIVIDERS

In most phasors, separate components are used to accomplish the power division function; but the same networks that control phase shift can also control the division of power, if desired. It is important to remember, when comparing different towers in an array, that the power delivered to a tower is not in proportion to the base current; a low base resistance

Given: Total power is 1000 watts with base current ratios and resistances as shown. I = Current in Tower 3.

Tower No.	Base Current Ratio	Base Resistance	Power
1	0.20	25 ohms	$(0.20\,I)^2 \times (25) = 1.0\,I^2$
2	0.70	130	$(0.70\,I)^2 \times (130) = 63.7\,I^2$
3	1.00	40	$(1.00\,I)^2 \times (40) = 40.0\,I^2$
4	0.50	10	$(0.50\,I)^2 \times (10) = 2.5\,I^2$

$107.2\,I^2 = 1000$ watts;
$I = 3.054$ amps

Tower No.	Base Current	Base Power
1	0.20 x 3.054 = 0.611 amps	$(0.611)^2 \times (25) = 9.3$ watts
2	0.70 x 3.054 = 2.138	$(2.138)^2 \times (130) = 594.2$
3	1.00 x 3.054 = 3.054	$(3.054)^2 \times (40) = 373.1$
4	0.50 x 3.054 = 1.527	$(1.527)^2 \times (10) = 23.4$
		1000.0 watts

Figure 6. Example of power division calculations.

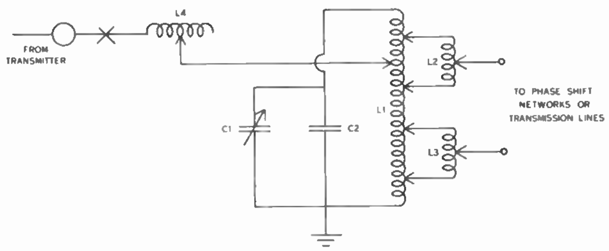


Figure 7. Typical tank-type power divider.

tower may consume very little power (I^2R) even if the current is relatively high. The actual power division is defined by the required base current ratios and the base resistances. An example of power division calculations is shown in Fig. 6.

Tank-Type Divider

In a tank-type power divider, as shown in Fig. 7, roller coils L2 and L3 (termed "jeep" coils in this configuration) adjust the power; the combination of L1, L2, and L3 is tuned close to resonance by capacitors C1 and C2. The input tap on L1 adjusts the common-point resistance; L4 adjusts the common-point reactance. This system is no longer popular because it is difficult to adjust and because bandwidth is limited by the excessive stored energy.

Shunt-Type Divider

A common type of divider is the "shunt" design shown in Fig. 8, which uses several adjustable coils in parallel, each of which controls the current to an individual tower. The reactance of the L1, L2, L3 combination is then tuned to resonance with C1. Component values are selected so that the impedance at point A is on the order of 400 ohms. A conventional T network then matches this load to the desired common-point impedance. This design is easy to adjust; however, the rollers on the variable coils may tend to bind and develop intermittent contacts after several years of use.

T-Network Power Division

A phasor can be designed to exploit the input impedance of conventional T networks to achieve control of both power division and phase, as shown in Fig. 9. In such a system, each T network presents a load at its input which is appropriate to divert the required power to its corresponding tower.

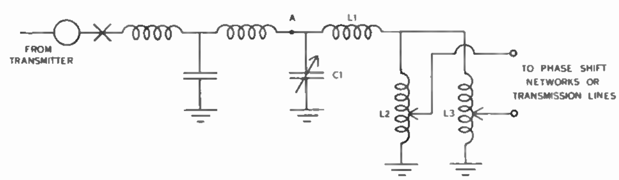


Figure 8. Typical shunt power divider.

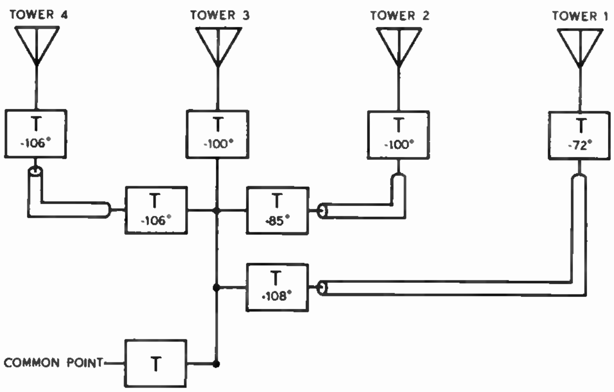


Figure 9. Typical phasor with "T" net power division.

Once such a phasor is properly adjusted, the small changes in phase and ratio that may later be required to maintain an array within tolerances are easily accomplished with very small adjustments in only the shunt and input or output arms of each individual tower network in the phasor. Although the phase and ratio adjustments interact (as with all phasors, due to the mutual coupling between towers), the shunt arms tend mostly to control power division, whereas the series input arms tend mostly to control phase. By noting which parameters are farthest out of tolerance, it is a simple matter to make small adjustments in the networks, observe the results of each change, and expeditiously restore such a phasor to the desired parameters.

Simplified Power Division

For two-tower directional systems, some very simple configurations can permit adjustment of both power division and phase shift with very few circuit elements. One such simplified system is a so-called "back-to-back" phasor; an example is shown in Fig. 10.

Components "C" and "L" are connected "back-to-back" at the output of the common-point matching network. In this example (which assumes matched

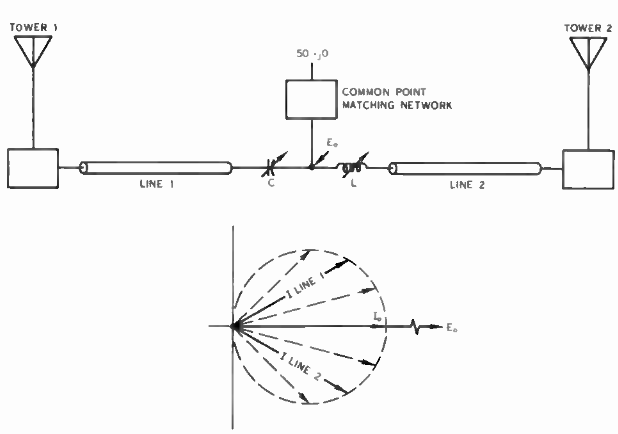


Figure 10. Example of a back-to-back phasor.

transmission lines with resistive inputs), reducing L to zero would yield a current into transmission Line 2 that is in phase with the applied voltage, E_0. As L is increased, the current into Line 2 diminishes and lags the applied voltage. As L is varied from zero to infinity, the locus of all possible Line 2 currents is the semi-circle on the lower half of the diagram. Similarly, as C is reduced, the amplitude of the current in Line 1 is reduced and its phase is advanced relative to the applied voltage. The locus of all the possible currents into Line 1, as C is varied from infinity to zero, is the upper semi-circle of the vector diagram. The ranges of adjustment of C and L need only be sufficient to reach the desired phase and ratio parameters, not the entire semi-circular gamut.

Independence of adjustment of phase and ratio is completely lacking in a back-to-back design; yet the phasor is easy to adjust. If one wants to change the ratio, an arbitrary change is first made in either component, phase is restored to the initial value by adjustment of the opposite component, and then the direction of ratio change is noted. It becomes a very simple cut-and-try process, not unlike that required to maximize transmitter efficiency by adjusting transmitter tuning and loading.

It is not always necessary to employ a coil and a capacitor in a back-to-back phasor. Often, the phase shift in the transmission lines and ACU networks can be tailored to provide the desired parameters with a series capacitor on the input to each line, rather than one capacitor and one coil. This arrangement permits a two-tower phasor in which both power division and phase adjustment are accomplished by only a pair of variable vacuum capacitors, offering excellent stability and resetability. This arrangement also provides excellent impedance bandwidth.

TRANSMISSION LINES

Many of the requirements for AM broadcast transmission lines differ from those used in FM and TV installations; the similarities and differences are discussed below.

Line Losses

The losses in transmission lines at AM broadcast frequencies are largely in the copper, not in the dielectric. For this reason, and for mechanical convenience, semi-flexible foam-filled lines are the most popular. Air dielectric lines are needed only when power levels exceed the ratings of available foam-filled lines.

Jacketing

All lines should include a black polyethylene outer jacket. Unjacketed lines can result in inadvertent poor contacts to other metallic conductors, which can be a source of minute arcs that can cause spurious emissions. Jacketed lines are particularly suitable for direct burial. Burying the lines greatly reduces the daily and seasonal temperature extremes they encounter, yielding greater stability relative to above-ground installations.

Although jacketed lines have an indefinitely long life when buried, they are vulnerable to attack by rodents, which can gnaw through the vinyl jacket and have been known to consume portions of the copper outer conductor, as well. The possible need to replace buried lines in the future should be considered at the time of initial construction. Protecting the lines in PVC conduits and/or installing spare lines are desirable options. At the minimum, PVC sweep elbows should be in place before concrete floors or foundations are poured for buried transmission lines and sampling lines that exit the transmitter building or doghouses.

Characteristic Impedance

A characteristic impedance of 50 ohms is the industry standard for radio frequency transmission lines. Such lines can be expected to be more easily available than other impedances in future years, should a replacement line be needed. The velocity of propagation in transmission lines is less than the velocity in air. Therefore, the electrical length of any line is somewhat greater than its physical length. This increase must be considered in any phasor design. Typical values of transmission line velocity constants are listed below:

Obsolete solid-dielectric flexible lines	68%
Original foam-filled lines	79%
Low-density foam (LDF) lines	88%
Semi-flexible air dielectric lines	90–93%
Modern rigid copper lines	99.8%

Phase-Stabilized Line

In its first year, a new directional array will usually experience small drifts in indicated phase and ratio parameters that will not be repeated seasonally thereafter. This initial drift is caused by the minor mechanical stresses remaining in the feed lines and sampling lines following manufacture and installation; these stresses are slowly relieved with temperature cycling. The result is a small initial change in characteristics following installation that will not be repeated. "Phase-stabilized" lines, which will reduce the initial drift, are available. Such lines have been temperature-cycled in an oven to relieve residual mechanical stresses. Within about a year following installation, no observable differences in stability exist between regular lines and those that were initially phase-stabilized.

Transmission Line Fittings

All modern transmission lines can be equipped with EIA-standard flanged fittings to mate with connecting equipment. Although the impedance continuity provided by such fittings is essential at FM and TV frequencies, such fittings are unnecessary for foam line antenna feeds at AM broadcast frequencies. However, adequate electrical bonds are essential between the transmission line conductors and the adjoining equipment. The dielectric in foam-filled lines is a better insulator than air. However, catalogs show such lines as having the same voltage breakdown rating as air-dielectric lines because of the limited air gap within the end fittings. For AM broadcast use, the power rating of foam lines can be increased by avoiding the use of end fittings and stripping back the outer conductor (while leaving the foam) to increase the air gap on both ends of each line.

Line Mismatch

At AM broadcast frequencies, the effects of a transmission line mismatch are quite different from the electrically-long transmission lines used for FM and television antennas, where mismatches can cause reflections back into a transmitter and result in crosstalk or ghosts. At AM broadcast frequencies, transmission lines are rarely longer than a wavelength. A mismatched transmission line yields an input impedance that is dependent on both the load impedance and the length of the line. Thus a transmission line can act as a simple impedance transformer and can be designed to exhibit a desired input impedance through proper choice of load impedance. This characteristic is occasionally exploited in designing phasing systems.

Perfect transmission line matches are not critical to overall system bandwidth. Although most phasing and coupling systems are designed with the intent that the transmission lines to the towers be matched, the consequences of substantially mismatched lines are usually trivial. Improving the match on any transmission line between the phasor and tower may improve or impair the bandwidth. Because the lines are electrically short, their behavior is more akin to that of a simple network than to the transmission lines in FM and television systems, which are often hundreds or thousands of wavelengths long. As a general rule, if lines to towers that handle 25% or more of the transmitter power are matched to a VSWR of 1.5:1 or better (at the carrier frequency) and any remaining lines are matched within 2:1, further efforts at improving matches may give only trivial reductions in line losses and trivial changes (for better or worse) in the bandwidth. Even though a line may be perfectly matched at the carrier frequency, it will never be perfectly matched at the sideband frequencies because no antenna load is absolutely constant over the desired bandwidth of ±10 kHz. The effects of the less-than-perfect sideband matches can be minimized by proper proportioning of the phase shifts between the phasor networks and the ACU networks. It is important to avoid confusion between transmission line VSWR at the carrier frequency and common-point VSWR at sideband frequencies.

Voltage and Power Ratings

With 100% modulation, the voltage on a matched lossless transmission line varies from zero on a modulation trough to twice the unmodulated value on modulation crests. This is quite unlike the situation in FM where every radio frequency cycle has the same amplitude. In FM and TV systems, the dielectric losses (because of the higher frequencies) cause the power rating to be limited by the heat generated within the dielectric as well as by the copper losses; power ratings are accordingly based on the ability of the transmission line to dissipate the heat resulting from the losses. At AM frequencies the dielectric losses are minor while the instantaneous voltages during modulation peaks can be very high. As a result, the power handling ability of AM transmission lines is usually limited by the voltage breakdown rating of the line. With 125% positive modulation, the maximum instantaneous peak voltage becomes more than three times the unmodulated RMS voltage and the instantaneous peak power becomes more than ten times the unmodulated power. An additional allowance must also be made to accommodate the increased voltage maximum along each line due to standing waves caused by any less-than-perfect transmission line match.

Under transient conditions (such as can be caused by lightning strikes or sudden component failures) the VSWR overload protection designed within all modern transmitters may not trip, and the power distribution among the towers may be drastically altered in unpredictable ways. This can lead to voltage breakdowns within buried transmission lines which are difficult locate and repair. For this reason it is good engineering practice to employ equal-size transmission lines to all towers that are capable of handling the full modulated output of the transmitter even though the power delivered to particular towers may be only a few percent of the total power during normal operation. This design factor will also be greatly appreciated in the event of an emergency requiring nondirectional operation on one of the low-power towers.

ANTENNA MONITOR AND SAMPLING SYSTEM

A directional antenna sampling system consists of current sampling transformers or sampling loops on each tower to provide a sample of the tower current, transmission lines which return the samples, and an antenna monitor which measures the amplitude and phase of each sample relative to that of the reference tower.

Sampling Loops

A sampling loop consists of a rigid single-turn coil permanently attached to each tower. Because sampling system stability is essential, the loops are typically made of galvanized or stainless angle iron or rigid copper water pipe. The loops must be at least ten feet above ground and may be insulated from the tower and kept at ground potential. However, it is more common practice to use loops at tower potential and return the sample to ground potential through an isolation coil (isocoil) formed by coiling up the sampling coax so as to form a high impedance across the tower base insulator while not disturbing the current sample carried within the coax.

Sampling Transformers

An alternative method of obtaining tower current samples is to use shielded current sampling transformers in the feeds to each tower. These have much to recommend them if the tower heights do not exceed approximately 130 degrees. Although a transformer will sample both the current going up the tower and a lesser component that flows to ground through the base insulator capacitance, the sampling error is usually inconsequential. The advantages to using sampling transformers are stability, a sampling device that is protected inside the ACU instead of being exposed to the weather, and elimination of the isocoils that are otherwise usually required at each tower.

Stability and Accuracy

Sampling system stability is vital, but accuracy is not critical because the proper phases and currents to generate a directional antenna pattern are determined by means of field strength measurements. The sampling system is required to detect changes from the original adjustments of indicated phase and ratio. The accuracy of a sampling system in measuring the absolute amplitude of radiation from each tower is inherently poor. Even though the sampling loops may be carefully adjusted so that each has exactly the same area in its single-turn coil and all are mounted at the same height, the samples will not be a measure of the relative radiation from each tower. The mutual impedances between towers seriously distort the sinusoidal current distribution which would otherwise exist throughout the height of each tower. Thus the indicated antenna monitor ratios are not an absolute measure of the relative radiation from each tower. The monitor ratios may also differ substantially from the base current ratios, particularly when the loops are a considerable distance above ground level or towers of unequal height are employed. The indications of phase from a good monitoring system are considerably more accurate than the ratio indications. Arrays with unequal-height towers usually require sampling loops that are all located approximately a quarter-wavelength below each tower top if the indications of phase and ratio are to be the best approximation of the actual radiation from each tower.

DETUNING AND DECOUPLING SYSTEMS

Detuning (Floating) Unused Towers

The most common occasion for detuning towers arises within directional arrays themselves wherein unused towers must be placed in a non-radiating condition. This is necessary in all arrays when nondirectional proof-of-performance field strength measurements must be made, where nondirectional daytime operation is permitted, or where one or more towers are not used for daytime or nighttime modes when two patterns are authorized. In more critical situations or with taller towers, more effective detuning is required than simply disconnecting the unused towers to reduce

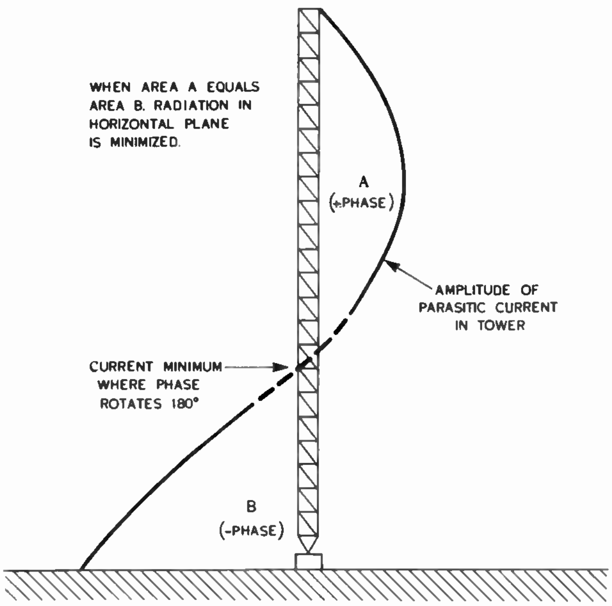


Figure 11. Current distribution on detuned tower.

the radiation (reradiation) caused by parasitic currents that flow in them due to the mutual coupling to adjacent driven towers. Disconnecting a tower will not eliminate the component of induced tower current flowing to ground through the capacitance of the tower base and base insulator. An inductive reactance of the proper value (often formed in part by the isocoil if sampling loops are employed) can be added across the base to resonate the base capacitance and effectively eliminate all current flowing to ground.

The optimum detuning (even for short towers, as well as for towers up to a wavelength tall) results when an inductive reactance across the base insulator produces a tower current distribution as shown in Fig. 11. Because there is a phase reversal at the elevation of minimum current, it is simply necessary to position

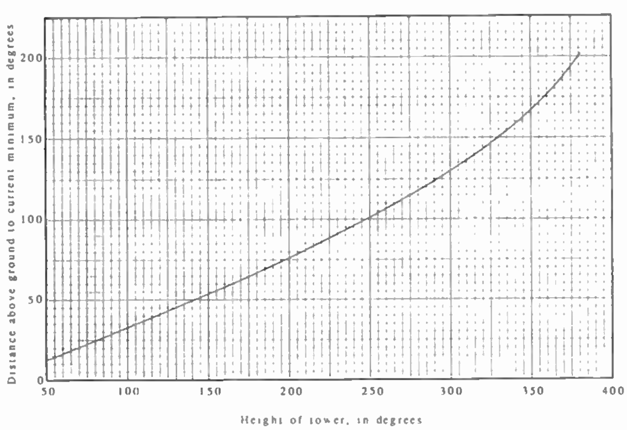


Figure 12. Approximate height of required current minimum to detune guyed towers.

a current minimum on the tower so that area A equals area B. In the horizontal plane, the radiation from the portion of the tower above the minimum will then cancel that from the portion below the minimum. The proper height for the tower current minimum (typically about one-third the tower height) is as shown on Fig. 12. The minimum can be achieved easily by climbing the tower with a field strength meter (or other current sampler and detector) and adjusting the reactance shunting the base insulator until a current null is obtained. Although the shunt reactance necessary to detune a tower is always inductive, the reactance may be so high as to not be within range of a coil of reasonable size. A vacuum variable capacitor in parallel with the coil then provides a convenient and precise way of adjusting the shunt reactance so as to float the tower.

Tower Sectionalization

Occasionally, when it is necessary to use a tower that is too tall, sectionalization of the tower is required. For example, it may be desired to share an FM tower with an AM station. If the overall tower height were to exceed about 225 degrees at the AM frequency, the tower would be a poor radiator, putting more AM signal up into the sky than along the ground. The preferred solution for such a tall tower is to include insulators in the tower legs (and an isocoupler in the FM feedline at the same height) so as to open-circuit the tower at the height needed to make it a good AM radiator. A parallel resonant circuit (an impedance pole) across the insulators can then resonate the insulator and stray capacitances to effectively decouple the upper portion of the tower. In cases where structural considerations prevent the sectionalizing of a tower by means of insulators, a workable solution can be to use detuning "skirts" to decouple the top portion of the tower. Examples of how skirts were used to sectionalize a tower at 1360 kHz and to detune it for protection of a nearby station on 600 kHz are shown in Fig. 13.

A tower skirt is an insulated wire cage constructed outside of a tower. The skirt is formed of wires that surround the tower and are spaced a foot or two away. All skirt wires are bonded to the tower at one end and can be connected in parallel at the other end so as to be tuned with a single capacitor. Electrically, the skirt can be considered as forming a shorted quarter-wave coaxial transmission line which is slightly foreshortened and loaded with a variable capacitor at its open end for easy adjustment. The tower forms the center conductor of the coax and the skirt wires approximate an outer conductor. Adjustment of the skirt tuning places a high impedance at the capacitor end of the skirt with a result that the tower is effectively open-circuited at that elevation. The effectiveness of such a tower skirt is determined by the characteristic impedance and losses in the transmission line section formed by the tower and the skirt. The most effective skirt will necessarily have a high Q, which is realized by using reasonably heavy skirt wires spaced as far as

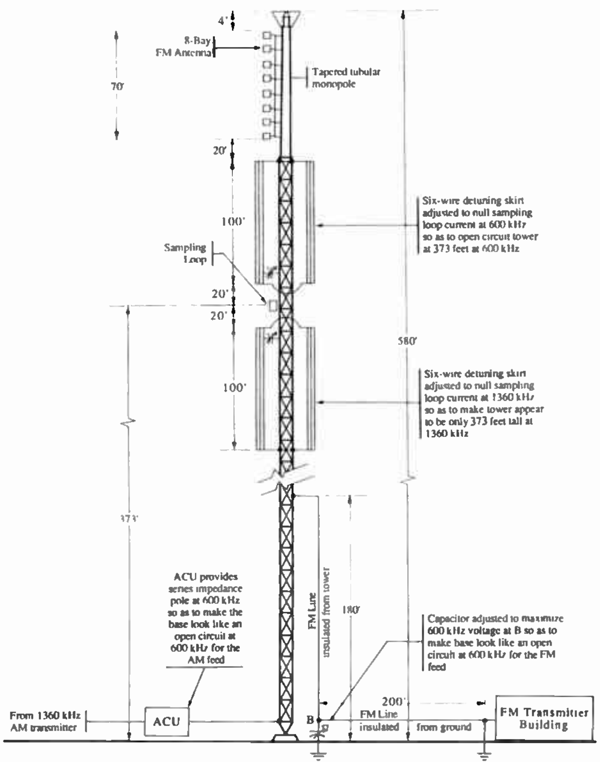


Figure 13. FM tower sectionalized to make an efficient 1360 kHz radiator simultaneously detuned at 600 kHz.

mechanically feasible away from the tower. For guyed triangular towers of uniform cross-section, combinations of three or six skirt wires are common. Inasmuch as the most effective skirt will necessarily have a high Q, it will best decouple the portion of the tower above its open end only at the frequency to which it is tuned. As a result, the tower base may exhibit a steep impedance versus frequency characteristic which can impair bandwidth. Another way of visualizing the operation of a skirt is to think of the tower and skirt wires as forming an inductive loop which is tuned to resonance so that the current flowing up the tower is effectively cancelled by a substantially equal and opposite current flowing down the skirt wires.

Detuning Power Lines

The detuning techniques described above can be applied to steel power line towers or to other tall metallic structures in order to reduce the radio-frequency currents in them, currents that would otherwise cause reradiation and distort the desired directional antenna pattern of a nearby radio station. Fig. 14 shows such an arrangement with each skirt wire on each tower leg separately tuned. Again, the effectiveness of the skirt depends on the length of the skirt section and the number, diameter, and spacing of the skirt wires.

Filter Systems

If it is necessary to control the current distribution on a tower at a different frequency in order to avoid disturbing the pattern of a nearby directional station, a filter system is needed to permit control of the impedance shunting the tower base at the other station's frequency. The control should be independent of routine phasor adjustments. Such an arrangement is shown in Fig. 15 for each tower of a station operating on 630 kHz. The 630 kHz station must control tower current distribution at 950 kHz in order to protect the directional antenna operation of the nearby station on that frequency.

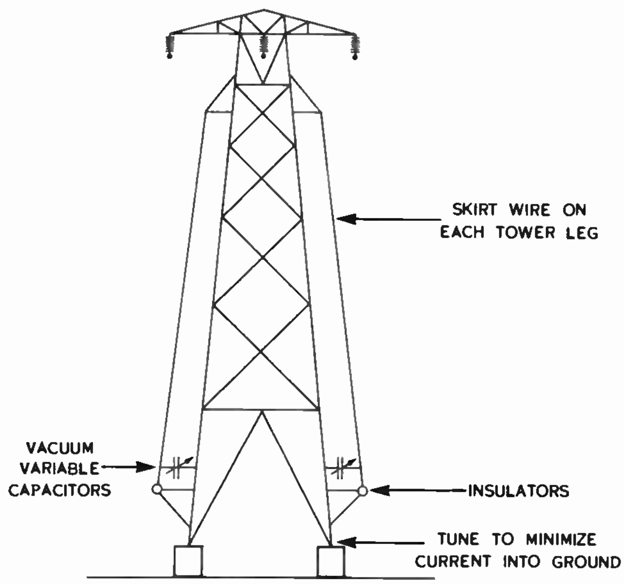


Figure 14. Skirts used to detune power-line towers.

For filters in series with antenna feeds (such as the C3/L7, C4/L8 combination), it is desirable to provide an impedance zero at the pass frequency as well as the impedance pole at the reject frequency. Even when the pass and reject frequencies are widely separated, failure to provide the impedance zero with a reasonably low L/C ratio can result in excessive losses.

Intermodulation

Filters to control spurious emissions resulting from intermodulation products generated within the final stage of a transmitter can have the same configuration

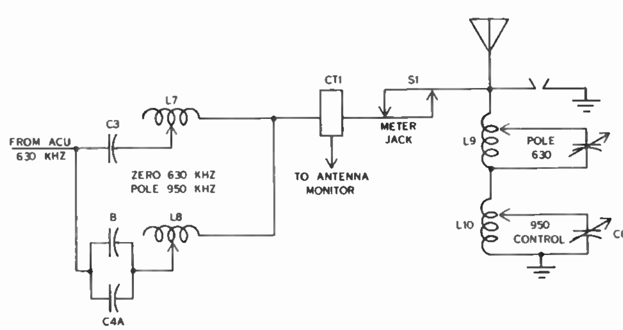


Figure 15. Independent control of tower current distribution at a second frequency.

as filters used to control tower-current distribution for a nearby station. Before such filters are designed, a knowledge of the frequencies that generate the intermod is essential. The second, third, and higher harmonic frequencies are strong components in a transmitter output stage and can generate intermod products there even though these harmonics are suppressed to acceptably low levels when measured at the transmitter output. In the general case, an unwanted incoming signal from some other station can be coupled into the final amplifier through the antenna system to mix with the fundamental or any harmonic to produce a spurious output frequency.

Effective filters can take the form of either traps to prevent an unwanted signal from getting into the transmitter or traps to prevent the resulting spurious emission from getting out (or both). Filters can be series-resonant shunt elements that short the unwanted incoming signal (or the spurious emission) to ground, or they can be parallel-resonant elements in series with the transmitter output to present a high impedance to either the unwanted incoming signal or the spurious emission.

One clever and expeditious arrangement, when a single spurious frequency must be eliminated, may be to modify the shunt arm of the common-point matching network to consist of a series-resonant L/C circuit that simultaneously has a very low impedance at the undesired frequency and the proper reactance at the station carrier frequency.

Many modern transmitters have exceedingly broadband output coupling networks which, together with improved antenna bandwidths, permit unwanted signals from other radio stations to get back into the final stage of the transmitter to produce spurious intermod products. Signals thus entering the transmitter need not be particularly strong. Many cases are known where the offending input that caused an objectionable spur originated from a radio station 20, 30, or even 40 miles away. The 500 kHz radio channel is an international distress frequency, so this frequency and channels adjacent to it should be checked for spurs whenever measurements of spurious radiation are undertaken on an AM broadcast transmitter.

Broadbanding

The term "broadbanding" can be quite confusing. Frequently it is applied to any circuit modifications that are intended to improve the impedance bandwidth of an antenna system. A "broadbander" is a circuit that in and of itself improves the impedance bandwidth, thus reducing the VSWR at the sidebands. A broadbander typically consists of the series combination of a coil and a capacitor which is inserted in series with the load to be improved (usually immediately before or after the common-point) and is adjusted to series resonance at the carrier frequency. Such a circuit is capacitive below the carrier frequency and inductive above the carrier frequency. Before a broadbander can function as intended, the load it sees must be inductive at the lower sideband frequencies and capacitive at the upper sideband frequencies. This condition is achieved by altering the phase shift in the common-point matching network (or inserting an additional phase-shifting network) so that the Smith Chart plot is "horns up", as is shown in Curve A, Fig. 16. Then the uncorrected load reactance (which is inductive for the lower sideband) can be canceled by the capacitive reactance of the broadbander and the capacitive reactance of the upper sideband load can be cancelled by the inductive reactance of the broadbander. The result is that the load reactance can be effectively canceled at any pair of sideband frequencies as determined by L/C ratio of the broadbander. For such a broadbander to function as intended, it is essential that it be effectively shielded so that there is no stray coupling between it and the other components in the phasor. When properly designed and adjusted, it can "wrap up" the common-point impedance, as shown in Curve B, Fig. 16.

Such a broadbanding circuit should not viewed as a cure-all, but it can provide a useful reduction in sideband VSWR once all other reasonable efforts at improvement have been applied.

TRANSMITTER LOAD OPTIMIZATION

A transmitter can deliver equal power to both sidebands when the upper and lower sideband load resistances are equal and the upper and lower sideband load reactances are equal. If the transmitter load VSWR at ±10 kHz from carrier is rather good (perhaps 1.2:1 or better), further adjustments to improve load symmetry may not be warranted. However, if the ±10 kHz sideband loads are rather poor (e.g., 1.5:1 or greater), sideband load symmetry can become an important matter. Symmetry is achieved by adjusting the phase shift in the common-point matching network so that the load resistances at the ±10 kHz sideband frequencies are essentially equal and the reactances are equal but opposite in sign as measured at the anode (or collectors) of the transmitter output stage.

Because different manufacturers use different output circuitry with different phase shifts, and because some have not published information as to optimum load orientation at the transmitter output terminal, the necessary measurements usually must be made on the load seen by the final amplifier stage anode or collector. Fig. 16 shows the Smith Chart orientation for the two conditions that satisfy symmetrical loading requirements. For tube-type transmitters (whose output stage functions substantially like a constant current generator), Curve C, which shows equal sideband resistances higher than at carrier, tends to emphasize the high frequency response. Curve A also shows equal sideband resistances, but lower than at carrier; this orientation tends to show reduced 10 kHz distortion. Either of the two conditions is preferable to intermediate conditions that yield asymmetrical load resistances at the two 10 kHz sideband frequencies.

Transmitter manufacturers are now more sensitive to the fact that all antenna systems provide less-than-

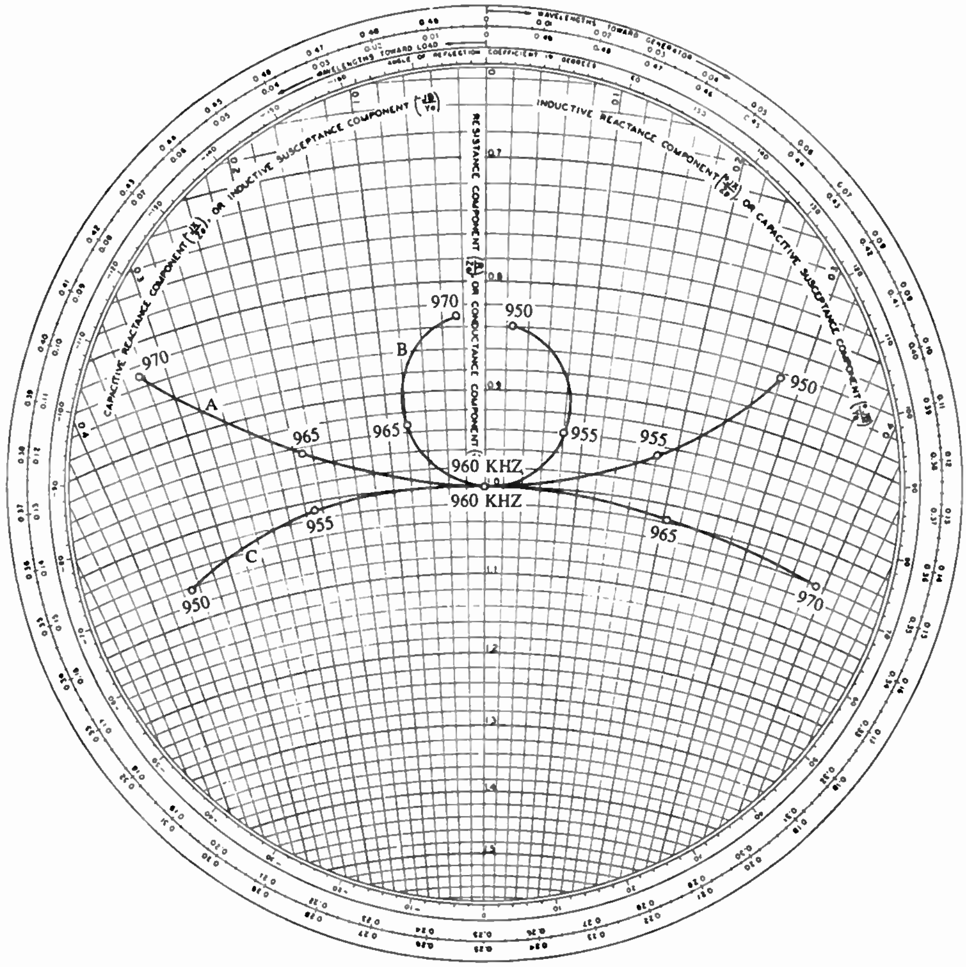


Figure 16. Proper load impedance at final amplifier plate or collector.

perfect loads at the sideband frequencies and that load orientation (as displayed on a Smith Chart) can be critical to overall system performance. Most manufactures can now supply information as to optimum load orientation (referred to the transmitter output terminals) for each specific transmitter. Experience has also shown that audio intermodulation distortion varies with load orientation, so intermod should be measured and minimized.

DIPLEXERS AND MULTIPLEXERS

Two, three, or more AM stations can be combined on a single tower, provided that sufficient frequency separation exists between the channels and the electrical height of the tower is not unreasonably short or tall at any of the channels. Depending upon the transmitter characteristics, only 35 dB or less of RF isolation is typically required between the transmitters to avoid generation of spurious emissions and to permit completely independent operation. Diplexer adequacy is assured when each station:

1. Has no objectionable spurious emissions.
2. Can conduct an acceptable audio proof, including distortion and noise measurements, while the other station is in operation.
3. Has no observable change in its indicated antenna current when the other station's transmitter is turned on and off.

The losses in diplexing and multiplexing filters can become crucial to a successful design. Since the tower impedance and required tower current at each frequency are known, the losses can be calculated with good accuracy. Capacitor losses are negligible. Coil losses can be calculated using an assumed coil Q of 300, which is a conservative approximation for typical silver-plated RF inductors in shielded enclosures. Diplexer losses can be minimized if the reactive component of base impedance can be eliminated by means of networks at the tower base. Then the shunt elements in the diplexing filters operate at lower voltage levels.

Two stations having directional antenna patterns can be diplexed on common towers provided the tower spacing and orientation permit acceptable directional patterns for both stations and the compatibility requirements previously mentioned for combining nondirectional stations are met.

Ground Systems

This material duplicates in part the information on ground systems in the previous chapters, information that is so important it bears repeating.

The current that flows into a tower does not simply disappear: it all flows back to ground through the capacitance that exists between each incremental length of the tower and the ground. (This makes the tower and ground system an integral part of the coupling and phasing system.) When this current reaches the ground, it must be collected and returned to the ground side of the ACU network to complete the circuit. The ground currents are usually strongest close to the base of the tower where the capacitance between the tower and incremental ground segments is greatest. If the ground resistance is substantial, the ground current (equal to the antenna current) will simply heat the ground and represents a power loss. To return the antenna current from the ground surrounding the tower to the ground terminal of the ACU network, a system of radial ground wires is employed. For nondirectional towers, a series of early-day experiments established that 120 radial ground wires uniformly distributed at three-degree intervals and extending a quarter-wavelength (90 electrical degrees) from the tower base was a reasonable limit beyond which the reduced losses to be realized from incremental increases in the number or length of the radials did not warrant the incremental increases in cost required for a more extensive ground system. As a result, ground systems consisting of 120 radials extending a quarter-wavelength from each tower (or foreshortened and bonded to a transverse copper strap where radials from adjacent towers overlap) have become the *de facto* industry standard, even though it is recognized that in directional arrays the ground currents do not always flow along radial lines, due to the effects of adjacent towers.

Near the base of each radio tower, there are both electromagnetic and electrostatic fields. With reasonably low base impedances (such as a quarter-wave tower), the electromagnetic field predominates. Near towers with a high base impedance (such as a half-wave tower), the electrostatic field predominates. The radial wire ground system minimizes the I^2R losses but not the E^2/R electrostatic losses which can occur in all materials near the tower base with less-than-perfect dielectric constants. The electrostatic losses are usually inconsequential in a nondirectional operation, but can become critical in a directional array. Dielectric materials in the vicinity of the tower base can include wooden fences, wooden or cement block tuning houses, and the earth that is on top of a buried ground system. The changing dielectric constant in these materials between wet and dry conditions can make enough difference in base impedance to cause substantial changes in base current ratios and phases. For these reasons the ground wires should be kept on top of the soil in the immediate vicinity of each tower base or covered with only a thin layer of gravel or paving. Beyond about 20 feet from each tower base the fields are less intense and the remaining portions of each ground radial are usually buried six to eight inches for mechanical protection. Deeper burial is called for only if the land is to be cultivated or is so soft that farm animals or machinery can sink into the soil and break the ground wires. Metal fences surrounding each tower base are quite superior to wooden fences, whose dielectric constant can change greatly when wet. If nonmetallic doghouses are employed in critical arrays, it may be necessary to provide an electrostatic shield of expanded copper mesh or Copperweld wire mesh up the side of the dog house adjacent to the tower and across its roof. The ultimate electrostatic shield takes the form of an elevated radial ground system which typically extends to a radius of 10 to 30 feet at a height of six feet or more above ground and effectively shields everything beneath it.

Number 10 AWG soft-drawn copper wire is the most common material for ground radials. Copperweld mesh ground screens, such as are employed for lighting protection under electric power substations, are suitable for the area immediately surrounding each tower, and the radial wires can be connected to the edge of such a screen. Frequently an additional 120 radials, each 50 feet long, are interspersed between the longer radials where a ground screen is not used. All connections within a ground system must be silver-soldered or brazed, as soft tin/lead solder deteriorates rapidly when buried.

OVERALL SYSTEM PERFORMANCE

Directional antenna pattern design and phasor design go hand-in-hand. Given a poor pattern design, no amount of clever circuit design in the phasor can provide an excellent system. Following are some considerations that can lead to an excellent antenna system.

In the Pattern Design

1. Avoid closely spaced towers; 90 electrical degrees or more should be the objective.
2. Avoid broadside minimums without ample spacing between towers; otherwise the driving point resistances can become very low.
3. Check to see that the individual tower radiation vectors add in-phase (or very nearly in-phase) in the pattern maximum; otherwise excessive circulating currents, poor bandwidth, and pattern instability can result.
4. Unless other considerations dictate, choose tower heights in the range of 100 to 130 degrees. Shorter towers exhibit lower driving point resistances and poorer bandwidth. Taller towers may require excessively large impedance transformations in the

ACU networks; this condition can add another bandwidth limitation.

In the Phasor Design

1. Minimize reactive power (I^2X) in the phasor components. This will avoid high Q circuits.
2. Avoid excessive use of series (trim) coils to adjust the reactance of fixed capacitors. Such coil/capacitor combinations necessarily have a steeper impedance versus frequency characteristic than a capacitor alone having the proper reactance.
3. Consider alternative phasor designs. The design with the fewest parts usually has the best bandwidth unless specific broadbanding circuits are incorporated.
4. Derate coil current ratings by 40% and mica capacitor current and voltage ratings by 50%. Remember that RF currents under heavy asymmetrical modulation can exceed 125% of the unmodulated current and that peak voltages can exceed 300% of the unmodulated RMS voltages.

A SIMPLE BANDWIDTH TEST

A simple test can evaluate the performance of the individual towers in an existing directional antenna system. The object is to determine the amplitude of each 10 kHz sideband component as radiated by each tower. This test is possible with any common field strength meter because the selectivity of such meters is just sufficient to resolve sideband components that are 10 kHz removed from the carrier. The procedure is as follows:

1. Keep the sampling lines terminated into the antenna monitor, but add "T" connectors to bridge off samples into a field strength meter operating as a linear tuned voltmeter. Be certain the samples do not exceed the safe voltage input for the field strength meter.
2. Modulate the transmitter 50% with 10 kHz sinewaves.
3. For each tower sample, first adjust the field strength meter gain so as to set the carrier level full scale (100%). Then tune in each sideband in turn and log each sideband amplitude as a percent of full scale. In a perfect system, this would result in sideband amplitudes equal to 25% of the carrier level in each tower.
4. Repeat the process for each tower.

In the best antenna systems, all 10 kHz sideband components will range between 23% and 27% with 50% modulation at 10 kHz. Poor systems have been observed to have individual tower sideband components as low as 5% or as high as 40% in this test.

In the main lobe of a directional antenna pattern, deficient 10 kHz sideband amplitudes do not usually result in measurable distortion because the radiation components from the individual towers add approximately in phase and the sideband deficiencies of any one tower are masked by the total radiation from all of the other towers. This test is useful in pinpointing limitations in existing systems.

SAMPLING SOURCE TO FEED A MODULATION MONITOR

The indications of a modulation monitor may be inaccurate at high modulation frequencies if the modulation sample is taken from a less-than-perfect antenna system. In an antenna system with a high VSWR at the 10 kHz sidebands, the voltage sampled for the modulation monitor may be higher or lower at the 10 kHz sidebands than at low modulating frequencies. The best measure of 100% modulation is at a remote point in the main lobe of the antenna pattern. Various sampling points within the transmitter or phasor can be tried until a sample source is found that matches the percentage modulation at high modulating frequencies observed in the far field of the main lobe.

CONCLUSION

When the complexities in the design of directional antenna patterns and circuitry are considered together, the need for specialists in this field is evident. However, these chapters should have increased the reader's understanding of the complexities and thus assisted in the maintenance and improvement of AM directional antenna systems.

2.5 AM Broadcast Antenna Systems

Part III: Maintenance

Edward Edison, P.E.
Hammett & Edison, Inc., San Francisco, California

Many stable directional arrays, being electrically passive devices, operate for ten or more years without any readjustment. However, maintenance attention will be needed to overcome the normal effects of age and deterioration (routine maintenance), to restore proper operation after catastrophic failures (such as caused by lightning), and to cope with new sources of reradiation that may develop within the environment in which the antenna system operates. Infrequently, readjustment and relicensing with new pattern parameters may be required.

ROUTINE MAINTENANCE

Routine mechanical and electrical maintenance is required to offset the effects of age and deterioration. Mechanical maintenance includes the attention necessary to keep the moving parts functioning and the tower painted and in good repair. Electrical maintenance may involve replacement or recalibration of base current meters; checks on the continuing proper performance of the phasor components, transmission lines, and ground system; and infrequent readjustment of the antenna phase and ratio parameters as necessary to ensure operation within licensed limits with the proper directional antenna pattern.

Mechanical Maintenance

The moving components within an antenna system may include RF relays, meter switches, dial drives, and the rollers on variable coils.

The contacts on RF relays that carry substantial currents may eventually wear out and require replacement. It is vital that such relays be interlocked so they cannot be switched while transmitter power is applied. Relay maintenance problems are reduced when vacuum relays are employed.

Meter switches, that disconnect thermocouple-type ammeters from the circuit when not being read, are subject to considerable mechanical wear and tear. Occasional lubrication and tightening of the components are necessary if the parts are to function properly. Replacement of thermocouple ammeters with modern transformer-type meters eliminates the need for RF meter switches and their associated maintenance requirements.

Dial drives on variable coils and capacitors, once properly adjusted, require little attention except infrequent lubrication. The thrust bearings on vacuum variable capacitors are factory lubricated and should not require maintenance attention. However, the rollers on variable coils are often a source of mechanical difficulties (particularly in older coils) and periodic maintenance is required if they are to operate smoothly and without excessive arcing. Networks that are designed to utilize variable vacuum capacitors instead of variable coils are superior in this regard.

Electrical Maintenance

Thermocouple RF ammeters that measure base currents and common-point currents are a frequent problem due to failure or changes in calibration. If such meters cannot be replaced with modern transformer-type units, an intermediate improvement can be effected by substituting make-before-break meter jacks at the base of each tower (in lieu of the original permanent thermocouple meters and associated protective switches). The meters can then be stored in a dry environment and taken to the towers only when base current readings are required. The effects of calibration errors are reduced if the same meter is used to measure all base currents (provided all currents fall within an acceptable range on the meter scale).

Thermocouple ammeters can be calibrated with 60 Hz current. A useful procedure is to remove all such meters from the system and to connect them in series for testing. A high-quality dynamometer or soft-iron-type meter should also be included as a calibration standard. A filament transformer having adequate secondary current rating can be used to drive current through all of the meters connected in series. Current amplitude is easily controlled by supplying the filament transformer from an adjustable AC supply such as a Variac.

Modern current-transformer RF ammeters cannot be calibrated with 60 Hz and, if defective, must be returned to the manufacturer for repair and recalibration. The transformer, meter, and interconnecting cable are calibrated as a unit. The cable length should not be changed without recalibration by the manufacturer.

Occasionally loose connections or deteriorating phasor components will produce abnormal heating. Failures from these causes can be anticipated by simply feeling all RF connections and components for warmth immediately after sign-off.

Transmission Lines

Air or foam dielectric lines in continuous lengths rarely develop troubles, but pressure should be maintained on air dielectric lines in order to prevent the accumulation of moisture. Breaks in the outer conductor of foam-filled lines that are buried may cause obscure symptoms and prove difficult to locate. In extreme situations, a time-domain reflectometer is the best tool to pinpoint the problem.

Ground System

Significant deterioration of a ground system can manifest itself in two ways. The first indications may be unusual changes in the phase and ratio parameters between wet and dry conditions. These changes depend upon the nature of the directional array and the character of the ground. When ground systems are installed in shallow tidewater or saltwater marshes, the inherent ground conductivity is so good that even drastic changes of the ground system may produce negligible effects.

A second, more obvious symptom of a deteriorating ground system is a reduction in antenna radiation and a reduction in base currents. It is not uncommon in older arrays to find that the absolute value of each base current has fallen off since the time of the initial installation, even though the base current ratios remain correct. With the available common-point power remaining constant, any increase in ground system losses must necessarily reduce the base currents.

It is essential that all joints in a ground system be made by brazing or silver soldering. Soft solder consisting only of lead and tin deteriorates quickly when buried.

The existence and continuity of buried ground radials is best checked by using an underground cable locator of the type used by utility companies to locate buried cables and pipes. These devices consist of a transmitter, which can be located at the tower base to place a modulated low-frequency signal on all connected ground wires and a receiver that detects the unique modulation. While in use, the receiver can be carried in a circle with a radius of 100 feet or more about each tower, so that the radial wires that are intact can easily be detected and counted.

Alternatively, if a cable locator is not available, a shielded pick-up loop held near the ground and connected to the external input of a field strength meter may suffice to detect the current in individual radial ground wires. Even when a radial is broken, the current beyond the break tends to concentrate in the wire instead of the adjacent earth. A most definitive check is to uncover the distant ends of adjacent radials and check for continuity between them with a DC ohmmeter.

Towers

Towers require occasional attention to ensure their structural integrity. Periodic inspections by an experienced tower contractor are desirable. Guyed towers should be checked to confirm that they are straight and plumb and that the guy tension is correct. Towers with tubular legs can rust from the inside out; a close inspection is necessary to detect this condition while repairs are still feasible. Towers with solid legs avoid the hazards of internal rust. Galvanized towers must be painted only if required for aeronautical safety. Nongalvanized towers should be avoided, not only because rusting makes the maintenance problem more severe but also because of their inferior electrical conductivity.

Lightning strikes and target shooters may destroy guy-wire insulators. Such failures can usually be discovered with the aid of field glasses or a telescope. The electrical field gradient can be very high adjacent to a tower, particularly near the top. Arc-overs and corona can be avoided by cascading guy-wire insulators (johnny balls) immediately adjacent to the tower. Under severe conditions, fiberglass rod insulators or dielectric guy cables are a preferred solution. In an existing system with marginal insulation, any dirt or salt deposits on the insulators may cause corona and arc-overs. Washing the insulators and coating them with silicone grease may provide a temporary remedy.

SUDDEN COMPONENT FAILURES

Lightning strikes are the most common cause of component failure in a well-designed antenna system. Lightning which strikes a tower should be conducted directly to ground through the ball gap that is typically part of each tower installation. Detailed information on lightning protection is provided in Chapter 2.2, "Lightning Protection for Broadcast Facilities," of this *Handbook*. However, lightning has been known to destroy components in what appear to be well-protected systems.

Damaged components can most usually be located by a careful visual inspection. Lightning currents flowing through adjacent coil turns tend to collapse them; therefore, coils should be inspected to confirm that all turns are in their original uniform alignment. The forces generated by lightning traveling through the coil strap and clip on a ribbon coil may dislodge the clip and leave the strap hanging in nearly its original position. Careful visual inspection is needed to locate dislodged coil clips.

Although vacuum capacitors are usually self-healing if subjected to minor over-voltages that cause internal arcing, the extremely high currents from a lightning strike can melt the plates together and short the capacitor, necessitating replacement.

If lightning damage is a recurring problem that warrants additional protection, the following steps are suggested:

1. Check the ball gap at the tower base. The balls should be arranged side-by-side so that dirt and water will not bridge the gap with the first rain of the season. The balls should be on the same radial line through the tower axis because towers tend to twist in high winds; this twisting could otherwise change the gap between the balls. The gap should be as narrow as possible.
2. Provide a two-or-three turn choke coil in the RF feed from the antenna coupling unit (ACU) to the tower if one is not already installed. The coil can be from 5 to 10 inches in diameter depending on the station power level and the feedline size. Its function is to present a high impedance to lightning strikes so they will jump the ball gap rather than enter the ACU. Readjustment of the output arm of the matching network will be required to offset the added inductance of the lightning choke.
3. Provide a horn-type lightning gap within the ACU similar to the design shown in Fig. 1. The relatively clean, dry, and insect-free environment inside the ACU permits a tighter gap setting and increased lightning protection than can be accomplished with an exterior gap. The path for lightning currents from the tower to the gap within the ACU should be very direct with minimum inductance and no sharp bends. Similar lightning gaps also can be installed at both ends of all transmission lines to help prevent lightning currents from getting into the phasor.
4. Adjust each gap individually by reducing the spacing until it arcs on the heaviest modulation peaks. The spacing should then be doubled for the permanent adjustment. This will be a much smaller gap spacing than is usually encountered on external gaps at the tower base.
5. Check to see that there is a DC path from each tower to ground, either through a phase sampling isolation coil, a separate static drain coil or resistor,

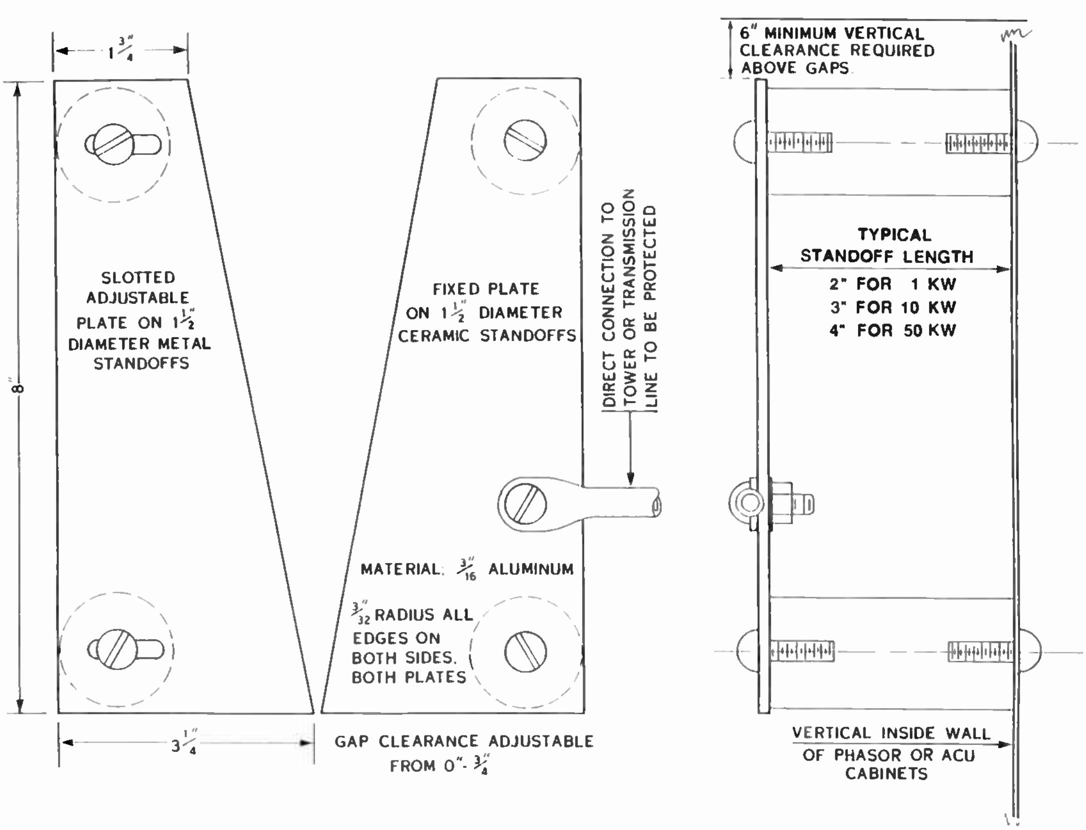


Figure 1. Suggested lightning gap design.

a spare winding on a tower lightning choke, or the network circuitry within the ACU itself. Without a static drain, static charges may build up on a tower and then discharge with a small arc that may be sufficient to trip protective circuits and take the transmitter off the air.

ENVIRONMENTAL CHANGES

The ideal environment for a directional antenna system is a flat plane of high conductivity with no tall structures in the general vicinity, such as buildings, power line towers, water tanks, or oil wells. Such structures (or any metallic objects of substantial vertical dimension) will have currents induced in them. These currents are dependent upon the size and shape of the structure, the frequency, and the ambient field strength with which they are illuminated. The currents in such objects make them parasitic antennas. The resulting reradiation can severely distort the pattern of a directional antenna array.

Reradiating objects located in the main lobe of an antenna pattern are of greater consequence than those located in or near pattern minimums where their illumination is substantially less. The reradiation from reasonably slim and electrically short objects (where the overall height is substantially less than a quarter wavelength) can be calculated with useful accuracy by means of the approximate formula shown in Fig. 2.

If feasible, a more accurate assessment of the reradiation can be obtained by measuring the RF current flowing in the reradiating object. If the object is reasonably thin (such as a guyed communications tower, a cable television headend tower, or the four individual legs of self-supporting power line towers) measurement of the current flowing from the tower to ground is not difficult. For this measurement, a toroidal pickup coil formed by the spiral steel wire in a length of clothes-dryer exhaust hose (or even better, some types of vacuum-cleaner hose) is useful. The hose section should be long enough to encircle the tower or current-carrying conductor to be measured and to have its ends attached to a field strength meter operating as

$$\left|E_r\right| = \frac{\left|E_i\right| \left[\dfrac{\cos(1.2\,hf) - 1}{\cos(1.2\,hf)}\right] h}{2000\left[\left(\ln \dfrac{2h}{a}\right) - 1\right]}$$

Where: E_r = Reradiated field in mV/m at 1 km
E_i = Incident field in mV/m at location of reradiator
h = height of reradiator in meters
a = effective radius of reradiator in meters
f = frequency in MHz
ln = natural logarithm
cos = cosine of angle in degrees

Figure 2. Formula for estimating reradiation from electrically short reradiators.

a linear voltmeter. Such a pickup coil can be calibrated by encircling a conductor carrying a known current, such as the feed, to the base of a radio tower.

It is not necessary to measure the current distribution throughout the height of electrically short objects. The current distribution can be assumed to be triangular, decreasing to zero at the top of the tower if it is freestanding; or the current can be assumed to be uniform throughout the entire height of the tower as a worst-case assumption where powerline towers exhibit heavy toploading due to the capacitance to ground of the conductors they support. Once the current is measured, the reradiation can be calculated by relating the "ampere-feet" of current in the reradiator to that in a quarter-wave tower at the same frequency. For example, a quarter-wave tower at 1000 kHz is 246 feet high, has a base resistance of approximately 50 ohms, and a base current of about 4.5 amperes for one kilowatt. With sinusoidal current distribution on the tower, the average current throughout its height would be 64% of the base current or about 2.9 amperes. Thus, at 1000 kHz about 700 ampere-feet of current would produce about 190 millivolts per meter of radiation at one mile or 300 millivolts per meter at one kilometer. This corresponds to radiation (or reradiation) in the order of 0.5 millivolt per meter at one kilometer for each ampere-foot of current in the reradiator. This rule-of-thumb as to the resulting reradiation from short towers is proportional to frequency when applied to channels removed from 1000 kHz.

The techniques for detuning reradiating objects are the same as those described in the previous chapter for detuning unused towers within a two-pattern directional array.

RADIATION HAZARDS

In 1986 the FCC promulgated rules limiting human exposure to radio frequency fields at all AM, FM, and TV broadcast sites. This subject is covered in detail in Chapter 2.9 of this volume. At AM broadcast frequencies, the potential for human exposure in excess of the radiation limits exists only in the immediate vicinity of radio tower bases and in the vicinity of phasors and antenna coupling units that are not in fully-enclosed cabinets. Phasors and ACUs employing open panel construction can exceed exposure limits at distances up to about five feet depending upon the station power level. Metallic fences surrounding each tower base form excellent shields at AM broadcast frequencies and greatly reduce the area exceeding exposure limits when compared to unprotected towers or towers within wooden fences. One neat arrangement, when weatherproof ACUs are employed, is to mount them so that their locked front door is in the plane of the fence surrounding the tower and is accessible without entering the fenced enclosure.

Compliance with radiation hazard limits effectively prevents maintenance or painting of a tower while it is in use. Therefore it is now highly desirable to provide for nondirectional operation on either of two towers so that every tower can be de-energized in turn when required for maintenance without taking the station completely off the air.

READJUSTING AN ARRAY

Any readjustment of a directional antenna array should be undertaken only by personnel who have a basic understanding of directional antenna theory and practice. At the minimum, the material in the preceding two chapters should be read and understood. Before any array adjustments are attempted, considerable information needs to be collected and analyzed.

Four conditions for proper array performance are explained below. All of these conditions should be met simultaneously in normal operation.

1. The antenna monitor indications should agree with the station license within the FCC tolerances of five percent for tower current ratios and three degrees for phase. A special case is a "critical" array which has been assigned tighter tolerances by the FCC.
2. Changes in base current ratios should be consistent with changes in antenna monitor ratios. Although the base current ratios and antenna monitor ratios are rarely in exact agreement (and will differ greatly in arrays with unequal height towers), these ratios should change by essentially the same percentage for small changes in the actual base current ratios. In other words, when the monitor indications match those shown in the license or the most recent proof, the base current ratios should also match what is shown in the same license or proof. Discrepancies are an indication of changes in the sampling system or in the base current metering and should be investigated.
3. The reradiation environment must be essentially unchanged from that in which the array was last in proper adjustment. Any substantial structures such as powerline towers or steel-frame buildings constructed recently in the vicinity of the array, particularly in the main lobe of the pattern, can cause trouble. Grounded structures approaching a quarter-wave high can distort critical antenna patterns even if located one or two miles away in the main lobe of radiation.
4. The monitoring points should be within the limits shown on the station license. A high monitoring point may result from a changed environment due to the construction of buildings or powerlines near the monitoring point. If a high monitoring point does not correlate with observed changes in antenna monitor parameters, the point is suspect and should be checked by measuring and analyzing ten or more of the other points on the same radial that had been measured and recorded in the most recent proof or partial proof. Readjustment to reduce excessive radiation is indicated only if the analysis of ten or more points confirms that the radiation in that direction is in fact in excess of the standard pattern limit.

Soil Conductivity

The analysis of a radial may be misleading due to changes in soil conductivity. Increased soil conductivity during wet seasons may cause substantial increases in signal strength at distant measuring points when compared to the same measurements during dry seasons. This phenomenon varies with distance, with season, and with differing geographical regions. These effects may not exist in some regions, but in extreme cases may show 2:1 changes in field strength at a distance of 20 miles. When analyzing field measurements, soil conductivity variations can be eliminated by comparing the present ratio of directional to nondirectional field strength at each measurement point to the same ratio in previous complete or partial proofs.

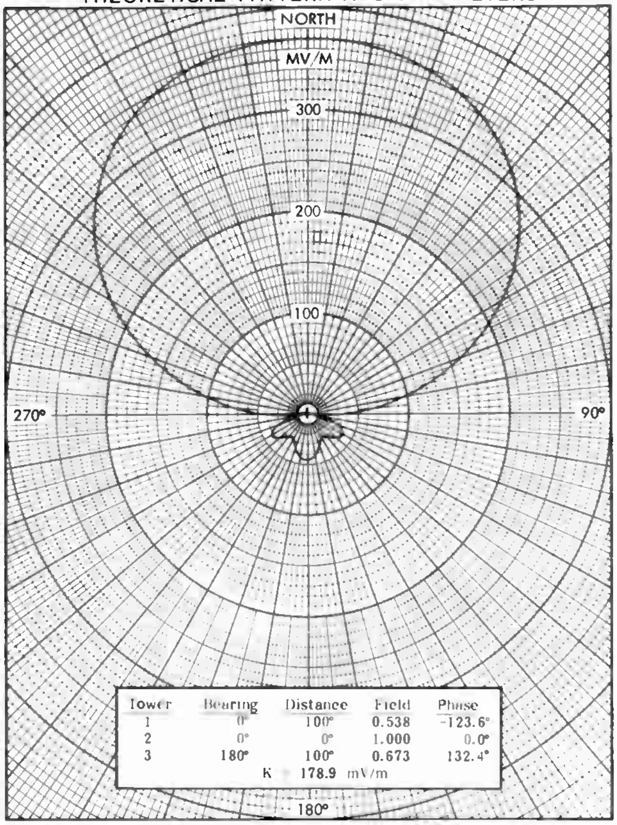
Because new nondirectional measurements may be needed at a future date, it is important to incorporate in the phasing system a switching capability to permit reversion to nondirectional operation at those stations which do not normally have a licensed nondirectional mode.

READJUSTMENT DATA AND ANALYSIS

Directional antenna arrays vary so much in pattern shape, suppression requirements, and phasor circuitry that precise readjustment instructions cannot be described. If possible, the consulting engineer who adjusted the array before its last proof of performance should be consulted on the problem. The following guidelines should prove helpful:

1. Carefully log all dial settings and all coil tap locations before making any changes.
2. Keep in mind what parameter you are trying to change before twisting any knobs or changing any coil taps.
3. Make changes only in small increments, typically never more than about one degree or one percent in phase or ratio.
4. Keep a step-by-step record of each change as it is made so that the array can be restored to intermediate or initial conditions, if necessary.
5. Remember that even small adjustments of phase or ratio can have a measurable effect on the common-point resistance. The power fed to the antenna system is only correct if the common-point resistance and current are essentially correct. A permanently installed common-point operating impedance bridge is a convenience not only for keeping track of the common-point impedance during adjustments but also for continuing confirmation of the licensed common-point value in day-to-day operation.
6. Construct vector diagrams for each important direction. Samples of such diagrams are shown in Fig. 3. The phase relationships must include both the electrical phase of the currents fed to the

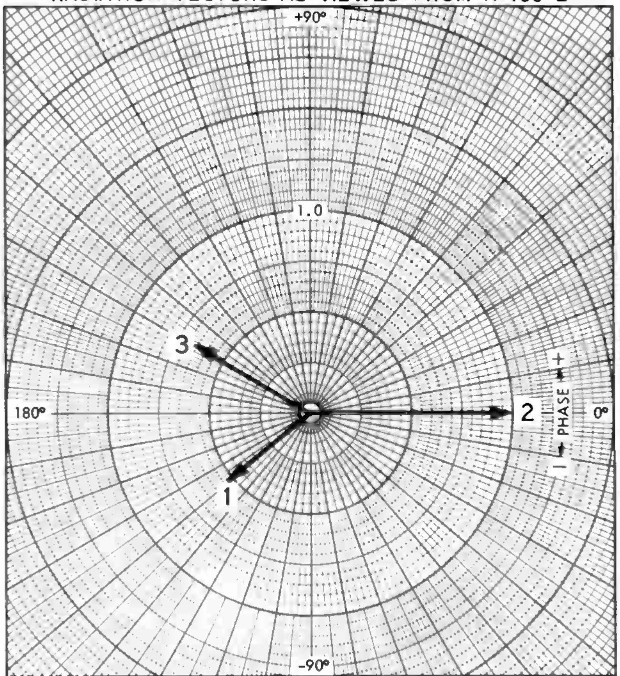
THEORETICAL PATTERN AND PARAMETERS



Tower	Bearing	Distance	Field	Phase
1	0°	100°	0.538	-123.6°
2	0°	0°	1.000	0.0°
3	180°	100°	0.673	132.4°

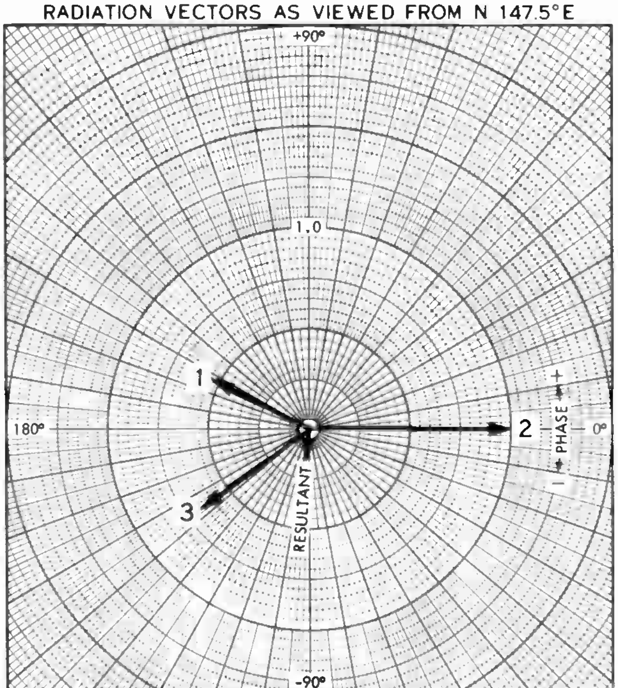
K = 178.9 mV/m

RADIATION VECTORS AS VIEWED FROM N 100°E



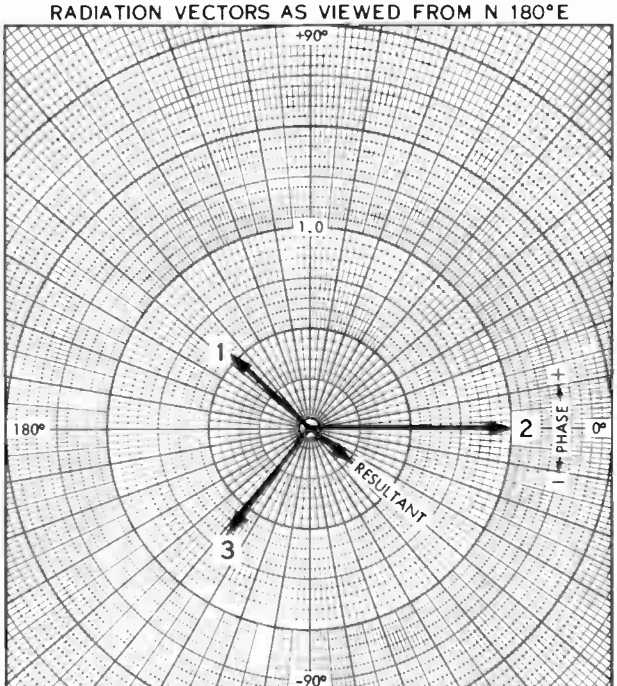
Tower	Field	Electrical Phase	Space Phase			
1	0.538	/-123.6°	+ 100° cos100°	0.538 /-141.0°	=	0.418 -j0.339
2	1.000	/ 0.0°	+ 0°	1.000 / 0.0	=	1.000 j0
3	0.673	/ 132.4°	- 100° cos100°	0.673 / 149.8°	=	0.582 +j0.339
				Resultant	=	0.0 j0.0
					=	0 /0°

RADIATION VECTORS AS VIEWED FROM N 147.5°E



Tower	Field	Electrical Phase	Space Phase			
1	0.538	/-123.6°	+ 100° cos147.5°	= 0.538 / -207.9°	=	-0.475 +j0.252
2	1.000	/ 0.0°	+ 0°	= 1.000 / 0.0°	=	1.000 j0
3	0.673	/ 132.4°	- 100° cos147.5°	= 0.673 / 216.7°	=	-0.540 -j0.402
				Resultant	=	-0.015 -j0.150
					=	0.151 /-95.7°

RADIATION VECTORS AS VIEWED FROM N 180°E



Tower	Field	Electrical Phase	Space Phase			
1	0.538	/ -123.6°	+ 100° cos180°	= 0.538 / -223.6°	=	-0.390 +j0.371
2	1.000	/ 0.0°	+ 0°	= 1.000 / 0.0°	=	1.000 j0
3	0.673	/ 132.4°	- 100° cos180°	= 0.673 / 232.4°	=	-0.411 -j0.533
				Resultant	=	0.199 -j0.162
					=	0.257 /-39.1°

Figure 3. Tower radiation vectors as viewed from different azimuths.

towers and the space phase component for each tower relative to the reference tower as viewed from each direction. Such diagrams will assist in visualizing what effect any contemplated changes in phase or ratio for any tower will have on the resultant field in each direction. The vector diagram amplitudes may be in terms of the individual tower field ratios (such as in Fig. 3) or in terms of the theoretical unattenuated radiation. The theoretical unattenuated radiation from each tower (expressed in millivolts-per-meter at one kilometer or one mile) is obtained by multiplying the theoretical pattern field ratio for each tower by the pattern constant K, which is also usually shown in each theoretical pattern.

7. If the pattern includes essentially zero minimums (which need to be as deep as possible), make talkdown adjustments at a series of the most distant points on the null radial from which adequate two-way communications can be obtained. At each talkdown location, adjust the phasor for any combination of tower currents and phases that gives the deepest possible minimum. (A deep minimum can usually be confirmed by rotating the field strength meter 90% on a vertical axis and noting that the signal received from scattered reradiators is stronger than that received from the station when the meter is in its normal measurement orientation.)

Construct the vector resultant for each talkdown on a vector diagram similar to that shown in Fig. 3, but use the antenna monitor indications of phase and ratio instead of the theoretical parameters. The vector resultant will probably not be zero because of errors in the antenna monitoring system, nonsinusoidal current distribution on the towers, and reradiation from objects external to the array. However, the resultants so plotted from a series of talkdowns along a radial will enclose an area in which the optimum adjustment has its vector resultant at its center. Knowing the desired resultant for each critical radial direction, one may then be able to infer a set of phase and ratio parameters to satisfy conditions in all directions simultaneously.

For simple arrays, a pragmatic approach that often yields effective results, is to station observers with field strength meters at each of the several monitoring points or in other critical directions. By using a two-way radio to learn the effect on field strength, cut-and-try adjustments of the phasor can yield a set of phase and ratio parameters that satisfy all conditions simultaneously.

Following readjustment, if any of the new phase and ratio parameters exceed the three degree phase and five percent ratio tolerances permitted by the FCC when applied to the parameters shown on the current station license, a partial proof-of-performance will be required before the FCC will license the new parameters. If the new parameters are within tolerance of the licensed values, a partial proof and application for changed parameters may still be desirable in order to avoid long-term operation uncomfortably close to the tolerance limits.

2.6 Antennas For FM Broadcasting

Peter K. Onnigian, P.E.
Consulting Antenna Engineer, Sacramento, California

Additional material provided by:
Eric Dye
Jampro Antennas, Sacramento, California

INTRODUCTION

This chapter is dedicated to broadcast engineers, technicians, and station managers who must make important decisions regarding FM transmitting antennas. To insure the best possible signal strength in the station's service areas, the site location, antenna height, antenna type, and propagation conditions must all be considered.

FM broadcasting was first authorized in the United States in 1940 by the Federal Communications Commission (FCC). The first FM station began operation in 1941. In 1945 the FM service was assigned to the 88 to 108 MHz band and divided into 100 channels, each 200 kHz wide. In 1991 there were over 5,400 commercial and 1,700 educational FM stations.

Most FM antennas are nonsymmetrical, that is they are mounted on one side of a steel supporting tower or pole. FM antennas outside the western hemisphere on the other hand are usually symmetrical, that is installed on all faces of a tower. However, both methods are capable of providing excellent omnidirectional azimuth patterns.

Antennas for FM broadcasting use horizontal polarization (H-pol), vertical polarization (V-pol), or circular polarization (CP).[1] Cross polarization is used as a means to prevent co-channel interference in some European countries but not in the western hemisphere. CP, together with its special form, elliptical polarization (E-pol), was introduced in the United States in the early 1960s as a means to provide greater signal penetration into the many different forms of FM receiving antennas, which are now found in the service area. H-pol is the standard in the United States. CP or E-pol may be used if desired. V-pol only is permitted for noncommercial FM stations seeking to limit interference to TV channel 6.

In the 10 year period between 1980 and 1990 there were over 300 million radios sold in the United States, of which over 140 million were automobile AM/FM radios.[2] FM radio receivers use a variety of antennas including extendable monopoles (whips), dipoles, and capacitive coupling to power lines and headphone leads.

Antennas for FM broadcasting must be chosen carefully in order to cover the service areas properly with adequate level and quality signals. For economic and technical reasons, the desired effective radiated power (ERP) should be produced with a balance between antenna gain and transmitter power. It is the purpose of this chapter to provide sufficient technical information for the broadcaster to achieve this.

The height of the antenna over the service area, distances to areas of population, the ERP, and the economics are items that must be considered.

Antennas currently available in the United States differ considerably from those to be found in Europe. The various American types are discussed so that the engineer will be informed on the subject. Considerable advances have been made in recent years in the design and fabrication of FM antennas. These improvements provide greater penetration of signals into automobile FM radios as well as popular small FM transistor radios of all kinds.[3] The newer FM broadcasting antennas must meet the more stringent requirements for FM stereo and subcarrier broadcasting. Most FM stations in the United States are using CP antennas.

PROPAGATION

FM broadcasting has some distinct advantages over AM (medium wave) broadcast service. These advantages stem from the propagation characteristics of FM frequencies as well as the modulation system.

There is essentially no difference between day and night FM propagation conditions. FM stations have relatively uniform day and night service areas. FM propagation loss includes everything that can happen to the energy radiated from the transmitting antenna during its journey to the receiving antennas. It includes the free space path attenuation of the wave and such factors as refraction, reflection, depolarization, diffraction, absorption, scattering, Fresnel zone clearances, grazing, and Brewster angle problems.

Propagation is dependent upon all these properties out to approximately 40 miles (65 km). Some additional factors enter the picture at greater distances. Radio wave propagation is further complicated because some of these propagation variables are functions of frequency, polarization, or both, and many have location and time variations.

The technical intent of the broadcaster is to put a signal into FM receivers of sufficient strength to overcome noise and to provide at least 20 dB carrier-to-noise ratio, which will provide at least 30 dB of stereo separation. The required RF signal level varies from about 2 μV/m (microvolt per meter) for high sensitivity FM stereo tuners in the suburbs to about 500 μV/m for less sensitive transistorized portables. Automobile receivers have wide ranged sensitivity values.

FM antenna manufacturers do not guarantee coverage. They supply antennas which meet certain radiation pattern requirements and gain. Many of the antennas assume an omnidirectional pattern. Although reasonable in free space, it is never fully achieved in practice due to sources of distortion (support structure, feed lines, etc.)

Some manufacturers in the United States provide azimuth pattern adjustment service to insure a horizontal plane pattern circularity of ± 3 dB when mounted on the side of a specific tower or pole. It must be pointed out that this radiation pattern and gain are for free-space conditions and may not relate directly to signal strengths measured at or near ground level, well away from the antenna.

Radiation pattern and propagation are two distinctly separate conditions. The pattern is the radiation which is transmitted by a given antenna in any given direction, without any propagation limitations, as measured on a good antenna test range. Propagation depends on path and environmental conditions existing between the transmitting antenna and the receivers.

The actual service area signal strength contours are based upon two probability factors. Contours are not solid signal areas. For example, the FCC signal coverage charts referred to as the F(50/50) curves, are based upon a probability of occurrence of certain voltage levels at 50% of the locations, 50% of the time. This means that at any given location, 30 ft above ground, the signal has a 50% chance to measure up to the predicted contour level. Furthermore, half the time at that location, it may reach or exceed the level predicted while at other times it may be lower in strength.

These FCC charts (FCC Rules, Section 73.333) are based upon the assumption that average propagation conditions exist. One or more of the conditions mentioned in the second paragraph under this heading may reduce the measured signal strength from the predicted values substantially.

Propagation Loss

The power radiated from a FM transmitting station is spread over a relatively large area, somewhat like an outdoor, bare light bulb on top of a tall pole. The power reaching the receiving antenna is a very small percentage of the total radiated power.

At 100 MHz and a distance of 30 miles (48 km) the figures indicate the free-space path loss to be 106 dB.[4] The formula used to compute free-space loss is:

$$FSL = 36.6 + 20 \log D(\text{miles}) + 20 \log F(\text{MHz})$$

Doubling the distance increases the space loss by 6 dB. The path loss does not attenuate the signal with distance as much as some other factors. Path loss between an earth station and a satellite is a classic text book example of a 6 dB loss every time the distance is doubled. But a typical FM station signal travels through a nearly perfect dielectric (air) and over the imperfect earth's surface (ground). Herein lies the FM radio propagation loss problem.

Refraction, diffraction, and reflection from scores of objects such as hills and buildings may occur in the propagation path between the transmitting and the receiving antennas. These, along with absorption, scattering, lack of Fresnel zone clearances, and other factors, all reduce the signal strengths.

Signal loss due to foliage has been well known to UHF TV broadcasters for many years.[5] This same condition exists to a lesser degree for FM broadcasting. Trees, shrubs, and other foliage on hills or smooth terrain affect the reflected as well as the lateral signal loss with distance. With average values of permittivity and conductivity in both foliage and ground, a loss of about 2.5 dB was found to exist in a ten-mile path, at FM frequencies.[6] The height gain factor is increased with heights above the foliage.

Considerable depolarization takes place because the transmission through or reflections from ground foliage is a diffracted field contribution.

Multipath Problems

The ideal reception condition is a strong direct single source signal. When energy from two or more paths reaches the receiver, (due to reflections) a condition called multipath reception occurs. Poor reception is experienced when there is insufficient strength difference between the direct and the reflected signals, because they can cancel each other where the geometry places them out of phase.

Nothing is more important in the way of broadcasting facilities than the location of the transmitting antenna. Great care must be exercised to find a suitable site. Poor selection of the transmitter point can result in very unfavorable signal propagation and negate the entire project. One very serious result of poor site selection is multipath propagation in some directions.

As an example, the transmitter should not be located so that strong reflections take place from nearby hills or mountains. This can happen when the transmitter is placed on one side of a large city and the other side of the city has a high mountain range. Radiation into the city directly from the transmitting antenna, as well as reflections from the nearby hills and mountains will create two or more signal paths to many receivers. These reflections can be so strong that only a 10 dB difference may exist between the direct and the reflected condition which causes severe multipath problems.

A TV station at this same location would experience unusable signals due to heavy ghosting, even with directional receiving antennas, which exhibit moderate signal pickup from their back. This is illustrated in Fig. 1 where a mountain range causes reflections back into a large city.

The multipath example shown in the sketch was an actual case. The site was chosen by the FM broadcaster, without proper engineering guidance, simply because the hill had a tower, building, power, and a road was in place. Later, the broadcaster learned why the original owner, a television station, had abandoned the site: the TV station had failed in part due to extremely heavy ghosting into the principal city.

A much better FM transmitting site was located on the hills between the high mountain range and the city. Using a directional transmitting antenna with very little radiation towards the high mountains, reflections were satisfactorily reduced, and the FM station is now operating successfully.

Multipath reflections are very easy to identify. On an automobile radio, the signal will drop out, sometimes abruptly, as the car moves. This effect may be rhythmic with distance while traveling slowly. It is sometimes called picketing as it acts like a picket fence alternately blocking and letting the signal pass. A field strength meter will usually reveal great variations of signal when moving, say, 100 ft (30 m) in a line with the transmitter. Cyclic variations over quite uniformly spaced intervals on the ground as great as 40 dB have been observed by the author.

This variation in signal levels is caused by the reflections adding and subtracting from direct and reflected signals caused by propagation problems existing in the path between transmitter and receiver. It usually has nothing to do with the qualities of the transmitting antenna. It is a function of site selection. This should not be confused with a similar effect observed near the base of the tower supporting a high-gain antenna. Nulls produced by stacking bays for gain are found near the antenna and may be filled-in if needed. (See Beam Tilt and Null Fill in this chapter.)

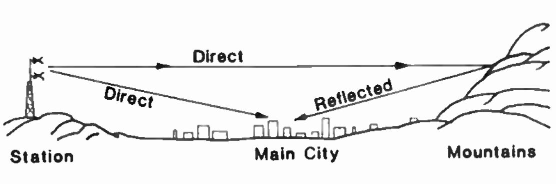


Figure 1. Example of poor station location causing severe multipath conditions due to delayed reflections from mountains on the right. Not to any scale.

Ground Reflections

In the elevation plane between transmitter and receiver, nearly all FM signal coverage lies between the horizon and 10° below. Called the *grazing angle*, it lies between the horizontal plane and the earth's surface. Generally the higher the transmitting antenna above the service area, the greater this angle will be.

The angle of incidence and reflection are not the same, as shown in Fig. 2. The depression angle and the grazing angle are not equal as would be the case for a flat earth. Reflections from these angles play an important part in the strength and the quality of the signal in FM broadcasting with circular polarization.

The ground which causes reflections at these grazing angles does not treat H-pol and V-pol in the same manner. The V-pol is attenuated considerably more than the H-pol as shown in Fig. 3. The phase of the V-pol changes substantially with angle, while H-pol remains nearly the same. At these useful low propagation angles, there is considerably less V-pol signal reflected than H-pol, when grazing takes place. Field measurements confirm this fact.[7] For this reason, it is impossible to measure accurately axial ratios in the service area. To be meaningful, the H-pol and V-pol ratios must be measured on a good antenna test range.

It is quite difficult to predict accurately the reflection coefficient (efficiency), which varies considerably as a function of polarization, frequency, grazing angle, surface roughness, soil type, moisture content, vegetation growth, weather and the season. There are complex formulas for predicting the ground conductivity at the frequency of interest. For 100 MHz, a value of 10 millimhos per meter ground conductivity is often used, with a permittivity of 25, as being about the average for the continental United States.[6]

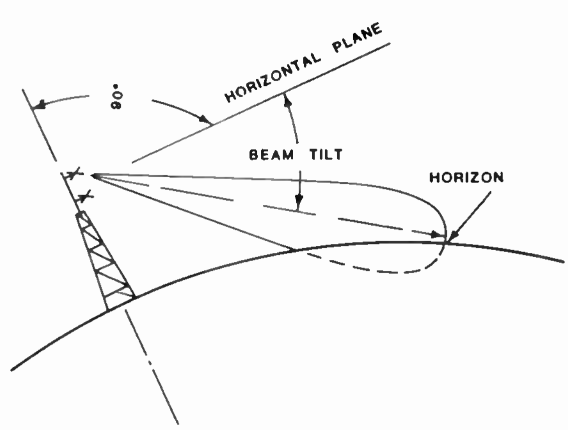


Figure 2. Beam tilt to radiate maximum ERP at the horizon. Not to any scale and exaggerated for illustration.

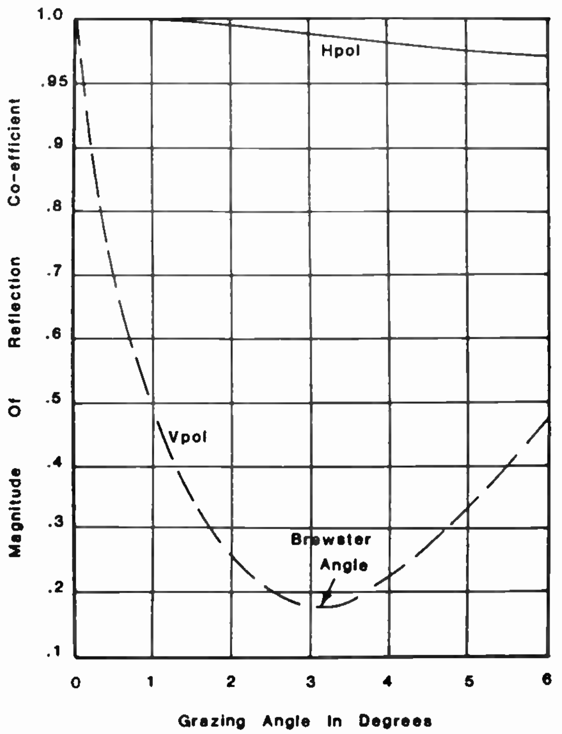


Figure 3. Magnitude of reflection coefficients showing differences for Hpol and Vpol, and the Brewster angle.

Brewster Angle

For polarization with the electric field normal to the plane of incidence, there is no angle that will yield an equality of impedances for earth materials with different dielectric constants but like permeabilities. An incident wave with both polarizations present will have some of the one polarization component but little of the other reflected. The reflected wave at this angle is thus plane polarized with the electric field normal to the plane of incidence, and the angle is the polarizing angle.

Notice that in Fig. 3 the minimum reflection coefficient occurs at a grazing angle of about 2°. Below this angle, the reflection coefficient rapidly increases to unity. The angle at which the minimum reflection coefficient occurs is called the Brewster or polarizing angle, after the Englishman who first discovered this phenomenon.

For ground reflections occurring near the Brewster angle, the reflection coefficient is much smaller for V-pol than the H-pol. Therefore, the reflected V-pol signal component of CP are attenuated considerably. The greatest attenuation for V-pol from ground reflection occurs at this angle.

Field measurement of V-pol signals will usually show a significant variability of H-pol to V-pol ratios due to this Brewster angle phenomenon. The Brewster angle is a function of soil conductivity and may change from place to place, as well as from season to season.[8]

It is important, then, that the antenna height above the service area results in grazing angles which are less than the Brewster angle. Otherwise the V-pol will be reduced and the radiation will be much more elliptical than circular in polarization.

Fresnel Zone Clearance

A much neglected consideration in FM transmitting antenna location and height is Fresnel zone radius clearance in the path to the service area. Microwave engineers always make certain that their signal paths have this important clearance.

The effect of clearance above ground or other obstacles was studied by August Jean Fresnel, a French scientist who first discovered this phenomenon in optics. Fresnel zones are circular areas surrounding the direct line-of-sight path of a radius such that the difference between the direct and the indirect path length to the zone perimeter is a multiple of half-wavelength longer than the direct path. This is illustrated in Fig. 4. The zone diameter varies with frequency and path length. The greater the path length, the larger the required mid-path clearance required for full signal.

Fresnel also discovered that the entire first zone radius is not required for full signal strength. Six-tenths of the first zone would suffice, which is fortunate since the radius is quite large at the FM frequencies. The equation for determining the first Fresnel zone radius for 4/3rd earth curvature is:

$$R = 1140\sqrt{d/f}$$

where d is the path length in miles, f is in MHz and R is in feet for the first radius.

In Table 1 the required 0.6 first Fresnel zone radii clearances at the middle of the path are shown for 98 MHz and service areas up to 52 miles (92 km) from the transmitter. The idea is to raise the height of the transmitting antenna so that the mid-path height is as high as or higher than shown in the table. Due to the geometry of the Fresnel zone, if the terrain is relatively flat, the mid-path radius will control and be larger than that required elsewhere along the path. If the mid-path clearance is less than the values shown, the FM signal will be attenuated in accordance with the curve shown in Fig. 5, presuming ideal reflection off the ground or obstructions.

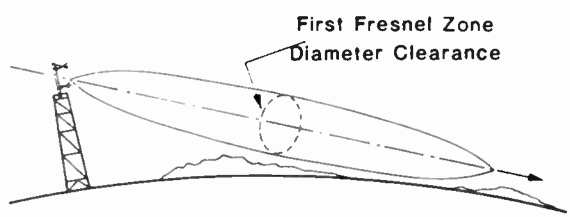


Figure 4. First Fresnel zone clearance occurs as shown above, but only six-tenths is required for full free space signal level. Not to any scale.

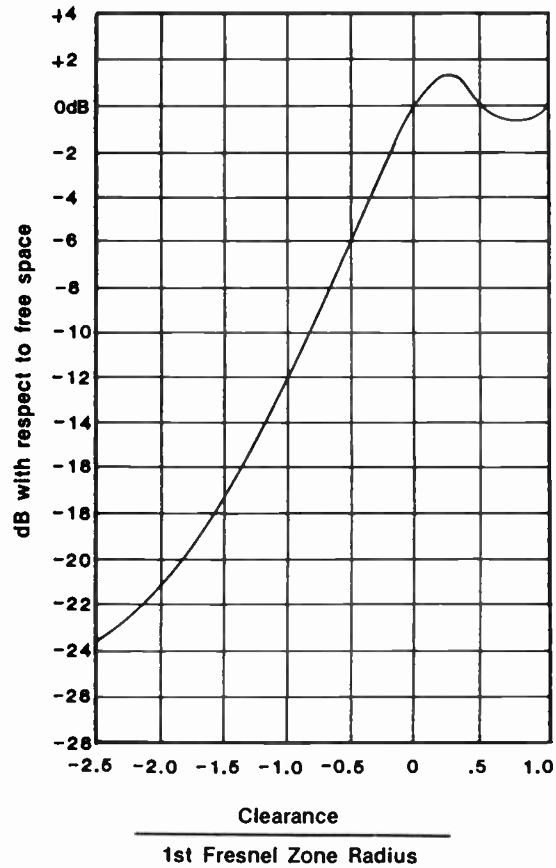


Figure 5. Attenuation of FM propagation when the path between transmitter and receiver lacks Fresnel zone clearance in the ratios shown.

The center-of-radiation heights of the antennas in Table 1 are actual and not height above average terrain (HAAT). Some of these recommended heights will reduce the allowable effective radiated power (ERP) in accordance with FCC 73.211 (b), depending on the class of station and the zone. However, it is better to have the Fresnel clearance than the maximum low height ERP values, as the higher heights will produce stronger signals.

It is a well known propagation axiom that greater heights are more useful in producing higher signal strengths far from the antenna than ERP levels, everything else being equal.

Without the first Fresnel clearance of 60% the signal level at the distant point may suffer. This reduction will follow the curve shown in Fig. 5 for different values of clearance and worst case reflection conditions.

In order for the FCC prediction curves to be valid, the recommended minimum antenna heights should be achieved. These heights not only provide line-of-sight conditions to the service limits but also proper Fresnel clearances. Both conditions are required for the FCC F(50,50) curves to be valid.

The values in Table 1 are for relatively flat terrain, but take into consideration the FCC suggested roughness factor of up to ±150 ft (50 m). Where the tower height is limited by HAAT values or other limitations, the signal strength will suffer due to those factors.

Soil Conductivity

The conductivity and permittivity of the soil, together with the vegetation on it, play a small part in the attenuation of FM signal strength. Average soil has a dielectric constant of about 15 millisiemens per meter at 100 MHz.[9]

TABLE 1

Recommended minimum antenna heights (for flat terrain and 98 MHz).

Service Area Radius Required		Fresnel Zone Six-tenths Clearance		Recommended Min. Antenna Height		Probable FCC 80-90 Class
Miles	Km	Feet	Meters	Feet	Meters	
5	8	155	47	310	95	A
7½	12	189	58	378	115	A
10	16	218	66	426	130	A
15	24	267	81	534	167	A,B,C-2
20	32	309	94	618	188	B-1,C-2
25	40	346	105	700	213	B-1
30	48	378	115	756	230	B
35	56	409	125	818	250	B
40	64	437	133	875	267	C-1
45	72	463	141	925	282	C-1
50	80	488	149	975	297	C-1
57	92	522	159	1,043	318	C,C-1

Linear Height Gain Effect

By raising the receiving antenna above the immediate effects of the soil, the signal level will be increased. Actual field measurements have proven a 9 dB increase in signal when the dipole was raised from 3.28 ft (1 m) to a level of 30 ft (9.1 m). This is due to reflection phenomena in the foreground of the receiver, not ground conductivity.

FCC Service Contours

From the FCC coverage prediction charts, it is possible to draw contours of the various grades of service for a given ERP and antenna height above average terrain. These predictions, at 50% of the locations, 50% of the time, constitute the basis for the service contours. The city grade contour is 70 dBμ (3.16 millivolts per meter) and primary service contour is 60 dBμ (1.0 millivolts per meter).

The FCC Rules, Section 73.333 charts for these predictions have a built-in terrain roughness factor, as explained above.

GENERAL COVERAGE STANDARDS

There are certain height and power levels fixed by the FCC for various classes of stations. The United States has been divided into three different geographical areas based on population density as well as propagation refractive index levels. These ERP and height values have been set to prevent co-channel and adjacent channel interference.

Zone I, generally speaking, is the northeastern part of the United States. Zone I-A includes Puerto Rico, Virgin Islands, and that portion of California lying below the 40th parallel. Zone II includes Alaska, Hawaii, and the remainder of the United States not in the above two zones. This is more fully described in FCC Rules Section 73.205.

Under Rules which resulted from Docket 80–90: "Modification of FM Broadcast Station Rules to Increase the Availability of Commercial FM Broadcast Assignments," in 1983, new ERP levels and additional classes of stations were created. The distance to the 60 dBu (1 mV/m) signal contour is the controlling factor so that the ERP based on the HAAT is adjusted to produce that level and no more at a specific distance for a particular class station.

Table 2 shows for each FM class station, the zone, the maximum ERP, the maximum HAAT, and the distance to the 60 dBu contour calculated by using the maximum ERP and HAAT, and then rounding to the nearest kilometer and mile. The FCC issued these in that docket and they are currently the standards.

Under the Rules, Class C stations are required to have at least 100 kW ERP and an antenna height of more than 984 ft (300 m) above average terrain. Class C-1 stations are now permitted a maximum of 100 kW ERP with an antenna maximum HAAT of 984 (300 m), while C-2 stations may go to 50 kW at a maximum HAAT of 492 ft (150 m). Higher HAAT may be used with reduced ERP values, in accordance with equivalent 60 dBu coverage. Class C and C-1 stations may thus share the same antenna and tower.

Stations may be upgraded using the easiest method, which is to increase existing location tower height. Such factors as local zoning laws and aircraft flight patterns may preclude this approach, however.

FM Signal Measurements

The signal strength received at 5 ft (1.5 m) above ground, which is about average for auto whip antennas, is several times lower in level than at the standard FCC measurement height of 30 ft (9.1 m). This fact should be taken into consideration when comparing low height measurements with the FCC Rules Section 73.333 prediction charts, which are based on a 30 ft receiving height, where signals are considerably stronger.[10]

Signal levels inside houses, apartments, offices, and other structures vary greatly. Levels depend on the type of building construction, but in nearly all cases will be lower than those outdoors. Reflections inside the building reduce stereo separation, and cause cross-talk problems with SCA channels. Outside FM receiving antennas generally provide good reception.

Field strength measurements should not be used to determine the transmitting antenna radiation pattern or efficiency except under carefully controlled conditions.

TABLE 2

Docket 80-90 FM station classes, zones, and ERP.

FM CLASS	ZONE	MAX. ERP In kW	MAXIMUM HAAT Feet	Meters	DISTANCE TO 60 dBu Miles	km
A	I, I-A	3	328	100	15	24
B	I, I-A	50	492	150	32	52
B-1	I, I-A	25	328	100	24	39
C	II	100	1,969	600	57	92
C-1	II	100	984	300	45	72
C-2	II	50	492	150	32	52

The propagation factors discussed previously camouflage the true antenna performance. The only technically acceptable way to determine the antenna's characteristics is on an antenna test range.

Elsewhere in the *Handbook* is an entire chapter dealing with the recommended methods and equipment to make FM and TV field strength measurements. This information may be used to determine the actual quality of service and the areas where useable signal levels in fact exist. Predicted contours may be considerably different from actual measured values.

Required Signal Strength

What is the minimum satisfactory signal strength? What is the maximum above which it is wasteful? The history of FCC proceedings provides some of the following levels:

34 dBu = 0.05 mV/m For rural areas
60 dBu = 1.00 mV/m Suburban areas
70 dBu = 3.16 mV/m Principal community
82 dBu = 12.64 mV/m Highest useful level

The first three levels were set by the FCC in the early 1950s when tube receivers and H-pol antennas were popular. Modern day transistor radios have much greater sensitivity. CP has added greater signal penetrating power than H-pol when the levels were first established.

The FCC defines two grades of signal contours on its applications. The first is based on the 70 dBu contour (3.16 mV/m) required to cover the principal community of license. The second is the 60 dBu contour (1 mV/m) which defines the primary service area.

The FCC also stated that, in rural areas, levels as low as 50 microvolts-per-meter were useful. Indeed current home stereo tuners and FM auto radios operate very well with only 25 microvolts-per-meter. In practice, 50 microvolts-per-meter (0.05 mV/m) provides good quieting in nearly all automobile and transistor radios receiving a stereo signal from a CP station antenna. Therefore 50 microvolts-per-meter should be considered the minimum useful signal level.

If the highest level of 3.16 mV/m is quadrupled, it will be 12.64 mV/m. This is a 12 dB increase, equal to increasing the FCC power level by more than 15 times. It can be safely said that this level of 12.64 mV/m is considerably more signal than necessary by any present day working FM radio. Any signal level higher than this at the receiving antenna has not proven to be of significant value.

Blanketing

Excessive RF signals can overload the front end of receivers and make satisfactory reception impossible. The FCC in Section 73.318 defines the 115 dBu (562 mV/m) level as the "blanketing contour," and adopted the free-space prediction method to predict how far this contour extends.

New or modified FM stations have the responsibility to satisfy all complaints at no cost to the complainant, of blanketing-related interference inside this contour within one year of commencement of operations.

The distance to the 115 dBu contour is determined using the following equation:

$$d \text{ (in kilometers)} = 0.394\sqrt{P}$$
$$D \text{ (in miles)} = 0.245\sqrt{P}$$

where P is the maximum effective radiated power (ERP), measured in kilowatts of the maximum radiated lobe, irrespective of vertical directivity. For directional antennas, the horizontal directivity shall be used.

ANTENNA CHARACTERISTICS

Antenna Gain

Gain can be increased by adding additional radiating elements (bays) to the antenna at the cost of narrowing the radiated beam. High-gain antennas concentrate the energy into such a narrow beam that often null fill must be employed to achieve the desired signal strength within the first few miles to the tower.

Directional antennas achieve increased gain over nondirectional antennas by limiting the radiated energy in various directions. Directional antennas are useful when the tower is located near a large body of water, mountain range, or other areas where energy radiated in those directions is otherwise wasted. They are also employed to avoid interference where stations are insufficient distances apart.

Antenna gain is expressed in power ratio or in dB. For example, an antenna with a power gain of 2 is also said to have a gain of 3.0 dB.

FCC Rules, Section 73.310(a) defines antenna gain as the inverse of the square of the root mean square value of the free-space field strength produced at one mile in the horizontal plane, in millivolts per meter for 1 kW antenna input power to 137.6 mV/m. (In metric units, 1 km and 221.4 mV/m).

Notice that this gain is in reference to a horizontally polarized half-wave dipole. For a CP antenna, the gain is half for the same input power.

A two-bay H-pol antenna has a power gain of approximately two. But a two-bay CP antenna in FCC terminology has a gain of about one because the other half of the power is V-pol and is not considered in the gain calculations. Only the horizontal polarization mode is used by the Commission. The vertically polarized energy must not exceed the H-pol (except for noncommercial, educational FM facilities attempting to minimize interference to TV channel 6 reception).

The power gain of an antenna is used with the transmitter gain and transmission line loss when determining the Effective Radiated Power (ERP). Consider for example a 10 kW transmitter and an antenna power gain of 5. Neglecting transmission line loss, the ERP is 10 kW × 5 = 50 kW ERP. If the antenna gain were 10 and the transmitter power was 5 kW, we would have the same ERP of 50 kW. (5 kW × 10 = 50 kW ERP)

Effective Radiated Power (ERP)

The FCC defines the term *effective radiated power* to mean the product of the antenna input power (transmitter output power less transmission line loss) times the antenna power gain. Where circular polarization is used, the term ERP is applied separately to the H-pol and V-pol of radiation. For allocation purposes, the ERP is the H-pol component of radiation only. The V-pol component power normally must not exceed the H-pol power.

Beam Tilt

FM broadcasting antennas are normally mounted on towers which are plumb, so the peak power beam in the elevation pattern is perpendicular to the tower axis. A standard FM antenna without any beam tilt radiates more than one half of the total radiated power above the horizon. All this power is lost.

The higher the antenna is above its average terrain, the larger the predicted coverage area. Since the earth is curved, the service horizon is bent lower than a perpendicular angle from the earth's surface. Thus the strongest portion of the signal is aimed above the horizon. It also follows that the higher the antenna above the terrain, the greater is the elevation angle down to the earth's horizon.

In order to strike the farthest service area from a high HAAT, the beam may need to be tilted down towards the earth. Electrical beam tilt lowers the beam angle equally in all azimuth headings and is chosen more frequently than mechanical tilting, which exhibits different effects in different directions. Choose enough tilt to position the center of the main beam on the furthest edge of the desired coverage area or just below the horizon, whichever is closer.

For low gain antennas (two to four bays), the main beam is very broad, and if the antenna height above average terrain (HAAT) is less than 500 ft there is little to be gained with beam tilt. On the other hand, beam tilt makes a large difference on high-gain antennas mounted on towers with a large HAAT.

Fig. 6 shows the comparison between elevation angle path and coverage distance. It incorporates the curvature of the earth. This chart can be used to determine the optimum beam tilt. Follow the curve which is closest to actual HAAT, and mark the point where it intersects the horizon or crosses the distance of furthest desired coverage area (vertical axis). Read the beam tilt on the horizontal axis. Round this value up to the nearest ¼°.

Consulting engineers, familiar with this problem, can easily work out the required amount of beam tilt, if it is necessary. Typical values are one-half to one degree of tilt, depending on the antenna height, distance to the far service area, and the antenna elevation pattern.

Beam tilt is usually accomplished electrically, by delaying the currents to the lower bays, and advancing the phase of the upper bay currents during the design and construction of the antenna at the factory.

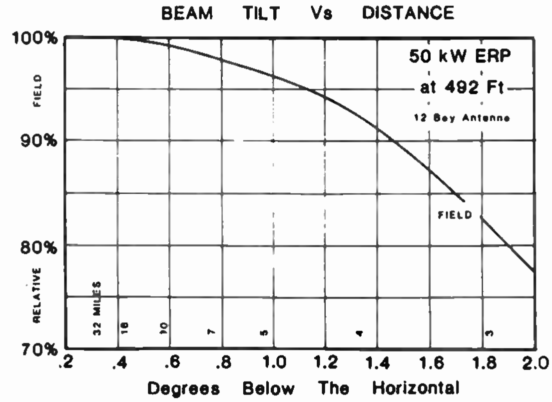


Figure 6. Twelve bay Cpol antenna with 0.3° beam tilt showing ERP distribution with coverage from 492 foot (150 m) tower over flat land. Degrees below the horizon are based on 4/3rd earths curvature. Horizon is −0.341° at 32 miles (52 km).

Null Fill

While the beam tilt puts more signal into the far reaches of the service area, it does not solve the problem sometimes caused by high-gain antennas within several miles of the transmitter. Elevation angle nulls common to all antennas with two or more bays appear farther and farther away from the antenna as its gain is increased with more bays.

When multiple bay arrays are employed, lobes and nulls occur in the elevation pattern. As the number of bays increases, the main beam narrows and the first null radius increases. The advantage of beam tilt and null fill varies depending on factors such as tower height, site elevation, number of bays, and relative locations of communities to be served.

A simple rule-of-thumb is that null fill is beneficial when there is desired service area within the radius of the first null. The equation below gives the approximate radius of the first null for multiple bay antennas.

$$\text{Null Radius (miles)} \cong \frac{H_{(ft)}}{\text{Number of Bays}/5280\ \text{ft/mi}}$$

where H is the height (ft) of antenna above radius ring.

In most FM applications, the null is very close to the antenna, thus a small amount of null fill (5–10%) takes care of the problem. Larger amounts of null fill are unnecessary and reduce the gain of the antenna. Note that null fill has no effect on distant coverage.

Refer to the "TV Antennas" Chapter 2.7 for information on null fill and beam tilt. The matter is essentially the same for FM and TV antenna systems.

VSWR Bandwidth

According to theory, the bandwidth of an FM signal is infinite if all the sidebands are taken into account. Also, at certain modulation indices, the carrier amplitude goes to zero and all the transmitted power is on frequencies (sidebands) other than the carrier frequency. Practical considerations in the transmitter and receiver circuitry make it necessary to restrict the RF bandwidth to less than infinity.

Prior to 1984 the maximum deviation for FM stations was 75 kHz, representing 100% modulation. In that year the FCC changed the maximum deviation to 82.5 kHz (110%) for those stations with 10% injection of subcarrier channels. This additional deviation requires greater antenna system bandwidth than previously needed.

System bandwidth is measured at the point in the antenna system where the transmitter is connected. This usually includes the harmonic filter, the main coaxial transmission line, and the antenna.

The significant sidebands are usually considered to be those whose amplitude exceeds one percent of the unmodulated carrier. With 110% modulation (82.5 kHz deviation) these side bands produce a bandwidth of 260 kHz.[11,12]

The VSWR bandwidth is the range over which the system under consideration has a reflection coefficient of less than five percent; a VSWR of 1.1:1.

Checking System VSWR

From time to time, the VSWR of the narrow-band antenna system should be checked and adjusted. If the exciter has thumb wheel exciter frequency adjustability in 10 kHz or 50 kHz steps, it can be used to change the frequency to check VSWR on different frequencies (with the transmitter operating at low power during the overnight experimental period). The reflectometer may be used as the indicator.

Alternately one of several methods for checking VSWR in coaxial line systems using test equipment may be used. These include a signal generator test setup, an impedance test set, or a network analyzer.

The VSWR should be measured to ensure that the reflection response is balanced to 130 kHz on each side of the carrier frequency. With transmission lines longer than 300 ft (100 m) it is suggested that the VSWR bandwidth be under 1.08:1 all the way out to ±130 kHz. The additional delay due to increasing line length becomes more of a problem, so the amplitude of the reflection must be reduced, for best operational results.

Importance of Low VSWR

The VSWR shown by the transmitter reflectometer does not increase or decrease the range of the signal. It has nothing to do with coverage. But VSWR values above 1.1:1 may decrease the final amplifier efficiency. Other definite negative effects of VSWR are increased intermodulation products and AM synchronous noise. Stereo separation is also degraded with increased VSWR.[13]

Intermodulation and SAM Distortion

Intermodulation distortion and synchronous AM (SAM) noise can be caused by narrow VSWR bandwidth in the antenna system, as well as by final amplifier circuitry all the way to and including the antenna.[14]

Synchronous AM (SAM) is extremely important in FM transmitter facilities employing subcarriers. SAM is AM modulation of the carrier caused by frequency modulation of the carrier frequency in the VSWR notch. At the notch the reflected energy is the lowest. As the deviation takes place, the greater the frequency swing, the greater will be the reflections, due to the VSWR notch. With a flat VSWR curve, SAM does not take place. If the VSWR curve is skewed, SAM will occur and intermod and stereo crosstalk will increase.

Directional Antennas

The FCC sometimes requires that the azimuth radiation pattern be directionalized to reduce normally allocated ERP towards a given short-spaced station, or for other reasons. See the FCC Rules Section 73.213, 215, and 316(b) and (c). To conform to these specifications, most broadcasters order antennas which are pattern adjusted, measured, and certified to the Commission's requirements.

Directional antennas are licensed for peak ERP values based on the azimuth pattern. The V-pol gain may not exceed the H-pol gain in a CP directional array nor may V-pol exceed the H-pol in the protection direction (except in the case of FM protection to TV Channel 6, mentioned elsewhere). The amplitude away from the null cannot climb more than 2 dB per 10° of azimuth.

Directional antennas are usually mounted on poles although some have been tower mounted. Since the support affects the pattern, they are specified and measured with the pole or tower on which they are mounted. Most firms will make the antenna meet the specific pattern requirements.

Directionalizing is a combination of the natural pattern resulting from sidemounting and the use of parasitic elements. Using the two factors, virtually any directional pattern can be produced.

Antenna gain is calculated differently for directional antennas. The azimuth directivity increases the gain value to correspond to the pattern. If all elements/bays are the same (the typical case), pattern multiplication can be used to determine gain. For linearly polarized antennas, the gain is simply the product of the azimuth directivity, the array factor, and the efficiency factor. The array factor is referenced to an ideal dipole. For directional CP antennas, the power distribution between polarizations must be taken into account.

$$\text{Antenna Gain (H-pol)} = g_H \times \text{array factor} \times \text{efficiency}$$

$$\text{V-pol} = g_V \times \text{array factor} \times \text{efficiency}$$

$$\text{where: } g_H = \frac{D_H \times D_V \times GA}{D_H + D_V \times GA}$$

$$g_V = g_H / GA$$

D_H and D_V are the directivities of the H-pol and V-pol azimuth patterns, while GA is the gain if the H-pol over the V-pol pattern.

Due to the gain of its azimuth patterns, directional antennas have gains which are typically 2 dB to 6 dB higher than their nondirectional counterparts of equal number of bays.

ANTENNA POLARIZATION

Horizontal and Vertical Polarization

Radio waves are composed of electric and magnetic fields at right angles to each other and to the direction of propagation. When the electric component (*E*) is horizontal, the wave is said to be horizontally polarized, as shown in Figs. 7B and 7D. Such a wave is radiated from a horizontal dipole. References are with respect to the earth plane. If the desired electric component is vertical as in Figs. 7A and 7E, a vertical dipole could be used to produce the vertically polarized wave.

Circular Polarization

When the two plane waves are equal in magnitude, and if one plane wave lags or leads the other by 90 electrical degrees, the field will rotate as shown in Fig. 7, at the speed of the carrier frequency and will be polarized circularly.

Only in the special case where the horizontal and vertical components are equal in strength with a 90° phase difference is the radiation said to be CP.

The direction of rotation shown by the vector arrows in Fig. 7 depends on the relative phase of the two components. Thus the polarization of the wave will appear to have either clockwise or counter-clockwise rotation, as shown. The FCC has set clockwise rotation as the technical standard, in order that similar sense of rotation antennas may be used for reception in the future.

Notice that in Fig. 7 the polarization rotates as the field propagates in time and space. Importantly, vertical and horizontal components are in quadrature phase. It is this rotation which enhances the signal penetrating qualities of CP, so useful in FM broadcasting.

The axial ratio as shown in Fig. 8 is that between the maximum and minimum voltage component at any orientation of the reference measuring test dipole which is placed perpendicular to the direction of propagation. An axial ratio of 1:1 (0 dB) is perfect. In

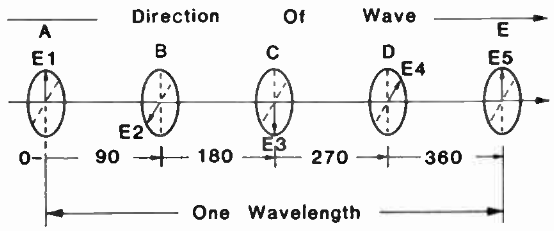


Figure 7. Circularly polarized wave propagation in one wavelength of travel, showing right hand rotation. Note vector rotation with wave travel.

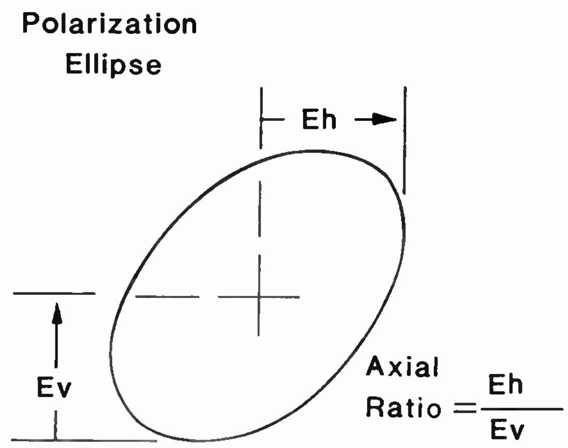


Figure 8. Axial ratio expressed in dB is the ratio of the larger polarized component divided by the smaller at any reference dipole orientation, where the maximum ratio occurs.

practice, axial ratios of 2 dB or better are considered to be excellent and commercially available. Axial ratios over 4.9 dB (1.75 to 1 voltage ratio) are considered to be elliptically polarized, a hybrid form and not as good in signal penetrating qualities as CP.

Since receiving antennas are linearly polarized, the introduction of CP does not increase the net power received since the vertical and horizontal components never occur during the same instant. Thus, CP does not necessarily mean an increase in coverage. However, the introduction of CP eliminates the requirement of the receiving antenna to have a specific polarization. Thus, CP allows more consistent coverage within the contour. Its rotating vector can penetrate areas where linear polarization is stopped, shadowed, or canceled due to out-of-phase reflections.

Summary

1. CP radiation does not increase the distance to the FCC contours.
2. Interference and allocation contours are not changed when switching from H-pol to CP.
3. The service contours are predictions based on the horizontal component of the CP radiation only.
4. The ERP from a CP station reaches it maximum H-pol power 90° *after* it reaches its V-pol peak.
5. H-pol and V-pol components do not add since they do not occur at the same instant, as shown in Fig. 7.

MATCHING COVERAGE AND ANTENNAS

Typical Class A Station Coverage

Table 3 shows the FCC predicted signal strengths for a typical Class A facility on a relatively flat plane, with the antenna center 328 ft (100 m) above average terrain (HAAT). A power of 3 kW is used. The first

TABLE 3
Transmitter power versus antenna gain.
Class A 3kW ERP — Zones 1 and 1-A
Maximum HAAT 328 Ft (100 m)

Service Miles	Distance Km	Vertical Angle	SIGNAL LEVEL IN mV/m 7.5kW Transmitter 1 Bay Antenna	1kW Transmitter 6 Bay Antenna
1	1.6	3.58	275	210
2	3.2	1.80	88	81
3	4.8	1.21	42	40
4	6.4	.92	24	22
5	8.0	.74	16	16
6	9.6	.63	11	11
7	11.3	.55	8.5	8.5
8	12.9	.49	6.2	6.2
9	14.5	.44	5.0	5.0
10	16.1	.41	3.7	3.7
12	19.3	.36	2.5	2.5
14	22.5	.33	1.8	1.8
16	25.7	.31	1.4	1.4
18	28.9	.29	1.1	1.1
20	32.2	.28	.85	.85
22	35.4	.28	.70	.70
24	38.6	.27	.55	.55
26	41.8	.27	.40	.40

two columns show the distances, with the farthest being the horizon from this height. The third column indicates the true earth angle from the antenna to the distances shown. From the elevation information, the ERP from each antenna was determined at each vertical angle. This ERP value was used to find the signal strength from the FCC F(50,50) FM prediction chart, FCC Rules Section 73.333, Fig. 1.

Under the signal level in millivolts-per-meter (mV/m) column, the predicted field strengths shown in this Table are based on the above procedure. From 5 miles (8 km) to the horizon, the signal strengths are identical. This is due to the shape of the antenna elevation pattern near the maximum.

Departure occurs as the depression angle to the receiver becomes larger. Beyond 4 miles (6.4 km), the one-bay antenna and the six-bay antenna produce nearly the same signal level.

Going towards the transmitter from 4 miles (6.4 km), the field increases in favor of the one-bay antenna. In this example, the table clearly indicates that the high-power transmitter low-gain antenna does not improve the signal strength available to the receivers beyond about 4.5 miles (7.25 km). The signal level starts to increase between 4 miles and 5 miles (6.4 and 8 km). Any increase above this level is useless because full limiting has certainly taken place in even the poorest FM receiver. (See Required Signal Strength in this chapter)

In Table 3 the same signal strength of (16 mV/m) at 5 miles (8 km) comes from either transmitter-antenna combination. This is due to the fact that the ERP power at the vertical angle of −0.74° is about the same from both antennas. The ERP at 0.0° elevation pattern will of course be exactly the same for both combinations. The field does not change measurably until observation is made beyond 1.5° from the peak 100% value in a six-bay antenna.

The signal strengths in this table were based on relatively flat terrain for an antenna 328 ft (100 m) HAAT. The true earth curvature distance to the horizon is 25.56 miles (41.23 km). Therefore the outer reaches of useful signal drop off very rapidly beyond this point in the typical Class A station.

There are no nulls in a one-bay antenna pattern. In a six-bay antenna the first null occurs at about −10°, approximately 0.37 miles (0.6 km) from the tower. Antenna arrays are never perfect so the null is never zero power. With a minimum radiation of five watts in the first null, the predicted signal would be 31 mV/m. The second null is closer to the tower and with the same five watts ERP would be even stronger in this example. So in practice there may be no need to fill in the nulls of the six-bay antenna.

Another consideration is that the nulls may fall very close to the tower and the number of people occupying the null areas may be small. Thus problems resulting from these close-in nulls would be minor.

Typical Class B and C-2 Station Coverage

The same comparisons of transmitter-antenna combinations can be made for Class B and the new Class

TABLE 4
Transmitter power versus antenna gain.
Class B, C-2 50kw EFP
Zone 1, 1-A & C-2

Service Miles	Distance Km	Vertical Angle	SIGNAL LEVEL IN mV/m 55kW Transmitter 2 Bay Antenna	10kW Transmitter 10 Bay Antenna
1	1.6	4.97°	900	140
1.55	2.5	3.25°	562	165
2	3.2	2.49°	310	230
3	4.8	1.67°	153	135
4	6.4	1.26°	92	88
4.6	7.4	1.15°	71	71
5	8.0	1.02°	57	57
7.5	12.	.70°	22	22
10	16	.55°	13	13
15	24	.41°	6.5	6.5
20	32	.36°	3.1	3.1
25	40	.33°	1.9	1.9
30	48	.328°	1.1	1.1
35	56	.332°	.7	.7

C-2 stations, operating with a HAAT of 492 ft (150 m) with 50 kW ERP. This is shown in Table 4. A 55 kW transmitter with a two-bay CP antenna would provide the 50 kW ERP, with high efficiency coaxial lines. It is compared with a 10 kW transmitter feeding a 10-bay CP antenna. The terrain flatness is assumed not to exceed ±150 ft (50 m).

Table 4 indicates that the signal levels are the same from 4.6 miles (7.4 km) out to 35 miles (56 km) under similar columns as for the Class A station comparisons. The FCC uses a receiving height of 30 ft (10 m) so the horizon is a bit further away, at 31.3 miles (50.5 km).

From the transmitter out to about 2 miles (3.2 km), the signal rises much more rapidly in the two-bay antenna than in the ten-bay, the latter being somewhat similar to a cosecant curve. There is surplus signal close in and more than is needed or can be tolerated.

Therein lies one of several problems with high transmitter power and low-gain antennas as seen in Table 4. With the two-bay antenna there is 900 mV/m at one mile (1.6 km) and 562 mV/m as far out as 1.55 miles (2.5 km). This is above, or at, the blanketing level of 562 m/Vm. (See "Blanketing" in this chapter) The high-gain antenna does not cause this type of problem under identical conditions.

The signals from both combinations are much more than necessary for present day FM receivers out to about 10 miles (16 km). There is no practical difference technically in useable signal strengths presented to receivers in the entire market area, from either antenna. There is, however, a great deal of savings in capital costs as well as operating expenses between the two combinations.

One antenna factor is not clearly indicated in Table 4. The two antennas have elevation pattern nulls. The two-bay antenna null at −30° falls 852 ft (260 m) from the base of the tower and can be disregarded. The ten-bay antenna nulls can be filled to as little as 2½% field, which will not affect its gain. This would represent a minimum ERP at the nulls of 31 watts. Although seemingly very small, it is very effective as shown in Table 5.

It is obvious that the ten-bay antenna nulls can easily be filled to produce signal levels in excess of those required. If the transmitter is located in a populated area, these high levels prevent the loss of stereo separation and noise in the SCA (if there are reflections from high level lobes in the built up areas). This problem is common to TV transmitters which produce ghosts from high signal level lobe areas reflecting into null areas. This problem is greatly and satisfactorily reduced with null fill as shown in Table 5.

MATCHING TRANSMITTER POWER AND ANTENNAS

Several available combinations of antenna gain and transmitter power will provide the necessary ERP. But which combination is the best? The choice is further complicated by the nature of the terrain in the service

TABLE 5
2½% Null fill in — 10 bay antenna.

Null	Angle	ERP	Distance	Field
First	−5.75°	31 w	4,800 Ft	31 mV/m
Second	−11.50°	31 w	2,240 Ft	70 mV/m
Third	−17.25°	31 w	1,512 Ft	109 mV/m

area. Is it all flat, some rolling hills, mountainous or a large valley? What are the regulatory limitations on the antenna supporting tower height?

Important considerations when choosing the transmitter power and the gain combination to produce a given ERP are as follows:

- Transmitter
- Feed system
- Antenna
- Transmitter tubes
- Tower
- AC power consumption

The transmitter, antenna, tower, and coaxial feed line are one-time capital costs for the station. Tube costs and commercial power use, however, are a continuing hour-by-hour cost factor. From the above it is apparent that a low-power transmitter is much more economical than a high-power transmitter. But is there a difference in signal strength?

The ERP is the product of the antenna power gain and the antenna input power. Many different combinations of power gain and input power will yield the same ERP. The azimuth pattern will be quite similar for many different antenna power gains.

The only difference in various combinations is the elevation pattern. As discussed in the Typical Class A and Class B Coverage Sections of this chapter, there is no momentous or important difference in serving listeners from very different transmitter/antenna ratios.

The signal strength at any given location is a direct function of the ERP from the antenna elevation pattern angle to that location, the height of the antenna, and the propagation path. The ERP at the pertinent angle is the product of the elevation pattern relative amplitude at that angle squared, times the maximum ERP. In practice there is no significant difference between a 3 kW ERP Class A station using a 7.5 kW transmitter and a one-bay CP antenna, or, one using a 1 kW transmitter and a six-bay CP antenna, all other factors being equal.

Normally, all the power radiated above the antenna elevation pattern to the horizon is wasted. It is the radiated power below the angle to the horizon that strikes the earth with all its FM receivers. Therefore only the radiated power towards the earth should be considered useful.

The ideal antenna system would put the same signal level from the base of the tower all the way out to the horizon. This requires an antenna whose elevation pattern is a cosecant curve, the normalized reciprocal of sine. It would be the most efficient antenna elevation pattern. Although this curve is impossible to achieve, it is approached as the antenna gain becomes greater.

ANTENNA SITE SELECTION

The transmitter location must be carefully chosen. Site economics should be secondary to the technical advantages of a particular site. Fresnel zone clearances and other factors outlined in this chapter should be considered. A site with an operating FM or VHF TV station makes an excellent source of signals to check propagation for a new station. If the existing station is FM, make certain that its antenna pattern has been optimized to provide as much circularity as possible.

A good field strength meter should be used to measure the actual signal from the existing station. Relative readings are important, not the absolute. Check for reflections as well as level changes within a short walking area of about 100 ft (30 m). Check for stereo separation. Using this information, the operation of a new station near the one being checked can be compared before moving or submitting the FCC application. The consulting engineer may find it useful to consider this information to evaluate the suitability of the new site.

High-Gain Antenna Contradictions

The many advantages of high-gain, low-power transmitter combinations to produce the ERP have been shown. Their superiority in relatively flat land applications cannot be disputed.

There is, however, the matter of unusual height over average terrain to be considered. As examples, if the transmitter is located on Mt. Wilson, in California, on a very tall building in Chicago, or New York, the elevation pattern problem can become serious. This is true particularly when there are listeners near the sites as is the case for these three locations.

Mt. Wilson which serves the greater Los Angeles metropolitan area is more than one mile (1.6 km) above most of its listeners. In fact coverage is required from 11 miles (17.75 km) out to the horizon which is −0.57° at 105 miles (168 km). Pasadena, the nearest city, is 13° below the horizon. A high gain antenna tilted down 0.5° would serve the far reaches well, but would not lay down a moderate signal at −13°.

Section 73.211 of the FCC's Rules limits the ERP for overheight antennas such as those on Mt. Wilson with 2,900 ft (884 m) HAAT. New stations using that height must reduce ERP in accordance with the equivalence calculation, so that the predicted signal at the 1 mV/m contour does not extend beyond 32 miles (52 km) for Class B stations.

In these situations a moderate gain antenna should be considered. From Mt. Wilson several existing four and five-bay antennas now provide excellent service.

TV CHANNEL 6/FM ANTENNA PROBLEM

Television channel 6 occupies the band from 82 MHz to 88 MHz. The FM broadcast band extends from 88 MHz to 108 MHz. Noncommercial educational FM stations are assigned from 88 MHz to 92 MHz. Interference can exist between the two, with the TV station viewers receiving interference from the FM stations. The FM receiver is relatively selective with a response to about 200 kHz, but the TV receiver has a bandwidth of at least 6 MHz. However, more than TV receiver selectivity is involved in this interference problem. See FCC Rules Section 73.525.

Three principal techniques can be employed to minimize channel 6 interference from FM stations: (1) collocation, (2) where collocation is not feasible, location of the FM station in an area of low population density, and (3) antenna cross polarization.

Collocation

The purpose of collocation (that is, placing the FM transmitter at the channel 6 transmitter site) is to achieve the same propagation path for both TV and FM stations, thus maintaining a nearly constant desired-to-undesired signal ratio in the service area. If possible, both antennas should be mounted on the same tower. If not, a maximum separation of 0.25 miles (400 m) between the two is still considered as collocation.

The horizontal and vertical plane radiation patterns of both antennas should be similar because the objective is to maintain a near constant desired-to-undesired signal ratio. The HAAT should be similar, thus the desirability of collocating on the same tower. The maximum ERP of the FM stations operating on this basis is in Section 73.525(d) of the FCC Rules.

Alternate Locations

The FM station may not be intended to serve the same community as the TV station, or collation may not be possible. In this event, the FM broadcaster should locate in an area of relatively low population density by imposing a limit on the population which may be included within that area where a particular undesired-to-desired protection ratio is exceeded.

Two ratios were proposed by a committee which studied this problem in 1983.[15] Their recommendation varied according to the educational station frequency separation from the channel 6 aural frequency of 87.75 MHz. In any event, the interference area should not have more than 3,000 people living in it. See FCC Rules Section 73.525(c) and (e).

Cross Polarization

Several organizations have made discrimination tests in the United States and in Europe with cross-polarized antennas. It has thus been well established that discrimination of 16 dB is to be expected in rural areas and 10 dB in urban areas between two stations one using the V-pol and the other using H-pol, and the receiving antenna being similarly polarized. This is sufficient in most cases to resolve the FM-Channel 6 problem.

While technically, cross polarization will help solve the problem, the Commissions Rules do not require it. This is left as an option for the FM applicant to use. Most TV channel 6 receiving antennas will remain H-pol, while automobile FM antennas will stay V-pol. So if the TV station remains H-pol, this interference problem may be cleared up if the FM station will switch to V-pol. See FCC Rules Section 73.525(e)(4).

Rejection Filters

The FCC believes that rejection filters installed at the TV receiver would be helpful, while others think this is not a satisfactory solution. Unfortunately, many viewers do not sufficiently understand this problem and are thus not motivated to have the necessary filter installed.

It is further complicated by the fact that many of the existing TV receivers still have balanced antenna inputs (300 ohms) and filters designed for them that do not usually provide the necessary amount of rejection. As more TV receivers with coaxial inputs are purchased by the TV viewing public, this situation could change. Coaxial (75 ohm) filters with 20 dB of attenuation of the FM signal are readily available.

COMMERCIALLY AVAILABLE ANTENNAS

There are several basic classes of antennas available for FM broadcasting. These and variations of them are made by several manufacturers in different models, gains, and input power ratings. They may be broken down into the following classes:

- Ring stub and twisted ring
- Shunt and series fed slanted dipole
- Multi-arm short helix
- Panel with crossed dipoles

These antennas have many things in common. For example, nonsymmetrical antennas are designed for sidemounting to a steel tower or pole. Radiating elements are shunted across a common rigid coax line. This has eliminated the problems associated with the older corporate feed system using semi-flexible solid dielectric low-power cables.

Shunting elements every one wavelength across a transmission line makes impedance matching simple. Bandwidth is limited by the VSWR of the individual elements and the use of an internal transformer.

With more than about seven bays, the first three of the above antennas have a more difficult task being matched and there is undesirable beam squint, since the elevation beam angle changes with frequency deviation by the transmitter. Antennas with more than seven bays are fed from or near the center, thus dividing the phase change in one-half and effectively eliminated the beam squint. Center feeding also simplifies the VSWR matching.

A means for tuning out reactance after the antennas have been installed on the tower is also common with all the antennas. Located at the input to the antenna, the VSWR tuner consists of adjustable location dielectric or metal slugs on the inner conductor of the main coax line. Several fixed-position variable capacitors, spaced one-eighth wavelength along the main feeder near the antenna input are also used on some side-mounted antennas to adjust the VSWR to very small values.

Another variety of antenna has curved radiating elements around a circumference whose diameter is determined by the number of element arms. Each radiator consists of two, three, or four such circular arms, depending on the model. Each element is fed through a shunt arrangement, and then shunted across the vertical rigid feed coaxial line.

Wideband panel antennas are becoming popular where high buildings, favorable mountain sites or high towers are available. Several firms make wideband panel antennas. Some have very wideband VSWR features in each radiator. Others with not so broad VSWR, use phase impedance compensation similar to the European scheme, which uses 90° phase quadrature impedance compensation.

Phase quadrature compensation makes it possible to cover the entire 88 MHz to 108 MHz band with a VSWR under 1.1:1 while maintaining excellent elevation, and azimuth patterns, together with very good axial ratios. Power ratings up to several hundred kilowatts are offered so that many FM stations can be diplexed into one such antenna.

Only the wideband community FM antenna design now uses a corporate feed system, while the others are shunt fed from a common rigid coax line. This corporate feed system, using air dielectric semi-flexible line at the lower power levels, is very successful. It splits the input power to many different dipoles at the correct amplitude and phase.

These four basic antenna classes are described in greater detail on the following pages.

Standard Sidemount Antennas

Standard sidemount antennas come in a variety of shapes and forms. They are currently used in the majority of applications. Their chief advantages are low cost, easy installation, availability of high gains, and low tower constraints. They are available in linear polarized configurations (H-pol or V-pol) or circularly polarized (CP).

Most sidemounts are comprised of a series of radiating elements, or bays, which are fed via a rigid inner-bay feed line. The most typical feed lines used are 1⅝-inch O.D. for applications with less than 10 kW antenna input power, and 3⅛ inch line for up to 40 kW. Most antenna elements come in high and low power versions. These antennas are mounted directly to the side of a tower or pole. Leg and face mounts are typical on tower structures.

Some manufacturers with test ranges offer sidemount antennas with custom directional patterns. The pattern shaping is accomplished by optimizing its mounting and adding parasitic reflectors which are on the order of one-half wavelength. Repeated range tests have shown that the sidemount antennas have largely distorted patterns due to feed lines, mounting structures, and other conductive media in the antenna aperture.

Sidemount antennas are inherently narrow banded. The bandwidth of a single element rarely exceeds 1 MHz. Although diplexing two stations on a single sidemount is occasionally done, the spread between stations must be small (a few MHz), and compensation tuning (such as long stubs must be used). In addition, none of these antennas are symmetrical, and each type has uncontrolled radiation from booms and feed lines in the aperture. This distortion deteriorates the antennas axial ratio and circularity.

Series-Fed V-Dipole Antennas

This antenna has similar bandwidth to its shunt fed counterpart, but the array is typically intentionally tuned high in frequency. The combination of this tuning scheme, and the internal protection of its feed allow this antenna to be somewhat resistant to light icing.

This model is larger in size and heavier than other types of sidemount antennas. Care should be taken in determining tower constraints. The antenna is typically field tuned for an optimized match. Careful placement of ceramic slugs can produce a good VSWR over the stations useful bandwidth.

Ring Radiators

There are several antennas that are simple adaptations of ring radiators and were designed and manufactured in the 1950s and 1960s for horizontal polarization. By adding vertical stubs to the ends of the radiator or twisting the ring, elliptical polarization (of sorts) is achieved.

Both the ring and the ring stub suffer from temperature variations which tend to change the spacing between the ring openings and thus the electrical capacitance and resonant frequency. The ring stub and the twisted ring are not really circularly polarized because the axial ratio varies considerably with azimuth. At best they may be said to be elliptically polarized.

Over the years the design has improved by adding a second horizontal ring and improving the feed. Reducing bay spacing reduces high axial side lobes. The antenna has very good circularity in free space, but like other types of sidemount antennas, it is strongly affected by its support structure and feed line.

The radiation patterns are strongly affected by the tower mounting environment. Being of relatively high Q design, they are more susceptible to detuning because of icing. Radomes and electrical deicers are available to overcome this problem. While the icing problems may be overcome, pattern optimization is not offered for these antennas.

Ring-Stub Antennas

The H-pol radiation from these antennas comes from the ring portion whose plane is parallel with the earth. There is a minor lobe from each radiator, which is strengthened with vertical stacking for additional power gain. This nadir-zenith lobe is the result of 360° stacking on the rigid coax feed line. It reduces the gain and presents a lobe at the tower base which is detrimental to low level audio equipment and personnel located in a building at the base of the tower.

In order to keep the cost down, like the twisted ring antenna, the ring-stub is manufactured in several radiator-to-radiator spacings across the FM band. This results in some minor beam tilt up or down depending on the frequency. Most higher priced slanted dipole and helix antennas are spaced exactly 360° and are usually tested to assure this spacing during production.

Shunt Fed Slanted Dipole Antennas

The slanted dipole antenna in its present configuration was developed and patented in 1970.[16] It consists of two half-wave dipoles bent 90°, slanted and fed in-phase.

The slant angle is critical as it is the factor which determines the ratio of vertically and horizontally polarized radiated power. The phase point center is at the feed insulator on the dipole support arm as shown in Fig. 9. When fed through a vertical support pole on which the antenna was mounted during initial development tests, the axial ratio varied less than 1 dB.

The commercial adaption uses a horizontal boom containing a step transformer. This boom supports two half-wave dipoles in which the included angle is 90°. The two sets of dipoles are rotated at 22.5° from the horizontal plane. Two opposite arms of the dipoles are delta matched to provide a 50 ohm impedance at the radiator input flange. All four dipole arm lengths may be adjusted to resonance by mechanical adjustment of the end fittings. Shunt feeding, when properly adjusted, provides equal currents in all four arms resulting in excellent azimuth circularity.

Figure 9. Three bay Cetec model JSCP-3 non-symmetrical FM antenna mounted on the tower leg. The guy cables are insulated fiber glass rods near the tower legs. (Photo courtesy Cetec Antennas)

Short Helix Antennas

A relatively recent asymmetrical radiator is the four-arm shunt fed helix. By using four dipoles, curved so that their circumference is about one wavelength, a CP antenna is produced.[17] Each dipole is about one-half wavelength and is shunt fed. These are supported on a four arm structure, one end of which is tied to the supporting structure. The dipoles overlap so that the current flow around the circumference is circular. The four feed arms are connected in shunt and the feed impedance is quite low, but may be brought up to useful values with an internal step transformer.

The CP quality of the four-arm side-fire short helix is good. Three and two arm models are also available, but their axial ratio is not as good as the four arm. Pattern circularity is ±1 dB for the four arm, together with an axial ratio of about 3 dB.

These radiators are stacked about one wavelength apart on a rigid coax feed line to obtain the necessary power gain. Like other asymmetrical FM antennas its patterns are strongly affected by the supporting structure. See "Pattern Optimization" in this chapter for the need and methods to circularize the azimuth pattern.

Electrical deicers using the stainless steel dipole arms as one half of the heating circuit are available. Heat is created by passing a large current at low voltage through each arm from voltage dropping transformers placed at each bay level. Plastic radomes are also available to keep snow and ice off the sensitive VSWR parts of the antenna.

Twisted Ring Antennas

This type consists of one or more rings, which have been partially twisted so that the open ends of the ring are about 10 inches (25 cm) apart. One semi-circular arm of the ring is fed with a small loop or by a direct tap on that arm. A number of these rings are fed in the same manner as the ring stubs, *and have the same zenith-nadir lobe problem.*

The mechanical twist is not the same when viewed in all the azimuth directions. Therefore the current is not the same, with the end result that in some directions there is much more elliptical radiation than in others.

These antennas are very simple and relatively inexpensive for single frequency use, but have some serious operational limitations for CP operation. They do not have the same signal penetrating effect as the slant dipole, short helix, or the flat panel antenna type of CP antennas.

Short Helix—Multi-Arm Antennas

The number of arms may be increased to four instead of the two in the slanted dipole variety. To provide CP, the arms are curved to form a one wavelength circumference. These short multi-arm helices are also stacked in the conventional manner, like the others in this series for power gains as desired. This design uses 2 gamma feed straps to feed all the elements in phase. This antenna is shunt fed, and is arrayed and mounted similarly to the slanted vee dipole antenna.

The azimuth pattern of all these non-symmetrical antennas is affected by the supporting steel structure. With pattern optimization, the pattern can be made quite omnidirectional. See "Pattern Optimization" in this chapter.

Series Fed Antennas

A similar arrangement of arms supported by a T arrangement may be series fed. That is, part of the outer end is insulated from the rest of the dipole and fed across the insulated break. To allow for adequate power handling capacity and to increase the VSWR bandwidth, three-inch (75 mm) diameter tubing is used.

The antenna has an VSWR bandwidth of about 1%, so it makes an excellent single channel FM antenna. The antennas are usually mounted on the side of a supporting tower or pole, and stacked vertically to achieve required power gain.

This antenna has greater wind loading than the shunt fed version due to its larger element diameters necessary to achieve useable VSWR bandwidth. It presents considerable large amounts of ac power for electrical deicing. Plastic radomes also present additional wind loading.

Flat Panel Antennas

Panel antennas for CP FM broadcasting are relatively new in the Western Hemisphere, although H-pol and V-pol have been used in Europe since the mid-1950s. This antenna was developed in Europe to provide a wide bandwidth for several collocated government stations without the need to change antennas when a new channel was added or the operating channels were changed from time-to-time.

Panels are from 7 ft to 8 ft (2,100 mm to 2,450 mm) square in the flat configuration. In the cavity style they are about 8 ft (2450 mm) in diameter and about 3 ft (1,000 mm) deep.

A heavy metal frame is often used over which large diameter wire mesh has been welded. The wire mesh screen openings vary from 4 to 12 inches (100 mm to 300 mm). Electrically they are considered nearly solid metal. These openings produce relatively low wind loads. The entire flat frame or cavity is strong enough to support a man on its mesh openings. Some manufacturers hot dip galvanize their steel after fabrication; others use stainless steel construction.

For FM use, two crossed dipoles are used as the illuminating source for each panel or cavity. Each dipole is fed in phase quadrature, that is one dipole receives its peak current 90° after the other, to produce CP. A typical set of electrical and mechanical specifications for a CP eight-bay cavity community antenna is shown in Table 6.

Flat panel antennas are typically sidemounted on large face size towers. The screen panels greatly reduce interaction and distortion between the antenna and tower. The panels are directional, thereby requiring 3 or 4 panels to be mounted around the tower to achieve acceptable azimuthal circularity.

These antennas are usually branch fed, and often the array's top and bottom halves are fed separately. This allows operation of either half of the antenna separately when it is necessary for maintenance or repairs. Circular polarization is achieved on each panel by feeding two perpendicular dipoles 90° out-of-phase. This phase offset helps this antenna achieve usable bandwidths on the order of 10 MHz.

By pulling the dipole back on its feed support arms, the arrowhead shaped dipoles control both V-pol and H-pol azimuth patterns. Rotating the dipoles 45° with the earth-ground reference improves the polarization ratios even further.

Round dipoles made of tubing as large as 6⅛-inches

TABLE 6

Typical measured community antenna performance.

Operational frequency range	88 to 108 MHz
Safe RMS input power rating	200 kW
Power gain ratio, each polarization	4.4 (6.43 dB)
Maximum VSWR any frequency between 88-108 MHz	1.1:1
Elevation pattern beam tilt	−0.5°
Polarization	Right hand circular
Axial ratio	Better than 2 dB
Azimuth circularity Vpol or Hpol	Better than ±2 dB
Antenna dead weight, less than	7,000 Lbs (3,183 kgs)
Active wind load, RS-222-C 50/33 PSF	8,000 Lbs (3,636 kgs)
Antenna input flanges, two, size	6-1/8 inch
Number of bays (stacks)	Eight
Radiator type	Circularly Polarized Cavity

(155 mm) in diameter are used along with a single line quadrature feed. This combined arrangement makes an excellent wideband CP panel to cover the entire FM band. Power splitters, dividers, and cables, along with a number of these panels, completes the antenna.

Even on large face towers, circularity in the H-pol can be quite good, on the order of ±2 dB. On standard configurations, the V-pol pattern is quite different. As a result, the axial ratio of this antenna ranges from good at some azimuth headings to rather poor at others. This is because the azimuth pattern of a H-pol dipole is like a figure eight, or cosine function, while the pattern for a V-pol dipole is not directional in its azimuth plane. Therefore, each polarization will react quite differently when mounted in front of a panel.

Dipoles on these panels are often mounted at 45° referenced to the ground. This has no effect on the axial ratio or pattern performance, it instead is done for tuning considerations to compensate for mutual coupling.

Variations to this design have reduced the differences between the patterns of the polarizations. These techniques are effective for applications requiring only a few MHz of bandwidth. One method optimizes the angle of the dipole bend as well as its distance to the panel. This design requires three panels to be mounted around a tower. Axial ratio and pattern circularity are improved at the cost of system bandwidth.

Another method uses four dipoles forming a square shape in front of the panel. By adjusting the spacing between the dipoles, the beamwidth of a panel can be controlled. Over a small bandwidth, the pattern performance is greatly improved. It is necessary to mount four panels around a tower for a circular pattern. A large amount of panel interaction and leakage are severe design limitations.

For projects that require wider bandwidth, skew mounting is often used. This physical configuration allows the panels to be fed in mod 1 (0°, 90°, 180°, 270° phase for 4 around) which can increase the bandwidth of the system at the input. Although skew mounting deteriorates pattern performance, the increase in bandwidth extends its applications.

Cavity Backed Panel Antennas

The use of a cavity screen instead of a flat panel has greatly improved axial ratios. The cavity acts as a resonator with little leakage towards the tower. The shape of the azimuth pattern in each plane becomes both controllable and symmetrical. System bandwidth is improved over the flat screen design. By adjusting the diameter of the cavity structure, beam widths can be altered to meet specific requirements. Mounting three cavities around a tower gives good pattern circularity. Axial ratios usually range from good to excellent.

The cavity antenna uses the reflective properties of the flat screen panel. In the cavity however, the illuminating dipoles are flat instead of round and all four arms are parallel to the plane of the cavity. Like the flat panel with its round dipole supporting balun, the cavity also holds its flat dipoles with a double coaxial balun.

The dipoles in the cavity get their wide VSWR bandwidth through the sleeve dipole principle.[18] Capacity is provided by a metallic ring close to all four dipole arms placed between them and the back of the cavity. Circulating surface currents flow on the dipole arms, which result in evenly radiated patterns in all polarization planes. The bandwidth of a single cavity can cover the full 20 MHz band with a VSWR better than 1.1:1. Therefore it is not necessary to skew mount these antennas for bandwidth considerations.

This antenna has the advantage over some other designs of greater VSWR bandwidth. It is considered closer to state-of-the-art due to better elevation and azimuth pattern control of both planes of polarizations by the shape of the cavity.

Cavities, and flat panels can be modeled using a computer. Factors such as tower size and orientation, as well as the phase and skew of the elements, can be modeled to determine optimum mounting and feeding. This is useful in projects which require a directional pattern. A station should take pattern constraints and desires to an antenna manufacturer to find out what is feasible.

Crossed Dipole Theory

Common to the flat panel and the cavity is the operation of the dipoles which generate CP. The dipoles are fed currents in phase quadrature, through a coaxially balanced balun, which provides equal currents to all four arms of the two dipoles. They excite the entire cavity or flat panel with a rotating RF field in a plane parallel to the dipoles. The RF field is thus CP and may be ideally represented by a rotating vector of constant magnitude revolving one revolution per wavelength of propagation distance. It is right hand polarized as the field rotation is clockwise as viewed from behind the screen, looking toward the direction of propagation, if the phasing between the two crossed dipoles is properly made.

Radiation patterns, associated beamwidth, and directivity are determined to a large extent by the size of the cavity or flat panel. The geometry of the dipole has less effect than the reflector size. The size and shape of the dipole controls the antenna impedance and the VSWR. The screen panel, be it flat or a cavity, fulfills the following five important electrical functions:

1. Isolates the radiating elements from the tower or the mounting structure, and reduces mutual coupling.
2. Provides sharper beamwidth and more gain than achievable with the dipoles alone.
3. Furnishes pattern control so that the beamwidth is nearly equal for both horizontal and vertical plane polarization.
4. With an effective balun feed system, the crossed dipole radiated pattern phase is very uniform as the amplitude changes normally with azimuth.

5. Computer aided designs are easily achieved in production for various width towers because the pattern is simply pure electrical geometry.

Antenna Element Spacing and Downward Radiation

Most FM antennas have elements that are spaced one wavelength apart (9 ft to 11 ft) for reasons such as gain considerations, mutual coupling effects and ease of feed design. There are cases, however, which require a different element spacing. High levels of downward radiation is the most common reason, although considerations such as aperture constraints and beam shaping also utilize nonwavelength spacing.

Radio frequency radiation (RFR) safety levels must be considered in nearly all site locations. Power radiated in the lowest sidelobe can cause a variety of problems, of which human exposure levels are most critical. Guidelines set by ANSI require that the level of radiation not exceed 1.0 mW/cm^2 (for FM) in areas of free human access. Depending on factors such as antenna type, height of antenna above ground level, and ERP, a station may require an antenna with short-spaced elements to meet the safety level.

$$S = \frac{(0.64)\,\mathrm{EIRP}}{\pi R^2}$$

where:

S = Power density
EIRP = Effective isotropic radiated power
R = Distance to the center of radiation

The above equation is taken from FCC OET Bulletin 65, which outlines a method of computation to predict the ground levels of RFR. Values of field strength can be obtained from the antenna manufacturer. If the calculated level points to a possible exposure problem, reduced inter-bay spacing might eliminate the problem. Realize that no equation can provide exact results. The only accepted method for correctly determining levels is by direct measurements at the site. Factors such as poor tower RF grounding and noninsulated guy wires may make actual levels substantially higher. See Chapter 2.9, "RF Radiation Hazards," for further instructions.

When antenna elements are arrayed, the resulting elevation pattern contains lobes and nulls. The furthest sidelobe from the horizon typically peaks between 70° and 90° below the horizon for full wave spacing. This lobe occurs since the physical path results in no phase cancellation in that direction, and thus each elements downward radiation is additive. Shortening the spacing changes the difference in the elements physical path length, and results in some phase cancellation.

Half wave spacing, for example, greatly suppresses the levels of the side lobes, while increasing the width of the main beam. Despite the lack of power in the side lobes, the extra width of the main beam causes the pattern to be less directive. Thus this exchange results in an overall gain reduction on the order of a third. Spacing the elements 0.8 wavelengths apart improves sidelobe suppression and the gain of the antenna is not greatly affected.

Short spaced antennas are fed either by a shunt or branch feed system. For half-wave spaced antennas, a shunt line delivers each element 180° out-of-phase with the next element. This phase distribution problem can be overcome by flipping every other element upside down, thus inducing a 180° phase shift in the feed. Other spacings, such as 0.8 wavelengths, require a branch feed, which can deliver equally phased signals to each element independent of spacing.

Problems With Sidemounted Antennas

Single station FM antennas are typically sidemounted on a pole or tower. Unlike panel antennas, the support structure greatly effects or distorts the radiation pattern. The resultant pattern may have large peaks and nulls that can result in coverage and reception problems.

In addition, the V-pol and H-pol patterns react quite differently to these distortions. Due to the geometric complexities of the CP radiating elements and tower structure, no computer model exists to accurately predict pattern effects. Therefore, the use of test ranges are required to determine how an antenna behaves when mounted on a tower section similar to the one on which the antenna will eventually be used.

For nondirectional stations, a test range can determine the proper mounting of an antenna to achieve an acceptable circularity. Depending on the tower size, the depths of nulls can be greater than 10 dB. An optimized mounting configuration can make the nulls less significant and oriented in areas where service to the primary coverage area is not hurt. Parasitic reflectors are often used to improve the circularity of nondirectional antennas.

When the top spot on a tower or structure is available, pole mounting is often preferred. A pole provides a stable and symmetrical support structure which has low interaction with the horizontally polarized component. In combination with the feed line, the pole typically induces a null in the V-pol pattern, directly opposite of the elements. Proper orientation of the element can reduce the effects of pattern distortion.

Mounting an antenna on the side of a tower can produce unpredictable results. The position of the peaks and nulls vary greatly from the orientation of the elements. As the face size of the tower increases, the distorting effects magnify. To compensate, many stations use smaller sections of tower at the top, where they plan to install the FM antenna. Eighteen-inch and smaller face towers tend to produce good results. With careful planning, the use of a 24 inch or larger face tower can also be successful.

Note that the patterns of each polarization react differently, and thus axial ratios can be quite poor. See the following section on pattern optimization.

PATTERN OPTIMIZATION

Single-station FM antennas are usually sidemounted on a pole or tower. This is economical and it frees the tower top for other possible uses. Unfortunately the pole or tower tends to distort the radiation pattern, seriously affecting station coverage in some directions.[19]

This problem can be serious if the FM antenna has been randomly attached to a support tower. FM antenna makers do not manufacture and sell towers. A few have made supporting poles on which the FM antenna has been affixed, adjusted, and pattern tested. TV antenna makers, on the other hand, always make the antenna as a complete self-supporting structure to be mounted on top of a support. They are usually not faced with this sidemounting problem. The logical but more expensive solution would be to make the FM antenna a self-supporting structure just like TV antennas.

Improper FM antenna sidemount installation on a tower can cause serious pattern problems. Measured patterns have indicated that, in some cases, the maximum radiation can actually be in the opposite direction from the desired direction.[20]

Need To Optimize

It is not wise to gamble with an FM station's antenna coverage. Nulls may be toward important service areas. Nulls as low as 1% of the RMS power have been measured with towers varying in width from 18 to 120 inches (0.5 m to 3 m).

Another problem is that with nulls come lobes. Lobes as great as 9.8 dB over RMS have been found. When used without pattern optimization, this lobe would produce an ERP in a given direction nearly ten times the FCC licensed value.

Translating this to a 50 kW ERP station there would be radiation in some directions of only 0.5 kW and others with 477 kW. This is a maximum to minimum ratio of 29.8 dB, and clearly not acceptable.

With CP came additional problems as the H-pol and the V-pol ratios are not always the same and vary moderately in any given azimuth. This ratio can be as great as 15 dB and must also be addressed in order to resolve the horizontal plane circularity problem. The axial ratio could be degraded causing the V-pol radiation in certain directions to be much stronger than the H-pol. This violates FCC's requirement that with CP, the V-pol must not be stronger than the H-pol component.

Section 73.316 of the Rules covers FM antennas but does not specifically address this problem of azimuth circularity. In fact the FCC assumes that FM nondirectional broadcast antennas have perfectly circular horizontal radiation patterns.[21] In actual practice, they seldom do.

In order to produce a horizontal plane pattern which even approaches a circle requires considerable work by the firm making the antenna. Since it is nearly impossible to produce a circular pattern with a nonsymmetrical sidemounted antenna, the term and technique, optimization (to do the best possible) is in common usage now.

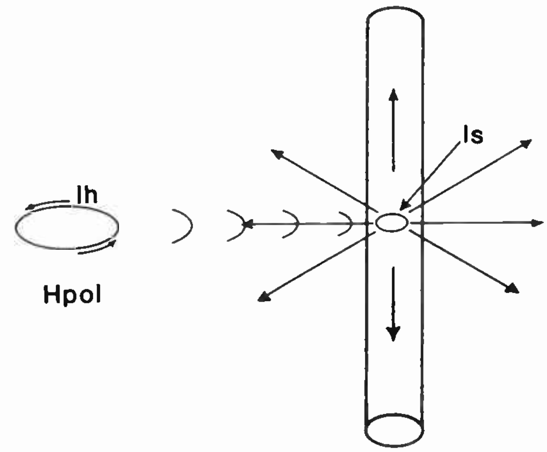


Figure 10. The effects of supporting steel towers or poles on one side of a non-symmetrical antenna are shown as horizontally polarized currents Ih flow on supporting members as Is and reradiate in all directions.

Theory of Optimization

Fig. 10 indicates how energy from the horizontal loop representing H-pol is intercepted by a pole and reradiated. Similarly, in Fig. 11 energy from the vertical dipole is intercepted and reradiated as V-pol.

In the first case, the pole diameter is small in wavelength in the direction of the electric field, thus, scattering is minimal and not much H-pol radiation can be expected. However when the V-pol dipole excites the pole, a large amount of energy is intercepted (Is) and reradiated because the large dimension of the pole

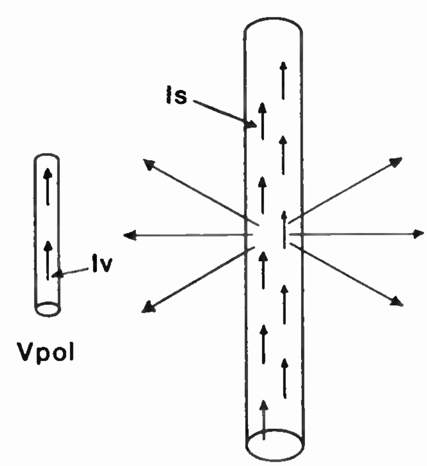


Figure 11. Vertically polarized current Iv flow over the supporting structure members as Is and reradiate in all directions.

(length) is parallel to the electric field. A similar effect is produced by the vertical transmission line which is common to the antenna itself, (IV in Fig. 11). The result is appreciable distortion of the vertically polarized azimuth pattern.

Fig. 12 shows the resulting H-pol and the V-pol patterns. The pole and/or vertical coaxial line have transformed the V-pol pattern from circular to a cardioid, while the H-pol pattern remains essentially omnidirectional. The null of the cardioid can be more than 7 dB down from the RMS value. This phenomenon is well known, and as a compromise, broadcasters generally install the antenna on the side of the tower support structure facing the main service area. There are many exceptions to this, as some measured patterns on triangular towers of standard construction have shown.

In Fig. 12 the V-pol is much stronger than the H-pol in the favored direction. In the opposite direction, there is little V-pol. The power ratio of H-pol to V-pol is 16 times. This makes for a very poor CP antenna in some directions.

Tower and poles under about 2 ft (0.6 m) in cross section will exhibit the same effects on the antenna patterns, as in Fig. 12. Towers greater than this size obviously will increase the complexity of scattering effects. Three or four tower legs, the horizontal and diagonal cross members, transmission lines, ladders, tower lighting, and deicer conduits, all will be excited by the vertical and horizontal currents from the radiators. All these surfaces will reradiate and affect the horizontal plan patterns.

In contrast to the simplicity of the antenna on the side of a pole, the tower supported antenna may be mounted on the face, or on a corner, at or between horizontal cross members, or tilted at various angles compared to the tower—all multiplying the complex factors affecting the patterns.

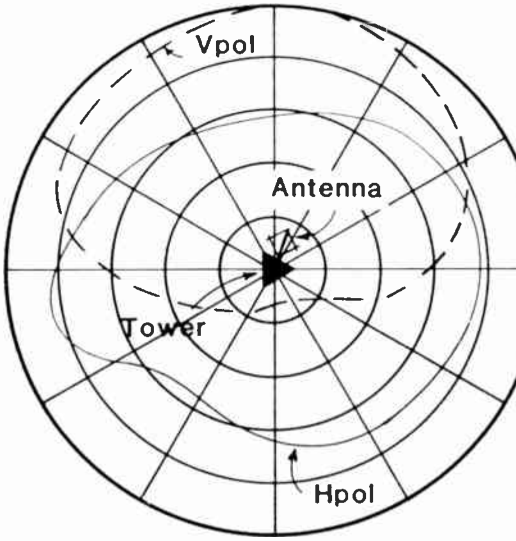


Figure 12. Combined results of Hpol and Vpol reradiation distort both patterns, producing more than 17 dB variation. This poor Cpol pattern should be pattern optimized to smooth out the azimuth pattern.

Optimization Methods

The most popular technique to achieve the desired pattern is through the use of Yagi antenna principles, wherein parasitic elements are placed in the field of the radiator to modify its radiation pattern. For example, a shortened dipole (director) placed in close proximity to a radiator reinforces radiation in the forward direction, and suppresses the signal in the opposite direction. If the parasitic element is longer than the radiator (reflector), the effect is reversed. The signal is suppressed on the side of the parasitic element and reinforced in the direction of the radiator.

Similarly, parasitic elements can be used with FM antennas mounted on the sides of towers or poles to produce pattern changes. As discussed here, both directors and reflectors may be used. Both are frequency sensitive. The effects of the supporting structure are also frequency sensitive.

Therefore, an arrangement of parasitic elements for a given FM frequency will not necessarily be the same for another, nor will the pattern be the same for a given arrangement, if it is moved up or down the tower by as much as 1.5 ft (0.5 m).

The resulting patterns cannot be predicted. There are many factors which affect the horizontal plan pattern. Only by actual antenna pattern range testing can the patterns be adjusted and properly measured. Therefore the cost for doing this is high, since it is time consuming, and requires qualified antenna technicians and the use of an antenna range. In addition, the final parasitic arrangement must be permanently fabricated and installed. However, the results are well worth it.

Pattern Service

There are two basic types of pattern service furnished by several antenna manufacturers in the United States. FM antennas may be adjusted for the best omnidirectional pattern possible or they may be adjusted to proven minimum ERP values in particular azimuth directions. The minimum required values, plotted on a polar chart by the broadcasters may be combined with the tower orientation. Using the customer's make and model tower, two or more bays of the antenna are fabricated, installed on a section of the tower, and put on the test range. Adjustments are made such as leg chosen, distance from the leg, and the orientation of the antenna with respect to that leg. Parasitic elements are then used to further improve and shape the pattern, so that the minimum ERP values will be achieved in the customers service area as given with the order on the polar plot.

For example, a Class A station may wish that a minimum of 3 kW be radiated in a pie from 90° to 120° and the remainder of the azimuth be no less than 1.5 kW. This would then require the antenna technicians

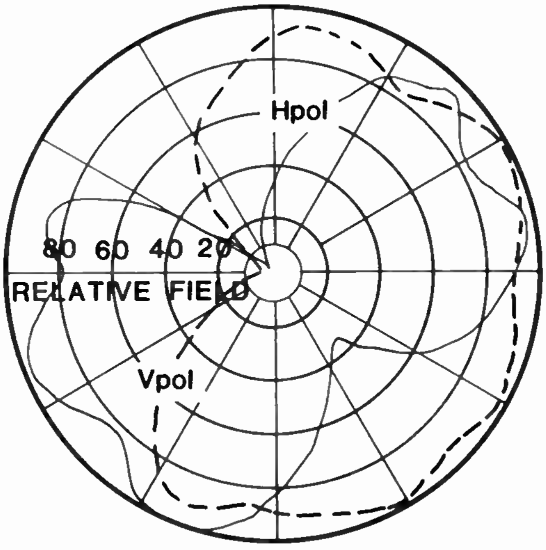


Figure 13A. Measured non-symmetrical Cpol pattern of tower side mounted antenna. Vpol variation is ±15 dB, while Hpol is ± 12 dB. The axial ratio was 24 dB. Antenna patterns are very poor.

to achieve a pattern without any field voltages less than 70% and that the vectors between 90° and 120° to be 100%. This sort of work has been done by many antenna manufacturers. See Figs. 13A and 13B for a typical before and after optimized pattern.

Various methods have been used to optimize FM antenna azimuth patterns. Some firms use models at twice the operating frequency. Others use theoretical methods, backed up by experimental proof. The final optimized antenna is match marked on the tower sections so that it will be assembled exactly as it was made and tested at the fabricator's plant, and tested on the antenna range.

A complete set of installation prints must be provided so that the antenna is assembled exactly as tested, with all the correct locations, and angles of all the parasitic elements.

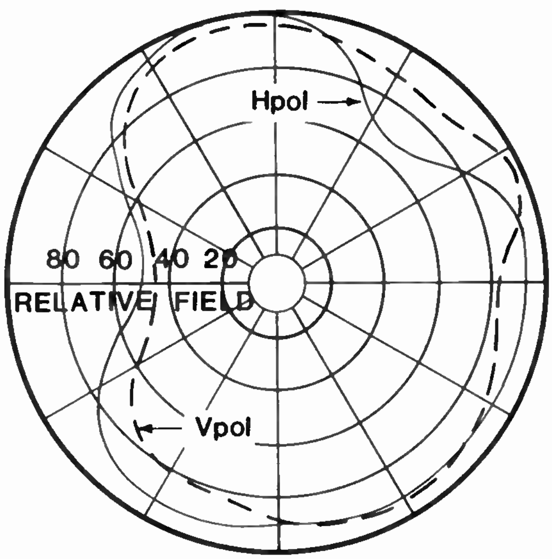


Figure 13B. Same antenna as 22A but much improved after pattern optimization. Hpol variation is ±3 dB and Vpol is ±3.4 dB. Axial ratio is quite acceptable at 2.9 dB.

WIDEBAND COMMUNITY AND MULTI-USER ANTENNAS

The most significant recent developments in FM antennas have come in broadband designs. Flat panels, cavities, helicals, and a new sidemount have given broadcasters a range of choices to fit their application. Factors such as pattern requirements, tower constraints, and system budget can narrow the options.

In order for an antenna to be useful throughout the 20 MHz FM band, its operation must be the same on any frequency. The VSWR on 89 MHz for example, must be just as good as on 108 MHz. The CP azimuth pattern should remain the same on one end of the band as the other, as must the axial ratio. This is a much more rigorous requirement than placed on the single-channel slanted dipole or the ring stub.

In the wideband antennas, several factors go together in order to meet these severe requirements:

- Basic wideband dipole radiators
- Screen-panel pattern control
- Quadrature phase distribution

By using these three principle parameters in a wideband antenna the radiation pattern, VSWR, and the gain can be nearly the same on any channel within the FM band.

Several methods are used to make the VSWR of the crossed dipoles as good as possible. The dipoles are usually fed with a folded balun or the split-tube type balun.[22] This improves the impedance match, phase as well as amplitude linearity of the resulting azimuth pattern.

The length to diameter (or width) ratio is usually about three. This not only reduces the Q but also increases the voltage flash over levels. The low Q also increases the bandwidth by decreasing the rate of reactance change with frequency.

A natural factor aiding the VSWR problem is the fact that in order to obtain CP from two crossed dipoles, there must be a phase quadrature of the two currents feeding the crossed dipole. The two reflections, as a result of VSWR, return back to the phasing device 180° out-of-phase with each other. Being the same amplitude, they cancel.

All of the above factors, plus two or three more levels of quadrature reflection cancellations, bring the overall system input at the antenna to under 1.08:1 across the band. This cancellation technique eliminates the need for electrical deicers or plastic radomes, as the VSWR is not affected by moderate ice coatings.

However radomes may be necessary with flat panels, to physically protect some radiators from falling ice.

These and other factors all contribute to make the panel-cavity antenna the best possible for either single channel or community antenna use.

Top Mount Antennas

There are a few broadband designs which require top mounting on a pole. The first design incorporates a series of dipoles which are mounted on a pole. Each dipole is branch fed individually. By rotating the mounting scheme, circular patterns can be achieved. The rather small radius between dipoles keeps nulls from crossover points to an acceptable level. Large bandwidths and good patterns have been measured on this system. The considerable amount of feed cables that run through the aperture, however, can cause problems if not properly RF grounded or shielded. Line damage, failure, and pattern distortion are possible results.

Spiral (or normal mode helical) antennas are now being used for FM applications. This traveling-wave-type antenna has bandwidth on the order of 15% to 20%. Fed from the top, a series of wires are wrapped around the conductive pole over an aperture of 2 to 3 wavelengths. The pitch angle of the wire wrap controls the pattern characteristics. Pattern circularities of ±1 over a 20 MHz band have been measured. Axial ratios better than 2 dB are considered typical. The feed system is enclosed by stainless steel feed cans which eliminates pattern distortion by the feed system, a common fault of other top mounts. In icing regions, heating elements inside the radiating wires stabilize its performance.

Multi-Channel Sidemount

The most recent design in broadband FM broadcasting antennas is the sidemount. This design incorporates dual baluns which insure a balanced and symmetric current excitation. This balance eliminates spurious radiation from currents on booms, feed lines, and other nearby conductors. The feed arms are of the skewed vee dipole configuration. Excellent circularities and axial ratios have been measured in free space. Interaction with feed lines and support structure still distort the pattern, though to a lesser extent.

A single FM element has been measured to have a bandwidth of 5.5 MHz with a VSWR of 1.1:1. This bandwidth does not limit the spread of two combined stations. By centering the band edges, the antenna can tune two stations over 8 MHz apart with a VSWR of 1.1:1 over each channel. This comes at a cost of poor performance at center frequencies, thus three stations over 8 MHz is not achievable. As a sidemount, this antenna has substantially reduced costs and tower constraints over a panel type.

Community Antenna Economics

The community antenna fits best in multiple station service. This allows sharing costs so all parties benefit from a superior antenna that each station independently could not economically justify.[23]

If enough planning is done in advance, it may be possible to install all the FM stations of one community on one tower, at considerable savings to all users. Some preclusions are lack of adequate mileage separations, the existence of excellent facilities, and FAA tower height limitations.

The break-even point appears to be with four stations. When five stations are involved, there is a 20% reduction in cost to each of them when compared to the costs of putting up their own individual single channel antenna and tower at the same height. See Table 7 for a break-out of costs for a wide-band community antenna system.

Community Antenna Examples

The costs used in Table 7 include an eight-bay CP omnidirectional community antenna with a gain of 4.40 mounted at the tower top. It also includes a dual run of 6⅛-inch (155 mm) coaxial transmission line, for upper-lower half feed system. The 1,020 ft (310 m) tower height puts the antenna center at approximately 984 ft (300 m) maximum HAAT level on flat terrain.

The guyed, lighted, and painted tower is EIA-RS-222-C rated for 85 mph (137 km) ground wind areas. The number of diplexers are equal to the number of stations so there is always one wideband port left for emergency use if one station loses its diplexer. The dual coax line permits using one-half of the antenna in the event of a problem in one line or one-half of the antenna.

The single channel comparison antenna system also has the same 984 ft (300 m) tower, but with one run of 3⅛-inch (79 mm) line. Diplexers are not used. An eight-bay single channel slanted dipole antenna with a power gain of 4.30 is sidemounted and has electrical deicers. Its pattern is optimized for a smooth circular azimuth pattern. The VSWR is field adjusted to be under 1.1:1 for ±250 kHz at the station carrier frequency.

All outdoor installation work was included in the costs shown. Both sets of costs have one antenna field technician for VSWR checking and adjusting the diplexers in the community example. Tower founda-

TABLE 7
Wide-band panel antenna costs (1985).

Number of Stations	Total Cost of System	% of Single System Cost
4	$1,410,000	Break Even
5	1,460,000	−20%
6	1,510,000	−31%
7	1,560,000	−39%
8	1,640,000	−44%
9	1,690,000	−49%
10	1,740,000	−52%

tions and buildings are not included as this is a variable cost item, and may be done by the owners.

In 1984, a group of Houston, Texas broadcasters formed the Senior Road Tower Group, and installed a 2,049 ft (625 m) tower supporting a twelve-bay community FM antenna system, with its HAAT at 2,000 ft (610 m).[24] This height permits the maximum service allowed.

Two runs of 8³⁄₁₆-inch (208 mm) diameter coaxial lines are used to feed the antenna in such a manner that power in both the lines causes right hand CP from the antenna for nine FM stations.

The nine stations use one diplexer each, all of which are housed in one 2,400 square foot (223 square meters) room. The 10 port modular diplexer has a total power handling capability of 350 kW. The insertion loss for each station is 0.80 dB (17% loss). The isolation between the various transmitters meets FCC spurious emission (intermodulation) requirements.

All the diplexers are monitored at a central operating rack which indicates each diplexer's forward, reflected, and rejected power. This permits trouble shooting in an orderly and rapid manner. Electrically operated coaxial switching permits each station to be connected to the dummy load for individual testing. Air conditioning and chilled water are used to remove heat produced during operation.

The tower also holds one UHF-TV antenna below the FM antenna, plus three levels of two-way radio communications antennas at the 800 ft, 1,200 ft, and 1,400 ft (244 m, 366 m, and 427 m) levels. In addition there are individual single-bay CP antennas for each station, fed with a 3⅛-inch (79 mm) line.

The income projected from the use of the two-way radio antenna facilities helps defray the operating costs of the entire plant including the electric power bills.

This community FM antenna project demonstrates the technical feasibility for a large system with nine FM users. It also demonstrates that the full intent of FCC Docket 80–90 can be met to the broadcasters satisfaction, as well as its wider range of listeners.

The technology is now available to combine multiple FM stations efficiently and without problems. The difficulty is in bringing broadcasters together to start, plan, fund, and complete such a large undertaking.

Technical Advantages

Besides the financial advantages cited under the economics heading, there may be the competitive advantage of protecting the channel classification and using the same height antenna as the competitor. Other advantages include its emergency upper-lower half feature for transmission line or antenna half backup. The flat VSWR curve is highly useful for SCA and stereo operation. Stations sharing this type antenna will all experience less intermodulation interference than if they had separate antennas, but closely placed.

STATION DIPLEXING

Diplexers are passive devices used to combine the power of two or more stations and feed the combined power to a common transmission line and/or a common transmitting antenna. This system of utilizing one well-sited, high-quality antenna has become popular, convenient, and economical.

Wideband panel antennas, although expensive for use by one station, are very cost-effective for two or more stations. These antennas maintain their omnidirectional horizontal plane patterns and VSWR throughout the FM broadcasting band from 88 MHz to 108 MHz. Thus, they make the ideal antenna for multi-station diplexing.

Diplexer Economics

Diplexers permit several stations to be combined into a single master antenna. New channels at any power level may be combined in any order by connecting another diplexer in tandem to the previous unit. The first channel may be fed into a wideband antenna without a diplexer. When two stations are to be combined, the first diplexer is put into service. When the third station is added to the system, a second diplexer is required and so forth. Fig. 14 shows this arrangement.

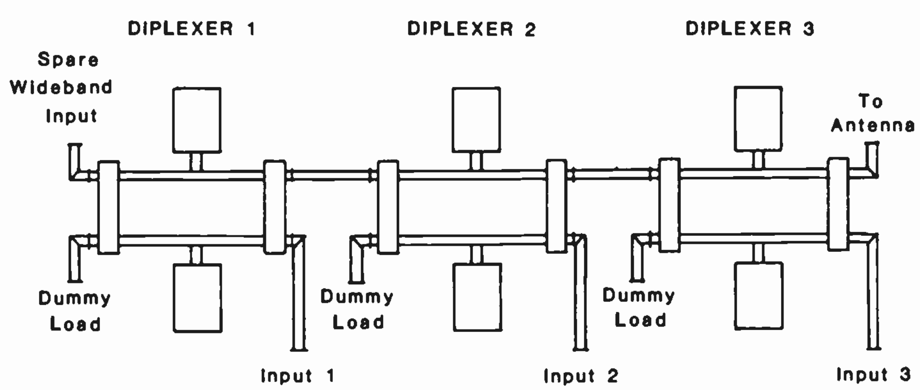


Figure 14. Three diplexers connected to provide combining facilities for three different stations for one community antenna. Note spare emergency wideband input port.

TABLE 8
Typical measured diplexer performance.
(Three Channels Into One Antenna)

Channel frequency	93.5 MHz	99.3 MHz	100.5 MHz
Input VSWR, plus minus 250 kHz	1.1:1	1.04:1	1.04:1
Isolation: From 93.5	—	− 80 dB	− 80 dB
From 99.3	− 59 dB	—	− 50 dB
From 100.5	− 80 dB	− 50 dB	—
From broad band port	− 56 dB	− 44 dB	− 39 dB
Amplitude Response:			
Carrier plus minus 250 kHz	0.15 dB	0.25 dB	0.35 dB
Group Delay, in nanoseconds:			
Carrier plus minus 250 kHz	25	27	38
Insertion Loss, carrier	0.23 dB	0.24 dB	0.23 dB

Each diplexer contains a pass through broadband input port and an injected frequency port. Nearly all uncombined diplexer power and other undesired products are absorbed by a dummy load connected to the fourth port. The third port goes to the antenna or to the next diplexer's broadband input port.

The combining of two different stations must be done without degradation to station. The important factors are amplitude, group delay, VSWR bandwidth, isolation, and insertion loss. Insertion loss is a continuing cost item as it consumes RF power that is generated by the transmitter at considerable expense. Table 8 shows the measured performance of a typical three-diplexer installation.

Diplexer specifications can be tailored to the specific requirements of the transmitters being used. Table 8 shows the measured performance of a typical three-diplexer installation.

Constant Impedance-Type Diplexers

The constant impedance-type takes its name from an operating characteristic of one of its components, the constant impedance of the broadband 3 dB hybrid. All of this type use two hybrids and cavities. Because of the use of hybrids, the constant impedance hybrid diplexer is sometimes called a *hybrid combiner*.

An efficient diplexer is not a complicated device, as it consists of two basic components, the two hybrids and two cavities. Coax is used to interconnect the components. Hybrids and cavities either have coaxial flange connections or are directly coupled. Some units are pressurized to keep moisture from tarnishing the cavities which may be silver plated or polished copper, in order to attain high Q.

A terminating load, rated at one-half the highest transmitter power being diplexed in that unit is connected to the reject port. In the event of failure of one or both cavities, half of the power will appear in this load before thermal and other sensors normally turn the affected transmitter off.

The hybrid, sometimes called a 3 dB coupler, is a four-port device. When power is fed into one port, it appears split 50% between two other ports, thus the name 3 dB coupler. Another very useful feature is that the two split powers are 90° apart in their phase relationship. The fourth port is isolated from the input port typically by 26 dB to 40 dB. If two of these hybrids are connected back to back, nearly all the power entering the first input port will appear in one of the output ports, and a second port will be isolated.

In TV use, a single hybrid is quite useful in feeding a turnstile batwing antenna, which requires split input power with a phase quadrature displacement. The old square quarter wave TV hybrid diplexers are similar in operation, but they are limited bandwidth devices and not suitable for wide-band use in FM community antenna diplexing.

FM diplexer hybrids are capable of extremely high power. They are usually made of large coaxial components with quarter wavelength coupling bars. Their large physical size and low Q greatly reduces power loss which rarely exceeds 0.05 dB (1.1%).

Many constant impedance diplexers use only reject cavities. Each diplexer contains a broadband input port and an injected frequency port. Nearly all uncombined

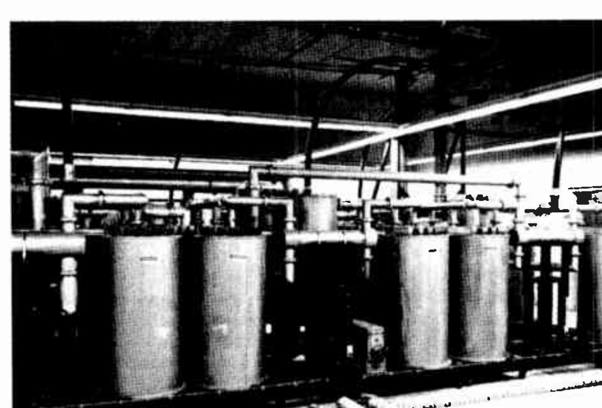


Figure 15. This Senior Road, Houston, Texas ten-port diplexing system, combines nine FM stations into one community antenna. It is capable of 350 kW total input power. (Courtesy Harris Corp.)

diplexer power and other undesired products are absorbed by a dummy load connected to the fourth port. The third port goes to the antenna or next diplexer's broadband input port.

A band reject diplexer is shown in Fig. 16. The signal from TX-1 splits equally in hybrid H-1. It passes by the cavities (C-1 and C-2) which are tuned to TX-2 frequency. The two signals of TX-1 combine in the antenna output port of hybrid H-2 since the two inputs to H-2 are equal in amplitude, but 90° out-of-phase, due to the action of hybrid H-1.

The signal from TX-2 splits equally in hybrid H-2 and are 90° out-of-phase with each other. Cavities C-1 and C-2 are tuned to the TX-2 frequency and present a short circuit to it. So the signal is nearly 100% reflected back to the inputs to H-2. These signals are 90° out-of-phase so they combine only in the antenna port of H-2 and go out to the antenna.

TX-1 frequency is not critical and can be any frequency within the FM band, as long as it is removed from the cavity frequencies by at least 1.0 MHz. However, it will work with some makes of reject diplexers when the separation is a minimum of 800 kHz.

The cavities tuned to TX-2 are not critical in their spacing from hybrid H-2. Because of the nature of the two hybrids, the VSWR remains low and is not affected by temperature variations. The isolation also remains high under these conditions.

Another advantage of the constant impedance notch diplexer is that any beat frequency or intermod product generated in the transmitter TX-2 will be absorbed in the dummy load since only TX-2 will be reflected. This is due to the frequency selectivity of C-1 and C-2.

When more than two transmitters are to be combined, the additional diplexers are merely added in series as shown in Fig. 14. Additional diplexers do not affect the performance of the rest of the system, and can be added at a later time if the need arises. Because of their simplicity, constant impedance diplexers are virtually maintenance free.

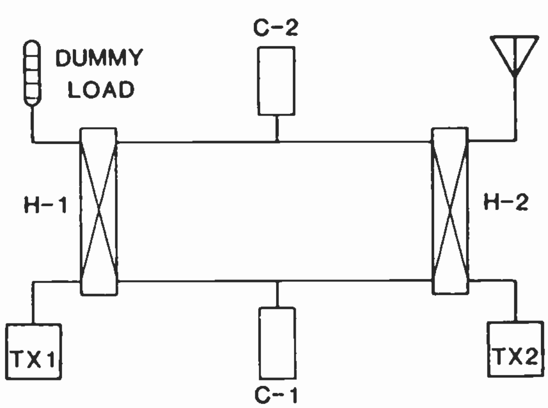


Figure 16. Basic constant impedance diplexer. See text for explanation of component functions.

Branched Starpoint Diplexers

This type does not use hybrids. Each transmitter feeds into a set of branching filters. In each set there is a bandpass cavity tuned to pass the operating frequency and at least one reject cavity tuned to the frequency of the adjacent channel to sharpen the skirt selectivity. If more than two frequencies are being combined, the middle frequencies have reject filters for both their adjacent frequencies.

All the filter sections connect to one common junction, or starpoint, which is the combiner's output. Coaxial cable lengths between each diplexer and the starpoint must be a defined electrical length for proper filtering action and low VSWR. This type of diplexer does not use an absorption load, as the small amount of rejected power is consumed by the rejected frequency transmitter and the antenna.

There is no broadband input, so adding channels requires that balanced modules be added on the output. If there is a failure of one branch, it will usually not take the others out of operation. Since there is no broadband port, the affected frequency cannot be switched as would be done with a hybrid diplexer system. However, the transmitter from the failed branch may be operated into one half of the antenna, while the output of the combiner, containing the unaffected channels feeds the other half, for an ERP reduction of 3 dB.

The principal advantages of a branched system are in space and price. There are fewer components, i.e., no dummy loads, no hybrids, and fewer cavities. They are frequently fan cooled, thus permitting the use of smaller cavities. This is all at the price of flexibility and, in some cases, performance.

Cavity Construction

In order for diplexers to work, a frequency selective electrical short circuit is required to be placed between the two hybrids. This is provided by high Q cavities.

During World War II, low-pass and high-pass filters were developed for VHF communications and radar. The need for microwave receivers with greater selectivity led to the development of magnetically coupled, quarter-wavelength long cavities. Improvements since the 1940s have made cavities with excellent loaded Qs and extremely low insertion loss.

Temperature stability improvement is the result of using invar steel, and the high Q results from silver plating and polishing the inside of the cavities. Cavity size for FM use varies due to the resonance mode selected by the design. Round cavities as small as 20 inches (508 mm) in diameter and 30 inches (760 mm) long have been used. Square types as long as 60 inches (1524 mm) using wave-guide modes have been used. Generally speaking, the larger the cavity, the smaller the RF loss through the cavity.

A very high loaded Q is necessary for a frequency spacing of 800 kHz, which is the closest frequency assignment in any specific community. Practical Q values vary from 1,000 to 12,000.

The power dissipated in heat will expand critical parts of the cavity, detuning it if the heat is not efficiently removed. Air blowers, cooling fins, or simple black paint can be used, depending on the amount of heat to be removed.

The cavities contribute most of the loss found in a diplexer. A diplexer with its two cavities tuned to the first possible channel 800 kHz away could have an efficiency of about 95% (0.25 dB loss). Efficiency goes up as the spacing between the passband and the reject frequency increases so that at 1.6 MHz it could be 96% and at 2.3 MHz, about 97%.

An extremely high Q would be excellent on the operating frequency but would not be useful on the FM sidebands. Using more cavities and stagger tuning them increases costs but does not completely solve the efficiency problem.

FM ANTENNA INSTALLATION ON AM TOWERS

The current trend is to locate FM transmitters in places where the best service may be rendered to the most listeners. This usually permits the maximum possible height to be used. Sometimes however, it may be economical and convenient to install the FM antenna on a tower used for AM broadcasting. If the steel AM tower is not base insulated but is grounded and shunt fed, the FM coaxial line may be connected to the tower, without any further problems.

Transformer Isolation

If the AM tower is insulated at the base, an isolation transformer may be used to couple the FM power across the base insulator without introducing objectionable mismatch into the FM antenna feed line. An isolation transformer is especially desirable for feeding high impedance AM radiators or AM radiators which are part of an AM directional antenna system which might be adversely affected by a quarter-wave isolation system.

These transformers have two tightly coupled RF coils which are resonant at the FM operating frequency. An air gap is provided for the AM power to pass through the two resonant loops. The capacity is quite low, resulting in a very high capacitive reactance placed across the tower base insulator.

Fig. 17 shows the internal basic construction of a typical isolation transformer. The insulation for AM under the top of the box may be high density polyethylene, teflon, or fiberglass. The metal top provides a rain shield as well as protection from dust, mud, or snow.

The use of these isolation transformers permits the AM tower to operate undisturbed by the presence of the FM antenna. It also allows the FM coaxial line to be connected in the usual manner, except for the placement of the isolation transformer. These have internal gas blocks and permit the passage of dry air pressure through the transformer via a plastic tube.

In addition to lower cost, the isolation transformer method has another advantage in directional AM tower use. It does not distort the AM radiator current distribution which may adversely affect the AM radiation pattern.

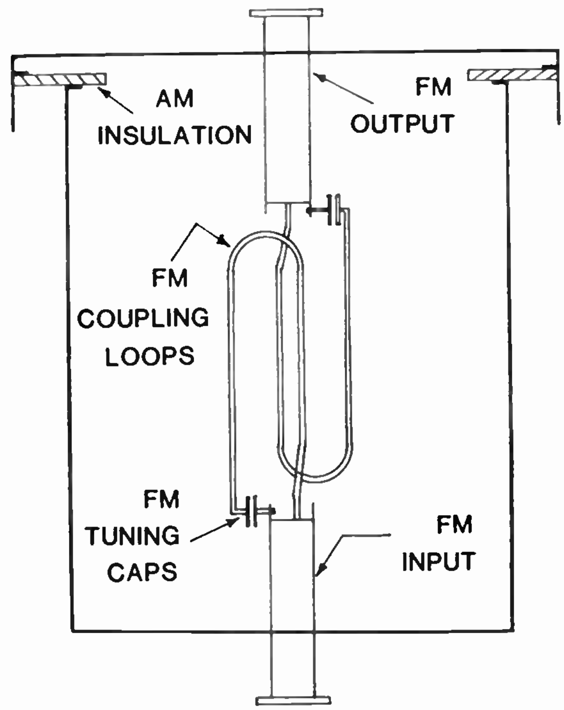


Figure 17. Typical AM-FM tower isolation transformer, used to decouple FM transmission line on a series fed AM tower.

Quarter-Wavelength Isolation

A less popular and older method is to use the technique of quarter-wavelength transmission lines. Simply stated, the opposite end of a shorted quarter-wavelength line has high impedance. This high impedance is placed across the AM tower base and may be successfully used to provide the necessary isolation. It is more difficult to physically accomplish this as the tower should be at least one quarter-wavelength high, and the FM antenna coax line must be insulated all the way down the tower. In practice the insulated part may be as short as 75° of line, as the line hangers and distributed capacity of the line tend to electrically increase the physically shorter line. For best results, the FM line should be placed within the tower body.

Guy Cable Considerations

The presence of continuous steel guy cables going through the FM antenna level on a steel supporting tower was studied by Jampro Antenna Company in 1968. It was found that guy cables had an effect of less than 0.6 dB in the azimuth pattern of a H-pol antenna. On V-pol the maximum variation was 1.8 dB on the azimuth pattern. The strongest effect is on CP where the azimuth pattern change was as great as 3.4 dB. The elevation pattern was also affected since the first and second nulls were filled as much as 4 dB.

In addition to pattern anomalies, the steel cables reradiate near the ground. This may cause RF feedback problems in some high power installations with low level audio equipment located in a building near the base of the tower.

Strong currents may be induced when the steel guys are in the immediate vicinity of the radiators. If that guy passes close to the side of the radiator, the field on that particular side of the antenna element will induce currents in the guy wire. The radiator currents can become unbalanced and the impedance of the element is disturbed, changing its VSWR and radiation pattern.

Because of these effects, it is common practice to break-up the guys using insulators, fiberglass rods, or plastic guy cable, within 10 ft (3 m) of the antenna radiators.

When a CP antenna is side-mounted on a guyed tower, the vertically polarized field will have an appreciable component parallel to the guy wire in its aperture and will induce currents in the wire. If the guys are continuous, a progressive wave traveling toward the ground will result, and will radiate most of its energy before reaching the ground. The energy will be radiated in cones concentric with the wire. A small amount of the V-pol power will thus be bled off.

Porcelain Insulators

If the FM antenna is side-mounted on an AM tower, which usually will have metal guys and porcelain break-up insulators, those insulators will probably be spaced several FM wavelengths apart. The induced currents will form standing waves on the sections between insulators and radiate multilobed patterns into space at many angles from the wire axis. If the sections between insulators happen to be of a resonant FM wavelength however, currents in the guy wires and their radiated fields will be considerable.[22] A single isolated piece of guy wire with its ends insulated can only resonate in multiples of one half wavelength, so this spacing of insulators must be avoided. In fact, with the capacitive end loading of the insulators, the resonant length of wire will be somewhat less than one half-wave, so ⅜ wavelengths should also be avoided. A quarter-wavelength is much better, but this would be quite expensive as it requires insulators every 30 inches (762 mm).

Alternatively, the guys through the FM antenna aperture on the tower could be replaced with plastic cable, which is transparent to RF energy.[25]

Nonconductive Guys

In order to eliminate FM pattern distortion, any guy cable going through the antenna level should be of nonconducting material. Plastic fiberglass (GRP) insulating rods as well as flexible plastic rope covered with a PVC plastic jacket can be used. The black jacket prevents deterioration due to ultra violet sunlight radiation, which may be injurious to the plastic strands of the rope. Plastic rope has been successfully used for more than 25 years.

Figure 18. A roll of plastic guy cable with eye and jaw end connectors in place. Diameter strength is equal to EHS steel stranded guying cable. (Courtesy Phillystran, Philadelphia Resins Corp.)

The idea is to remove metallic RF conducting steel guy cables from within the antenna aperture. The rest of the guy may be of steel construction. The length of the steel guy from its attachment point near the antenna to a point well below the antenna is simply replaced with an equal length of fiberglass rods, or plastic rope.

Plastic rope is available in continuous lengths of up to 1,000 ft (304 m) and kits are available for installing the end fittings. The cable is quite flexible as the Fig. 18 shows a 225 ft (69 m) length coiled up with its end fittings installed at the factory.

The cable may be purchased in strengths exceeding similar diameter steel guy cable. These strengths are shown in Table 9 for corresponding size of commonly used EHS (extra high strength) steel guy wire. Sizes smaller and larger than shown in the table are available.

INSTALLATION PROCEDURES

If the installation is not properly planned and carried out, there may be unwarranted delay and cost associated with putting the FM antenna on its support tower. The following suggestions are offered to avoid unnecessary delays and expenditures.

Planning The Installation

Because of the extremely high cost of rigging services, it is essential to carefully plan the installation. Make sure that all parts are on hand.

The installation of the antenna should be planned by a technically qualified person who must supply accurate tower construction information to the antenna manufacturer. If this information contains errors, these

TABLE 9
Phillystran type HPTG plastic guys.

Outside Diameter		Break Strength		Jacketed Weight		EHS
Inches	mm	Pounds	Kgs	1,000 Ft	300 m	Equivalent
0.20	5.1	4,000	1,815	18	8.2	3/16
0.29	7.4	6,700	3,039	31	14.1	1/4
0.42	10.7	11,200	5,080	55	24.9	5/16
0.46	11.7	15,400	6,975	69	31.3	3/8
0.53	13.5	20,800	9,435	93	42.2	7/16
0.58	14.7	27,000	12,247	115	52.2	1/2
0.63	16.0	35,000	15,876	142	64.4	9/16
0.68	17.3	42,400	19,235	167	75.8	5/8
0.73	18.5	58,300	26,445	195	88.5	3/4

will be carried through the design and fabrication of mounting hardware, and finally show up in the field to plague the installing crew, wasting time and money at every stage of the process.

The station should consider hiring a tower rigging firm that is financially qualified and mechanically well equipped to do the work. A written contract should exist between the station and rigging firm, with a fixed price. The rigging contractor should be licensed as a contractor in your state, and should post a completion bond. He should also supply an insurance policy holding the station harmless, and making the station and its personnel co-insured. Only in this manner will the broadcaster be protected in the event of injury or property damage.

The riggers should be knowledgeable about antennas and coax line, and should inspect the tower and check out the mounting design of the brackets before the full rigging crew arrives.

If any factors are discovered which appear to negate the installation design, contact the factory immediately. Particular attention should be paid to the following:

1. Fit of mounting brackets to tower member.
2. Freedom from interference of the mounts with gussets, leg flanges, guys, and their attachment points, tower face members, and obstruction lights.
3. Compatibility of transmission line and antenna input coax terminals.
4. Location of transmission line run relative to antenna input terminals.
5. Use of fiberglass guys on the tower in the region occupied by the FM antenna; refer to the AM Guying section in this chapter.
6. Availability of proper voltage, current, and cable size for deicers if required.
7. Adequacy of tower to carry the windload placed upon it by the antenna, particularly where radomes are used. This radome/antenna load should be checked by a competent structural engineer, as all antenna installations should be checked. This is usually required by the company carrying insurance on the tower.

Receiving And Unpacking

The boxes are usually numbered and the total number is indicated on each box; contact the shipper if not all boxes are delivered, or if equipment is received damaged. Do not store the material outdoors, boxed or otherwise.

As soon as the antenna is received, open and examine for shipping damages so that any necessary claims may be filed with the shipping company immediately. Check the material against the parts list and installation drawing.

The box with the installation drawing and instructions is usually marked. Open it first, so that the balance of the items may be easily identified and counted. Contact the factory immediately if any material appears to be missing or is damaged during transportation.

Do not call the riggers until all antenna and coaxial line at the site. Otherwise, unnecessary delays and costs may result.

Installation Tips

Broadcast antenna manufacturers furnish detailed installation procedures with their products. These instructions should be closely followed by the rigger. Together they will ensure a perfect installation saving time and money.

The following items are specifically called to the attention of the broadcast engineer (in addition to all those stated before) to permit proper installation and good performance for many years.

1. Follow manufacturer's instructions. See that the riggers also read these instructions.
2. Do not leave antenna parts where rain or moisture can enter. Store indoors and keep units capped as received.
3. Do not allow dirt or other foreign matter to enter any coaxial part.
4. Protect all antenna parts from physical damage and abuse.
5. Hoist antenna members carefully, with a tag line to prevent damage by striking against the tower.
6. Install on the tower as indicated by the manufac-

turer's instructions, remembering that bay number 1 is the uppermost top unit.

7. Riggers should lubricate "O" rings with a small amount of silicone grease before mating flanges.
8. The full complement of flange bolts must be used and they should be as tight as possible.
9. Tuners or individual element devices, if used, should be adjusted only after the entire antenna and tower installation has been completed.
10. Rigid transmission lines should be properly installed with two hangers per 20 ft (6 m) length, and with the inner conductor retaining pin on the top of each section.
11. If semi-flexible cable such as Heliax or Wellflex is used, it should be firmly tied down at least every 5 ft (1.5 m) for 3 inch (76 mm) line, and every 3 ft (1 m) for 1⅝-inch (43 mm) coax line. The line manufacturer's hangers should be used. The line should not be attached to the tower using wrap-lock straps.
12. After physical installation has been completed in accordance with the manufacturer's recommendations, the main transmission line should be pressurized with dry air through a dehydrator, air pump, or by using dry nitrogen gas. See the "Air Pressurization" section in this chapter for more information.
13. Dry air or gas pressure should be maintained at all times. Most antenna warranties are not valid unless this is done. It is the riggers responsibility to make certain that the entire coax and antenna holds air pressure.
14. The antenna system should be checked by a qualified rigger every time the obstruction lights are replaced, or if lights are not used, at least once a year. The rigger should look for vibration and storm damage, loose or broken coax hangers, and signs of arcing across exposed insulators. A dry rag soaked in 91% isopropyl or other solvent alcohol or equal should be used to wipe clean all exposed insulators in each antenna element. (DO NOT USE CARBON TETRACHLORIDE!)

STRUCTURAL CONSIDERATIONS

Most FM antennas in the western hemisphere are installed on the sides of a steel tower, between 18 and 60 inches (45 cm to 152 cm) wide. The antenna and its transmission line together with all mounting brackets introduce wind loading, in addition to their dead weight. The live wind loading is a result of the amount of physical surface presented to the wind. It is sometimes called the *wind catch area*. This consists of either flat or round antenna members, coaxial lines, mounting brackets, and hardware, all represented as surfaces which are exposed to the wind.

The dead weight of the antenna system is fixed and is always present on the tower. The live load is a variable, depending on the wind velocity, and is added to the dead load for the total amount present.

The standard wind load starts with an assumed wind velocity of 87 mph (139 m), which will produce a push of 35 pounds per square foot. (170.8 kgs/sq.m) With lesser wind speeds the wind push is less, and more with higher velocities. Various building codes determine the rated winds to be considered in the design of the tower system. While most of the United States has a 35 pound per square foot minimum rating, some parts of the country have higher requirements due to higher wind velocities. Some insurance companies may require even higher safe wind ratings.

ANTENNA POLE MOUNTING

Nonsymmetrical antennas may also be installed on a round pole, made of various diameters of steel pipe. Several antenna manufacturers supply these as a complete system and will optimize the horizontal plane pattern. The advantage of pole mounting on top of a tower or building is that the pattern may be more easily contoured. This provides more signal in the service area since the antenna orientation is not limited by a fixed triangle formed by a guyed tower.

TRANSMISSION LINE SYSTEMS

Two types of coaxial transmission lines may be used to feed FM antennas. One uses rigid coaxial line sections, each 20 ft (6.09 m) long and requires elbows, flanges, spring hangers and other devices.

The other has a semi-flexible coaxial line which is available in either air or foam dielectric, with fixed hangers. Semi-flexible cable is available on a spool whose diameter depends on the line size. EIA end flanges mate to the antenna flanges as well as other RF equipment. A typical semi-flexible coaxial transmission line layout is shown in Fig. 19.

A supply of dry air should be used with all air dielectric transmission line. Its purpose is to keep out moisture, which may find its way into the line. Dry air not only keeps the moisture from covering the internal plastic insulation, and thus arcing is greatly reduced. It also keeps the internal copper from corroding. Copper which has been exposed to moisture will turn dark brown and oxidize after time. Rigid as well as semi-flexible line in the 3 and 3⅛ EIA size which had not been pressurized have been measured with increased attenuation of nearly 4 dB per 100 ft (30.4 m) at 100 MHz. This is a power loss of 60%. See the section on "Air Pressurization" on how this may be accomplished.

Popular types and sizes of coaxial cable transmission line are shown in Table 10. The most popular are the semi-flexible air cables.

The safe power ratings shown in Table 10 are for a perfect VSWR of 1.0:1. It is considered good engineering practice to derate this by dividing the power in kilowatts by the expected VSWR. For example three inch Heliax rated at 36.8 kW at 100 MHz and divided by 1.1 derates it to 33.45 kW. Other increases

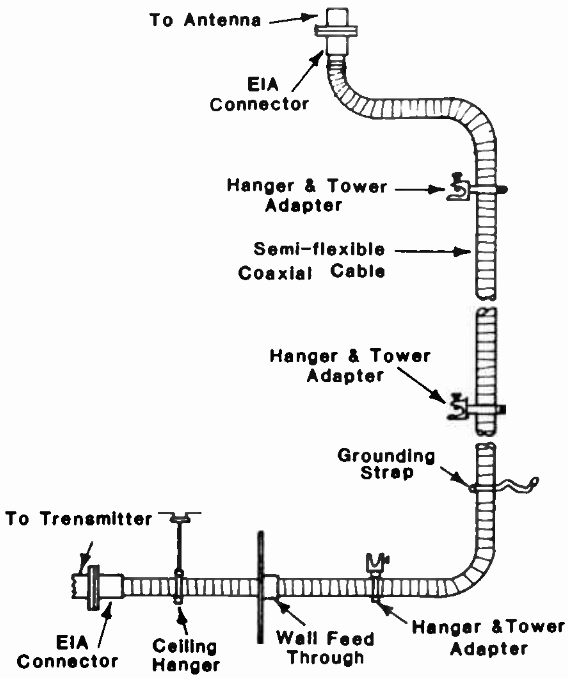


Figure 19. Typical semi-flexible coaxial cable installation on a tower. Number of various pieces of hardware depend on length and the diameter of the cable.

and decreases of coaxial cable ratings may be found in the catalogs of most manufacturers of transmission lines.

Ratings are for a VSWR of 1.0:1; ambient temperature of 75°F (24°C) with one atmosphere of dry air pressure. Rigid line is rated at 122°F (50°C). The LDF and HJ series are manufactured by Andrew Corp. of Orland Park, Illinois while the rigid line is by Shively Labs., of Howell Labs. Inc. Bridgton, Maine.

AIR PRESSURIZATION

If the antenna is operated without positive pressure of dry air or nitrogen, the manufacturer will not assume responsibility for failure under power. Moisture or even the accumulation of water within the coaxial transmission line is a very serious matter. Its presence causes the VSWR to rise. When a sufficient amount of moisture is present, arcing will take place burning the line or antenna radiating elements. High humidity or moisture will cause the inside of the coaxial transmission line to corrode over time, thereby increasing the line loss. For this and other reasons, the entire antenna system must be dry air pressurized.

After the antenna is installed and the transmission line connected, the system is purged with dry gas or dry air to remove trapped moisture before RF power is applied. A manually opened or pressure actuated purge valve is installed in nearly all FM antennas made by American firms. When the gas pressure is raised to 10 psig (0.68 atmospheres), the automatic pressure relief valve (if the antenna has one) will open up letting moist air out. The complete system purge requires a considerable volume of dry gas.

Before expending this amount, it is good practice to perform a quick check for major leaks. The system pressure is raised to a point below the relief valve setting, such as 8 psig (0.48 atmospheres) the source of supply shut off. A pressure gauge should be installed on the antenna side of the shut-off valve. The pressure, when corrected for temperature, should not fall to less than half its initial value in a 24-hour period.

If the pressure loss is more than this, the system should be checked with a leak detector, or soap suds,

TABLE 10
Coaxial transmission line.
(Characteristics At 100 MHz)

Line Type & Number	Nominal EIA Size	Attenuation In dB	100 Ft Eff. %	Safe FM Power-kW
-Foam-				
LDF5-50	7/8	.385	91.5	5.2
LDF7-50	1-5/8	.231	94.8	13.1
-Air-				
HJ5-50	7/8	.373	91.7	6.4
HJ7-50A	1-5/8	.205	95.4	14.2
HJ8-50B	3	.142	96.7	36.8
HJ11-50	4	.115	97.4	56.1
HJ9-50	5	.078	98.2	73.9
-Rigid-				
1213-1	1-5/8	.191	95.7	10.0
1313-1	3-1/8	.096	97.8	40.0
1413-1	4-3/8	.070	98.4	81.0
1613-1	6-1/8	.050	98.8	161.0

to locate the leak. A pinched or missing O-ring is the most common cause for large leaks.

Once the system is known to hold pressure, it should be purged with dry air or gas. Either must be dry enough to have a dew-point well below the coldest temperature expected to be encountered. When using nitrogen, it should be of the "oil-dried" type, to remove nearly all the moisture from the gas.

Five to eight psig (0.34 to 0.48 atmospheres) should be maintained in the system at all times to ensure that no moisture will be able to enter. Very small leaks will pull in moisture, if the pressure is lower than suggested above, when the transmitter is turned off nightly. This is due to the pumping action due to expanded dry air/gas pressure cooling down, and contracting below the outside air pressure, during cold ambient temperatures.

PROTECTION FROM ICING

High Q antennas are subject to increased VSWR ratios as well as pattern distortion, with light to moderate coatings of ice. Low Q antennas such as the panel type are usually not affected in this manner. Where climatic conditions cause sufficient ice or in some cases snow to affect the antenna's performance, there are two remedies. The radiating element may be covered with a plastic cover, or, it may be electrically heated to melt or prevent the formation of ice on its sensitive surfaces.

Electrical Heaters

By far the more popular method of deicing is to order electrical heaters at the time the antenna is ordered. Electrical deicing equipment is supplied as an option and is factory installed. Kits are furnished for interbay connections, but the broadcaster must supply power from the building to the center of large arrays, or the bottom element on smaller antennas. Local electrical codes of course must be followed.

While a thermostat may be used with small total deicer wattages, a power relay operated by such a device is required and furnished with most electrical deicer kits by the antenna supplier. Due to high power costs, a sophisticated deicer control, which operates when both temperature and humidity conditions produce sleet or icing, is often required.

Most deicers use a resistance heating element which is inserted inside the antenna radiator arms. One manufacturer, however, uses a different method, dropping the 230/240 volts to a few volts with a transformer located at each bay level. The low voltage is passed through the ice sensitive arms of the radiator and connected to the far ends by a heavy teflon coated wire. The current return is by the stainless steel antenna element, whose ohmic resistance is sufficient to produce enough power heat loss to melt or keep the ice off. This method is becoming obsolete as the transformers are expensive and heating costs are rising, as hourly electrical rates go up. The voltage dropping transformers are not as efficient as direct heaters.

A word of caution when selecting a FM antenna with electrical heaters. Some deicers use 1 kW of power for each bay as described above and increase the wind loading by 225%. Others have a switchable power option feature using 125/500 watts per element, with only a 15% increase in windloading, when compared to an antenna without electrical deicers. The continuing cost of electrical deicers is a consideration of the operational cost of the station and should not be overlooked.

Automatic deicers are those with a thermostat for mounting near the antenna for accurate temperature sensing of the actual ambient temperature. The temperature zone of +20°F to +35° (−7°C to +2°C) is generally the most likely icing range, depending on humidity conditions. Deicers should be turned on at +35°F, prior to ice formation, because it is better to prevent icing than to remove it once it is formed. Power should be turned off when the temperature goes below +20°F since ice does not usually form below this temperature. Fig. 20 shows electrical deicers being wired.

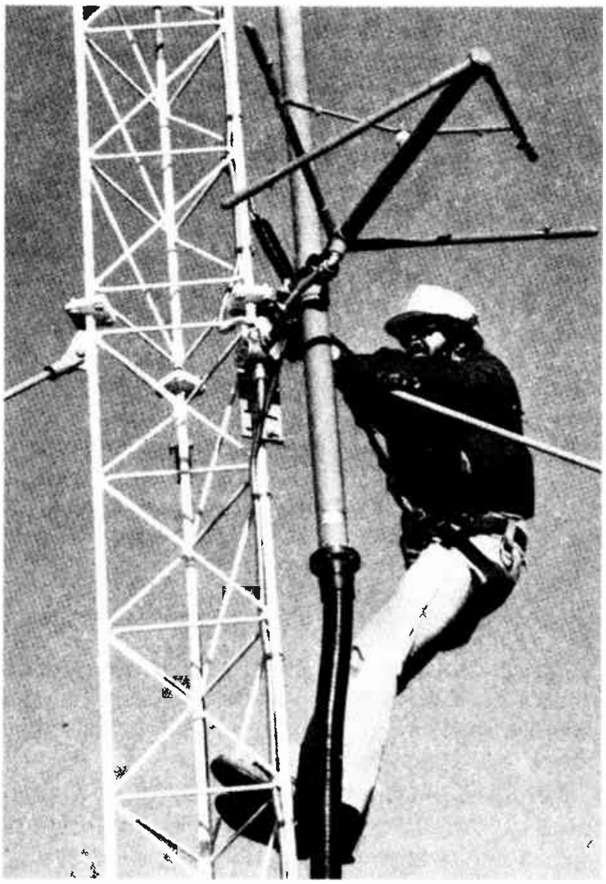


Figure 20. Rigger finishing electrical tie in box for deicers. Looped stainless steel heaters were factory installed in the four arms, plus one in the feed support boom. (Courtesy of Cetec Antennas)

Radomes

A radome is a protective dielectric housing for an antenna radiating element. Its function is to protect the antenna not only from ice, but snow and physical damage due to ice dropping from above. Radomes also protect the radiating element from environmental corrosive atmospheres.

The primary purpose of using radomes on FM antennas is to prevent the VSWR from rising with the formation of ice, if the site and height causes icing to occur during the winter months. Ice formation detunes high Q radiators, increases the VSWR, and causes vertical plane pattern changes. Fig. 21 shows a typical radome enclosing a radiator.

Ice may form on the radome but does not particularly affect the operation of the radiator if that ice is kept at least 0.05 wavelengths from the sensitive portions of the antenna element.

Radomes are particularly desirable in heavy icing environments where deicers are not adequate even with very high heat density. They are also useful in protecting antenna elements from falling ice when they are so exposed.

Radomes are cost effective with single channel high Q antennas where electrical deicer heating power costs are expensive. The deicer power cost is a continuing one, while radomes are a one time capital investment, which may be depreciated over time.

Radomes are generally composed of low-loss dielectrics with low values of dielectric constants and loss tangents. Laminated fiberglass, using glass cloth reinforcements, has a constant of about 4.1 and a loss tangent of about 0.15. Water absorption by the radome increases its dielectric constant and loss tangent. Materials which do not easily absorb water or those treated with a protective gel coat are often used to shed water and prevent the adhesion of ice.

Good radome designs take into consideration operating temperature, a relative humidity of 100%, safe wind pressures, ice, hail, snow loads, rain adhesion, wind and supporting tower vibration, fire retardant plastic, and the ability to safely withstand air contaminants over the useful life of the antenna. All these factors increase the cost, but are necessary for long, useful life. Radome shapes are dictated by the form of the radiating element in most instances.

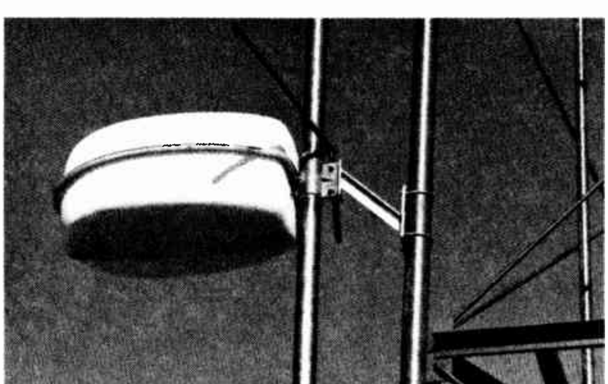


Figure 21. Radome over radiating element. Transparent to RF, it keeps ice from detuning high Q radiator. (Courtesy of Cetec Antennas)

In all cases radomes are supplied by the antenna manufacturer, and usually supplied in two pieces which are bolted together with stainless steel fasteners.

LIGHTNING

Because FM towers are usually located on high ground, hilltops, or high buildings, they require lightning protection since they are likely recipients of lightning strikes. The type of damage that can be caused by lightning to a FM tower varies. Smaller coaxial lines will usually melt; larger coax (1⅝, 3⅛) will also melt in some cases, and others will conduct the heavy current into the transmitter building to do damage there.

The FM antenna itself may heat, arc, melt, and otherwise be damaged. Holes in the outer conductor, burns, and melting at flanges are common. Teflon or polyethylene insulation will burn, depositing a film of carbon, causing further damage if RF from the transmitter continues after the strike.

Protection of the FM antenna system may be provided to some degree, by taking several precautions. The top of the tower should have a lightning rod, about one foot (0.3 m) higher than the uppermost obstruction light part. The FM antenna itself should be firmly grounded to the tower. If the coaxial cable is buried between the tower and the transmitter building, it must be at least six ft away from any tower base grounding system copper wire or strip.[26]

A ground system should be located immediately around the base of the tower. This should have a direct current loss of less than 10 ohms to earth ground. This low resistance may be obtained by using ground wires buried in the soil. Six radials spaced 60° if possible, buried as deep in the soil as possible and running out up to 150 ft (46 m) each, should provide a suitable ground of less than 10 ohms, even if the soil is shallow or rocky.[26]

Guyed tower anchors should also be grounded. This is covered in the chapter on "Design, Erection, and Maintenance of Antenna Structures." It is important to install the proper number of ground rods and/or copper wire radials in order to obtain a connection to earth ground of less than 10 ohms. In any event these ground rods or radial wires must be tied together with number AWG 4 or larger copper wire, or copper strap two inches (5 cm) minimum width. This is to provide for thousands of amperes of current flow for less than one second in the event of a direct lightning hit.

If the FM antenna is located on an AM insulated base tower, then the spark gap should be set at the lowest point providing protection for the highest AM modulation peak voltage.

Another way to protect the FM transmission line isolator (if one is used), as well as the tower and FM antenna, is to use a RF choke across the insulated tower base. This tends to reduce the static build-up voltages due to passing thunderstorm clouds, snow, hail, or dust storms. Arcovers due to these sources usually do not cause damage, but may trip the FM transmitter reflectometer since they will create a current flow through the reflectometer circuitry.

If the base insulated tower supporting the FM antenna is located in an area of regular thunderstorms, another way to protect both antennas from lightning is to ground the AM tower at the base, and shunt feed it. Several excellent methods exist. The folded unipole method not only grounds the tower for lightning purposes, but improves the VSWR bandwidth, so necessary for AM stereo. See the chapter on "AM Antennas" for several suggested methods.

FM SCA MULTIPLEXING

With a 92 kHz subcarrier and 110% modulation intermodulation products may be created due to mixing of the various subcarriers with their own harmonics within nonlinear devices. One such nonlinear device can be the antenna system.

Antenna linearity is determined by its VSWR response curve versus frequency. Phase delay in the antenna system is also important. In the past, with 67 kHz being the highest SCA frequency, the ±100 kHz bandwidth was considered sufficient. Now with a 92 kHz subcarrier and 110% deviation (82.5 kHz), the minimum bandwidth is ±130 kHz under 1.1:1 VSWR.[27] See "VSWR Bandwidth" in this chapter for more specific requirements and recommendations.

Tests have shown that 92 kHz is the frequency of choice for a new aural SCA service after 67 kHz.[28] The 92 kHz subcarrier produces lower intermodulation product levels and less interference to the main channel stereo service than 67 kHz.[29] It may be successfully operated in addition to stereo and existing SCA services.

Other nonlinearities in the exciter, and transmitter, plus multipath reception, receiver misalignment, and user mistuning are contributions to the received intermodulation distortion of the baseband signals. In addition these products can cause small levels of audible swishing beat notes in some types of FM receivers.

SPURIOUS FREQUENCIES

Interference to other stations within the FM broadcast band as well as to other services outside the band can be caused by RF intermodulation product energy developed between two or more FM broadcast transmitters. It may be due to coupling through a diplexer or coupling between two antennas. This phenomenon has been well documented.[12]

Detailed information on the susceptibility of various types of transmitters to interference from other collocated transmitters has not been thoroughly investigated. A method has been devised by which the mixing loss between two transmitters can be accurately characterized.

When RF energy from two or more transmitters is combined, new spectral components are produced by mixing the fundamental and the second harmonic of each of them. The dominant intermod product generated by each transmitter is at twice the transmitter's frequency. For example, 101.1 and 102.7 transmitters would produce two intermod signals appearing on 99.5 MHz and 104.3 MHz.

Second harmonic traps or low-pass harmonic filters in the transmission line of either transmitter prior to the diplexer have little effect on the generation of intermod products. This is because the harmonic content of the interfering signal entering the transmitter output circuit has much less effect on intermod generation than the harmonic content within the nonlinear device itself. The resulting intermod falls within the passband of the low-pass filters and outside the reject band of the second harmonic traps, so these devices offer no attenuation to intermod products.

Even the perfect diplexer by its very nature will reflect some of the undesired energy back to each transmitter, generating intermod products. The key to this problem is to keep that undesired power level as low as possible using proper transmitter output circuitry and tight diplexer specifications. Diplexed transmitter installations should be routinely checked for the presence of excessive intermod products.

HARMONIC FILTERS

FCC Rules Section 73.317(d) calls for the harmonics of FM transmitters to be up to 80 dB or more below the transmitter output. This requirement is usually met by using a low-pass filter which passes the station carrier frequency power but attenuates its harmonics.

The transmitter provides some harmonic attenuation of course, and is usually 25 dB to 38 dB for single ended amplifiers. The worst case harmonic is the third. Harmonic filters by several firms provide a minimum of 50 dB for harmonics from the second through the tenth. Adding the transmitter attenuation to that of the filter normally provides more than the required level.

The high level of rejection is made possible by using high-impedance (inductance) and low-impedance (capacity) coaxial sections for m-derived three to five section filters, with half-pi end sections. Harmonic filters are commonly made in three production schedules. One firm actually adjusts them to the customers operating frequency, so that there are no attenuation gaps in the higher harmonics. They are not tunable outside the factory as the insertion loss, and attenuation along with pass band VSWR are closely related. It requires sophisticated knowledge and equipment to properly adjust.

The insertion loss in the pass band varies from 0.05 dB to 0.08 dB while the rejection from the second through the tenth harmonic is from 50 dB to 60 dB, depending on the number of internal filter midsections. The VSWR in the pass band varies from 1.05 to 1.1:1.

Harmonic rejection is due to the very high VSWR on the harmonic frequencies which may be as high as 15 to 1. This rejected power is passed back to the transmitter amplifier.

Harmonic filters are available in straight rigid coaxial line sections and may sometimes be pressurized. Power capacity varies from 10 kW for 1⅝ EIA line size to 50 kW for the 6⅛.

ACCESSORY ANTENNA SYSTEM EQUIPMENT

Several other devices are associated with the antenna system. The dry air pressurization of coaxial transmission line was discussed under that heading in this chapter.

Reflectometers

The reflectometer is a device for detecting the ratio of power flow from the transmitter to the antenna (forward) and the rejected power back from the antenna (reverse). A short coaxial line section about 12 inches (305 mm) contains diode detectors, coupling loops and terminations to produce dc current, which drives a suitable VSWR meter.

Reflectometers are wide band devices and therefore must be placed AFTER the harmonic filter. Putting them between the transmitter and the filter causes them to read the rejected harmonic power along with the reflected, thus giving an erroneous reading.

Dummy Loads

A very useful test device in an antenna system is the terminating load. At least one is needed when two amplifiers are combined and fed to the antenna, or when a number of diplexers are used in a community antenna arrangement. Dummy loads are available in several power levels up to 50 kW and are cooled by air. Water cooling types are also available.

RF Switches

Often used with a dummy load, coaxial line switches are available to provide electrical or manual switching of transmitter power to diplexers, antennas, standby transmitters, etc. They are not pressurized.

REFERENCES AND ENDNOTES

1. H-pol, V-pol, and E-pol are unofficial abbreviations used for clarity and brevity in this chapter.
2. Broadcasting Yearbook 1991. Broadcasting Publications, Inc. Washington, D.C.
3. Onnigian, Peter: "A Study into the Effects of Vertically Polarized Radiation in FM Broadcasting," Technical monogram. Jampro Antennas, Inc., 1965.
4. Freeman, Roger: "Telecommunications Transmission Handbook," John Wiley & Sons, New York, pp. 180–186, 1975.
5. Head, Howard: "Influence of Trees on TV Field Strengths," Proceedings of the Institute of Radio Engineers, Vol. 48, pp. 1016–1020, June 1960.
6. Armstrong, A.: "Study of Electromagnetic Wave Propagation at 112 MHz." Proceedings of the IREE Australia, pp. 105–110, April 1969.
7. Moeller, Adolph: "Effects of Ground Reflections on Antenna Test Range Measurements." Microwave Journal, pp. 47–54, March, 1966.
8. Reed, Russel: "Ultra High Frequency Propagation." 2nd edition, Boston Technical Publishers, pp. 223–238, 1964.
9. Skolnik, Merril: "Radar Handbook." McGraw Hill Book Company. New York, New York, pp. 39–8 to 39–9, 1970.
10. Saveski, Peter: "Radio Propagation Handbook." TAB Books Inc. Blue Ridge Summit, Pennsylvania. pp. 148–159, 1980.
11. Gray, Lawrence: "Radio Transmitters." McGraw Hill Book Company. New York, New York, pp. 181–186, 1961.
12. Mendenhall, Geof. "Study of RF Intermodulation Between FM Broadcast Transmitters Sharing Diplexed Antenna Systems." Technical Monograph. Broadcast Electronics, Inc. Quincy, Illinois, 1983.
13. Onnigian, Peter: "Stereo Degradation as a Function of Antenna System VSWR" Audio Engineering Society Annual Meeting, Audio Engineering Society, New York, New York, 1976.
14. Mendenhall, Geof: "The Composite Signal - Key to Quality FM Broadcasts." Technical Monograph. Broadcast Electronics, Inc. Quincy, Illinois, 1984.
15. Cohen, Jules: "Proposed Solutions, Channel 6-Educational FM Broadcast Interference Problem." Proceedings of the 38th Annual Broadcast Engineering Conference, NAB, Washington, DC, 1984.
16. Onnigian, Peter: Circularly Polarized Antenna, U.S. Patent 3,541,470.
17. DuHamel, Ray: "TV and FM Transmitting Antennas." Antenna Engineering Handbook. Johnson and Jasik, 2nd edition. McGraw Hill Book Company, New York, New York, Chapter 28, pp. 8–9, 1984.
18. Bock, E.: "Sleeve Antennas." VHF Techniques, Vol. 1, McGraw Hill Book Company, New York, New York, Chapter 5, pp. 119–137, 1947.
19. Knight, Peter: "Re-radiation from Masts at Radio Frequencies." Proceedings of the IEEE, Vol. 114, pp. 30–42. January, 1967.
20. Jampro Antennas, Inc. International Technical Communication. Sacramento, California.
21. FCC Public Notice. "Criteria for Licensing of FM Broadcast Antenna Systems." Notice 84–437 25004, September 14, 1984, Federal Communications Commission, Washington, D.C.
22. Rudge, A.:"Handbook of Antenna Design." Vol. 2, Peter Peregrinus Ltd., London, England, pp. 917–922, 1983.
23. Onnigian, Peter: "Multi-Station FM Antennas." Paper presented at the 23rd Broadcast Symposium,

IEEE Group on Broadcasting, Washington, DC 1973.
24. Fisk, Ronald.:"Design and Application of a Multiplexed Nine-Station FM Antenna, Senior Road Tower Group." Technical Monograph, Harris Corporation, Quincy, Illinois, 1983.
25. Gregorac, L.: "Electrical and Mechanical Analysis of Plastic Guys of Broadcast Towers." Technical Monograph. Radio-Television Ljubljana, Yugoslavia. 12 pp. 1973.
26. Marshall, J.L. "Lighting Protection." Canadian Broadcasting System, John Wiley & Son, New York, New York, pp. 53–54, 1973.
27. Kean, John: "Distortion of FM Signals Caused by Mismatched and Limited Bandwidth Transmitting Antennas." National Public Radio, Washington, D.C. pp. 37–42, 1984.
28. McMartin, Ray B.: "Super Eight" Proceedings of the 38th Annual Broadcast Engineering Conference, NAB, Washington, D.C. pp. 160–166, 1984.
29. Denny, Robert: "Report on SCA Operation." Proceedings of the 37th Annual Broadcast Engineering Conference, NAB, Washington, D.C. pp. 187–196, 1983.

2.7 Antennas for Television Broadcasting

Thomas J. Vaughan
Micro Communications, Inc., Manchester, New Hampshire

INTRODUCTION

This chapter begins with a review of technical terms and FCC considerations. This is followed by a review of antenna design considerations as they relate to these demands, as well as a review of specific applications. The chapter concludes with a review of current broadcast antennas and antenna systems.

TECHNICAL TERMS

Definitions of the following terms, consistent with FCC and industry usage, are important to the discussions in this chapter.

Effective Radiated Power

The product of the antenna input power and the antenna power gain is the effective radiated power (ERP).

Polarization

The polarization (the orientation of the electric field) can be linear, horizontal, or circular. If circular is used, the term ERP applies to the horizontal and vertical components separately.

Azimuth Pattern (Horizontal Plane)

An azimuthal pattern is a plot of the free-space radiated field intensity versus azimuth angle at a specified vertical angle with respect to a horizontal plane (relative to smooth earth) passing through the center of the antenna.

A horizontal pattern is an azimuthal pattern when the specified vertical angle is zero.

For higher gain antennas where beam tilt is employed, the azimuthal pattern at the specified beam tilt is significant. In general it has been customary to determine television broadcast antenna radiation by an azimuthal pattern at the specified beam tilt and a sufficient number of vertical plane patterns taken at various frequencies in the channel.

An omnidirectional antenna is defined as one that is designed to radiate the same signal in all azimuthal directions. Antennas with variations up to ±3 dB have rendered satisfactory service and are considered to be omnidirectional.

A directional antenna is one which is designed to radiate more signal in one azimuthal direction than in other azimuthal directions.

Vertical Pattern (Elevation Plane)

The vertical pattern is a plot of free-space radiated field intensity, measured in the far field versus vertical angle in any specified azimuth plane, which contains the center of the antenna and the center of the earth.

The far field or Fraunhofer region, as usually defined, extends beyond a point where the distance between the transmitting and receiving point is

$$\frac{2a^2}{\lambda}$$

where "a" is the length of the radiating portion of the antenna and λ is the wavelength.

A requirement for the broadcast service is that a free-space radiated field should not be influenced by the proximity of the earth in such a way as to set up a nonuniform field over the antenna aperture. Proper precautions must be taken to accomplish this.

Antenna Gain

Gain is the ratio of the maximum power output at any angle from the subject antenna, to the maximum power from a thin, loss-less, half-wave, horizontally polarized dipole having the same input power.

Antenna gain depends on several factors, including:

- The amount of power concentrated in the maximum direction.
- Losses in the antenna, which include ohmic and other losses such as energy radiated at polarizations other than the desired one.

The amount of power concentrated in the maximum direction can be determined by a comparison with a reference antenna or by integrating the total power flow through a sphere, which is done by taking a sufficient number of vertical and azimuthal patterns.

Both methods are capable of giving accurate results when the proper precautions are taken. Ohmic losses are taken into account in the comparison method or can be calculated when using the power integration method. Cross-polarized radiated energy can be measured. The measurement of gain must be carefully done with a full knowledge of all the problems that are involved. The measurement of gain used in the calculation of ERP for a circular polarized antenna must be made relative to a horizontal dipole. The directivity gain of a ½-wavelength dipole antenna relative to an isotopic antenna is 1.64. The gain of a circular polarized transmitting antenna relative to a "like" circular polarized receiving antenna will be 3 dB higher than that of a horizontally polarized dipole receiving antenna.

Gain requirements for a television broadcast antenna depend on the transmitter power, economics, and field-intensity requirements as determined by the terrain and population distribution.

The maximum ERP (effective radiated power) currently permitted by the FCC is:

Channels 2 to 6	100 kW
Channels 7 to 13	316 kW
Channels 14 to 69	5000 kW

Economics is a factor in antenna choice. As a general rule, for a required ERP the combined costs of transmitters and antennas are less when a higher gain antenna is used.

Impedance

Input impedance is the complex impedance looking into the antenna terminals throughout the television channel.

Most antennas are designed for the same input impedance as the standard transmission line at the antenna terminal. Impedance matching requirements for television antennas are generally more severe than for other types of service because reflected energy, which would occur when the antenna does not terminate the line properly, can cause ghosting and voltage peaks on the transmission line, which could result in damage to the antenna, couplings, and transmission line.

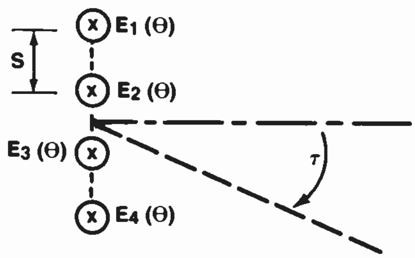


Figure 1. Elevation pattern calculation.

RADIATING CHARACTERISTICS

The television broadcast transmitting antenna should have an omnidirectional pattern in the azimuth plane and a narrow beam in the elevation plane. For an omniantenna the gain is approximately one per wavelength per like polarization. Most broadcast antenna elements are spaced one wavelength apart, therefore, the gain of a linear polarized antenna is approximately equal to the number of elements.

Elevation Pattern

The elevation pattern is the product of the element pattern times the array pattern.

$$F(\theta) = E_n(\tau) \sum_{n=1}^{N} A_n e^{j[(2\pi ns)\sin\tau + \delta_n]}$$

where: $E_n(\tau)$ = element elevation pattern
$\approx \cos^m(\Theta)$
s = element spacing
τ = elevation angle
δ_n = phase of nth element
A_n = amplitude of nth element

The resulant antenna elevation pattern for a four-element array pattern spaced 1.02 wavelength with a $\cos^2(\theta)$ element pattern is shown in Fig. 2

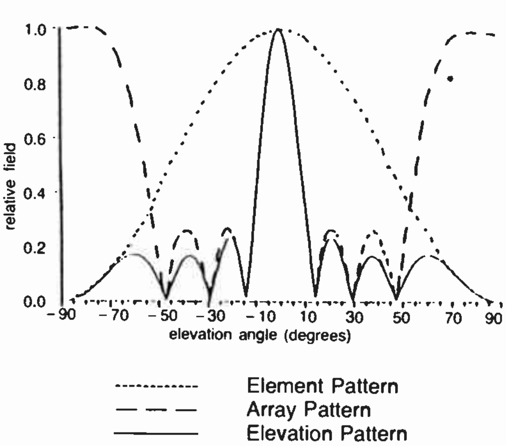


Figure 2. Elevation pattern for a four-element array with $\cos^2(\theta)$ element pattern and 1.02λ element spacing.

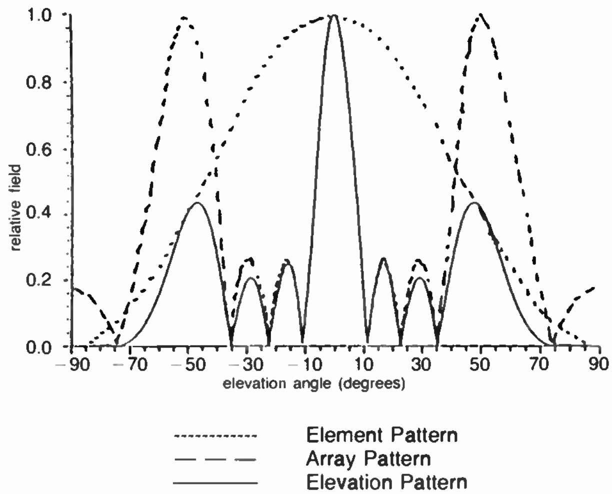


Figure 3. Elevation pattern of four-element array with $\cos^2(\theta)$ element pattern and 1.30λ element spacing.

As the spacing between the elements in the array pattern is increased, "grating" lobes begin to move closer to the main beam. The element pattern will no longer drive the grating lobe to zero at ±90°, resulting in high side lobes.

Fig. 3 shows the patterns for a four-element array with 1.30 λ spacing. Although the main beam will narrow as the element spacing increases, energy will be lost in the high side lobe resulting in an overall loss of gain. The optimum spacing is approximately one wavelength for a cos θ element pattern. Fig. 4 shows power gain versus element spacing for multiple bay antennas.

Azimuth Patterns

The ripple content of the azimuth pattern is dependent on the element pattern and the distance to the phase center as illustrated in Fig. 5.

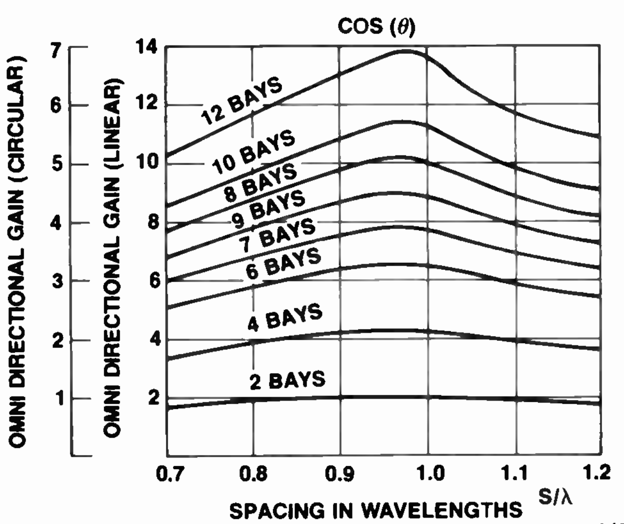


Figure 4. Power gain versus element spacing.

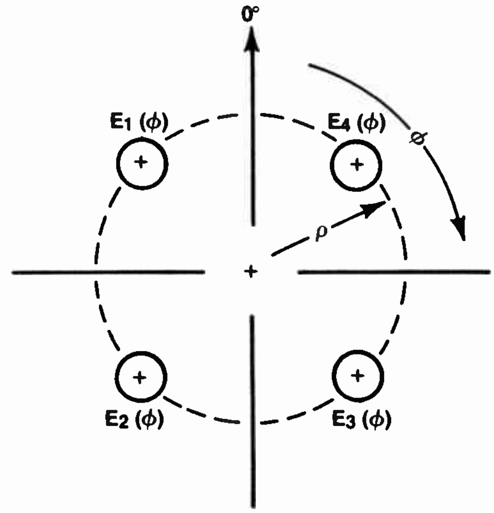


Figure 5. Azimuth pattern calculation.

$$F(\phi) = \sum_{n=1}^{N} E_n(\phi)\, A_n e^{\,j[(2\pi)\cos(\phi - \phi_n) + \delta_n]}$$

where: $E(\Theta)$ = element pattern
ϕ = azimuth angle
δ_n = phase of nth element
A_n = amplitude of nth element
ρ = distance to phase center (in wavelengths)

The value of ρ for panel antenna is dependent on the panel width. The value for a coaxial slotted antenna is dependent on the size of the coax line which in turn is dependent on the power handling capability of the line.

The value of ρ for waveguide slotted lines is dependent on the mode of operation which is a minimum of a half wavelength. This makes waveguide fed slot antennas practical only at UHF frequencies.

For some antennas, such as cross dipoles or turnstiles, the value of ρ is zero. Others such as panels, or slotted coax, or waveguide have a value of ρ from 0.25 λ to 2.0 λ. Fig. 6.

The ideal element pattern for a three-sided tower is E = cos (θ); and for a four-sided tower it is E = $\cos^2(\theta)$. The ripple content of four panels with a $\cos^2(\theta)$ element pattern and phase center spacing of ρ = 0.5, 1.5 is shown in Fig. 7.

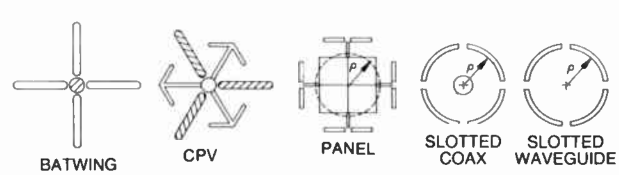


Figure 6. Azimuth pattern of four-side panel antenna viewed from the end.

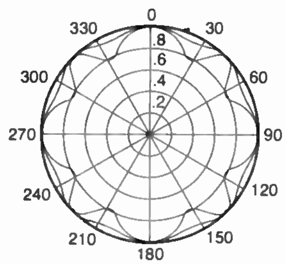


Figure 7. Arrangements of radiating elements. Azimuth versus relative field.

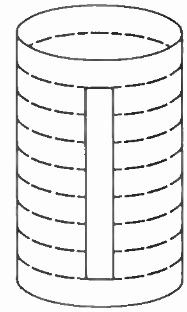


Figure 9. Slot in cylindrical sheet.

RADIATING STRUCTURES

Slot Antenna

The slot antenna is similar to a dipole.

Currents in the slot antenna spread out over the entire sheet (in which the slot is cut), and radiation takes place from both sides of the sheet.

The resemblance between the two becomes even more pronounced when it is recognized that the field patterns of the two will be equivalent if the physical dimensions of the slot and the cross-section of the dipole are the same.

Furthermore, the impedance of the slot is proportional to the admittance of the dipole of the same dimensions by the relationship

$$Z\,\text{slot} = \frac{35{,}467}{Z\,\text{dipole}}$$

and the bandwidth characteristics are essentially the same for both.

Actually, the above discussion is rigorously accurate only if the sheet is of infinite extent, but it is substantially correct if the edge of the sheet is half a wavelength from the slot.

Bending the sheet into a cylinder results in another form of slot antenna which also takes on characteristics of a stack of coaxial rings as shown in Fig. 9.

The slotted cylindrical tube can have an inner conductor and be coaxial in form or have no inner conductor and be a cylindrical waveguide.

Panel Antennas

A panel antenna has either dipoles or loops mounted $\lambda/4$ in front of a reflecting sheet. These antennas will have a directional [$(\cos(\theta)$ or $\cos^2(\theta))$] pattern in the azimuth and elevation planes depending on how it is mounted. Fig. 10 shows typical panel antenna configurations.

The pattern of a panel antenna is directional since the energy radiated from the element sees its image in the ground plane resulting in a $\lambda/2$ spaced end-fire radiator. The dipoles in the *doublet panel* are spaced $\lambda/2$ and are effective in eliminating downward radiation.

Traveling Wave Antenna

Traveling wave slot antennas can be either resonant (*standing wave*) or nonresonant (*traveling wave*). Most are bottom-fed. When the far end is shorted, an infinite standing wave is set up in the transmission line (either coax or waveguide). When the slots are appropriately sized and spaced, energy will be coupled out. The amplitude and phase developed across the aperture will result in a high-gain antenna. The bandwidth requirements for broadcast limits this technique to narrow band or single channel applications. Several traveling wave antennas are illustrated in Fig. 11.

The most common bottom-fed traveling wave array is the nonresonant slot array. The number of half-wavelength slots arranged about the periphery is dependent on the shape of the pattern desired. The slots, circumferentially arranged around the antenna structure, are usually displaced by $\lambda/4$ and spaced vertically one wavelength apart. The slot coupling from the bottom to the top increases, since the power

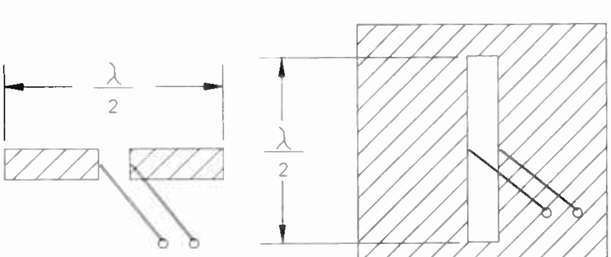


Figure 8. Dipole/slot equivalence.

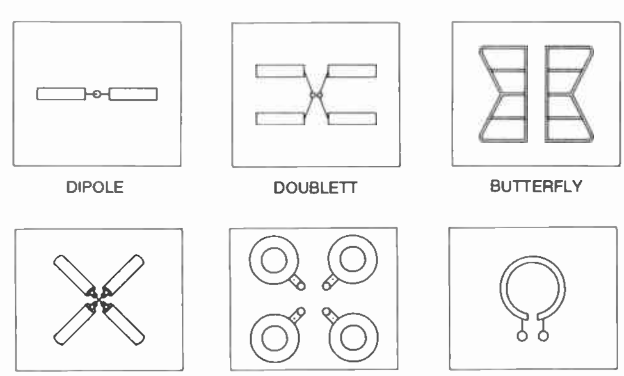


Figure 10. Panel configuration.

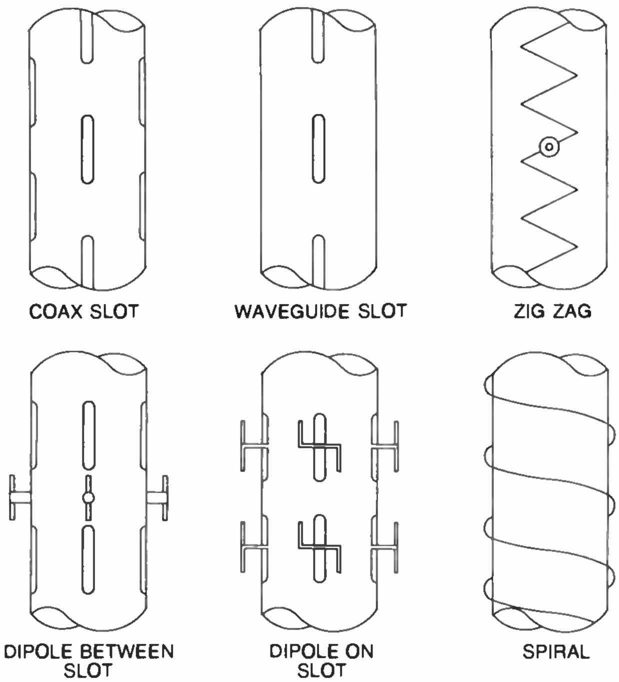


Figure 11. Traveling wave antennas.

radiated progressively reduces the power on the line. This method produces an amplitude taper across the aperture that provides null fill. With all nonresonant arrays, there is a small amount of energy left over. This is radiated in specially designed end-loaded slots or in the opposite polarization.

Circular polarization can be obtained by orienting successive slots 90° and exciting them in phase quadrature, or exciting a vertically polarized radiator located near the slot.

The linear polarized spiral (helix) is essentially a strip transmission line mounted over a ground plane; the spacing controls the radiation. Each element is a $\lambda/2$ radiator, and because of the orientation, the vertical components cancel. The choice of pitch angle and length along the circumference of the spiral, will result in very good horizontally polarized broad-side radiation.

The circular polarized spiral or multi-arm helix consists of multiple wires mounted around a large cylinder forming a current sheet. The optimum parameters are a complex function of the number of arms, the length of one coil, and the spacing to the support pole.

Excellent circularity can be obtained since in any heading the radiation is essentially that from a current ring.

ERP

In determining the ERP required to serve the area under consideration, there is a trade-off between using a low-power transmitter and a high-gain antenna, or high transmitter power and low antenna gain.

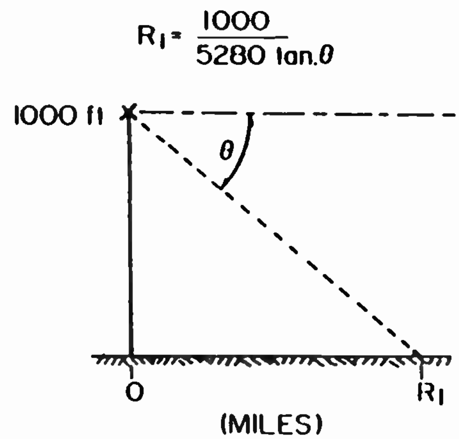


Figure 12. Elevation angle patterns for VHF antennas.

For VHF antennas, the transmitter power to antenna gain ratios are fairly well established. For channels 2 to 6, antennas usually use gain values from about four to six depending upon the length of the transmission line run. For the channels 7 to 13, band gain values vary from 12 to 18. For UHF antennas, higher antenna gains are required. Because of the much shorter wavelength at UHF frequencies, antenna with gains on the order of 20 to 30 can be built economically. At $0.10 per kW/hr, it costs $1,500/yr to generate 1 kW of RF power for a typical UHF TV station.

However, the higher gain results in narrowing the main beam. For a given transmitter input, the high gain antenna may sacrifice local coverage for more distant coverage. As the gain of the antenna increases, the beamwidth narrows. The position of the nulls and side lobes will move out from the tower.

The distance in miles from the tower is:

$$R = \frac{H}{5{,}280 \tan\theta}$$

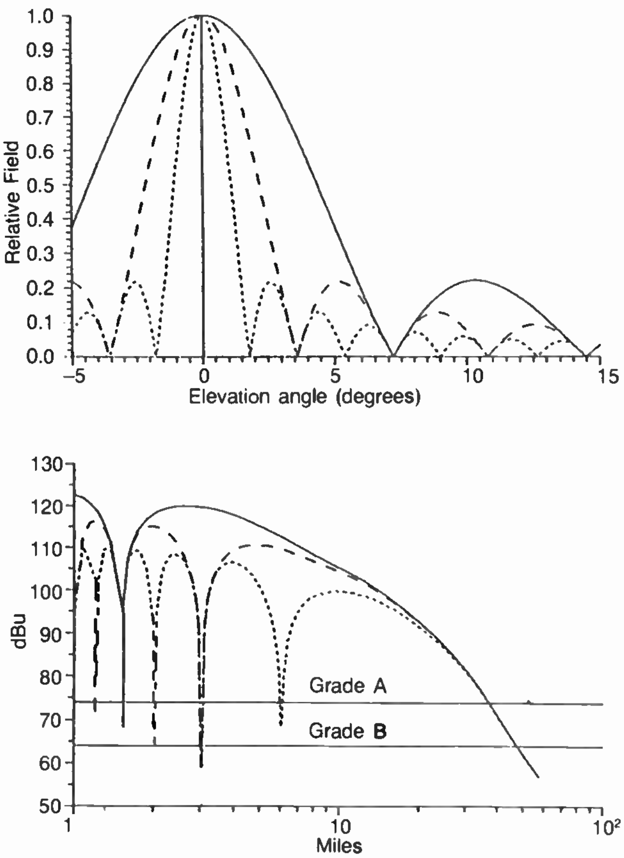
where: H is the height in feet and
θ is the elevation angle

Field Strength

The elevation pattern and single strength versus distance are shown in Fig. 13 for a UHF antenna on a 1,000 ft tower with constant ERP but different amounts of antenna gain.

Note that if the ERP and height above average terrain (HAAT) are held constant, there will be no change far in the far field signal strength. Fig. 13 shows there is no change in relative field for the three normalized antenna patterns from 0° to 0.5°. This corresponds to the radio horizon to 20 miles.

As the gain of the antenna is reduced, the main beam will widen from 1.6° (gain 40) to 7.2° (gain 10). The null will move from six miles to 1.6 miles. The signal strength near the antenna will increase by 10 to 20 dB. It should be noted that this increase makes a very strong signal (100 dBμ) much stronger (120 dBμ).



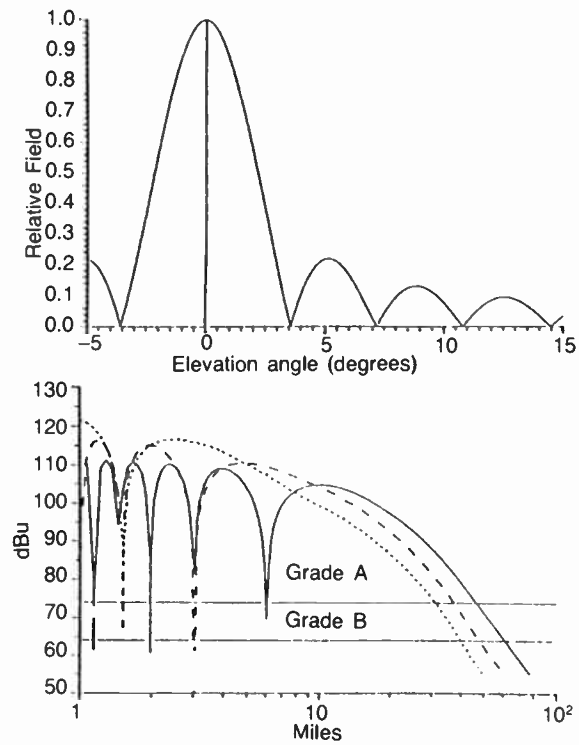
Curve	Transmitter Power	Antenna Gain	ERP
....	60 kW	40	2,400 kW
— — —	120 kW	20	2,400 kW
———	240 kW	10	2,400 kW

Figure 13. Signal versus antenna gain at UHF.

Therefore, if a higher gain antenna is contemplated, the local field intensities should be calculated, using the FCC (50, 50) propagation curves. It is desirable to maintain a 100 dBμ level over the important local area to be covered. Most UHF antennas are designed to accomplish this with an ERP of the order of 2.5 megawatt at 1,000 feet. In hilly terrain, it may be desirable to increase the field by 10 dB or more; and in heavily populated cities with large buildings, by 6 dB or more. The field intensity specified by the FCC for the different grade levels is shown in Fig. 14.

	LoV	HiV	UHF
Channel	2–6	7–13	14–80
City Grade	74 dBμ	77 dBμ	80 dBμ
Grade A	68 dBμ	71 dBμ	74 dBμ
Grade B	47 dBμ	56 dBμ	64 dBμ

Figure 14. Field intensities for specified contours.



Curve	Transmitter Power	Antenna Gain	ERP	HAAT
....	120 kW	20	2,400 kW	500 ft
— — —	120 kW	20	2,400 kW	1,000 ft
———	120 kW	20	2,400 kW	2,000 ft

Figure 15. Field intensity versus antenna height.

If fields of this order cannot be achieved with a high gain antenna, the transmitter power should be increased.

Relative field and field intensity versus distance are shown in Fig. 15 for a fixed antenna gain on a tower of 500, 1,000, and 2,000 ft.

POLARIZATION

Polarization is defined by the plane of the electric vector (E). Elliptical polarization is the most general form.

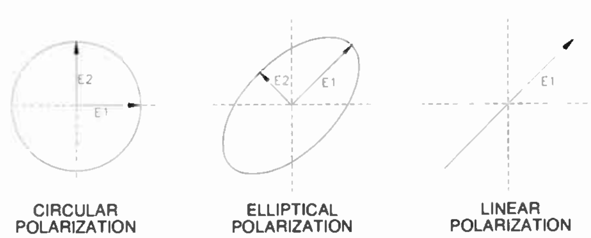


Figure 16. Polarizations.

In general, an elliptically polarized wave may be expressed in terms of x and y components given by:

$$Ex = E_1 \sin(\omega t - \beta z)$$
$$Ey = E_2 \cos(\omega t - \beta z + \delta)$$

The condition for linear polarization is when either E_1 or E_2 is zero with the orientation usually horizontal. The condition for circular polarization is when $E_1 = E_2$ and δ equals 90°, i.e., the two components must be quadrature in both space and time.

The transmitted wave of FM and TV broadcasting signals is specified by the FCC to be linear and in the horizontal plane. AM radio on the other hand is linear and in the vertical plane.

FCC regulations permit the radiation of circular polarization (CP) provided the component does not exceed the licensed ERP.

The possible advantages of using circular polarization for television are:

1. Less critical receive antenna azimuth orientation; permitting good reception on all types of indoor antennas including rabbit ears, whips, and rings.
2. Improvement in service and penetration because the two orthogonal fields (horizontal and vertical) have the same power, giving a power density twice that for a horizontal signal alone.
3. Improved coverage at the fringe area, due to twice the power density. Improvements on the order of 3 dB can be expected, moving the Grade B contour out by a calculated five to eight kilometers.

Many of the stations that changed to CP also moved their transmitting plants and changed the heights of their antennas, making a direct measured signal comparison between horizontally polarized (HP) and CP not possible. Experiences reported by stations using circular polarization indicate:

1. Receiving antenna orientation is less critical, specifically for rabbit ear and whip antennas.
2. Fringe area coverage has improved.
3. No ill effects have taken place.
4. No change in adjacent or co-channel interference has been reported.

BEAM TILT

Beam tilt is sometimes used to aim the main vertical beam tangential to the earth (toward the radio horizon), if the relative height of the antenna is such that substantial energy is being radiated above the horizon. The distance to the radio and optical horizon can be determined from Fig. 17.

Note that the height over the service area may not necessarily be the height over average terrain, especially in mountainous areas. Also, if a body of water limits the service area, or as in the case of Los Angeles, with the transmitter on nearby Mt. Wilson, it may be desirable to mechanically tilt the main beam to a point somewhat below the radio horizon.

In some cases, beam tilt may be desirable to improve

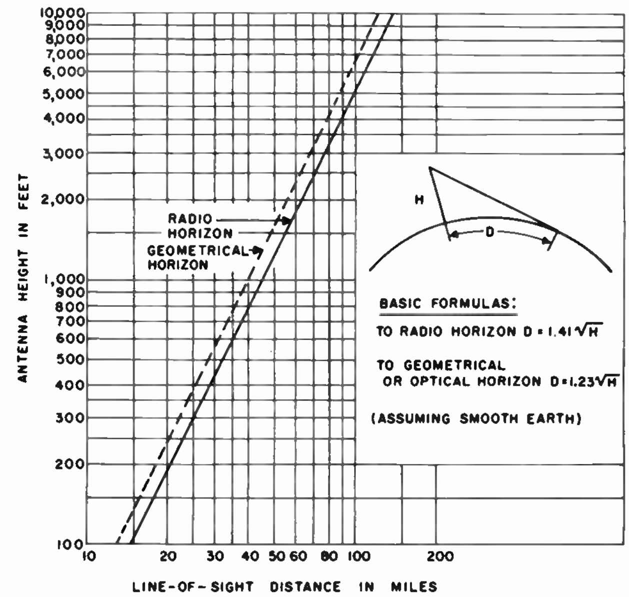


Figure 17. Distance to radio horizon versus height.

local coverage. Electrical beam tilt can be used but it reduces the power gain, especially for a higher gain antenna. However, the increase in local coverage is generally a more important consideration than the slight loss of power gain. Electrical beam tilt is accomplished by delaying the energy fed to upper elements or the upper section of a split-feed antenna, as shown in Fig. 18.

NULL FILL

The amount of null fill, and the number of nulls that need to be filled, depends upon how close the populated area is to the transmitter site. Allowance should be made for population movement towards the site in the future.

If the transmitter site is in the center of the population area or on the edge of it, consideration should be given to having null fill.

The exact amount of first and second null fill is not critical. Anything greater than 5% will usually result in signals much greater than *city grade*. The null depth (dB) in signal strength will correspond to the difference between the first null and first side lobe (dB).

The effects of null fill on signal strength is shown in Fig. 19.

A combination of null fill and beam tilt is shown in Fig. 20. A beam tilt of 1° will usually result in a 10 dB improvement in near-in coverage. Note that null fill and beam tilt have little effect on distant signal coverage.

POWER CAPABILITY

Power in television broadcast transmission systems is usually in terms of *peak power*, which is the

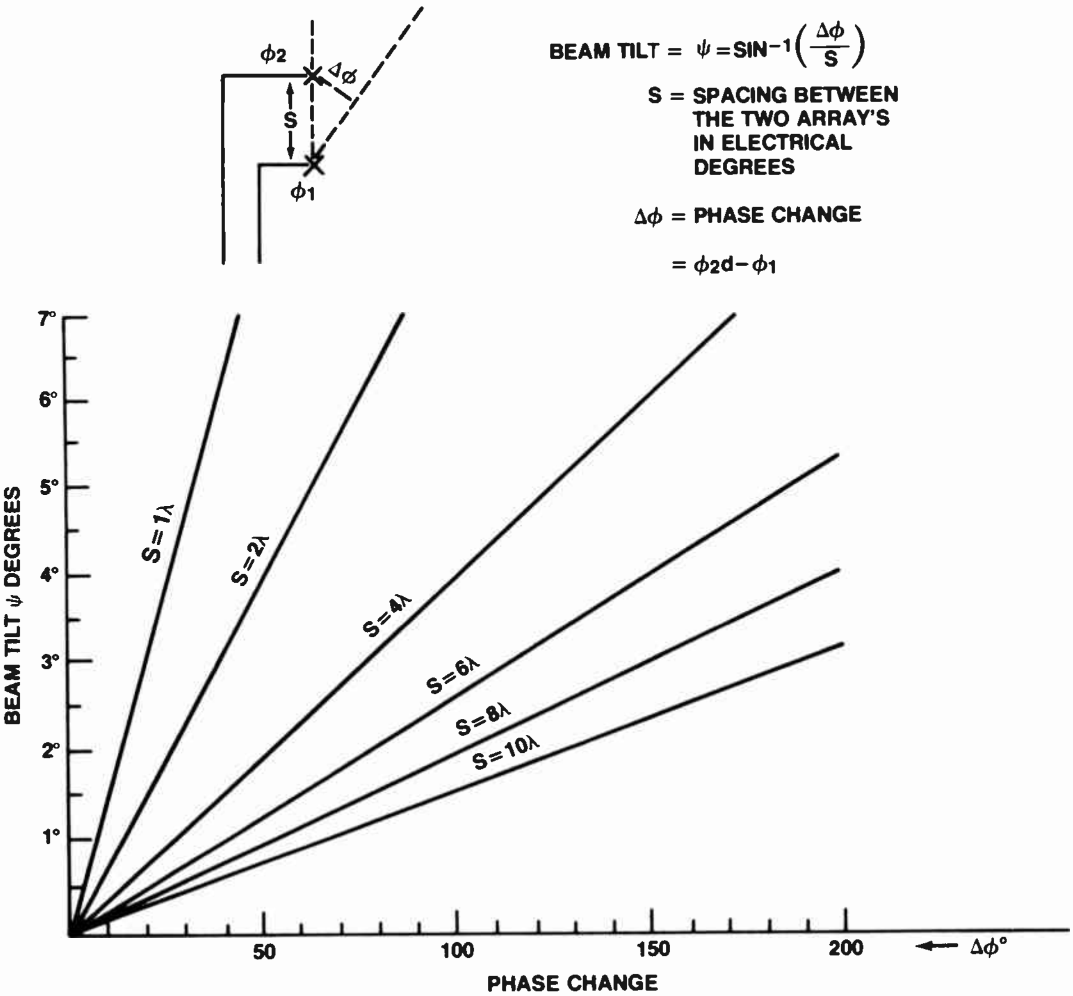


Figure 18. Beam tilt versus element phase change (Δ,φ).

instantaneous power developing during the peak of the synchronizing pulse in the visual transmitter. Since the black level signal is 0.75% of the total voltage value of the pulse, the black level power for a totally black picture is 56% of the peak-to-peak sync power. The duty cycle of the synchronizing pulses, both horizontal and vertical, adds about 4% to this power so that black level power is 60% of the peak TV power. Since the aural transmitter is usually 10% of the peak TV power, the total heating or continuous wave (CW) power of the TV signal is 66% of the peak TV power. The average power level (APL) of the video signal with typical program measures over a long period of time, 4.32 dB (37%) below the peak TV power.

The design of all TV antennas should allow for a sufficient safety margin to handle the peak-to-peak sync level, imperfections in the transmission line, VSWR, changes in pressure, and the aural power. Long transmission lines feeding the antenna will usually attenuate this figure by 10% to 20%.

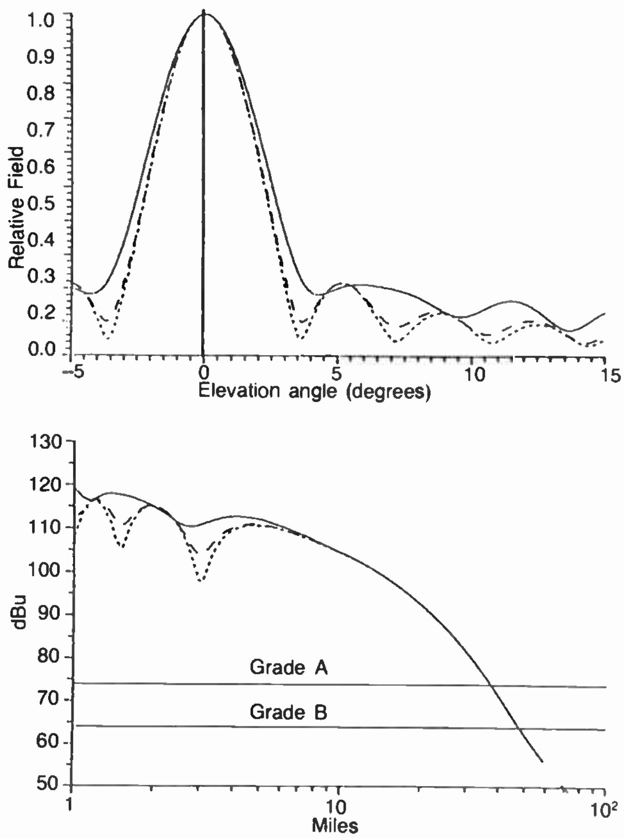
Unwanted Antenna Radiation

The antenna pattern is the product of the array pattern times the element pattern. The array pattern for one wavelength spacing has a downward and upward lobe as large as the desired main lobe. If the element pattern is not a $\cos(\theta)$ function, the element pattern will not drive the array downward lobe to zero, resulting in a very large signal in the vicinity of the tower base. Fig. 21 shows the resultant antenna pattern with a $\cos^{1.0}(\theta)$, where N = 0.0, 0.5, and 1.0.

Care should be exercised in making sure the element pattern has a distribution in the elevation plane of $\cos(\theta)$ or $\cos^2(\theta)$.

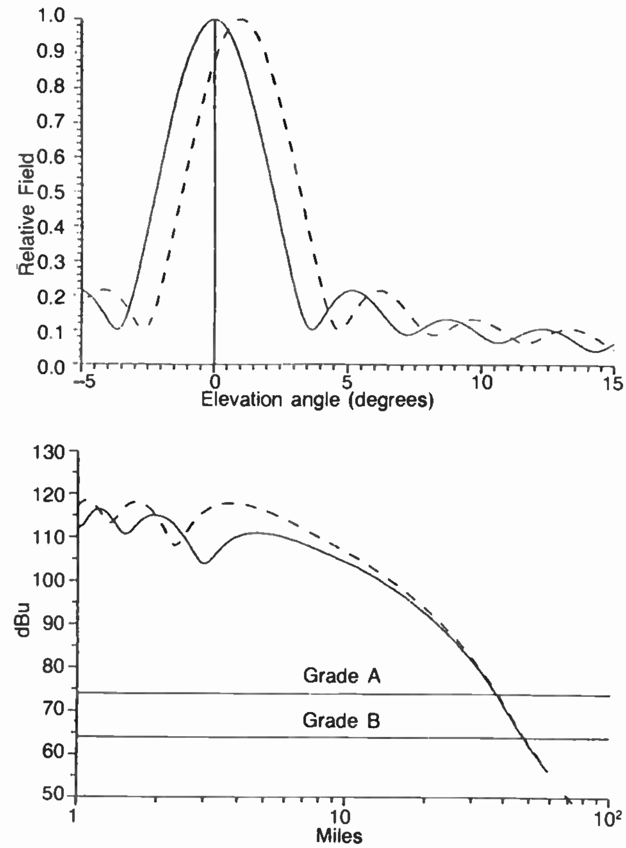
A well designed antenna system should not create a nonionizing radiation problem. The tower height is usually selected for maximum signal coverage and should therefore be high enough above the ground so the downward radiation is well below the safe ANSI levels.

Equally important would be the power density level



Curve	Null Fill	Transmitting Power	Antenna Gain	ERP
------------	5%	120 kW	20	2,400 kW
— — —	10%	120 kW	20	2,400 kW
________	20%	120 kW	20	2,400 kW

Figure 19. Null fill.



Curve	Beam Tilt	Null Fill	Transmitting Power	Antenna Gain	ERP
— — —	1	10%	120 kW	20	2,400 kW
________	0	10%	120 kW	20	2,400 kW

Figure 20. Null fill and beam tilt.

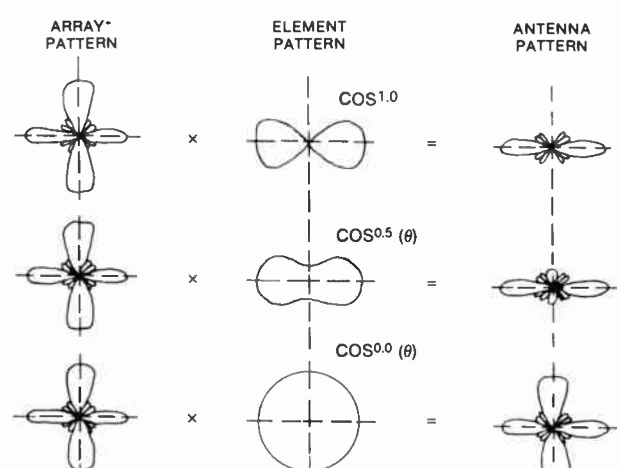
on nearby towers, which could result in tower maintenance personnel being exposed to high energy levels when in the main beam of the antenna on a nearby tower.

The distance in feet along the main beam to the 1 mW/cm^2 contour for maximum ERP and the height above ground to the 1 mW/cm^2 contour is listed in Fig. 22.

ANTENNA CHARACTERISTICS

Antenna Input Impedance

The primary purpose of an input VSWR specification is to insure a good match to the transmission line. If the mismatch is a great distance away, the reflected power may be of such magnitude that it travels back to the transmitter where it is will be re-reflected back to the antenna and radiate. It will then appear as a secondary image or ghost on the television picture.



*Six Bays, One Wavelength Spacing

Figure 21. Pattern multiplication.

Service	Low V	High V	UHF
ERP	100 kW	316 kW	5,000 kW
EIRP	164 kW	518 kW	8,200 kW
Distance to Nearby Tower	89 ft	163 ft	624 ft
Minimum Height Above Ground	43 ft	72 ft	305 ft

Figure 22. Distance to 1 mW/cm² level.

The image is delayed by twice the length of the transmission line.

Subjective experiments have established that the reflection at the carrier should be no greater than three percent of the incident voltage.

Empirically derived data of objectional reflection VSWR that will cause ghosting as a function of distance in microseconds from the transmitter to the antenna is shown in Fig. 23.

The dashed line represents the minimum detectable reflection readable on the window pulse.

Due to the concentration of energy, the VSWR at the picture carrier should be kept fairly low at this frequency. The values below visual carrier are not as critical since the slope in the receiver cuts off most of the energy below the carrier. The VSWR in the visual pass band should be as shown in Fig. 24.

There are two locations of reflections that will cause ghosts: the lower elbow complex and the combination of the upper elbow complex and the antenna. This is summarized below.

The upper elbow complex and the antenna reflection would be more objectionable than the lower elbow complex, because of its distance and displacement (1/△f) rather than its magnitude. A 5% (VSWR 1.10)

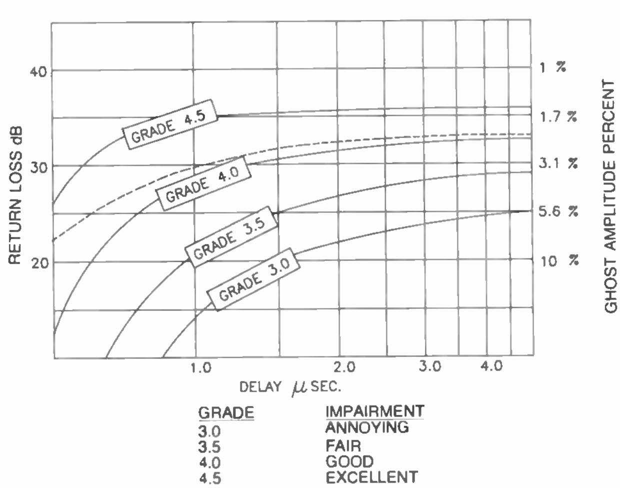


Figure 23. Signal impairment grades.

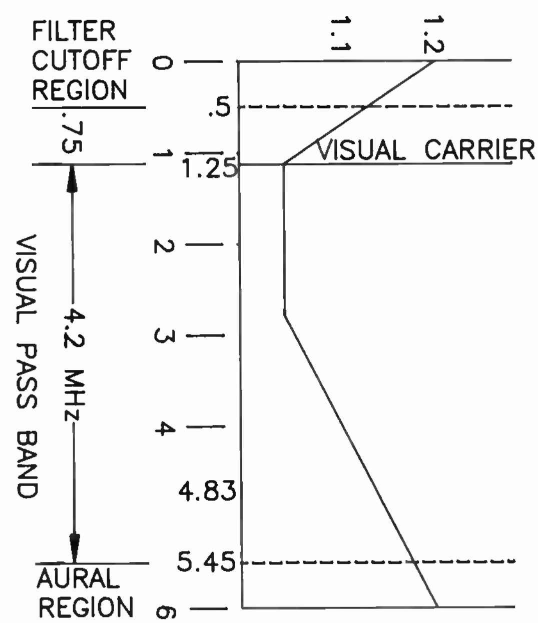


Figure 24. Acceptable VSWR levels.

echo would not be as objectionable at 0.2 μ sec as it would be at 2.0 μ sec, as shown in the following table.

Location of Reflection	Lower Elbow Complex	Antenna
Distance	100 ft	1,000 ft
VSWR	1.08	1.04
Reflection r	4%	2%
△ f MHz	4.96	0.49
Displacement in μ sec on TV Receiver	0.2	2.0

Transmission Line and Components

Testing the antenna system determines:

1. That the transmission line and components are properly assembled.
2. That the reflections from the antenna and other components at or near the tower top are sufficiently low so that no visible ghost occurs.
3. That the impedance presented to the transmitter will result in the maximum transfer of energy.

For an extremely broadband device, like a coaxial transmission line which is usually designed to cover the entire, or at least, a large portion of the TV band, the time-domain reflection (TDR) is the most effective test to determine if the line and components have been properly assembled.

Note that VSWR measurements made at the input to the transmission line measure the combined input impedance at that point only, and not the discreet reflections that can cause ghosts.

The TDR, on the other hand, will measure the reflection on a time rather than frequency basis so the exact position and magnitude of the reflection can be determined.

A high VSWR at this point is an indication of a mismatch between transmission line and antenna. For example, a VSWR of 1.22 is a return loss of 20 dB e.g., 99% of the power is transferred to the antenna system and 1% is reflected.

There are usually two sources of reflection assuming that the line itself is reflectionless:

- Lower elbow complex—V1
- Upper elbow complex and antennas—V2

When measured at the input to the system these can be thought of as two vectors. Each vector individually will rotate about the preceding vector (Fig. 25) at a frequency corresponding to the distance from the point of measurement.

The frequency of rotation can be calculated as follows:

$$\Delta f = \frac{496^*}{L} \quad \text{coax}$$

$$\Delta f = \frac{496^*}{L} \quad (\lambda_o/\lambda_g) \quad \text{waveguide}$$

*Half the velocity of light. L in feet.

Close reflections (approximately 100 feet) will have a period equivalent to 5 MHz; distant reflections (approximately 1,000 feet) will have a period equivalent to 0.5 MHz. The individual and combined reflections are shown in Fig. 26.

The use of a *network analyzer* with TDR is effective in determining the location and magnitude of the antenna system VSWR and sources of reflection. Measurement of VSWR in both frequency and time domains for an antenna and 1,700 feet of transmission line (waveguide) is shown in Fig. 27 A and B. The near and far reflections are quite apparent in the frequency domain.

At UHF frequencies where the transmission line guide wavelength is two to three feet, a 50°C temperature change on a 1,000 foot tower can result in electrical length changes of λ/4 to λ/2 wavelength. This is significant in dual feed systems where the differential change can cause beam tilt.

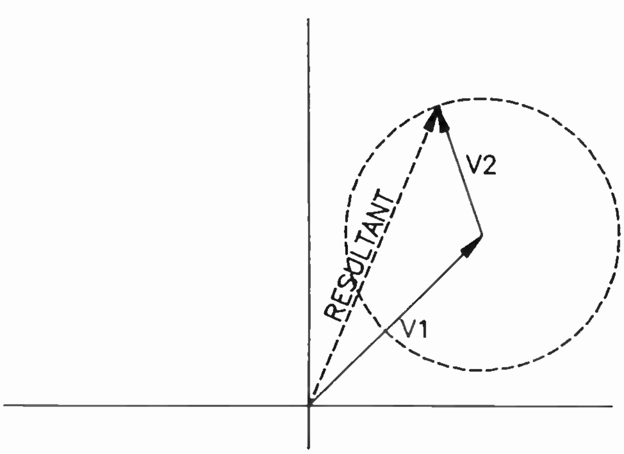


Figure 25. Vector relationship.

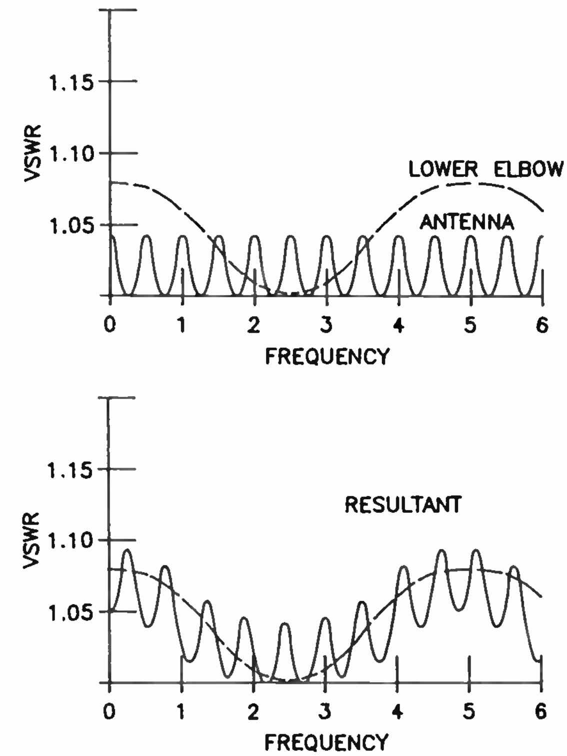
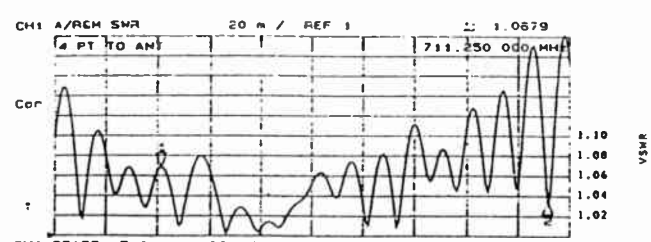
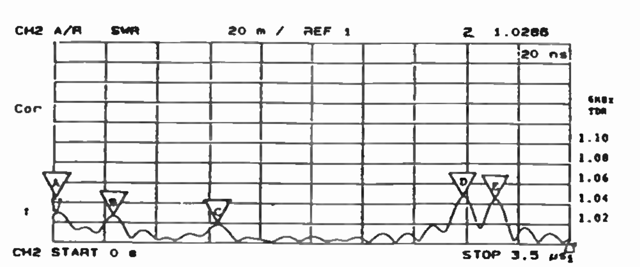


Figure 26. Example of combined reflections.



A. Frequency Domain.



B. Time Domain.

A - Gas Barrier
B - Lower Elbow Complex
C - Transition
D - Upper Elbow Complex
E - Antenna

Figure 27. VSWR of antenna and waveguide (1,700 ft) frequency domain and time domains.

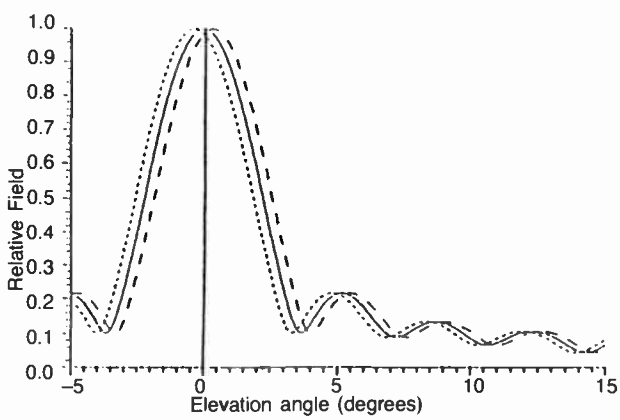
Deflection and Wind Load

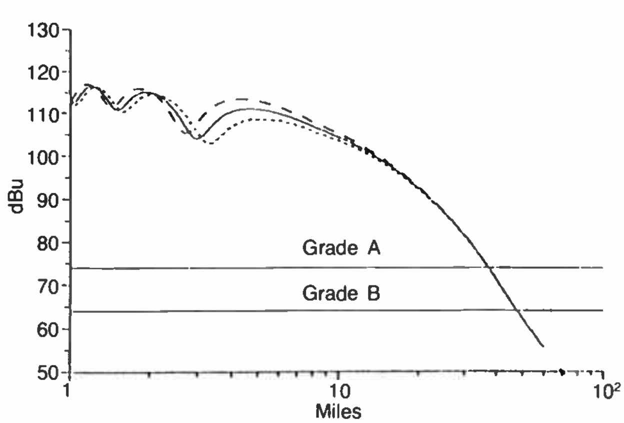
Guy tension in guyed towers is usually adjusted so that the tower deflects as a straight member.

Towers for broadcast service, when so specified, are designed for maximum deflection of 0.5°, which means that the top plate will deflect this amount for the maximum wind velocity. For instance, a 40-pound tower will thus deflect 0.5° for a 100 mph wind. Since tower deflection varies as the square of the wind velocity, the deflection will be 0.125° for a 50 mph wind.

Structurally a free-standing antenna can be considered as a cantilever beam in which the deflection increases toward the end. Antenna deflection is stated as the angle from the vertical of the chord that connects the base to the top of the antenna.

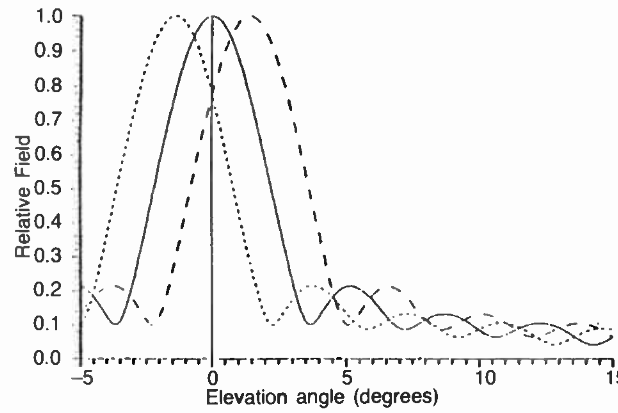
The movement of the antenna beam and the variation in signal strength for a 50 mph and 100 mph wind is shown in Fig. 28 and summarized in Fig. 29.

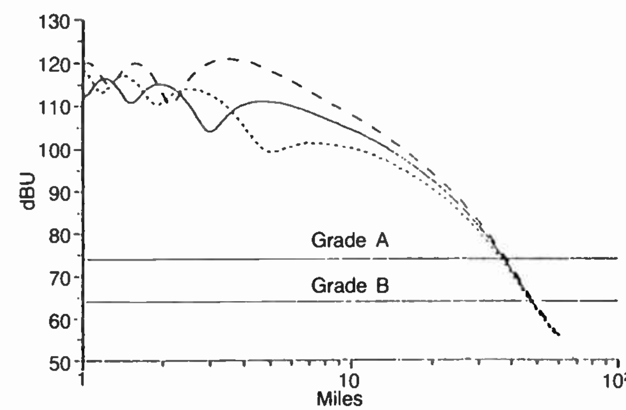




Curve	Wind	Transmitter Power	Antenna Gain	ERP
------------	Away	120 kW	20	2,400 kW
— — —	Forward	120 kW	20	2,400 kW
————	0	120 kW	20	2,400 kW

Figure 28. Field strength changes with a 50 mph wind.





Curve	Wind	Transmitter Power	Antenna Gain	ERP
------------	Away	120 kW	20	2,400 kW
— — —	Forward	120 kW	20	2,400 kW
————	0	120 kW	20	2,400 kW

Figure 29. Field strength changes with 100 mph wind.

The 50 mph wind condition is one that may occur 25 times a year at a 1,000 ft. elevation, and about four times a year at a 500 ft. elevation (Fig. 30).

The 100 mph wind is a design limit figure which rarely occurs, and one during which there would probably be little television viewing. Most outdoor receiving antennas would probably be severely damaged in such a wind, and power service seriously curtailed.

Wind Velocity (mph)	50	100
Wind Load (psf)	10	40
Deflection (in) Antenna at Top	4.9	19.6
Deflection (degrees) Chord Bottom to Top Tower Antenna and Tower	 0.227 0.125 0.352	 0.914 0.500 1.414

Figure 30. Antenna and tower deflection.

Antenna Feed Systems

The feed system of a television broadcast antenna is that portion of the transmission system having its input at the antenna terminal (which is at the top of the vertical run of coaxial transmission line in the tower), and its output at the radiating elements.

Most antenna gain specified by the manufacturer takes the losses of the feed system into account. Therefore, when system gains are calculated, the feed-system loss can be excluded.

In the television broadcasting field, three types of feed systems are in wide use. They are the branching, standing wave, and traveling wave feed systems. Each meets a specific need.

Branching Feed Systems

Branching feed systems are necessary for all-band antennas (i.e., an antenna that covers the full VHF or UHF band), and since they are center-fed, they eliminate beam steering or beam tilt.

This feed system progressively divides the power as shown in Fig. 31. It is used when the radiators are individual elements, each with their own terminal—such as a dipole or panel. The system shown in (Fig. 31A) will have a narrower impedance bandwidth than (Fig. 31B) since, for economy, one eight-way power divider is used. The junction impedance is $Z_j/8$. If the element impedance is 50 Ω, the power divider must transform 6.25 Ω to 50 Ω. The system shown in (Fig. 31B) uses two-way power dividers and is sometimes called a "corporate feed." Although it has a broader bandwidth, it is less economical since there are seven power dividers with additional interconnecting cables. Null fill and beam tilt is accomplished by changing the length of the feed cables or using unequal power dividers, or both.

A problem with branching feed systems is the presence of the feed line in the aperture of the lower elements. The feed lines can cause reradiation and distort the azimuth pattern. The branching feed system can be more effectively used with panel antennas that require a center support tower or mast where the transmission lines are behind the antenna.

Standing Wave Feed Systems

A coaxial or waveguide transmission line can be shorted at the far end, resulting in standing waves along the length of the line. If slots or coupling probes are appropriately sized and positioned, the RF energy can be radiated and a desired amplitude and phase distribution across the aperture can be obtained. This resonant array structure has a desirable feature; all coupling parameters are the same and equally spaced. Its disadvantages are a narrow bandwidth, and it can only be used at high VHF and UHF frequencies. (Fig. 32)

Traveling Wave Feed Systems

The traveling wave feed system operates on the principle of a gradual attenuation (radiation) of the input signal as it progresses from the input along the aperture of the antenna. An application of this principle is the *slot antenna* or *spiral antenna*. (Fig. 33)

Fig. 33A shows the principle of this feed system using short rod radiators to illustrate the theory. A number of uniformly spaced radiators per wavelength are loosely coupled to a coaxial line. Because of the number of radiators and the relatively slight reflection between them, the effect is essentially that of a uniform loading. The result is a uniformly attenuated traveling wave in the line. Since a traveling wave has a linear-phrase characteristic, the excitation of each successive radiator will be lagging from the previous one by an amount which depends on the spacing between the radiators and the velocity of propagation in the line. If the radiators are alike, their currents will have the same phase relationship as the excitation. Thus, the radiating currents will be successively lagging, and repetition of phase occurs after every guide wavelength.

To obtain an omni-directional pattern, the radiators, instead of being in line, can be moved around the

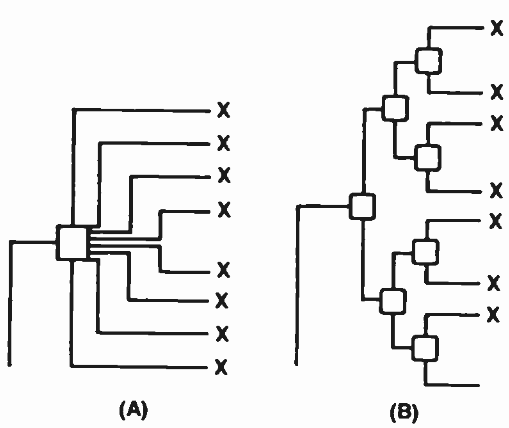


Figure 31. Branching feed systems.

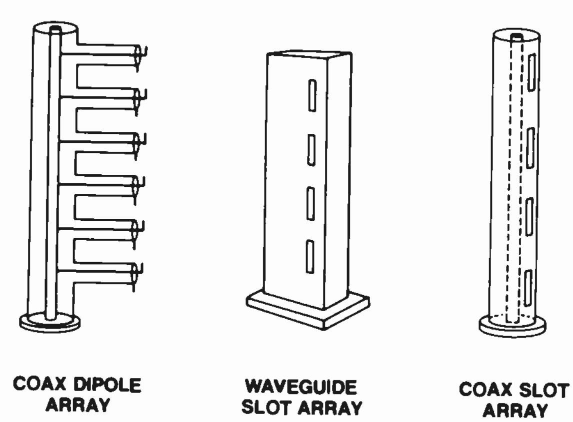


Figure 32. Slotted arrays.

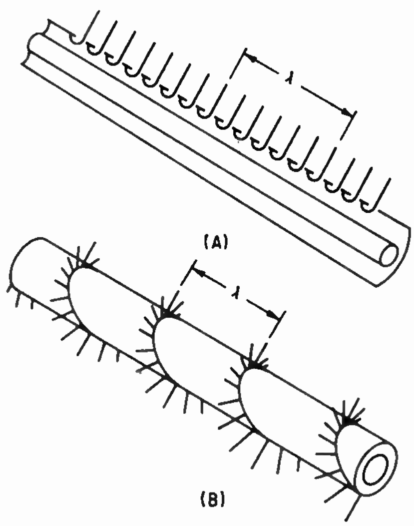


Figure 33. Traveling wave excitation.

periphery to form a "spiral" as shown in Fig. 33B. For a horizontal main beam, the pitch of the spiral has to be equal to the guide wavelength in the transmission line. In this arrangement all the radiators in any one vertical plane on one side are in phase, and the phase difference between radiators in different planes equals the azimuth angle difference between the planes. That is, the phase rotates around the periphery. The rotating phase produces a rotating field. Because of the relatively small amount of current change from layer to layer, an onmi-directional pattern is produced.

ANTENNA TYPES

The types of antennas available can be grouped into the following categories:

- Top or side mounted
- Linear or circular polarization
- High power, low power

	Horizontal Polarization		Circular Polarization	
	Top Mtd	Side or Tower Mtd	Top Mtd	Side or Tower Mtd
VHF Low Band	Batwing	Butterfly Doublett H-Panel	TDM	
VHF High Band	Batwing Traveling Wave	Butterfly Doublett H-Panel	Spiral Slot&Dipole Slot&Director	TCP CBR Arrowhead Ring
UHF	Slot Coax Wave Star Trasar	Doublett	Slot&Z Dipole Slot&Dipole Slot&Director	Slot&Dipole Slot&Dipole Slot&Director

Figure 34. Commonly used television transmitting antennas.

The prime purpose of the tower is to support the antenna. The antenna should have the required radiation characteristics to deliver a satisfactory signal to the viewer. Since it is desired to cover as large a viewing area as possible, tall towers are used. It is not uncommon for antennas to be mounted on 1,000 feet, 1,500 feet, or 2,000 feet towers. These tall towers are guyed triangular towers with face dimensions of five to ten feet.

The cost of a tower is heavily dependent on the wind load presented by the tower, the transmission line, and the antenna. For VHF antennas, the antenna wind load is the most significant parameter. For UHF antennas the transmission line or waveguide is the more significant factor.

The length of the antenna is related to the channel, or wavelength, and the gain requirements. (Fig. 35)

The ideal omni-directional antenna would be a small-diameter, infinitely stiff pole. The smallest cross section that can resonant is $\lambda/2$. Therefore, the antenna cross-section is also dependent on the wavelength.

At low VHF frequencies the resonant half-wave element of eight feet is sufficiently large so that a support pole can be used in the center with the antenna mounted outside (such as a batwing antenna). On the other hand, at UHF frequencies the resonant half-wavelength of one foot is so small that the resonant elements must be outside the support pole, or the support pole itself must be a radiator like a slotted coax or waveguide array.

Ideally, omni-directional antennas should be fed from the bottom, so there are no feed lines in the aperture of the antenna to distort the pattern. If feed lines are required for the upper elements, they should be on the inside. Of course, this is not possible for waveguide arrays. A conflicting requirement is that the diameter of the feed line be large enough to satisfy the power handling requirements, yet small enough not to create pattern distortion.

To minimize pattern distortion by the tower, the antennas should be mounted on top of the tower rather than on the side.

The wavelength of high VHF and UHF antennas is such that the resonant half-wavelength is too small to be a satisfactory structural member.

The radiating element for many UHF antennas is either the slot or dipole excited from an internal source such as coaxial or waveguide.

	Wavelength	Length for Maximum Gain	Maximum Gain
LowVHF	16 ft	100 ft	6
HighVHF	5 ft	90 ft	18
UHF	2 ft	120 ft	60

Figure 35. Dimensions of typical antenna.

Most slotted arrays are traveling wave structures which are bottom fed and contain many slots. The greater the number of slots the narrower the bandwidth, and the greater the beam steering or tilt between the visual and aural frequencies.

Some slotted arrays use the outer conductor (coax), or the waveguide itself as the structural member.

Panel antennas are fed from the back of the panel and require a secondary structural member to support the panels. This can be the tower itself. Three or four panels must be mounted around the tower to produce an omnidirectional pattern.

Panel antennas, when fed with a corporate type feed harness, can preserve the wide bandwidth of the dipole and cover the full high VHF band (174–216 MHz) or UHF band (470–800 MHz).

VHF ANTENNAS

Linear Polarized

Superturnstile/Batwing

The first antenna developed for commercial service was the Superturnstile. It consists of a central sectionalized steel pole upon which is mounted the individual radiators, or *batwings*. These radiators are mounted in groups of four around the pole in north-south and east-west planes to form a "section," and the sections are stacked one above the other to obtain the desired gain. Fig. 36 illustrates this construction.

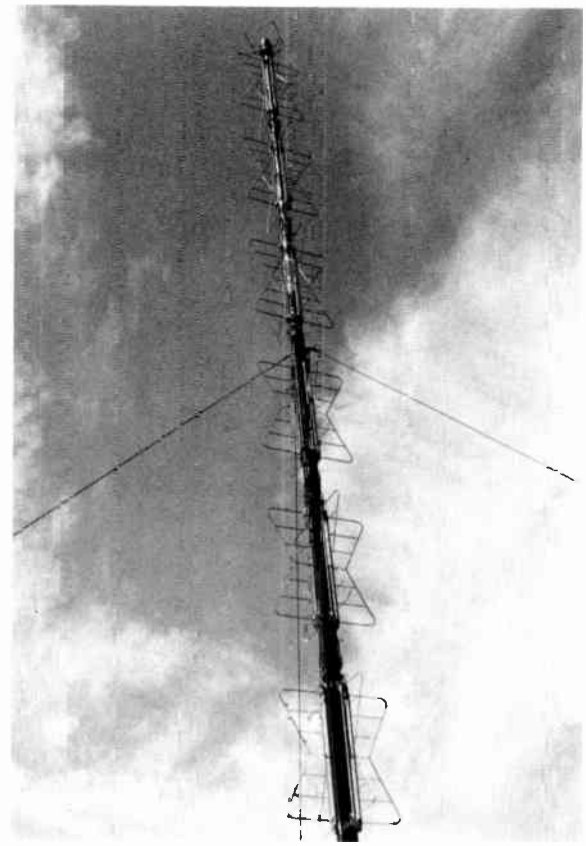


Figure 36. Superturnstile/Batwing antenna. (Courtesy of Dielectric)

Each of the radiators at the batwing antenna is fed by its own feed line. The impedances are carefully matched. The feed lines, in turn, are combined at junction boxes; which perform the dual function of feeding power simultaneously to all feed lines, and of transforming the combined impedance of these lines to that of the transmission line. This latter function is achieved by the use of three-stage transformers immediately below the junction box.

At the base of the antenna, a combining network is used when there are more than two junction boxes. These networks accomplish power division between portions of the antenna, if desired. Batwing antennas are manufactured in various gains from three to 12 for channels 2 to 6 and, gains of six to 18 for channels 7 to 13. They can also be designed for various types of null fill and they have been used in stacked and candelabra installations.

The batwing antenna is wideband and can be used for two channels. A number of them are operating at channels 4 and 5 or channel 6, an FM channel, and also in various combinations in the channel 7 to 13 range.

Traveling Wave

The traveling wave (TW) antenna is a slot antenna with a traveling wave feeding the slots. (Fig. 37).

The TW antenna is a coaxial line, with pairs of slots in the outer conductor spaced at intervals of a quarter wavelength throughout its length. Probes at the center of each slot distort the field within the line to place voltages across the slots. These, in turn, drive currents on the periphery, setting up a radiated field. Attenuation of the signal, by withdrawal of a portion of the power at each slot, reduces it to a very low value at the upper end of the antenna. There, a special pair of slots, designed to match the line, radiates the remaining power.

Operation of the TW antenna can be better understood if the section of the aperture having pairs of slots are recognized as being, in effect, dipoles.

Successive pairs of slots are alternately in one plane and in another at 90° to it, so that the antenna can be simulated by stacked dipoles with a 90° angle between successive layers.

In a given plane, reversal of the direction of feed every half wavelength (by placing the probes on opposite side of the slots), together with the half-wave change in phase of the signal as it passes along the aperture through this distance, results in all the "dipoles" in that plane being fed in phase.

The same action takes place in the other plane, except that they are fed 90° out of phase with the first plane, owing to their 90° displacement along the antenna.

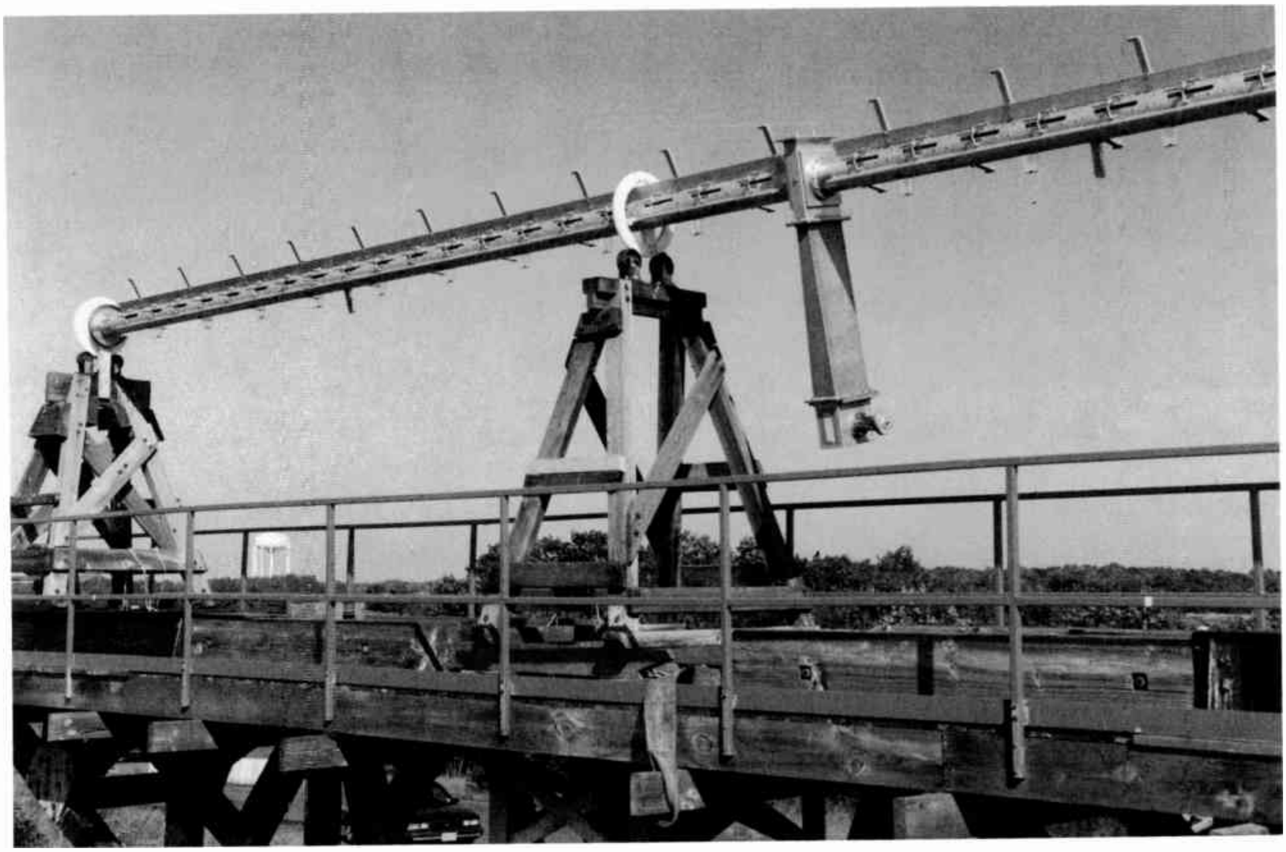


Figure 37. Traveling wave antenna. (Courtesy Dielectric)

Each plane of dipoles radiates essentially a figure-eight pattern. Since the planes are fed in quadrature, addition of the patterns results in an omnidirectional pattern. Because of the circular cross-section and the lack of obstructing radiators, the resulting horizontal pattern is almost nearly circular.

Panels

The panel antenna is designed to wrap around the tower. Four panels are needed for a square tower where each panel must radiate a $\cos^2(\theta)$ element pattern. For a triangular tower, each element must radiate a $\cos(\theta)$ element pattern for an omnidirectional pattern.

A wide variety of azimuth and elevation patterns can be obtained by using fewer panels or changing the power division to the panels.

The panels are 0.7λ to 0.9λ in vertical length and spaced approximately one wavelength.

The radiating elements may be either single dipoles like the H-panel *rhombus* in Fig. 38A or *delta dipole (butterfly)* in Fig. 38B which is essentially a folded back batwing or dual dipoles.

The *dual dipole* in Fig. 38C is designed to minimize downward radiation and consists of a pair of dipoles spaced a half wavelength apart.

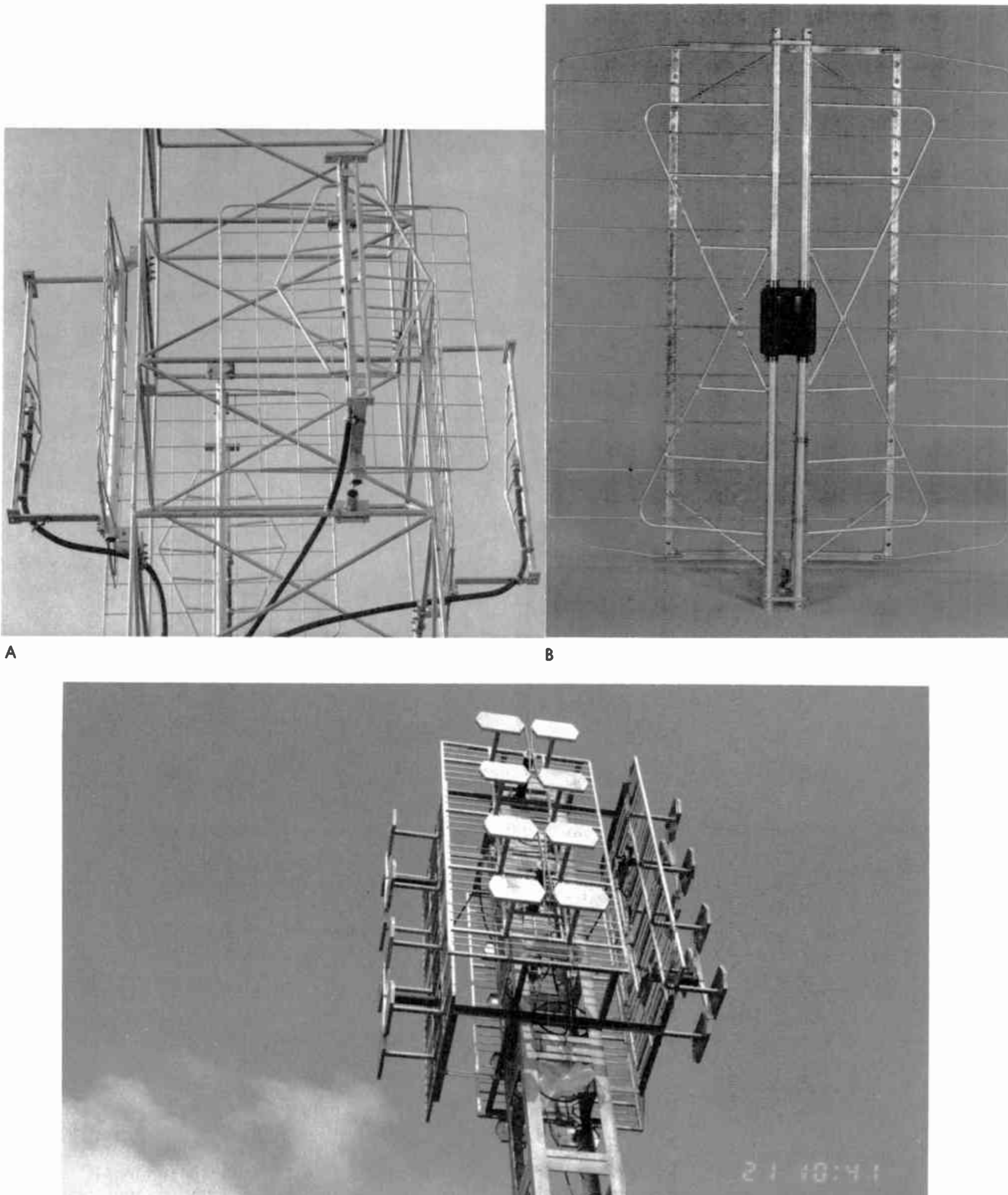
The impedance bandwidth at the panel antenna is very wide and capable of handling more than one TV channel. The bandwidth can be further improved with the four-side configuration by using phase rotation. That is, 90° between each panel in the bay. The impedance at the four-way power divider will be conjugate for every two panels, resulting in wider bay impedance bandwidth than the element itself.

Circular Polarized

Transmission Dual Mode (TDM)

The TDM shown in Fig. 39 is a circularly polarized antenna for channels 2 to 6, and is designed for tower top mounting. It is capable of replacing an existing six-bay Superturnstile without any increase in tower windloading.

The TDM antenna utilizes seven layers of radiators in a slanted dipole configuration, with three radiators mounted symmetrically around the pole per layer. Each of the three radiators is fed in phase by a single feedline. Only 21 feedlines are required for the entire antenna. A branch type feed system is used to achieve excellent vertical pattern stability. One junction box feeds the upper four layers, and another feeds the lower three layers, each box fed by a 3⅛" line. This



A B C

Figure 38. Panel antennas. (A) H-panel Rhomous (Dielectric), (B) Delta dipole "Butterfly" (Harris), (C) dual dipole (MCI).

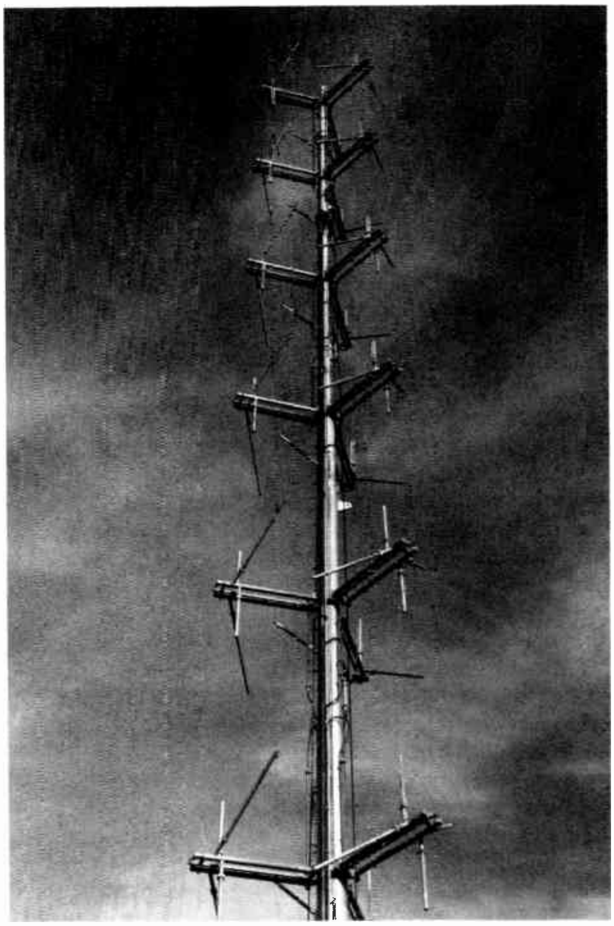


Figure 39. TDM antenna. (Courtesy Dielectric)

Figure 40. CPV antenna. (Courtesy Harris)

feature allows for standby capability in the event of weather-related or other damage. The TDM can be supplied with deicers and/or radomes depending upon environmental requirements.

The unique design of the TDM provides for excellent pattern circularity and axial ratio. Axial ratio measurements are performed on a complete full scale as-built antenna standing vertically at the factory. Since all elements are excited, the mutual effects of adjacent elements are considered.

Circularly-Polarized "V" (CPV)

The CPV antenna shown in Fig. 40 is a circular, polarized, top-mounted antenna consisting of three cross "V" dipoles mounted at 120° intervals around a vertical mast. The dipoles are segmented by three vertical grids like a corner reflector, used both for isolation and to shape the element for good circularity. The cross dipoles are fed in phase quadrature and radiate circular polarization from each element.

A branching feed system is used with the lines fed up the mast. Null fill and beam tilt are accomplished by changing the electrical length of the feed cables.

Panel

The cross dipole panel antenna consists of dual dipole feed in both space and time quadrature, a necessary condition for circular polarization.

The TCP (transmission circularly polarized) series uses dipoles (shown in Fig. 41A), and are in the form of a clover leaf mounted on front of a ground plane or panel. The four elements generate the required $\cos^2(\Theta)$ element pattern. By tilting the ground plane edges forward, as a partial corner reflector, further shaping can be obtained.

The CBR (circulary basket reflector) series in Fig. 41B consists of fat dipoles mounted in a basket. The element pattern is controlled by the diameter and depth of basket. The ring panel antenna in Fig. 41C is a large resonant loop mounted against a ground plane. The ring is approximately one wavelength circumference and hence radiates a circular polarized wave in all planes perpendicular to the panel.

The arrowhead dipole in Fig. 41D is similar to the other cross dipoles. The shaping of the element pattern

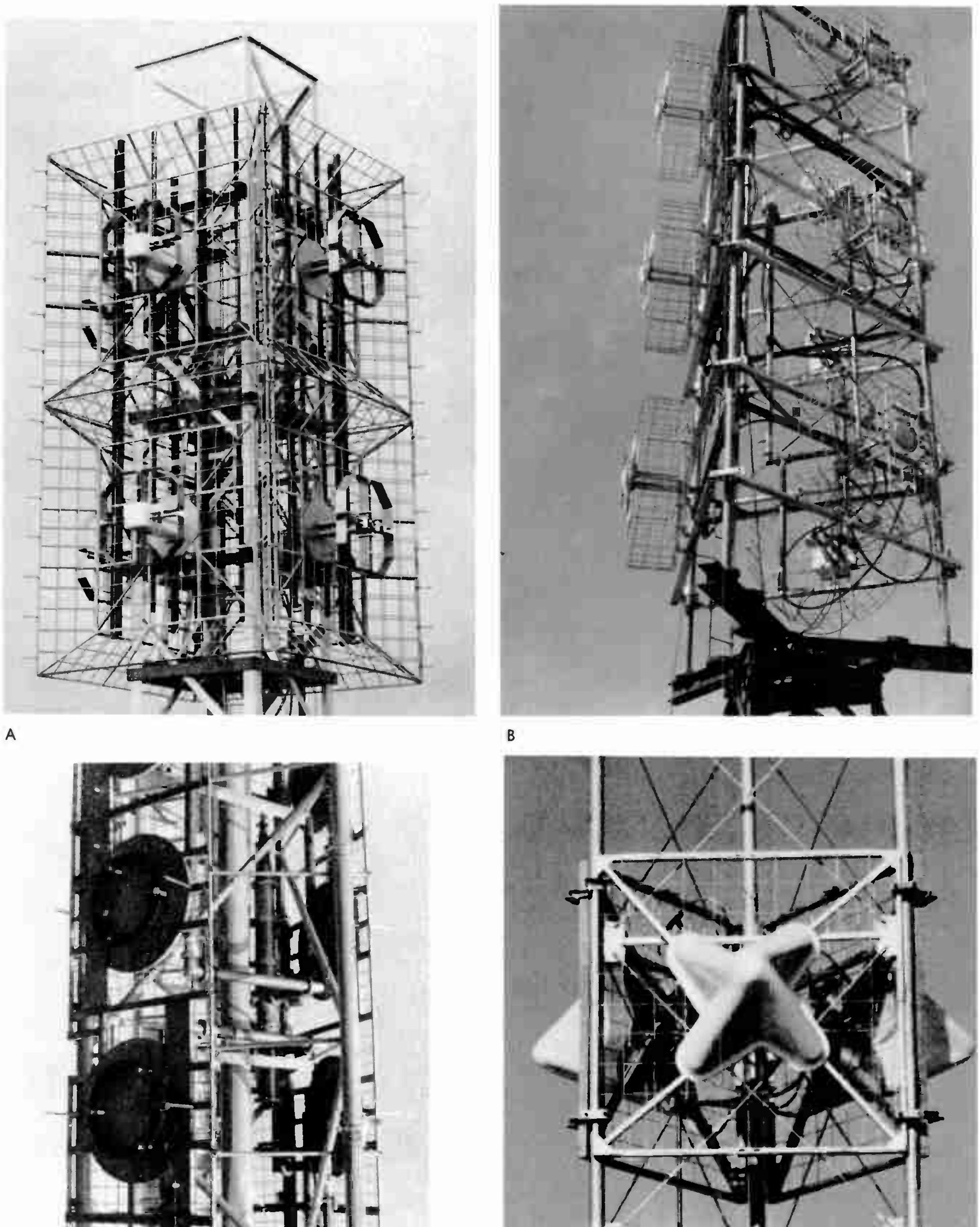


Figure 41. Panel antennas. (A) TCP (Dielectric), (B) CBR (Harris), (C) ring panel (Lampro), (D) arrowhead dipole (MCI).

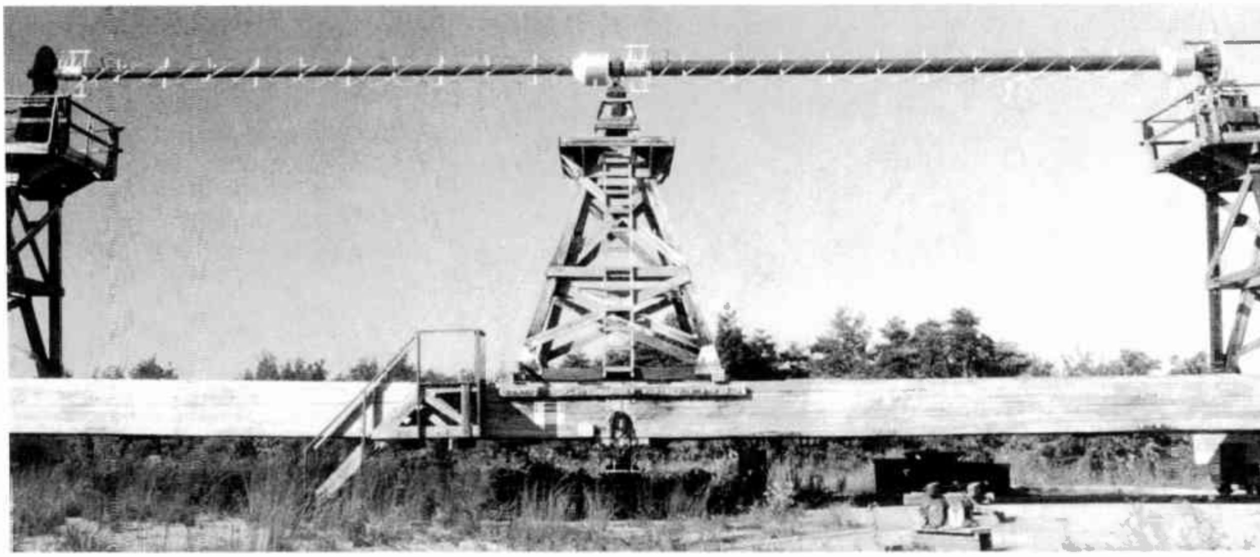


Figure 42. Spiral antenna on test range. (Courtesy Dielectric)

is accomplished by tilting back the dipole to broaden the E-plane pattern.

Spiral

The spiral antenna is designed for tower top mounting. The antenna is composed of sections, each operating under the principals of a traveling wave antenna. Each section consists of two or more stainless steel coils (radiating elements) wound around the supporting pole. A "beltline" feed system is used in each section to excite the radiating elements. The phase initiating the traveling wave is a function of the number of coils used.

Each section is terminated with radiating end loads on each coil to radiate the remaining energy. These loads minimize the reflections of energy from the far end of the radiator back toward the input, which would affect the traveling wave illumination and distort the pattern.

Vertical pattern characteristics such as gain, null fill, and beam tilt are determined by a number of factors: the number of wavelengths "wrapped" around the pole, the spacing of the radiators off the pole, the pitch (angle of wrap) around the pole, and the phasing between the vertical sections. The antenna is deiced by low-voltage, high-current dc to heat the radiating coils. A spiral antenna under test is shown in Fig. 42.

Slotted Coax

The circular polarized slotted antenna in Fig. 43 is similar to the linear polarized slotted antenna with the addition of vertical radiators.

The vertical radiators are located outside the array and the energy coupled to them from the slot is in phase quadrature.

Any polarization from elliptical to circular can be obtained by adjusting the length of the radiators.

UHF ANTENNAS

Linear Polarized

Slotted Coaxial

The UHF *pylon antenna* in Fig. 44 is a coaxial transmission line with radiating slots in the outer conductor. The number of slots (per layer) around the

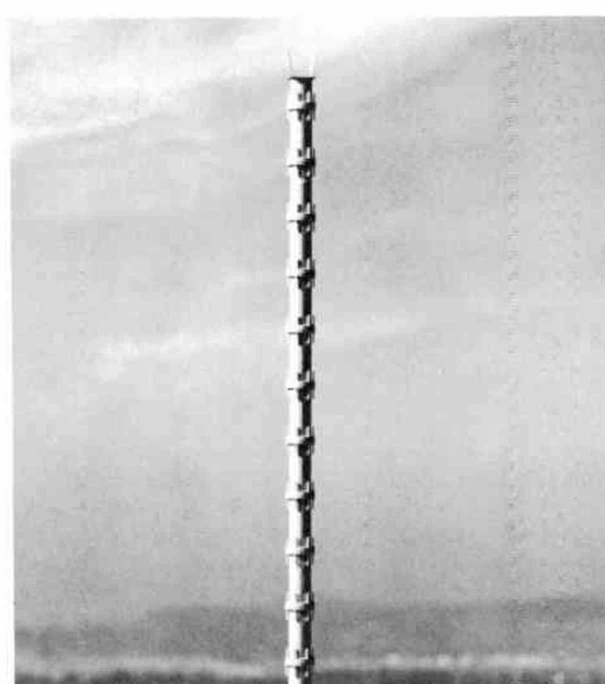


Figure 43. CP slotted antenna. (Courtesy Andrew)

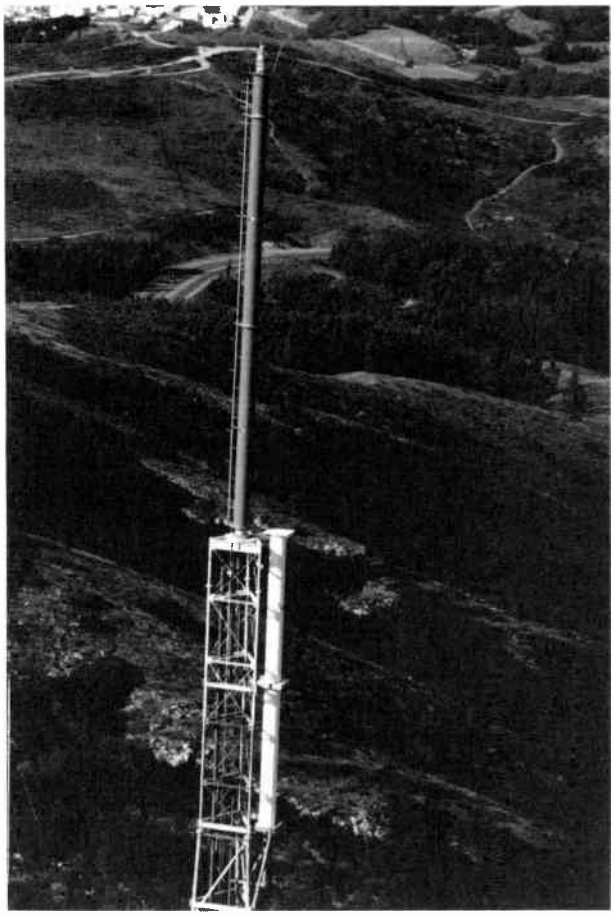


Figure 44. UHF Pylon antenna. (Courtesy Andrew)

circumference is determined by the horizontal pattern; such as one slot for a skull shaped pattern, two for a peanut shaped pattern, three for a "trilobe" pattern and four or more slots, depending on outer cylinder diameter, for an omnidirectional pattern.

The layers are located at one wavelength spacings along the antenna, with the number of layers determined by the vertical gain and pattern. The radiation parameters of phase and amplitude are determined by a combination of slot length and coupler bar diameter. This allows discreet control of the illumination along the antenna aperture at every wavelength resulting in vertical pattern control and shaping. It also allows for maximum aperture efficiency and, in conjunction with the extremely low cross-polarized radiation component of a slot, produces the highest vertical gain for a given antenna length. The antenna is a bottom fed traveling wave resonant antenna. Some antennas launch the energy into the coax radiating section at the center. Others feed the coax radiating section at the center. The bottom coax feed is located inside the radiating coax feed. The pylon uses a radome to cover the radiating slots only.

One version is an omnidirectional bottom fed traveling wave waveguide slot antenna. Since it is waveguide, no inner conductor is required. The signal propagates in the TM mode. The resultant current rings on the inside wall are interrupted by slots cut in the wall. The slots have radome covers.

Another version is designed to radiate a cardioid pattern and has a single row of slots. This is also a bottom fed traveling wave waveguide array operating in the TMOI mode. This is built by exciting the fields in a rectangular waveguide and rolling the waveguide into a cylinder.

Slotted Director

The *slot director antenna* in Fig. 45 is linearly polarized with a single row of slots with directors to fill in the low signal on the opposite side of the slots.

The slots are one wavelength long on one wavelength centers. Each is fed directly at the slot from a high-power strip line feed. Each section is up to 8λ's and four sections are usually stacked for the desired gain.

Each section is branch fed while the slots in each section are fed in a traveling wave manner.

Doublet Array

The doublet array panel antenna in Fig. 46 consists of two pairs of dual dipoles, each pair spaced slightly greater than one wavelength apart.

The dual dipole antenna is two broadband dipoles spaced a half-wavelength apart.

The half-wavelength spacing reduces the coupling

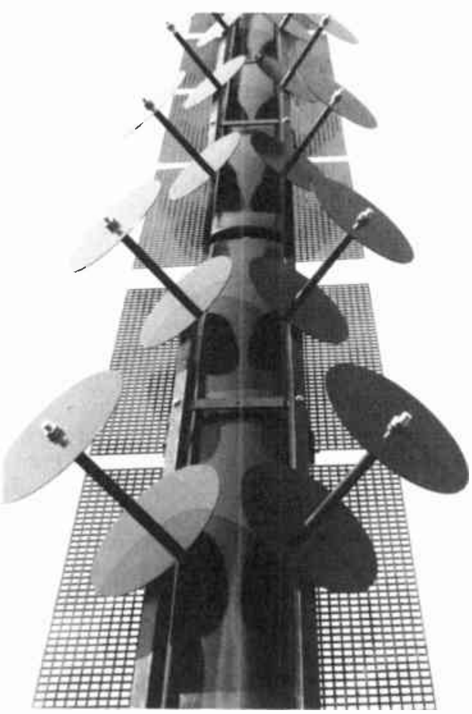


Figure 45. Slotted director antenna. (Courtesy Cablewave)

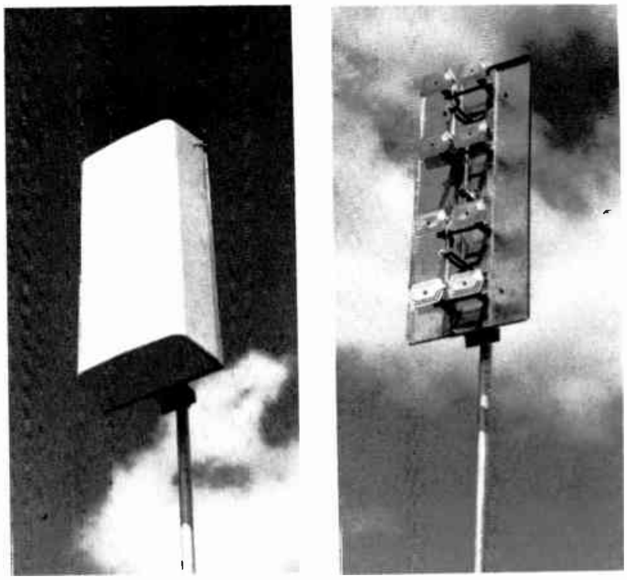


Figure 46. Doublett array antenna. (Courtesy MCI)

and cancels the downward radiation, thereby increasing the gain.

The doublet array has a VSWR less than 1.10:1 from 470 to 800 MHz.

Arrays composed of doublet arrays can be arranged to produce omni or directional patterns, and can handle four or more high power channels.

Circular Polarized Antennas

Slotted Coax

The UHF circular polarized slot antennas are similar to the linear polarized slot antenna.

The slot cut in the wall will radiate a horizontally polarized signal. If a vertical dipole is placed above or near the slot, energy will be coupled to the dipole and reradiated as a vertically polarized component.

The number of slots about the periphery of the cylinder can be varied to obtain omni or directional patterns.

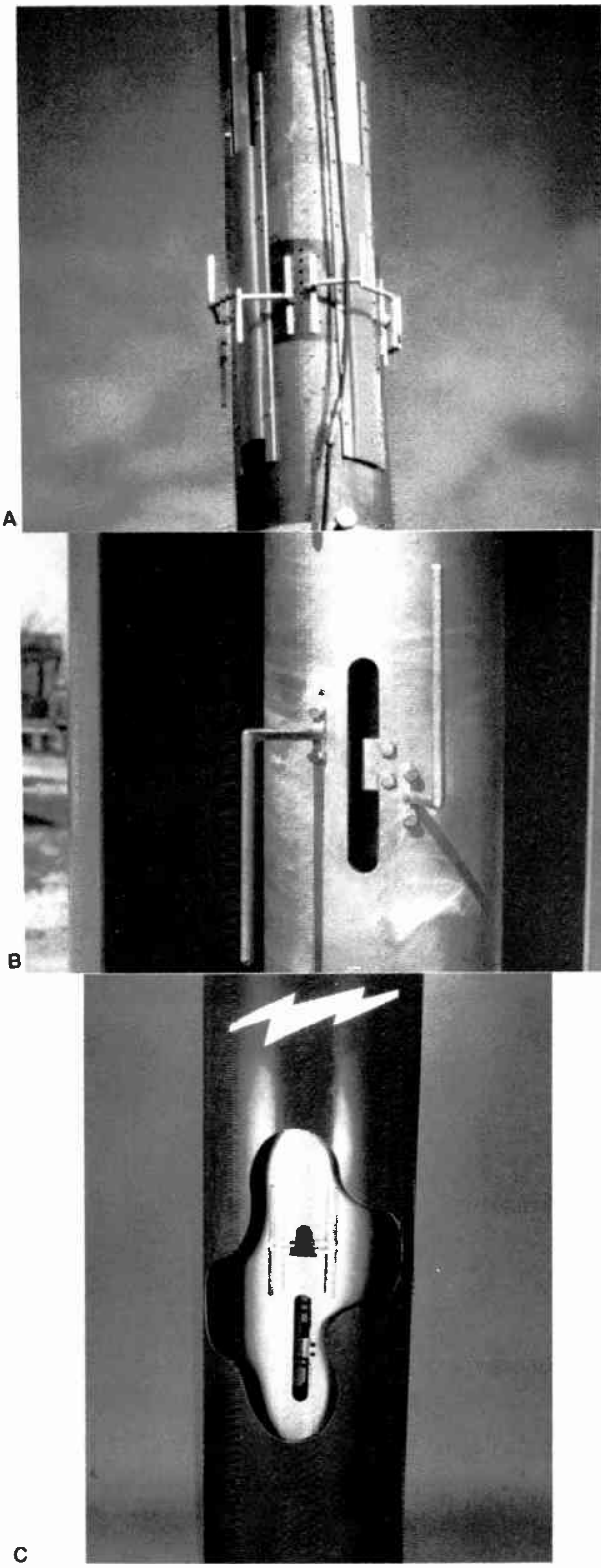
The circular polarized pylon antenna in Fig. 47A uses a Z dipole located directly above the slot, and radiates a vertical component in phase quadrature with the same elevation pattern as the horizontally polarized slot. The size and spacing of the graduated dipole can be used to control the amount of vertical component radiation.

A vertical dipole configuration is shown in Fig. 47B.

Another configuration consists of a series of slotted arrays that uses a vertical dipole, located between the slots, as shown in Fig. 47C.

The pylon circular polarized type of antennas can be either end fed or center fed. The center fed array can be fed internally for a top mounted location on externally fed for side mounted installations.

The antennas in Figs. 47B and 47C are traveling



A

B

C

Figure 47. Circular polarized antennas. (A) Z dipole antenna (Dielectric), (B) Wave Star vertical dipole antenna (Harris), (C) Trasar vertical dipole antenna (Andrew).

wave circular polarized and each have a TM_{01} mode circular waveguide feed system for very high-power applications.

All of the traveling wave slotted arrays have a cylindrical radome covering the full array.

ANTENNA TESTS

Antennas must be tested to determine if the necessary requirements for impedance and patterns are met. Impedance tests are usually run on all production antennas. Pattern tests are normally run on prototype and custom antennas.

Custom antennas are always impedance tested before shipment. These measurements should be made with the antenna completely assembled and in an area free of reflections. Using an impedance plotter or network analyzer, a determination of the antenna reflection characteristic can be made.

A pattern measurement test is conducted for two reasons: (1) to determine the gain as compared with a dipole for which a substitution method could be used; and (2) to determine the amount of radiation at all vertical and horizontal angles which have an influence on the coverage.

Pattern tests can be conducted on full or partial scale models of the final antenna. Scale models have an advantage of reduced size that permits high gain antennas to be tested in an anechoic chamber, free of reflections (from the ground or nearby objects that occur when conducting full scale tests on a range).

REFERENCES

1. Matti Siukola: "Evaluation of Circular Polarized TV Antenna Systems," IEEE, BC-24, March 1978.
2. J. Giardina, T. Vaughan, and J. Neuhaus: "True APL Levels," IEEE Broadcast Technology Society, Broadcast Symposium, 1984.
3. J. Stenberg and W. Pries: "Advances in RF System Management Techniques," NAB Engineering Conference Proceedings, NAB, Washington, DC 1989.
4. T. Vaughan, J. Windle: "Tall Towers for Super Power TV," NAB Engineering Conference Proceedings, NAB, Washington, DC 1988.
5. R. W. Masters: "The Superturnstile Antenna," Broadcast News, January 1946.
6. a. Matti Siukola: "Predicting Performance of Candelabra Antenna by Mathematical Analysis," Broadcast News, October 1957.
   b. R. H. Wright and J.V. Hyde: "The Hill-Tower Antenna System," RCA Engineer, August-September 1955.

2.8 Radio Wave Propagation

Martin H. Barringer
Frederick G. Griffin, PC, Lynchburg, Virginia

Additional material provided by
Kenneth D. Springer
National Association of Broadcasters, Washington, District of Columbia

Radio wave propagation is the study of the transfer of energy at radio frequencies from one point, a transmitter, to another, a receiver. Radio waves are part of the broad electromagnetic spectrum that extends from the very low frequencies which are produced by electric power facilities up to the extremely high frequencies of cosmic rays. Between these two extremes are bands of frequencies that are found in every day uses: audio frequencies used in systems for the reproduction of audible sounds; radio frequencies; infrared and ultraviolet light;, and X-rays.

All electromagnetic waves propagate at the same velocity, regardless of the frequency. Light is an electromagnetic wave, and thus the propagation velocity is often referred to as "the speed of light" (c), which for a vacuum is approximately 3×10^8 m/sec. The velocity of any wave is dependent upon the medium in which it is travelling, but for simplicity is usually considered with respect to a vacuum. The frequency of a wave is defined in terms of the number of cycles per second or Hertz (Hz) and is related to the wavelength (λ) by the expression $f = c/\lambda$. Fig. 1 shows the ranges of various bands within the electromagnetic spectrum in terms of frequency and wavelength.

Radio frequencies are generally confined to that portion of the electromagnetic spectrum below the infrared frequencies. At present, the practical upper limit of radio frequencies is roughly 100 GHz.[1] Within the radio frequency spectrum are bands of frequencies that have been allocated to the broadcast service. The following discussions and methods will apply particularly to these bands of the radio frequency spectrum.

The AM frequency allotments are contained in what is referred to as medium frequencies (MF), 300 kHz to 3 MHz. The FM frequencies and a portion of the TV band are contained in the VHF band which extends from 30 MHz to 300 MHz. The remaining TV allocations are contained in the UHF band of 300 MHz to 3 GHz. Allocations for broadcast auxiliary services such as remote pickup, studio/transmitter links, intercity relays, MDS, and ITFS are interspersed within the MF, VHF, UHF, and SHF (super high frequency) bands. Table 1 illustrates some of the allotments assigned to the broadcast service. The allocations for auxiliary services may change from time to time as the needs of various services for radio frequencies change and as technology for equipment improves. As such, Table 1 is included as an illustration of the wide band of frequencies allocated to the broadcast industry. For a discussion of the origin of the frequency allocations for broadcasting and broadcast auxiliary services the reader is referred to Chapter 1.4 of this *Handbook*, "Frequency Allocations for Broadcasting and the Broadcast Auxiliary Services." In order to keep abreast of new frequency allocations, the most current FCC Rules and Regulations should be consulted.

QUANTIFYING PROPAGATION

The energy that is emitted from a transmitter may take many different paths before it is received. The path that the radio wave will take depends on many factors, some of which include: frequency, antenna type and height, atmospheric conditions, and terrain. Radio waves that propagate along the surface of the earth are commonly referred to as ground waves. All radio waves have some ground wave component, however, because the earth is a lossy medium, it severely attenuates the radio wave. This attenuation increases with frequency, so this mode of propagation is useful only for frequencies below 30 MHz. To achieve significant distances, the atmosphere is preferred over ground waves as a transmission medium. The atmosphere is comprised of several different layers, as depicted in Fig. 2. The troposphere is the layer that extends from the earth's surface up to about 16 km. This layer is

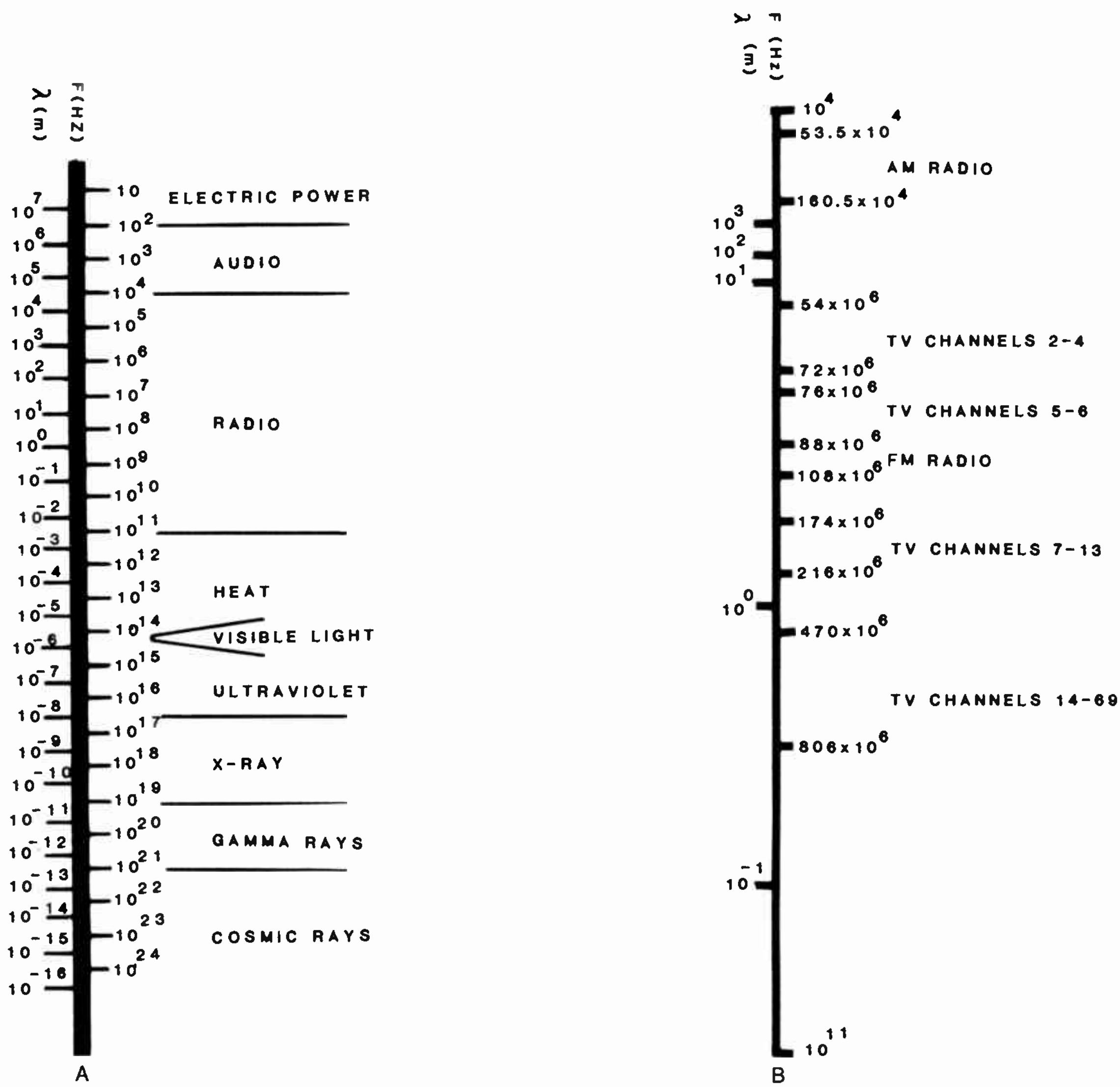


Figure 1. Electromagnetic spectrum and broadcast radio spectrum.

TABLE 1

Broadcast frequency allocations.

FM 300 kHz—3 MHz	
AM:	525 kHz—1605 kHz
VHF 30 MHz—300 MHz	
FM:	88 MHz— 108 MHz
TV:	54 MHz— 72 MHz Channels 2–4
	76 MHz— 88 MHz Channels 5–6
	174 MHz— 216 MHz Channels 7–13
UHF 300 MHz—3 GHz	
TV:	470 MHz— 806 MHz Channels 14–69
AM-FM STL:	947 MHz— 952 MHz
MDS:	2150 MHz—2162 MHz
ITFS:	2500 MHz—2686 MHz
Auxiliary Services:	2000 MHz—3000 MHz
SHF 3 GHz—30 GHz	
Auxiliary Services:	6.425 GHz— 7.125 GHz
CARS:	12.700 GHz—13.250 GHz
TV STL:	17.700 GHz—19.700 GHz

the chief mode of propagation for frequencies above about 30 MHz, and propagation through this layer is dependent upon weather conditions. The next layer is the stratosphere which extends to about 40 km above the earth. This layer has no major effect on the propagation of radio waves. The ionosphere extends upwards of 400 km above the surface of the earth. This region is a charged environment where the air is sufficiently ionized, mainly by the sun's ultraviolet radiation, to reflect or absorb radio waves below about 30 MHz. The ionosphere is constantly changing and is usually considered as consisting of the following sublayers.[2]

1. D layer—This layer exists at heights from about 50 km to 90 km and is present only during daylight hours. The electron density is directly related to the elevation of the sun. This layer absorbs medium and high frequency waves.

2. E layer—This layer exists at a height of about 110 km and is important in the nighttime propagation of medium frequency waves. The ionization of this layer is closely related to the elevation of the sun. At certain times irregular cloud-like areas of high ionization may occur. These areas are known as sporadic E and occasionally prevent frequencies that normally penetrate the E layer from reaching higher layers. The sporadic E layer is prevalent during the summer and winter months. The sporadic E layer formed during the summer is the longest lasting from May to August, and the winter layer lasts about half as long beginning in December. During the mid-summer months when the electron density is at its greatest levels, TV signals in the lower VHF band can be transmitted over distances of hundreds or thousands of kilometers.[4]
3. F1 layer—This layer exists at heights of about 175 to 200 km and is present only during the day. Waves that usually penetrate the E layer (3 to 30 MHz) will penetrate this layer and be reflected by the F2 layer. This layer introduces additional absorption of these waves.
4. F2 layer—This layer exists at the upper boundaries of the atmosphere, 250 km to 400 km, and is present at all times; though the height and electron density will vary from day to night, with the seasons, and over sunspot cycles. During the night the F1 layer merges with the F2 layer at about 300 km. This, in addition to the reduction of the D and E layers, causes nighttime field intensities and noise to be generally higher than during the day.

FREE SPACE PROPAGATION

To evaluate and compare radio wave propagation under various conditions, it is convenient to establish a reference standard. It is customary to consider as a standard the theoretically calculated loss for waves propagated in free space between two idealized antennas. The simplest case to investigate is the radiation emitted from an isotropic source: an ideal antenna which radiates energy with uniform intensity in all directions. An analogy to an isotropic antenna is a point source of light, such as a candle. The intensity of the energy varies proportionally to the inverse of the distance squared from the source, the "inverse square law." The power flux per unit area P_a(W/m^2) at

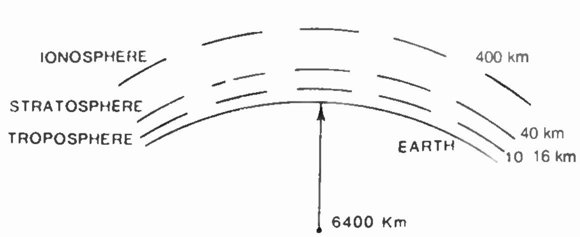


Figure 2. Atmosphere layers.

a distance d(m) from a loss free isotropic antenna radiating a power P_t(W) is given by:

$$P_a = P_t / 4\pi d^2 \quad [1]$$

where $4\pi d^2$ is the surface area of a sphere at a distance d(m) from the source. The power available from a loss free antenna P_r is the product of the power flux per unit area (P_a), and the effective aperture area of the receiving antenna (A_e). This area is related to the gain of the antenna by the expression:

$$A_e = G\lambda^2/4\pi \quad [2]$$

Aperture areas and gains for a specific antenna can be found in the sixth edition of the *NAB Engineering Handbook*, page 121.[3] For a loss free isotropic antenna, $G = 1$, the basic free space transmission loss is defined as:

$$L_{bf} = P_t / P_r = (4\pi d/\lambda)^2 \quad [3]$$

where d and λ have the same units. This equation can be rewritten in its more common form, expressing the loss in dB, as:

$$L_{bf}(dB) = 32.44 + 20\log(F) + 20\log(d) \quad [4]$$

where F is the frequency in megahertz (MHz) and d is the distance between the antennas in kilometers. In the above equation it should be remembered that ideal loss-free isotropic antennas are considered. In real world systems, antenna gain is a significant factor. The transmission loss, L, incorporates the antenna gains and is defined as:

$$L = L_{bf} - (G_t + G_r + L_d) \quad [5]$$

where G_t and G_r are the free space antenna gains with respect to isotropic for the transmitting and receiving antenna respectively. The term L_d is the aperture-to-medium coupling loss or polarization coupling loss between the antennas. The term L_d will have a value of 0 dB when the transmitting and receiving antenna have the same polarization.

In considering the potential service area coverage for a broadcast station, it is usually more desirable to express measurements in terms of field strength rather than transmission loss as previously presented. The root mean square (RMS) field strength, E(V/m), at a point where the power density of a plane wave is P_a(W/m^2) is given by:

$$E = \sqrt{120\pi P_a} \quad [6]$$

where the term 120π is the impedance of free space. The field strength is related to the power available from a loss free isotropic antenna by combining Eqs. [1], [3], and [6] above as:

$$E = \sqrt{480\pi^2 P_r / \lambda^2} \quad [7]$$

A more useful form of the free-space field can be expressed in logarithmic terms above 1 microvolt per meter (dBu) when F is in megahertz and P_r is expressed in decibels above 1 kW (dBK):

$$E\,(\mathrm{dBu}) = 107.2 + P_r + 20\log(F)\,\mathrm{dBu} \quad [8]$$

The electric field produced by a transmitter radiating a power P_t(W) at a distance d(m) in free space can be derived from Eqs. [1], [3] and [6] and is given by:

$$E = \sqrt{30P_t/d^2} \qquad [9]$$

or, in logarithmic terms, where P_t is expressed in decibels above 1 kW (dBK), d is in kilometers, and a transmitting antenna has a gain G_t in decibels above isotropic:

$$E\,(\mathrm{dBu}) = 105 + P_t + G_t - 20\log(d) \qquad [10]$$

Using the same units, the field strength E(dBu) for nonfree-space environments can be related to the basic transmission loss by:

$$L_b\,(\mathrm{dB}) = 137 + 20\log(F) + P_t + G_t - E \qquad [11]$$

These equations form the basis for characterizing propagation. They do not, however, take into account such real world factors as the presence of the earth, atmosphere, or obstructions. To adequately describe an actual radio system, additional losses will need to be added to the free-space equations derived above.

PRESENCE OF EARTH

When the transmitting and receiving antennas are placed over ground, the propagation of radio waves is modified from the free-space models presented above. Radio waves that strike the earth are partially absorbed and partially reflected. Waves that are reflected by the earth experience changes in the phase of the wave, which affects the distribution of available energy. The extent to which the waves are reflected or absorbed is dependent upon frequency and the ground constants: conductivity and permittivity.

Propagation Over Plane Earth

The geometry of the idealized situation of propagation between two antennas placed above a plane earth is shown in Fig. 3. This geometry is valid for antennas that are sufficiently closely located so that the curvature of the earth is not a factor, yet far enough apart from each other so that the energy may be described as a plane wave, and ray theory can be applied. The resultant received electric field can be represented as the sum of the direct and reflected rays:

$$E = E_d\,[1 + |\mathrm{R}|e^{j(\phi_\Delta + \phi_r)}] \qquad [12]$$

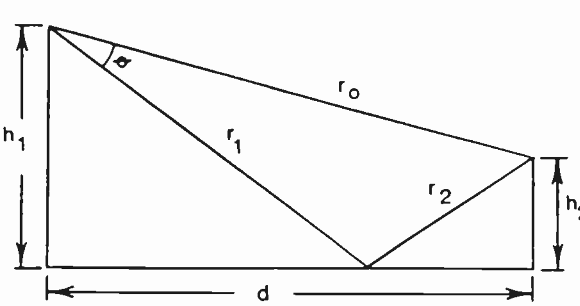


Figure 3. Path rays for plane earth.

This equation is valid for small angles of Θ and deserves some additional explanation. The term E_d is the free-space electric field that is produced at a distance d(m) by the direct ray. The terms $|\mathrm{R}|$ and ϕ_r are the magnitude and phase of the complex reflection coefficient. This term is dependent upon the nature of the surface (i.e., conductivity (δ) and permittivity (ϵ_r)), the angle between the surface and incident wave, the wavelength of the radio wave, the polarization of the wave, and the curvature of the earth. The magnitude of the reflection coefficient varies between −1 and +1. Several sources have derived equations for the reflection coefficient and plotted the effects of changing variables, and the reader is referred to these for further study.[4,5] The term ϕ_Δ is the phase delay due to the longer path that must be taken by the reflected wave, and has the form of:

$$\phi_\Delta = \frac{4\pi h_1 h_2}{\lambda d} \qquad [13]$$

It is often sufficient to assume the ground approximates a large flat surface. In such a case, a sufficiently accurate expression is given by:[4]

$$E = 2E_d \sin\left(\frac{2\pi h_1 h_2}{\lambda d}\right) \qquad [14]$$

Some cases of special merit that can be derived from Eq. [14] are:

Case I	$h_1 h_2 = d\lambda/2$	$E = 0$
Case II	$h_1 h_2 = d\lambda/4$	$E = 2E_d$
Case III	$h_1 h_2 = d\lambda/12$	$E = E_d$

Therefore, depending on the antenna heights, distances, and wavelength, it is possible to totally cancel out the field at the receiver or magnify the wave to a field strength double that which could be achieved from a free-space field. The variation of signal strength due to multipath effects can be minimized in point-to-point applications through the use of antennas with narrow beamwidths.

When considering the case of VHF antennas that are close to the ground, the effective antenna heights h_t(m) and h_r(m) will need to be substituted for h_1 and h_2 respectively for Eq. [14]. The new antenna heights h_t and h_r allow for the effects caused by the relative permittivity (ϵ_r), and conductivity (δ) of the ground. The effective antenna heights are related to the physical antenna heights above ground level by:[4]

$$h_t = \sqrt{h_1^2 + h_0^2} \qquad [15.1]$$

$$h_r = \sqrt{h_2^2 + h_0^2} \qquad [15.2]$$

the term h_0 is dependent upon the type of polarization being considered.

Vertical Polarization

$$h_0 = (\lambda/2\pi)\,[(\epsilon_r + 1)^2 + (60\lambda\sigma^2)]^{1/4} \qquad [16.1]$$

Horizontal Polarization

$$h_0 = (\lambda/2\pi)\,[(\epsilon_r - 1)^2 + (60\lambda\sigma^2)]^{-1} \qquad [16.2]$$

TABLE 2
Ground conductivity and dielectric constants.

Terrain	Conductivity (S/m)	Dielectric Constant (esu)
Sea water	5	80
Fresh water	8×10^{-3}	80
Dry sandy, flat coastal land	8×10^{-3}	10
Marshy, forested flat land	8×10^{-3}	12
Rich agricultural land, low hills	1×10^{-2}	15
Pasture land, medium hills and forest	5×10^{-3}	13
Rocky land, steep hills	2×10^{-3}	10
Mountainous	1×10^{-3}	5
Residential area	2×10^{-3}	5
Industrial area	1×10^{-4}	3

Table 2 lists values for conductivity and permittivity for various soil conditions. As a way of example, assume that an antenna is placed 3 meters (9.8 feet) above dry, sandy, flat coastal land ($\delta = 2 \times 10^{-3}$ S/m, $\epsilon_r = 10$) and operates at a frequency of 100 MHz ($\lambda = 3$ m). Then h_0 in [16.1] and [16.2] will be 1.59 m and 0.16 m respectively. The effective height of the antenna will then be increased to 3.4 m (11.1 feet) for vertical polarization and will remain unchanged at 3 m (9.8 feet) for horizontal polarization. As the frequency increases above VHF, the wavelength becomes increasingly small and the distinction between true antenna height and effective height is immaterial.

MEDIUM FREQUENCY PROPAGATION

As stated earlier, medium frequency waves lie in the frequency range of 300 kHz to 3 MHz and are characterized by their long wavelengths, 1,000 meters to 100 meters. The standard AM broadcasting band frequencies are within this range. The transmitting antenna is located right at the surface of the earth, and the receiving antenna is very close to the earth's surface with respect to a wavelength. In this case the direct and ground reflected waves cancel, and the transmission is by means of the ground wave (also known as the surface wave) and the sky wave.

Ground Waves

These waves are characterized by the fact that they are guided along the earth's surface, similar to a transmission line. The field is attenuated in this propagation mode by losses in the ground. Therefore the composition of the soil, ϵ_r and δ, have a direct bearing on the amount of attenuation the wave will experience, and subsequently how far reliable communications can be established. The attenuation is also dependent upon the frequency and polarization type. The attenuation factor, A, is a measure of the amount of attenuation present and can be determined for a ground wave using the chart of Fig. 4. The term ρ is known as the

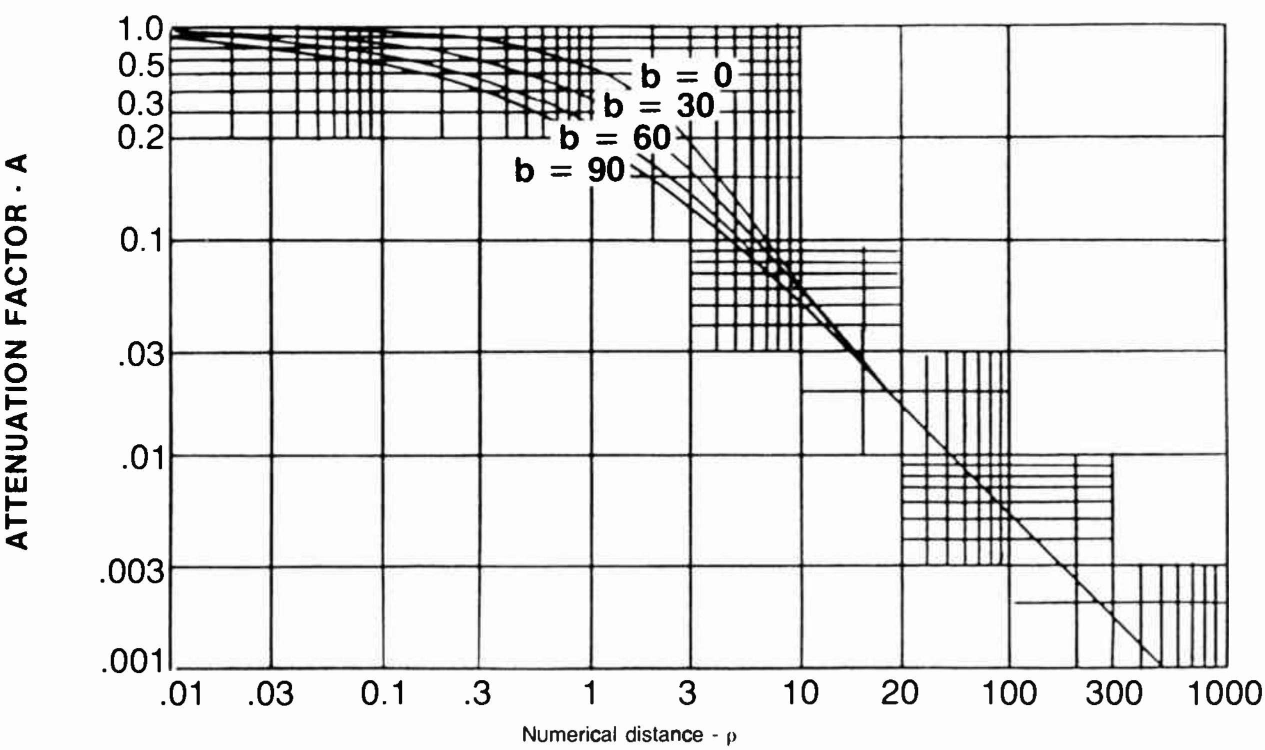


Figure 4. Attenuation factor of ground waves.

numerical distance and b is the phase constant. Values for these terms can be calculated from the following equations:[5]

$$\rho = (\pi d/\lambda x)\cos(b) \qquad [17.1]$$

$$b = \arctan[(\epsilon_r + 1)/x] \qquad [17.2]$$

$$x = 18*10^3\sigma/F \qquad [17.3]$$

To determine the electric field strength, the attenuation factor must be added to Eq. [12].

$$E = E_d[1 + Re^{j(\phi_\Delta + \phi_r)} + (1 - R)Ae^{j(\phi_\Delta + \phi_r)}] \qquad [18]$$

It is interesting to note that the same earth which acts as a conductor at very low frequencies will act as a small-loss dielectric at very high frequencies. It is also noteworthy to observe that the losses for horizontally polarized waves are much greater than for vertically polarized waves. Thus, for practical applications only, vertically polarized waves should be considered. For more detailed and accurate representations of the effects of ground wave, the works of Norton and as expanded upon by Jordan should be consulted.[6,5]

Sky Waves

While the ground wave provides the major path for medium frequency propagation, the wave attenuates relatively quickly with distance and is reliable for distances of only a few hundred kilometers. To achieve greater distances, the waves propagate via the ionosphere and are known as sky waves, and can provide sufficient signal strength at distances up to a few thousand kilometers.

The ionosphere is a constantly changing environment that begins approximately 65 km (40 miles) above the earth and extends to about 400 km (250 miles). This region of the atmosphere is composed of three major sublayers D, E, and F as discussed earlier. These layers are not present at all times. For example, the D layer is present only during the day and is a major absorber of medium-frequency waves. The E layer is a principal reflector of medium-frequency waves. Thus during the day the majority of the medium-frequency waves are absorbed by the D layer, but at night the D layer is not present, allowing the medium-frequency waves to be reflected by the E layer.

Interference Between Ground Waves and Sky Waves

Interference to a receiver may occur from co-channel stations located many kilometers from the desired station. Because of the sky wave, sufficient signal strength may be received to interfere with the local station. This effect has been minimized by the FCC by limiting two factors in the operation of some AM stations: the operating power and time of operation.

Multipath interference occurs when the waves from a transmitting antenna reach a receiver from different paths in such a manner as to cancel or severely interfere each other. This can happen at distances where both the ground wave and sky wave are sufficiently strong to interact. The geometry of this is similar to that shown in Fig. 3, except the direct ray will be a result of the ground wave and the reflected wave will be from the ionosphere. At distances relatively close to the transmitter, the ionosphere will not reflect waves back to the earth, so the ground wave is predominant. At distances beyond a few hundred kilometers, the sky wave will dominate and the ground wave will be too weak to interfere. Multipath interference can also occur where the sky wave follows more than one path to the receiver.

Effects of Solar Activity

Interference to medium-frequency waves can also be caused by solar activity such as sunspots and flares which manifest an increased or reduced emission of radiation from the sun. The changes in solar radiation levels can cause changes in the ionospheric layers that may result in unusual sky wave propagation conditions called "skip" which, in turn, can cause inter-station interference. The effects of such solar activity will have their strongest effect on propagation in the AM band during the first five to ten days after the start of a storm. This has the effect of reducing sky-wave field strengths. The effect has been observed to increase with frequency.[13]

PROPAGATION ABOVE 30 MHz

Smooth Earth Conditions

At frequencies above about 30 MHz the principal propagation mode is tropospheric. The surface wave is attenuated too severely to be of any practical long distance use and, though attenuated, the sky wave is usually passed through the ionosphere to space.

For waves that propagate close to the earth's surface the curvature of the earth will introduce additional effects that must be included in the plane earth model that was considered earlier. Fig. 5 shows the geometry of a smooth earth model. First the reflection coefficient R of the reflected wave has different characteristics than for a plane surface. Since the wave is reflected

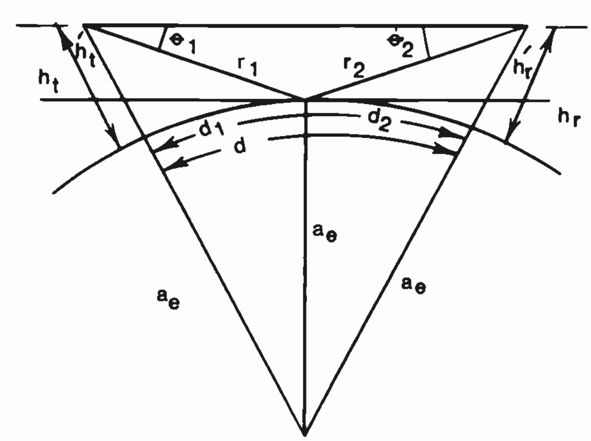


Figure 5. Reflection from smooth earth on line-of-sight path.

against a curved earth the energy diverges more than is predicted by the inverse square law and the reflection coefficient R, in Eq. [12], must be multiplied by the divergence factor D, given by:[4]

$$D = \sqrt{\frac{1 + 2d_1d_2}{2a_e(h'_t + h'_r)}} \qquad [19]$$

It should be noted that for smooth earth conditions, the heights h'_t and h'_r for the transmitting and receiving antennas above the plane tangent to the earth at the point of reflection are less than the antenna heights h_t and h_r above the surface of the earth.

Under normal propagation conditions, the refractive index of the atmosphere decreases with height so that radio waves near the surface of the earth travel more slowly than at higher altitudes. This variation in velocity as a function of height results in a bending of the radio waves. This may be represented as a modified earth radius commonly known as the effective earth radius, a_e, which allows the radio waves to be represented as straight lines. The ratio of the effective earth radius to true earth radius is commonly known as the k factor. Values of k can vary between from 0.6 to 5.0 depending on the climate being considered. For temperate climates the average value of k is 1.33, and most works refer to this as the 4/3 earth model when used in calculation.[2]

Beyond Line of Sight Conditions

In order to determine when conditions exist where propagation is considered to be beyond line of sight, the respective distances from the transmitter and receiver to the radio horizon must be calculated. The radio horizon is the distance the horizon appears from an antenna, as defined by a plane from the antenna to the tangent of the earth's surface and is depicted in Fig. 6. The equation for the radio horizon in terms of d_{lt}(km) and h_t(m) and the k factor is of the form:

$$d_{lt} = 3.57\sqrt{h_t k} \qquad [20]$$

When the sum of the distances to the radio horizon for the transmitter and receiver is less than the total distance of the path under consideration, then a beyond line of sight condition exists. Diffraction makes it possible for radio waves to travel beyond that possible for line of sight transmission, though an additional loss term must be added to the free-space loss. The amount of attenuation can be determined by diffraction methods. The geometry of beyond line of sight propagation is shown in Fig. 7.

The exact calculation of the field strength at any

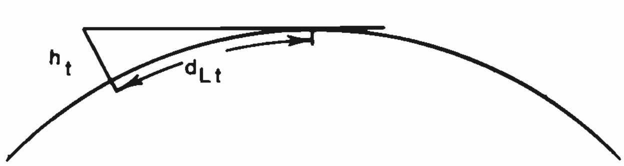


Figure 6. Distance to radio horizon.

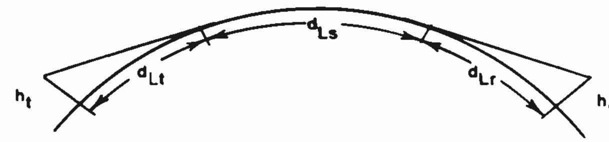


Figure 7. Beyond line of sight.

point beyond the line of sight for a smooth earth is rather complex and the presentation of such a method is beyond the scope of this text. However, nomograms have been developed that apply to a large number of cases. For the reader interested in the actual prediction of the losses to be expected for smooth earth diffraction, the National Bureau of Standards publication may be consulted.[9]

Fig. 8 is a nomogram that can be used to determine the loss that must be added to the free-space loss. In order to use the nomogram, the distance d_{lt} must be less than d_{lr}. The total loss (L) is the sum of the three losses L1, L2, and L3. By way of example, assume a system that has the following parameters: h_{lt} = 14 m (45.9 feet); h_{lr} = 178 m (583.8 feet); F = 100 MHz; k = 4/3; total path length of 85 km; and the wave is vertically polarized over land. The distances are calculated to be: d_{lt} = 15 km (9.3 miles), d_{lr} = 55 km (34.2 miles), and d_{ls} = 15 km (9.3 miles). The total loss relative to free space is thus L = L1 + L2 + L3 = 22.9 + 3.8 + 4.3 = 31.0 dB.[7]

Effects of Obstacles on Propagation

In the previous sections, a perfectly smooth sphere was assumed for earth. Only the effects of the atmosphere were accounted for in the k factor. The assumption of a perfectly smooth earth allowed for a relatively simple calculation of the expected field strengths and transmission losses at various points within the line of sight and regions beyond the line of sight. However, the real world is much less than ideal, and the presence of hills, buildings, foliage, as well as the atmosphere all have a bearing on the computation of field strengths. These obstacles have a complex effect on the propagation of radio waves which makes it virtually impossible to predict the field strength or transmission losses at discrete points close to these obstacles. However, the path being considered may be quantized by use of earth profiles, and through the use of some simplifying assumptions, predictions of the field strength which are more accurate than smooth earth approximations can be performed.

Hills

Perhaps the most common obstructions that will appear in the path of a radio wave are hills. The amount of attenuation the hill will introduce into a path is a function of the distance from the antenna terminals to the hill, and the height of the hill above or below the line-of-sight ray between the transmitting and receiving antennas. The hill's height and distance from the antenna can be determined by constructing a path profile and plotting the terrain features on special graph

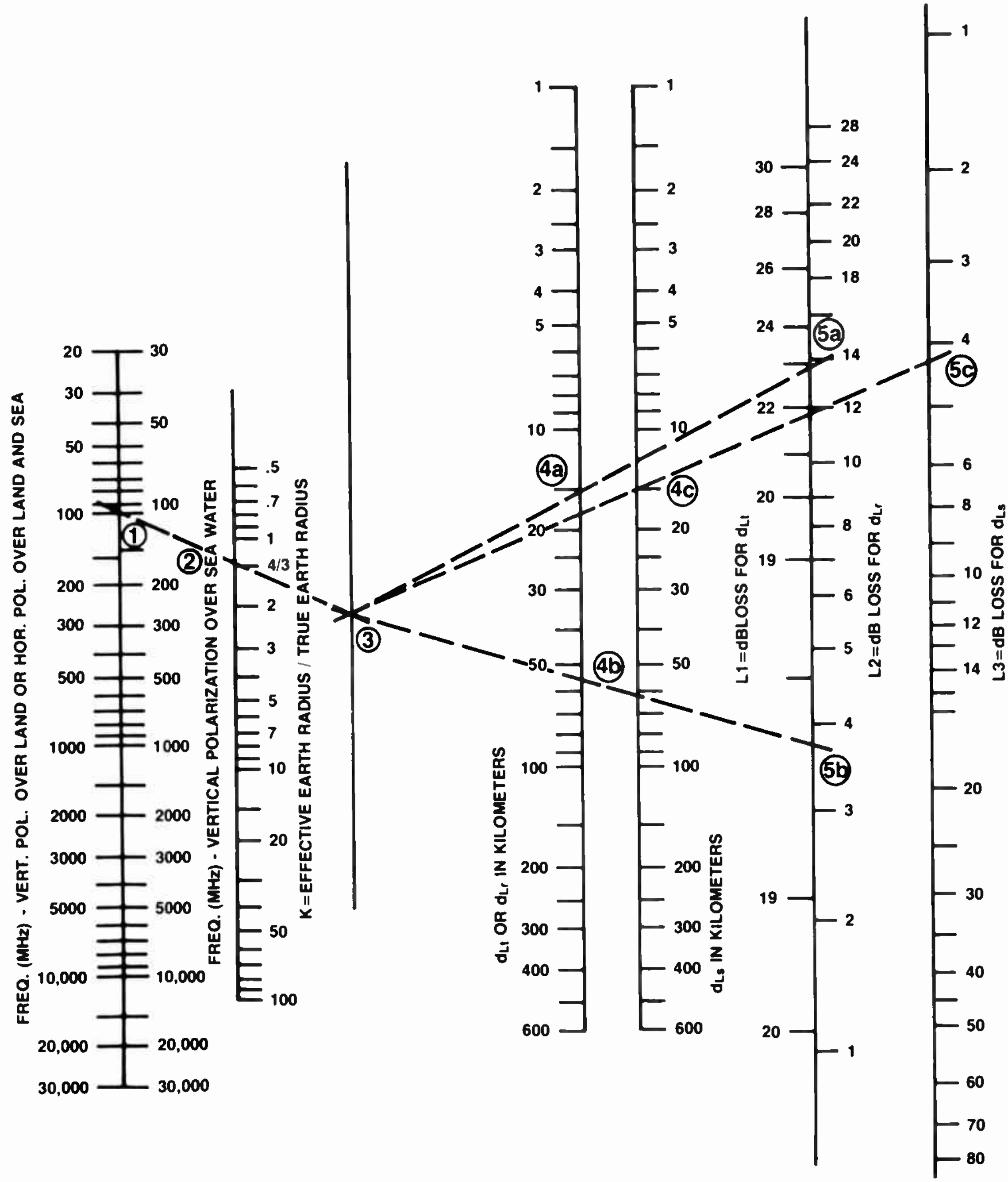


Figure 8. Diffraction loss over smooth earth.

paper that includes the effect of refraction. The most common charts are defined for a factor of k of 4/3. A typical path is shown in Fig. 9. Terrain elevations necessary to construct a path profile can be obtained from topographical maps of the area, or from computer databases. To obtain the best accuracy in using topographical maps the smallest scale available should be used, 1:24,000. These maps can be obtained from the United States Geological Survey in Denver, Colorado or from any of its branch dealerships. Computer models exist that can map the terrain along a propagation path and are available from many sources.

In order to determine when a hill is sufficiently removed from a path to allow free-space conditions to exist, the Fresnel zone clearance equation can be used. This equation was initially developed to explain the diffraction of light around knife edged obstacles, and has since been applied to radio theory. This equation

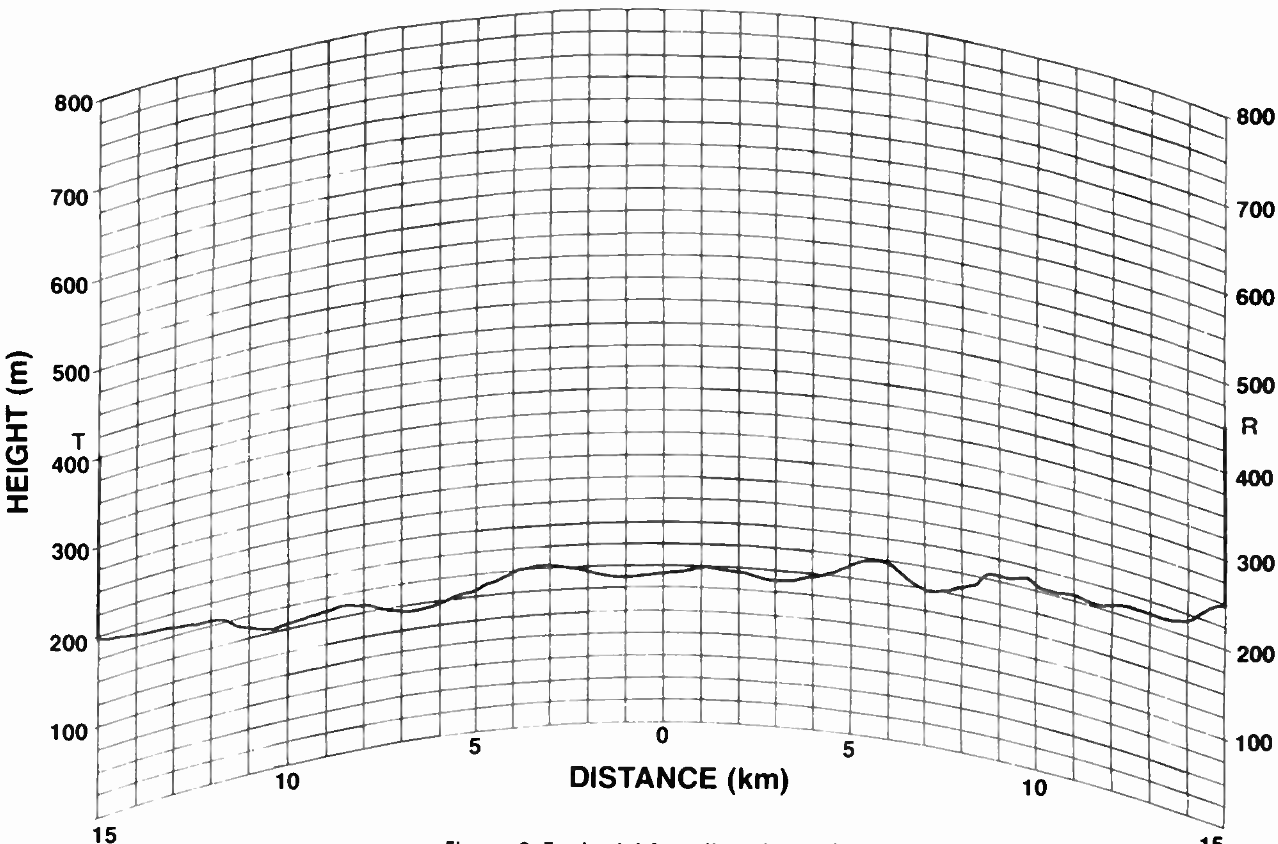


Figure 9. Typical 4/3 earth path profile.

describes a radio path as an ellipsoid with the transmitting and receiving antenna located at the focal points of the ellipse. As Fig. 10 depicts, the curves for various reflection coefficients intersect at 0 dB from free space when the clearance is equal to six-tenths of the distance to the first Fresnel zone clearance. Thus free space conditions exist when obstacles are outside the $0.6F_1$ zone radius. This distance can be calculated by:

$$h = 0.6F_1 = 328.6\sqrt{d_1 d_2/Fd} \qquad [21]$$

where the height of the $0.6F_1$ zone is in meters; d_1 is the distance from one antenna to the obstacle in kilometers; d_2 is the distance from the second antenna to the obstacle in kilometers; d is the total path distance in kilometers; and F is the frequency in megahertz. When determining whether a path clears a hill, an additional height, typically 15 meters, should be added to the height of the hill to account for any trees that may be present.

If the hill lies within the calculated Fresnel zone radius then nonfree-space conditions exist and additional losses will be present. When the frequency is high enough for the hill to be considered as a sharp ridge, and the transmitter and receiver are far enough away from the hill, then the loss may be calculated using diffraction from a knife edge, as shown in Fig. 11. The height of the hill H(m) is measured from the line joining the centers of the two antenna to the top of the ridge. The amount of attenuation or shadow loss with respect to free space may be read from the graph shown in Fig. 12. The value of ν, the diffraction parameter, can be calculated, with respect to the distances measured in kilometers and the frequency F(MHz) from:[4]

$$\nu = 0.00258H\sqrt{dF/d_1 d_2} \qquad [22]$$

When considering paths that are obstructed by hills that appear rounded rather than knife edged, the attenuation can be calculated using diffraction around a cylindrical surface, as depicted in Fig. 13. This condition predominates when the elevation of the hill changes drastically within a wavelength. This can occur when considering paths at the lower end of the VHF spectrum that pass over older mountain ranges, i.e., Appalachian, Blue Ridge, Catskill.

The amount of attenuation can be found from the chart of Fig. 12. The term ρ from the chart is a dimensionless quantity known as the index of curvature of the cylinder's radius R and is calculated from:[4]

$$\rho = 0.83R^{1/3}\lambda^{1/4}\sqrt{d/d_1 d_2} \qquad [23]$$

where all distances are in the same units. For those interested in incorporating the calculation of losses due to diffraction over knife edge and rounded obsta-

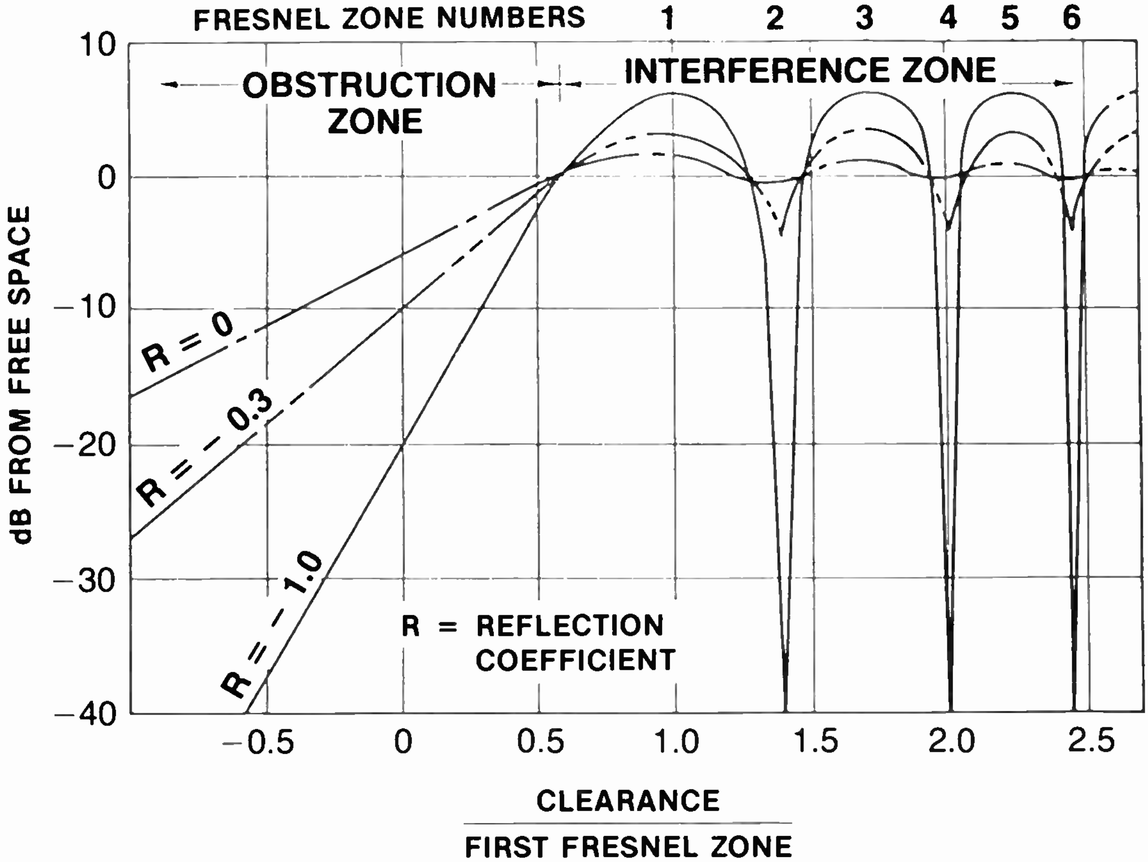


Figure 10. Effect of path clearance on radio propagation.

cles into digital computers more exact equations can be found in Rice, et al.[9]

While the method for calculating the loss due to a single obstacle is relatively straightforward, there are times when successive obstacles are present in a radio path as shown in Fig. 14. In order to determine the loss associated with multiple diffraction regions, an approximation method has been developed based on an extension of single-edge diffraction. The obstacle which would singly produce the greatest diffraction loss is determined using the methods discussed above. Lines are drawn joining the summit of this obstacle to the transmitter and receiver antenna locations. The additional attenuation caused by the remaining obstacles should be calculated using their heights, H_a and H_b, above these lines. These additional losses are then

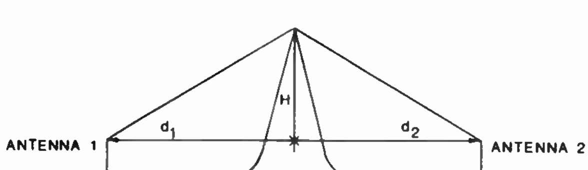


Figure 11. Ray path for knife edge diffraction.

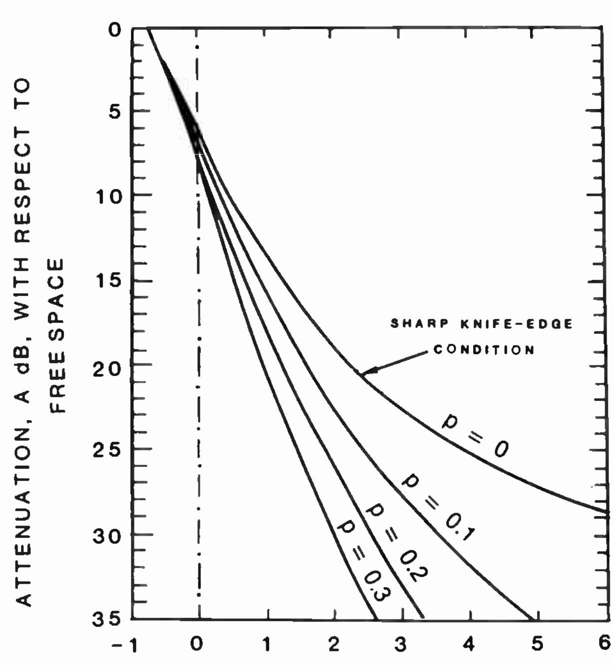


Figure 12. Attenuation due to various diffraction conditions.

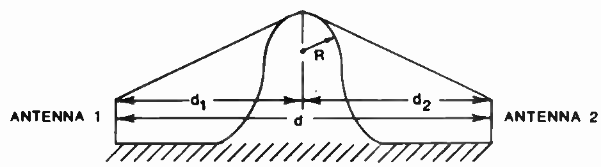


Figure 13. Diffraction due to cylinder.

added to the loss due to the main obstacle in calculating the nonfree-space loss. It is important to note that even if H_a and H_b are slightly negative, i.e., below the lines, they may still produce a small amount of attenuation due to the Fresnel zone clearance requirements.[4]

Buildings

When planning for radio locations within built-up areas of cities or residential areas, buildings will have an effect on radio propagation. For radio relay stations such as studio-to-transmitter links it is the normal practice to select sites that will be clear of buildings. However, where this is not feasible and the path geometry is known, i.e., height and location of buildings, then the diffraction methods discussed for hills may be applied. In planning for broadcast systems, it is not practical to relate attenuation measurements made in built-up areas to the particular geometry of buildings. Therefore it is more conventional to treat the losses in a statistical manner, dividing the general classifications of building types into loss groups, so that a loss can be derived for a particular type of building, i.e., multi-story made of concrete and steel or single-story residential made of wood.

Within built-up areas there is much more back scatter than in open country. Additionally, due to the fact that buildings are more transparent to radio waves than the earth, there tends to be less shadow loss caused by buildings. However, the angles of diffraction due to buildings are usually much greater than in open country for natural terrain and thus the loss resulting from the presence of buildings tends to increase. Measurements indicate that at 100 MHz the median field strengths are 4 dB to 6 dB below that expected for a plane earth and drop off to about 10 dB for 200 MHz.[3] These measurements were made in areas containing some large buildings and open areas, but mainly consisting of residential areas. Some recent measurements conducted in the 850 MHz band indicate field strengths 20 dB to 34 dB below that expected for free space for path distances of 1 km to 25 km.[10]

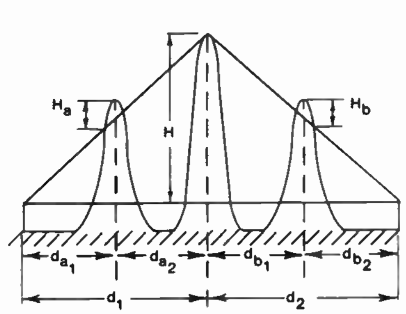


Figure 14. Multiple knife edge diffraction.

Vegetation

Among the many factors that have an effect on the determination of the losses present in a propagation path, vegetation is sometimes the most overlooked. Depending on the type of terrain in consideration, i.e., open or forest, the effect of vegetation can add several dB loss to the system. The amount of attenuation present is dependent upon the frequency and polarization of the wave, see Fig. 15. As can be seen, the attenuation for a horizontally polarized wave for frequencies below about 1000 MHz is much less than for a vertically polarized wave. At around 1000 MHz, trees that are thick enough to block the field of vision can be modeled as an almost solid obstruction and the attenuation over or around these obstructions can be predicted from knife edge diffraction methods.[3]

The effect of vegetation on a radio path varies seasonally in the case of deciduous trees. During the winter months the losses due to shadowing and absorption are less than those during the spring and summer. It is interesting to note that the greatest losses will occur during the spring since new growth has more sap and moisture content which adds to the absorption losses. When the antenna is raised above trees and other vegetation, the prediction of field strengths depends upon the estimation of the height of the antenna above areas of reflection and the reflection coefficients. For areas of fairly uniform growth and for angles of incidence approaching grazing, the reflection coefficient will approach -1 at about 30 MHz. Even low growth that is uniform, i.e., a wheat field, may yield a value of -0.3 for the reflection coefficient.

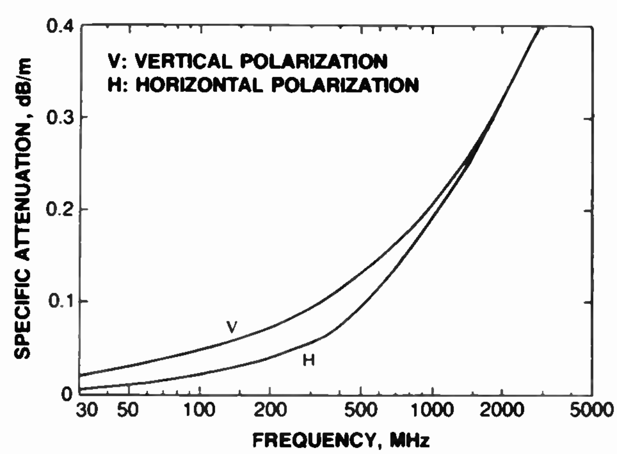


Figure 15. Attenuation through vegetation.

Atmosphere

As was discussed earlier the troposphere is the major medium for propagation at VHF frequencies. The refractive index (n) of air has a value near unity (typically 1.00035). The index is dependent upon the dielectric constant, and can vary depending on the pressure and temperature of the air and on the amount of water vapor present. Therefore the refractive index changes with weather conditions and with the height above the earth. The velocity of radio waves is dependent on the refractive index of the atmosphere. As a general rule the velocity of a wave is slower at the earth's surface than at higher altitudes. So a horizontally polarized wave will be refracted back towards the earth, though unusual atmospheric conditions may change this. Because there are numerous constantly changing variables involved in the prediction of field strengths due to atmospheric conditions, some simplifying assumptions are generally needed to obtain a solution under known meteorological conditions.

Ducting

Changes in the index of refraction of only a few parts per million can have dramatic effects on radio waves. Therefore it is usually more convenient to refer to the refractive index in terms of the refractivity, N:

$$N = (n - 1) \times 10^6 \qquad [24]$$

Under meteorological conditions where the refractive index decreases rapidly with height over a large horizontal distance, radio waves can become trapped and experience low propagation loss over long distances. This phenomenon is known as ducting. Although ducting is frequent with some locations and meteorological conditions, due to its randomness and long range unpredictability it is not a reliable mode for communications. However, due to the strong fields over the horizon caused by ducting, inter-station interference can result. In addition, line of sight paths may be affected by severe fading.

In order for atmospheric ducts to occur, two conditions must exist. First the refractive index gradient must be equal to or more negative than −157 N/km. The refractive index gradient is a measure of the change of the refractivity across a vertical height h, dN/dh. When this condition is present, the radio waves will remain close to the earth's surface beyond the normal horizon. Secondly, the refractive index gradient must be maintained over a height of many wavelengths. The duct may be thought of in the same manner as a transmission line waveguide. However, unlike metallic waveguides, natural ducts do not have sharp boundaries, although there is a wavelength cut-off above which waves will not propagate. Since the duct does not have sharp boundaries, the thickness (t) will not be rigid. Therefore the cut-off wavelength (λ) will not be fixed, but an estimate can be obtained from:[4]

$$\lambda = 2.5 \times 10^{-3}\, t^{2/3} \sqrt{\frac{\delta N}{t} - 0.157} \qquad [25]$$

where the wavelength and thickness are in meters. The term δN represents change in refractive index across the duct. As an example, a duct near the ground that is 25 meters thick and has a refractive index change of 10 N, i.e., −400 N/km, will have a cut-off wavelength of 0.15 m (2 GHz). However, a duct with the same refractive index gradient will have to be about 87 meters thick to propagate a wavelength of 1 meter (300 MHz).

A duct spreads the energy within it in the horizontal direction, but is constrained in the vertical direction as the distance from the transmitter is increased. Thus, in principle it is possible for the field strength within a duct to be greater than the free-space field at the same distance. However, a duct will 'leak', or allow energy to escape at the boundary, adding to the transmission losses so that field strengths are seldom greater than free-space values.[4]

There are typically two types of ducts: ground based and elevated. A ground-based duct, as its name implies, forms close to the earth's surface. Energy is propagated in this duct by being refracted back to the earth, reflected off the earth, then refracted again, see Fig. 16A. An elevated duct forms above the earth's surface and is generally very short lived. Energy in an elevated duct is refracted back and forth between boundaries without coming in contact with the earth, similar to the way coherent light propagates in a graded index optical fiber. (See Fig. 16B.) Shadow regions are formed along the area outside of a duct where, due to the nature of the duct, radio waves are not present. Receiving antennas placed in such a region will experience a loss of signal. As can be seen from Figs. 16A and 16B these regions can form not only above the earth's surface from a ground-based duct, but can also form along the earth's surface in the case of an elevated duct. Therefore a shadow region, that can result in

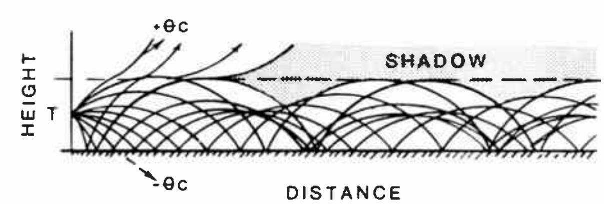


Figure 16A. Ray propagation in a ground based duct.

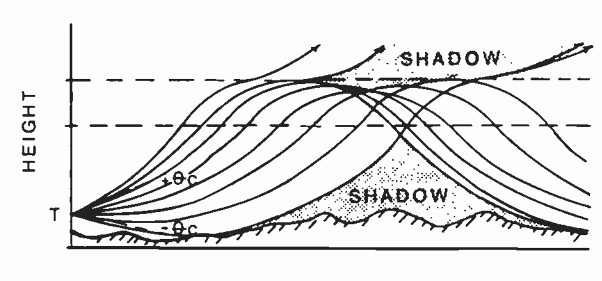


Figure 16B. Ray propagation in an elevated duct.

loss of communications, can form at a receiver that is located relatively close to the transmitter.

Radio waves that leave the transmitting antenna at an angle greater than a certain angle, the critical angle, will not become trapped in a duct. These radio waves will propagate through the boundary of the duct, though they will experience some bending due to the change in the index of refraction at the duct's boundary.

Atmospheric Absorption

Radio systems using frequencies above 1 GHz experience another loss that must be accounted for when planning the system: atmospheric absorption. Relay links and STLs that are inadequately engineered may experience outages during periods of heavy rainfall due to this loss. The amount of attenuation due to rain is dependent upon three factors: (1) the rate the rain is falling, (2) the frequency of the wave, and (3) the length of the rain cell a wave must propagate through. If the path length is only several kilometers long, it is usually adequate to approximate the length of the rain cell by the total path length. The average rainfall rate varies from one section of the country to another, however typical rainfall rates are given in Table 3. The specific attenuation $\lambda_r \left(\frac{dB}{Km}\right)$ is given by[9]:

$$\lambda_r = KR_r^{\alpha} \quad [26]$$

where R_r is the rainfall rate in millimeters per hour, and the terms K and α are found[9] as:

$$K = [3(F-2)^2 - 2(F-2)] \times 10^{-4} \quad [27.1]$$

$$\alpha = [1.14 - 0.07(F-2)^{1/3}] \times [1 + 0.085(F-3.5)e^{(-0.0066F^2)}] \quad [27.2]$$

where F is the frequency in GHz. These equations give a good approximation to attenuation curves published by the CCIR for frequencies below 50 GHz.[14]

In addition to rain, attenuation can also be caused by water vapor and oxygen that is present in the air. The attenuation due to water vapor and oxygen is less than that for rain and usually can be neglected. However, as radio systems employ higher microwave frequencies, attenuation due to these losses will become significant. A description of how these losses can be calculated has been presented in previous works.[9]

COVERAGE AREAS

Engineering a radio or television broadcast station using the methods presented previously are too cumbersome to be of any practical use in determining the service area of the station. While radio waves actually behave in the manner described in the previous sections, it would be too involved to use these methods at every point surrounding a station. Therefore other quantitative methods are needed to determine field strengths quickly and reliably. Considerable work has been conducted in this area and is still being carried out.

TABLE 3
Rainfall amounts.

Characteristics	Rate
Drizzle	0.25 mm/hr
Light Rain	1.00 mm/hr
Moderate Rain	4.00 mm/hr
Heavy Rain	16.00 mm/hr
Very Heavy Rain	100.00 mm/hr

As can be seen from the previous sections, the received field strengths are subject to natural and manmade phenomena. These can cause the field strengths to vary over periods of time and from one location to another. These changes can be long term such as seasonal changes, i.e., weather, temperature, and foliage, or short-term changes such as weather disturbances, i.e., storms and fronts, and vehicles passing in front of the receiver. These variations have an effect on radio systems that is difficult to account for when determining service areas. Thus it is appropriate to describe the field strength statistically, by what percentage of locations will receive a particular field strength for what percentage of time. By describing field variations in this manner, it is possible to determine the service area of a station. However, the terrain still needs to be defined. In preparing propagation curves, this is accomplished by incorporating a terrain roughness factor h. The terrain roughness factor is a generalization of the local terrain and is defined as the difference in elevation between the levels exceeded for 10% and 90% of the terrain along a path. The average value of h for the United States is 50 meters.[11] In using the propagation curves found in the FCC Rules and Regulations for FM and television stations, the local terrain is accounted for by determining the height of the antenna above average terrain along a radial.[12]

Field Strength Prediction

To simplify field strength prediction, curves have been developed to determine the service area of a station. These curves are generally developed using measured values taken from different geographical areas over certain periods of time. The median values are incorporated into a family of curves that describe the field strengths for various antenna heights, frequencies, and distances. A detailed discussion of how to perform field strength measurements may be found in Chapter 2.11 of this *Handbook*, "FM and TV Field Strength Measurements." The FCC curves for FM and television broadcast stations were derived in this manner.[11] The curves used by the FCC for FM and TV describe the field strengths for service at 50% of the locations for 50% of the time. These curves are referred to as F(50,50) and are based on an effective power of 1 kW radiated from a half-wave dipole in free space. The F(50,10) curves used by the FCC describe the field strength for 50% of the locations for 10% of the time. These curves can be used in conjunction with the method described by Allen to estimate the service provided by FM and television stations.[3]

Computer Databases

Through the use of personal computers, field strength estimates can be made quickly, allowing designers to try more options and see the effect on the service area. The designer can change transmitter locations, power levels, and tower heights to optimize the station. There are typically four modules which comprise a computer simulation package: (1) the environmental data base, (2) the equipment data base, (3) the propagation loss module, and (4) the graphical output module.

The environmental data base defines the conditions in which the radio station must operate. For stations operating in the AM broadcast band, this data base is a digitized version of ground conductivity constants for an area. Programs are available which allow a user to search an area to determine appropriate ground conductivity values. For stations operating in the frequency range above 30 MHz this is a topographical data base. The U.S. Geological Survey (USGS) and National Geophysical Data Center (NGDC) provide computer tapes containing digitized terrain elevation data. Several third party sources have adapted these computer tapes to personal computer disk format. The most common resolution provided for terrain elevation provides data spaced at 30 arc-second intervals. The 3 arc-second data base is beginning to become widespread with the acceptance of high density CD-ROM drives for personal computers. The 30 arc-second data base still provides acceptable results for area of coverage predictions, but should be used with caution for STL or microwave planning as it may underestimate significant terrain features. For detailed point-to-point profiling, the 3 arc-second data base is recommended.

The equipment data base defines variables relating to equipment used in the radio system. This includes transmitter output, receiver sensitivity (or field strength), and feed line losses. More sophisticated packages will allow the user to define additional losses such as foliage and buildings to be accounted for by the propagation loss program. In addition, these packages may allow a user to define additional standard deviation values to Rayleigh fading losses. Many packages account for transmitter antennas in varying degrees of sophistication. The simplest accounts for only omni-directional patterns. More in-depth packages allow the user to define directional antenna patterns.

The propagation-loss module uses the information contained in the environmental and equipment data bases to calculate the propagation loss for the area. There are two typical methods for calculating the propagation losses, theoretical and empirical. Theoretical methods are based on well understood physical equations as presented earlier in the chapter. These methods typically do not account for such clutter factors as foliage and buildings. These losses must be accounted for by the user. Empirical methods are based on measured observances of actual propagation characteristics for a particular frequency band and geography. These methods typically employ measurements taken throughout a range of environments, and hence account for such factors as foliage and buildings.

The graphical output module presents the data derived from the propagation loss module in a form easily understood by the user. The output may take two forms. The data can be presented on the computer's CRT terminal. In addition, it may be possible to obtain a hard copy output from a plotter. The plotter output is usually on a clear or translucent film at a scale suitable for overlay to a topographic map.

REFERENCES

1. *IEEE Standard Dictionary of Electrical and Electronic Terms*, ANSI/IEEE Std 100–1977, Wiley-Interscience.
2. *Reference Data For Radio Engineers*, 6th ed., 1982, Howard W. Sams & Co.
3. Allen, Edward W. Wave Propagation, Radiation, and Absorption, NAB Engineering Handbook, 6th ed., p. 119, 1975.
4. Hall, Martin. *Effects of the Troposphere on Radio Communications*, Peter Peregrinus, New York, 1979.
5. Jordan, Edward C. *Electromagnetic Waves and Radiating Systems*, Englewood Cliffs. Prentice-Hall, Inc., 1950, pp. 608–688.
6. Norton, K.A. *Ground Wave Intensity Over a Finitely Conducting Spherical Earth*, Proceedings of the IRE, p. 623, December, 1941.
7. Bullington, K. Radio Propagation Variations at VHF and UHF, *Proceedings of the IRE*, p. 27, January, 1950.
8. "Wave Propagation," *The ARRL Antenna Book*, 14th ed., pp. 1- 19, 1982, American Radio Relay League.
9. Rice, Longley, Norton, and Barsis. "Transmission Loss Predictions for Tropospheric Communication Circuits," National Bureau of Standards Technical Note 101 (Rev.)
10. Okumura. "Field Strength and its Variability in VHF and UHF Land-Mobile Radio Service," (Tokyo) Review of the Electronics Communications Lab. 16, 1968, pp. 825–873.
11. Damelin, Daniel, Fine, and Waldo. "Development of VHF and UHF Propagation Curves for TV and FM Broadcasting," FCC Report No. R-6602.
12. Federal Communications Commission, Rules and Regulations, Section 73.
13. Wang, John C.H. "A Skywave Propagation Study in Preparation for the 1605–1705 kHz Broadcasting Conference," *IEEE Transactions on Broadcasting*, vol. BC-31, pp. 10–17, March 1985.
14. "Attenuation and Scattering by Rain and Other Atmospheric Particles," CCIR Report 721, Vol. V.

2.9 Human Exposure to RF Radiation

William F. Hammett, P.E.
Hammett & Edison, Inc., San Francisco, California

Jules Cohen, P.E.
Jules Cohen & Associates, P.C., Washington, District of Columbia

Richard A. Tell
Richard Tell Associates, Inc., Las Vegas, Nevada

INTRODUCTION

The possibility of biological damage from strong radio frequency (RF) fields has been brought to the attention of broadcasters in relatively recent years. Although the scientific community had been conducting experiments on the effects of RF exposure on animals, tissue, and cell cultures for a number of decades, broadcasters were not alerted to a need to consider such effects until the Federal Communications Commission (FCC) issued a *Notice of Inquiry* in 1979.[1] *A Notice of Proposed Rule Making*[2] followed the *Inquiry* and, in 1985, rule changes were made incorporating criteria for exposure to RF fields.[3] In the absence of a federal RF exposure protection guide, the FCC included in its Rules the criteria set forth in the current standard of the American National Standards Institute (ANSI),[4] the best documented consensus standard available.

Commencing with the effective date of January 1, 1986, broadcasters have been obliged to certify to the FCC that their operations comply with the radiation exposure criteria of the ANSI Standard. Such certifications are required at the time application is made for new facilities, for changes in existing facilities, and for renewal of station license. Certification may not be made without specific knowledge of the facts derived either from appropriate calculations or measurements, both of which will be discussed in this chapter.

In addition to the absolute need to comply with the FCC Rules, broadcasters have become aware of increasing interest on the part of both their employees and the general public to possible hazards from exposure to RF energy. Books, magazine articles, newspaper stories, and radio and television reports have both educated and misled the public. Much of the information disseminated has presented a distorted viewpoint not supported by credible scientific evidence. The existence of such articles and reports puts a special burden on broadcasters to know the facts about RF, including the protection guides, in addition to ANSI, that have been adopted by responsible standards-setting organizations.

The objectives of this section of the *Handbook* are to provide an understanding of applicable RF protection criteria, to define those situations where calculations are appropriate or where measurements are necessary, to provide guidance in the planning of new facilities, and to describe both calculation and measurement procedures.

RF EXPOSURE STANDARDS

In 1966 the American National Standards Institute (ANSI) issued its first standard that recommended maximum exposure values for electromagnetic field intensities. This early guide on exposures was relatively simplistic in that it set a power density limit of 10,000 μW/cm² across the frequency range of 10 MHz to 100 GHz. Since 1966, the ANSI Standard has been twice revised and is presently in the stages of being revised again to reflect new research findings that are relevant to the setting of realistic and accurate RF exposure standards. In addition to the ANSI Standard, there have been developed numerous standards, guidelines, ordinances, exposure policies, and recommendations by federal, state, and local governments. This section is intended as an overview of some of these documents to provide a perspective on what presently exists in the regulatory world of RF fields and some of the trends that are occurring in regulation development.

Virtually all modern RF exposure guides are now frequency dependent; i.e., the recommended maximum exposure levels vary according to the frequency of the exposure fields. This fundamental commonality is based on the finding that the body absorbs RF energy from electromagnetic fields differently at different frequencies. While research continues to provide a better understanding of how RF fields interact with the human body, there have been no substantially different findings which argue for a totally different approach to evaluating the physiological significance of exposure to RF fields. Thus, the rationales that have been developed for different standards or protection guides are merely the result of different judgements about what represents an adequate margin of safety, reflecting a difference in opinion over the magnitude of the margin of safety that is appropriate for the general public rather than any fundamentally different evaluation of what represents a hazardous effect in laboratory animals.

Coupling of Fields to the Body—SAR

In the quest for finding a suitable dosimetric parameter to use for quantifying biological effects of RF fields, the concept of the specific absorption rate (SAR) has been used extensively for many years. This parameter, the result of the complicated manner by which RF fields interact with the body, is the rate at which energy from the incident fields is absorbed, expressed in units of watts per kilogram (W/kg) of body mass. Fig. 1 shows how the SAR for a prolate spheroidal model of the human body exhibits a peak in the so-called body resonance range, that frequency range in which the body is approximately one-half wavelength in height. Three different curves in Fig. 1 correspond to conditions of alignment of the body's long axis with the electric field (curve labeled E), alignment with the magnetic field (curve labeled H), and alignment with the direction of propagation of the RF field (curve labeled K).

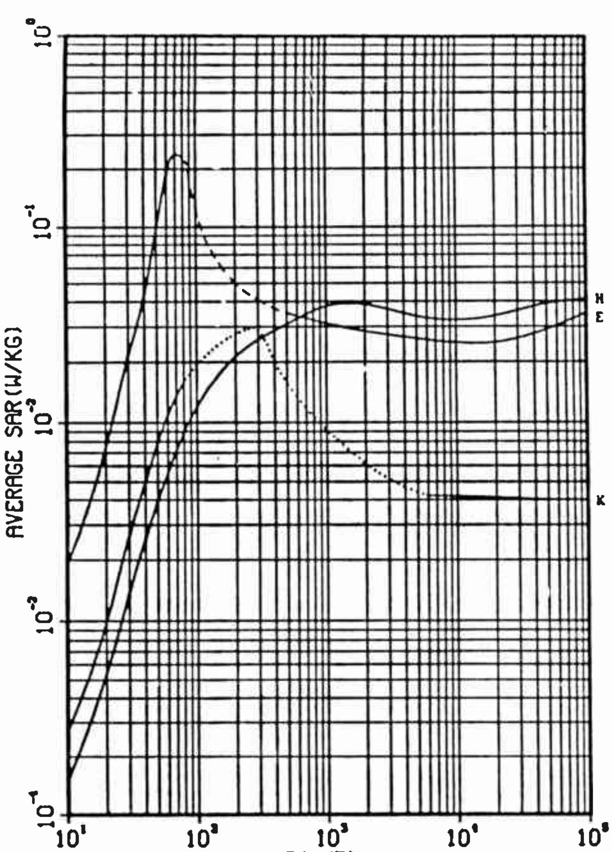


Figure 1. Specific absorption rate (SAR) for electromagnetic fields for an average adult sized human for an incident RF field power density of 1 mW/cm². Curves E, H and K refer to conditions of alignment of the electric field, magnetic field or direction of propagation with the long axis of the body with the assumption of the body being suspended in free space (not grounded). (From Durney et al [1986].)

ANSI Radiation Protection Guide

The current ANSI Standard, C95.1–1982, is a frequency-dependent radio frequency protection guide (RFPG) that specifies maximum values of RF exposure for individuals. Fig. 2 illustrates this frequency dependence. The varying limits are due to the fact that the human body exhibits a frequency response much like a radio antenna, absorbing RF energy more effectively in its resonance range and not absorbing as well at other frequencies. The Standard specifies RF field intensities in two different ways: the permitted electric and magnetic field strengths (actually, the ANSI Standard limits the squares of field strength) and the power densities that would be associated with plane waves having the equivalent electric and magnetic field strengths (plane-wave equivalent power density).

In deriving the RFPG, the ANSI subcommittee elected to assume that man would respond much like the laboratory animals, for which energy absorption rates greater than four to six watts per kilogram, if maintained for protracted periods of time, could eventually be hazardous. To take into account the many variabilities in translating from animal to man, and the range of susceptibilities that individuals might exhibit to RF field exposures, a tenfold margin of safety was introduced to obtain the working limits of the ANSI RFPG. Thus, the specified limits in the ANSI RFPG for RF fields are designed to keep the energy absorption rate, when averaged over the whole body, to less than 0.4 W/kg, ten times less than the presumed hazard level. Since absorption of RF fields within the body is highly nonuniform, ANSI also prescribed limits intended to limit the local SAR value to no more than 8 W/kg in any one gram of tissue.

The plane-wave equivalent power density of a RF field and the associated electric and magnetic fields are related as shown in the following equation:

$$S(W/m^2) = E^2/_{377} = 377\,H^2$$

where S is the power density, expressed here in watts per square meter, and E and H represent the electric and magnetic field strengths, in units of volts per meter and amperes per meter, respectively. The factor of 377 is the impedance of free space in ohms. However, ANSI chose an arbitrary value for the impedance of free space equal to 400 ohms rather than the empirical value of 377 ohms. For this reason, in the ANSI

Frequency Range (MHz)	Electric Field Squared (V^2/m^2)	Magnetic Field Squared (A^2/m^2)	Equivalent Far-field Power Density (mW/cm^2)
0.3 - 3	400,000	2.5	100
3 - 30	4,000 x ($900/f^2$)	0.025 x ($900/f^2$)	$900/f^2$
30 - 300	4,000	0.025	1.0
300 - 1,500	4,000 x (f/300)	0.025 x (f/300)	f/300
1,500 - 100,000	20,000	0.125	5.0

Note: f = frequency (MHz)

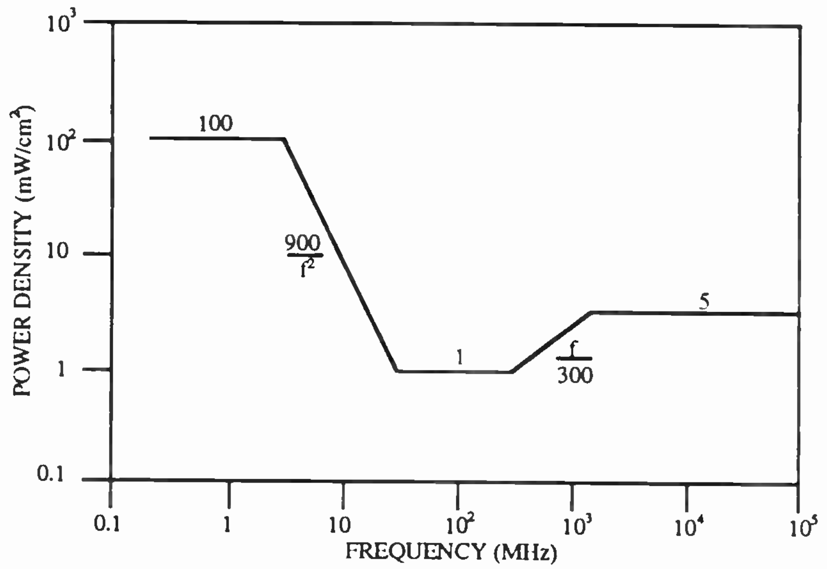


Figure 2. Summary of ANSI C95.1-1982 RF Protection Guide.

standard, 1 mW/cm^2 is equated to a squared electric field strength of 4,000 V^2/m^2 rather than 3,770 V^2/m^2. The equation for power density then becomes:

$$S(mW/cm^2) = E^2/_{4000} = 40\,H^2 \qquad [1]$$

Fig. 2 shows that in the very-high-frequency (VHF) part of the spectrum (30–300 MHz), the ANSI guide for exposure is the most stringent; i.e. it limits the fields to the lowest value anywhere in the electromagnetic spectrum. This is because the adult human body exhibits a whole-body resonance peak at around 70–90 MHz, while the body of a small child resonates at a considerably higher frequency, nearer 300 MHz. Thus the flat limit line between 30 MHz and 300 MHz protects individuals of all sizes from absorbing RF energy at excessive rates.

There are three additional qualifications of the limiting values of RF exposure given in Fig. 2. First, RF fields measured as being equal to the ANSI limit will normally incorporate a significant degree of additional safety since it is generally unlikely that the body will be exactly aligned with the polarization of the incident fields. A second qualification is that, while the RFPG power density limits were based on the concept of a human body being uniformly exposed to a given power density (that value that would result in a whole-body averaged SAR of 0.4 W/kg), the Standard does not explicitly state that the field must be uniform over the body to be of significance, nor does it explicitly state that exposure of only part of the body to a field greater than the specified power density limits is acceptable. Hence, a conservative view is that exposure of any part of the body to RF fields that exceed the RFPG for sufficiently long times represents noncompliance with the Standard.

The third qualification of the limiting values of RF exposure given in Fig. 2 is that they represent the time-averaged value over any six-minute period of time. Thus, the limits given in the ANSI RFPG are applicable to long-term exposure of indefinite extent. But, when the exposure duration can be controlled to periods shorter than six minutes, then higher levels are permitted. This means that RF fields higher than the values shown on Fig. 2 are acceptable, providing that the exposure time is controlled in such a way as to keep the average value over any six-minute window of time to no more than that given by the graph. For example, at 100 MHz, if the exposure time can be controlled to only three minutes in any six-minute period, then the power density can rise to 2 mW/cm^2. During the remaining three minutes of the six-minute period, however, no exposure would be permitted. This provision is helpful in those instances where transient exposures may occur as in passing momentarily through elevated field areas. In practice, determining the actual time-averaged exposure level is complicated because of having to keep track of both the RF field level and the time duration of exposure to each level.

During the development of the ANSI RFPG, it became apparent that controlling exposure to low-powered sources would be extremely difficult. Therefore, based on the concept that only limited SARs would result from exposure to low power systems, ANSI introduced an exclusion clause in the guide which exempts sources of 7 W or less power which operate within the frequency range of 300 kHz to 1 GHz from falling within the compliance requirements of the guide. In this context, most handy-talkies and studio-transmitter-link transmitters are excluded devices because of their low powers.

National Council on Radiation Protection and Measurements

The National Council on Radiation Protection and Measurements (NCRP) is a congressionally chartered, nonprofit corporation established in 1964 to collect, analyze, develop, and disseminate in the public interest information and recommendations about radiation and radiation measurements, particularly those concerned with radiation protection. The NCRP has developed exposure criteria for RF fields but has distinguished maximum recommended RF levels for those occupationally exposed from levels that are applicable to the general public, a distinction not contained in the present ANSI RFPG. The concept introduced by NCRP to recommend more stringent limits for the general public is based on a concern that a greater margin of safety is appropriate since a wider range of health conditions may exist among the public (i.e. there may be individuals that are more susceptible due to age or other factors) and those working with RF fields are more likely to be aware of the potential for hazards and can take preventive measures to limit their exposure.

The NCRP criteria for occupational exposures are the same as the ANSI criteria, shown in Fig. 2. However, the criteria for the general public are equivalent to one-fifth of these values. In the VHF band (30–300 MHz), for example, the NCRP general public exposure limit is 0.2 mW/cm^2. However, the averaging time associated with the NCRP general public criteria is 30 minutes rather than the six minutes used in the occupational exposure criteria and the ANSI RFPG. By selecting a 30-minute averaging time, the total energy absorption (equal to the product of power density and time) is the same for both the general public and workers.

NCRP took a slightly different approach to the issue of fields produced by low power devices: in this case, NCRP states that the "Use of such devices is permitted, as a personal decision by the individual, provided that the devices are designed and used as designed so that the exposure of the individual does not exceed the recommended occupational guidelines and provided that, in using the devices, the individual does not expose other persons above the population guidelines." As such, the NCRP guidelines do not contain a specific power limit for such devices.

IRPA Guidelines

The International Radiation Protection Association (IRPA), an international association of radiation protection societies, in 1988 revised its earlier guidelines for RF exposure to reflect advances that have been made in biological effects research, particularly in the field of dosimetry.

TABLE 1

IRPA occupational exposure limits for RF fields.

Frequency (MHz)	E(V/m)	H(A m)	P_{eq}(mW/cm^2)
0.1–1	614	1.6/f	—
1–10	614/f	1.6/f	—
10–400	61	0.16	1
400–2,000	$3\sqrt{f}$	$0.008\sqrt{f}$	f/400
2,000–3,000	137	0.36	5

However, the limits for magnetic and electric field strengths indicated above for frequencies above 10 MHz may be exceeded for the case of near-field exposure, provided that:

$$\tfrac{5}{6}(E^2/_{377}) + \tfrac{1}{6}(377H^2) < P_{eq}(W/m^2)$$

where E is the electric field strength (V/m) and H is the magnetic field strength (A/m). IRPA recommended limits for the general public are more restrictive, with P_{eq} being one-fifth of that for occupational exposures.

Both the occupational and general public exposure limits recommended by IRPA employ an averaging time of six minutes during a 24-hour day. In this sense, the IRPA limits are more restrictive than the NCRP criteria since relatively high but momentary exposures will be more curtailed according to the shorter time over which they can be averaged. The IRPA guidelines also contain an exclusion clause, similar to ANSI, permitting the use of low powered transmitting devices which operate with powers no greater than 7 W that may produce RF fields exceeding the specified field strength limits.

OSHA RF Standard

The Occupational Health and Safety Administration adopted the 1966 ANSI standard for occupational exposure in 1971, permitting an exposure of up to 10 mW/cm^2 for frequencies between 10 MHz and 100 GHz. The OSHA standard was determined to be merely advisory in 1976 and was implemented by OSHA under the General Duty Clause of the Occupational Safety and Health Act until March 1982, at which time OSHA proposed to revoke its standard. In February 1984, OSHA determined to retain its standard because it provided useful advice for employers. Thus, while the OSHA standard remains on the books even today, it is ten times less restrictive in the 30 MHz to 300 MHz frequency range than the current ANSI RFPG.

ACGIH RF Threshold Limit Values

The American Conference of Governmental Industrial Hygienists (ACGIH) has developed it own threshold limit values for RF fields. The limiting values are similar to the ANSI RFPG in that they are frequency dependent, although, since it is assumed that workers are adults, the power density limits have been tailored to control the SAR in the adult size body and consequently exhibit a different variation with frequency. An averaging period of six minutes is used, similar to ANSI.

Environmental Protection Agency

The Environmental Protection Agency, after almost ten years of work, issued a Notice of Proposed Recommendations for public comment in 1986. This document, "Proposed Alternatives for Controlling Public Exposure to Radiofrequency Radiation," contained four alternative approaches to limit the public's exposure to RF radiation. Three of these options were regulatory and suggested possible numerical exposure limits for RF fields. The fourth option was nonregulatory: information and technical assistance programs would be conducted in lieu of adopting federal guidance.

The options on which EPA sought comment would limit whole-body average SARs to 0.04, 0.08, or 0.4 W/kg. These limiting SARs correspond to RF field power densities in the 30 MHz to 300 MHz frequency range of 100, 200, and 1,000 μW/cm^2 respectively. Thus, the least restrictive option would correspond to the ANSI RFPG while the middle option would be similar to the NCRP criterion for public exposure. Recently, though, the EPA announced that it had determined to phase out its RF program and not to issue Federal Guidance as it had intended to do, citing budgetary constraints to complete projects in apparently higher priority areas and a lack of hard scientific evidence of harm due to exposure to RF fields in the environment. The EPA Office of Radiation Programs has tasked the agency's Human Health Assessment Group (formerly called the Cancer Assessment Group) to perform a comprehensive review of the cancer literature. This project, begun in 1986, seeks to determine whether the literature, when subjected to the same guidelines used by EPA's cancer experts to judge chemicals, supports the classification of electromagnetic fields as a possible carcinogen. The question of an association of exposure to RF fields and cancer is presently unclear and the literature in this area has not been found useful by EPA, ANSI, or NCRP for the development of guidelines. The report from the EPA has reached the review draft stage and is, as of this writing, being reviewed by an external EPA Science Advisory Board committee. The review draft seems to support the contention that it is unrealistic that the cancer issue can be resolved for some years to come in terms of performing quantitative risk assessments.

Selected Foreign Standards

Numerous countries have developed recommendations, if not regulatory limits, on maximum RF fields to which workers and/or the general public should be exposed. Table 2 is intended as a brief summary for some selected countries. Most of these standards reflect the same general type of frequency dependency that the ANSI, NCRP, and IRPA recommendations are based upon, and consequently they can be relatively complex to describe:

TABLE 2

Foreign Nation	Most Restrictive Worker Limit (mW/cm^2)	Public vs. Worker	Time-Average Period
United Kingdom	1	no difference	6 minutes
Australia	1	5 x tighter	60 seconds
Canada	1	5 x tighter	6 minutes
Germany	2.5	no difference	6 minutes
USSR	0.025	2.5 x tighter	2 hrs. or 20 min.
Poland	0.20	2 x tighter	intermittent

It should be noted, however, that the very restrictive limits of the Eastern Bloc countries, such as Poland and the Soviet Union, are different from the others in that there seems to be no scientific rationale for the standards. This is in contrast to those guidelines in most western countries in which relatively clear documentation exists of the underlying database of effects, knowledge, and thinking that went into the development of the standard. The lack of this documentation makes the evaluation of such guidelines or standards more difficult and calls into question their technical validity.

Local Standards

Several states, counties, and cities have enacted, or attempted to enact, standards or ordinances which set upper limits on exposure of the public to RF fields. Some of these standards have relied on the ANSI standard and others have been developed in a manner more similar to the NCRP exposure criteria: examples include Arizona, Connecticut, Massachusetts, New Jersey, and Texas. Counties and cities active in this local regulatory area include Avon, Connecticut; Onondaga, New York; Multnomah County and City of Portland, Oregon; Jefferson and Boulder Counties, Colorado; New York City; and King County and the City of Seattle, Washington. The local codes vary from location to location, as well as within a community over time, which means that the interested broadcaster will need to research the particular restrictions that apply to him.

Pending Revision to ANSI

ANSI standards are subject to regular revision. At five year intervals, unless authority is granted for an extension, standards must be reviewed and either modified, reaffirmed, or deleted. In the case of ANSI C95.1–1982, the revision process began shortly after the issue date. At the time of preparation of this *Handbook*, a revised standard had been approved by a majority of the Subcommittee preparing the revision and by IEEE Standards Coordinating Committee 28. The IEEE Standards Board is expected to approve the revision and send it to ANSI sometime in 1991 to be issued as a revised ANSI standard.

After the revised standard receives approval by ANSI, the FCC is expected to issue a Notice of Proposed Rule Making bringing its environmental rules in line with the revision. Therefore, the expectation exists that broadcasters, at some future date, will be required to show compliance with a different Standard, one that may make a distinction between controlled and uncontrolled environments and one that may introduce new measurement techniques as a means of establishing compliance. If a trend were to be identified in the limiting levels of RF fields which exist in standards settings activities, it would seem to be comparable to the NCRP or IRPA exposure criteria for general population exposure, i.e., approximately one-fifth of the ANSI RFPG levels, at least in the VHF spectrum (30–300 MHz) and higher in frequency.

Serious thought is being given to the question of whether there is a technical basis, in fact, for differentiating between individuals who are occupationally exposed and those in the general population. Such thinking is predicated on the observation that almost all biological effects of RF field exposure appear to be threshold effects; this means that effects do not occur until the SAR is above some threshold value. If this is the case, some argue, then RF radiation is not a zero-threshold agent as ionizing radiation is considered to be. Consequently, more protection of the public is not warranted since lower level exposures are not cumulative in causing effects. RF field biological effects are believed to be principally dose rate phenomena, i.e., they depend on the rate at which energy is absorbed from the field, not the total energy absorbed as is the case with nuclear radiation. This rate concept is demonstrated in the time-averaging provisions of the standards, in which exposure to the RF field may be distributed over some period, typically between 6 and 30 minutes. The averaging time specified in various standards is not related to any cumulative biological effect, but rather to the thermal time constants of the body.

The use of induced and conducted body currents as surrogates for SAR is evolving. Currents in tissues of the body can be related to the SAR in the tissue, and thus, can be used to evaluate conformance of RF exposures with the exposure criteria upon which many protection guides are based. Often, currents, either induced from the field itself or as a consequence of contacting an object immersed in a RF field, can offer a more meaningful indication of potential hazard than the fields themselves. Measurement of contact currents can, in many instances, show that the SAR which may result from touching an object which exhibits strong surface RF fields that exceed the field strength limits of various protection guides does not, in fact, exceed the SAR limits of the guide.

CONSIDERATIONS SPECIFIC TO BROADCAST SERVICES

FCC rules implementing the National Environmental Policy Act of 1969 encompass much more than RF exposure, the subject matter of this *Handbook* chapter.[5] However, this discussion will relate only to the environmental aspects of RF exposure.

To avoid the requirement for an Environmental Assessment,[6] the broadcaster must be prepared to demonstrate by calculation or measurement that their existing or proposed operation subject to results in RF exposures not in excess of the protection guide set forth in ANSI C95.1–1982. Compliance with the protection guide (and with the other aspects of environmental considerations) establishes that Commission actions granting the license renewal or construction permit for new or changed facilities would "have no significant effect on the quality of the human environment and are categorically excluded from environmental processing."[7] Broadcast facilities sites should be selected, or the equipment designed, to achieve "categorical exclusion". Failure to meet the RF protection guide is virtually certain to result in refusal by the Commission of a requested construction permit or license renewal.

AM Stations

In AM broadcast systems, the relatively long wavelengths employed result in the fact that all RF exposures of interest from a compliance viewpoint are in the "near field". The near field is characterized by the fact that the electric (E) and magnetic (H) components are at random phase rather than in phase quadrature as in the far field. The far-field, plane-wave characteristic relationship, $E/H = 377$ ohms, does not apply. Consequently, at any point either component of the electromagnetic field may be high and the other low. Both electric and magnetic fields must be determined independently at each location of interest and checked against the separate exposure criteria for the two fields to determine whether or not compliance has been achieved.

Calculation and measurement procedures for AM will be covered in more detail below. A cautionary note may be appropriate here, however. Measurements are not always what they seem. The most commonly used meters for AM measurements, the Potomac Instruments FIM series, are calibrated in units appropriate for the electric field. However, the shielded loop antenna employed actually responds to the magnetic field. Construction of the scale has been based on an assumption that the meter is being used in the far field, where E(V/m) = 377H(A/m) and where that relationship can be used to translate meter readings to the true magnetic field that is being measured. Measurement of the electric field in the near field must be made with instruments described in the measurement section below.

Critical areas on AM transmitter sites include not only the towers themselves and their proximity, but also the vicinity of inductors used in phasing and matching networks. The highest magnetic fields in an AM transmitting plant are likely to be found near the ends of such inductors. Particularly in high-power operations, where inductors are likely to be carrying large currents, entire tuning houses may be found to exceed the present standards for continuous exposure to magnetic fields.

Transmitters, themselves, are not likely to exhibit significant field leakage from their cabinets unless door seals have been damaged. Nevertheless, leakage should be checked periodically. Windows in transmitters and phasing cabinets must be checked as possible sources of excessive exposure to operators using them for observation of internal components. At any time that work must be performed on the final amplifier stage of a transmitter, high voltage must be removed, not only to avoid excessive RF exposure, but also to reduce the danger of life-threatening shock.

FM Stations

The highest levels of RF exposure near ground level from an FM antenna may not be directly under the antenna, but some tens of feet away. Older FM antennas, in particular, may have a strong downward lobe placing the strongest field at the base of the supporting tower, but antennas of modern design can avoid emission of that "grating lobe." Depending on the number of bays, the radiation characteristics of the individual elements, the spacing between elements, and the extent of null fill and beam tilt, the magnitude and location of maximum field strength from secondary antenna lobes can vary widely.

Near ground level, and except for the effect of reflections from the ground or nearby objects, far-field conditions obtain in that the electric and magnetic fields are in quadrature phase relationship. However, reflected energy affects that relationship, generally requiring that the electric and magnetic fields be considered separately. Another factor requiring consideration is the shape of the radiation pattern. At locations relatively close to the antenna, i.e., at distances less than ten times the antenna length, irregularities not anticipated by the manufacturer's published radiation pattern may occur. Hence, maximum field strength may be found at locations not suggested by study of the radiation pattern.

Transmitter considerations are similar to those in AM except that the final amplification stage is likely to be much better shielded than in the AM case. Nevertheless, caution must be exercised in being certain that the transmitter is inoperative and that the transmission line is properly terminated when work is to be done within the cabinet.

TV Stations

Television antennas are, more often than not, on tall towers or relatively tall structures on building tops. Furthermore, the antennas used have relatively little radiation downward. Consequently, even though substantial power is generally involved, particularly in the UHF band, radiation levels near ground level are usually well below protection criteria. The result is that the major consideration may be related to tower and antenna maintenance rather than ground-level exposure.

Preferably by measurement, the limits of areas conforming to the protection guide should be defined for different power levels of the facility. As described later in this section, safe work rules should be adopted based on those determinations.

Auxiliary Services

Broadcast auxiliary services are covered by Part 74 of the Commission's Rules. Only Experimental (Subpart A), Low Power TV, TV Translators and TV Boosters (Subpart G), and FM Boosters with transmitter output power in excess of 10 watts (Subpart L) are subject to the requirement to certify compliance with the RF radiation protection guide. All other auxiliary services are categorically excluded from routine consideration, because of the low likelihood of excessive exposure problems, although no exclusion is established from the basic need for all licensed facilities to comply with the ANSI Standard.

Multi-User Sites

Joint use by multiple transmitting facilities of a single site involves special considerations relative to determinations of compliance with RF radiation criteria and the need for cooperation among user parties. The Commission has clarified responsibility with the following statement: "In such situations [exceeding of guidelines due to emissions from multiple transmissions], actions necessary to bring an area into compliance shall be the shared responsibility of all licensees, not otherwise categorically excluded, whose transmitters contribute more than 1% of applicable exposure limits."[8] ". . . [I]n the case of a newcomer, not categorically excluded. . .. whose additional contribution of more than 1% of the guidelines would cause a location previously in compliance to become non-complying, the responsibility for corrective action would rest with the newcomer."[9]

The shared responsibility of avoiding exposure in excess of the applicable exposure limits extends to workers required to enter areas of normally high signal strength for the purpose of maintaining equipment. The cooperation of all parties with transmitters at multiple-use sites is mandatory. Although the FCC can be called upon to adjudicate differences among its licensees, a far better procedure is to achieve voluntary agreements without the intercession of the Commission. The Commission has suggested that complaints in this regard should be directed to the Mass Media Bureau; the Field Operations Bureau is not expected to be prepared to intercede in specific cases.

MEASUREMENT PROCEDURES

In anticipation of the effective date for the adoption of the ANSI RFPG by the FCC, the FCC's Office of Science & Technology published Bulletin No. 65, "Evaluating Compliance with FCC-Specified Guidelines for Human Exposure to Radio-Frequency Radiation", October 1985. This document, prepared by Dr. Robert F. Cleveland, forms the basis for all submissions to the FCC regarding compliance with ANSI.

Whenever practical, field measurement of the RF power density levels is to be preferred over calculation, as the definitive method of determining those levels. Care must be taken, however, when taking measurements for the purposes of determining compliance with the ANSI RF exposure guidelines, in order that the examination of the site be complete. The principal concern when taking these measurements is to identify all areas requiring access restrictions, in order to protect fully the liability exposure of the FCC licensee(s) at the site. Failure to identify these areas begs the question of licensee candor in self-certification to the FCC at license renewal, as well as the issue of negligence in the event of legal action based upon "excessive" exposure to RF energy. It is important that the measurement process should be accurately described to allow for a high degree of repeatability in future measurements by others.

Physical Parameters

Power density is most commonly expressed in units of milliwatts per square centimeter (mW/cm^2). Instruments designed for electromagnetic field hazard surveys are typically calibrated in terms of either field strength, field strength squared units, or plane wave equivalent power density. As will be seen, many instruments use detectors that are actually responsive to the square of the field strength, either electric or magnetic, but provide meter readings in plane wave equivalent power density; the meter indications are obtained via an electronic manipulation (Eq. 1). In other cases, instruments designed to sense the magnetic field are calibrated in terms of plane wave equivalent electric field strength.

Other parameters of importance are the frequency or frequencies of the principal source(s) and the duration of the exposure. Virtually all current standards for electromagnetic field exposures are frequency dependent; this dependence takes into account the frequency response of the human body in terms of energy extraction from the incident fields. Because the body acts much like a radio antenna absorbing energy from the field better at certain frequencies, most recommended limits for exposure levels vary in accordance with this frequency selectivity of the body, with the most stringent controls in the body resonance range and less restrictive levels at frequencies both lower and higher than this range. Therefore, when performing a survey of exposure, knowledge of the frequency of the field is necessary to relate to permitted exposure limits and the instrument being used for the measurements must be capable of accurate response at the frequency of exposure.

Electromagnetic fields have a wide range of waveforms depending on their use. For example, a pulsed radar signal typically will exhibit a very high peak to average field strength ratio due to the use of very narrow pulse widths but very high peak powers. In the case of normal AM radio broadcasting, again the average value of the field will be dependent on the programming material and adjustments to the transmitter. TV broadcasting involves the transmission of synchronization pulses which have considerably greater peak powers than the associated video programming material. The waveform of some fields may be of interest because of their rapid changes in time, which, for magnetic fields, may be important from the standpoint of induced currents in the body.

Broadband Instruments

In terms of convenience, broadband instruments are considerably more popular than narrowband instruments for evaluating exposure to electromagnetic fields. Broadband instruments, as their name implies, respond over a wide range of frequencies, possessing nearly flat responses independent of frequency within their frequency pass band. Most common broadband meters feature isotropic probes, consisting of three mutually orthogonal detecting elements connected electronically in such a manner that the output of the meter becomes essentially independent of probe orientation within the field. Isotropy is typically achieved to within ±0.5 to 1 dB, but no information on the electrical phase of the fields is retained.

Probe detection elements are normally either thermocouples or diodes. Thermocouples exhibit the useful property of being average-responding detectors. When placed in a modulated field, such as a pulsed radar field, they will accurately respond to the average value of the field squared or average plane-wave equivalent power density. This same property, i.e., true root mean square (RMS) detection, of thermocouples, makes them useful in multiple frequency electromagnetic field environments since they can accurately respond to the sum of the average fields present at the probe. Diode detectors, on the other hand, tend to be peak detectors rather than average detectors unless operated within limited field strength ranges where they can exhibit so-called square law responses, i.e., they will respond correctly to the power density or square of the fields. In some instruments, special circuits are used to correct for the nonsquare law operation of diode detectors. This technique works well for single frequencies, but cannot properly correct readings when multiple signals are being measured. In summary, while diode detectors have certain disadvantages, they also have several advantages over thermocouples, including durability, ability to withstand overloads from strong fields, significantly better thermal stability, and high sensitivity.

While isotropic survey instruments have separate probe assemblies connected by cable to the meter, some survey instruments have single or multiple sensing probes mounted directly on the instrument package. Those instruments equipped with one probe element are responsive to only one field polarization component at a time. The absence of a connecting cable between the probe and the readout package, in some cases, is responsible for superior performance when used in low-frequency fields; it is common for cables to "pick up" or couple with the electric field at lower frequencies, typically below about 3 MHz, often leading to erroneous indications of field strength. Cableless instruments, or ones equipped with nonconductive, fiber optic cables, are preferred for electric field measurements at low frequencies.

Broadband instruments can exhibit flat frequency responses over large frequency ranges. For example, electric field probes can be purchased with specified passbands as wide as 200 kHz to 40 GHz. Magnetic field probes have narrower passband characteristics, such as 0.5 MHz to 300 MHz, and are not generally available for UHF measurements. Care must be used in conducting broadband magnetic field surveys in the presence of UHF broadcast stations, since the loop sensors of the instrument may exhibit high frequency resonances leading to erroneous indications of fields.

The most common broadband isotropic field strength meters, well suited to measurements taken for purposes of ANSI compliance determination, are shown in Table 3.

Most consulting engineering firms own meters of this type to support their work in this field, and some of the meters are available, as well, from certain equipment leasing companies. The total accuracy of these meters is typically ±2 to 3 dB, due to variations in frequency response, isotropicity, and repeatability. Other errors can be introduced, as well, but none which would cause the meter to read low, according to the manufacturers' literature. Therefore, while the

TABLE 3

Make & Model Series	Probe Freq. Range[10]		Suitable Applications
IFI RFH	E-field: 0.01–220	MHz	AM
Holaday HI-3000	E-field: 0.5–6000	MHz	FM, VHF, UHF, some microwave
	H-field: 0.3–300	MHz	AM, FM, VHF
Narda 8700	E-field: 0.003–40	GHz	FM, VHF, UHF, microwave
	H-field: 0.3–300	MHz	AM, FM, VHF

measurements are expected to be conservative, leeway should be allowed when establishing the locations of ANSI-threshold contours.

There is also a measurement probe that accounts for the frequency variation of present ANSI RFPG by weighting strength of the fields according to frequency such that it indicates in units of percentage of the exposure criterion permitted. As with any broadband type of instrument, however, it does not indicate how the resultant field magnitude is distributed across the frequency spectrum.

Narrowband Instruments

Broadband instruments, while convenient for conducting area surveys and rapidly assessing exposure levels of electromagnetic fields, do not yield information on the frequency of the fields at a measurement point. Because of the frequency dependence of most RF field protection guides, however, such information is often desired, particularly where frequencies are present for which different exposure limits apply. If total RF power density exceeds 1 mW/cm^2 in such a situation, the contribution that each field makes to the total must be determined in order to know whether the protection guide is actually being exceeded.

A tuneable receiver is the most common instrument for this purpose, allowing each emission frequency to be individually tuned and measured. A spectrum analyzer, which repetitively scans between a lower and upper frequency, producing a graphical display of the measured signal strengths in that range, permits rapid assessment of the relative signal amplitudes across a wide range of frequencies. It is particularly effective when the frequency of a specific source is unknown or when the presence of the field is intermittent.

In either case, however, narrowband meters and their companion sensing antennas are not as portable as broadband meters. It can be very inconvenient or simply impractical to conduct the measurement using the large antennas typically employed with narrowband instruments. This is a serious consideration when performing field measurements under difficult circumstances, such as on tall broadcast antenna towers, and in other situations where it is impractical to carry narrowband equipment. In these cases, arrangements must be made to operate the various sources individually, so that their individual contributions can be assessed with a broadband type of instrument.

Special Adaptations or Applications of Instruments

In the last several years, data processing hardware has been introduced that can be configured with portable RF survey instruments to simplify the data collection task. These devices, which have taken the form of microprocessor based data-loggers or analog devices for simply integrating exposure over some prescribed time interval, now allow direct assessment of the time-averaged exposure to electromagnetic fields. The attractiveness of such devices is that one can use them to actively manage personnel exposures in areas where the exposure levels may reach substantial values momentarily. By constantly monitoring the time-averaged value, the exposed individual has the knowledge to remove himself from the field to keep the average level within acceptable limits.

An area of instrumentation highly developed in the ionizing world of radiation, personal dosimetry, has not been successfully developed for the general case of electromagnetic fields. This has been the case because of the difficulty in relating radio frequency field strengths determined very close to the surface of the body to equivalent, unperturbed values from which SAR may be derived. At lower frequencies, however, personal dosimeters have been developed for use in power frequency fields, i.e., 50 Hz and 60 Hz. At these frequencies, especially for magnetic fields, body perturbation effects can be accounted for in a simpler manner than for higher radio frequencies. Some of these dosimeters contain advanced digital circuitry which allow the retention of data on exposure levels versus time of day. Hence, very complete exposure histories can be determined for subjects wearing these devices. Recently, new developments in the form of personal RF monitors useful in the microwave frequency range have been commercially announced.

Measurement Techniques

There are several different approaches to using broadband exposure meters, depending on the type and number of stations at the site of interest and the purpose of the measurements. In general, a top-down approach is recommended, ensuring that no areas of the site are neglected while also ensuring that potential problem areas are examined in sufficient depth.

The basic technique is to make a quick pass first, throughout the site, of ambient fields, to determine where to concentrate and what fields/boundaries to locate and record. It may be, for instance, that an area exists in which the ambient fields exceed the ANSI RFPG; this is the type of area that would be examined next. There, the location of the ANSI-threshold contour should be defined carefully, for fencing or marking at the time the mitigation measures are determined.

Next, the localized measurements are done, looking for the so-called "hot-spots." The FCC adopted under Docket 88–469[11] a relaxed distance requirement for measuring hot-spots. In the concluding paragraph (42) of the *Report and Order*, the following statement by the Commission provided somewhat conflicting directions:

> "In summary, it is recommended that during routine measurements of radio frequency fields for compliance purposes, a minimum separation distance of 20 cm be maintained between a reradiating object and the closest sensing element of a probe. However, as a precautionary measure, it is also recommended that consideration be given to the presence of intense, localized fields in the range of 10 to 20 cm from a reradiating object."

This paragraph has been interpreted by some with emphasis on the first sentence, so that measurements are generally made no closer than 20 cm from reradiating objects, but with sufficient measurements made at a distance of 10 cm to determine whether the RFPG is exceeded where "intense, localized fields" are present. Others interpret the paragraph more conservatively, reasoning that, without consistently measuring as close as 10 cm, one cannot give "consideration" to the presence of intense, localized fields in the 10 cm to 20 cm range. Nevertheless, this increase in measurement spacing, from 5 cm specified in the ANSI RFPG, has alleviated the worries at many sites about localized fields, which tended to occur frequently atop wire-mesh fences or at gates, locations where whole-body SARs would probably remain low but where reproducible measurements indicate fields in excess of the RFPG.

Finally, measurements are made inside the transmitter buildings at the site, looking for both ambient and localized fields and searching with particular care, since these are generally areas of prolonged exposure. Additional grounding will often remove or reduce the severity of localized fields. It is also often helpful for later reference, when determining mitigation measures, to temporarily ground possible sources of reradiation to evaluate the effectiveness of grounding in that specific case.

CALCULATION METHODS

Bulletin OST-65 contains in its Appendices B, C, and D tables of calculated "worst-case" distances from broadcast antennas, beyond which one may assume compliance with ANSI. These tables should be the first "calculations" one performs, and they will, in many cases, suffice. At many sites, exposure conditions are so low that no calculations other than table look-ups are required. The OST-65 tables assume only a single station at a site. In some cases, it may still be possible to use these tables where there are several stations at the site, by conservatively assuming all of the sources to exist at the lowest antenna height and using the least directive antenna pattern.

For those more complicated situations involving FM and TV stations where reference to the tables is not adequate to establish compliance, OST-65 specifies in its Section II the formulas by which predicted RF fields are to be calculated. The general form of this equation is as follows:

$$\text{Power Density } S = \frac{2.56 \times 1.64 \times RFF^2 \times [0.4 \times VERP + AERP]}{4\pi D^2} \quad \text{in mW/cm}^2$$

where: $VERP$ = Total peak visual ERP (all polarizations), in kilowatts
$AERP$ = Total aural ERP (all polarizations), in kilowatts
D = Distance from the center of radiation to the point of calculation, in meters
RFF = Relative field factor at the direction to the actual point of calculation, a unitless value between 0 and 1.

The factor of 2.56 accounts for the increase in power density due to ground reflection, assuming a field reflection coefficient of 1.6 ($1.6 \times 1.6 = 2.56$). The factor of 1.64 is the gain of a half-wave dipole relative to an isotropic radiator. The factor of 0.4 converts peak visual ERP to an average RMS value; for FM stations, of course, the value of VERP is zero. Note that the ERP is total power, so stations employing circular polarization must double their authorized ERP when using the formula. Finally, the factor of 100 in the numerator converts to the desired units of power density.

This formula contains all of the factors necessary to predict the RF power density at any point relative to the source of radiation. In practice, there are several complications which arise, including how to determine the RFF, where to calculate the RF levels, and how to handle multiple sources. The RFF factor is the product of the azimuth and elevation plane relative fields at the specific azimuth and depression angle toward the calculation point. The azimuth plane pattern is generally well defined, but elevation plane patterns beyond about $-20°$ were not published for many of the older FM and TV antennas, and may require supplemental information from the manufacturer for the specific installation being studied. In any case, there is generally enough scatter due to the antenna mounting configuration that no elevation plane relative fields less than 15% should be used for calculations.

The RF levels should be calculated, as a minimum, at locations on the site where public access is a possibility. This often means determining ground elevations at the site with greater precision and resolution than previously needed, determining the slant distance to the antenna, and determining the relative field at the appropriate depression angle for each point to be calculated. Many FM antennas have significant lobes at steep depression angles, so one should beware of the common trap of calculating at points along a radial line from the tower and assuming that decreasing fields will not increase further out when the next lobe upward is encountered. In addition, calculations should be performed for ground-level areas and at on-tower locations where occupational access is required. When calculating conditions on a tower supporting an FM or TV antenna, it must be noted that the OST-65 tables and formulas assume that the source of radiation is at a single point. For close approaches to a multibay antenna, this assumption does not hold, and the antenna should be broken down into its several bays for modeling, each with its proportionate share of the total power.

At a multi-user site, OST-65 states that the sum of the ANSI-fraction contributions from all sources must not exceed unity. This requires, of course, that the frequency-sensitive ANSI guideline for each source be used to normalize the results of the power density calculations. Also, when calculating (or measuring) RF exposure conditions at a multi-user site, attention must be paid to the effect of any auxiliary (standby) facilities that may exist, as these will often create worse exposure conditions than the corresponding main facilities. Even where a measurement program is to be followed, prior calculations can help to identify those combinations of main and auxiliary facilities which present the worst cases in need of measurement.

Computer programs can be developed for performing these calculations, using the OST-65 formulas and incorporating all of the special considerations discussed above. In particular, knowing the make and model of the antenna should be adequate to determine the elevation plane relative field pattern, from research of the FCC's original source documents. In addition, the program should have the capability of handling nonplanar-site surfaces, so that the calculations may be performed over the entire site at once. Fig. 3 shows sample results from such a program. Note the ring of higher fields, due to the lower lobe of the antenna pattern, and the rectangular shaped extension at the southwest edge of that ring, corresponding to higher calculated fields on the roof of the transmitter building.

One final note when relying on calculations to demonstrate compliance with ANSI: 'hot-spots' can not realistically be modeled. The calculation method assumes that the point of calculation is in the far-field of the radiation source. However, intense localized fields of reradiation can occur, particularly at FM and TV sites. A useful rule-of-thumb is that, when the calculated fields are within 50% of the ANSI guideline, measurements should be performed.

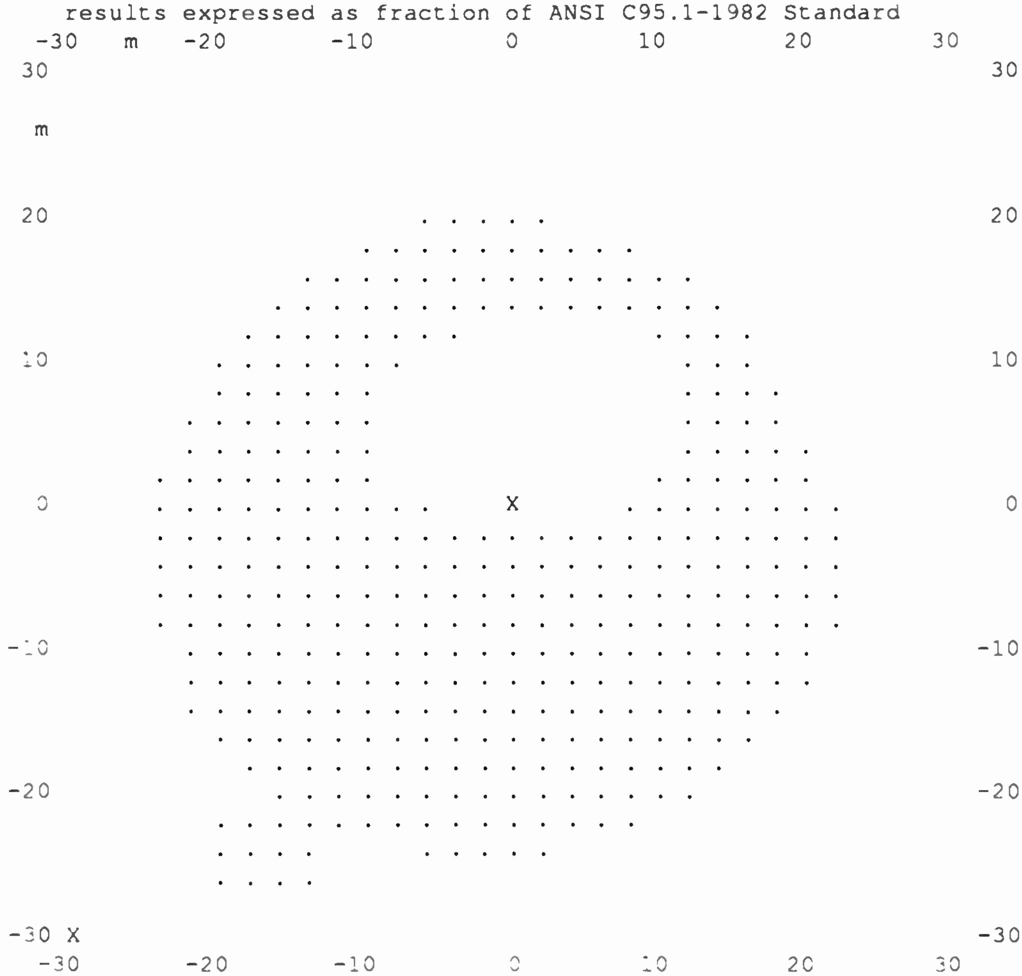
MITIGATION OF HAZARD

There are three distinct exposure conditions to consider when assessing whether a particular facility complies with ANSI. These are:

- Public exposures
- Ground-level occupational exposures
- On-tower occupational exposures

Compliance can be achieved in different manners for each condition, particularly since personnel author-

```
   RADIO-FREQUENCY ELECTROMAGNETIC FIELD POWER FLUX DENSITY
    calculations performed according to OST Bulletin No. 65
    results expressed as fraction of ANSI C95.1-1982 Standard
```




```
  calculation height above site: 2 m        site contour file: NAB.ELV
call  ch   location   ANSI   kW ERP, x F  ht,m  patterns H/V   @base   @ max
WFM   250    0    0   1.00    35.00 2.00  72.0  OMNI /FM8-2    0.034   0.054
WTV   62   -30  -30   2.54  1000.00 0.50  56.0  OMNI /UHFSTD   0.037   0.040
                      ©1990 Hammett & Edison, Inc.
legend for fields:  below .05|.05 to .5|.5 to 1.0||1.0 to 5 | above 5
(fraction of ANSI)    blank  |    .    |    :    ||    o    |    O
```

Figure 3. Example of RF hazard calculation program results.

ized to be at a site can be instructed as to the meaning of certain markings and as to the nature of certain procedures to be followed. In those cases where measurements or calculations show that there are no potential problems with excessive exposure to RF radiation, then no special measures need to be taken. However, in almost all cases there is at least the issue of on-tower occupational exposure, therefore much of this section will be applicable to all facilities.

When action is necessary to bring a site into compliance with ANSI, whether determined by measurement or by calculation, such action can have one of two goals: (1) to limit access by the public or by workmen to the fields in excess of ANSI, or (2) to reduce those fields to levels less than ANSI. All compliance techniques involve one or the other of these objectives. The FCC *Public Notice*, "Further Guidance for Broadcasters Regarding Radio Frequency Radiation and the

Environment", dated January 28, 1986, outlines a number of situations and provides guidance on which of the above approaches is the more appropriate compliance mechanism. It is a helpful aid for deciding the extent of action required in any particular case.

Access to areas of high RF fields may be limited by means of physical barriers, such as fences, to preclude access by the public. In combination with warning signs, barriers are the preferred method of achieving compliance for public exposure. The Notice does recognize that the remoteness of a site and the presence of natural barriers may be pertinent; essentially, if one cannot build a scenario in which the public ("including trespassers," the Notice warns) would have occasion to be exposed to high RF fields, then the fields probably do not require barriers, although the Notice implies that signing is always required.

For occupational exposure considerations, lesser measures may suffice to restrict access to high RF fields. In particular, these would include procedures to take advantage of the time-averaging provisions of ANSI when only intermittent access is required for maintenance activity, such as reading AM base current meters. Areas in which specific procedures must be followed can be marked with contrasting paint or by low fencing. The meaning of these markings can be explained in a document called an Occupational Exposure Guide. This document is an 8½ × 11 inch, plastic-laminated sheet which describes the ANSI Standard and its time-averaging provisions; explains the significance of the markings and fences; defines the maximum time allowed in the marked areas; gives the reduced power factors for on-tower work; and provides several telephone numbers for obtaining additional information. Multiple copies of the OEG for a particular site should be available, so that one is displayed on the inside of each exterior door at the site. Several should also be kept at each studio location for review with any workers, such as HVAC repairpersons or two-way communications technicians, who may be authorized for occasional access to the site.

The other compliance mechanism, reduction of the fields themselves, clearly applies when considering on-tower maintenance work. The power cutback requirements for such work should be a part of the OEG. It may be, however, that changes to the antenna system(s) can remove large contributions to the ANSI limit from particular areas and achieve ANSI compliance in those areas. Typically, these changes can apply at FM or TV sites and involve raising an antenna to place it higher above the site or replacing the antenna with a model having reduced downward radiation, such as is now available from most antenna manufacturers. Yet another factor to consider is the establishment of separately-located auxiliary facilities, to which the station(s) could switch operation while maintenance work is required at the main facilities. Too often, the auxiliary facilities provide only emergency service because their operation creates RF fields above the ANSI limit at the main antenna.

One particular advantage to calculating RF levels occurs when considering on-tower conditions at a multi-user site. Here, if maintenance is to occur on any tower, the licensees must know the exposure conditions at various heights on that tower and, in order to reduce power as necessary, must know how much each station is contributing to the total fields at those heights. To acquire this information by measurement is not practical, at least not without exposing the person taking the measurements to fields in excess of ANSI. Generally, for even a few stations at a single site, the number of power-reduction combinations is so large that attempting to effect even some fraction of them is difficult and, without knowing beforehand which combinations are likely to achieve ANSI compliance, the attempt is often futile.

With regard to on-tower occupational exposures, the use of "bee-keeper" suits, such as the Lion Uniform M-Guard model, has received no specific approval from any government or industry group for use in reducing RF exposures at broadcast frequencies. Therefore, even though the material is believed to provide significant attenuation at some frequencies, it is imprudent at present to rely upon these suits to achieve compliance with ANSI.

There are, then, several good rules of thumb to follow for achieving compliance with ANSI: (1) fence all tower bases to limit access, (2) post RF warning signs so that they are visible from all approach angles, (3) establish a written policy for compliance with ANSI (have station employees sign it, and give copies to all contractors), (4) send letters (if at a multi-user site) to all other users, expressing your commitment to comply, and (5) develop an auxiliary facility that does not limit access to the main antenna.

REFERENCES

1. *Notice of Inquiry*, General Docket No. 79–144, adopted June 7, 1979, released June 15, 1979, 44 Fed. Reg. 37008 (1979), 72 FCC 2d 482 (1979).
2. *Notice of Proposed Rule Making*, General Docket No. 79–144, adopted January 28, 1982, released February 18, 1982, 47 Fed. Reg. 8214 (1982), 89 FCC 2d 214 (1982).
3. *Report & Order*, General Docket No. 79–144, adopted February 26, 1985, released March 14, 1985, 50 Fed. Reg. 11,151 (1985), 100 FCC 2d 543 (1985).
4. *Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 300 kHz to 100 GHz*, ANSI C95.1–1982, approved July 30, 1982, The American National Standards Institute, 1430 Broadway, New York, N.Y. 10018.
5. See FCC Rules, Section 1.1307(a).
6. 47 CFR 1.1307(b).
7. 47 CFR 1.1306(a).
8. *Report & Order*, General Docket No. 88–469, Paragraph 1, adopted December 20, 1989, re-

leased January 18, 1990, 55 Fed Reg 2380. See also 47 CFR 1.1307(b), Note 2.
9. *Ibid*, Paragraph 22.
10. Note that these frequencies correspond to the overall range for which probes are available; i.e., more than one probe may be needed to cover the entire indicated range.
11. *Ibid*, Paragraph 42.

APPENDIX

Further Guidance For Broadcasters Regarding Radio Frequency Radiation And The Environment

(Federal Communications Commission Public Notice, January 28, 1986)

"The National Environmental Policy Act of 1969 (NEPA), 42 U.S.C. Sections 4321–4361, requires all federal agencies to ensure that the environment is given appropriate consideration in agency decision-making. In a *Report and Order* in General Docket No. 79–144, 100 FCC 2d 543 (1985), the Commission decided that human exposure to radio frequency (RF) radiation was a proper environmental concern of this agency and specified that the guideline for determining the significance of such exposure will be the "Radio Frequency protection Guides" adopted in 1982 by the American National Standards Institute (ANSI C95.1–1982). As of January 1, 1986, all applications for new facilities, modifications to existing facilities, and renewals must contain *either* a specific indication that the RF radiation of the particular facility or operation will not have a significant environmental impact *or* an environmental assessment which will serve as the basis for further Commission action.[1] *See* Part 1, Subpart I of the Commission's Rules for specific regulations regarding environmental matters.

Most broadcasting facilities produce high RF radiation levels at one or more locations near their antennas. That, in itself, does not mean that the facilities significantly affect the quality of the human environment. Each situation must be examined separately to decide whether humans are or could be exposed to high RF radiation. Paragraph 37 of the *Report and Order* points out that accessibility is a key factor in making such a determination. As a general principle, if areas of high RF radiation levels are publicly marked and if access to such areas is impeded or highly improbable (remoteness and natural barriers may be pertinent) then it may be presumed that the facilities producing the RF radiation do *not* significantly affect the quality of the human environment and do not require the filing of an environmental assessment.

Because we wish to avoid burdening applicants with unnecessary work, expenses and administrative filings, we offer the following guidance as to how we will view typical situations. The term "high RF level" means an intensity of RF radiation, whether from single or multiple sources, which exceeds the ANSI guidelines.

Situations

A. High RF levels are produced at one or more locations above ground level on an applicant's tower

- If the tower is marked by appropriate warning signs, the applicant may assume that there is no significant effect on the human environment with regard to exposure of the general public.

B. High RF levels are produced at ground level in a remote area not likely to be visited by the public

- If the area of concern is marked by appropriate warning signs, an applicant may assume that there is not significant effect on the human environment with regard to exposure of the general public. It is recommended that fences also be used where feasible.

C. High RF levels are produced at ground level in an area which could reasonably be expected to be used by the public (including trespassers)

- If the area of concern is fenced *and* marked by appropriate warning signs, an applicant may assume that there is not significant effect on the human environment with regard to exposure of the general public.

D. High RF levels are produced at ground level in an area which is used or is likely to be used by people, and to which the applicant cannot or does not restrict access

- The applicant must submit an environmental assessment. This situation may require a modification of the facilities to reduce exposure or could lead to a denial of the application.

E. High RF levels are produced in occupied structures, on balconies, or on rooftops used for recreational or commercial purposes

- The applicant must submit an environmental assessment. The circumstances may require a modification of the broadcasting

1 In applications for new and modified facilities, the requirement for a specific indication is satisfied by answering the question on the form regarding environmental matters. An environmental assessment is the narrative statement described in Section 1.1311 and elsewhere in the Commission's rules.

facility to reduce exposure or could lead to a denial of the application.

F. High RF levels are produced in offices, studios, workshops, parking lots, or other areas used regularly by station employees

- The applicant must submit an environmental assessment. The circumstances may require a modification of the facilities to reduce exposure or the application may be denied. This situation is essentially the same as E. We have included it to emphasize the point that station employees as well as the general public must be protected from high RF levels. Legal releases signed by employees willing to accept high exposure levels are not acceptable and may not be used in lieu of corrective measures.

G. High RF levels are produced in areas where intermittent maintenance and repair work must be performed by station employees or others

- ANSI guidelines also apply to workers engaged in maintenance and repair. As long as these workers will be protected from exposure to levels exceeding ANSI guidelines, no environmental assessment is needed. Unless requested by the Commission, information about the manner in which such activities are protected need not be filed. If protection is not to be provided, the applicant must submit an environmental assessment. The circumstances may require corrective action to reduce exposure or the application may be denied. Legal releases signed by workers willing to accept high exposure levels are not acceptable and may not be used in lieu of corrective measures. The foregoing also applies to high RF levels created in whole or in part by reradiation.

A convenient rule to apply to all situations involving RF radiation is the following:

1. Do not create high RF levels where people are or could reasonably be expected to be present, *and*
2. Prevent people from entering areas in which high RF levels are necessarily present.

Fencing and warning signs may be sufficient in many cases to protect the general public. Unusual circumstances, the presence of multiple sources of radiation, and operational needs will require more elaborate measures.

Intermittent reductions in power, increased antenna heights, modified antenna radiation patterns, site changes, or some combination of these may be necessary, depending on the particular situation.

For further discussion see Office of Science and Technology Bulletin No. 65, "Evaluating Compliance with FCC-Specified Guidelines for Human Exposure to Radiofrequency Radiation," October 1985. Copies of this bulletin may be ordered from the National Technical Information Service, (800) 336–4700, Order No. PB 86–127081.*

For further information regarding this *Notice*, applicants should contact the AM Branch, FM Branch, TV Branch, or LPTV/Translator Branch, as appropriate."

---

* NOTE: A limited number of copies are also available from the FCC, (202) 653–8169.

2.10
AM Field Strength Measurements and Proof-of-Performance

Donald G. Everist
Cohen, Dippell, & Everist, P.C., Washington, District of Columbia

INTRODUCTION

An AM directional antenna is composed of at least two or more radiating elements. The amplitude and phase relations of RF energy fed to each element and the physical spacing between each element produces a predicted directional pattern shape. Each directional pattern is designed and constructed for a specific application, to provide the optimum service to the desired area while providing adequate protection for other, cochannel and adjacent channel stations.

The design and construction of the elements of an AM directional antenna are covered in Chapter 2.1, "Design, Erection, and Maintenance of Antenna Structures," and Chapter 2.5, "AM Broadcast Antennas."

The obligations of broadcasters regarding possible human exposure to RF radiation are discussed in Chapter 2.9, "RF Radiation Hazards." This chapter should be reviewed to ensure that the station is in compliance with guidelines of the Federal Communications Commission.

Quick Reference Guide
Selected FCC Rules Applicable to AM Directional Antennas

Antenna Resisitance Measurements	73.54
Antenna Testing During Daytime Hours	73.157
Directional Antenna Monitoring Points	73.158
Directional Antenna System Parameter Tolerance	73.62
Emergency Antennas	73.1680
Equipment Test	73.1610
Operational During Modification of Faciltiies	73.1615
Operating Power	73.51
Partial Proof-of-Performance	73.154
Proof-of-Performance	73.151
	73.186
Special Temporary Authority	73.1635

(This list is not comprehensive. A more extensive and general list of FCC requirements is included in Chapter 1.2, "FCC Field Operations Bureau," Appendix A.)

Monitoring Points

An important, initial step in setting up a directional antenna is the selection of monitor points along the radials specified in the construction permit. These points must be reachable in inclement weather, and free of possible reradiation objects such as underground pipes and overhead wires. Schoolyards, churches, and cemeteries often provide useful locations since the surrounding area is usually clear.

Prior to submission to the FCC, data should be taken at proposed monitoring points over a period of time until the accuracy and stability of measurements can be established. Readings must be in the direction of the station. This can be established by using a U.S. Geological Survey topographic map or by switching the antenna array to the nondirectional mode (if available).

Once the directional array has been adjusted and its performance documented, one of the often overlooked items is the monitoring of the area for new construction. Vertical structure construction or new construction authorized by governmental authorities is often very difficult to detect in the advanced stage of planning. New power-line construction and construction of other communication structures can act as reradiators if located sufficiently close to the array. Such reradiated energy can affect the pattern so as to reduce interference protection to other stations.

PROOF-OF-PERFORMANCE

The proof-of-performance is used to establish the initial operation of the directional pattern at the time of licensing; it is a condition imposed by the FCC before licensing can occur; it is used to update the performance of the antenna system at such times as may be necessary or may be directed by the FCC; and it is used as a reference for subsequent partial proof-of-performance.

Reference Proof-of-Performance

A *reference proof* is the latest complete proof-of-performance accepted by the FCC, and is the proof to which *partial proof* measurements must be referenced. The reference proof-of-performance specifies, for each monitoring point, the point number, the distance in miles from the transmitter site, the radial and the point location. Section 73.151 of the FCC Rules defines the general field strength measurement requirements to be made on the construction permit and nonconstruction permit radials for the reference proof.

A reference proof defines the nulls, suppression, and major radiation areas of the directional pattern. A reference proof-of-performance requires the taking of *nondirectional* and *directional* measurements along each radial under similar environmental conditions. While each proof has its own requirements, in general, nondirectional measurements begin (as specified by the FCC Rules, Section 73.186) at a distance of five times the height of the nondirectional antenna. They are to be made at approximately equal sets of intervals such that, as a goal, 15 to 20 measurements are established within the first three kilometers (two miles), eight to ten measurements within the next seven kilometers, (miles two through six) and 14 measurements to extend to a distance of 24 to 32 kilometers from the transmitter site (miles 6 through 15 to 20). Directional measurements are to be made under similar environmental conditions beginning at ten times the widest spacing between the elements of the antenna system. The instantaneous changing of patterns from nondirectional to directional at either a prearranged time or by two-way communications can be beneficial. This technique permits the acquisition of the measurement data in the minimum amount of time and permits the continuous measurement observation without moving the field strength instrument. The accumulation of measurement data whereby both the nondirectional and directional measurements are made at a given point at the same time will eliminate relocation and time differences.

Measuring Instrument to be Used

When field strength measurements are made in support of the partial or reference proof-of-performance, a portable instrument made solely for this purpose and of known accuracy is utilized.

For directional antenna systems, the taking of a proof-of-performance and the availability of the measurement instrument are conditions of the construction permit as provided below:

THE AUTHORITY GRANTED IS SUBJECT TO THE FOLLOWING CONDITIONS:

> Field measuring equipment shall be available at all times and, after commencement of operation, the field strength at each of the monitoring points shall be measured at least once every seven days and an appropriate record kept of all measurements so made.

A complete nondirectional proof-of-performance, in addition to a complete proof on the directional antenna system shall be submitted before program tests are authorized. The nondirectional and directional field strength measurements must be made under similar environmental conditions.

Authorization to Use an AM Directional Antenna System

The directional antenna system requires FCC approval before its operation can commence. Unlike nondirectional operations, directional operations cannot be instituted upon a notice and a promise of subsequent FCC submission to fulfill license requirements. The FCC program test authority is a telegraphic authority. The program test authority indicates that the FCC is in essential agreement that the antenna system is in compliance with the FCC Rules and indicates that permittee has met its basic obligations and authorizes that operation can commence as described in its license application. However, specific requests by the Commission may accompany the program test authority and those requests must be satisfied before the FCC will issue a license. The FCC requires for each directional operation after submission of the proof-of-performance and receipt of the program test authority from the FCC a thirty day set of measurements. This request for the data will accompany the issuance of the antenna parameter data and has language similar to the following:

> Program test should be conducted with the directional antenna system adjusted in accordance with the enclosed corrected specifications pending further action on the license application. It is requested that you check the values of field strength at each of the monitoring points at least once weekly during the next thirty (30) days, and submit this information in tabulated form to the Commission together with the following meter readings extracted from the transmitter log at the time the monitoring points are checked: (1) common point current, (2) base currents and their calculated ratios, (3) antenna monitor sample current ratios, (4) phase indications, (5) final amplifier plate voltage, and (6) plate current.

PARTIAL PROOF

Section 73.154 of the FCC Rules governs the taking of a partial proof-of-performance as presently required by the FCC Rules. A partial proof must be done, for example, when there is a change of directional operating parameters; when changes are made above the base insulator of the antenna (such as the addition or alteration of an FM antenna or transmission line mounted on the tower), when there is an increase of an existing monitor point license value, or when changes in the environment of the array dictate the necessity of demonstrating compliance with the station's instrument of authorization.

The partial proof-of-performance measurements are to be made at the same locations as specified in the reference proof, but measurements do not have to be taken at all locations. The partial proof-of-performance must contain a minimum of ten measurements at points along each of the radials defined in the reference proof. The partial proof must contain an arithmetic or logarithmic analysis of the measurement data so obtained. It must demonstrate that the antenna system is operating within its instrument of authorization. Generally speaking, the measurements should be made at an interval of three to 16 kilometers (two to ten miles). A statement that the impedance of the common point has been measured and is unchanged from the licensed value prior to the making of the measurements should be provided. A change in common point impedance at the operating frequency requires an appropriate submission to the FCC and is to be requested using FCC Form 302.

The partial proof-of-performance in those areas where ground conductivity is not constant should be made under similar environmental conditions as the reference proof. More accurate information can be obtained when the directional partial proof-of-performance is based upon current nondirectional measurements.

GENERAL REQUIREMENTS FOR MEASUREMENTS

All measurements should be made during the daylight hours in the absence of interference, and special temporary authority may be required prior to the commencement of measurements for a new station. For established stations, the FCC Rules permit considerable flexibility in operation during periods of making field intensity measurements on the antenna system.

FCC policy has been that the measurement observation to be recorded and utilized as a basis of analysis of the inverse distance radiation is that observed with the field set oriented towards the station. The field set maximum indication can have an orientation away from the transmitting source. This phenomenon can be affected by many factors, particularly the depth of a null. Other factors vary from local effects surrounding or adjacent to the measuring point, nonuniform conditions inherent in the propagation path, and the position of the monitoring point in a region where there is a sharp change in the pattern.

A record must be kept of the measurement data including the point number and description, the field strength observation, date and time, the pattern under investigation, the name of the individual taking the measurements, the general weather conditions, and the field strength instrument and date of last calibration. A sample form is provided for tabulating the field measurement data.

GRAPHICAL ANALYSIS

The inverse distance field or attenuated field intensity at a reference distance (1 km) is that radiation predicted if the earth were to behave as a perfect conductor. As the wave energy travels away from the antenna, this energy reduces in value. The value of the radiated field is reduced by the inverse proportion to the distance from the antenna. For example, if the value of the unattenuated field at 1 km is 100 mV/m, its value at 2 km will be one-half that value or 50 mV/m, and at 10 km, its value will be one-tenth or 10 mV/m.

FIELD STRENGTH MEASUREMENTS

FOR ____________

DATE: __/__/__

PAGE NO: ________

ENGINEER: ________

METER: ________

N ° E

POINT NO.	POINT DIST.	DATE	TIME	___kW ND MV/M	DATE	TIME	___kW DA-___ MV/M	RATIO DA___/ND	POINT DESCRIPTION

Figure 1. Field strength measurement log.

After the distance from the transmitting antenna to each of the measuring points has been determined and tabulated opposite to the observed field strength values, a measured field strength can be plotted on log-log graph paper. The ordinate is expressed in mV/m and the abscissa is expressed in distance. The FCC, in its conversion to metric, redetermined the frequency curves.

There are 20 frequency charts that encompass the frequencies from 540 kHz to 1700 kHz. Each graph shows the uppermost portion with conductivity curves normalized for 100 mV/m/km from 0.1 km to 50 km and the bottom portion reflects the conductivity curves for 10 km to 5,000 km.

Data can be plotted on groundwave field intensity graph paper available through NAB's Station Services Department (800-368-5644).

For the logarithmic coordinate system, (log-log graph paper) the inverse distance line plots as a straight line. Each of the curves is drawn for the case of an inverse distance field of 100 mV/m at 1 km, and its use is not limited to that value. If an inverse distance field is 200 mV/m (twice the reference value) or 50 mV/m (one-half the reference number) or some other value at 1 km, and if all points on the curve are multiplied by that ratio, this would be the equivalent of moving the curve by that amount on the logarithmic coordinate paper. This is the basis by which measurements are analyzed and the appropriate graph for the frequency involved is made by matching the abscissa of the data with that of the FCC graph. By sliding the ordinate information data vertically, the best fit is obtained. By this method, both the unattenuated field at 1 km and the conductivity value along the radial path have been determined. The use of a light table or window or TV set tuned to an unused channel will serve as a back light for visual analysis of the data.

An individual attempting to analyze measurement data for the first time or not having benefit of supervision can find this a frustrating experience. One approach is to take log-log graph paper for the appropriate frequency (either the regular or expanded scale) and plot the measurement point values normalized to 100 mV/km. For example, if the nondirectional 0.25 kW operation is expected to possess an RMS field at 1 kW of 91 mV/m (70° [0.194 of a wavelength] electrical height tower with a normal ground system—see Figure 8, Section 73.190 of the FCC Rules), it has a field 91/100 less than the FCC log-log conductivity graph. Therefore, multiply all values (divide all values if the expected field is greater than 100 mV/m) of the measurement data by the ratio of 100/91 to normalize it to 100 mV/m. Plot the normalized data. The plotted values can be viewed in relation to the conductivity values if the assumption of the inverse distance field is correct. If the normalized data appears to be over the inverse distance line than the radiation value is higher than assumed and conversely if the normalized data appears abnormally low, the assumed radiation value selected is too high.

This approach can be useful when the nondirectional measurements out to 3 km in the various directions have been taken and a quick evaluation of the conductivity values/radiation efficiency around the site is desired. It also will help to assess whether or not the nondirectional radiation pattern is being influenced by other adjacent towers in the directional antenna system.

AUTHORIZATIONS

Special Temporary Authority

Special temporary authority will be generally required with any operation which has not received prior authorization from the Federal Communications Commission. Section 73.1635 of the FCC Rules provides that special temporary authorization (STA) be requested in writing (an original and two copies) to the FCC. The authorization request should delineate the station frequency, location, a complete description of the proposed operation, and the necessity of the STA. The letter should indicate the capacity of the requesting individual and their phone number so that the FCC can request additional information informally, if required.

For licensed operations, the FCC Rules provide that a nondirectional operation that has been authorized in a proof-of-performance for a daytime only. A station with a single or more than one directional pattern for day and night, etc., can utilize nondirectional power set forth in the latest proof-of-performance without further authorization from the FCC. However, this privilege is permitted only in the commission of field strength measurements. In addition, the FCC permits without further authorization the nighttime pattern's being operated during daytime hours when field intensity measurements are being taken.

Antenna Monitor System Approval

For each authorization providing for a new station, details of the antenna monitor system components and installation must be contained in the proof-of-performance in order to obtain the necessary recognition from the FCC that the antenna monitor system conforms to the FCC Rules. An antenna monitor system request accompanying a partial proof must be, for prompt consideration, in a separate submission. During construction or revision of the monitoring system, special temporary authority for variance of parameters may be required for existing stations. This authority can be obtained by submitting a request (original and two copies) to the FCC. The purpose of the request as well as its duration should be provided. For specific situations, reference should be made to Section 73.68 of the Rules; however, with the revision in Docket 85–90, the FCC is less specific as to the detail of sample-system construction. The FCC indicates as a matter of policy that the procedures methods outlined in the Rules as modified by MM Docket 83-16 would receive continued FCC acceptance. Other less conventional methods may be subject to rigorous scrutiny including observations over a period of time and a partial proof-of-performance. For convenience, Sec-

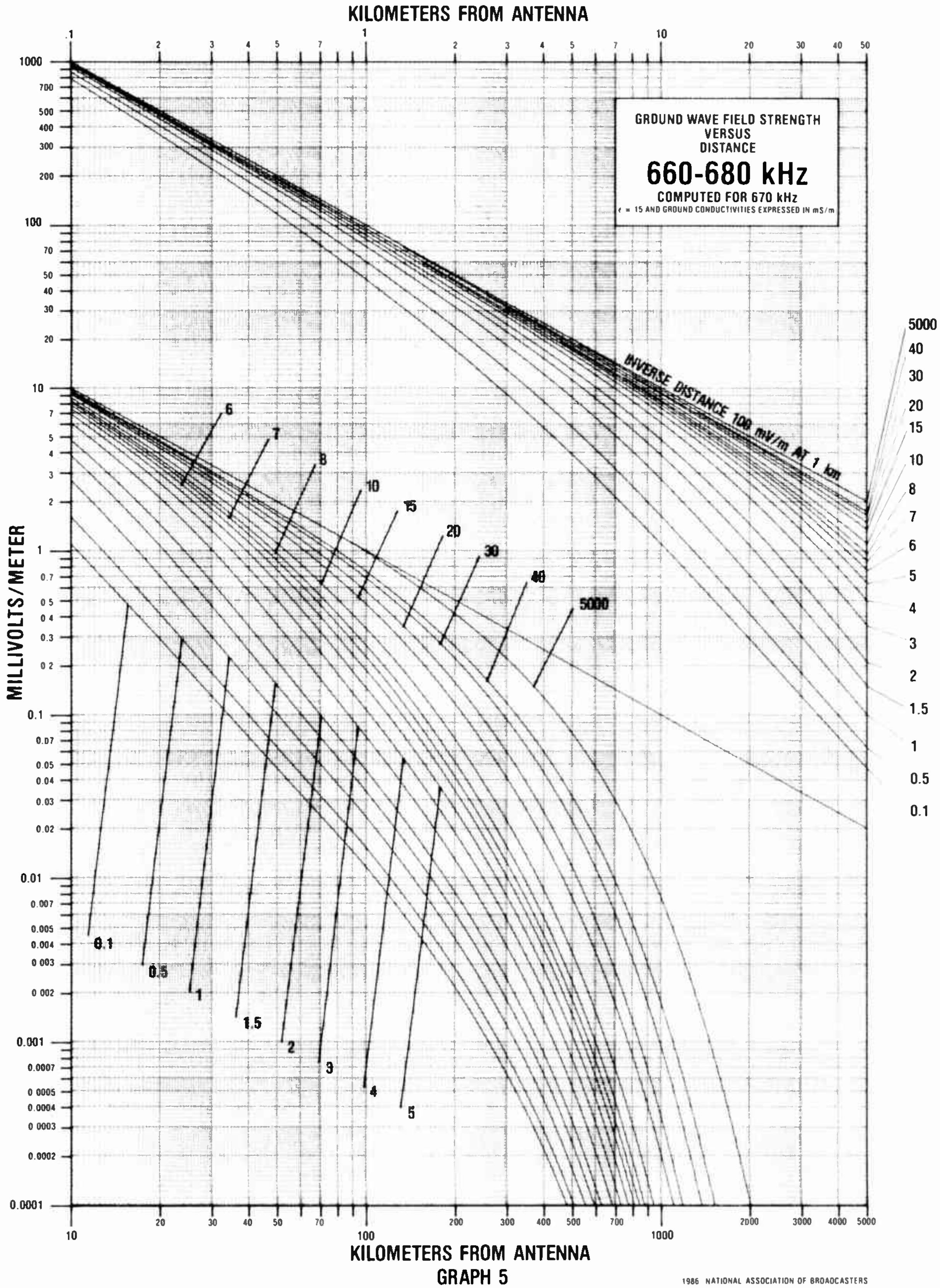


Figure 2. Ground wave field strength versus distance 660-680 kHz. (Graph paper is available from NAB Services, 1-800-368-5644.)

tion 73.68 of the FCC Rules is provided prior to alteration by Docket 85–90.

Whether contained in the proof-of-performance or otherwise, the request for approval must indicate information sufficient to determine compliance with the FCC Rules concerning the method of sample, type of sample line, and its electrical length and whether the sample lines are under similar environmental conditions. The submission should incorporate other descriptions as may be required to demonstrate compliance with the FCC Rules and for FCC Policy, and must be accompanied by a signature of authority from the station organization. A station with an approved sampling system is permitted to establish its own schedule of monitor-point measurements based on its operating conditions.

FCC Form for License

Each proof-of-performance must include the information requested in FCC Form 302. It must comply with the provisions and intent of the FCC Rules including Section 73.186 of these Rules regarding the number of measurement radials, the number of nondirectional and directional measurements made along each radial, the mathematical or graphical analysis, if utilized, the plot of the field versus distance measurements on semi-log or log-log graph paper, and the reproduction of the quadrangle maps (or other maps as required) showing the radials along which measurements were made and locations used in making each of these measurements. Section 73.186 of the FCC Rules also dictates the form and substance concerning the submission of nondirectional antenna resistance as well as the common point impedance measurements with the required graphical plots.

Furthermore, descriptions of the monitoring points complete with the monitor-point photographs and a route diagram as specified by the construction permit must be supplied (see Section 73.158 of the FCC Rules). The station must also comply with other provisions of the construction permit.

A diagram of the RF feed system as constructed (including the phasor, transmission lines, and the tower matching networks) is to be provided. A plot of the inverse distance field at 1 km for the nondirectional as well as each directional mode is to be supplied, based upon the interpretation of the measurement data. The directional pattern must not exceed the authorized pattern in any direction and must have the requisite RMS. In certain situations for new stations or revisions of existing facilities requiring a new reference proof-of-performance, the FCC will permit an adjustment of directional power which can be effected at the time of the license application.

For each directional operation, the parameters as indicated by the antenna monitor for both loop and phase (including SIGN) as well as the base currents and the ratios are to be furnished. Each field strength instrument and its type number, make and model, and date of the last calibration should be listed. If more than one instrument is utilized, a comparison of the accuracy observed for each instrument should be made.

License

The FCC license provides, among other things: the licensee name, specifies the term of the license; the station location; the main studio location, if not at the transmitter or within the boundaries of the principal community; the remote control location; the transmitter location and its coordinates; the type of antenna and ground system, if nondirectional; the frequency; the nominal power; the hours of operation and any special conditions. For stations using a directional antenna system, the second page will provide: a description of the directional antenna system; the spacing; orientation; and height of the towers and a description of the ground system. Also provided are the theoretical and operating specifications determined from the most recent partial or full proof-of-performance. The description of the field strength of the monitor points and the maximum limits that the point must not exceed are contained on the following page(s).

Changes and modifications of any of these items require appropriate notification to and concurrence by the FCC. When receiving a new license, it should be inspected for correctness as compared with that used as a basis for the license application. The operating parameters must be maintained in accordance with Section 73.62 of the FCC Rules and the directional antenna must be maintained with indicated relative amplitude of the antenna base currents and antenna monitor currents within 5% of the values specified in the license, unless other tolerances are specified. In addition, the directional antenna relative phase angles must be maintained within three degrees of the values specified in the license unless other tolerances are required.

Monitor-point values must be maintained within the values specified in the license. An increase in an existing monitor point(s) value can only be accomplished by submission of a partial proof-of-performance to the FCC. A change in monitor point location requires submission of a photograph, route diagram, description of the new monitor point, and a minimum of ten measurements including the newly-designated monitor point between two and ten miles that is shown in the latest reference proof. This information is to be submitted to the FCC.

APPENDIX A

FCC Policy Statement Entitled "Criteria for Approval of Sample Systems for Directional AM Broadcast Stations" Dated December 9, 1985

On October 31, 1985, the Commission adopted a Report and Order in MM Docket 85–90 concerning the antenna sampling systems and proofs-of-performance for directional AM broadcast stations. The new rules are based upon performance standards in terms of accuracy and stability rather than upon construction specifications. This Notice clarifies the information required for directional AM sampling system approval under the new provisions of Section 73.68(a) of the Rules. As before, stations constructing new antenna systems pursuant to a construction permit must obtain approval of their sample system when filing for a covering license. Existing stations may obtain approval by informal request to the FCC in Washington, D.C.

To obtain antenna system approval, applicants may follow either of the procedures set forth in Paragraphs A or B below:

A. Demonstrate that the system complies with the provisions of Section 73.68(a) of the Rules in effect prior to January 1, 1986.

B. Demonstrate stability of operation by submission of the following information:

1. A detailed and complete description of the antenna monitoring system installation.
2. Field strength readings taken on a monthly basis at each of the monitoring points specified in the instrument of authorization for a one year period prior to the date of the application.
3. The following readings taken daily for each directional pattern use during the thirty-day period prior to the filing of the application:
   a. Common point current
   b. Base currents and their calculated ratios
   c. Antenna monitor sample current ratios
   d. Antenna monitor phase readings
   e. Final amplifier DC input voltage and current
4. The results of either a partial proof-of-performance (Section 73.154) or a full proof (Section 73.186) conducted no longer than 3 months prior to the filing of the application and the common point impedance at the operating frequency measured at the time of the proof.

Additional sampling system components and configurations found by the Commission to be accurate and stable over a wide range of environmental and operational conditions will be included as acceptable under Paragraph A above and announced periodically via public notice.

Questions concerning sampling system approval may be directed to John Sadler (202)632–7010 and questions concerning the Report and Order in MM Docket 85–90 may be directed to John Reiser (202)632–9660.

2.11 The Measurement of FM and TV Field Strengths (54 MHz—806 MHz)

Joseph W. Stielper
du Treil, Lundin & Rackley, Inc., Washington, District of Columbia

MEASUREMENT OBJECTIVES

The purpose of this chapter is not merely to provide a cookbook describing the actual taking of television and FM broadcast field strength measurements; most field strength meter instruction books are adequate for this purpose. Taking field strength measurements is relatively easy, but planning a program and analyzing the resulting data is more difficult. That is the emphasis of this chapter.

But first, why take measurements? Television and FM field strengths may be measured to accomplish several objectives. These include: (1) determination of measured coverage contours or coverage of a particular area instead of relying only on predictions of coverage; (2) evaluation of the performance of transmitting systems; (3) measurement of spurious emission; and (4) special studies to evaluate the effect of factors such as terrain and vegetation on field strength. Special studies that require field strength measurements are also used for purposes such as interference reductions and making frequency allocation decisions.

MEASUREMENTS VERSUS PREDICTIONS

It is not necessary or even usual to take measurements to estimate the coverage of a station or even of the stations' coverage of a particular area, its city of license, for example. There are numerous propagation models available that may be used to predict field strength and coverage.[10,11,12,20,22,23] FCC Rules specify the propagation model that must be used to determine coverage for filing with the Commission, although additional showings using alternate models may be made. The FCC Rules also contain measurement and analysis procedures that must be followed if the results of measurements will be submitted to the Commission. For in-house use there are no constraints, and any propagation model or measurement scheme may be used. The engineer is thus faced with the choice between predictions or calculations and measurements. The choice is not always clear. Predictions may be made in the comfort of one's office; there is no concern over equipment availability or weather. The choice hinges on the answers to the questions: (1) which is more accurate or reliable? (2) will the results be filed with the FCC? (3) what are the relative costs?

At a given location, field strength varies with time, which limits the accuracy of measurements. Determination of the long term median field strength is usually desired. The variation may be negligible on short paths but may contribute the greatest uncertainty on long paths, particularly near the radio horizon. Time variability is usually neglected in the analysis of measurements. The effect of time variability is discussed in detail later in this chapter. Also, the effects of time variability are treated more readily with predictions than with measurements. Except for the unusual case of measurements designed to evaluate time variation factors, the effect of weather or climate on measurements should be minimized. Variations of field strength with time are generally greatest near or just beyond the radio horizon. A more detailed discussion of this problem is included in a later section.

In addition to varying with time, field strength varies with location. The variations are caused by factors such as terrain, manmade structures, vegetation, and weather. The effect of each of these on the measurements should be considered when designing the program.

Measurements are not always more accurate or reliable than predictions although the FCC tends to prefer measurements. Thus, in contested cases, measurements are generally desirable. Alternate propagation models often yield conflicting results. The FCC has difficulty resolving this conflict; hence, the preference for measurements.

All propagation models consider the terrain and the environment between the transmitting and receiving antennas in a simplified fashion. In other words, complex terrain and environmental situations involving multiple obstacles and reflections are reduced to much simpler models such as one or two diffracting obstacles. In many cases it is not clear which of the many propagation models is appropriate. Because predictions often yield contradictory answers, there is a tendency to conclude that taking measurements will resolve the problem. However, unless care is taken in the planning and execution of the measurement program, the results may not be an improvement on predictions. Generally, a haphazard measurement program is less useful than a well planned, careful prediction regime.

Many prediction routines do not consider the effect of environmental clutter such as buildings and trees. Such clutter may be more significant than terrain, particularly at UHF frequencies. The effects of clutter are described in a number of references.[8,9,21]

In summary, the methodology of any prediction or measurement program must be critically reviewed before basing decisions on the results.

PLANNING A MEASUREMENT PROGRAM

Before the start of measurements, a plan should be designed considering the objectives of the field strength survey. For example, the constraints imposed by a survey to determine coverage of a station are different and generally less restrictive than those imposed by propagation studies. Design of a program to evaluate the performance of transmitting systems is likely to be the most difficult and require even more careful planning than programs for other purposes. Suggestions for designing measurement programs for each of the above objectives are presented in the following sections, as well as a discussion of the techniques of actual field strength measurement.

There are several basic considerations affecting all types of measurement programs that should be addressed. These include the choice of antenna height for the measurements and allowance for factors such as weather that are beyond the control of the engineer.

The FCC Rules require measurements using a nine meter (30 foot) receiving antenna height. This height should be used unless there are very good reasons for using another height. The bulk of reliable data was taken at 30 feet, and the broadcast coverage contours are based upon the nine meter height. Thus, a nine meter receiving antenna height must be used for filing with the FCC, for direct coverage measurement, or for direct comparison with most other data. Use of a nine meter high antenna raises several practical problems including safety. These are discussed in a later section. For in-house use, another antenna height may be appropriate. For example, to evaluate FM reception by car receivers, a low height should be used. In many areas the height of few, if any, FM or TV receiving antennas approximates the standard nine meter height.

Measurement of Coverage

The coverage of a broadcasting station and the technical quality of the service provided are determined by the quality of the received signal including the field strength. Currently available methods of estimating field strengths within the service ranges of FM and television stations are only approximate, and even the best methods of calculating field strengths often fail to take into account variations due to important local conditions. In general, for making decisions affecting a station's operation, the most useful determination of station coverage is provided by properly made field strength measurements. This section describes measurement programs for measuring field strengths to determine coverage of FM and television broadcast stations.

The quality of service is related to field strength by considerations of receiver sensitivity and noise figure, receiving antenna gain and transmission line loss, and tolerable signal-to-noise ratios. The required fields vary with the class of service and frequency assignment. Table 1 lists the frequencies employed by television and FM broadcast stations. Interfering signals from other transmitters on the same or adjacent channels may limit service to higher values of field strength.

Table 2 lists values of median field strength required for various grades of FM and television service in the absence of interfering signals as established by the Federal Communications Commission's Technical Standards.[1] It also includes revised estimates of the fields required in the television bands to provide acceptable grades of service based on the practical experience of operating stations and the findings of the Television Allocations Study Organization (TASO).[2] The latter has not been officially adopted by the Commission and may well be obsolete. A number of changes in the definition of television coverage grades based upon the TASO studies and other data have been proposed from time to time since the adoption of the present definitions. A review of the rationale for the VHF contour definitions and some potential revisions is presented in the FCC Report FCC/OCE RS 77.01.[3]

TABLE 1
Frequencies Employed for FM and Television Broadcasting.

Service	Frequencies MHZ	Channel Nos.	Channel Bandwidth
TV	54–72	2–4	6 MHz
TV	76–88	5–6	6 MHz
FM	87.9–108	200–300	200 kHz
TV	174–216	7–13	6 MHz
TV	470–806	14–69	6 MHz

TABLE 2
Median Field Strengths Required for Various Grades of Service in the Absence of Interfering Signals.

FM Broadcasting (All Channels)

Grade of Service	μV/m	dBμ[a]
Principal City	3,160	70
Urban	1,000	60

Television Broadcasting
(FCC Technical Standards)

Grade of Service	Ch. 2–6		Ch. 7–13		Ch. 14–83	
	μV/m	dBμ	μV/m	dBμ	μV/m	dBμ
Principal City	5,000	74	7,000	77	10,000	80
Grade A	2,500	68	3,500	71	5,000	74
Grade B	225	47	650	56	1,600	64
(Based on TASO Data)						
Primary	250	48	1,400	63	7,500	75
Secondary	50	34	200	46	630	56
Fringe	20	26	55	35	180	45

Service is defined in Table 2 in terms of the median field strength, with respect to both location and time, at a receiving antenna at a height of nine meters above ground. Thus, as noted above, field strength measurements to be filed with the FCC must be taken using a nine meter receiving antenna height. It may not be necessary or even desirable to employ a nine meter antenna height for measurements that will not be filed with the FCC. For example, if comparative coverage of several stations is desired and receiving antenna heights in the area are less than 30 feet, it may be appropriate to use a lower height. In these frequency bands, field strength usually varies appreciably with antenna height, generally tending to increase with increasing antenna height. However, the variation in field with height may not follow simple laws. Measurements taken at a lower height may be, if necessary, adjusted to reflect the standard nine meter height. Because of the uncertainty of the magnitude of the adjustment, the adjusted results are ordinarily less precise than measurements taken at nine meters. Adjustment factors are discussed more fully in a subsequent section.

The presence of trees, buildings, and terrain irregularities[4-9] often results in considerable variation in field strength from one location to another, even within relatively small areas. The variation in field strength with location must be taken into account in measuring field strength as well as in specifying service. Service is usually defined in terms of the median value of field strength, which is the value exceeded for at least 50% of the time at the best 50% of the receiving locations.

The results of field strength coverage surveys are customarily presented as contour maps, showing lines of constant median field strength which represent the outer limits of various grades of service. A typical map of measured television station coverage is shown in Fig. 1. Methods of preparing contour maps are described in detail under the heading "Analysis of Measurements to Depict Coverage."

Much of the present knowledge of wave propagation in these frequency bands has been derived from field strength coverage surveys on operational FM and television stations. The information gained from these commercial coverage surveys has added to the body of scientific knowledge, but field strength measurement surveys employing special techniques are often needed to supply data for special problems. Examples of such special techniques are discussed under other headings in this article.

A measurement program to determine coverage may be laid out according to the radial route method of the FCC Rules as described for propagation studies. Unless the results will be filed with the FCC, there is no need to maintain the precision required by the Commission. However, following the FCC Rules will minimize the effect of unintended biases that may affect the data. If the program planner is uncertain about the effect of the choice of measurement location, it would be conservative to use the FCC Rules procedure. The precision required by the FCC Rules is primarily intended to insure statistical randomness for propagation studies and to assure that the propagation path is a true radial. For in-house coverage studies, there is generally no need for the measurement location to be exactly on a true radial route at random two mile increments as required by the FCC Rules. Since many FM and TV transmitting sites are located near the center of cities it is often convenient to take measurements along more or less radial roads. Selection of measurement locations need not be random but may be influenced by population density or the desire to obtain data in particular areas.

Field strength is only part of the answer that is usually sought when there are questions concerning coverage. Actual picture or sound quality is only partially correlated with field strength. For example, multipath may cause ghosts in television pictures and distortion in FM, particularly to stereo signals. Such effects can only be determined accurately by subjective observations. Care must be taken to insure that the receiving test equipment and environment represent actual conditions. Observations made with high gain receiving antennas at 30 feet will lead to erroneous conclusions if reception on low gain indoor antennas essentially at ground level predominates. Actual picture or sound recordings should be made rather than only evaluating quality using one of the grading scales. Recordings tend to significantly reduce problems such as variations among observers, repeatability, and inaccurate reporting of the type of impairment.

The following are scoring grades commonly used for visual and aural signals. The visual scale follows the TASO[2] six point system. The aural scale is taken from CCIR Recommendation 262. These first two scales are used for the subjective rating of one aural or visual signal. Comparative scales may also be used where

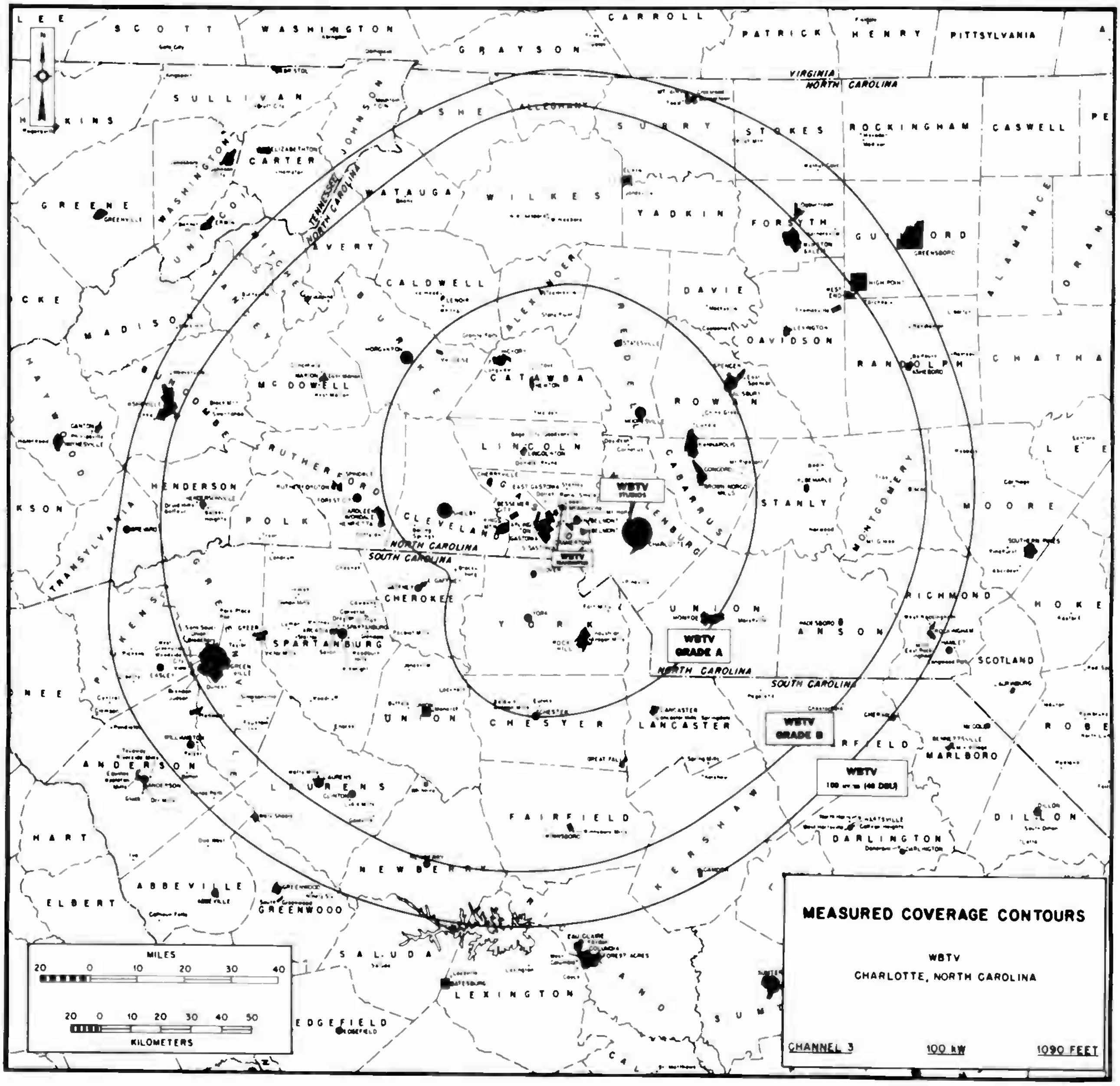


Figure 1. Map showing measured service contours for an operating television station. (Courtesy of Jefferson-Pilot Broadcasting Company)

appropriate. Comparative tests tend to yield data that is more reliable and repeatable than tests on one signal.

TASO PICTURE QUALITY SCALE

Number	Name	Description
1	Excellent	The picture is of extremely high quality—as good as you could desire.
2	Fine	The picture is of high quality, providing enjoyable viewing. Interference is perceptible.
3	Passable	The picture is of acceptable quality. Interference is not objectionable.
4	Marginal	The picture is poor in quality and you wish you could improve it. Interference is somewhat objectionable.
5	Inferior	The picture is very poor but you could watch it. Definitely objectionable interference is present.

6	Unusable	The picture is so bad you could not watch it.

CCIR AURAL SCALE

Quality	Impairment
5 Excellent	5 Imperceptible
4 Good	4 Perceptible but not annoying
3 Fair	3 Slightly annoying
2 Poor	2 Annoying
1 Bad	1 Very annoying

Obviously, with minor editing, the TASO scale may be adapted for aural use. The definitions would aid in improving repeatability in evaluation of the quality of aural signals. The following is the seven grade CCIR aural comparison scale which may also be used for visual signals.

CCIR COMPARISON SCALE

3	Much better
2	Better
1	Slightly better
0	The same
−1	Slightly worse
−2	Worse
−3	Much worse

Analysis of Measurements to Depict Coverage

If a measurement height of other than nine meters was used, the received fields must be adjusted to the field expected at a receiving antenna height of nine meters above ground to obtain the location of standard contours. An antenna height of three meters (10 feet) is often used when the results will not be filed with the FCC. It has been common practice to assume the field strengths to increase linearly with antenna height, as indicated by classical plane earth propagation theory. For this assumption the relationship between the field, E_{30}, which would be expected at 30 feet, and the field E_H, measured at a receiving antenna height H_r, is $E_{30}/E_H = 30/H_r$. For example, the ratio of the field at 30 feet to the field at 10 feet is $30/10 = 3.0$, or 9.5 dB.

The application of the linear height-gain function discussed above is recommended only in relatively flat terrain. In rolling or rough terrain the following height-gain factors were recommended by TASO to convert from three meter to nine meter fields. The values were preliminary and are not based upon measurement programs designed for this purpose.

Channel	Smooth Unobstructed Terrain	Rolling Hilly Terrain	Rough Terrain
2-6	9.5 dB	8 dB	7 dB
7-13	9.5 dB	7 dB	5 dB
14-83	9.5 dB	5 dB	2 dB

Median fields, as established in accordance with a radial or modified radial procedure described above, are plotted as a function of distance from the transmitter, and a smooth curve is drawn through the plotted points. Fig. 2 is a typical graph showing the plotted field strengths as a function of distance from the transmitter, together with the smooth curve through the plotted points. The dashed curve in Fig. 2 is the predicted field strength calculated using the propagation curves and prediction methods specified in the FCC Television Broadcast Technical Standards.[19]

Individual graphs of median field strength versus distance as shown in Fig. 2 are prepared for each of the directions along which the measurements were made; the distances to the desired field strength contours, selected from Table 2, are then plotted on a suitable map, and contours are drawn to produce a finished map such as shown in Fig. 1. In some cases a mountain ridge may cause an abrupt decrease in field strength and the curve-fitting procedure may not be appropriate.

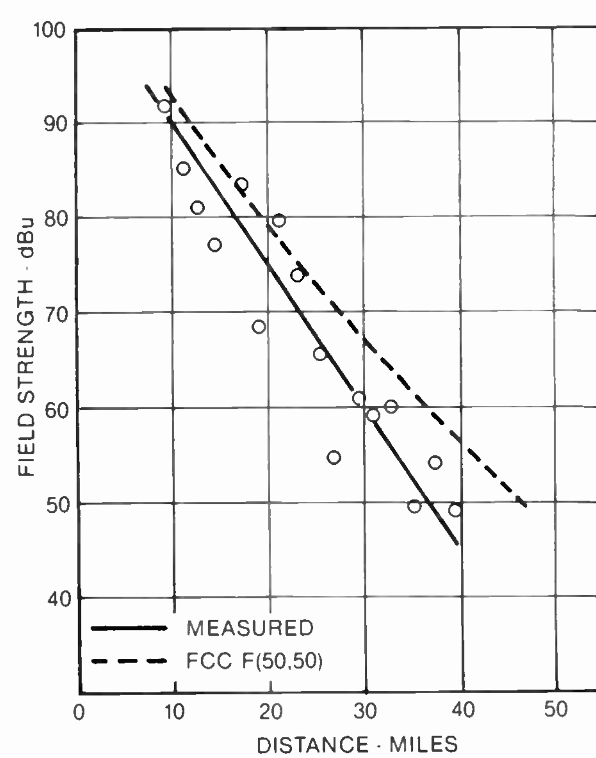


Figure 2. Graph of measured field strength versus distance for a typical radial series of measurements. The results of each mobile run are shown. The solid line is best fit curve through the points. The dashed line is the predicted field strength from the FCC curves.

Transmitting System Evaluation

Field strength measurements are occasionally used to assess the performance of a station's transmitting system, particularly the operation of the antenna. The difficulty in this case is that of separating the effect of propagation factors from the effect of the transmission system. Ideally, comparison of transmission from the antenna to be tested should be made with that from a standard antenna. The standard antenna would be a low-gain antenna of known characteristics mounted near the antenna under test. This procedure is described in detail in the TASO report.[2] The procedure is rarely feasible.

Measurements on two or more stations operating in the same frequency band are often used to determine if one is not performing properly. Differences in facilities and location must be considered in the analysis of the data. Vertical plane pattern effects may be minimized by avoiding locations near any of the stations. The uncertainty in the results is proportional to the degree of the adjustments that are made. Numerous measurements should be taken to reduce the uncertainty.

Comparisons of measured data with standard propagation curves such as the FCC Rules curves or smooth earth curves must be used with caution. The standard deviation of measured data used to draw the FCC curves is 7.7 dB at low VHF, 6.8 dB at high VHF, and 9.3 dB at UHF after adjustment for terrain.[10] Even in smooth terrain field strength is affected by other factors such as trees, buildings, and atmospheric variations.

In hilly terrain, the effect of clutter, terrain, and atmospheric variations can be minimized by selecting sites near the transmitting antenna to minimize atmospheric effects and where ray path clearance over all obstacles can be obtained. This clearance should be at least 0.6 Fresnel zone. The radius of a full Fresnel zone is given by:

$$H_O = 2280 \sqrt{\frac{d_1 d_2}{f(d_1 + d_2)}} \qquad [1]$$

where:
- H_O = Fresnel zone radius in feet
- d_1 = Distance from transmitting antenna to obstacle in miles
- d_2 = Distance from receiving antenna to obstacle in miles
- f = Transmitting frequency in MHz

Application of this criterion is illustrated by Fig. 3.

Even if such ray clearance over all obstacles can be obtained, reflections from terrain may often modify field strength. Allowance for reflections can only be achieved by taking measurements at a large number of locations to "average out" reflection effects.

Field strength measurements intended to evaluate system performance are sometimes taken in an aircraft in an attempt to reduce terrain effects. However, there are several problems associated with aircraft measurements, including: (1) Determination of the actual location of the aircraft relative to the transmitting system under test, (2) constraints on the design of the program imposed by the aircraft's performance, (3) mounting of the receiving antenna on the aircraft to minimize the effect of the airframe on the receiving antenna pattern without creating a hazard, and (4) uncertainties, such as actual path loss, that limit the accuracy of the measurement of absolute gain of the transmitting system under test.

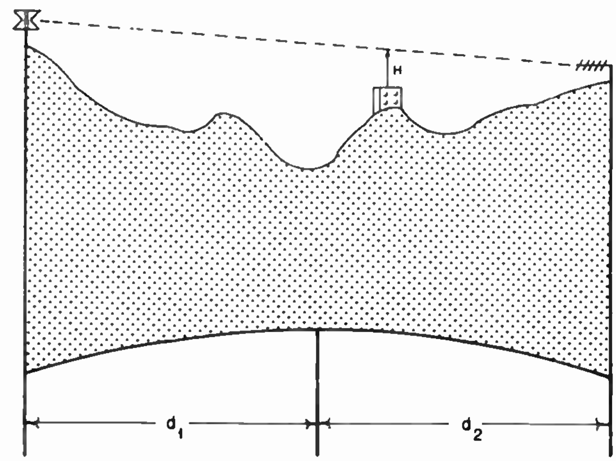


Figure 3. Illustration of required ray clearance to avoid obstacle loss.

The methods for determining the location of the aircraft or maintaining the desired location vary from simple manual techniques to control using electronic location measures such as loran. In the most basic case, an altimeter is used to maintain the desired height. A more or less circular route can be flown by flying over known ground locations. At the other end of the scale a completely electronic system can be used. Such systems are used in Europe because proof measurements are routinely required after construction.

Reflections are a limiting factor in antenna measurement whether measurements are taken on a test range, on the ground, or on an aircraft. In the latter case the standard procedure is to fly close to the antenna under test so that the reflecting ground will likely be illuminated by a relatively low ERP. The distance should not be selected arbitrarily, but should be based upon the vertical pattern of the antenna. Ideally, a helicopter should first be used to measure the antenna's vertical plane pattern at a number of locations to select the radius of the horizontal plane flight. If this is not feasible, a range measured or theoretical pattern should be used. If possible, the radius should be selected so that the depression angle of a ground-reflected ray falls in a null of the vertical plane pattern. For the case of a well smoothed vertical pattern, the flight radius may not be optimized, but the engineer should estimate the uncertainty so that excessive validity is not attached to the data. Error estimates may be made based upon theoretical reflection coefficients found in NBS Technical Note No. 101.[12]

If measurements are taken on vertically-polarized signals, care should be exercised to avoid flying at distances which permit a ground-reflected ray to fall at or near the Brewster angle. At this angle a vertically-polarized wave changes phase. Such a phase change will contribute to the uncertainty of the data. Even if Brewster angle problems are avoided and an aircraft flight path is carefully controlled, ground reflections will likely cause apparent but unreal variations in measured antenna patterns. For example, at an aircraft and antenna height of 300 meters (984 feet) and a flight radius of 2 kilometers (1.2 miles), a change of 10 meters in ground elevation at a reflecting point changes the reflected ray path length by approximately 5 meters. This is more than one half wavelength at all FM and TV broadcasting frequencies. Thus, relatively small changes in reflecting point elevation can completely change the phase of a reflected ray. In this example, if the magnitude of a reflected ray is only one-tenth of that of a direct ray, the expected variation is approximately ±0.9 dB.

Perhaps the most difficult aspect of aircraft measurements is the accurate determination of the gain of the receiving antenna as mounted on the aircraft. The aircraft and the antenna transmission line can change antenna gain significantly from the nominal "free space" value. Unless receiving antenna gain is accurately known, only relative antenna patterns can be measured. The best method would be to measure receiving antenna gain by measurements on a transmitting installation of precisely known ERP. Because of the effect of ground reflections, measurements at varying distances should be taken to average out these effects.

Analysis of Measurements to Evaluate Antenna System Performance

The following description applies to data taken on land at locations where at least 0.6 Fresnel zone clearance obtains over all obstacles. Data from aircraft measurements may be analyzed in a similar fashion. For aircraft measurements, essentially continuous data is analyzed; in ground measurements, discrete data is used.

If a station uses an omnidirectional antenna, the data may first be analyzed by neglecting noncircularity or adjusting measured data for an assumed horizontal pattern. Relative field strength may then be determined assuming free space propagation for comparison with the antenna's theoretical or range measured vertical plane pattern. The result of such an analysis is shown in Fig. 4. Data may also be analyzed by comparing measured to predicted field strengths. If the antenna is performing properly the median difference should be small and approach the accuracy of the test procedure. The median difference between measured and predicted strength for the data shown in Fig. 4 is 1.2 dB; that is, measured field strength exceeded predicted field strength.

A comparison of measured and predicted field strengths may also be used to analyze comparative measurements on different stations. Measured data may be compared with predicted values using the Commission's curves. Adjustments to allow for fairly small differences in height and location can be made directly to the measured ratio of the field strength. In decibels, a factor of 20(log(height ratio)) can be used. For example, to adjust for a height difference between 800 and 1,000 feet, a factor of 1.9 dB is added to the lower height data. For distance adjustment, a factor of 40(log(distance ratio)) may be used. In this case (for example for distances of 9 and 10 miles), 1.8 dB would be added to the data at 10 miles. By using this technique it is not necessary to calculate predicted field strength.

PROPAGATION AND OTHER SPECIAL STUDIES

Sections 73.314 and 73.686 of the FCC Rules describe the procedures to be employed for field strength measurements used in propagation studies. This procedure is basically the technique developed by TASO. The actual field strength measurement procedure is discussed later in this article. The FCC procedure is intended primarily to yield data that may be analyzed to study the effect of terrain and other local influences on field strength. Measurement programs intended for other purposes may require substantially different plans. For example, measurements primarily intended to determine the effects of clutter, such as trees and buildings, should be taken if possible, in smooth terrain to eliminate the effect of terrain.

Measured field strength for propagation studies may be compared to calculated field strength for a number of propagation models.[10-13]

Measurement programs have been conducted to evaluate field strength time variation factors. In this case field strength measurements are recorded at fixed locations over a period of time.

Field strength measurements may be an integral part of special field tests, particularly with regard to interference allocations and changes in FM and TV operation. Examples of such field tests in the past include those on VHF TV-Land Mobile channel sharing, tests relating to the problem of educational FM interference to channel 6 and on circular polarization for television. Extensive measurements were made in southern California in the study of propagation between Santa Barbara and San Diego with regard to the proposed allotment of television channel 10 to Santa Barbara. Detailed discussion of special projects is beyond the scope of this article; however, the basic techniques of field strength measurement discussed herein are valid for special programs. In addition, many of the topics discussed in planning coverage, transmission equipment evaluation and pure propagation tests are appropriate for use in special tests.

Measurement of Spurious Emission

It is occasionally necessary to measure spurious radiated field strength from FM and TV broadcast

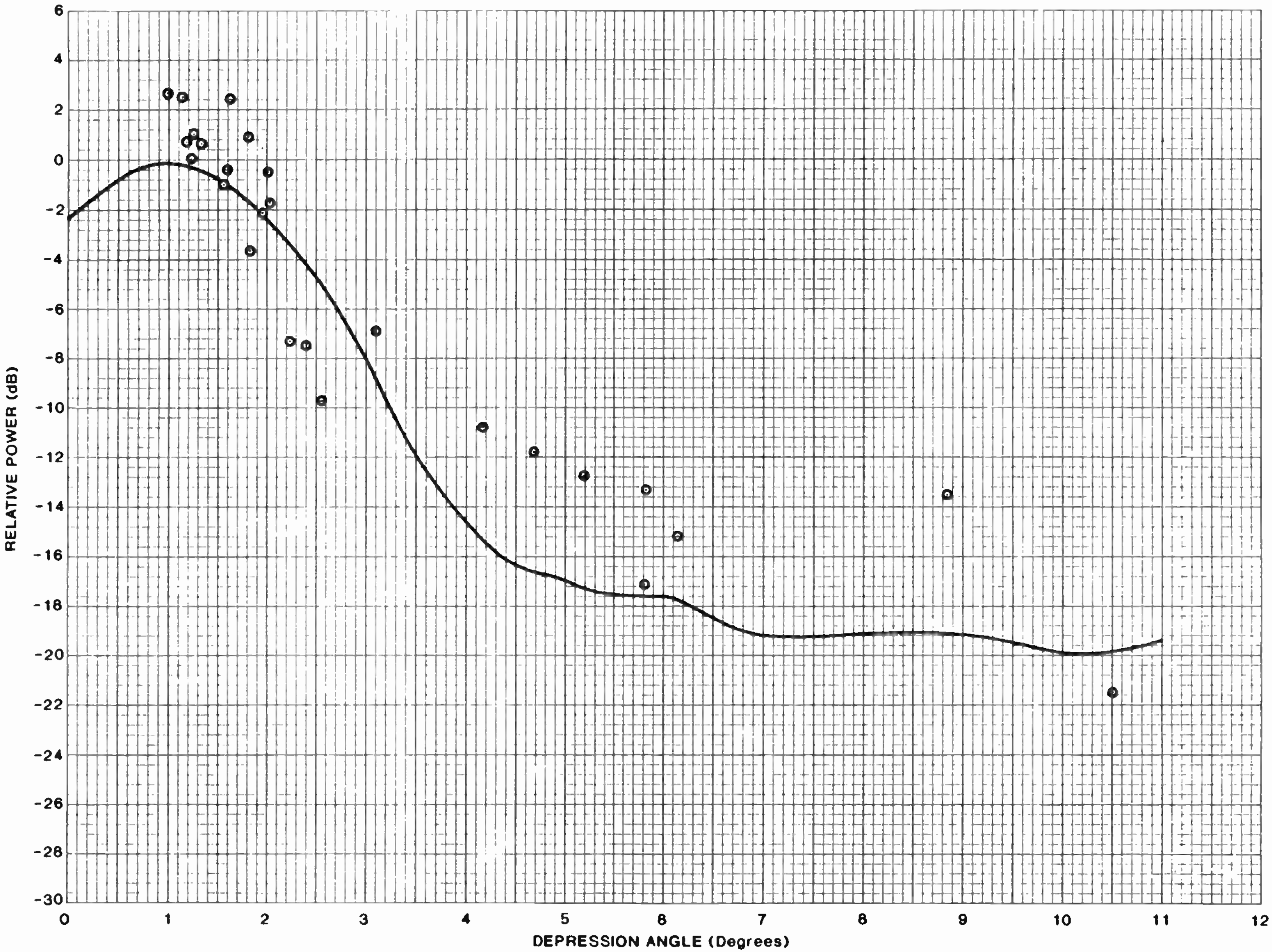


Figure 4. Comparison of data based upon field strength measurements with antenna vertical plane pattern.

facilities to show compliance with FCC Rules after installation. Requirements regarding the level of spurious signals relative to carrier level is specified in Sections 73.317 and 73.687 of the FCC Rules for FM and TV transmitting systems, respectively. Such field strength measurements are nominally regulated by Section 2.993 of the Commission's Rules.

This section is quite vague and it is recommended that the person intending to take measurements determine acceptable procedures from FCC Laboratory personnel before undertaking the measurements.

BASIC EQUIPMENT PRINCIPLES

Field strengths in the VHF and UHF bands (30 to 3000 MHz) are ordinarily measured by determining the voltage which the field induces in a half-wave dipole. The basic relationships can be expressed in several forms. The power transferred between two half-wave dipoles in free space separated by a distance, d, is given by

$$P_r/P_t = 2(1.64)(\text{£}/4\pi d)^2 \qquad [2]$$

where: P_r = Received power
P_t = Transmitted power
£ = Wavelength in same units as d

In terms of the field at the receiving dipole, the power delivered to a matched load by a half-wave dipole in a field of E volts/m is

$$P_r = (0.0186E\text{£})^2 \text{ watts} \qquad [3a]$$

where £ is expressed in meters.

Or, alternatively, in dB relative to one watt (dBW)

$$P_r = F - 20\log f - 105.1 \qquad [3b]$$

where f is the frequency in MHz and the field strength, F, is in dBμ.

For a resistive load of R ohms, the voltage V developed across a matched load by a dipole in a field E is

$$V = (E\text{£}\sqrt{R})/53.7 \text{ [volts]} \qquad [4]$$

The fundamental problem presented, therefore, is that of measuring the developed RF voltage by a practical instrument of acceptable accuracy.

The voltage measuring device is ordinarily separated from the antenna by a length of cable. The cable may introduce losses, and any impedance mismatch must be sufficiently small that calibration errors are not introduced by differences between the antenna and cable impedance and the internal impedance of the calibrating oscillator.

The following paragraphs describe the basic components used for field strength measurement using the standard FCC methods. The system described is the conventional system using chart recorders and manual analysis of recorded data. Digital sampling may be used to replace, entirely or at least in part, manual techniques. At this time an engineer, particularly a station engineer, conducting field strength measurements is most likely to use conventional equipment for reasons of cost and availability. Therefore, the following paragraphs will focus on conventional equipment. However, the principles of the measurement procedure also apply to digital equipment and a discussion of conventional equipment will illustrate these principles more clearly.

Practical Field Strength Meters

Field strength meters are calibrated receivers that fall into three basic categories: (1) receivers that contain a precision oscillator and attenuator and use direct comparison between meter readings produced by the received power and the output of the oscillator, (2) receivers that use a precision oscillator to adjust receiver gain to produce a direct reading meter, and (3) receivers that are direct reading but do not contain a calibrating oscillator. Field strength meters of Type 3 are not considered suitable for precision field strength measurements. In addition, spectrum analyzers may be used for field strength measurements, the use of such instruments is discussed below.

Fig. 5 is a block diagram of a practical Type 1 field strength meter. Type 2 meters are more common now; however, the purpose off the following discussion is to illustrate measurement principles. The Type 1 meter is more suitable for this purpose. The antenna delivers its received power to a transmission line leading to the receiver input. If the receiver input is unbalanced to ground, a balance-to-unbalance transformer (balun) is required. The transmission line between the antenna and the receiver is shielded to avoid stray pickup.

The RF attenuator shown serves two purposes: to avoid overloading of the receiver input on strong signals and to improve the impedance match when the receiver input impedance is substantially different from the characteristic impedance of the transmission line. It is frequently omitted when not required for either of these purposes.

The signal at the receiver input is amplified and converted to the intermediate frequency. Amplification and attenuation at the intermediate frequency permit operation over a wide range of field strengths; further range is provided by the receiver gain control. The rectified receiver output operates the indicating meter.

In operation, the attenuators and gain control are adjusted to provide an on-scale reading of the indicating meter. The receiver input is then switched between the output of the transmission line and the output of the calibrating oscillator, which is tuned to the frequency being measured. The output of the calibrating oscillator is adjusted to a predetermined fixed value using the RF power monitor, and the calibrated attenuator is adjusted until the indicating meter deflection is the same as that obtained from the antenna and transmission line.

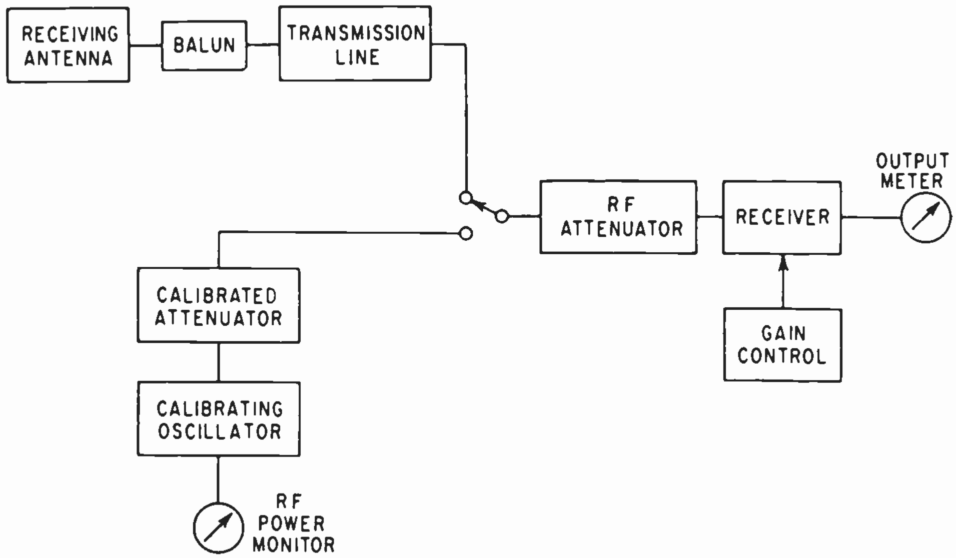


Figure 5. Block diagram of practical field strength meter.

For this condition, the voltage at the output of the calibrated attenuator is the same as that from the antenna and transmission line. By taking line and balun losses into account and applying Eq. 4 above, the field at the antenna required to produce this voltage can be determined. The relationship between field strength and receiver input voltage is usually expressed as $E = KV$, where K is a function of frequency. Fig. 6 is a typical graph showing values of K for a VHF field strength meter.

A typical commercial field strength meter of professional quality is shown in Fig. 7. The instrument shown is a Potomac Instruments type FIM-71, covering the VHF, FM, and television band from 54 to 216 MHz. A companion instrument, similar in appearance, covers the UHF television band from 470 to 806 MHz.

Accurate instrument calibration is essential in measuring RF fields. During use, the calibration of the instrument described is provided by the calibrating RF voltage source, which is usually an integral part of the field strength meter (see Fig. 7). The calibration of the oscillator and the calibration of the instrument as a whole must in turn be established and maintained by reference to laboratory standards.

The most direct laboratory calibration of the complete field strength meter is established by generating a known standard field in which the receiving antenna is placed. Standard field ranges have been developed and constructed at both UHF and VHF[14-15] and are sometimes used in primary calibration of field strength meters. Most commercial laboratory calibrations, however, are made by removing the dipole elements from the standard antenna and applying a known RF voltage at the proper frequency to the dipole terminals in series with an impedance equal to the receiving-antenna impedance. The calibration of the balun, line, and receiver is established in terms of this applied voltage, which is then related to field strength through Eq. 4.

The calibration of the internal reference oscillator section includes the calibration of both the oscillator proper and the variable-output attenuator, if employed. The attenuator is usually of the inductively coupled piston type,[16] which depends only on its dimensions for proper functioning. This can be checked against the correct dimensions or against a laboratory standard attenuator. The oscillator can be compared with a standard oscillator, or its output can be measured with a laboratory standard such as a bolometer bridge.[17] This calibration is normally, but not necessarily, performed by the manufacturer.

If measurements are made on the visual carrier of a television station, the difference between the peak and average powers of the transmission must be taken into account. This can be done by establishing a calibration in terms of average power for a still scene (such as test pattern or black picture), or a peak-reading voltmeter can be employed to indicate the level of the synchronizing peaks. Such peak-reading voltmeters are an integral part of professional commercial field strength meters such as the one illustrated in Fig. 7.

Spectrum analyzers may be used to replace standard field strength meters to permit the simultaneous measurement of multiple signals. The multiple signals must generally be arriving from a common direction, for example from the same transmitting site. Otherwise, an omnidirectional receiving antenna must be used. In the case of television measurements, care must be exercised to insure that peak of sync is measured. The total power of an FM signal is constant, so the level of an FM carrier diminishes under modulation. Thus, to obtain accurate measurements (particularly mobile measurements) on FM signals, the spectrum analyzer

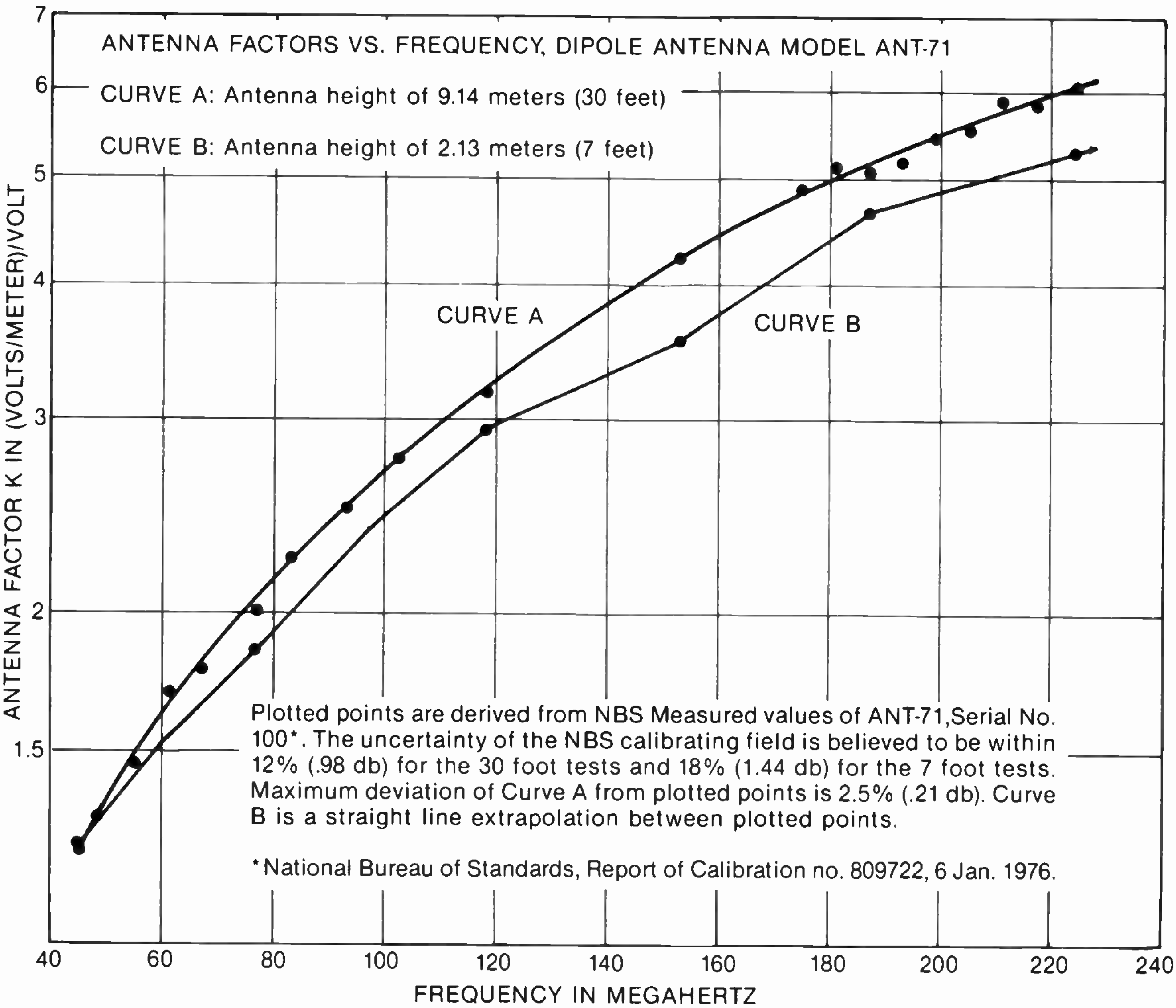


Figure 6. Graph of K for a typical VHF field strength meter. (Courtesy of Potomac Instruments, Inc.)

must be adjusted to "see" the bandwidth of each signal. In addition, intermodulation signals may be generated within the spectrum analyzer itself. If there is any doubt, the validity of signals should be checked.

In addition to the field strength meter, several accessory items are needed when making a field strength survey. The principal items and their use are described in the following paragraphs and include (1) a special receiving antenna, (2) an antenna supporting mast, (3) a chart or digital recorder, and (4) power supplies. Fig. 8 shows the field strength meter, chart recorder, and some additional equipment for a survey. The size and weight of the equipment usually dictate that it be mounted in an automobile or light truck. As discussed in Appendix A, all of this equipment, including the field strength meter, should be grounded to the vehicle frame. Fig. 9 shows a van with elevated mast supporting a UHF antenna and containing the mounted equipment of Fig. 8.

In addition to taking field strength measurements, it is often desired to use other equipment such as the television monitor shown in Fig. 8. A vehicle devoted to general field test programs should have sufficient space available for special equipment including monitors, wave-form scopes, and magnetic recorders.

Receiving Antennas

The measurement survey can be made by employing the standard dipole antenna furnished with the field strength meter, or other antennas can be utilized. It is often desirable to use an antenna other than the standard antenna for actual measurements. Standard antennas are usually not mechanically sturdy. If measurements are to be taken at a large number of

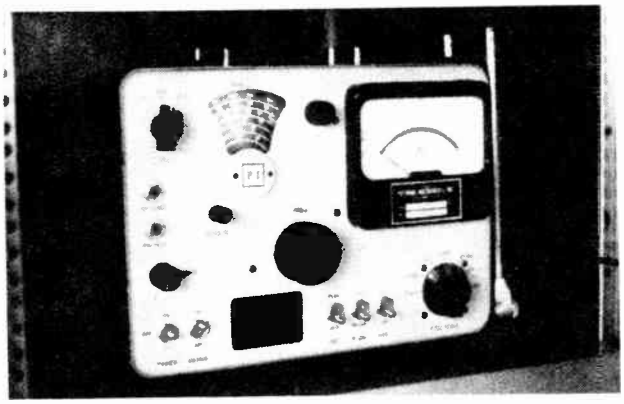


Figure 7. A VHF meter of professional quality.

locations, especially with extensive driving between locations, it is desirable to use a rugged or readily replaceable antenna and thus minimize the probability of damaging the standard antenna. For subjective evaluations the antenna should be representative of antennas actually used in the area by the public.

An antenna which is essentially omnidirectional in the horizontal plane does not require orientation as the vehicle is moved; however, it is generally desirable to use a receiving antenna with directivity. Directional receiving antennas possess gain which is useful principally for UHF measurements, and their directivity is useful to eliminate unwanted interfering signals. Directional receiving antennas also tend to discriminate against reflected signals, thus reducing apparent variation in measured field strength.

The antenna employed for the measurements must be calibrated on the measurement vehicle because of ground- and vehicle-proximity effects.[18] The difference in the calibration curves of Fig. 6 is caused by ground-proximity effects. The received field is first measured using the standard dipole antenna mounted on the vehicle at the measurement height. The antenna to be used in making the survey is then mounted on the vehicle, at the height to be employed in making the

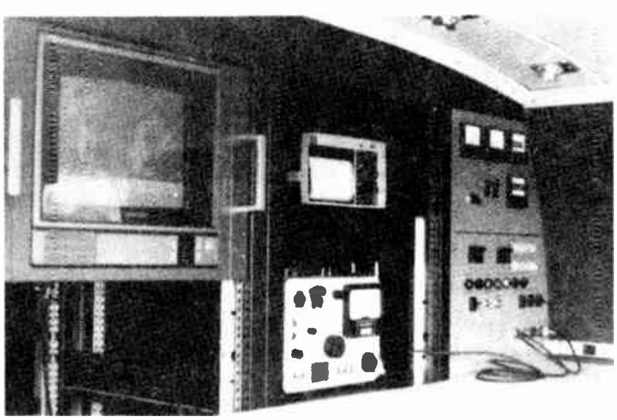


Figure 8. Equipment set up for field strength measurement.

Figure 9. Van equipped for making field strength measurements with 30 foot antenna height.

survey, and the receiver input voltage is determined with the receiving antenna at the same spot in the field. If an omnidirectional receiving antenna is employed, the circularity of the pattern of the antenna as mounted on the vehicle must be determined. The antenna pattern is best established by rotating the vehicle with the antenna mounted as for measurements, and recording field strength and measuring antenna gain as above. If the vehicle cannot be rotated, an alternate comparable procedure must be followed.

The gain of a directional service antenna can be established relative to the dipole antenna by means of measurements with the antennas stationary, but more consistent results are often obtained by making short mobile runs over identical paths and recording the signals from the two antennas. In either case a location essentially free of standing waves should be used. For either procedure, the voltage gain of the service antenna G_s relative to the standard dipole antenna G_d is $G_s/G_d = V_s/V_d$, where V_s and V_d are the voltages delivered to the receiver input using the service and standard dipole antennas, respectively.

If the transmission line or balun between the antenna and receiver is different from the standard cable and balun supplied with the instrument, the antenna calibration must include the measurement system cables and baluns.

Horizontally polarized receiving antennas are normally used for field strength measurements. At the present time, the FCC Rules require that FM and TV stations operate with horizontal only, or dual or circular polarization. Exceptions are translators and noncommercial educational FM stations. The latter may operate with only vertical polarization when there is potential for interference to a channel 6 television station. If measurements on a vertically polarized field are desired, precautions must be taken to reduce the effect of coupling between the vertically polarized antenna and the supporting mast and vertical run of transmission line. In the case of horizontal polarization, the transmission line and mast are at right angles to the receiving antenna and thus exert negligible effect on reception. For vertical polarization the line and mast are parallel and generally close to antenna elements. The receiving antenna pattern and gain may thus differ significantly from those of free space. The standard antenna furnished with the field strength meter is used for measurement or calibration of another antenna, preferably directional. The measurement antenna may then be mounted close to the mast and the gain is determined compared to the standard antenna which is offset from the mast. The standard antenna should be mounted at least one and preferably two wavelengths from all metallic vertical components. This requirement presents a mechanical problem at the lower VHF frequencies. This measurement antenna is then calibrated against the standard antenna cantilevered out from the mast about two wavelengths. Thus, in the vicinity of the standard antenna the transmission line is at right angles to the antenna. Since the mast and vertical run of transmission line are relatively distant from the antenna, the calculated variation in gain is only about ± 0.5 dB when the standard antenna is cantilevered out by two wavelengths. The result is that the effect of the mast and transmission line on the measurement antenna is compensated for in the calibration.

Antenna Supporting Mast

The receiving antenna is ordinarily supported at a height of two to nine meters (six to 30 feet) above ground, depending on the measuring technique employed. For the short heights, a simple mast of metal tubing can be used. For the standard nine meter (30 foot) height, a special mast is normally used to raise and lower the antenna, and the mast arrangement should permit the vehicle to move over limited distances with the mast elevated.

The measuring unit shown in Fig. 9 employs a telescoping mast constructed of aluminum tubing elevated by compressed air or nitrogen; the mast descends under gravity when the pressure is relieved. A handle inside the vehicle permits the mast to be rotated to orient the receiving antenna. Electronic newsgathering (ENG) vans are normally equipped with a telescoping mast. Such masts may have tilt heads that may be used for antenna orientation.

Operation with an elevated antenna involves a number of safety hazards posed by the measurement vehicle. Power lines are the principal overhead obstruction of concern. The avoidance of these hazards, equipment grounding, and operation after encountering a hazard are discussed in Appendix A. These safety procedures were taken from the TASO Report[2] with minor modifications that update the procedures and description.

Recorder

For measurements made with the vehicle in motion, a chart or digital recorder is employed. A chart recorder can be driven from the vehicle speedometer or a clock drive motor. Excitation of the recorder is provided by a dc amplifier, which usually is built into the field strength meter or may be a separate accessory.

When the chart recorder is employed, the recorder pen element must be calibrated against the receiver output indicator of the field strength meter. The dc recorder is adjusted for balance at the ends of the meter scale, and a calibration curve is prepared for intermediate values.

Instead of or in addition to a chart recorder, an analog-to-digital converter can be used to sample field strength to permit computerized analysis of the measured data. A relatively simple technique is to use a converter from a radiation hazard power density meter. Portable laptop computers can be used to store and process the measurement data.

Power Supplies

The power drain of the measuring equipment can be fairly substantial, especially if much accessory equipment is employed. It is usually preferable to provide a power source for the measuring equipment separate from the vehicle battery. This may consist of a separate battery bank to operate the meter and accessories, or a separate 115V ac alternator may be mounted in the vehicle. The latter is employed to operate the ancillary equipment shown in Fig. 8.

MEASURING PROCEDURES AND TECHNIQUES

The FCC FM and TV Technical Standards prescribe measuring methods to be employed in making measurements to be submitted to the Commission. These or similar methods are also usually employed in making other surveys such as for the measurements on station coverage. Variations from the official procedure are frequently taken; some of these variations are discussed under earlier headings. The following paragraphs summarize the present requirements of the Commission's Standards.

FCC Standard Method For the Collection of Propagation Data

The following discussion summarizes the FCC procedure in making field strength measurements for propagation studies. The forms for recording and submission of data and other technical requirements are presented in Section 73.314(b) and 73.686(b) of the Rules for FM and TV respectively.

The Commission's Technical Standards require field strength measurement surveys to be made with mobile equipment along at least eight radial lines from the transmitter. The radials need not be laid out along bearings separated by 45°, beginning with true North as is standard for contour prediction. Measurements are required to be taken from 16 kilometers (10 miles) in increments of 3 kilometers (2.0 miles) in each direction. If it is desired to establish contour location, the distance should extend somewhat beyond the field strength contour which the engineer desires to establish. The routes are selected to encounter representative terrain and to permit interpolation between adjoining radials. A precise radial line is laid out from the transmitter on topographic maps to the distance to which measurements are to be made. Along this radial line, measuring locations are marked at exact 3-kilometer (2-mile) intervals, beginning at exactly 16 kilometers (10 miles) from the transmitting antenna. The actual measurements are made precisely on the radial, at locations as close as possible to the exact 3-kilometer (2-mile) marks established as described. The ground elevation of the actual measurement location should be the same as that of the intended location.

FCC Standard Method For Measurement of Coverage of Communities

Sections 73.314(c) and 73.686(c) of the FCC Rules describe the procedures for the measurement of service to specific communities for FM and TV stations, respectively. These rules outline a measurement pattern that is in the form of a rectangular geographic grid overlying a map of the community. Measurements are made at locations as close as possible to the intersecting points on the grid.

The number of measurement points must be at least 15 or $0.1\sqrt{P}$ (whichever is greater) where P is the population of the community in thousands. Additional requirements describing documentation and calibration are contained in the rules.

The Rules also contain a statistical procedure to analyze field strength. This method fits the measurement data to a normal distribution and yields the median or average field strength in a community. The analysis may be graphical or numerical. This result is not completely compatible with other FCC Rules. The principal city coverage rule, for example, is based upon the determination of the location of a coverage contour. This difficulty can be eliminated by using a grid that is not truly rectangular but consists of radials and orthogonal arcs. This plan permits both radial and grid analysis. Since the resulting grid is not perfectly rectangular, use of this plan should be cleared with appropriate FCC personnel before undertaking measurements when the results will be filed with the Commission.

The individual measurements consist of short mobile runs (at least 30 meters [100 feet] along the road) at each location so chosen, with the receiving antenna at the nine meter (30 foot) height. If measurements are made on multiple stations for comparison purposes, it is desirable to mark the beginning or end points for each run to insure that the run is made over identical paths on each channel. Before making the measurement run, the gain of the field strength meter is adjusted to the meter reading and chart recorder indicating the initial calibration.

The chart or other recorder is used to record the field strength meter output; and the median, minimum, and maximum values of the field for each recording are determined from the chart recording. When using a chart recorder precise determination of the median is usually made after completion of the survey; however, it is useful to use the meter's calibrating oscillator to make a trace of an estimated median value on the chart paper after each run. The median of the run is estimated and the output of a precision attenuator may be adjusted to the estimated median. This provides an additional calibration check and permits a rapid preliminary calculation of field strength that is often useful to monitor the progress of a survey.

Fig. 10 shows a sample of a typical chart recording obtained by this method. With field strength meters such as the one shown in Fig. 7, there is no built-in adjustable attenuator; however, there is an oscillator output which may be fed through an external attenuator to the input to achieve this calibration.

The chart median values are then converted to received field strength by combining the individual

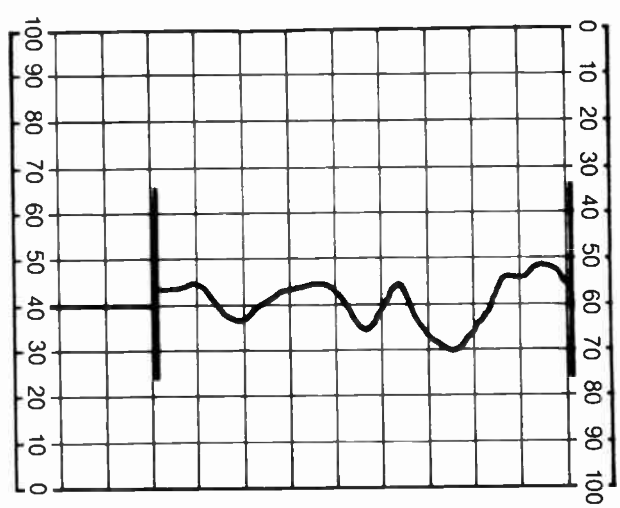


Figure 10. Sample of field strength chart recording of a short mobile run showing traces marking beginning, ending and calibration.

calibrations of the antenna, transmission line, field strength meter, dc amplifier, and chart recorder as discussed above.

PRACTICAL PROBLEMS ENCOUNTERED IN MAKING FIELD STRENGTH SURVEYS

Before any field strength measurement survey is undertaken, the radiated power of the transmitting installation must be established as closely as possible. The transmitter output power should be determined by means of the dummy load and maintained as closely as possible to the proper value throughout the survey. The radiated power is established from the measured transmitter output power, taking into account the antenna power gain and the transmission line and diplexer losses.

The use of a nine meter (30 foot) receiving antenna mounted on a vehicle requires special permission from police or highway authorities in most states. These requirements vary among the individual states, but full details can be obtained from the state police or highway headquarters in the various state capitals.

The operation of such a nine meter (30 foot) mast presents safety hazards which require the exercise of utmost caution. In addition to proper grounding (discussed earlier) the TASO field strength measuring specification includes a special appendix (included here as Appendix A) dealing with overall safety requirements. When measurements are made with an elevated antenna, the need for caution must be borne in mind at all times.

Time Variation of Field Strength

Fairly substantial variations in field strength with time are frequently noted, particularly near and beyond the radio horizon, although significant variations occur at other distances. (Factors for time variation are included in several references.[10-12]) These variations may be relatively rapid, occurring over a period of a few minutes, or slow variations may appear over periods of several hours. There are also long term seasonal variations. Average field strengths in this region are usually lowest during winter afternoons, and higher average fields may be observed during the evening hours and during summer. The variations in field strength with the passage of time must be taken into account in planning and making field strength coverage surveys.

The observed fluctuation of the field near the horizon is believed to be due principally to variations in the refractivity gradient of the lower atmosphere, which in turn is determined by the temperature, humidity, and barometric pressure gradients. Measurements for coverage surveys should not be made during changing weather conditions or if weather fronts are known to be in the area.

The variations in field strength with time often result from causes which are not readily apparent, and it is frequently difficult to determine whether typical propagation conditions prevail. One method which has been proposed and tried with some success is that of establishing fixed recording stations in one or more directions, at locations near the expected outer limit of the measurement program and recording the received signal over a period of several days. These recordings will give an indication of the signal to be expected under average conditions; the coverage survey measurements beyond the horizon can be made during a period when recordings indicate propagation conditions to be typical. Measurements should not be made on days when these recordings indicate excessively high or excessively low field strengths.

APPENDIX A—SAFETY

The following recommended safety precautions are based upon those developed by TASO for Mobile Field Strength Measurements With Antennas Elevated 30 Feet Above Ground. At the time that TASO began, the field strength measurement procedure contained in the FCC Rules specified use of an antenna height of 10 feet despite the fact that the 30 foot standard for predicted fields was in the Rules. As discussed above, the field strength height gain factor for adjusting data from 10 feet to 30 feet is not constant. Accordingly, TASO decided to use a 30 foot height and a series of surveys were begun. However, it became apparent after a series of unfortunate accidents that the measuring of field strength at television frequencies can also be extremely dangerous where observations are made at the 30 foot level. The danger here is not with the field strength measuring equipment but with contact with primary electrical power circuits and, to a lesser extent, potential traffic hazards. In view of the dangers associated with field strength measuring, there are certain precautions everyone should take in addition to the specific precautions included within this memo for safety in operation of the field strength measuring vehicle. These general precautions include the following:

1. Never take chances.
2. Do not service or work on equipment when the power is turned on.
3. Never work on electrical equipment containing high voltages unless there are at least two persons present.
4. If in doubt, keep one hand in a pocket at all times.
5. Everyone who has occasion to work on such electrical equipment must have a knowledge of the best methods of artificial respiration.

I. EMERGENCY PROCEDURE

If the procedures outlined in this memorandum for the safe operation of field strength measuring are followed, there should be no need for emergency procedures. In spite of this, it is felt worthwhile to outline some emergency procedures to be followed in the event of some unforeseen accident. Assuming that some accident has arisen involving overhead obstructions, the following precautions should be observed:

1. The obstruction which you are entangled with may carry **HIGH VOLTAGE.**
2. The vehicle you are in may be at a high potential with respect to ground, so **STOP AND THINK.**
3. Under no circumstances should the transmission line (associated with the field strength measuring gear) be touched. The grounding connections may have broken due to mechanical strain or have burned up due to extremely high currents.
4. The precautions you have taken in inspecting the vehicle before starting the day's work should prevent any high voltage from entering the vehicle.
5. Study your predicament carefully to determine your best course of action. It may be one of the following:
   a. Back the vehicle up.
   b. Drive the vehicle ahead.
   c. Lower the mast.
   d. Raise the mast further.
   e. Get away from the car — **IF YOU DO THIS, REMEMBER THE CAR MAY BE AT A HIGH POTENTIAL WITH RESPECT TO GROUND SO DO NOT TAKE CHANCES—JUMP CLEAR.**
6. Be sure no one else approaches the scene or comes in contact with the vehicle.

II. CONSTRUCTION SAFETY PRECAUTIONS

Since there is a remote possibility of the mast or antenna of the measuring vehicle coming in contact with extremely high voltages, the construction of the vehicle should be such as to reduce to the absolute minimum the possibility of these electrical voltages entering the vehicle. The secure electrical bonds referred to herein must be made with a view towards the hundreds or thousands of amperes that may be involved in the event of an accident. The first of these precautions have to do with external features of the vehicle.

1. All antenna elements which are directly connected to the transmission line must have a secure electrical bond to the mast.
2. The outer conductor of all transmission lines used with the mast must have a secure electrical bond to the top of the mast.
3. The vehicle shall be operated with blinking hazard lights when the mast is elevated.

The additional constructional details to be observed with the vehicle are as follows:

1. The outer conductor of all transmission lines must have a secure electrical bond to the vehicle as soon after entering the vehicle as possible.
2. All electrical equipment (field strength meters, recorders, receivers, and signal generators) must have secure electrical bonding to the vehicle.
3. The vehicle shall be equipped with a light readily visible to the vehicle driver which indicates that the mast is under pressure or up and/or the vehicle shall have a window in the roof from where the driver and engineer, if feasible, can directly view the mast and antenna.

The vehicle shall also be supplied with certain safety equipment, as follows:

1. A pair of high voltage rubber gloves with protecting leather gauntlets. The rubber gloves should be tested at least once a year and a memo including the date of test and the testing organization shall be included in the carrying box for the gloves. All major power companies have provision for making these tests.
2. A nonmetallic safety pole 8 foot minimum for handling hot wires.
3. A CO_2 fire extinguisher. This should be checked annually.

III. OPERATIONAL SAFETY PRECAUTIONS

The foregoing equipment safety precautions are believed to be adequate to make the inside of the vehicle safe in the event of some unforeseen accident. These precautions will only be effective as long as the equipment is in good working condition. Therefore, it is important that all safety precautions outlined in the foregoing paragraph be checked each morning before beginning the day's work. In addition to having a safe vehicle, it must be operated in a safe manner. Therefore, the following precautions must be observed when the car is used:

1. No night work is permitted without prior written approval. In the event this permission is given, it will include additional detailed precautions for the specific job for which approval is given.
2. The mast must not be elevated unless two operators are present.
3. The location for elevating the mast must be carefully chosen both to prevent contact with overhead obstructions and to avoid being a traffic hazard. The mast must not be elevated on busy urban streets or heavily traveled rural highways. The chosen area must be reasonably level and, if a mobile run is contemplated, the vehicle must traverse the path before the mast is raised.
4. Having selected the location for measuring, the following procedure shall be used for elevating the mast. The driver and the engineer will both step out of the vehicle and examine the overhead area for obstructions. The engineer may then return to the vehicle and elevate the mast. If a mobile run is to be made, the driver must walk ahead, examining the path for overhead obstructions and leave a marker at the end of the chosen path. If repeated measurements are to be made along the same path, the starting point should also be marked so that neither end of the examined path is passed.

The vehicle may then be driven between the markers.

5. When the measurements are completed, the driver shall, without stepping out of the vehicle, determine that the overhead area is free for the lowering of the mast. The driver must also make a personal observation that the mast has been fully retracted before driving to the next measuring area.

ACKNOWLEDGEMENT

The assistance of the following present and former associates in writing and editing this article is sincerely appreciated: Howard T. Head, Louis R. du Treil Jr., and Timothy Z. Sawyer.

REFERENCES

1. Federal Communications Commission: FM Technical Standards, Sec. 73.311(c); TV Technical Standards, Secs. 73.683(a) and 73.685(a).
2. "Engineering Aspects of Television Allocations," Report of the Television Allocations Study Organization to the Federal Communications Commission, March 16, 1959.
3. Kalagian, G.S. "A Review of the Technical Planning Factors for the VHF Television Service FCC/OCE RS 77-01," March 1, 1977.
4. LaGrone, Alfred H.: "Forecasting Television Service Fields," Proc. IRE, vol. 48, no. 6, pp. 1009–1018, June 1960.
5. Brown, G.H., J. Epstein, and D.W. Peterson: "Comparative Propagation Measurements: Television Transmitters at 67.25, 288, 510 and 910 Megacycles," RCA Rev., vol. 9, no. 2, pp. 177–202, June 1948.
6. Bullington, K.: "Radio Propagation Fundamentals," Bell System Tech. J., vol. 36, no. 3, pp. 593–626, May 1957.
7. Peterson, D.W., and J. Epstein: "A Method of Predicting the Coverage of a Television Station," RCA Rev., December 1956, vol. 17, no. 4, pp. 571–582.
8. Head, H.T.: "The Influence of Trees on Television Field Strengths at Ultra-High Frequencies," Proc. IRE June 1960, vol. 48, no. 6, pp. 1016–1020.
9. Kinase, Akira, "Influences of Terrain Irregularities and Environmental Surroundings on the Propagation of Broadcasting Waves in the UHF and VHF Bands," Japan Broadcasting Corporation (NHK) Tech. Monograph No. 14, March 1969.
10. Damelin, J., W.A. Daniel, H. Fine, and G.V. Waldo, "Development of VHF and UHF Propagation Curves for TV and FM Broadcasting," FCC Report No. R-6602, September 7, 1966.
11. Norton, K.A. "The Calculation of Ground Wave Field Intensity over a Finitely Conducting Spherical Earth," Proc. IRE, vol. 29, December 1941.
12. Rice, P.L., A.G. Longley, K.A. Norton and A.P. Barsis "Transmission Loss Predictions for Tropospheric Communication Circuits," NBS Technical Note No. 101, January 1, 1963.
13. Reed, H.R. and Russell C.M. "Ultra-High Frequency Propagation," John Wiley and Sons, 1953.
14. Greene, Frank M.: "Calibration of Commercial Radio Field Strength Meters at the National Bureau of Standards," Natl. Bur. Standards Circ. 517, December 1951.
15. Greene, Frank M., and Max Solow: "Development of Very-high Frequency Field Intensity Standards," Natl. Bur. Standards Research Paper RP2100, vol.44, May 1950.
16. Terman, F.E.: "Radio Engineers' Handbook," McGraw-Hill Book Company, Inc., New York, 1943.
17. Schrack, R.A.: "Radio-frequency Power Measurements," Natl. Bur. Standards Circ. 536, March 16, 1953.
18. Greene, Frank M.: "Influence of the Ground on the Calibration and Use of VHF Field Intensity Meters," Natl. Bur. Standards Research Paper RP2062, vol. 44, February 1950.
19. Federal Communications Commission's Rules, Sec. 73.684.
20. Longley A.G. and P.L. Rice: "Prediction of Tropospheric Radio Transmission Loss - A Computer Method," ERL 79-ITS 67, Envir. Sci. Services Admin., July 1968.
21. Longley A.G.,"Radio Propagation in Urban Areas," OT Report 78-144, U.S. Dept. of Com., April 1978.
22. Hufford, A.G.Longley and W.A. Kissick: "A Guide to The Use of The ITS Irregular Terrain Model in The Area Prediction Mode," U.S. Dept. of Com., April 1982.
23. Okumura Y., Oiimori, E. Kawano, T and Fukuda, K.: "Field Strength and Its Variability in VHF and UHF Land Mobile Radio Services," Rev. of Elect. Comm. Lab. vol. 16, no. 9–10, Sept.–Oct., 1968.

3.1
AM Transmitters

George W. Woodard, P.E.
Radio Free Europe/Radio Liberty, Inc., Munich, Federal Republic of Germany

Broadcasting to the general public began with the process known as amplitude modulation near the start of the second decade of this century. Then, as now, the system of modulation chosen for transmission was heavily dependent on practical and economic aspects of receiver technology. The evidence is clear that amplitude detection was the only known practical method of signal demodulation when the ideas of radio communication began to be formulated in the late 1800s and early 1900s. It is apparent from some of the earliest technical writings on radio communications, that a general mathematical knowledge of angular modulation (i.e., frequency and phase modulation) did not exist until the mid-1920s.[1] These modes of transmission and the necessary receiver technology were not to be proven practical until long after amplitude modulation had become the standard of the radio communication and broadcast technological art.

As time and technology progressed, engineers could no longer ignore the problems of transmission efficiency. When, in the 1930s, the telephone industry was planning a major switch from double-sideband, full carrier, amplitude modulation (DSB-FC-AM) in favor of single-sideband, suppressed carrier (SSB-SC) for both its long distance wire and wireless services because of its higher transmission efficiency and channel capacity,[2] the broadcast industry was necessarily committed to the continuance of conventional AM because of the need for economic receiver compatibility. Therefore, broadcast engineers' attention naturally turned to improvement of transmission standards such as fidelity, efficiency of transmission, reliability of transmission, co-channel and adjacent channel interference. Of these, efficiency and reliability of transmission have seen and continue to see the greatest improvement in the views of many transmitter designers and users. Transmission fidelity, always important, reached a plateau in the late 1940s of approximately 2% to 3% total harmonic distortion (THD), 10 kHz to 15 kHz audio bandwidth, and −55 dB noise. This was not significantly improved upon until the early to mid-1980s when totally solid-state transmitters began to appear on the AM transmitter market, yielding nominally 1% or less THD, and −65 dB or better noise. In addition, pre-transmitter program processing equipment and philosophy, which influence perceived reception fidelity, continue to improve and change through use of modern linear circuit and digital design techniques.

AMPLITUDE MODULATION THEORY

Amplitude modulation in broadcasting is today more accurately referred to as DSB-FC-AM, an acronym for *double sideband-full carrier-amplitude modulation*. However, the abbreviation "AM" has been generally accepted to describe the mode of transmission currently used throughout the world. From almost the very beginning of broadcast technology, it was known that AM was not a very efficient mode of transmission, either in spectrum usage or in transmission of intelligence. The pioneers in AM broadcast transmitter and receiver engineering technology had a deep sense of responsibility toward developing this "marvelous medium" to enhance the lives of the general public with informative, educational, and entertaining programs. Some early communication technology pioneers in the U.S. and abroad had great vision for both the technical and programming aspects of public broadcasting.[3]

Amplitude modulation of a radio signal occurs in at least two basic forms, coded (digital) and linear (analog). The first amplitude modulation processes for long distance communication involved on/off keying of a radio carrier wave. The pattern or "code" of the on/off keying process determined the content of the information being transmitted. Linear or quasi-linear undulations are normally used to transmit the analog information present in speech and music. The radio carrier wave signal onto which the analog amplitude variations are to be impressed is expressed as:

$$e(t) = AE_c\cos(\omega_c t) \qquad [1]$$

where:

- $e(t)$ = Instantaneous amplitude of carrier wave as a function of time (t)
- A = A factor of amplitude modulation of the carrier wave
- ω_c = Angular frequency of carrier wave (radians/second)
- E_c = Peak amplitude of carrier wave

If A is a constant, the peak amplitude of the carrier wave is therefore constant and no modulation exists. Periodic modulation of the carrier wave exists if the magnitude of A is caused to vary with respect to time as, for instance, a sinusoidal wave:

$$A = 1 + (E_m/E_c)\cos(\omega_m t) \qquad [2]$$

where E_m/E_c is the ratio of modulation amplitude to carrier amplitude leading to:

$$e(t) = E_c\left[(1 + (E_m/E_c)\cos(\omega_m t)\cos(\omega_c t)\right] \qquad [3]$$

This is the well known basic equation for periodic (sinusoidal) amplitude modulation, and when all multiplications and a simple trigonometric identity are performed the result is:

$$e(t) = E_c\cos(\omega_c t) + (M/2)\cos(\omega_c t + \omega_m t) + (M/2)\cos(\omega_c t - \omega_m t) \qquad [4]$$

where: M = the amplitude modulation factor E_m/E_c

Eq. 4 can be represented graphically in three familiar ways: in the time domain representation shown in Fig. 1; in the frequency domain shown in Fig. 2; and as relative vectors shown in Fig. 3. The graphical representations shown are for a single-tone modulation index (M) of 0.7: i.e., the peak modulating voltage is 70% of the peak carrier wave voltage ($E_m/E_c = 0.7$). Fig. 2 shows the occupied bandwidth of an AM signal with single tone modulation. From this figure and its defining Eq. 4, it is clear that the bandwidth of an AM signal is equal to twice the highest modulating frequency when no system distortion is present. High quality music reproductions include frequency components as high as 15 kHz or higher and therefore the required theoretical bandwidth of a DSB-FC-AM signal capable of high quality music reproduction is at least 30 kHz. There are practical limits to the transmitter bandwidth which are related to the allocation requirements for stations on adjacent channels, however. Regulatory limits on transmitted bandwidths are discussed in later sections. Harmonic and intermodulation system distortion have the effect of widening the effective occupied bandwidth of the system. However, most modern transmitters have sufficiently low distortion characteristics that bandwidth stretching, due to system nonlinear distortions, is not normally a significant problem. The occupied bandwidth characteristics of an AM broadcast transmitter are discussed in more detail in the sections on "Factory Tests" and "Audio Processing and Preemphasis."

RADIO FREQUENCY

Power Amplifiers

The most important system component which is common to all amplitude modulation systems is the high-power final radio frequency (RF) power amplifier. High-power amplifiers that produce 0.25 kW to 50 kW of carrier power are common in AM broadcast transmitters in North America. Carrier power levels of one megawatt and higher are common in other parts of the world for medium wave broadcasting. The most common amplifiers used to meet the demands of high output power and high efficiency are the class-C, C/D, or D amplifiers, which may be either vacuum tube for large output powers from a single amplifier circuit or solid-state for large output powers derived from several combined lower power amplifier circuits. Class-B RF power amplifiers are still used occasionally at lower power levels and for driver stages of final class-C or D stages. Solid-state class-D amplifiers of up to five kilowatts are becoming more common as driver stages for final vacuum tube amplifiers and for the final power amplifier modules in transmitters of up to 50 kW carrier power. The most common major concerns for both manufacturers and users of modern AM broadcast transmitters are operating reliability and efficiency, both affecting operating cost. Therefore, the stages which consume the most power, the modulator and the final RF power amplifier stages, must be designed for the highest possible operating efficiency.

The basic tuned-anode vacuum tube amplifier is described in graphical form in Fig. 4. The vacuum tube can be either a triode, tetrode, or pentode. Tetrode final amplifiers are most common in modern high power transmitter designs. The RF excitation voltage is supplied to the grid of the power amplifier tube and the ratio of dc grid bias voltage to peak RF excitation voltage (shown sinusoidal in Fig. 4) determines the conduction angle (Θ_c) of anode current, given as:

$$\Theta_c = 2\arccos(E_{cc}/(E_g - E_{cc}) \qquad [5]$$

where the exciting grid signal (E_g) is sinusoidal as shown in Fig. 4A.

The shape of the anode current pulse is determined by the vacuum tube transfer characteristics and input wave shape. The pulse of current thus generated, Fig. 4C, is supplied by the DC power supply, E_{BB}, and passed through the resonant anode tank circuit shown in Fig. 4D. The resonant anode tank circuit is assumed to have sufficient operating Q to force anode voltage, e_p, to be essentially sinusoidal and of the same periodic frequency as the RF excitation voltage and resultant anode current pulse. The instantaneous anode dissipation, shown in Fig 4E, is the product of instantaneous tube anode voltage drop and anode current. The tube

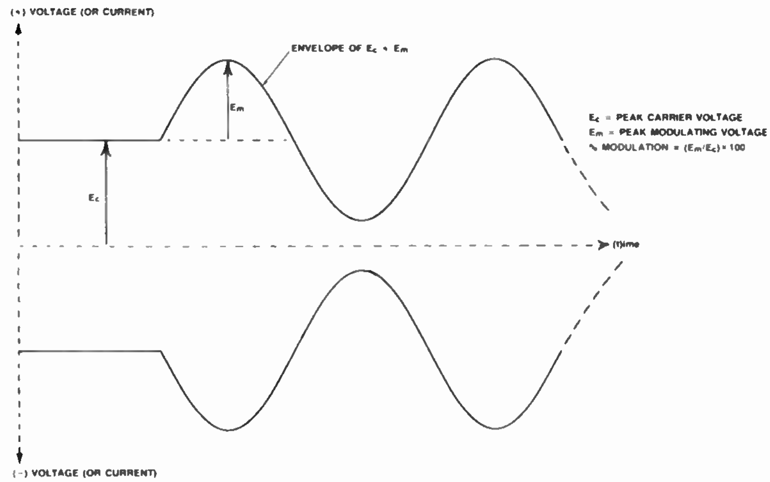


Figure 1. Time domain representation of a carrier wave signal amplitude modulated by a sinusoidal audio signal to a peak modulation depth of 70%.

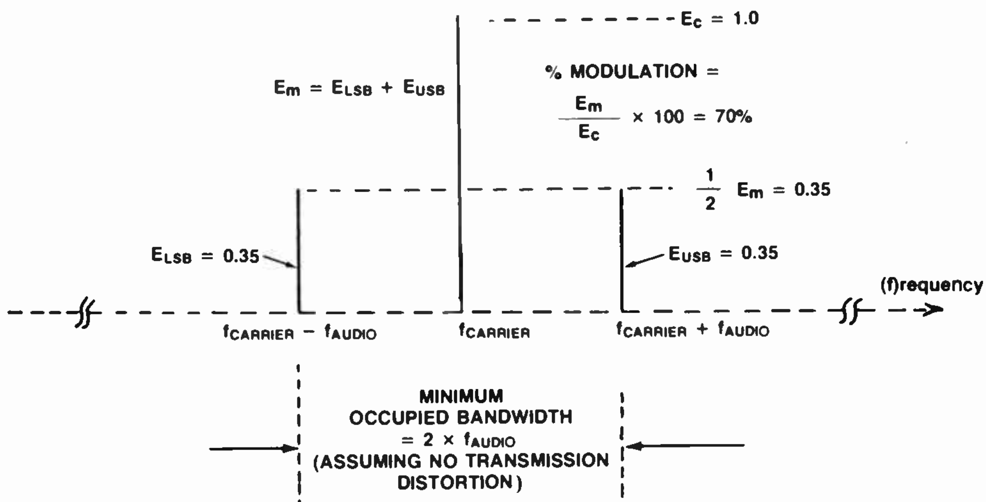


Figure 2. Frequency domain representation of an amplitude modulated signal showing the carrier wave signal and two resultant modulation sidebands at 70% modulation.

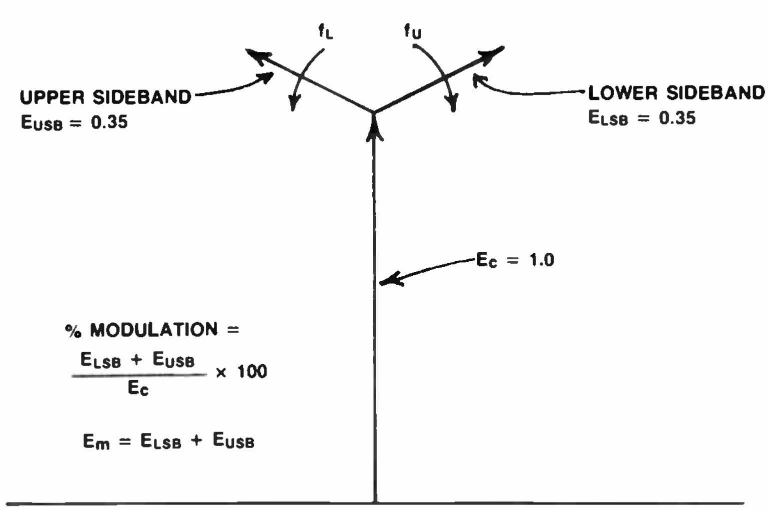


Figure 3. Vector representation of an amplitude modulation signal showing the carrier wave signal and two resultant modulation sidebands at 70% modulation.

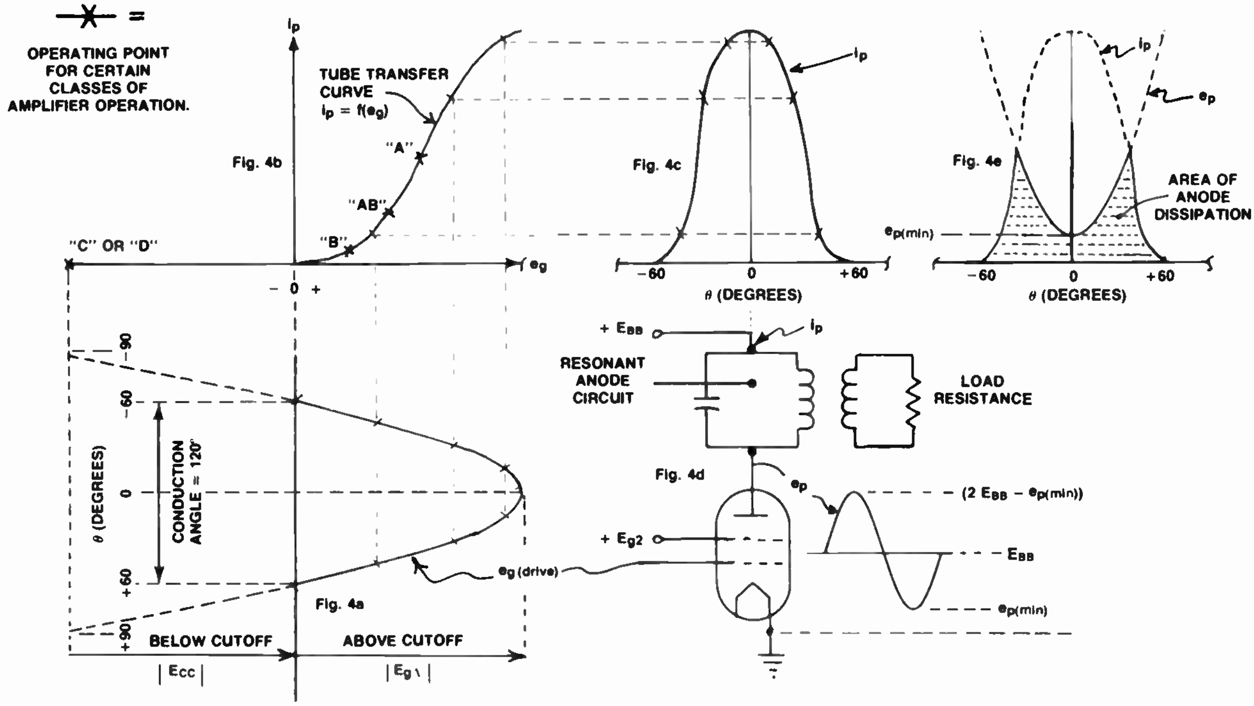


Figure 4. Classical vacuum tube class "C" amplifier with sinusoidal grid drive, 120° anode current conduction, and resonant anode load.

transfer characteristic is a variable dependent upon many tube factors, as well as maximum drive signal E_g. The exact shape and magnitude of the current waveform is normally obtained from a load-line plot on constant current characteristic tube curves supplied by the tube manufacturer. The resonant anode load impedance is chosen and adjusted to allow e_p(min) to be as low as possible without causing excessive screen grid (in the tetrode case) or control grid dissipation.

Some manufacturers increase anode efficiency beyond the limits for typical class-C amplifiers by using a circuit employing a third harmonic resonator between the output anode connection and the fundamental resonant circuit. In some cases, fifth harmonic resonators are also employed. This has the effect of squaring up the anode voltage waveform (e_p) thus causing the integral of the $e_p \times i_p$ product, or anode dissipation, to be smaller; resulting in lower anode dissipation for a given RF power output. An amplifier employing the third harmonic anode trap is commonly referred to as class-C-D: suggesting an efficiency rating somewhere between conventional class-C operation (nominal 120 degree conduction angle) and true class-D operation with rectangular anode or collector voltage waveforms. Anode efficiencies can be increased typically to values of 90% for transmitters up to approximately 10 kW carrier power and approximately 85% for transmitters higher than 10 kW carrier power by using the third harmonic trap technique. Table 1 shows a comparison of anode efficiency for six classes of high power tuned RF amplifiers.

DISCUSSION OF BASIC AM SYSTEMS

High Level Anode Modulation

The first practical method of generating the amplitude modulation signal was Heising constant current modulation,[4,5] a method of applying audio modulation to the anode supply voltage of a class-C RF amplifier. This general class of modulation is now known as high level anode (or plate) modulation. The Heising modulator was used at least as early as 1920 and was usually used to modulate a low-power RF amplifier or master oscillator stage, which was followed by several linear amplifier stages, until the desired power level was attained. In some cases the Heising modulator was used to modulate the final RF amplifier stage of lower power transmitters. The Heising shunt modulator operated in the class-A mode and therefore was low in operating efficiency. The early linear amplifiers were tuned class-B amplifiers operating with carrier level anode efficiencies of 30% maximum. Heising and similar systems of audio amplification were also used to modulate the grid bias level of RF amplifier stages in order to obtain the AM signal to be used for further linear amplification. Heising constant-current anode modulation was very popular in military and aviation radio sets used through the end of World War II.

Chireix "Outphasing" Modulation

Outphasing modulation was originally described in the literature by its inventor Henry Chireix in 1935.[6] It is a unique and ingenious method of obtaining an

AM signal by use of counter-phase modulation and vector addition of two separate radio frequency signals. It was marketed for many years by RCA under the trade name "Ampliphase," and many of those transmitters and others of European manufacturers are still on the air today at power levels of 100 kW and higher. This system of modulation is described graphically in Fig. 5. Two RF signals are derived from a common excitation source and then split into two separate channels. Each channel is shifted in phase, one positive and the other negative. The two channels are then each phase modulated by the modulating signal, again in opposing polarity. The two channels are amplified and then recombined in a vector-additive network which has the effect of producing the desired amplitude modulation. The main advantage, as with all systems previously discussed, is in operating efficiency. The two independent channels contain only phase modulated RF signals and, therefore, each can be amplified to the desired power levels in high efficiency class-C or D amplifiers. The actual modulation process takes place both at low level, in the phase modulators, and at high level, in a passive output network combiner.

There are two major disadvantages to this system of modulation. First, the efficiency of the output power amplifiers is not quite as high as the simple description above would imply, due to the fact that, at all instantaneous levels of modulation except one, the anode circuits must work into a reactive load. Secondly, output carrier power setting is sensitive to tuning of any stage in the chain; operators of early versions of this type of equipment soon learned of the potentially disastrous consequences of "trimming" a tuning control of some lower power stage. Later designs used broadband amplifying stages for the lower levels to circumvent the problem of tuning sensitivity.

A major disadvantage of the Chireix system in older transmitters was the generation of incidental phase modulation (IPM) in excessive amounts when the two low-level phase modulators were not well balanced in opposing PM characteristics. It was common for such an imbalance to exist, producing as much as 12° to 18° of peak IPM at either 100% modulation crest or trough. Modern solid-state circuitry using digital phase modulation techniques would completely remove this disadvantage (were the system still being commercially produced). This system ceased to be produced commercially when RCA discontinued AM transmitter manufacturing in the mid-1970s.

TABLE 1

Comparison of tuned RF amplifier efficiencies.

Amplifier Class	Conduction Angle (degrees)	Anode Efficiency (%)	Defined Conditions of Operation
A	360	30	$E_{b(min)} = 0.10 \times E_{BB}$
A-B	240	60	$E_{b(min)} = 0.10 \times E_{BB}$
B	200	67	$E_{b(min)} = 0.10 \times E_{BB}$
C	120	8	$E_{b(min)} = 0.05 \times E_{BB}$
C-D	120	90	$E_{b(min)} = 0.05 \times E_{BB}$
D	120	95	$E_{b(min)} = 0.05 \times E_{BB}$

Class-B High Level Anode Modulation

Historically, the most popular method of applying an audio modulating voltage to the anode circuit of a class-C RF power amplifier used a high-power, push-pull class-B audio amplifier. With the final RF power amplifier operating at approximately 80% anode effi-

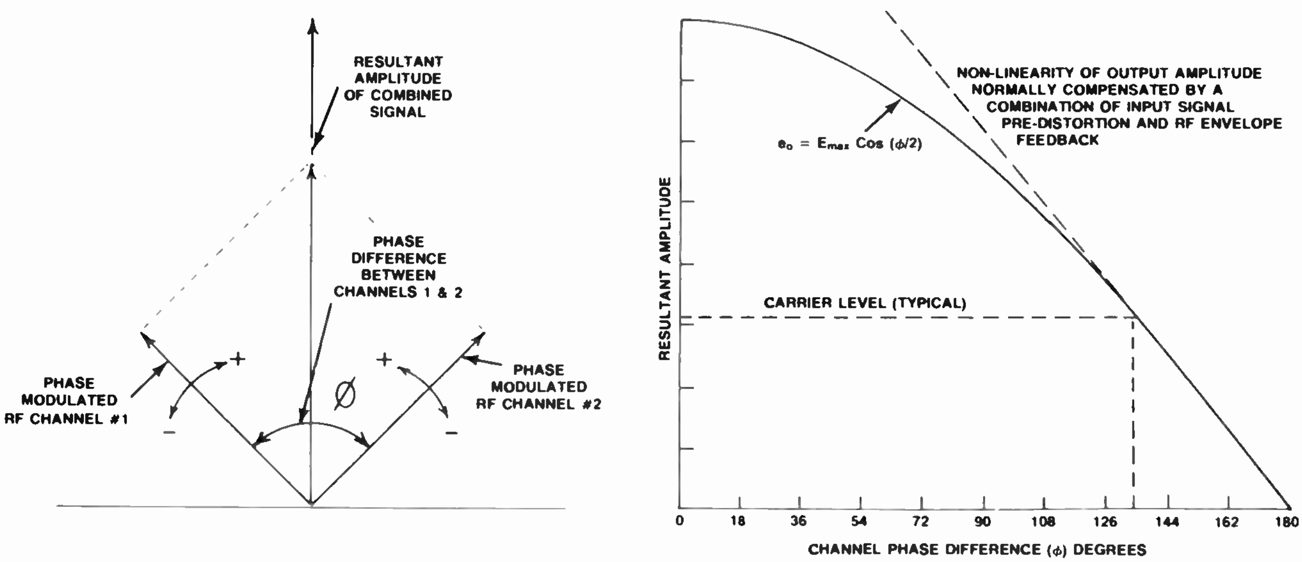


Figure 5. Principle of Chireix "outphasing modulation."

ciency and the class-B audio modulator total static currents approximately one-tenth that of an equivalent Heising modulator, total anode efficiencies at carrier level rose to approximately 72% compared to 37% for the Heising system and 30% for conventional linear amplification. A simplified drawing of a typical high level class-B anode modulation system is shown in Fig. 6. The vacuum tubes shown in Fig. 6 may be either triodes, tetrodes, or pentodes. The output circuit of the class-B modulator shows the output modulation transformer (MT), an audio coupling capacitor (C), and a dc shunt feed inductor (L).

This arrangement was used in all high level class-B high-power broadcast transmitters until about 1960, because of a transformer design constraint that would not economically allow unbalanced direct currents to magnetize the transformer core material without poor low frequency distortion. Advanced-technology core materials and careful magnetic transformer design allowed elimination of the coupling capacitor and the dc feed shunt inductor. Many of the more advanced modern AM broadcast transmitter designs still using high level class-B anode modulation have eliminated the extra C and L components from the modulator output circuitry, and the dc current to the modulated RF amplifier anode flows directly through the secondary of the output modulation transformer. This is necessary for optimum operation of modern AM stations. With the extra C and L components, the modulator output is effectively a three-pole, high-pass filter which causes low frequency transient distortion to be generated when driven with the complex waveforms that are produced by many modern and popular program audio processors. Eliminating the C and second L component causes the output modulator circuit to become a single-pole high-pass filter, greatly reducing low frequency transient distortion.

Another problem with transformer-coupled high-level class-B anode modulation is high frequency audio transient distortion. Stray internal winding capacitances and leakage inductances form multi-pole low pass filtering at the high frequency end of the audio spectrum. This equivalent multi-pole, low-pass filter generates transient overshoot distortion when driven by the processed complex program waveforms mentioned above. Transient overshoot up to 12% is typical for square-wave modulator input signals and results in a required modulation level reduction of the same 12% in order to maintain peak modulation levels within FCC allowed limits.

High frequency transient overshoot distortion can be effectively minimized by filtering the audio input to the transmitter with linear phase filters, resulting in somewhat lower high frequency audio response, by careful control of the modulation transformer equivalent circuit yielding more linear audio phase characteristics for the entire modulator circuitry, or by both methods. Balanced modulator negative feedback is used to reduce modulator nonlinear distortion and noise. Negative feedback, however, usually worsens high audio frequency transient distortion characteristics.

Doherty High-Efficiency Linear Amplifier for High-Power Facilities

The Doherty high-efficiency linear amplifier was first described in the technical literature in 1936 by its inventor, W.H. Doherty.[7] The Doherty high-efficiency

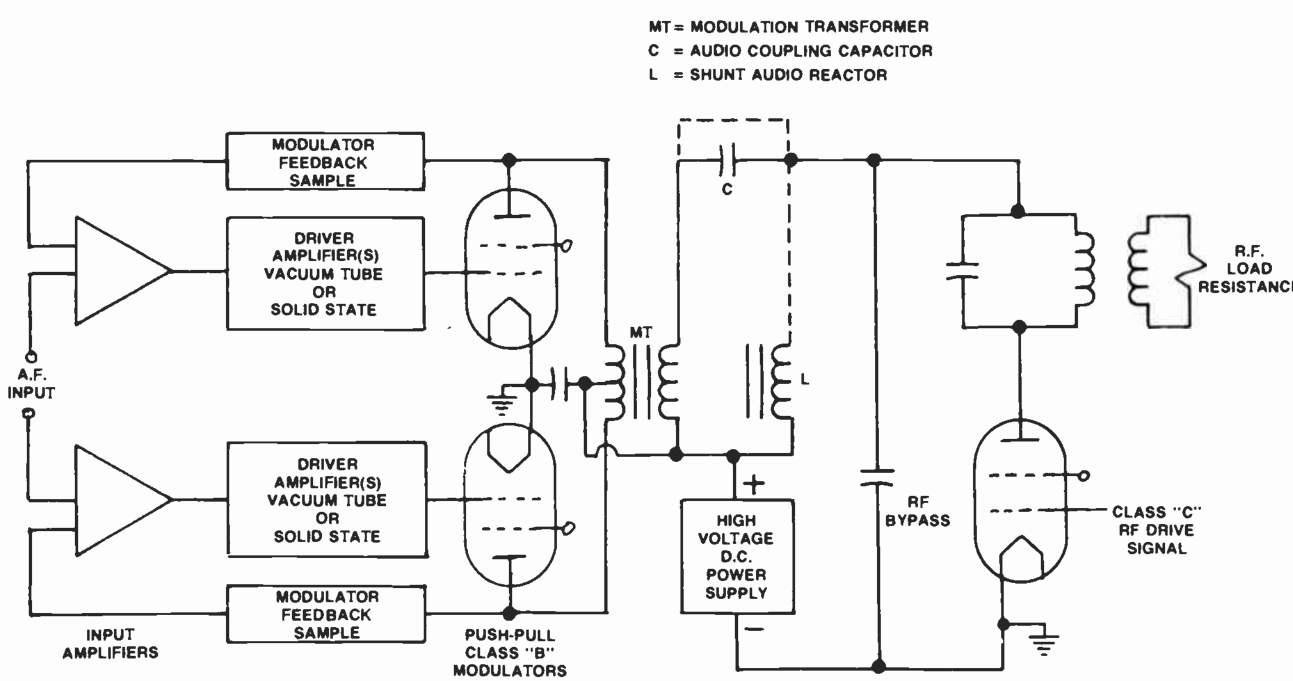


Figure 6. High level anode modulation employing transformer coupled push-pull class-B modulators.

linear amplification system has been used at power levels up to 500 kW carrier power, in both the original and in the patented Weldon[8] modified form, in many installations throughout the world on the medium wave broadcast and international short-wave broadcast bands, as well as on the long wave broadcast band in Europe. The Doherty linear amplifier is described graphically in Fig. 7. As with conventional linear amplifiers, the AM signal is generated at low levels and applied to the input of the final amplifier stage. The Doherty system employs two output amplifier stages, one defined as the carrier amplifier and the second as the peak amplifier.

The outputs of the two stages are combined in phase at the anode of the peak amplifier tube. At carrier level, the carrier tube is operated as a nearly saturated class-B amplifier and thus delivers almost all of the carrier power at class-B efficiencies, (approximately 70% anode efficiency). At carrier condition, the peak tube is biased and driven just above cutoff and therefore supplies a small amount (approx. 2% to 6%) of carrier power. The anodes of the two tubes are connected through a 90° impedance-inverting RF network. As the modulated signal increases in the positive direction to both peak and carrier tubes, the current supplied to the output load by the peak tube increases. The saturated voltage drop at the anode of the carrier tube remains constant over the entire positive modulation half-cycle, thus causing the current at the output of the inter-anode 90° network also to be constant during the same positive modulation half-cycle. The rising current from the peak tube anode has the effect of raising the impedance presented to the inter-anode network. Since the current from the network is constant, the net effect is an increase in output power from the carrier tube: i.e., $I^2 \times R_L$ increases because R_L increases.

At the 100% positive modulation crest, both tubes are producing exactly twice carrier power to the load, satisfying the requirement that peak envelope power (PEP) = $4 \times P_{carrier}$. During the negative half-cycle of modulation, the peak tube is cut off and the carrier tube behaves as a normal linear amplifier, allowing the envelope power output to drop linearly to zero at the 100% negative modulation trough. The anode efficiency of the Doherty high-efficiency linear amplifier at carrier level is more than twice the efficiency of conventional AM class-B linear amplifiers.

The Doherty linear amplifier also has two other important advantages in high power broadcast transmitters. First, and most important, the peak anode voltage at either tube is only about one-half that required for an equivalent carrier power high-level pulse-width modulation (PWM) or class-B anode modulated transmitter, thus allowing reliability and usable tube life to increase significantly. Secondly, no large modulation transformer or special filtering components are used in the final amplifier stages to cause transient overshoot distortion as previously discussed for class-B anode modulation or pulse-width anode modulation.

The main problems of the Doherty linear amplifier are nonlinear distortion and increased complexity of tuning. The major sources of nonlinear distortion are the nonlinearity of the carrier tube at or near the 100% negative modulation crest, and the nonlinearity of the peak tube at or near carrier level when it is just beginning to conduct. Both sources of distortion are effectively reduced by use of moderate amounts of overall envelope feedback. The tuning complexity problem is usually overcome by experienced and trained operators, by simplifying tuning procedures, and by built-in test equipment.

Pulse-Width High-Level Anode Modulation

Pulse-width modulation (PWM) of the dc anode voltage of a class-C RF amplifier was the first commercially successful attempt to improve significantly upon the efficiency of the popular high-level class-B modulation system by applying and improving basic PWM concepts described by Heising.[9] Since this first success, pulse-width modulation has become the method of high-level anode modulation preferred by many broadcast engineers, and is employed in broadcast transmitter designs by several manufacturers worldwide. The basic pulse-width modulation system and two ingenious improvements to the basic system are shown in Fig. 8. The circuit in Fig. 8A graphically describes the basic principle of PWM.

An inherent practical deficiency in the basic concept, Fig. 8A, is caused by the relatively high shunt circuit capacitance of the modulator tube filament transformer plus stray capacitances. Though special low-capacitance isolation transformers can be used to supply modulator filament and auxiliary power to minimize capacitive switching losses, typical realizable values of capacitance can cause excessive switching losses and audio distortion in lower-power transmitters.

For example, the switching losses at carrier level of a typical PW modulator for a 5 kW transmitter can be higher than the quiescent modulator losses of an equivalent power class-B modulator even when the stray modulator tube capacitances are as low as 100 pF. The power lost per modulator switching cycle is:

$$P_{modsw} = (CV^2/2) + P_{td} \qquad [6]$$

where:

- C = The shunt modulator filament transformer plus stray capacitance to ground
- V = The pulse switching voltage to ground at the cathode of the tube
- P_{td} = The saturated tube and diode losses during the respective on and off conduction states

Besides causing switching losses, this high stray capacitance to ground also causes modulator distortion at high negative modulation indices.[10] The circuit shown in Fig. 8B is one ingenious way to overcome the stray modulator capacitance problem. It is identical to the circuit in Fig. 8A, in principle, except that the high-voltage, pulse-modulated wave is at a point in the circuit where shunt capacitances to ground are inherently minimized. The circuit in Fig. 8C is basically

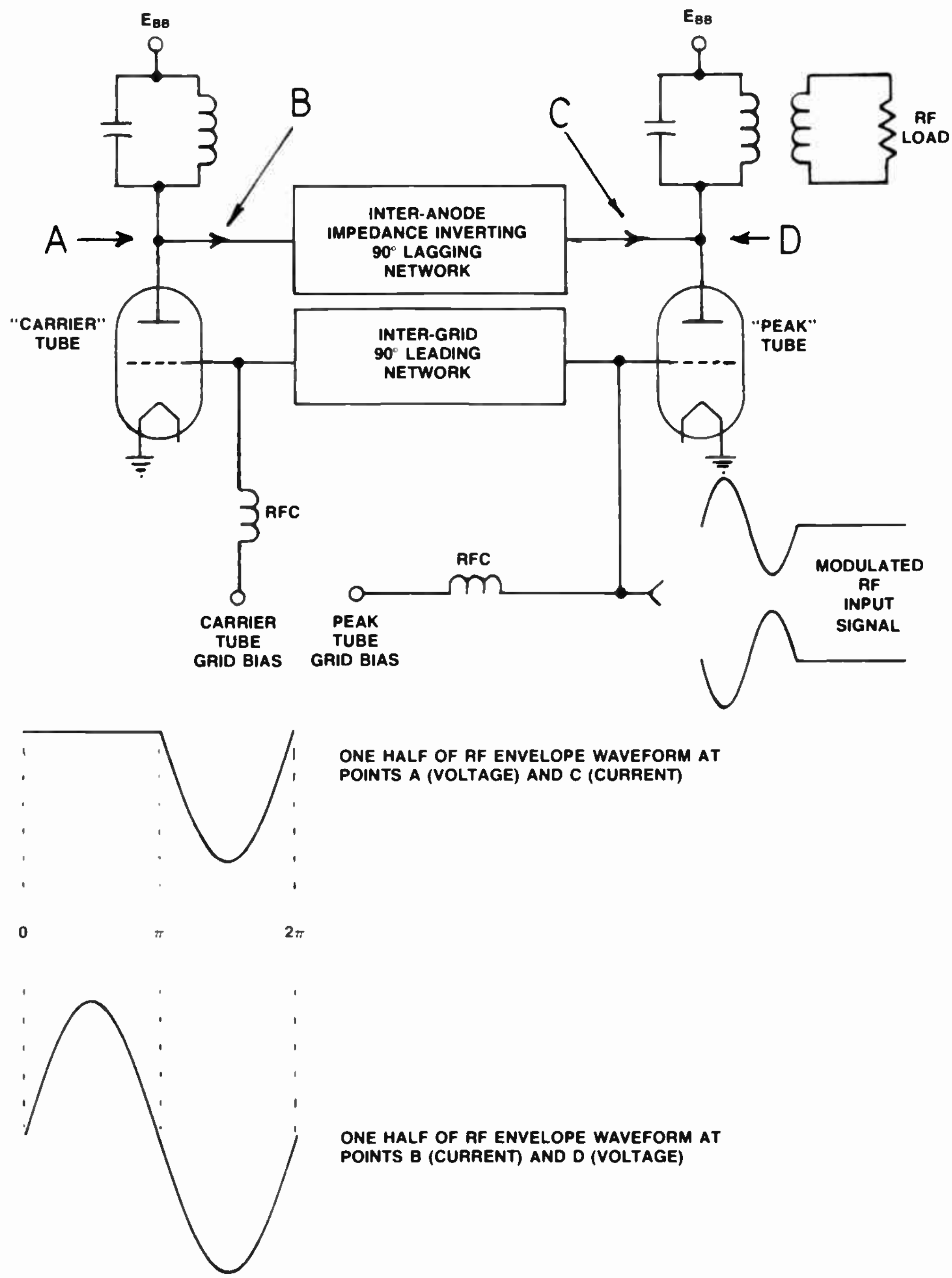


Figure 7. Doherty high efficiency linear amplifier for modulated waves.

the same as in 8B except that the system ground has been placed at another point in the circuit. PW modulator anode efficiencies approach 92% in some higher power transmitter designs yielding a combined modulator and carrier tube anode efficiency of approximately 74% at all levels of modulation.

An added efficiency advantage of PWM over high-level class-B anode modulation is that a PWM transmitter may have only two high power vacuum tubes, one modulator and one final RF amplifier, thus eliminating the filament heating power of one large vacuum tube required in push-pull modulator designs.

A significant disadvantage of PWM, as described in Figs. 8A and 8C, is that the cathode and grid circuits of the modulated RF amplifier are operated at high voltage levels off ground, adding complications to the circuitry that are avoided with the classical PWM circuit of Fig. 8A and conventional high-level class-B anode modulation. Another major disadvantage of PWM in any form is transient distortion caused by the phase nonlinearity of the multipole PWM filter (similar to that previously discussed for high-level class-B anode modulation). The switching frequency is typically 70 kHz for most transmitters manufactured in North America. This frequency is chosen to ease compliance with FCC regulations, which require that all spurious radiation more than 75 kHz removed from the carrier frequency be 80 dB or more below the carrier level.

In order to meet the spurious-output requirements of paragraph 73.44 of the FCC regulations, a very steep cutoff low-pass filter is required at the output of the PW modulator. Most manufacturers of PWM transmitters attempt to maintain linear phase characteristics in this filter to the highest audio frequency possible, but the laws of nature prevent required linear phase characteristics to the highest third audio harmonic from being achieved within the major attenuation constraints noted above. As a result, transient response to a square wave input can result in peak overshoot of approximately 6%. As with class-B modulation, this overshoot can be eliminated effectively by the use of linear phase (Bessel) filtering of the input modulation signal at the expense of a small reduction in modulation frequency response capability.

High-Efficiency Screen/Impedance Modulation

In 1938, Terman and Woodyard[11] described a modification to the basic Doherty high efficiency linear amplification system previously described. In the new system, the grid bias level of two tubes operating class-C is varied at the audio modulation rate, thus creating a higher efficiency system of amplitude modulation rather than amplification while still using the impedance inverting properties of the inter-anode network described by Doherty. The Terman-Woodyard system of modulation, however, was not extensively used in commercially successful high-power transmitter designs.

High efficiency screen/impedance modulation[12] is a method in which the audio modulating signal is applied to the screen grids of two tetrode vacuum tubes operating as class-C carrier and peak amplifiers. Invented by J.B. Sainton in 1965, the screen/impedance modulation system exhibits significant improvement because RF excitation voltages and audio modulating voltages are isolated from each other, thereby eliminating a troublesome source of tuning versus modulation interaction.

The screen/impedance modulation system is shown and described graphically in Fig. 9. The peak and carrier tubes are biased and driven in quadrature as class-C amplifiers from the continuous wave RF drive source. At carrier level, the screen voltage of the carrier tube is adjusted so that the carrier tube is near anode saturation and delivering approximately 96% of the carrier power. The screen voltage of the peak tube is adjusted so that the peak tube is just into conduction and supplying the remaining approximate 4% of carrier power. The combined anode efficiency at carrier level is better than 77% as shown in Eq. 7.

$$n_{at} = 1/(pc/n_{ac}) + (pp/n_{ap}) \qquad [7]$$

$$n_{at} = 1/(0.96/0.8) + (0.04/0.40) = 0.77$$

where:

- pc = Percent carrier power supplied by carrier tube (as a decimal)
- pp = Percent carrier power supplied by peak tube (as a decimal)
- n_{ac} = Carrier tube anode efficiency at carrier level (as a decimal)
- n_{ap} = Peak tube anode efficiency at carrier level (as a decimal)
- n_{at} = Total anode efficiency at carrier level (as a decimal)

Modulation of the RF carrier wave occurs when the screen voltage of the peak tube begins to rise during the positive modulation half-cycle, thus causing the peak tube to supply more RF current to the output load. This increase of current into the output network causes the resistance seen by the inter-anode network to increase and, due to the impedance inverting characteristic of the 90° inter-anode network, causes a proportional decrease in the load impedance presented to the carrier tube anode.

The carrier tube resonant anode voltage drop is fully saturated over the entire positive modulation half-cycle, and therefore is effectively a constant voltage source. The power output of the carrier tube thus increases during the positive modulation half-cycle, due to the modulated decreasing impedance at its anode until both peak and carrier tubes deliver twice carrier power at the 100% positive modulation crest. During the negative modulation half-cycle, the peak tube is held out of conduction while the carrier tube output voltage decreases linearly to zero output at the 100% negative modulation trough. Simple alterations to the inter-anode 90° network characteristic impedance gives the screen/impedance modulation system the capability of full FCC-allowed positive modulation (up to 125% peak) without added distortion.

The advantages of screen/impedance modulation are the same as mentioned for the Doherty linear amplifier except that screen/impedance modulation has higher efficiency at all depths of modulation, and is less critical to misadjustment of RF amplifier tuning. Screen/impedance modulation has been successfully used in medium-wave transmitters throughout the world at the two megawatt carrier power level and up to 250 kW in automatically-tuned transmitters for international shortwave broadcasting.

SOLID-STATE BROADCAST TRANSMITTERS

Solid-state technology came to high-power medium and shortwave AM broadcast transmitters about 1984. The new designs and products are proving themselves

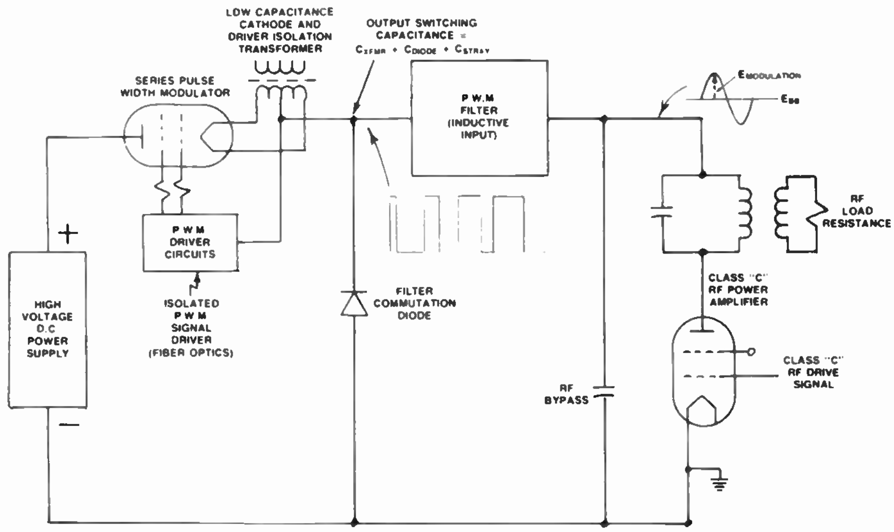


Figure 8A. Basic classical high level anode pulse-width modulation of a vacuum tube class "C" amplifier.

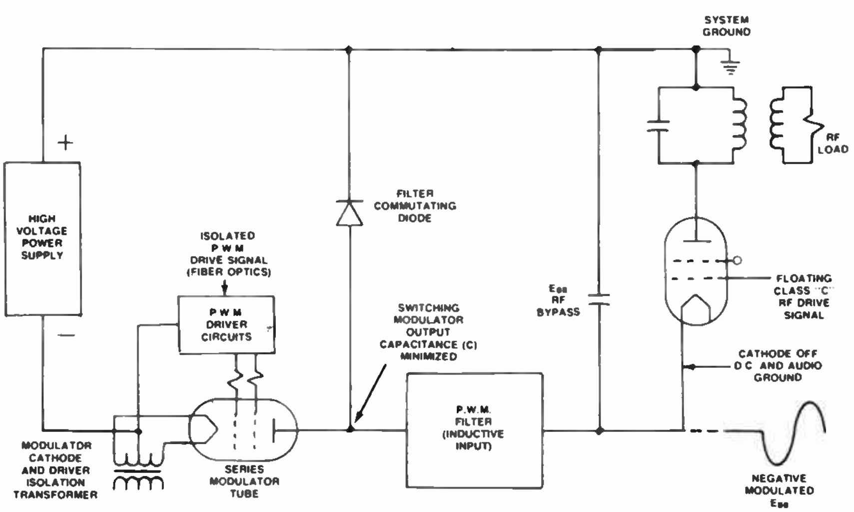


Figure 8B. Collins Radio/Continental Electronics patented modification to basic high level PWM System to minimize modulator lossed by minimizing switching modulator output capacitance.

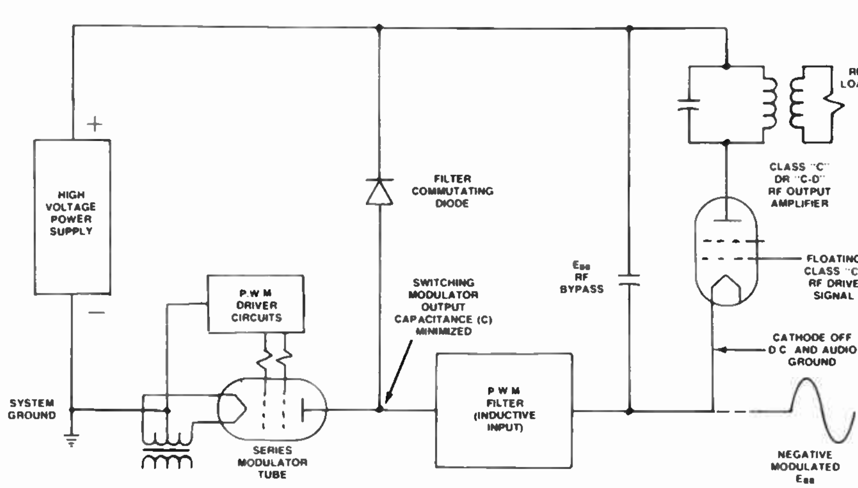


Figure 8C. Harris Corp. patented modification to basic high level PWM System to minimize modulator losses by minimizing switching modulator output capacitance.

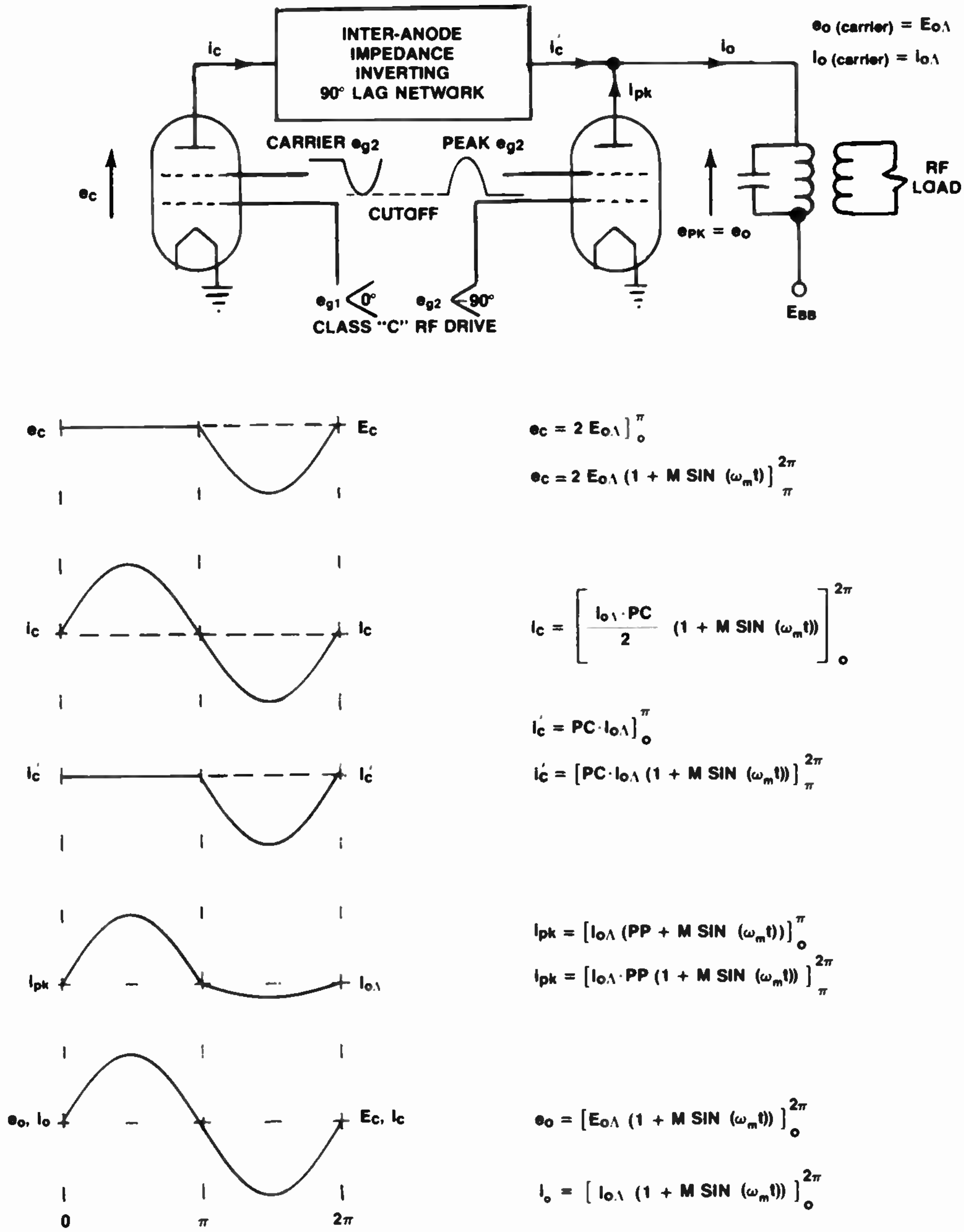


Figure 9 Principles of Sainton high efficiency screen/impedance modulation.

worthy in the harsh environment of high-power transient disturbances caused by gas arcs in associated vacuum tube circuitry and lightning strikes on antennas, power lines, and buildings. Products now available on the commercial market offer solid-state modulators for high-level anode modulation of a vacuum tube final RF amplifier up to 600 kW on medium wave and 500 kW for shortwave transmitters. The new all solid-state designs, up to 50 kW carrier power levels, are achieving great commercial success because of their higher operating efficiencies, reliability, size, weight, and superior modulation characteristics over competitive vacuum tube designs. Products and designs employing solid-state technology in the modulator circuitry for high-level anode modulation of power grid vacuum tubes are also enjoying great commercial success worldwide.

SOLID-STATE EQUIPMENT AND CIRCUITRY

Solid-state designs of AM broadcast transmitters basically fall into three categories: (1) solid-state modulator circuitry supplying the audio modulating power for high-level anode modulation of a power grid vacuum tube RF power amplifier, (2) solid-state, high-level collector modulation of solid-state RF power amplifiers, and (3) a design in which the amplitude modulation process is achieved by pseudo-digital pulse-code modulation taking place directly in the RF power amplifier stages.

The main advantage of solid-state circuitry is operating efficiency. It is still yet to be determined whether the reliability of solid-state designs is distinctly advantageous over vacuum-tube circuitry, but the efficiency advantages are clear and proven. The efficiency of a single power-grid vacuum-tube pulse-width modulator (including filament, control-grid, and screen-grid power losses), is approximately 90% maximum for a 50 kW transmitter. Pulse-step width modulator efficiency for a solid-state modulator at the same power level is approximately 95%.

When solid-state circuitry is also employed in the RF amplifier circuitry, the efficiency is even greater. The typical maximum anode efficiency of a power-grid vacuum-tube class-C amplifier is 84% (neglecting output network circuit losses). This can be improved to approximately 90% by anode-voltage wave shaping techniques, such as third and fifth harmonic traps in the output anode network. Therefore, the all-tube pulse-width modulated transmitter overall modulator/RFPA anode efficiency is 90% × 90%, or approximately 81%. Transmitter designs employing solid-state circuitry in the RF power amplifiers typically use class-D amplifier circuitry which achieve approximately 95% RF collector efficiency at medium wave frequencies. Therefore, an all solid-state pulse width modulated transmitter can have overall collector efficiency of 95% × 95% = approximately 90%, 9% greater than the all-tube transmitter. With energy costs at record high levels in all of the world and constantly rising, the efficiency of high-energy consuming products, such as broadcast transmitters, will continue to be of prime and ever-increasing importance.

Asea Brown Boveri (ABB) Pulse-Step Modulator

The Brown Boveri Corporation of Switzerland, who later merged with the Swedish electrical giant ASEA, was one of the first companies to market a workable very high-power solid-state modulator for use with their family of low frequencyh (LF), medium frequency (MF), and short wave (SW) transmitter products, from 100 kW to 600 kW of carrier power. The circuit employs several low voltage power supplies in series which are switched into operation in steps to provide the required anode modulation voltage, hence the term "step modulation." The concept was first employed in a 500 kW, shortwave transmitter design and later scaled down at LF, and MF, as well as other shortwave products of lower power levels to 100 kW, minimum. Typically the number of steps employed are approximately 28. Each "step" is pulse-width modulated to achieve the necessary total linearity required for broadcast transmitter modulation linearity characteristics. Each step represents approximately 1,000 volts in modulator output, yielding a total modulator output capability of approximately 28,000 volts at the 100% positive modulation crest.

The solid-state devices used to control the 1,000 volt steps are gate-turn-off (GTO) thyristors or SCRs. The transmitters were designed to comply with the CCIR international shortwave broadcast standard of 4.5 kHz audio bandwidth which is also compatible with the requirements for audio bandwidth of the European Broadcasting Union for Medium- and Long-Wave transmissions. Recently, ABB changed their high power switching devices from GTO Thyristors to IGBT (insulated gate bi-polar transistors) because of the latters superior high speed switching characteristics resulting in even higher modulator efficiency and modulation bandwidth capability.

To achieve high overall transmitter efficiency, ABB engineers concentrated not only on the solid-state modulator circuitry, but also on techniques to maximize the anode efficiency of the power-grid vacuum tube final RF power amplifier and associated circuitry. Working closely with the Power-Grid Vacuum Tube Division of ABB, they designed a tube with optimum circuit efficiency characteristics which operates with minimum grid drive power requirements an anode conduction angle of approximately 110° (rather than the more conventional 120° standard for class-C operation) which yields a total anode efficiency of approximately 87%, excluding circuit losses. Medium-wave transmitter designs by ABB, using pulse-step modulation, employ similar technologies with the addition of anode wave-shaping circuitry, resulting in anode efficiencies exceeding approximately 92%. Overall shortwave transmitter efficiency has been measured at greater than 76% on some international shortwave bands. Overall efficiency of ABB medium-wave transmitters has been measured at greater than 77%.

Continental Electronics Corporation (CEC) Pulse-Step Modulator

Continental Electronics Corporation introduced their own version of a high-power pulse-step modulator system in 1990. The system is similar to the ABB system described above in basic operating principle. The technique used to generate each step and the subsequent pulse-width modulation of each step is different than the system employed by ABB, but the net result is similar.

The CEC system for the basic step generation uses a modified flash A/D converter with an additional digital "ring" modulator system that electrically rotates the active steps equally among all available modules. This rotation provides equal power distribution among all system modules. Each step is pulse-width modulated by a linear mixing of a triangular "dither" signal to each A/D flash comparator.

The CEC system employs typically 48 discrete steps of approximately 1,500 volts each, and has been built for power levels of 100 kW, 250 kW, and 500 kW transmitters, both medium wave and shortwave. The CEC design employs high speed IGBT (insulated gate bi-polar transistors) producing typical modulator efficiencies of 95% or higher, overall transmitter efficiencies of greater that 72%, and modulation bandwidth capabilities exceeding 10 kHz.

Nautel Corporation—AMPFET Transmitters

The Nautel Corporation of Nova Scotia, Canada was one of the first to introduce all solid-state transmitter designs on the commercial broadcast market. Initially available for only 1 kW, 5 kW, and 10 kW carrier powers, they have since added capability up to 50 kW.

The Nautel design employs modules of approximately two kilowatts, each, with class-D RF power amplifiers using pulse-width modulation of the class-D amplifier collector voltage. Each module contains its own RF power amplifier (RFPA) and pulse-width modulator and a number of modules, depending on the power level, that are combined in an RF-combining network to achieve the desired carrier-power level. The collector efficiency of each class-D RFPA is approximately 95%, and the efficiency of the combining networks is greater than 99%, yielding a total RF system efficiency of approximately 94%. The efficiency of the pulse-width modulators are approximately 95%, resulting in an overall collector efficiency of better than 89%. Nautel engineers then carefully selected and designed the remaining transmitter components to require no more than approximately 10% of nominal carrier power for all auxiliary transmitter functions, such as control circuitry, exciter and driver circuitry, cooling equipment, etc., thus achieving an overall transmitter efficiency of greater than 80%.

The basic circuit of the Nautel design is shown in Fig. 10.

Polyphase Pulse-Width Modulation

In the late 1960s, Harris developed pulse duration modulation (PDM), a technique that offers improved efficiency and audio quality while eliminating some large and costly components from the transmitter. In the original configuration, audio information to be transmitted is sampled at a nominal 75 kHz rate. This creates a 75 kHz pulse train with its pulse width rate-of-change being equal to the audio frequency. Studies have determined that the time duration of the sample pulse is not critical. The sample rate is based on an upper frequency of 15 kHz sampled at five times that frequency to ensure broadcast-quality bandwidth and distortion. In practice, most transmitters manufactured for use in North America employ a switching frequency of approximately 70 kHz, to simplify compliance with the FCC requirement for suppression of spurious emissions beyond 75 kHz from the carrier frequency.

The technique of polyphase pulse-width modulation (PPDM) was a further development to overcome the switching transition losses of the higher sampling frequency. Fig. 11 shows the relations of the four-phase system operating at a fundamental rate of 60 kHz. The effective sample rate is four times that or 240 kHz.

When extracting the long and short term DC average information from an encoded single-phase PDM signal, a sharp filter roll off characteristic is required. Such a characteristic is shown by the solid line curve in Fig. 12. This sharp attenuation characteristic generates overshoot due to the inherent group delay characteristic of the sharp roll off filter. The effective 240 kHz sampling rate of the Harris four-phase sampling system, however, allows some improvements in filter design that significantly reduce overshoots.

The resultant 240 kHz four-phase switching frequency being farther removed from the audio baseband than the 75 kHz single-phase PDM sampling rate has facilitated the design of a phase linear PDM filter. This new Harris PPDM filter reduces the group delay just mentioned. The resulting softer roll-off characteristic is depicted by the dotted line response curve of Fig. 12. Reference to Fig. 13 shows the relative distortion amplitudes and spectra of 75 kHz and effective 240 kHz sampling rates. Inspection of this figure shows the added advantage of less distortion energy present in the demodulated audio baseband, which is due to the effective sampling rate being farther removed from that baseband than previously.

Attenuation of 60 kHz energy occurs within the PDM filter in two ways: the normal filter attenuation mechanism, as well as a capacitive coupling technique. The capacitors C1, C2, C3, and C4 emanating from each of the PDM filters in Fig. 14 tie together at a common terminal. This configuration in effect connects the 60 kHz energy present in the individual filters to a common summation point. As the spectral energy created by the sampling process creates energies that are 180° out of phase with each other, and assuming that a balanced amplitude among sampling phases exists, the capacitive coupling of 60 kHz energies will result in a high degree of cancellation.

The output of the PDM filter is a combination of long-term and short-term average information. Transmitter output power is determined by the long-term average voltage; modulation power is determined by the short-term average voltage.

Harris Broadcast Division, DX Series Pulse-Code Modulation

One of the newer improvements to medium-wave AM broadcast transmitters employs direct digital techniques to generate the amplitude modulated wave. The technique has no "modulator" as such and the complete AM wave, carrier and modulation sidebands, are produced by direct digital pulse-code control of several low-power solid-state RF amplifier modules.

Invented by Hilmer Swanson of Harris Broadcast Division, the DX system is employed in transmitters ranging in power from five to 50 kW for the domestic U.S. market, and has been sold at 100 kW power levels for the CCIR Region 2 and 3 markets in which carrier powers up to one and two megawatts are not uncommon.

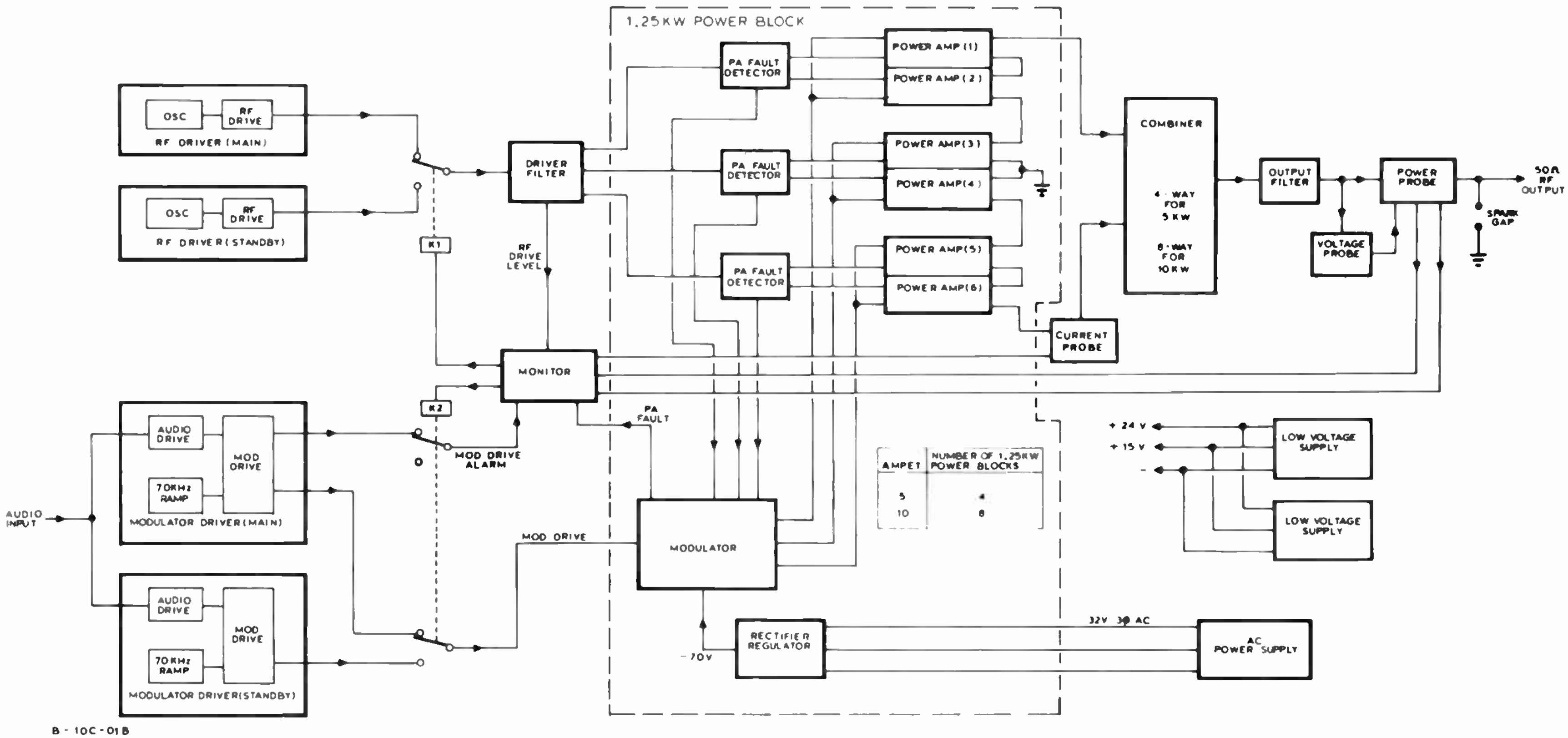


Figure 10. Block diagram, Nautel Model AMPFET 5 and AMPFET 10, 5 & 10 kW AM broadcast transmitters.

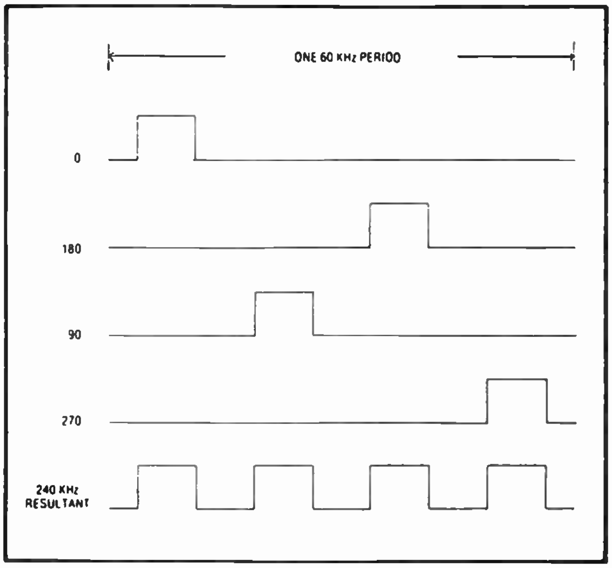


Figure 11. Effective 240 kHz sampling rate.

The system has many inherent advantages over conventional AM modulation systems:

1. It is solid-state and employs no separate modulation of the final RF amplifiers; i.e., the amplitude modulated wave is achieved directly digitally by selecting series-connected RF output modules.
2. Having no audio modulator, it is capable of high quality modulated wave characteristics with lower modulation noise, distortion, and transient modulation characteristics.
3. Being basically a digital system, it is capable of receiving digital audio input directly from any variety of digital audio sources.

The Harris DX system is shown in Fig. 15. The incoming analog signal is converted to digital code by conventional sample-and-hold and pulse-code-modulation (PCM) techniques and circuitry. The top six (most significant) bits of a twelve-bit analog-to-digital PCM system are converted into 64 equal steps. The top 22 steps are not used, leaving 42 equal steps which control 42 equal-output series RF amplifier modules. This is equivalent to keeping the lower 5.4 of the 6 top bits of the PCM digital code; i.e., $2^{5.4} \approx 42$. The bottom 6 (least significant) bits are then used to control six half-step RF amplifiers; i.e., amplifiers with outputs of ½, ¼, ⅛, 1/16, 1/32, and 1/64 of the major 42 equal step amplifiers. In effect, this creates a digital reconstruction of the analog input waveform of 2709 discrete steps; i.e., $(42 + 1) \times 63 = 2709$. This is roughly equivalent to a pulse-code-modulation system with a total resolution of 11.4 bits; i.e., $2^{11.4} \approx 2702$.

Such a PCM system[Panter13] yields a theoretical harmonic distortion value of approximately 0.03% and audio noise of less than −70 dB referred to 100% modulation. Typical measured values of total harmonic distortion (THD) are approximately 10 to 30 times this theoretical figure due to the practical difficulty in achieving balanced output contribution from all of the series output modules. However, even these levels of

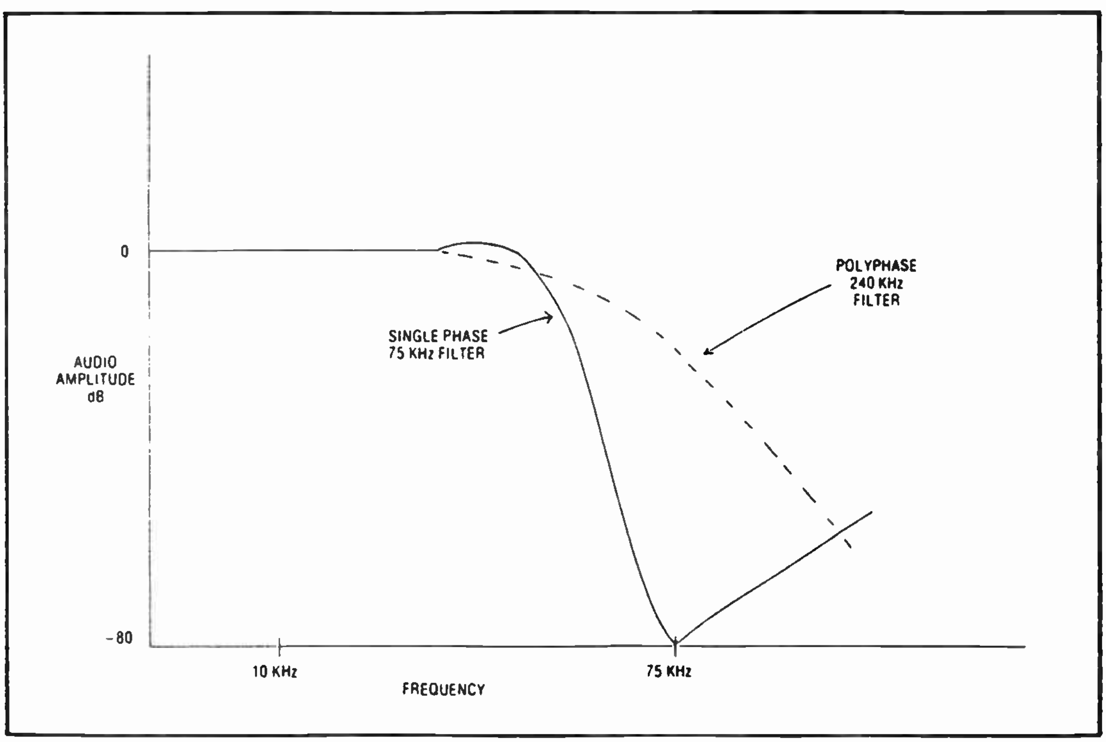


Figure 12. Comparative PDM filter responses.

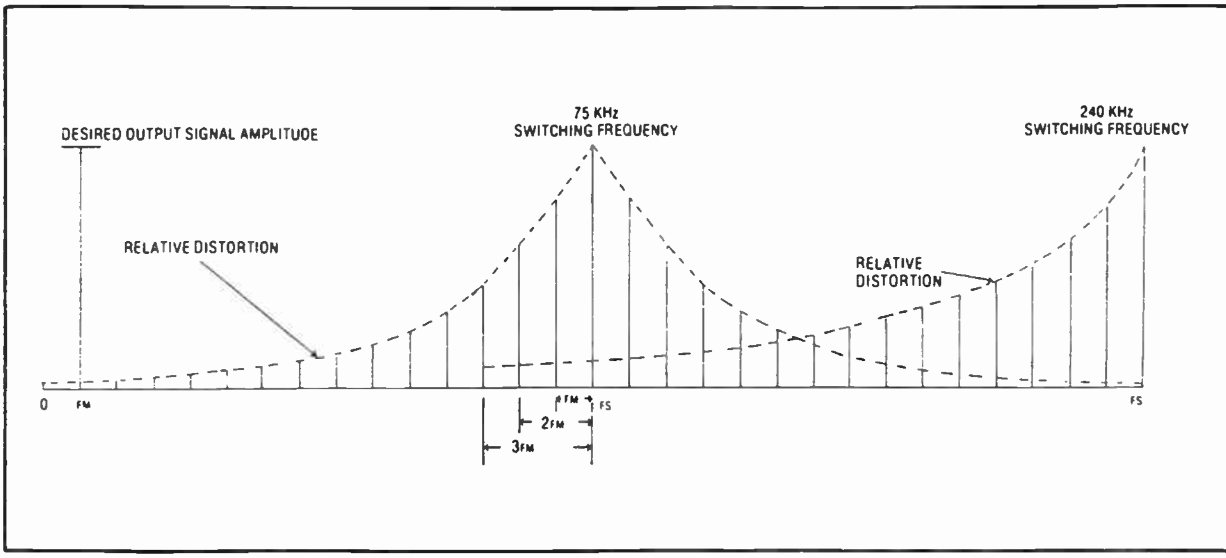


Figure 13. Switching modulator spectrum.

harmonic distortion (less than 1% maximum) are well below accepted norms for modern high-power AM broadcast transmitters. Typical values of measured audio noise are very close to the theoretical maximum of −70 dB referred to 100% tone modulation at full carrier power.

Each high-power amplifier module is a solid-state amplifier operating in the class-D mode, having approximately 95% power efficiency, including output circuit losses. Since there are no additional modulator losses (because the modulation process takes place in the RF circuitry itself, as described above) the auxiliary

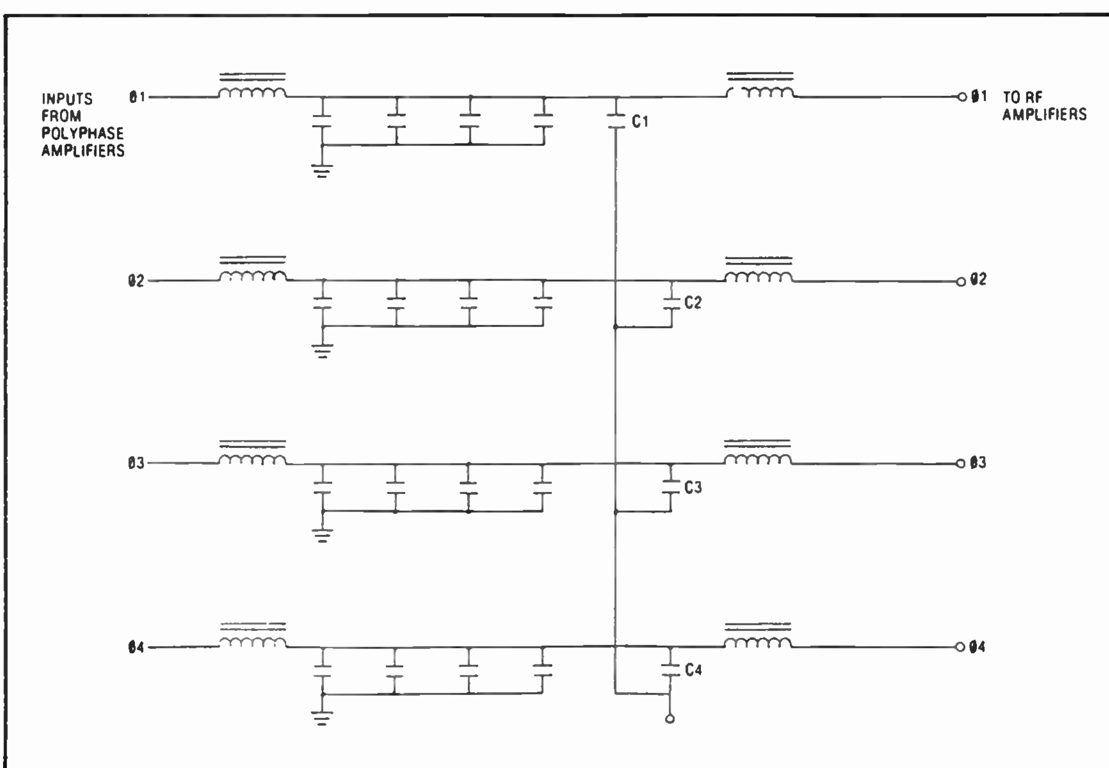


Figure 14. Polyphase PDM filter.

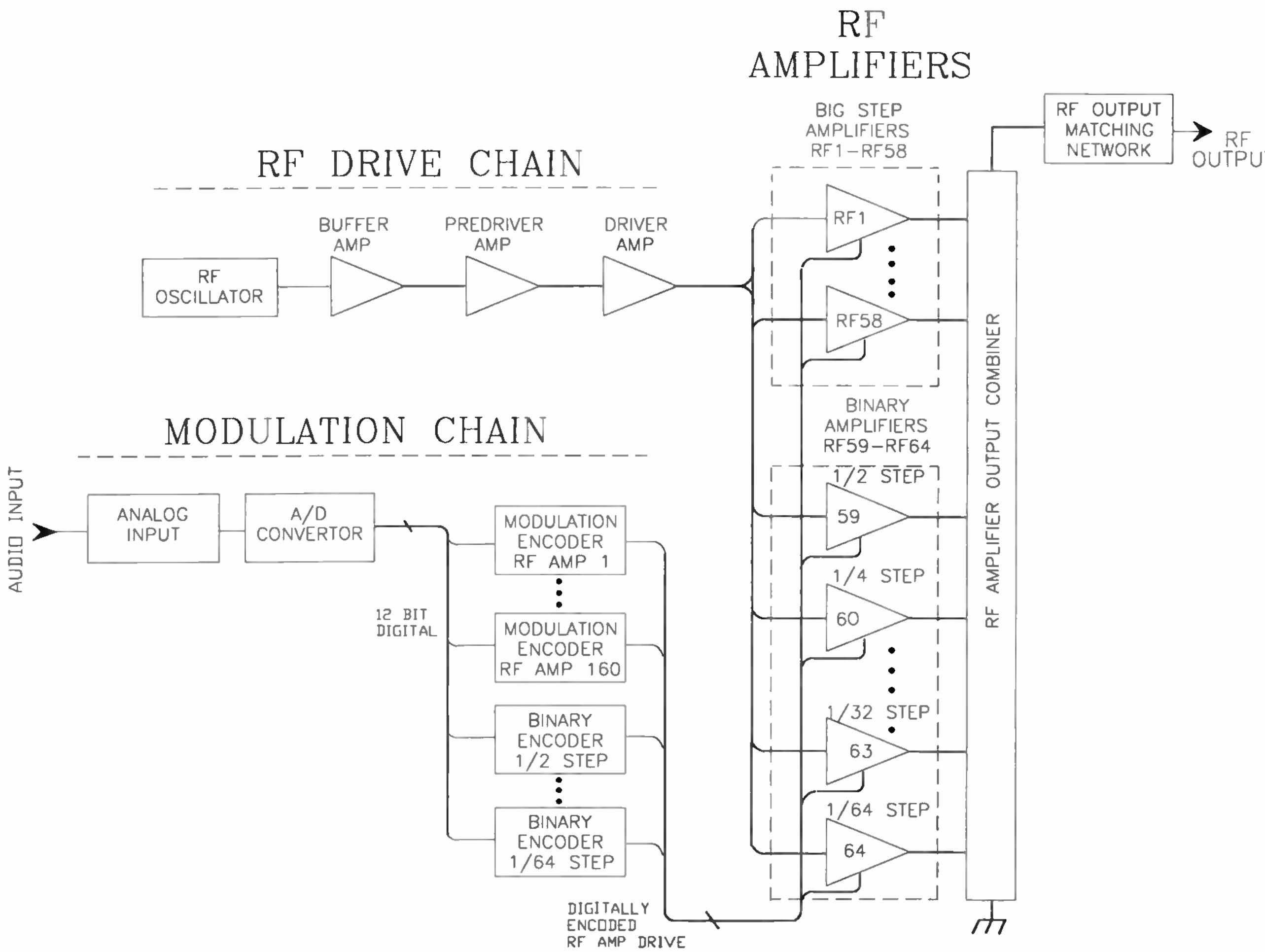


Figure 15. Harris DX system.

circuits for cooling, control, excitation, and drive represent the only other input power requirements. The total auxiliary input power of a 50 kW transmitter is approximately 5 kW. This means that the total input power of a 50 kW transmitter at carrier-only power (no modulation) is approximately: 50/0.95 + 5 = 57.63 kW, yielding an overall efficiency of approximately (50/57.63) x 100% = 86.76%. Harris Broadcast Division claims 87% or greater overall efficiency for all their DX series transmitters greater than 5 kW carrier power.

Kahn Power-Side System

"Power-side" refers to an asymmetric modulation system wherein the power in one sideband is three times that in the other, and the other sideband is reduced to maintain the maximum envelope modulation of +125%. All preemphasis is removed from the stronger sideband and concentrated in the weaker sideband. Since the strong sideband is free of preemphasis, listeners with continuously-tunable receivers or digitally-tuned radios with less than 3 kHz steps may *sideband tune*.

The advantages of sideband tuning is that, by tuning for the loudest signal, the listener may tune away from adjacent-channel interference. Also, the receiver's IF passband is centered more closely on the stronger sideband than on the carrier, increasing the apparent fidelity of the receiver. Furthermore, the proper tuning range for a power-side signal is less sensitive, allowing over five times the normal angular displacement of tuning. All types of receivers, including 10 kHz step digital receivers, should benefit from this asymmetry under adverse conditions such as selective fading, distorted antenna nulls, distortion from reradiation from buildings and power lines, and so-called *carrier-beat co-channel* interference.

TRANSMITTER CIRCUITRY

Detailed transmitter circuit design is as varied as the individual designers. There is, however, some basic circuitry common to all transmitter types and models will be briefly discussed in this section.

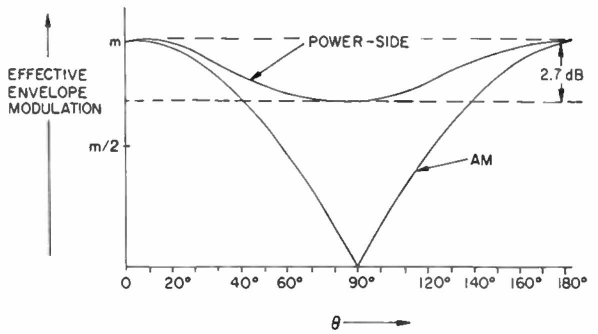


Figure 16. Envelope modulation of Kahn Power-Side compared to conventional AM.

Carrier Frequency Generator/Exciter

The stable frequency source for virtually all AM transmitters has been the quartz crystal oscillator. The quartz crystals used in older model transmitters were large cuts of natural quartz, vacuum-sealed in glass envelopes, similar to small-power vacuum tube envelopes, and occasionally mounted in temperature-controlled ovens to obtain the required FCC carrier frequency stability. In modern designs the quartz crystals are enclosed in small, hermetically sealed metal cans made popular and proven reliable by the military and commercial communications equipment industries. These lack special temperature controlled circuitry. Modern quartz crystal manufacturing technology and solid-state oscillator designs allow these types of small metal-sealed units to maintain adequately the FCC's current requirement of ±20 Hz carrier frequency tolerance. Stability of the quartz oscillator circuits is normally adequate over the full lifetime of the equipment. Should frequency adjustment ever be required to bring a unit back inside the FCC limits, mechanically stable adjustment components, usually a glass or ceramic piston-type of capacitor, are provided for use by a qualified station engineer using calibrated frequency measuring equipment. (FCC Rules Sections 73.1540 and 73.1545.)

Exciters for all proposed AM-stereo systems provide the carrier frequency excitation for the transmitter as well as the stereo generating circuitry. Some manufacturers of AM-stereo exciters generate the desired carrier signal with frequency synthesizer techniques. This method is generally equal to or better than the discrete quartz oscillator method with regard to frequency tolerance, but can produce higher phase modulation noise if the circuits are improperly designed or adjusted.

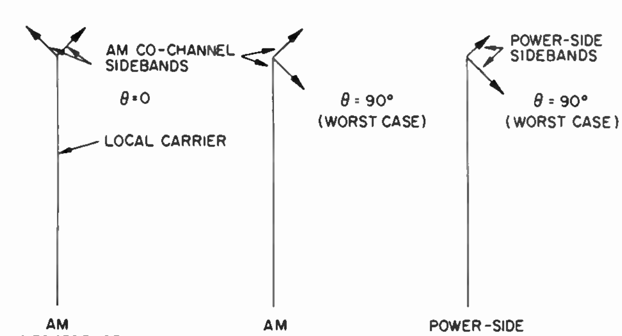


Figure 17. Interference phase relations of Kahn Power-Side and conventional AM.

RF Power Amplifiers

As discussed in earlier sections, all modern AM broadcast transmitters using power-grid vacuum tube final RF amplifiers employ class-C amplification. Some employ designs using third and fifth harmonic trap circuitry yielding quasi-class-D operation for improved efficiency. All solid-state transmitters currently available in the North American Market use class-D RF final power amplification.

RF Output Networks

The purpose of the RF output network is to match the impedance of the load (the common-point impedance of one or more antenna matching and combining networks) to the impedance required by the final RF power amplifier tube(s) or transistor(s) in order to produce the desired carrier and sideband power. The output network circuit is also designed to provide the attenuation characteristics necessary to meet FCC requirements for spurious and harmonic output. There are many techniques to accomplish these basic tasks. One simple and effective method is shown in Fig. 18. The terminating load impedance for the network shown in Fig. 18 is defined by the Smith chart representation of Fig. 19, which is a typical common point impedance characteristic of a multi-tower AM broadcast directional antenna array.

Fig. 20 shows the impedance-versus-frequency characteristics at the input of the network (at the anode of the final RF amplifier tube). The shape of the impedance curve in Fig. 20 differs significantly from the shape of the terminating impedance curve shown in Fig. 19 due to the narrowing of bandwidth caused by the output matching network. The voltage standing wave ratio (VSWR) (normalized to the resistive carrier impedance) at ±10 kHz is about 1.2 in both Figs. 19 and 20, but at ±50 kHz the VSWR at the input to the network has increased from about 2.4 to approximately 3.0. The shape of the impedance-versus-frequency curve at the anode(s) of the RF output amplifier tube(s) or transistor collector(s) greatly affects the high audio frequency performance characteristics of the transmitter, such as frequency response and harmonic distortion.[14] The shape of the impedance versus frequency curve should be symmetrical about the resistive axis of the Smith Chart and should yield lowest practical VSWR values at the highest expected fundamental sideband frequencies.

Fig. 21 shows the result of an attempt to lessen the magnitude of the mismatch at the input to the matching network by the addition of the dotted components

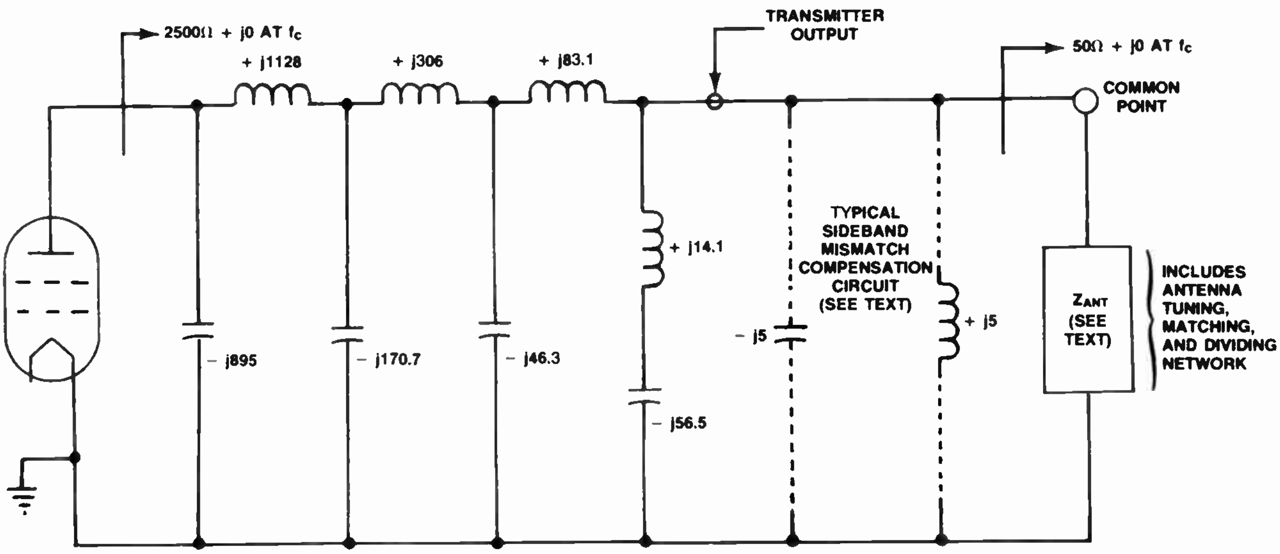


Figure 18. Typical output matching network yielding 80 dB or better harmonic attenuation and a compensated symmetrical sideband load at the anode(s) of the RF output amplifier.

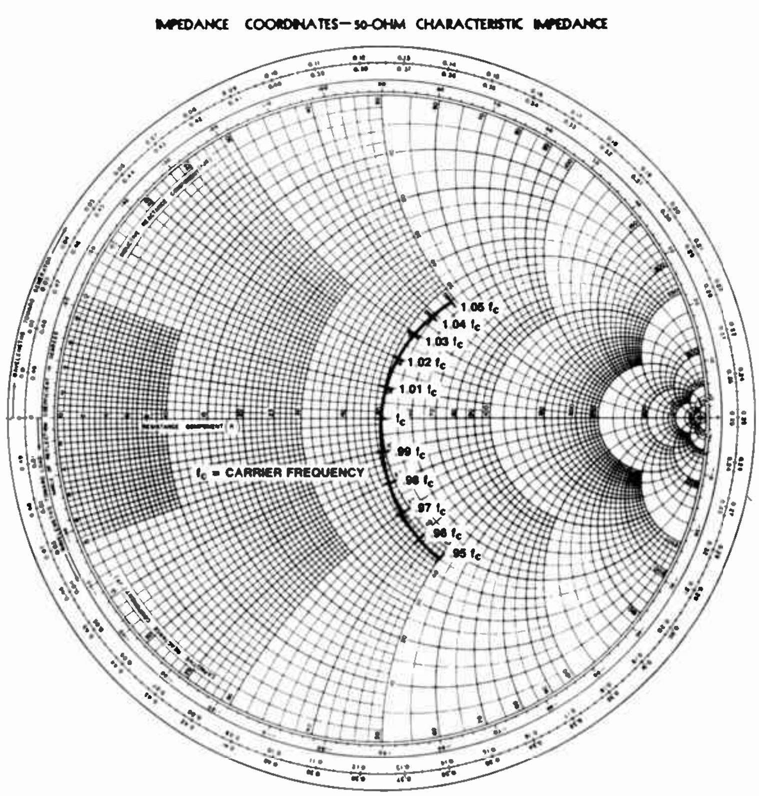


Figure 19. Typical impedance versus frequency characteristic of antenna common point.

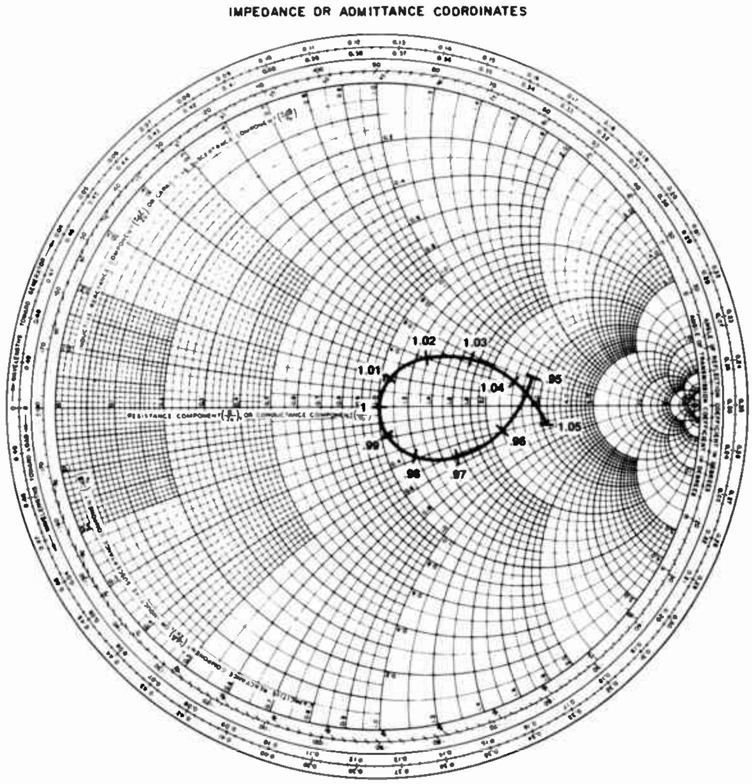


Figure 20. Impedance versus frequency characteristic at anode of final RF power amplifier.

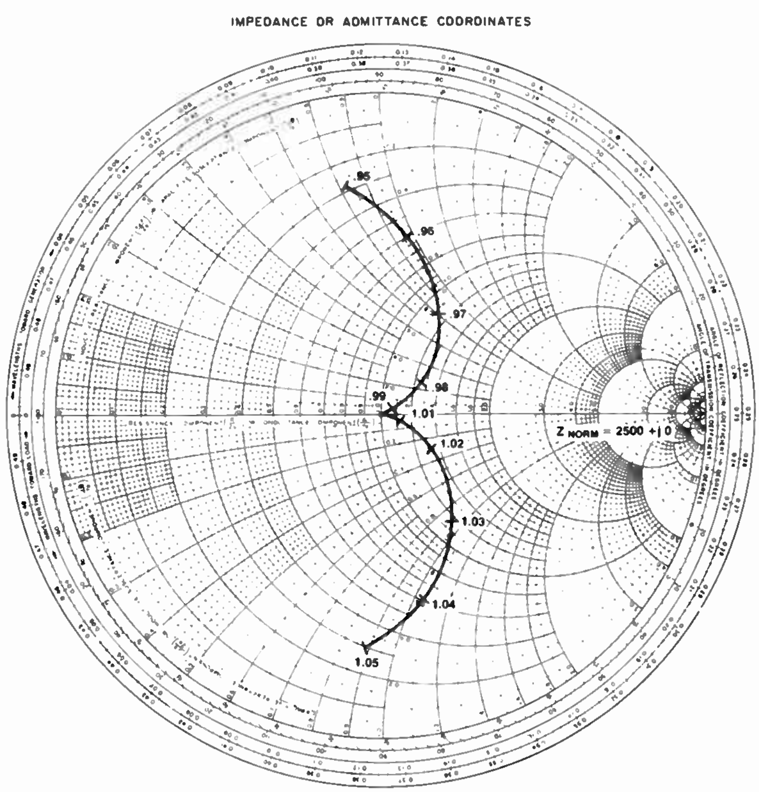


Figure 21. Impedance versus frequency characteristic at output anode with sideband mismatch corrective network added.

shown in Fig. 18. The result is a significant reduction in tube anode load mismatch at sideband frequencies of ±10 kHz; from about 1.2 VSWR in Fig. 20 to approximately 1.1 VSWR in Fig. 21. The reduction of mismatch at ±10 kHz, however, comes at the expense of an increasing mismatch at sideband frequencies greater than ±30 kHz. The increasing mismatch at higher sideband frequencies is normally not a problem and often tends to improve transmitter harmonic distortion performance, acting as an output bandpass filter to higher-order harmonics. There is, theoretically, no limit to the amount of sideband mismatch correction that can be achieved with more complex network circuitry.

The most important criterion is symmetry about the Smith Chart resistive axis. Station engineers or their engineering consultants should work with the transmitter manufacturers and antenna system designers to determine and achieve the desired impedance-versus-frequency curve at the output RF amplifier tubes anode connection. The example given above for load sideband mismatch correction is for illustrative purposes only. It is quite common for the complex impedance at the sideband frequencies to present a more severe mismatch at the antenna common point than that shown in Fig. 19. More severe mismatch and impedance asymmetry requires more complex sideband-mismatch corrective networking.

Transmitter Control and Monitoring

AM broadcast transmitter control circuitry is normally very basic and uncomplicated. It is now common to find transmitter control performed with discrete digital IC logic circuitry instead of simple relay control logic. Some manufacturers are incorporating microprocessor technology in their latest equipment designs to replace discrete digital logic circuitry. The aim of manufacturers is normally to provide the operating engineers and technicians with the most basic, reliable, and easy-to-maintain and troubleshoot transmitter possible. Experience has shown that well-designed relay control logic, discrete digital IC logic, and microprocessor-based logic are all about equal in terms of reliability and ability to perform the required transmitter control functions.

Future microprocessor-based transmitter control logic promises: (1) to provide self- and remote-assisted diagnostics of transmitter problems and remote interrogation of transmitter operating parameters; (2) to alter the basic characteristics of a transmitter control system through software control; and (3) to allow the basic transmitter design to be more easily customized to individual users' operating requirements. Because of the high voltage and high current faults that can exist in any high-power transmitter component, extra care must be taken by designers and manufacturers to prevent the potential destructive energy in these faults from affecting the performance and operation of the relatively delicate solid-state control logic circuitry.

High speed vacuum contactors and solid-state regulator/controllers are used in modern designs to control the high voltages and currents encountered in all levels of high power AM transmitters, which previously had been controlled by slower, though equally reliable, air-magnetic contactors and relays.

Remote control systems are commercially available that will allow remote control into almost any transmitter, old or new. These systems, like the basic transmitter control system, use relay, discrete digital IC, microprocessor-based logic circuitry, or combinations of these, depending on the manufacturer of the control equipment and the complexity of the remote control functions desired. Many transmitter equipment manufacturers also provide limited built-in remote control functions and circuitry for their products.

The FCC requires manufacturers of AM broadcast transmitters to provide certain basic monitoring functions for proof of performance parameters. These requirements are contained in the FCC Rules Section 73.1215.

High-Voltage Power Supplies

High-voltage power supplies in AM broadcast transmitters must provide acceptable performance in two basic areas: (1) power supply ripple, which affects transmitter hum and noise output, and (2) dynamic regulation, which affects low frequency modulation transient response. It is typical for transmitters of five kilowatts carrier power and lower to operate from single-phase ac power sources, usually 240 volts in North America. Transmitters with carrier power ratings of 10 kW operate with single-phase or three-phase power source, depending upon the manufacturer of the transmitter. Transmitters with carrier power ratings of greater than 10 kW operate only from three-phase power source, usually 480 volts for power levels up to 100 kW and 4,160 volts or 11,000 volts in North America for higher carrier power levels.

Three-phase power has the advantage of being easier to filter and usually provides better dynamic regulation of critical modulator voltages than single-phase supplies. Single-phase power is more readily available, which is the only reason why it is used at the lower transmitter power levels, because installation cost would be disproportionately increased were three-phase power required. Single-phase rectifier power supply systems generally require inductance-capacitance (L/C) filtering to provide low ripple output for low transmitter hum and noise specifications. However, L/C filtering creates power supply resonances in the audible to sub-audible range of modulating frequencies and therefore is a source of poor dynamic power supply regulation when the modulator circuitry is excited by vowel sounds or musical percussion sounds. It has been common since about 1970 for higher-power transmitters to use special high voltage supply transformers to generate a six-phase ac supply from the basic three-phase power source. The six-phase supply, when full wave rectified, yields a 12 pulse rectified dc waveform that has both lower ripple content and a higher ripple frequency than conventional three-phase full wave rectification. As a result, the output of the rectifier can be sufficiently filtered with no additional filter inductors, thus improving low audio modulating frequency dynamic power supply regulation and hence, low frequency transient distortion.

TRANSMITTER PERFORMANCE MEASUREMENTS

AM transmitter performance parameters should be measured on a periodic basis in order to assure that certain minimum broadcast quality standards are provided to the listening public. Excluding the performance standards for AM stereo, these are listed below. Some are specifically required by the FCC Rules, while others are merely a matter of good engineering practice.

- Operating power [Section 73.51]
- Carrier output power delivered to the antenna system [Section 73.54]
- Modulation capability
- Total audio frequency distortion
- System frequency response
- Carrier-amplitude regulation (carrier shift)
- Hum and noise output level
- Carrier frequency tolerance [Section 73.1545(a)]

All audio measurements should be made from a demodulated voltage or current sample at the antenna system common point.

Output power measurement (the measurement of transmitter power output by the direct method described in Sections 73.51 and 73.1215) is subject to more than 13% error if the measurement inaccuracies allowed by the FCC are taken to the limit. For example, assume that a common point impedance of 50 ohms resistive can be measured to within 2% accuracy with a RF impedance bridge (a realistic tolerance). Further, assume a direct-reading RF ammeter having a full scale reading of 100 amperes and an FCC allowed tolerance of ±2% of full scale is used and indicates 33.33 amperes of common point current, just meeting the minimum FCC accuracy and indication requirements in Sections 73.1215(b)(2) and (3). Under these conditions, the actual power delivered to the antenna is between the limits of (35.33 × 35.33 × 51) = 63.65 kW, as a maximum, and (31.33 × 31.33 × 49) = 48.10 kW, as a minimum; yielding a total measurement error of approximately +14, −13%. Using an RF ammeter with a 50 ampere full scale reading, the same FCC allowed inaccuracies, and the same meter indications as above, would result in power output measurement errors of approximately ±8%.

FACTORY TESTS

When buying an AM broadcast transmitter it is advisable to ask manufactures for specific, detailed test and performance data at the start of the investigation. Before a final decision is made, detailed tests should be made at the manufacturers factory under strict control of an experienced engineer or engineering consultant. Most manufacturers of AM broadcast equipment welcome this kind of intelligent approach. Hints on what kind of tests and a discussion on the details of each test are given below.

Audio Frequency Response

AM transmitter frequency response characteristics, referred to a reference frequency of either 400 Hz or 1,000 Hz, typically is well within ±1 dB from 50 to 10,000 Hz at any depth of modulation, exclusive of antenna system characteristics which can and often does have a deleterious effect on many audio performance characteristics. Practically all mass production AM receivers made in any country have (IF) intermediate frequency and audio amplifier bandpass characteristics that limit receiver −3 dB high-end audio frequency response to between approximately 2500 Hz and 5000 Hz, with 2500 Hz the more common of the two figures. Typical low end −3 dB frequency response of consumer AM radios is between 100 Hz and 300 Hz, with approximately 200 Hz a common value.

Audio Harmonic and Intermodulation Distortion

AM broadcast transmitters typically produce less than 2% total harmonic distortion (THD), up to 90% modulation at any frequency of modulation between 50 Hz and 10,000 Hz for monophonic transmission. Most of the modern all solid-state digital systems produces typical audio distortion of less than 1% THD.

Intermodulation distortion (IMD) has been known for years and has been documented in technical audio journals to be more disturbing than harmonic distortion, though both are important. The CCIR method of IMD measurement is preferred for radio broadcast transmitters. With this method, two equal audio tones separated by 170 Hz are fed to the transmitter, and the peak modulation level is adjusted to between 85 and 95% modulation. The level of odd and even-order products are measured using an audio-wave or spectrum analyzer connected to the test output terminals of a high quality modulation monitor.

Two IMD measurements should be taken, one with the two tones near mid-audio band (for example 400 Hz and 570 Hz) and one with the two tones near the upper audio end (for example 7,000 Hz and 7,170 Hz). High quality broadcast transmitters should produce IMD products more than 30 dB below the level of either of the two modulating tones. The RSS value of all IM products, relative to the level of either modulating tone, should also be less than 7% at 90% peak modulation levels.

Residual AM Hum and Noise

AM noise of 60 dB below the 400 Hz/100% modulation level can be achieved by most current production AM broadcast transmitters. The bandwidth of noise measuring equipment should be 20 kHz. Typical modulation monitor demodulated audio bandwidth is approximately 25 kHz to accommodate FCC bandwidth requirements.

Residual PM Hum and Noise

Residual phase modulation (PM) noise is normally not a problem for modern AM broadcast transmitters designed for monophonic broadcasting. Quartz crystal oscillator circuits and even moderately-careful RF component mechanical designs produce quite acceptable phase noise characteristics. However, for AM-stereo applications, where a frequency synthesizer may be used for RF signal generation, is it wise to test the purity of the synthesizer circuitry. Excessive residual PM noise can convert to AM noise over certain nighttime propagation paths and appear to distant listeners as objectionable AM hum and noise.

An acceptable value of PM noise is −25 dB RMS relative to one radian peak, measured in a 15 kHz bandwidth, for monophonic, medium wave band AM transmission. Transmitters with quartz crystal oscillator exciters typically exceed this recommendation by 25 dB or more. Transmitters used for international short wave broadcasting (4 MHz to 26 MHz) require PM noise levels of approximately −45 dB relative to a one radian peak because of more severe sky wave PM to AM conversion at higher frequencies.

Incidental Phase Modulation (IPM)

Like residual PM, IPM is more important in stations using or planning to use AM stereo transmission. IPM is defined as the peak phase deviation of the carrier frequency (in radians) resulting from amplitude modulation. IPM values of several radians were common in the very early days of broadcasting. Typical values of IPM for modern transmitters that have not been specifically designed or adjusted to minimize IPM range from about 0.1 to 0.5 radians peak (approximately 6° to 30°). A maximum acceptable value of IPM required for present or future AM stereo operation is generally considered to be approximately 0.03 peak radians with a desired value of better than 0.012 peak radians. State-of-the-art modulation meters such as the Hewlett Packard model HP-8901B and other similar instruments are preferred for accurate PM and IPM

Carrier Amplitude Regulation (Carrier Shift)

Carrier-level shift in a given transmitter has more importance to overall transmitter performance than many broadcast engineers realize. Large values of negative carrier shift can have as much effect on effective transmitted sideband power as poor transient overshoot distortion. The term carrier shift may be somewhat confusing, especially to newcomers in radio broadcasting who often equate the terminology with frequency shift instead of level or amplitude shift; hence, the new terminology of carrier amplitude regulation. The CCIR refers to the same characteristic as carrier-level shift, which appeals to many engineers because of its closer adherence to the original terminology, but without the ambiguities. Carrier-level shift is the effective shift in apparent carrier-level due to the amplitude modulation process.

Carrier-level shift can be caused either by poor power supply regulation or by modulation even-order harmonic distortion (which generates a dc offset component in the modulated RF envelope) or by both. Carrier-level shift can be either positive or negative, although it is usually negative. This is because power supply regulation is most often the major source of carrier-level shift and power supply regulation is most generally negative in sign (lower voltage output at higher current loads). Poor power supply regulation is not always caused by the transmitter power circuitry; it can also be, and often is, caused by poor supply-line voltage regulation or, more generally, a combination of the two.

Another common misconception regarding carrier-level shift is that the defined level shift is direct shift in carrier power. Actually, carrier-level shift is defined as the shift in effective carrier *voltage or current* due to modulation. This means that a carrier-level shift of −5% is equivalent to a carrier power shift of approximately −10% i.e.,

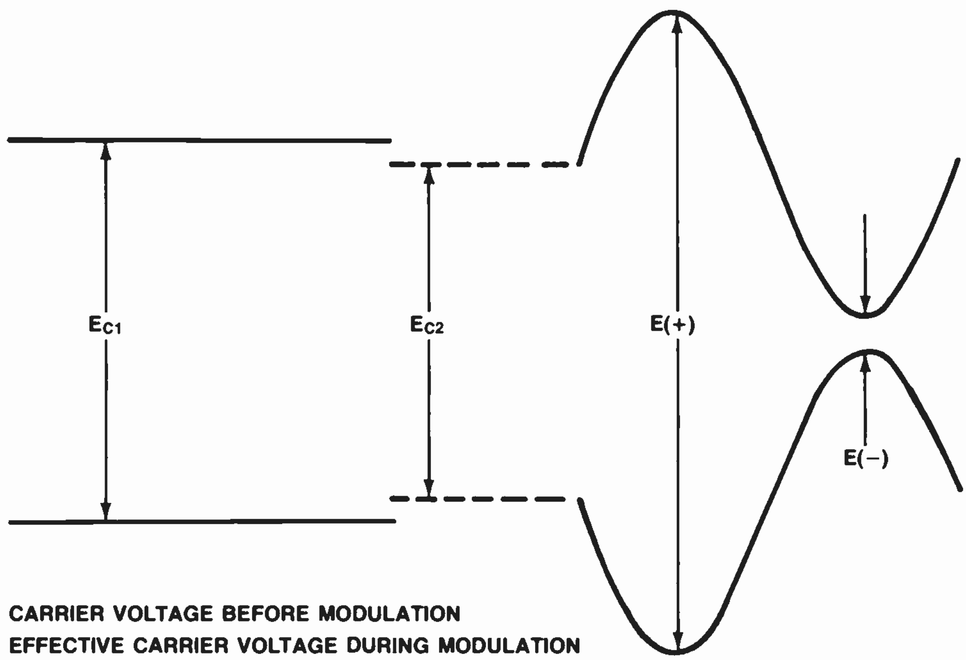
$$P_{carrier\,(mod)} = \text{Pcarrier} \times (1 - 0.05)^2$$
$$= 0.9025 \times \text{Pcarrier}$$

Transmitters having no carrier-level shift produce an average output power of 1.5 times the carrier power level with 100% sinusoidal tone modulation. Transmitters that exhibit 5% carrier level shift produce an average output power only 1.35 times the carrier level at the same conditions of symmetrical sinusoidal tone modulation. A broadcast station engineer or engineering consultant should completely understand these and other equally important transmitter characteristics in order to make an intelligent buying decision based upon measurable and proven technical merit. Fig. 22 gives a graphical representation of carrier-level shift.

Audio Phase Linearity

Proper attention is given to phase linearity by most station engineers and engineering consultants. Station managers and engineers are usually concerned about the "sound" of their stations. Loudness, or perceived loudness, is a common criteria of quality in many stations with diverse programming formats. However, it is not uncommon for station engineers to spend more time researching the merits of program processing equipment than the one characteristic in their transmitter purchase that could partially neutralize any advantage from a program processor. That characteristic is audio phase linearity.

Audio phase nonlinearity and its major detrimental result, transient overshoot have been discussed in previous sections of this chapter. Modern programming philosophy and program processors have made trans-



E_{C1} = CARRIER VOLTAGE BEFORE MODULATION
E_{C2} = EFFECTIVE CARRIER VOLTAGE DURING MODULATION

$$\text{CARRIER LEVEL SHIFT (CS)} = \frac{\dfrac{E(+) + E(-)}{2}}{E_{C1}}$$

$$\text{\% MODULATION} = \frac{E(+) - E(-)}{E(+) + E(-)} \times 100$$

Figure 22. Graphical representation of carrier level shift and formulas for calculation of carrier level shift and % modulation.

mitter audio phase linearity characteristics an important performance criteria.

In the late 1960s, the FCC authorized 125% positive program modulation, allowing AM transmissions to accommodate certain naturally-occurring asymmetries in voice and music, thus achieving a gain of 2 dB of real loudness or actual program side-band power (20 log(1.25) = 2 dB). As stated earlier, some AM broadcast transmitters in present use exhibit as much as 12% overshoot of a square wave input due to phase nonlinearity. This has the effect of taking away one of those two dB (20 log(1/1.12) = −1 dB).

A simple way to determine the effects of phase nonlinearity is to require the transmitter manufacturer to demonstrate rectangular-wave modulation characteristics of the transmitter. When such a test is performed, the overshoot will be directly visible and measurable on the oscillographic display of the output envelope. The station engineer should also investigate the effects that antenna system phase nonlinearity may have on the total system transmission transient characteristics.

Occupied Bandwidth

Occupied bandwidth of an AM broadcast transmitter can best be measured with a band-limited, colored Gaussian noise source (similar to pink noise used in certain acoustical tests) to provide a continuous wideband modulating signal. Pink noise has equal energy-bandwidth ratio (equal energy per octave, per third octave, per tenth octave, etc.). White noise has equal energy per bandwidth (equal energy per Hz). Both noise signals can have a gaussian or pseudo-gaussian probability density function which closely resembles that of voice and music. Fig. 23 shows a block diagram of the test procedure. The reader is also referred to CCIR recommendation 559–1 (Vol X, Part 1, Geneva, 1982).

The measurement of occupied bandwidth is a dynamic test which effectively summarizes two important transmitter static parameters, audio nonlinear distortions (which are the source of IMD and THD) and incidental phase modulation (IPM).

Harmonic and Spurious Output

Harmonic and spurious output of an AM broadcast transmitter or transmission system can only be measured effectively in two ways:

1. Use a sample of the transmitter RF output when operating into a dummy load with a known or accurately-measurable impedance characteristic

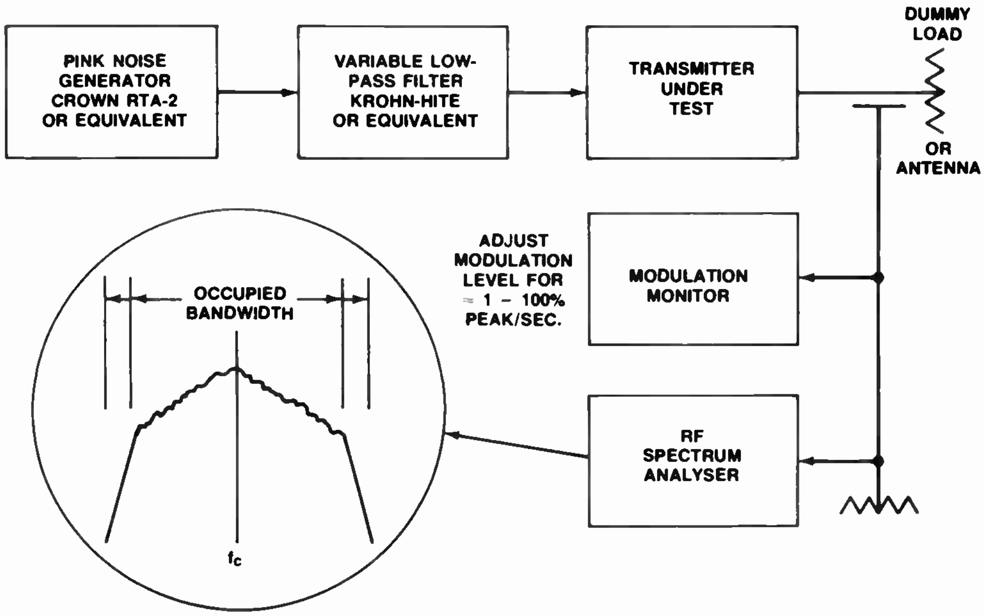


Figure 23. Block diagram test procedure for determining transmitter occupied bandwidth.

out to approximately the tenth RF harmonic. Then, with a calibrated measuring system, take measurements and compute the output power of each harmonic according to the formula:

$$P_n = V_n^2 \times R_{pn} \qquad [8]$$

where:

P_n = The power at the nth harmonic
V_n = The corrected measured voltage of the nth harmonic at the calibrated impedance point
R_{pn} = The parallel resistive component of the load impedance at the nth harmonic

2. Measure the actual power radiated from the antenna system at each harmonic frequency or suspected spurious output frequency using standard field intensity measurement techniques. This is the most meaningful of the two techniques because the measurement is made under actual conditions of operation with all systems interconnected. It is the transmitter manufacturer's and user's joint responsibility to correct any actual interference problems to other broadcast or nonbroadcast communication services, even when the interfering signal meets the standard FCC requirements (FCC Rules Section 73.44.c).

Carrier Output Power

The most accurate method of measuring the RF output power of a transmitter is by the calorimetric method, which uses the very accurately known and measurable physical and thermal characteristics of water or other liquids. This measurement is usually done only in a transmitter manufacturer's factory because the capital investment required to purchase and maintain calibration of this kind of infrequently-used equipment is usually not justified for AM broadcasting operations.

Water is known to have a thermal capacity very close to 4.186 Joules per degree C per gram weight at a mean temperature of 60°C. A Joule is equal to one watt-second. Therefore, the capacity of water to absorb power is 69.8 watts per degree C per liter of water flow per minute, or

Power (kW) = Flow (lpm) × DT°C × 0.0698 for water flow measured in liters per minute (lpm)

OR

Power (kW) = Flow (gpm) × DT°C × 0.2641 for water flow measure in U.S. gallons per minute (gpm)

The flow of water can be measured with an accuracy of approximately ±1%, by even the most common methods. Differential temperature measurement accuracy of approximately ±0.1°C is commonly practical, which, for temperature differentials of 20°C, is equivalent to ±0.5% accuracy. Using calorimetric measurement techniques, the output power of AM transmitters can be measured with total accuracies better than ±2%. The RF output amplifier efficiency factor F (referred to in the FCC Rules Section 73.51(e)(1) can then be determined for future operating and proof-of-performance reference. However, even with this method of determining the factor F, an accuracy of less than ±2% cannot be maintained over the life of the equipment. FCC-required transmitter voltage and current meters have a ±2% accuracy, in addition to the multiplied power accuracy of approximately ±4%. When accuracy levels are combined, this yields a total uncertainty of ±6% for the factor F. Still, this is considerably better than the accuracies obtainable by the direct measurement technique.

Operating Efficiency and Input Power

Measurement of transmitter input power should be done under actual or simulated operating conditions. This requires program or simulated program modulation during the power input measurement. The preferred method of measuring ac input power uses a standard rotating-disk watt-hour meter. The accuracy of these familiar meters is typically better than 0.5%, better than four times the accuracy of any other conventional direct ac power measurement technique, and they can be obtained with accuracies better than 0.1%. The watt-hour metering system should be connected in the main power feed line to the transmitter.

Sinusoidal test signals, useful for other tests of transmitter performance, are not recommended for power consumption tests because of the distinctly opposite statistical characteristics of sinusoidal signals and typical voice and music program material and the effect this difference has on actual operating power consumption measurements. This difference between periodic sinusoidal signals and mathematically random types of signals such as human voice and/or music has been known for decades. The effect this difference produces in AM broadcast transmitter power consumption and operating efficiency measurements, however, was first documented in 1980, by investigators in Europe[15] and was later verified and further explained by investigators in the U.S.[16,17] and other countries.

The critical difference between sinusoidal signals and typical program types of signals is explained in Fig. 24, which shows the amplitude-density characteristics of a sinusoidal waveform (U-shaped curve) and an amplitude density characteristic of typical broadcast program modulation (pyramid curve). The totally different shapes of the two curves in Fig. 24 cause the measured transmitter efficiency at identical RMS modulation levels to be quite different. The pyramid shaped curve in Fig. 24 was generated in 1984 from measurements taken off the air of five differently-formatted FM radio stations located in Dallas, Texas. Data was taken continuously for 24 hours for each station and averaged for presentation as shown in Fig. 24. FM stations were used because of their consistent day/night signal levels, symmetrical modulation characteristics (for easier comparison to the sine wave), and omni-directional emission pattern. Some transmitters,

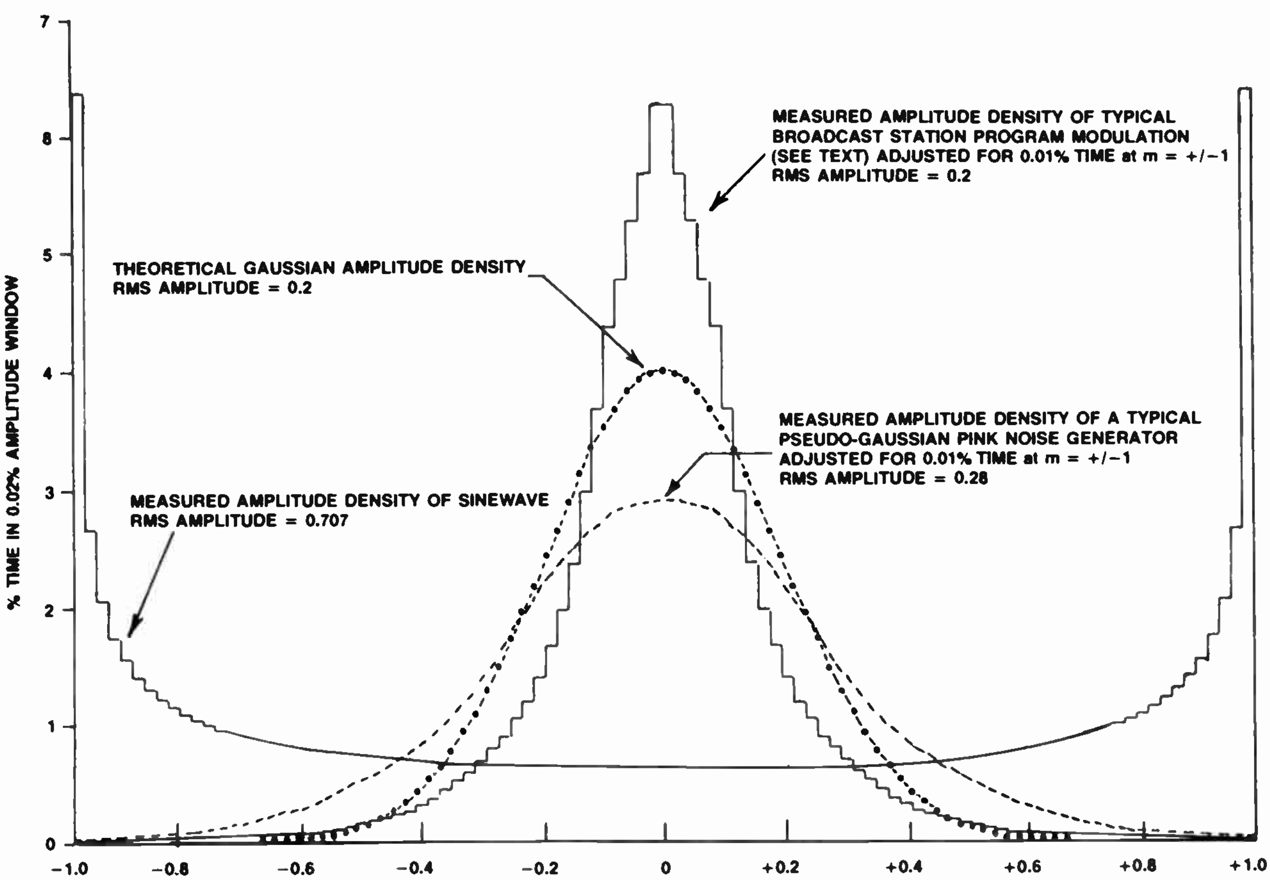


Figure 24. Amplitude density of sinusoidal test signals and typical program or noise test signals. Program or noise test signals recommended for transmitter efficiency measurements (see text).

or, more correctly, modulation systems, are more efficient with program modulation than with equivalent RMS sinusoidal test signals while others have poorer efficiency with program modulation.

Modulation systems or techniques that yield higher carrier efficiencies also yield higher program efficiencies (lower program power consumption than with an equivalent RMS sinusoidal modulation test signal). It is obvious that the program input power consumption test is the one most meaningful to broadcasters who broadcast programs, rather than sinusoidal tones. In this discussion, the terms "program efficiency" and "program input power consumption" have generally been given equal value. It is correct to equate these parameters for transmitter comparisons if it is assumed or defined that average transmitter output power is constant for a defined program input. Such an assumption is correct, except that transmitters with excessive transient overshoot will have correspondingly less average modulated RF power output for given peak levels of processed program modulation.

Experimental methods have been devised to measure average long-term transmitter output power (RF kilowatt-hour output).[18] With such equipment it would be convenient to compute accurately actual program efficiency knowing both kilowatt-hours input and kilowatt-hours output. Until such equipment becomes commercially available, however, only the transmitter-energy input can be measured accurately with single-phase and multi-phase kilowatt-hour instruments.

Fig. 24 also shows a third curve (dashed curve) which is the amplitude density function of a popular pink noise generator. It is recommended that either a pink noise signal with amplitude density characteristics similar to the curve in Fig. 24 or a recording of actual program modulation be used as a program source for these tests. Program processing equipment, if used, and peak modulation levels should be adjusted to normal station operating procedures. Transmitter input power measurements are then taken during a 30 minute (minimum) segment of the program material. Refer to FCC Rules Section 73.1570(a) and (b) for modulation setup procedures.

The average input power determined by this method will be very close to the transmitter power consumption of the transmitter during its operating life and therefore can be used to predict accurately actual operating energy costs. This is the only method that will provide accurate energy consumption forecasts.

IMPORTANT AM-STEREO TRANSMITTER CHARACTERISTICS

Incidental Phase Modulation

Some AM-stereo systems use a form of phase modulation for encoding the stereo signal on the AM carrier. For this reason, the most important transmitter characteristic affecting AM-stereo operation is incidental phase modulation (IPM), which has been discussed previously. Excessive IPM can affect stereo separation, monophonic distortion, and the occupied bandwidth of a stereo transmission; with the most significant of the three being monophonic distortion.

There are many potential sources, but the most common is incorrect neutralization of either a final modulated RF amplifier or a lower-power driver stage. The solution is, of course, better neutralization of the offending amplifier. This is easier said than done in most cases. Since the inception of AM-stereo broadcasting, manufacturers of transmitters have paid more attention to IPM and in most cases have reduced it to acceptable levels for AM-stereo operation.

For especially difficult problems, field engineers from the factory must make special on-site adjustment of transmitters which are particularly susceptible to excessive IPM. Several engineering consulting firms have developed special knowledge in the AM-stereo field and have collected data on many transmitters in current use (some of which have been out of production for several years). Such a firm should be contacted to solve particularly difficult neutralization problems.

Phase Noise

Residual phase-modulated noise cannot normally be detected by a standard AM broadcast receiver employing envelope detection, but stereophonic AM receivers are sensitive to PM noise. PM to AM conversion can occur on the medium wave band over multi-ionospheric "hops" (nighttime skywave propagation) and can then be detected by receivers employing standard envelope detection.

Phase-modulated noise in modern transmitters is virtually nonexistent due to the use of high-quality quartz crystal oscillator and synthesizer circuitry. Phase noise modulation of 0.6° (0.01 radians) average is fully acceptable for monophonic AM broadcasting. Phase noise modulation of approximately 0.2° (0.0032 radians) average is usually considered acceptable for AM stereophonic broadcasting. As with measurements of IPM, an HP-8901, (or equivalent), *modulation analyzer* is recommended for phase noise modulation measurements.

Stereophonic Phase/Gain Equalization

Standard production exciters for AM stereo systems may incorporate circuitry designed to match approximately the phase and gain characteristics of the normal monophonic transmitter transmission path to the transmission path for the encoded stereo signal. Transmitters which have excessive in-band nonlinear phase characteristics may require special "out-boarded" phase/gain equalization networks to achieve optimum stereo performance.

INTERNATIONAL SHORTWAVE BROADCAST TRANSMITTERS

Shortwave Transmitters

Shortwave broadcast transmitters are similar in many respects to medium wave transmitters and very different in others. The similarities are in methods of modulation, control, and monitoring. The differences are, generally, that shortwave transmitters are higher in power, more complex in tuning, and are more difficult to operate and maintain than medium wave "standard" broadcast transmitters. Although there are numerous exceptions, the general rule is that the minimum usable carrier power level on shortwave is 50 kW, and the maximum economical carrier power level from a single transmitter is 250 kW to 500 kW.

It is not unusual for a shortwave transmitter to operate on five to ten separate frequencies in one broadcast day. Schedules of the prestigious broadcasting organizations are very tight, necessitating built-in automatic frequency-changing circuitry which allows the transmitter to tune to several programmed frequencies in 10 to 30 seconds, with minimum operator intervention. The trend in shortwave transmitter operation is toward unattended or minimum attended sites, with program and frequency changes done either by remote or computer control or by both.

Single-Sideband Operation on the International Shortwave Bands

The International Telecommunications Union (ITU) and its radio broadcasting special committee, the International Radio Consultative Committee (CCIR) encourage the adoption of a form of single-sideband (SSB) as the standard modulation system for international shortwave broadcasting. This is due to the ever-increasing congestion in the shortwave broadcast bands, the competitive increases in transmitter power levels (which contributes to congestion), and appreciation of the increased efficiency of SSB transmission. A World Administrative Radio Conference Committee, meeting in Geneva in January and February 1984, adopted a detailed 20-year plan for conversion to SSB on the international shortwave broadcast bands. The committee report addressed necessary changes in both transmitter and receiver technologies. These changes can double the available channel space and improve reception quality in the current international broadcast bands.

AUDIO PREEMPHASIS

In the 1980s, transmitter and audio processor equipment manufacturers, broadcasters, and receiver manufacturers came together in a forum called the National Radio Systems Committee (NRSC), jointly sponsored by the National Association of Broadcasters (NAB) and the Electronic Industries Association (EIA). Out of this forum came definition and formal recognition of certain system inconsistencies that were impeding the growth of AM broadcasting.

The major such inconsistency in AM broadcasting had been known for some time. The offending parameter could be called transmitter/receiver interface. Because of the nature of nighttime skywave propagation on the medium-wave AM band, and the adjacent (and sometimes next-adjacent and even next-next-adjacent) channel interference created by this propagation, manufacturers had severely restricted the intermediate frequency (IF) and audio bandwidth of receivers produced for the AM band. The bandwidth restriction was often severe; −20 dB to −30 dB frequency response at 5 kHz was not unusual. Yet, stations continued to transmit wideband audio (out to at least 10 kHz and often 15 kHz) in an effort to provide the highest quality signal possible.

The committee addressed this transmitter/receiver interface mismatch and produced a recommended standard which restricted transmitted bandwidth to 10 kHz hoping this would encourage receiver manufacturers to widen the IF and audio-frequency response of the receivers correspondingly.

The solution was the only one that could be made in light of the present congestion in North America on the AM band. The standard that the committee proposed was based on the compliance of a majority of North American broadcasters and manufacturers of AM receivers for the North American market. More than 1000 private AM broadcasters in the U.S. and Canada participated, prompting the FCC to make the NRSC standard a requirement by all U.S. AM broadcast stations. The FCC requirement took effect on June 30, 1990. The standard recommends an AM preemphasis curve is to be used in audio processing equipment or in the transmitter itself. The curve is called a "Modified 75 micro-second AM Preemphasis Standard Curve." It exhibits approximately 1 dB boost at 1 kHz and approximately 10 dB boost at 10 kHz, followed or accompanied by a single pole low-pass filter with a break frequency of 8.7 kHz to reduce the peak boost at high frequencies. The matching deemphasis curve in the receiver can be achieved in the IF and audio stages.

The recommendation by the NRSC and following supportive action by the FCC is at least an important

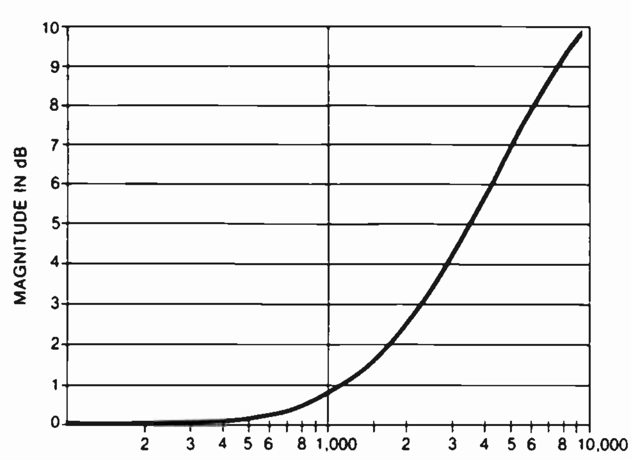


Figure 25. Modified 75 micro-second AM standard preemphasis curve.

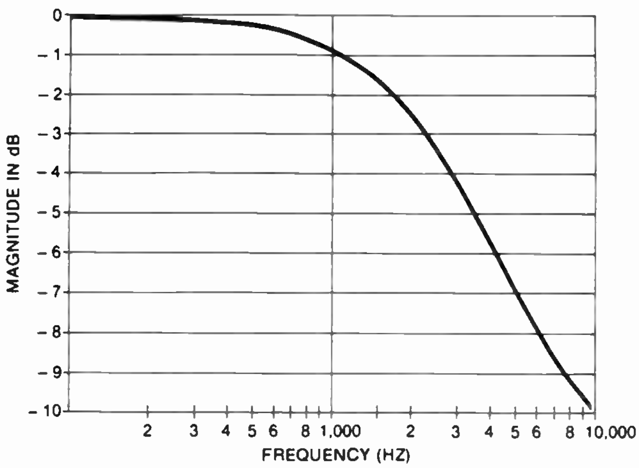


Figure 26. Modified 75 micro-second AM standard deemphasis curve.

first step in achieving improved bandwidth on the North American AM broadcast band. The issue of high audio-frequency preemphasis is not without some controversy. Some broadcast engineers believe the standard is good in concept but does not go far enough, that high-frequency boost should be limited to somewhat lower than 10 kHz to further reduce or possibly even eliminate adjacent channel interference. The European Broadcasting Union (EBU), in conjunction with the CCIR, recommends moderate high-frequency preemphasis to 4.5 kHz (European AM channels are spaced at 9 kHz intervals), followed by a sharp-cutoff low-pass filter at frequencies above 4.5 kHz. An identical standard of 4.5 kHz maximum sideband width is also recommended by the CCIR for shortwave broadcasting where the channel spacing is 5 kHz.

The NRSC also generated a second standard pertaining to measurement of actual occupied bandwidth of a licensed AM broadcast station. The FCC has ordered that all U.S. AM broadcast stations comply with this second new standard, NRSC-2, by June 30, 1994.

The full text of NRSC standards 1 and 2 follow as Appendix A and B.

REFERENCES

1. "Notes on Modulation Theory," John R. Carson, Proceedings of the Institute of Radio Engineers, Volume 10, No. 1, June 1925, p. 57.
2. "Production of Single Side-Band for Trans-Atlantic Radio Telephone," R.A.Heising, Proceedings of the Institute of Radio Engineers, Vol. 13, No. 3, June 1925, p. 291.
3. Inaugural Address of Dr. Lee DeForest upon becoming President of the Institute of Radio Engineers., Proceedings of the Institute of Radio Engineers, No. 7, Jul. 1930, p. 1121.
4. "The Equivalent Circuit of the Vacuum Tube Modulator," John R. Carson, Proceedings of the Institute of Radio Engineers, Vol. 9, No. 4, Aug. 1921, p. 243.
5. "Modulation in Radio Telephony," R.A. Heising, Proceedings of the Institute of Radio Engineers, Vol. 9, No. 3, June 1921, p. 305.
6. "High Power Outphasing Modulation," Henry Chireix, Proceedings of the Institute of Radio Engineers, Vol. 23, No. 11, Nov. 1935, p. 1370.
7. "A New High Efficiency Power Amplifier for Modulated Waves," W.H.Doherty, Proceedings of the Institute of Radio Engineers, Vol. 24, No. 9, Sept. 1936, p. 1163.
8. "Amplifiers," J.O. Weldon, U.S. Patent No. 2,836,665, May 1958.
9. "Transmission System," R.A. Heising, U.S. Patent No. 1,655,543, January 1928.
10. "Distortion in Pulse Duration Modulation," Ernest R. Kretzmer, Proceedings of the Institute of Radio Engineers, Vol. 35, No. 11, November 1947, p. 1230.
11. "F.E. Terman and John R. Woodyard: A High Efficiency Grid-Modulated Amplifier," Proceedings of the Institute of Radio Engineers, Vol. 26, No. 8, Aug. 1938, p. 929.
12. "High Efficiency Amplifier and Push-pull Modulator," J.B. Sainton, U.S. Patent No. 3,314,024, April 1967.
13. "Panter, Philip F: *Modulation, Noise, and Spectral Analysis*, McGraw-Hill, 1965, Chapter 20, pp. 640–641."
14. "Operation of AM Broadcast Transmitters into Sharply Tuned Antenna Systems," W.H. Doherty, Proceedings of the Institute of Radio Engineers, Vol. 37, No. 7, July 1949, p. 729.
15. M. Sempert and W. Tschol: "Efficiency of High Power Broadcasting Transmitters in Regular Programme Service," BBC Review, July 1980.
16. G.W. Woodard: "Efficiency Comparison of AM Broadcast Transmitters," IREE Journal (Australia), Vol. 2, No. 2, June 1982.
17. G.W. Woodard: "Simulating Typical Program Modulation Measurements of Operating Efficiency and Modulation Capability of AM Broadcast Transmitters," Radio and Electronic Engineer (UK), Vol. 53, No. 9, p. 325.
18. "Measurement of the efficiency of high-power AM transmitters with regular broadcast programmes," A.W. Jaussi and W. Tschol, EBU Review, No.194, August 1982, p. 169.

NATIONAL RADIO SYSTEMS COMMITTEE

Electronic
Industries
Association

National
Association
of Broadcasters

INTERIM
VOLUNTARY NATIONAL STANDARD

1. 75 µS AM Broadcast Transmission Preemphasis
2. Complimentary 75 µS AM Receiver Deemphasis
3. 10 kHz AM Transmission Bandwidth
4. Five-year Review Provision

January 10, 1987

Sponsored by the Electronic Industries Association and the National Association of Broadcasters

NRSC Co-convenors:

Charles T. Morgan
Susquehanna Radio Corp.
York, Pennsylvania

Bart Locanthi
Pioneer North America, Inc.
Duarte, California

Alan Boyer
Sony Corp. of America
Kansas City, Missouri

NRSC Subgroup Co-convenors:

John Marino
NewCity Communications
Bridgeport, Connecticut

William F. Gilbert
Delco Electronics Corp.
Kokomo, Indiana

TABLE OF CONTENTS

§ 1. Scope 1

§ 2. Introduction 1

§ 3. Basic Definitions 1

§ 4. AM Transmission Preemphasis 3

§ 4.1. In General 3

§ 4.2. Description of the Standard Preemphasis Curve 3

§ 4.3. Methods for Determining Performance 3

§ 4.3.1. Use of Audio Tones 3

§ 4.3.2. Location of Measurement 3

§ 5. AM Receiver Deemphasis 3

§ 5.1. In General 3

§ 5.2. Description of the Standard Deemphasis Curve 4

§ 5.3. Methods for Determining Performance 4

§ 5.4. Notch filters 5

§ 6. 10 kHz Bandwidth for AM Transmission 5

§ 6.1. In General 5

§ 6.2. Description of the Standard 5

§ 6.3. Method for Determining Performance 6

§ 6.3.1. Location of Measurement 6

§ 6.3.2. Use of Standard Test Signal 6

§ 6.3.3. Use of Standard Measurement Devices 6

§ 7. Five Year Review Provision 7

§ 8. Effective Date 7

Figure 1: AM Preemphasis Curve
Figure 2: AM Deemphasis Curve
Figure 3: 10 kHz Bandwidth Specification
Figure 4: Spectra of USASI Noise
Figure 5: Pulsed USASI Noise Generator
Figure 6: USASI Noise Weighting/Pulser Circuit

§ 1. SCOPE

The National Radio Systems Committee (NRSC) is a joint Committee composed of all interested parties including representatives of AM broadcast stations, AM receiver manufacturers, and broadcast equipment manufacturers. This document describes an interim[1] voluntary national standard that specifies the preemphasis of AM broadcasts, the deemphasis of AM receivers, and the audio bandwidth of AM stations prior to modulation. The standard applies to AM monophonic and AM stereo L+R transmissions, and to dual bandwidth and single bandwidth AM receivers. Compliance with the standard is strictly voluntary. To the NRSC's knowledge, no industry group or entity is or will be adversely affected by issuance of this document. Every effort has been made to inform and accommodate any and all interested parties. The NRSC believes that implementation of the standard will reduce AM interference, increase useful AM service areas, and encourage the production of higher fidelity AM receivers.

A five year review provision is established.

§ 2. INTRODUCTION

On September 5, 1985, the NRSC adopted a resolution to study proposals to standardize AM transmission preemphasis and AM receiver deemphasis with the objective of establishing an industry-wide AM preemphasis/deemphasis voluntary standard. After twelve months of study, on September 10, 1986 the NRSC released a draft voluntary standard that proposed a specific AM preemphasis/deemphasis curve as well as a 10 kHz standard AM bandwidth. The bandwidth specification evolved from NRSC deliberation on the causes and cures of AM interference, and ways to technically encourage the production of higher fidelity AM receivers. After a three month public comment period, the NRSC, on January 10, 1987, formally adopted this standard and authorized its publication by the National Association of Broadcasters and the Electronics Industries Association.

The purpose of the NRSC voluntary standard is to create a transmission/reception system where (1) AM broadcast stations will know, with certainty, the likely audio response characteristics of AM receivers, and (2) AM receiver manufacturers will know, with certainty, the likely audio response characteristics of AM broadcasts. A "matching" of preemphasis and deemphasis is expected to improve the consumer's overall satisfaction with the technical quality of listening to AM radio. The NRSC believes that the public interest is served by establishing a compatible transmission/reception system and the accompanying improvement of AM broadcasts and AM receivers.

This document also describes a specification for the maximum audio bandwidth transmitted by AM broadcast stations. Implementation of a bandwidth specification will reduce second-adjacent channel interference and thereby lead to (1) a significant reduction of second-adjacent channel interference as perceived on "wideband" AM receivers; (2) a corresponding increase in the interference-free service areas of AM stations; and (3) an incentive for the further building of dual bandwidth AM "wideband" receivers.[2] Analysis by a subgroup of the NRSC has shown that there would be little if any detrimental effect on today's "narrowband" AM receivers upon the implementation of this voluntary standard.

§ 3. BASIC DEFINITIONS

A. Preemphasis. The boosting of high audio frequencies prior to modulation and transmission.

B. Deemphasis. The attenuation of high audio frequencies during the process of reception and demodulation.

C. "Narrow" receivers. A subjective term to describe receivers with typical combined RF, IF and AF response characteristics of -10 dB at 4.2 kHz, -20 dB at 6.0 kHz. Response characteristics of narrow AM receivers are known to vary widely.

D. "Wideband" receivers. A subjective term to describe receivers with typical combined RF, IF, and AF response of -6 dB

1.The standard is described as "interim" until it is submitted and approved by the American National Standards Institute.

2.First Adjacent channel interference considerations may continue to discourage the building of single bandwidth "wideband" receivers; however, the extent and nature of this form of interference has not been fully studied by the NRSC.

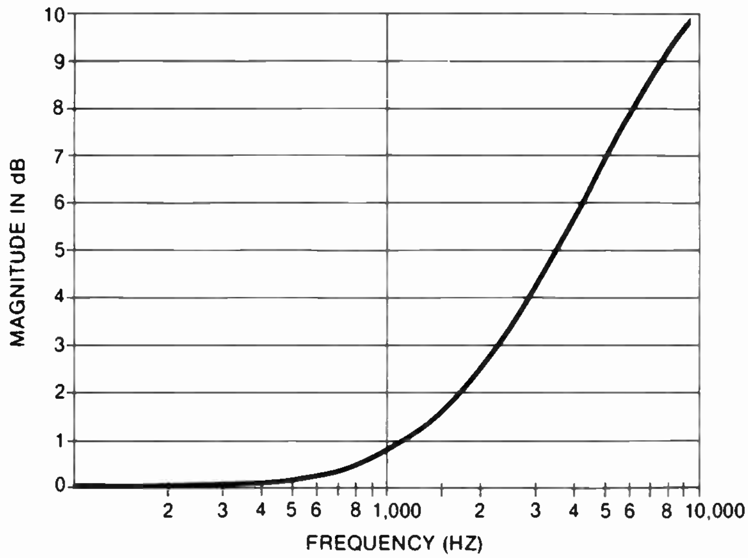


Figure 1. Modified 75μs AM Standard Preemphasis Curve

Technical Information

Frequency	Magnitude (dB)	Phase (deg)	Group Delay (sec)	Frequency	Magnitude (dB)	Phase (deg	Group Delay (sec)
50	0.00	1.0	-5.6669E-005	5000	6.92	37.1	2.3048E-006
100	0.01	2.0	-5.6547E-005	5500	7.41	36.6	3.3525E-006
400	0.14	8.0	-5.4175E-005	6000	7.85	35.9	4.0592E-006
700	0.42	13.7	-4.9467E-005	6500	8.24	35.2	4.5169E-006
1000	0.81	18.7	-4.3318E-005	7000	8.58	34.3	4.7926E-006
1500	1.63	25.5	-3.2247E-005	7500	8.89	33.4	4.9357E-006
2000	2.54	30.4	-2.2343E-005	8000	9.16	32.5	4.9823E-006
2500	3.44	33.6	-1.4509E-005	8500	9.41	31.6	4.9595E-006
3000	4.28	35.7	-8.6612E-006	9000	9.62	30.8	4.8871E-006
3500	5.05	36.9	-4.4133E-006	9500	9.82	29.9	4.7801E-006
4000	5.75	37.4	-1.3702E-006	10000	10.00	29.0	4.6495E-006
4500	6.37	37.4	7.8900E-007				

at 6 kHz, -10 dB at 8 kHz. Response characteristics of wide AM receivers are known to vary widely.

E. "Excessive" Preemphasis. Preemphasis that produces no discernable benefit when received by a "narrow" receiver but increases interference to adjacent channel AM stations.

§ 4. AM TRANSMISSION PREEMPHASIS

§ 4.1. In General. AM preemphasis is the boosting of high audio frequencies prior to modulation and transmission. Today, most AM stations use preemphasis to varying extents. This preemphasis is employed in an attempt to compensate for the "narrow" response of most AM receivers. If AM preemphasis is not controlled, one station may interfere with AM receivers listening to neighboring stations located on adjacent AM channels. Whether such interference is objectionable will depend on (1) the response characteristics of the AM receiver, (2) the amount and nature of transmission preemphasis, (3) the extent to which the AM station is employing compression/ limiting techniques, and (4) whether the AM transmission system is bandlimited in the audio processor, transmitter or antenna.

Preemphasis is useful for improvement of the AM transmission-reception system audio response only to a limited extent for receivers using IF transformers. Many receivers using ceramic filters with narrow response characteristics can not be improved by use of excessive preemphasis. These receivers can not "hear" the transmission of preemphasized high audio frequencies. But excessive preemphasis will foster adjacent channel interference and cause wideband radios to sound shrill or strident.

§ 4.2. Description of the Modified 75 uS Preemphasis Curve. Each AM broadcast station shall broadcast with audio preemphasis as close as possible (within the capabilities of the station's transmission system) to the recommended standard, without exceeding it. The curve applies for audio frequencies up to 10 kHz.

The NRSC proposed standard AM transmission preemphasis curve is shown in Figure 1. The curve describes the recommended net transmission system static audio response of an AM station.

The recommended preemphasis curve is a single zero curve with a break frequency at 2122 Hz. It is similar to the 75 uS curve used for FM broadcasting. To reduce the peak boost at high frequencies, a single pole with a break frequency of 8700 Hz is employed. NRSC analysis has shown that the proposed curve is compatible with most existing AM receivers.

§ 4.3. Methods for Determining Performance. The NRSC AM preemphasis curve is a static curve, and cannot be measured dynamically. NRSC studies have shown that the dynamic and non-linear functions performed by most AM station audio processors will modify any given preemphasis curve. In addition, it is the audio response of the entire AM transmission system that indicates performance in accordance with the standard. For these reasons, measuring a station's preemphasis curve for the purpose of determining compliance with the NRSC standard shall be performed in accord with the following specifications:

§ 4.3.1. Use of Audio Tones. Compliance with the curve shall be measured by sweeping the station's transmission system with audio tones. The dynamic functions of the AM station's processor, but not the frequency shaping circuits, must be disabled (i.e., in "proof" mode).

§ 4.3.2. Location of Measurement. The net transmission system audio response is best measured by detecting the over-the-air signal. This will ensure that the AM transmitter and antenna combination is faithfully reproducing the preemphasized audio.[3] Alternatively, if the transmitter and antenna combination is reasonably broadband, performance can be determined by static measurement of the audio signal prior to modulation.

§ 5. AM RECEIVER DEEMPHASIS

§ 5.1. In General. Receiver deemphasis results from the selectivity

---

3.However, the deemphasis characteristics of the device used to demodulate the AM transmission must be accounted for. Additionally, some AM stations with transmitter or antenna problems may not be able to pass preemphasized audio without introducing "splatter" interference and/or overmodulation. If a particular AM station transmission system cannot "handle" the NRSC recommended curve, it is suggested that a lower amount of preemphasis be used until the system problems are corrected to allow the NRSC curve to be faithfully implemented.

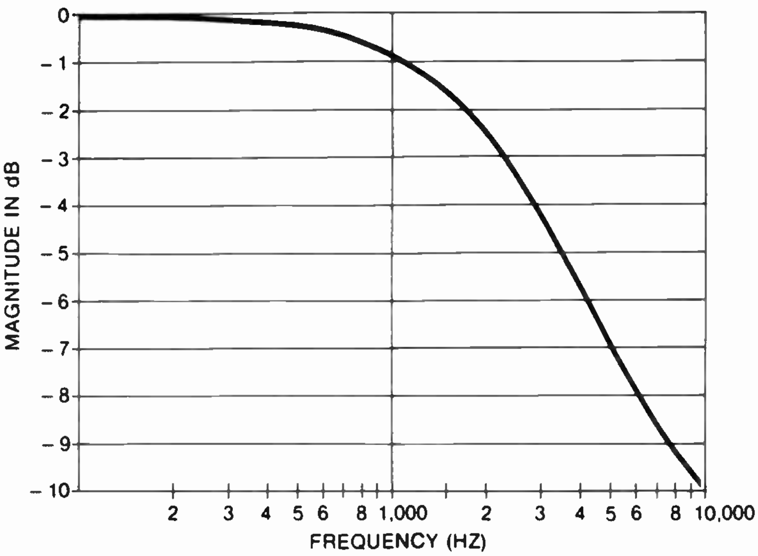


Figure 2. Modified 75μs AM Standard Deemphasis Curve

characteristics of a receiver's RF and IF stages and the response characteristics of the receiver AF section. A standard deemphasis curve permits AM stations to know, with certainty, the likely overall response characteristics of AM receivers.

§ 5.2. Description of the Standard Deemphasis Curve. AM receivers shall complement the recommended transmission preemphasis characteristic described in § 4 by incorporating a net receiver system audio response described in Figure 2. (The net system audio response of an AM receiver is the combined RF, IF, and AF audio response.) The NRSC deemphasis curve is characterized by a single pole at 2122 Hz and a single zero at 8700 Hz. It is the precise complement of the preemphasis standard described in § 4. The preemphasis/deemphasis voluntary standards apply only for audio frequencies below 10 kHz; the implementation of preemphasis/deemphasis standards produces a transmission/reception system that is essentially flat to nearly 10 kHz and limited only by the AM receiver's choice of bandwidth.

§ 5.3. Methods for Determining Performance. The deemphasis characteristic shall be determined by measuring the overall frequency response in accordance with International Electrotechnical Commission ("IEC") Publication 315.3, Clause 11.2:

(1) The receiver is brought under standard measuring conditions and the reference audio-frequency output voltage is noted. The modulation frequency is then varied and the output voltage at each frequency is noted and expressed in decibels relative to the reference voltage.

The modulation depth is adjusted at each frequency in accordance with the preemphasis characteristic of AM broadcast transmission. To avoid overmodulation at some frequencies it may be necessary to use a modulation factor of less than 30% at some frequencies.

(2) If overloading of the AF section of the receiver occurs, either the volume control attenuation should be increased or the modulation factor reduced, and a corresponding factor applied to the results.

(3) The measurements may be repeated with other values of RF input signal level and frequency.

The frequency response shall be measured for both monophonic and stereophonic reception, in accordance with the definition of the particular AM stereo system. For dual bandwidth receivers, the frequency response shall be measured in both bandwidth positions.

Results may be presented graphically, with modulation frequency plotted logarithmically as abscissa and the output in decibels as ordinate.

The frequency response can be stated as follows:

Selectivity	Frequency Response
Narrowband	50 Hz to 5000 Hz +/- X dB
Wideband or Stereo	50 Hz to 10,000 Hz +/- X dB

(Where X is the maximum deviation from the recommended frequency response, and 5000 Hz and 10,000 Hz are example frequency specifications.) The deviation X shall be of as low a value as practical. If a notch filter is used while the AM receiver is under test, the stated frequency response should be modified accordingly. Suggested modifications include (1) adding an appropriate footnote to the frequency response specification; and/or (2) lowering the upper limit value to the above "wideband" audio response specification.

§ 5.4. Notch Filters. A notch filter is a very selective filter that attenuates the spectrally pure carriers of first adjacent channel AM stations. Although an optional enhancement to an AM receiver, using notch filters is recommended. If used, the notch filter should (1) have as high a "Q" as is practical, (2) adequately suppress the interfering carriers, and (3) not unduly degrade the desired bandwidth performance of the AM receiver.

§ 6. 10 KHZ BANDWIDTH FOR AM TRANSMISSION

§ 6.1 In General. Each AM broadcast station shall modulate its transmitter with an audio bandwidth described by the specification in Figure 3. Appropriate and carefully designed audio low-pass filters as the final filtering prior to modulation can be used to implement this specification. The purpose of the bandwidth specification is to remove interference by controlling the occupied RF bandwidth of AM stations.[4]

§ 6.2. Description of the Standard. The audio bandwidth transmission standard is specified in Figure 3. The audio

4.It should be noted that the operation of non-linear AM Stereo systems theoretically may produce phase modulation components outside the desired RF bandwidth. The NRSC will examine this phenomena with the goal of determining whether such components exist and, if they do exist, whether they are objectionable. For a discussion of the detrimental effect of high audio frequencies on occupied RF bandwidth, see Klein, Modulation, Overmodulation, and Occupied Bandwidth: Recommendations for the AM Broadcast Industry, Proceedings of the 1987 NAB Engineering Conference, Dallas, Texas (April, 1987).

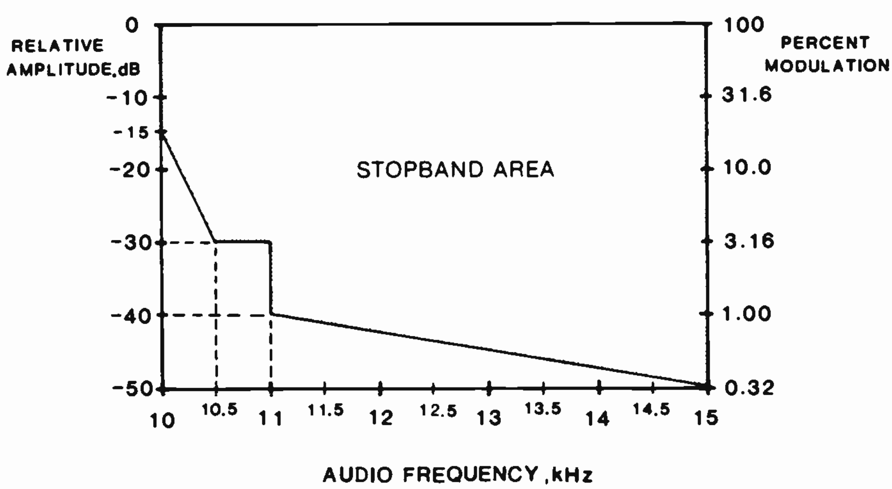


Figure 3. NRSC Stopband Specification
(Audio Envelope Input Spectrum To AM Transmitter)

envelope input spectrum to the AM transmitter shall be -15 dB at 10 kHz, smoothly decreasing to -30 dB at 10.5 kHz, then remaining at -30 dB from 10.5 kHz until 11.0 kHz. At 11.0 kHz, the audio bandwidth shall be -40 dB, smoothly decreasing to -50 dB at 15 kHz. Above 15 kHz, the audio bandwidth shall remain at least -50 dB. The reference level is 1 dB above a 200 Hz sine wave at 90% negative modulation. See Figure 3.

§ 6.3. Method for Determining Performance. An AM station is determined to be in compliance with the NRSC bandwidth characteristic by measurement of the station's audio bandwidth in accordance with the following parameters:

§ 6.3.1. Location of Measurement. Audio bandwidth measurements shall be obtained at the audio input terminals to the AM transmitter. For AM Stereo stations, audio bandwidth shall be measured at the L+R audio input terminals to the RF modulator. Note that the NRSC bandwidth standard characterizes an audio bandwidth that represents station program material that has been modified by possibly non-linear circuits in the station's audio processor. For this reason, the NRSC recommends use of a test signal that adequately characterizes typical audio program material, rather than relying on static audio test tones. However, it may still be useful to measure bandwidth statically at the time that AM preemphasis is measured.

§ 6.3.2. Use of Standard Test Signal. Audio bandwidth shall be measured using a test signal consisting of USASI (United States of America Standards Institute) noise that is pulsed by a frequency of 2.5 Hz at a duty cycle of 12.5%. See Figure 4. USASI noise is intended to simulate the long-term average spectra of typical audio program material. Pulsing of the noise is intended to simulate audio transients found in audio program material. USASI noise is a white noise source[5] (i.e. noise with equal energy at all frequencies) that is filtered by (1) a 100 Hz, 6 dB per octave high-pass network and (2) a 320 Hz, 6 dB per octave low-pass network. See Figure 4. A pulsed USASI noise generator is shown in Figures 5 and 6. Using the attenuator pad, the ratio of peak-to-average amplitude shall be 20 dB at the audio output of the pulser. The station's audio processor must be in normal operating mode.

§ 6.3.3. Use of Standard Measurement Devices. A suitable swept-frequency or FFT (Fast Fourier Transform) spectrum analyzer shall be used to measure compliance with the NRSC bandwidth specification.

(a) Spectrum Analyzer Setup. When a swept-frequency audio spectrum analyzer is used to measure compliance with the NRSC bandwidth specification, the analyzer's setup shall consist of:

a. 300 Hz resolution bandwidth.[6]
b. 2 kHz/horizontal division.
c. 10 dB/vertical division.
d. Reference: 1 dB above 200 Hz (sine wave) 90% negative modulation.
e. Display: maximum peak hold (or equivalent function).

The analyzer's operating span and sensitivity are adjusted as necessary to determine compliance.

(b) Fast Fourier Transform Analyzer. When a FFT analyzer is used to measure compliance with the NRSC bandwidth specification, the analyzer's setup shall consist of:

a. Reference: 1 dB above 200 Hz (sine wave) 90% negative modulation.
b. Window: Hanning.
c. Horiz. span: 20 kHz.
d. Dynamic range: 80 dB or available range.
e. Display: Maximum peak hold (or equivalent function).

5.Acceptable white noise sources include GenRad Models 1382 and 1390B; Bruel & Kjaer Model 1405; and National Semiconductor IC No. MM 5837N.

6.Note: if the audio bandwidth under test fails to meet the NRSC specification when a 300 Hz resolution bandwidth is employed, a narrower resolution bandwidth, such as 100 Hz, may be used to determine compliance; however, the sweep rate and the video bandwidth of the analyzer must be adjusted according to the manufacturer's instructions in order to assure accurate representation of the resolution bandwidth employed. If in doubt, check with the analyzer operating manual or the manufacturer. Further Note: the NRSC may suggest a different or more precise measurement standard as the industry gains experience. Spectrum analyzers that are capable of 300 Hz resolution bandwidths at audio frequencies include, but are not limited to, Tektronix Models 5L4N and 7L5; Hewlett-Packard Models 3580A, 3582A, 3585A, 8553A or B, 8566A or B, 8568A or B, 71100A; Marconi Models 2370, 2382; Rhode/Schwarz/Polarad Model CSA-240M; and a Techron Model TEF System 12.

§ 7. FIVE YEAR REVIEW PROVISION

It is the goal of the NRSC to increase the fidelity of the AM transmission and reception system from its present state to a quality level that approaches the quality available via FM broadcasting. Towards this end, the voluntary standards described in this document shall be in effect for five (5) years from the effective date of this standard. During the interim five year period, this voluntary standard will be reviewed at least once a year to determine whether the fidelity goals of the NRSC are being realized.

§ 8. EFFECTIVE DATES

These dates serve only as objectives. AM Broadcast stations and AM receiver manufacturers are expected to make a good faith effort to implement the NRSC standard.

A. AM Broadcast Stations. The effective date of this standard is January 10, 1987.

B. AM Receiver Manufacturers. The effective date of this standard is January 10, 1988.

§§§§

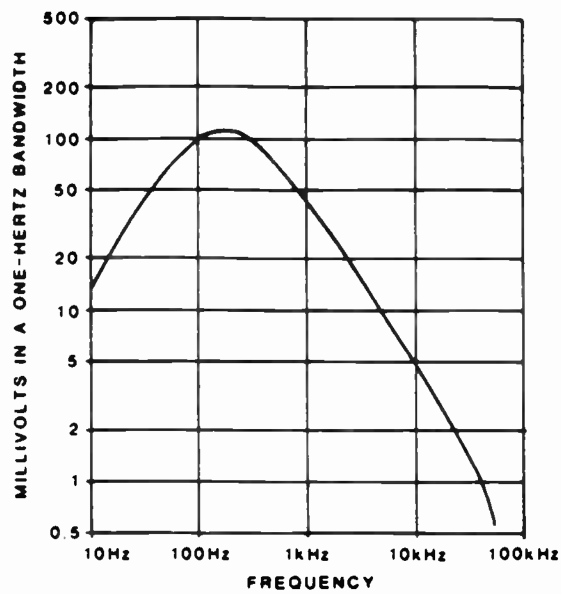


Figure 4. Spectra of USASI Noise

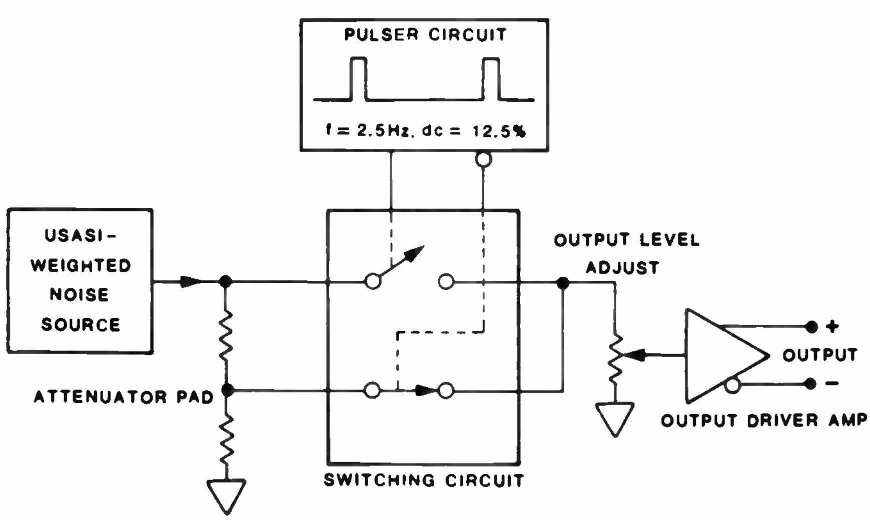


Figure 5. Pulsed-USASI Noise Generator

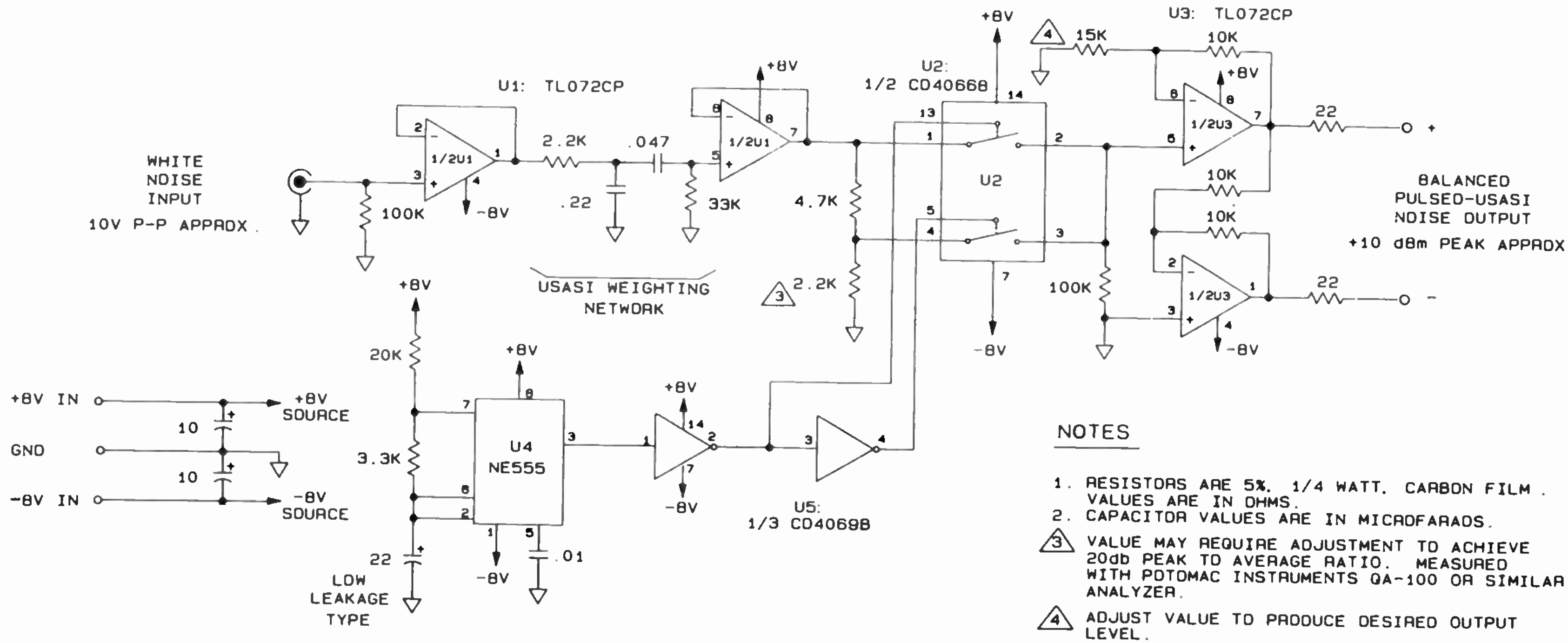


Figure 6. Application Circuit: USASI Noise Weighting/Pulser Circuit

For more information about this document, call:

NAB Department of Science and Technology
(202) 429-5346 or

EIA Engineering Department
(202) 457-4975

NATIONAL RADIO SYSTEMS COMMITTEE

2001 Eye Street, N.W.
Washington, D.C. 20006
(202) 457-4975

National Association of Broadcasters
1771 N Street, N.W.
Washington, D.C. 20036
(202) 429-5346

June 1, 1988

INTERIM
VOLUNTARY NATIONAL STANDARD
EMISSION LIMITATION FOR AM BROADCAST TRANSMISSION

NRSC Chairman:

Charles T. Morgan
Susquehanna Radio Corporation
York, Pennsylvania

NRSC AM Subgroup:

John Marino
NewCity Communications
Bridgeport, Connecticut

William F. Gilbert
Delco Electronics Corporation
Kokomo, Indiana

Voluntary Standard No. NRSC-2

Sponsored by the Electronic Industries Association and the National Association of Broadcasters

TABLE OF CONTENTS

§ 1. Scope 1

§ 2. Introduction 1

§ 3. RF Maximum Occupied Bandwidth Specification 2

§ 3.1. Purpose 2

§ 3.2. Maximum Occupied Bandwidth 2

§ 3.3. Measurement Procedure 3

§ 3.3.1. Use of Ordinary Program Material 3

§ 3.3.1.1. Use of Audio Tones 3

§ 3.3.2. Use of Spectrum Analyzer 4

§ 4. RF Occupied Bandwidth Testing and Control Standard 4

§ 4.1. Purpose 4

§ 4.2. Maximum Occupied Bandwidth 4

§ 4.3. Measurement Procedure 5

§ 4.3.1. Standard Noise Test 5

§ 4.3.1.1. Monophonic Conditions 5

§ 4.3.1.2. Stereophonic Conditions 5

§ 4.3.2. Use of Spectrum Analyzer 6

§ 5. Splatter Monitor 6

§ 6. Effective Date 6

Figure 1 : AM Broadcast RF Emission Limits
Figure 1A: AM Broadcast RF Emission Limits (Expanded Scale)
Figure 2 : Noise Generator for a 3dB (L+R) to (L-R) Ratio

EMISSION LIMITATION FOR AM BROADCAST TRANSMISSION (NRSC Voluntary Standard No. NRSC-2)

§ 1. Scope.

The National Radio Systems Committee (NRSC) is a joint committee composed of all interested parties including representatives of AM broadcast stations, AM receiver manufacturers, and broadcast equipment suppliers. This document describes an interim[1] voluntary national standard that specifies radio-frequency spectrum occupancy for AM broadcast stations. The standard applies to both AM monophonic and AM stereo transmissions. Compliance with the standard is strictly voluntary. To the NRSC's knowledge, no industry group or entity is or will be adversely affected by issuance of this document. Every effort has been made to inform and accommodate any and all interested parties. The NRSC believes that implementation of this standard will lead to reduced AM interference, thus providing increased service for all AM stations and an increase in quality of service to present and future AM listeners.

§ 2. Introduction.

On January 10, 1987, the NRSC authorized the National Association of Broadcasters and the Electronic Industries Association to publish an interim voluntary national standard specifying AM preemphasis, AM deemphasis and a 10 kHz AM audio bandwidth (Standard No. NRSC-1). The NRSC-1 audio standard applies to the audio signals that are intended to modulate the AM transmitter. Its purpose is to reduce second-adjacent channel interference by band limiting AM stations to a nominal

[1]The standard is described as "interim" until the test methods contained within the document can be fully verified through field tests. In addition, field test data of "splatter monitor" technology will be evaluated to determine the correlation between results obtained with such devices and the methods described within this document (see § 5.).

20 kHz occupied radio-frequency (RF) bandwidth (twice the 10 kHz audio bandwidth presented to the transmitter's modulation circuits). Implementation of the NRSC-1 audio standard alone largely achieves this purpose. However, there remain characteristics of the AM transmission process that may cause the RF occupied bandwidth to exceed a nominal 20 kHz. This document accommodates these transmission characteristics. It is in two Sections. Section 3 is a voluntary standard maximum RF occupied bandwidth of AM broadcast transmissions. Section 4 consists of a voluntary RF occupied bandwidth testing and control standard designed to insure repeatability and consistency of RF occupied bandwidth test measurements.

§ 3. RF Maximum Occupied Bandwidth Specification.

§ 3.1. Purpose. The purpose of an RF maximum occupied bandwidth specification is to control modulation products, desired or undesired, that fall outside the specified RF occupied bandwidth.

§ 3.2. Maximum Occupied Bandwidth. The maximum occupied RF bandwidth voluntary standard represents the maximum peak output of a swept-frequency spectrum analyzer IF over a minimum ten minute period.[2] The specification encompasses all spectral components caused by direct programming and all ancillary or data communications. AM broadcast stations shall occupy spectrum according to the following maximum specifications:

(see next page)

[2]It is recognized that the output of the spectrum analyzer depends on the shape of the analyzer's IF filters.

Table 1

Frequency Band Relative to Carrier (+/- kHz)	Attenuation Relative to Carrier (dB)
0 to 10	0
10 to 20	greater than -25[3]
20 to 30	greater than -35
30 to 60	greater than -(5 + 1 dB/kHz) from carrier[4]
60 to 75	-65[4]
above 75	-80[4]

(See Figures 1 and 1A (solid line) attached).

§ 3.3. Measurement Procedure.

§ 3.3.1. Use of Ordinary Program Material. Measurements of AM station spectrum occupancy shall be conducted using ordinary program material. All audio processing used in the AM station shall be in normal operating modes. The audio signal input to the AM transmitter shall conform to the NRSC audio standard adopted January 10, 1987.

§ 3.3.1.1. Use of Audio Tones. Sweeping a transmission system with audio tones is a widely accepted and respected method for gauging spectrum occupancy and for troubleshooting and adjusting AM transmission systems. The NRSC endorses audio tones for these purposes, but urges caution in the use and selection of audio tones particularly with AM stereo transmission.[5] It should be noted, however, that it is

---

[3]Note: the slope of occupied bandwidth in the transition region between 10 and 11 kHz is expected to parallel the NRSC-1 audio standard. Accordingly, attenuation levels in the region shall be 6 dB greater than described in the audio standard to adjust for carrier level reference.

[4]For carrier power levels between 50 and 5000 Watts, the maximum limit shall be $-(43 + 10 \log P_W)$ dB (where P_W is the carrier power in Watts) or as indicated in Table 1 and/or the attached Figure 1, whichever is lesser attenuation. For carrier power levels below 50 Watts, a -60 dBC maximum limit shall be used.

[5]The manufacturer of the particular AM stereo system employed should be consulted for the appropriate tone frequencies/modulation levels for "worst-case" condition testing.

difficult to infer the dynamic response of a transmission system while observing it in a steady-state condition.

§ 3.3.2. Use of Spectrum Analyzer. A suitable swept-frequency RF spectrum analyzer shall be used to measure compliance with the NRSC RF occupied bandwidth specification. The analyzer shall measure the over-the-air RF spectrum occupancy as perceived in the far field (i.e., at least 10 wavelengths from the antenna center). Some caution should be used in measuring spectrum occupancy with directional antennas.[6]

The analyzer's setup shall consist of:

a. 300 Hz resolution bandwidth.
b. 5, 10, or 20 kHz/horizontal division (as appropriate).
c. 10 dB/vertical division.
d. Reference: carrier peak.
e. Peak Hold: 10 minute duration minimum.
f. No Video Filter.

§ 4. RF Occupied Bandwidth Testing and Control Standard.

§ 4.1. Purpose. The NRSC recognizes that for the purposes of troubleshooting, design, and adjustment it may be desirable to use an occupied bandwidth emission standard that is coupled to a standard test signal and measurement procedure. This second measurement technique can also be utilized by transmitter manufacturers and broadcasters to provide results that may be directly correlated with each other. For these purposes, the NRSC proposes the following interim voluntary national test and control standard:

§ 4.2. Maximum Occupied Bandwidth. The RF occupied bandwidth test and control standard is a stored peak specification. With the standard test signals and measurement procedures specified below, AM stations shall occupy a maximum RF bandwidth that conforms to the following maximum specifications:

---

[6] *See* Klein, *Modulation, Overmodulation, and Occupied Bandwidth: Recommendations for the AM Broadcast Industry* (NAB, September 11, 1986) at 18-23.

Table 2

Frequency Band Relative to Carrier (+/- kHz)	Attenuation Relative to Carrier (dB)
0 to 10	0
10 to 13.5	Minimum is defined by a line with endpoints found at -25 dB/10 kHz and -35 dB/13.5 kHz[7]
13.5 to 54.5	Minimum is defined by a line with endpoints found at -35 dB/13.5 kHz and -65 dB/54.5 kHz[8]
54.5 to 75	-65[8]
above 75	-80[8]

(See Figures 1 and 1A (dotted line) attached).

§ 4.3. Measurement Procedure.

§ 4.3.1. Standard Noise Test. Measurements of AM station spectrum occupancy shall be conducted using a standard noise test signal described in the January 10, 1987 NRSC audio standard.[9] All audio processing employed in the AM station or test configuration shall be in a normal operating mode. The audio signal input to the AM transmitter shall conform to the NRSC audio standard adopted January 10, 1987.

§ 4.3.1.1. Monophonic conditions. The noise source is unmodified.

§ 4.3.1.2. Stereophonic conditions. Two independent but equivalently designed USASI-weighted noise sources are employed. Pulsing of the sources is controlled by a single control signal. The pulsed output of one noise generator is defined as L+R (mono, sum information) where the other is attenuated by 3 dB to

[7]See footnote 3.

[8]For carrier power levels between 50 and 5000 Watts, the maximum limit shall be $-(43 + 10 \log P_W)$ dB (where P_W is the carrier power in Watts) or as indicated in Table 2 and/or the attached Figure 1, whichever is lesser attenuation. For carrier power levels below 50 Watts, a -60 dBC maximum limit shall be used.

[9]*See National Radio Systems Committee, Interim Voluntary National Standard (NRSC-1)*, § 6.3.2 and Figures 4, 5, and 6 (January 10, 1987).

provide L-R (stereo, difference information). The signals are then matrixed to provide left and right channel information to be applied to the audio input terminals of the stereophonic audio processor employed.[10] See Figure 2 (attached).

§ 4.3.2. Use of Spectrum Analyzer. A suitable swept-frequency RF spectrum analyzer shall be used to measure compliance with the NRSC RF occupied bandwidth testing and control standard. The analyzer's setup shall consist of:

a. 300 Hz resolution bandwidth.
b. 5, 10, or 20 kHz/horizontal division (as appropriate).
c. 10 dB/vertical division.
d. Reference: carrier peak.
e. Peak Hold: 10 minute duration minimum.
f. No video filter.

§ 5. Splatter Monitor. It is understood that the NRSC does not anticipate a Spectrum Analyzer will be available for mask standard measurements in most AM radio stations. However, the current development of a low cost "splatter monitor" device may allow economical continuous monitoring of compliance with RF mask characteristics.

Such a device accommodates the following factors: (1) NRSC deemphasis, and (2) amplitude detection of in-phase and quadrature signal components found in the defined stopband region.

As these devices are evaluated and correlated with current spectrum analysis measurement techniques, a determination shall be made by the NRSC with respect to possible modification of RF mask compliance measurement methods.

§ 6. **Effective Date. June 1, 1988.**

§§§§

[10]The signal provided by the audio processor to the transmitter left and right audio input terminals shall not exceed single channel modulation limits as dictated by the constraints of the particular stereo system employed.

FIGURE 1

AM BROADCAST RF EMISSION LIMITS

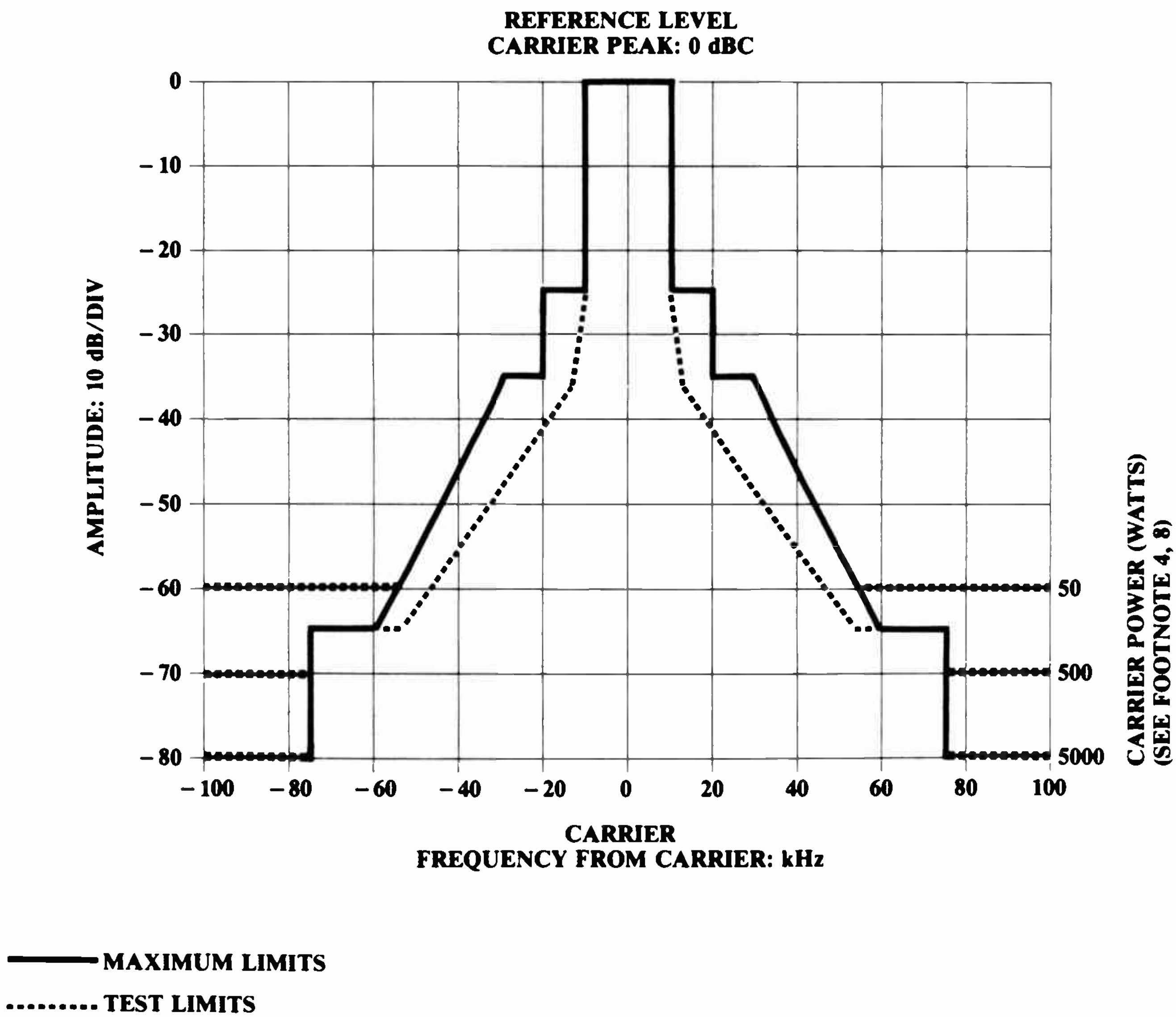
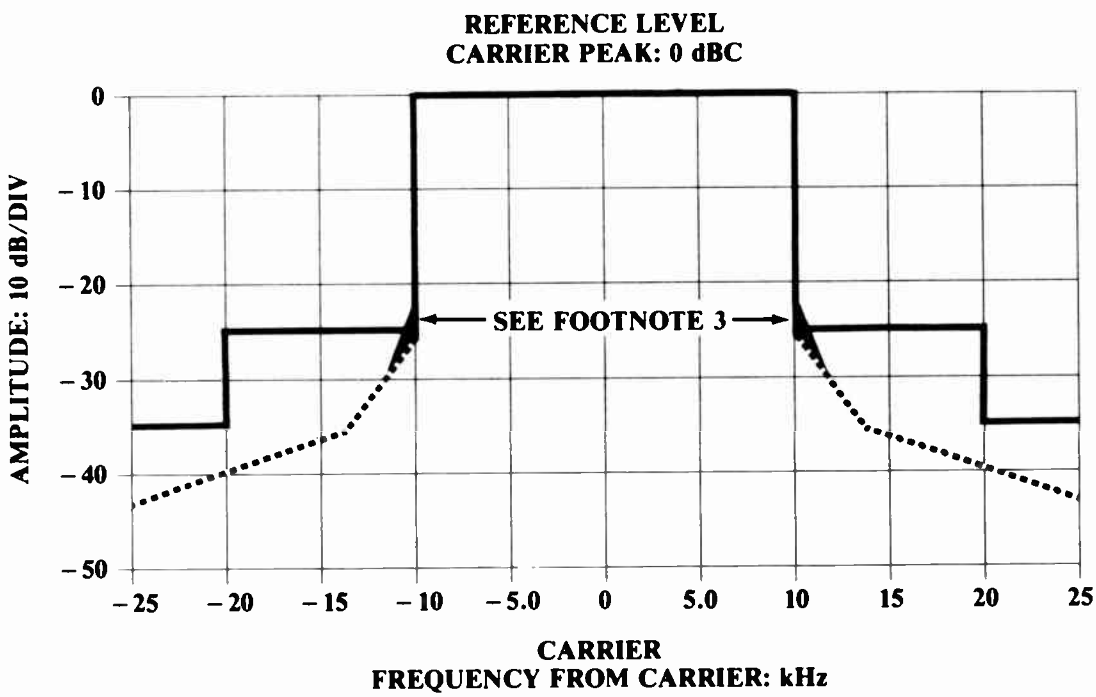


FIGURE 1 A

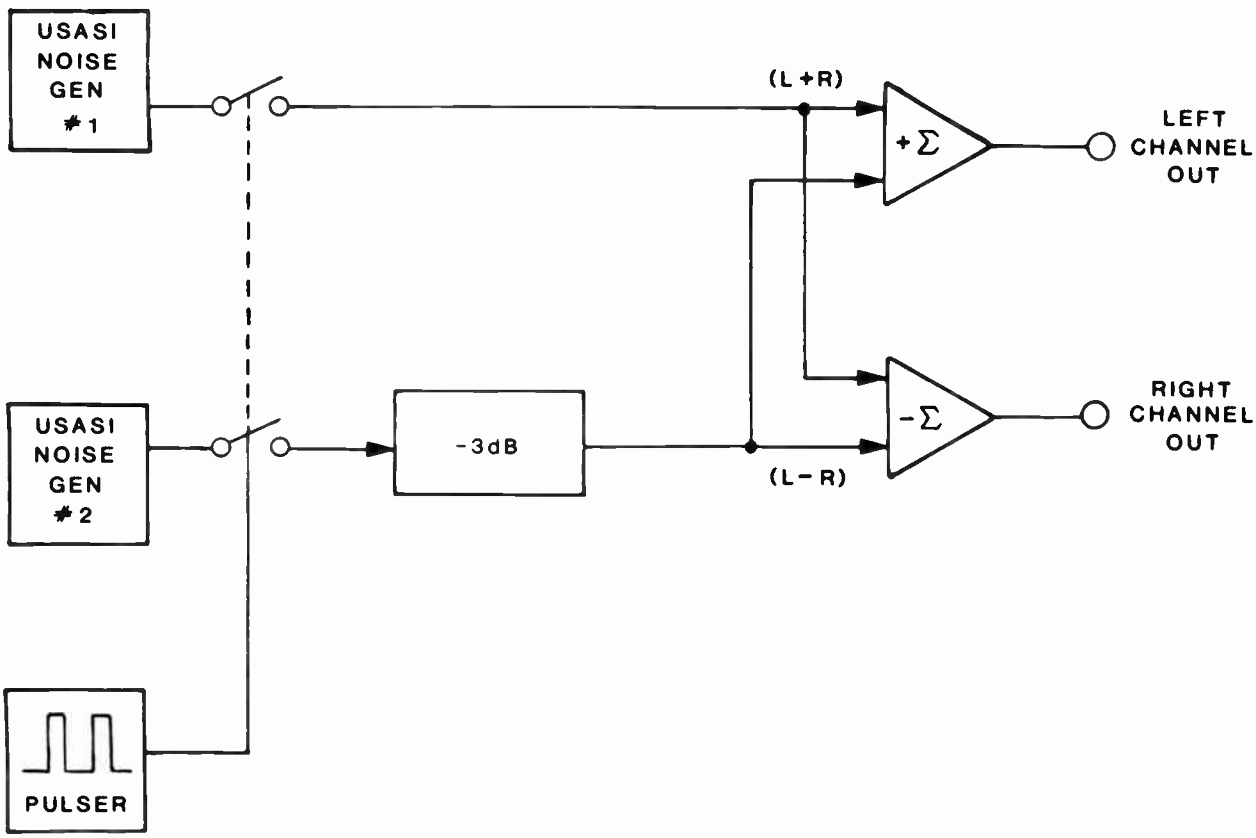
AM BROADCAST RF EMISSION LIMITS

(EXPANDED SCALE)



——— MAXIMUM LIMITS

.... TEST LIMITS


USASI NOISE GEN #1
(L+R)
+Σ
LEFT CHANNEL OUT
USASI NOISE GEN #2
-3dB
-Σ
(L-R)
RIGHT CHANNEL OUT
PULSER

For more information about this document, call:

NAB Department of Science and Technology
(202) 429-5346 or

EIA Engineering Department
(202) 457-4975

3.2
AM Stereo Broadcasting

Edward J. Anthony
Broadcast Electronics, Inc., Quincy, Illinois

INTRODUCTION

The acceptance and implementation of AM stereo broadcasting in the United States has been a slow and difficult transition. Beginning with the now famous FCC marketplace decision in 1981, the field of prospective AM stereo system proponents has been reduced from five (Belar, Harris, Kahn Communications, Magnavox, and Motorola) to two, namely the Kahn Independent Sideband System (ISB) and the Motorola C-QUAM System.

Recent years have shown a quiet growth in AM stereo conversion, with less controversy than was experienced in the early and mid-1980s. The broadcaster's interest has turned more toward proper installation and maintenance of AM stereo than on which system to use. Both remaining methods of stereo transmission have proven to be effective, and it is not in the scope of this chapter to discuss the relative merits of one system over the other. A separate discussion of the C-QUAM and ISB systems may be found in Appendices A and B.

The proper operation of stereo transmission for AM is relatively more difficult than its FM counterpart. Anyone familiar with installing AM stereo will attest to this. It is intended, therefore, to discuss the more practical aspects of AM stereo preparation, installation, and maintenance to aid the AM broadcaster in achieving a high quality stereo transmission.

STATION PREPARATION FOR AM STEREO

Depending on the particular station, part or perhaps all areas of the operation will be affected during the AM stereo installation process. It is best to consider each section of the transmission path separately, so as to make a more thorough and logical conversion.

The Audio Chain

Unless the current AM studio is outfitted with stereo equipment, it will be necessary to install new stereo sources such as cart machines, CD players, consoles, and other distribution facilities. If good engineering practices are employed during the installation and layout phases, this portion of the process should pose no major hurdles.

One area to be concerned about, however, is to insure proper audio phase and amplitude matching of both channels throughout the facility. Without adequate tracking between left and right channels, proper monophonic frequency response and stereo imaging will not be maintained.

Fig. 1 shows the resulting loss in monophonic frequency response due to improper phase relationship between the two audio channels. Another consequence of poor channel matching is a rapid degradation in L+R (monophonic) to L−R (stereo), and L−R to L+R crosstalk. This parameter typically degrades with frequency, reducing high frequency monophonic coverage and altering the high frequency stereo image.

Fig. 2 graphically shows the degree of amplitude and phase matching required to obtain an arbitrary amount of crosstalk. For instance, a 0.1 dB amplitude error combined with 1 degree of phase mismatch will limit L+R to L−R and L−R to L+R crosstalk to about 40 dB. As we shall see later, this relationship holds true for stereo separation during transmitter equalization.

AM Stereo Processing

Due to the nature of AM stereo transmission, the audio paths are different than its FM multiplex counterpart. Rather than operating in a discrete left and right channel mode, AM stereo transmission necessitates the use of matrix mode for transmission, since the amplitude modulation information comes from the combination of L+R (monophonic) audio and the

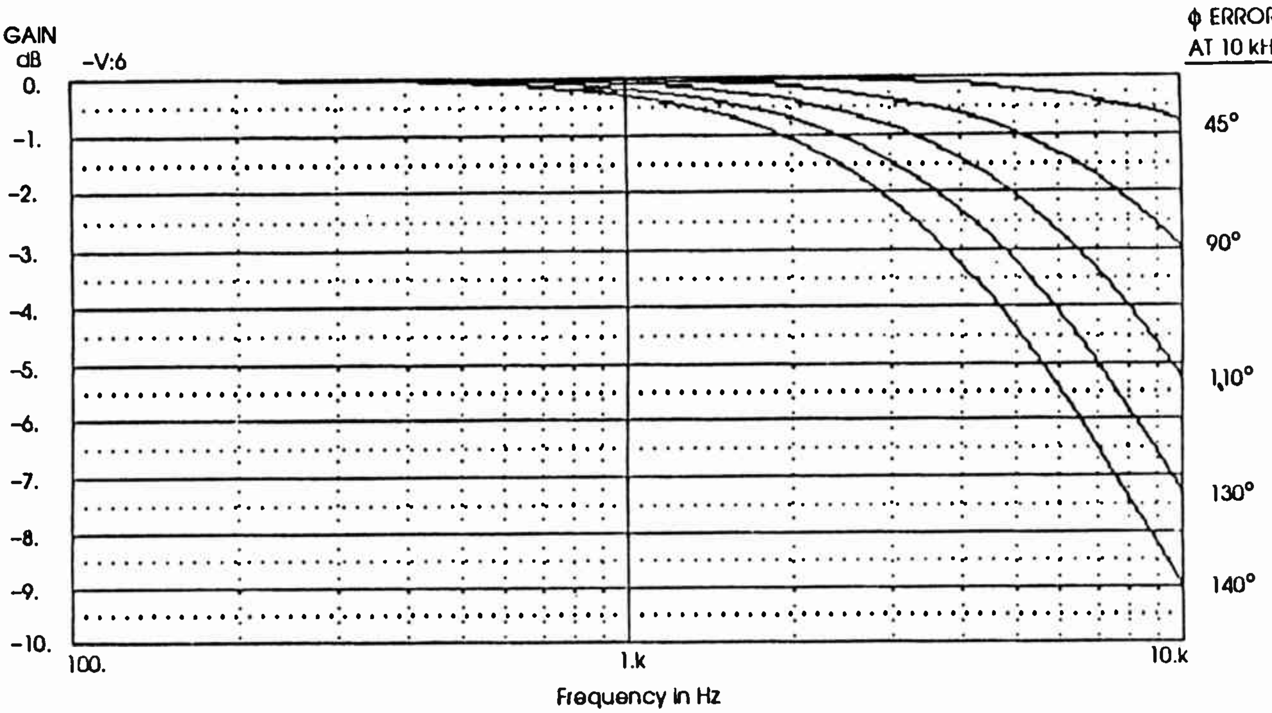
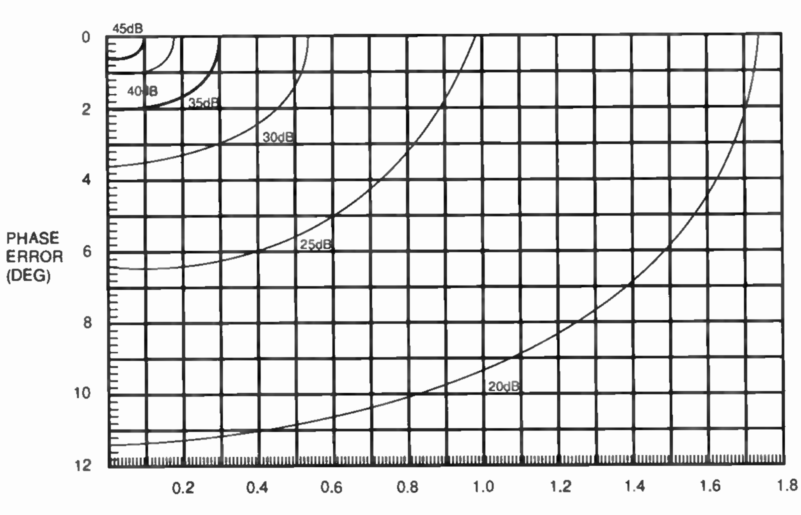


Figure 1. Loss in monaural frequency response due to phase errors.

RF carrier employs some form of phase modulation representative of the L − R (stereo) information. Unlike composite FM stereo, these paths are independent of each other and may also be processed independently. In fact, since the two audio paths are combined in matrix form before the actual modulation occurs, processing for AM stereo is most effective if it is also done in matrix form.

Fig. 3 is a representation of a common and very effective tool used in AM stereo installation and maintenance: the X − Y or Lissajous pattern of a two channel oscilloscope. Standard operation is with the X input driven by the left channel, and the Y input driven by the right. This will produce a straight line display of 45° angle in the first and third quadrants for total L + R modulation, and a straight line display of 45° in the second and fourth quadrants for total L − R information.

If conventional left and right channel processing is used, the limiting must be done symmetrically, and at



General Equation:

$$\text{Separation }(A,\theta) = \left[\frac{(\cos\theta + A)^2 + (\sin\theta)^2}{(\cos\theta - A)^2 + (\sin\theta)^2}\right]$$

Where: $A = \frac{L-R}{L+R}$ Amplitude ratio

$\theta = \frac{L-R}{L+R}$ Phase error in degrees

Figure 2. Crosstalk (separation) vs. amplitude and phase errors in a matrix system.

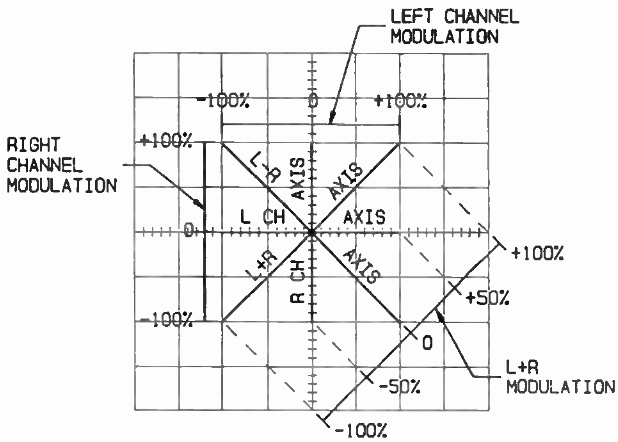


Figure 3. Conventional "X-Y" oscilloscope display.

a level equivalent to 50% L + R and L − R modulation if only one channel were being applied. In this manner, proper 100% limiting will occur when L = R for the envelope (amplitude) modulation, and when L = (− R) for the phase modulation. However, under single channel conditions, there is a 6 dB loss in both monaural loudness and coverage. This undesired effect can be improved if the processing is done in matrix form.

Fig. 4 shows the classic diamond shape produced by full matrix stereo processing. With this method, left and right channels are increased under single channel conditions, and under heavy stereo conditions, to maintain full envelope modulation. This mode of operation is called full monaural support matrix stereo limiting.

Unfortunately, instead of the 6 dB loss in monaural loudness associated with the conventional processing method, the full matrix method results in a 6 dB *increase* in stereo single channel loudness. However, this has been found to be less objectionable than the alternative, and combined with the infrequent nature of single channel conditions, is a more desirable side effect for AM stereo.

There is still one aspect of full monaural support matrix stereo limiting which makes it unacceptable. The unique decoder requirements necessary for the C-QUAM AM stereo system necessitate that the left and right channels be limited to − 75%, where − 100% is equivalent to full envelope modulation caused by a single channel input. In addition, the Kahn ISB system has nonlinear effects caused by single channel modulation greater than 75%. Therefore, all professional AM stereo matrix processors include a single channel limiter to prevent these problems. The new operational area is shown in Fig. 5. Notice the protection areas which have been included to prevent nonlinear stereo operation. This style of AM stereo processing has been referred to as **modified** monaural support matrix stereo limiting.

Proper alignment and operation of any matrix AM stereo processor requires a thorough understanding of these departures from conventional FM stereo processing, as well as the practical limitations imposed by current AM stereo hardware designs. Without a proper grasp on these concepts, the processing can do more harm than good, and has often been the source of great frustration during installation and maintenance.[1]

Studio-to-Transmitter Link

Many AM stations operate studios at locations other than that of the transmitter. Therefore, there must be some form of studio-to-transmitter link to carry the audio information to the transmission facility. Traditionally, this has taken one of two forms: equalized phone lines, or a microwave radio link.

Stereo transmission requires a second link to be installed, and if not done carefully, can be a serious limitation to proper stereo operation.

Phone Lines

Installing a second phone line is often the first choice for many AM stations, especially if they are currently

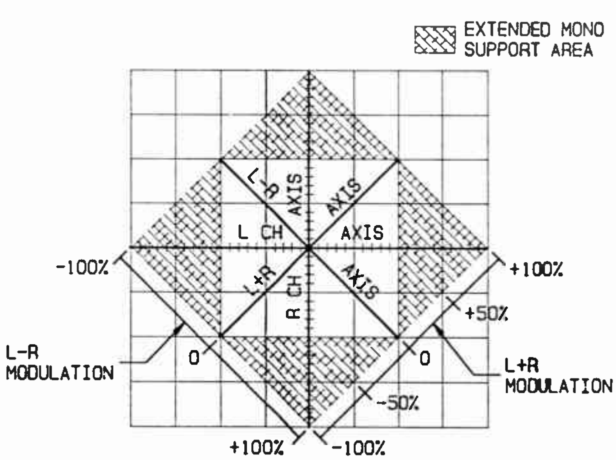


Figure 4. Full matrix stereo processing.

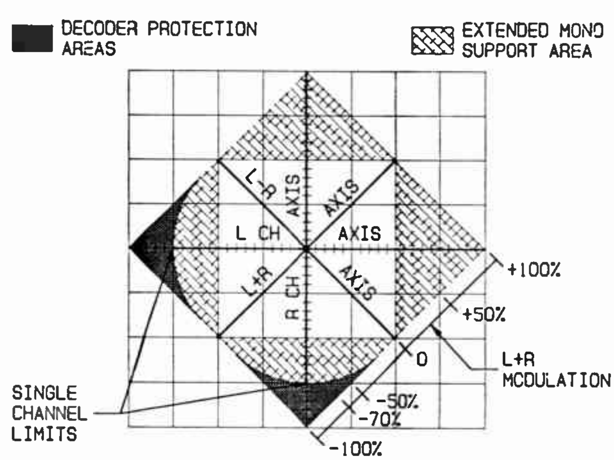


Figure 5. Modified matrix stereo processing.

using one. Two of the more common complaints with this approach are the increase in monthly fees associated with the second line, and the troublesome problem of maintaining the amplitude and phase matching between the two lines.

Equalized phone lines typically have complex filters and frequency shaping equalizers associated with them. This makes them difficult to match, and also prone to drift with time and temperature. The complexity of the circuits can also increase the total possible phase shift, making complete monophonic cancellation, or combing, a problem under extreme conditions. This same phenomenon will also cause complete mono to stereo image rotation, which will cause unexpected stereo imaging and excessive phase modulation.

Unfortunately, these paths are not under the direct control of the station, and one is often left to the mercy of the local telephone company for support.

Radio Links

Installing a new radio link is generally considered a more favorable alternative, providing the frequency allocation is available. It also relieves the monthly financial burden associated with rented phone lines. However, discrete left and right STL transmission can suffer some of the same problems with amplitude and phase matching, but usually not to the same degree as phone lines. The STL transmitters and receivers will have audio, RF and IF filters, and depending on the design, construction and alignment, may not be well matched. It is best to consult the manufacturer to get an assurance that the two systems will track adequately.

There is a third approach which has been used in some locations with great success. An FM stereo generator can be used to create a baseband composite FM stereo signal which can then be transmitted on a single wideband STL link. The signal is then decoded back to left and right at the transmitter location using a high-quality stereo decoder. This approach can be more expensive than discrete radio links, but results in a very high quality, reliable, and stable way to provide the two channels to the transmitter.

Stereo Synthesizers

A final alternative which will be mentioned involves the use of a stereo synthesizer at the transmitter location. This is perhaps the least expensive way to add stereo, since it requires no studio or STL modifications. Rather, a stereo synthesizer is placed at the transmitter to provide a pseudostereo image. This approach will still require the use of a matrix audio processor at the transmitter location, and the user runs the risk of monaural coverage problems due to combing if too much stereo synthesis is employed. However, it is mentioned here as a low cost way to get started in AM stereo.

Transmitter Preparation

Now that the studios are stereo ready, and the two channels are present at the transmitter site, the real job of installing AM stereo begins. Depending on the age and model of the current transmitter, preparing it to accept AM stereo can range from minor to the nearly impossible.

Unfortunately, there are no absolutely defined procedures for making a transmitter stereo-ready. It is highly recommended that a competent and reliable broadcast consultant be used to prepare the transmission chain. A reputable consultant with AM stereo installation experience can save time, trouble, and in the long run, money, since he will know what is required for your unique set of circumstances and can implement them correctly.

In general, the work which will be required can be broken down into a few basic categories.

General Maintenance

The first step to high-quality AM stereo is to have high-quality AM monaural. Many AM transmitters have been neglected for years. A monaural proof of performance will show if the system is up to specifications. If not, it needs to be fixed so that it produces a decent proof. This is definitely money well spent, since it will improve your on air sound to all listeners, both monaural and stereo. It will also make the stereo installation easier, since there will be enough other problems to compensate for without the added burden of poor monaural performance.

Factory Modification Kits

Most manufacturers of AM transmitters will provide support to help upgrade their transmitters for AM stereo. Past experience has shown them where the stereo problems will be for any given model. In addition, some manufacturers have standard modification kits available, with instructions, to prepare the transmitter for stereo.

IPM Reduction

Incidental phase modulation, or IPM, is broadly defined as any undesired angular phase shift of the RF carrier. It can be further broken down into two subcategories: IPM caused by power supply ripple (or induced magnetic fields), and IPM caused by envelope modulation.

The first form of IPM most often results from insufficient power supply bypassing. Before the advent of AM stereo, transmitter manufacturers were not concerned about IPM, since it was not decoded by conventional envelope detectors and generally did not effect the AM signal. It was simply not worth the additional cost to provide the extra bypassing.

Its main effect is to limit the decoded L − R signal-to-noise ratio (SNR), and is often the dominant stereo noise component. Any improvement in this area will improve the decoded stereo signal-to-noise ratio. It should be possible to reduce this form of IPM to −45 to −55 dB below 100% L − R modulation.

The second form of IPM is much more troublesome to reduce, and more detrimental to quality AM stereo. This mode of IPM is caused by amplitude modulation of the carrier, and can be caused by several mechanisms. For tube type transmitters, it is often the result of poor neutralization of the final PA stage, and can result in equivalent L − R modulation levels in excess of 25%.

The interelectrode capacitances found in vacuum tubes can cause instabilities. In particular, the plate to grid feedback capacitance provides a path for positive feedback. Fig. 6 shows the static interelectrode capacitances found in a tetrode. The C_{g-c} and C_{g-p} are the two capacitances of importance for proper neutralization. The intent of neutralization is to cancel the C_{g-p}, so as to minimize its effect, both on stability and on IPM. The most popular method of neutralization is the capacitance bridge. Fig. 7 shows a simplified tetrode configuration employing classic bridge neutralization. C_{g-p} and C_{g-c} are the grid-to-plate and grid-to-cathode interelectrode capacitances, respectively. C is comprised of any input capacitance, including stray capacitance, and is required for complete neutralization of C_{g-c}. C_n is the neutralization capacitance required to balance C_{g-p}. For proper balance (proper neutralization), the capacitor ratios must satisfy the following relationship:

$$\frac{C_n}{C} = \frac{C_{g-p}}{C_{g-c}}$$

With proper neutralization, IPM levels should be able to be reduced, typically, to −35 to −45 dB below 100% L − R modulation.[2,3]

For the solid-state transmitter, IPM is most often caused by the nonlinear output capacitance of the solid-state device. The value of this capacitance is a function of the applied collector or drain voltage. Since most solid-state transmitters use high level modulation,

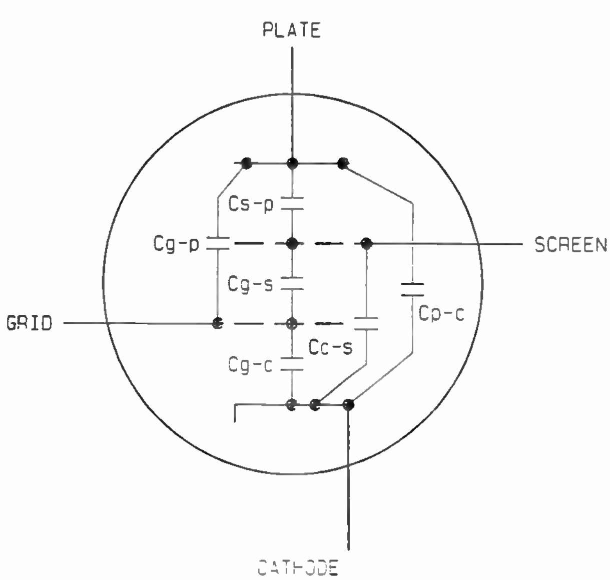


Figure 6. Static interelectrode capacitances of a tetrode tube.

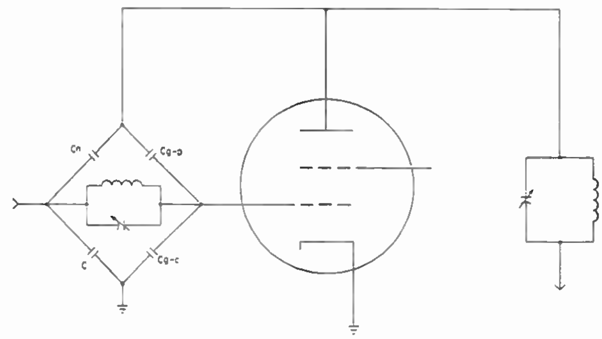


Figure 7. Bridge neutralization.

this voltage varies at the audio rate, which will effect the RF phase angle.

Fortunately, most solid-state transmitters are relatively recent designs, and steps have been taken to reduce this effect to an acceptable level. Many new transmitters are shipped stereo-ready, requiring no additional effort to reduce IPM.

Tuned Circuits

Some older transmitters used narrow bandwidth tuned RF circuits as input, interstage, and output coupling networks. As we shall see later, proper stereo operation requires that the amplitude and phase relationship of the L + R (envelope) information match that of the L − R (phase modulated) information. These narrowband tuned networks have three undesired effects. First, they introduce a nonuniform time delay to the RF carrier, resulting in complex phase equalization problems. Second, they result in L − R high frequency response problems, further aggravating the equalization requirements. Finally, the narrowband nature of the network alters the sideband structure of the phase modulation, resulting in increased distortion.

It is a fairly common practice to increase the bandwidths of these circuits to improve the stereo signal. A word of warning, however: reducing the Q of a tuned stage will result in a decrease in voltage at the output of the network. Care must be taken to insure the resulting signal is adequate to drive the following stage. It is wise to consult the factory or a knowledgeable consultant before attempting this modification.

Antenna and Phasor Alignment

All AM stereo transmissions are sensitive to amplitude and phase nonlinearities to one degree or another. For best performance, the antenna system should be as broadband as possible, and as symmetrical as possible. The C-QUAM system especially requires good transmission bandwidth and phase linearity.

If required, a consultant should be retained to broadband the antenna system and insure its fitness for AM stereo transmission.

Monaural Proof-of-Performance

One final step should be carried out before the actual stereo conversion begins. A complete monaural proof-of-performance should be done to insure the system is ready to accept stereo, and as a record to compare final monaural performance after installation.

STEREO EXCITER INSTALLATION

With a properly operating monaural transmission chain in place, the actual transmitter conversion to AM stereo may begin. Mount the AM stereo exciter close enough to the transmitter to allow proper interconnection. A general rule is the RF interconnection cable should be no more than 30 feet from the transmitter.

RF Interfacing

Picking the RF Insertion Point

Since AM stereo transmission uses some form of phase modulation of the RF carrier, the AM stereo exciter will have a phase-modulated RF output which will replace the transmitter's internal RF oscillator. The rated output power varies from exciter to exciter, from a few watts up to ten watts.

It will be necessary to pick a suitable RF insertion point to break the internal RF chain and replace it with the RF from the exciter. Stereo-ready transmitters will be outfitted with a BNC connector, usually requiring a transistor-transistor logic (TTL) level (5 volts, peak-to-peak) signal for proper operation. Other transmitters require modification to provide a way into the RF sections.

If the particular type of transmitter uses narrow tuned RF stages, it is a good idea to pick the farthest point in the transmitter which the exciter can still provide adequate RF drive.

Many transmitters can accept a TTL level signal but are not outfitted with the necessary interface. Fig. 8 shows a general purpose interfacing network which may be used in these transmitters.

Providing the Proper Termination

The RF output of an AM stereo exciter is designed to operate in a 50 ohm system. This is required due to the extended distances from the RF source and the transmitter. For proper operation, the transmitter end of the 50 ohm coax *must* be terminated with 50 ohms to prevent reflections on the line. This is especially important if the RF output is a square wave, which it is most likely to be. Improper termination will cause reflections on the line, which can cause excessive ringing and, for TTL inputs, improper logic operation. If the ringing is excessive, it is possible to damage the input stage due to excessive voltages. For TTL compatible input levels, a ¼-watt resistor is adequate, provided that the input to the transmitter is AC coupled, then DC restored (Fig. 8). If higher power levels are required, select the proper resistor power rating based on the following formulas:

$$P_{diss} = \frac{(V_{rms})^2}{50} \text{ (for sine wave input)}$$

$$P_{diss} = \frac{(V_{peak})^2}{50} \text{ (for square wave input)}$$

Internal RF Oscillator

Since the transmitter's original RF oscillator has been replaced by the AM stereo exciter, some possible problems may need to be dealt with. First, there may be some coupling between the original oscillator and the new RF signal. This coupling manifests itself most often in the form of a beat frequency equal to the difference in the two oscillators frequency calibration.

The most common cure is to simply remove the original crystal from the transmitter, or disabling the power supply to the oscillator. This will prevent beat frequencies, and is minor enough modification to allow quick replacement should the original oscillator be needed in an emergency.

If the old oscillator should be required due to a failure in the AM stereo exciter, it would be nice to be able to quickly switch back to the original oscillator and return to the air, in monaural, until the problem can be resolved. This can be done fairly easily with a relay. Fig. 9 shows one possible implementation of an RF switching circuit. Notice the extra relay contact used to remove the power from the original oscillator during normal stereo operation. Again, this eliminates the possibility of beat frequencies.

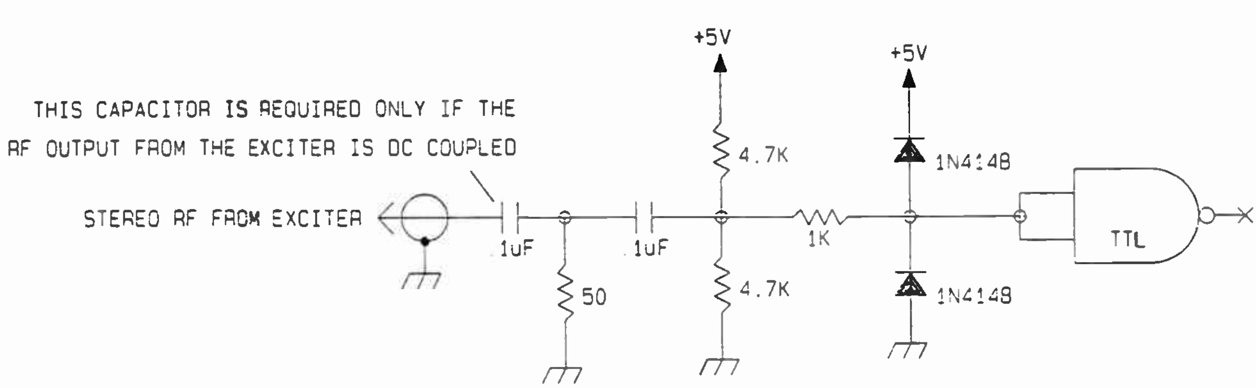


Figure 8. TTL RF interfacing.

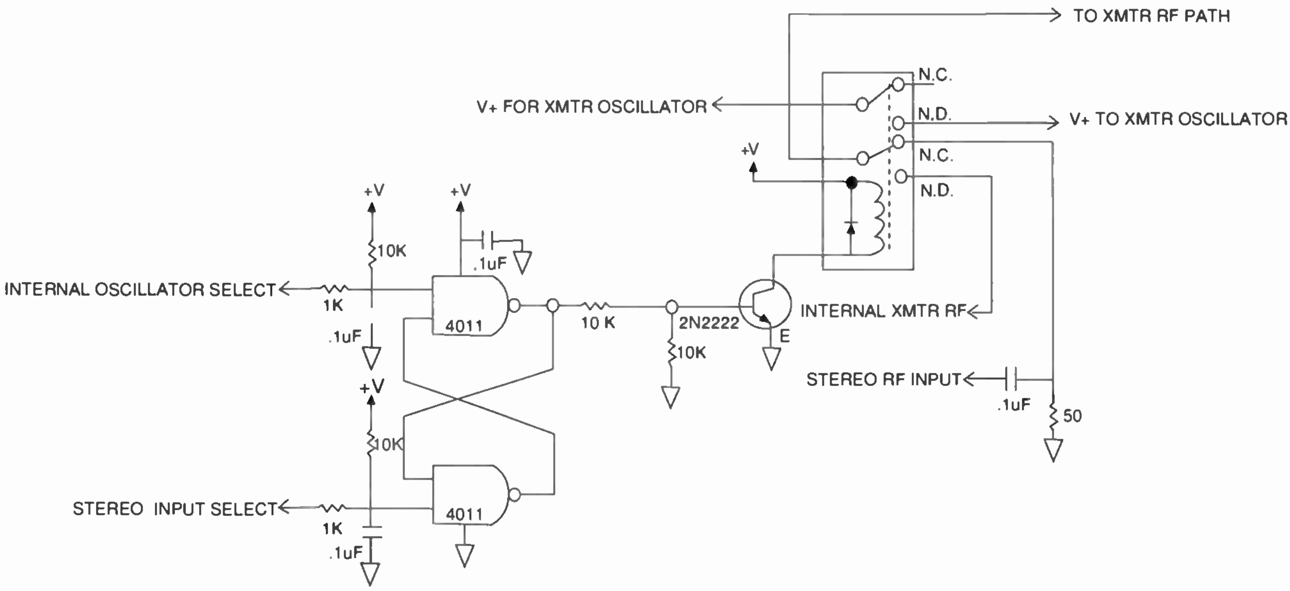


Figure 9. Mono/stereo emergency switching.

Audio Interfacing

The left and right channels must be connected to the exciter, and the levels set according to the manufacturers recommendations. The existing monaural input to the transmitter should be replaced by the monaural output from the AM stereo exciter. All audio connections should be made using a high quality shielded cable, such as Belden 8451. This is important to prevent RF pickup on the cable. To prevent ground loops, ground the shield at only one end.

AM STEREO EQUALIZATION

Now that all the necessary pieces are in place, the final part of the AM stereo installation can begin. Since half of the stereo information is contained in the monaural (L + R) connection to the transmitter, and the other half is in the form of phase-modulated RF (L − R), they must arrive at the modulator portion of the transmitter at the same time. This is required so that proper left and right channels may be recovered in the receiver after dematrixing. The amount of amplitude and phase (delay) matching required at the transmitter's modulator to produce a given amount of stereo separation after dematrixing can be seen by referring back to Fig. 2.

Without proper audio equalization, both stereo separation and distortion would suffer. Unfortunately, the amount of delay and amplitude equalization required for one type of transmitter is different from any other. In addition, the circuits in the transmitter which cause the problems are generally complex, so the best that any AM stereo exciter equalizer can do is to approximate the response shape. The effectiveness of the equalizer is dependant on the type of transmitter and the complexity of the antenna pattern at any given station.

For these reasons it is clear to see why there are no firm rules to follow during equalization. It is often a trial and error routine, especially at the higher audio frequencies where the response shape is most complex. If an AM stereo exciter were designed specifically for one brand of transmitter, it would be possible to equalize the two paths almost perfectly, but since it is a generic conversion tool, it can, at best, only come close.

Types of Equalization

There are three different equalization sections in an AM stereo exciter. Each is intended to fix a particular frequency range. Some installations require all three, others only two, and very rarely, only one will be needed.

Group Delay

The first section which is almost universally required is commonly labeled the group delay equalizer. Its purpose is to provide a constant time delay to either the L + R or L − R paths, depending on which one arrives at the modulator first. It is primarily used to equalize the low to mid frequencies (up to a few kilohertz). If the response shape of the offending path in the transmitter is low pass (or band pass for the RF path), the time delay for the lower frequencies will be constant, so the alternate path can include this equalizer to compensate adequately.

A common example would be any AM transmitter employing pulse width modulation (PWM), in which case the L + R path would include a PWM low-pass filter in the modulator just ahead of the RF amplifier. The L + R audio will have significantly more delay than the L − R (RF phase modulation) path. In this case, the group delay equalizer would be added to the L − R path to slow it down so that it arrives at the RF amplifier at the same time the L + R signal does.

On the other hand, some transmitters will have more low/mid-frequency delay in the L − R path due to several RF interstage band-pass filters. In this instance, the group delay equalizer would be added to the L + R path to compensate.

Low-Frequency Equalization

The second type of equalizer found in the exciter is the least often used section. It is only necessary if the low-frequency response of the L + R path is nonuniform. This is found primarily in older plate modulated transmitters where the reactance of the plate transformer causes phase, and to a lesser degree, amplitude nonlinearities.

Some transmitters will also include an active audio high-pass filter in the monaural input. This can also be effectively corrected by the low-frequency equalizer.

High-Frequency Equalization

The third audio equalizer is designed to approximately complement the complex high frequency response shape of most AM transmitters. The amount and shape of high frequency correction varies greatly from transmitter to transmitter, which makes this section the most difficult adjustment of the installation process.

Since this section is only meant to approximate the high-frequency amplitude and delay characteristics, it will not correct at all frequencies. It is not uncommon to be caught in a seemingly endless loop of adjusting for one frequency, only to find out you have made another frequency worse. The key to a successful installation is the ability to recognize a good overall compromise, and once you have, to quit. This is one area where a good engineering consultant with AM stereo installation practice will be helpful.

Fig. 10 shows the various equalization adjustments found on one C-QUAM AM stereo exciter. Notice the additional set of equalization labelled "Night." This is useful for stations utilizing a different transmitter, antenna pattern, or both for part of the broadcast day.

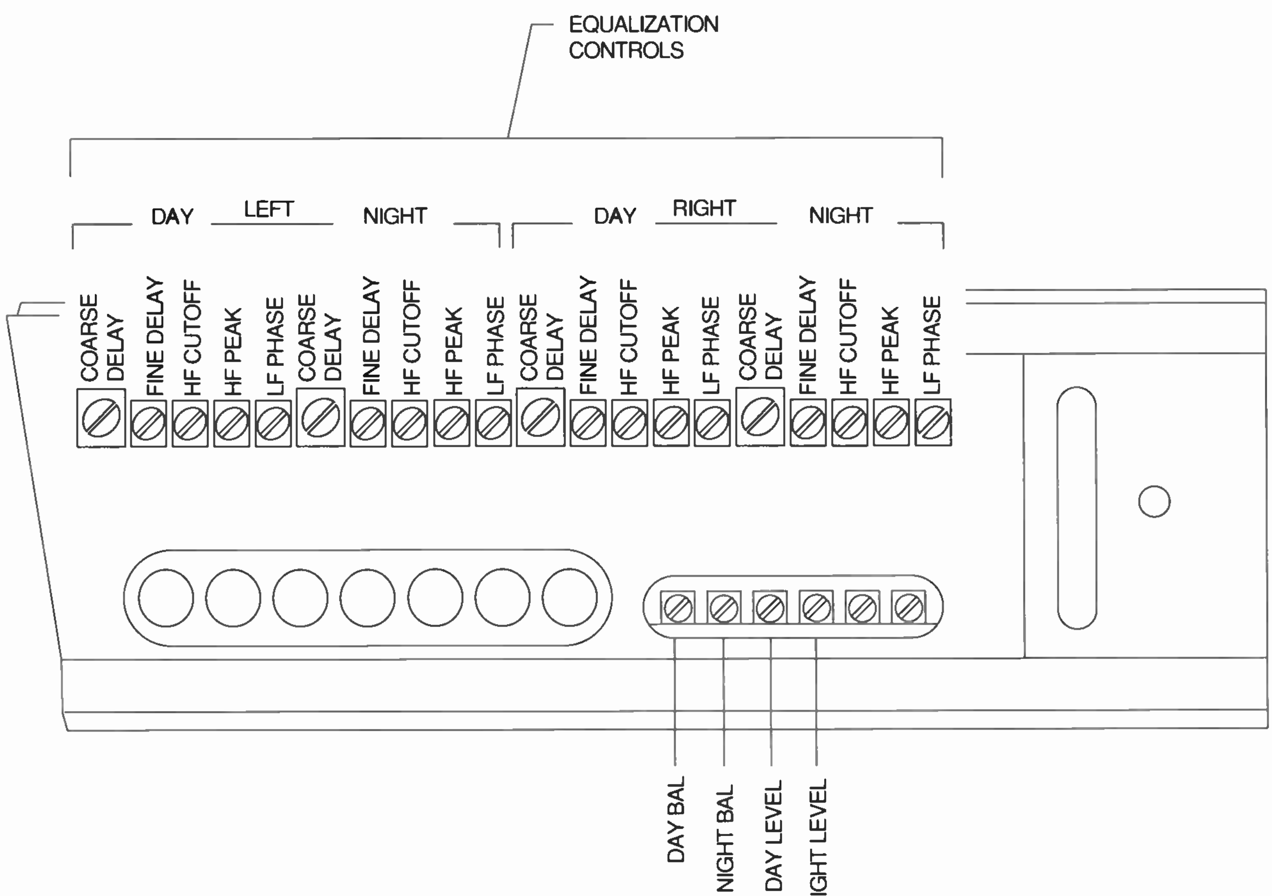


Figure 10. Broadcast Electronics Model AX-10 stereo equalization.

It can also be used to equalize an emergency backup transmitter.

Equalization Path Selection

The proper alignment of the three equalization sections is the key to high-quality AM stereo. However, before they can be adjusted, they must be inserted into either the L + R path or the L − R path, depending on where they are needed.

For someone with an intimate knowledge of the design of a particular transmitter, it may be possible to know for sure which path needs which equalizer. However, it is also possible that unique antenna characteristics could change what would be necessary into a 50 ohm resistive load. What results is a "try it and see" method. It may be necessary to change paths during the installation procedure as other equalizer sections are included. This will happen when, for instance, the inclusion of the high-frequency equalizer in series with the group delay equalizer actually forces the constant delay to be added to the other path to compensate for low- and mid-frequency delay found in the high-frequency equalization.

Fig. 11 shows one method of selecting the proper path for any equalizer. This matrix switch is a flexible patch bay approach to inserting equalizers in series or parallel into either the L + R or L − R paths.

Equalization Procedure

Fig. 12 outlines a typical test setup required for AM stereo equalization. Test equipment required for an adequate installation can be as simple as a low distortion audio oscillator/analyzer, an AM stereo modulation monitor, and a dual trace oscilloscope with X − Y capability. In addition, a spectrum analyzer capable of resolving audio sidebands at the AM band frequencies can be very useful, but is not mandatory.

The X-Y Oscilloscope

The fundamental problem in AM stereo is determining the type and path of equalization required. In

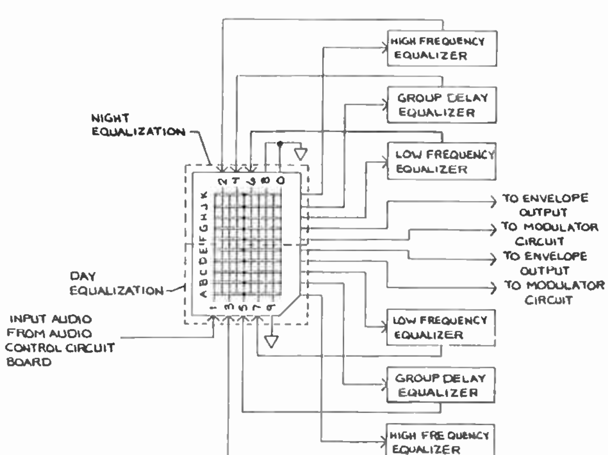


Figure 11. Broadcast Electronics Model AX-10 matrix switch.

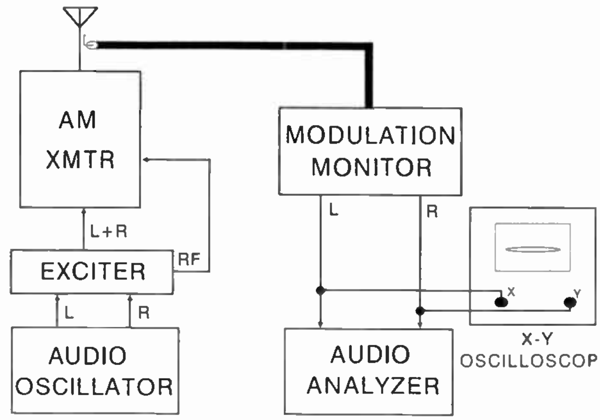


Figure 12. Stereo equalization test setup.

addition, it is necessary to determine if the system requires phase or amplitude equalization. A dual trace oscilloscope operating in the X − Y or Lissajous pattern mode is capable of resolving between amplitude and phase errors (or a combination thereof), as well as showing a multitude of additional information such as harmonic distortion, IPM, negative limiting, modulation compression, and so on. To the trained eye, it is one of the most useful tools in AM stereo installation. Without this tool, it would be almost impossible to equalize the system.

AM Stereo Modulation Monitor

The final performance results of a stereo installation will only be as good as the equipment used to measure them. One of the most critical pieces of test equipment for proper installation, and just as important, routine maintenance, is the AM stereo modulation monitor. It should be a high quality piece of test equipment, and simple to use.

Fig. 13 is a picture of one commercially available monitor that incorporates a high-quality C-QUAM decoder for high-performance stereo decoding. In addition, the use of an RF AGC insures constant decoder performance over fluctuations in RF input level. Front panel audio monitoring connectors and autoranging meters simplify the job of installation and maintenance.

An Equalization Example

AM stereo exciters will have similar equalization adjustments, but may be labelled differently. The *day* adjustments are used, however, the *night* procedure is identical.

The first adjustment to be made is to set the exciter's L + R, or envelope, level to get the correct amplitude modulation for a corresponding L = R input. This is, at this stage, only a coarse adjustment. It can be done at any convenient modulation percentage. For instance, input 1 kHz L = R to the exciter such that the L + R meter on the exciter reads 80%. Then adjust the envelope output level of the exciter until the transmitter is modulating 80% as measured on the modulation monitor.

Figure 13. Broadcast Electronics Model AS-10 C-Quam AM stereo modulation monitor.

Next, apply a 1 kHz tone to the input of the left-channel only, at a level equivalent to 50% amplitude modulation. The oscilloscope should show a horizontal line (assuming the 'X' input is the left channel). Most likely, it will be neither exactly on the X axis, nor a perfectly straight line. Fig. 14A shows a typical display. If a display similar to Fig. 14B is seen, the polarity of the L + R connection to the transmitter is wrong. This can easily be fixed by reversing the plus and minus terminals at the transmitter input. Fig. 14B actually shows the high second-harmonic distortion produced when the polarities of the amplitude modulation and phase modulation are not correct.

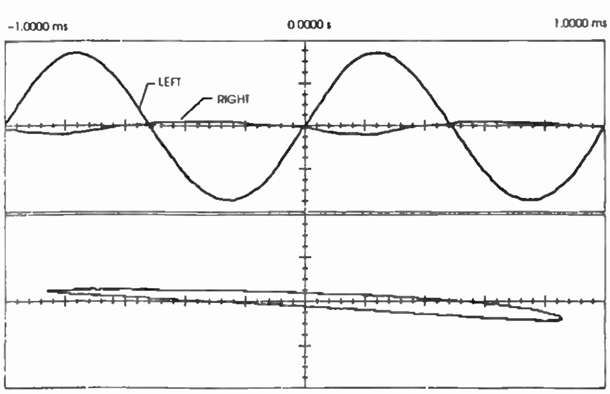


Figure 14A. Typical left channel only display before equalization.

Next, the fine adjustments to maximize separation can be made. While monitoring both the residual right channel separation and the oscilloscope, insert the left channel group delay in either the L + R or L − R path (Fig. 11) and adjust the coarse group delay. If the display opens up (Fig. 15) rather than closes, and the separation worsens, it has been inserted in the wrong path. Reverse the location and adjust the coarse and fine group delay until the display closes (Fig. 16). Then, fine tune the envelope level until the display lies exactly on the X axis and the left to right channel separation nulls (Fig. 17).

Apply a 1 kHz tone to the input of the left-channel only, at a level equivalent to 50% amplitude modulation. The oscilloscope should show a vertical display with only phase error. Insert the right channel group delay in the same path as the left channel and adjust until the display closes. The residual separation in the left channel will also null. Theoretically, the envelope level adjustment should not have to be adjusted for the right channel, but often the right to left separation can be improved slightly by fine tuning the level, at the expense of left to right separation. It is best to

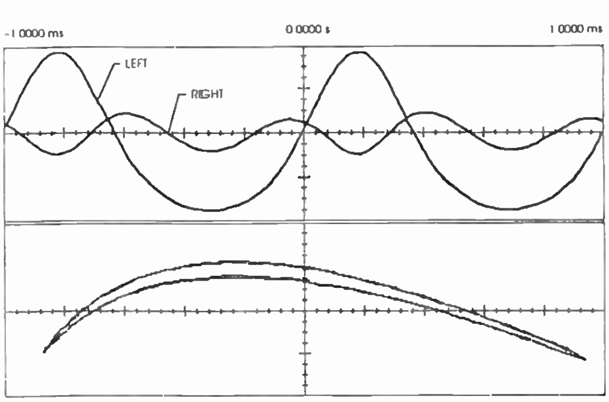


Figure 14B. X-Y display caused by improper L + R polarity at transmitter.

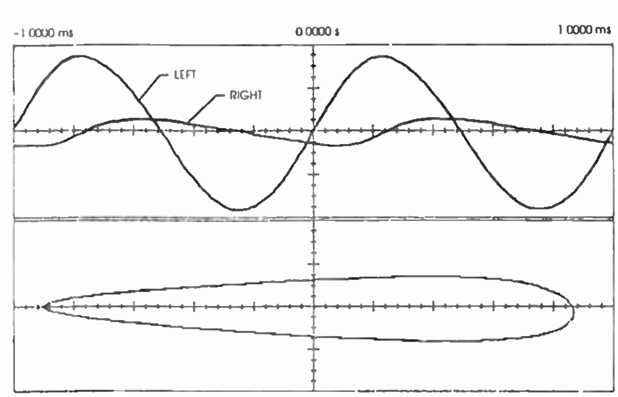


Figure 15. Left channel only with phase equalization error.

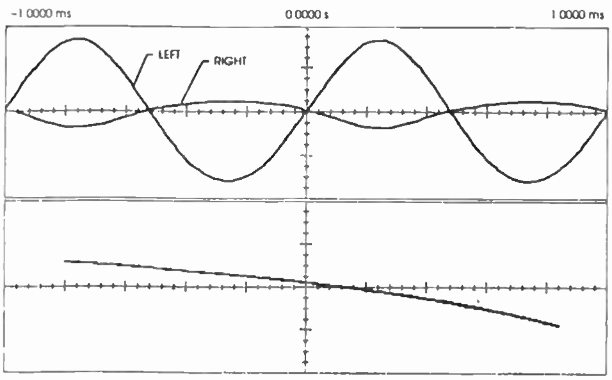


Figure 16. Left channel only with amplitude equalization error.

compromise the two to obtain the same amount of separation left to right as right to left. Some exciters are equipped with a balance control to help this problem. Refer to the individual service manual for the correct adjustment procedure.

The next step is to check the low-frequency separation. A good frequency to look at is 100 Hz. If the separation has degraded significantly, the low frequency equalizer can be used to correct the problem. Again, this is usually only necessary in plate modulated transmitters, or ones which employ low level audio input high-pass filters. The procedure is very similar to the group delay adjustment.

If low-frequency equalization is used, recheck separation at 1 kHz and adjust the group delay, if necessary, to compensate.

At this point, the transmitter should be equalized for good separation (30 dB to 40 dB) from 50 Hz to a few kilohertz. A quick check of separation at 5 kHz to 10 kHz will most likely show a rapid degradation in separation performance.

With the left channel high-frequency equalization controls set at minimum, insert the HF equalizer in one of the paths. Input 7.5 kHz left channel at a level sufficient to produce 50% envelope modulation. Observe the oscilloscope display while adjusting the HF cutoff control. If the phase degrades (display opens), then switch the paths. If it improves, adjust until it closes. Input 1 kHz and readjust the group delay to compensate for the added mid-frequency delay, and repeat until 1 kHz and 7.5 kHz are equalized for phase.

If the display at 7.5 kHz is not on the X axis, adjust the HF peaking and then the HF cutoff until the display lies on the X axis with minimum phase error. Again, adjust the group delay control at 1 kHz to compensate and repeat the procedure until maximum separation is achieved at both 1 kHz and 7.5 kHz. Repeat the procedure for the right channel.

Once this is done, spot check the separation performance from 1 kHz to 10 kHz. Using good judgement, it may be necessary to compromise the performance at 7.5 kHz to improve it elsewhere. The procedure is identical with the exception of X kHz in place of 7.5 kHz. While there is no such thing as a typical proof, it should be possible to achieve greater than 20 dB separation out to 10 kHz.

The Stereo Proof

When satisfactory separation performance is obtained, the final step in the AM stereo installation is to run a full proof of performance, both monaural and stereo. In particular, pay attention to high-frequency distortion. It may be necessary to fine tune the equalization to reduce distortion at the expense of a few dB of separation. An acceptable proof should show less than 3% THD and greater than 20 dB separation at all frequencies. This proof is a valuable tool during routine maintenance to gauge performance and to assess whether the equalization needs adjusting.

ROUTINE MAINTENANCE

Any change in the transmission system amplitude or phase response will cause a degradation in stereo performance. These changes will occur for several reasons; including seasonal environment changes, routine transmitter tuning, and vacuum tube variations over its lifetime.

It is a good practice to include a spot check of stereo performance as a routine maintenance item. This should be done at an interval of no more than every six months. Left unattended, these normal system variations can cause a serious degradation in audio quality, but a small time investment will insure continued high-quality AM stereo performance.

SUMMARY

Properly installed, AM stereo can offer many advantages to the AM broadcaster with the ever increasing pressure for high fidelity audio. Good planning, execution and follow up to the installation, are essential to avoiding the creation of new problems.

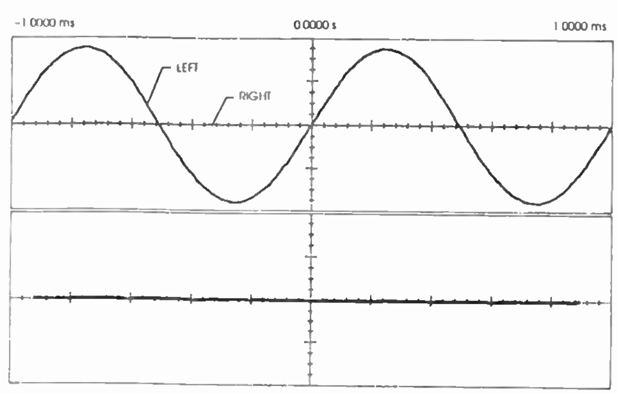


Figure 17. Left channel only, properly equalized.

REFERENCES

1. "Stereo Processing," Circuit Research Labs, Inc., Tempe, Arizona, 1983.
2. Stanley, James S., "Engineering Report: Wilkinson AM-1000B Stereo Conversion," Stanley Broadcast Engineering, Tempe, Arizona, 1987.
3. Orr, William I., *Radio Handbook*, 20th Edition, H.W. Sams & Co., 1975.
4. Orban, Robert, and Oganowski, Greg, "AM Broadcast and Receiver Standards—Towards Opening a Dialogue," SAE Technical Paper Series, Warrendale, Pennsylvania, 1985.
5. "Electronic Circuit Analysis 2.50 MS-DOS Version," Copyright 1990 by Tatum Labs, Inc., Ann Arbor, Michigan.
6. Anthony, Edward J., "Second Generation Techniques for AM Stereo Exciter Design," Broadcast Electronics, Inc., Quincy, Illinois, 1984.

APPENDICES

Appendix A: Independent Sideband AM Stereo

Appendix B: Motorola C-QUAM AM Stereo System

These appendices were prepared by the proponents of the systems. NAB takes no position on their relative merits and recommends contacting the proponents for further information.

APPENDIX A

INDEPENDENT SIDEBAND AM STEREO

Leonard R. Kahn
Kahn Communications, Inc., Carle Place, New York

INTRODUCTION

Under good reception conditions, and with wideband AM receivers, AM stereo systems can provide performance virtually equivalent to FM.

AM broadcasting's major strength is its superior coverage (for Class I and II stations), which is due to its frequency band and the modulation method which allows acceptable reception under poor signal-to-noise and signal-to-interference conditions. Wideband FM trades bandwidth for improved signal-to-noise performance under good signal-to-noise and interference reception conditions. FM is not, however, noted for its good performance under poor reception conditions.

The ISB stereo system is a frequency separation system and is free of phase sensitivity. It is able to utilize a stereo decoder that remains in stereo under all conditions of reception.

FUNDAMENTALS OF ISB AM STEREO

The Kahn/Hazeltine AM stereo system is an independent sideband (ISB) system wherein the L stereo channel information is transmitted by the lower sideband, and the R stereo channel information by the upper sideband. The envelope of the wave is a linear function of L+R. A suitable stereo receiver for the Kahn/Hazeltine system is essentially an ISB receiver, and stereo separation is determined by the frequency of sidebands relative to the carrier. Thus, the Kahn/Hazeltine system is properly designated as a "frequency separation" system. The stereo separation characteristic of the Kahn/Hazeltine system is not a function of the phase of the carrier component relative to the sidebands, as is the case in other AM stereo systems, which may be described as "phase separation" systems.

Basic structure:	**Independent sideband**
Stereo information provided by:	Frequency separation Left Lower SB Right Upper SB
Pilot:	15 Hz, 0.1 radian
Envelope function:	L+R
Preferred receiver stereo decoder:	Uses inverse modulation by L+R function

One of the basic disadvantages of an AM wave (and, indeed, FM waves) is that the sidebands must be properly phased relative to the carrier. Accordingly, any phase anomaly in the path of the AM wave, such as narrowband antennas, antenna null areas, close-in fading, reradiation from power lines and buildings, off-tuned receivers, and co-channel interference can create a phase shift between the carrier and the sidebands. This can cause (1) loss of desired envelope modulation, and (2) envelope distortion.

ISB AM stereo signals can be received under fading conditions, in antenna nulls, and also under co-channel interference, in some cases over that of conventional AM monophonic transmission. The reason for this is that, for an undistorted, fully-modulated envelope, sidebands need not be symmetrical in an ISB stereo transmission.

Referring to Fig. A-1, it is seen that for a conventional AM wave, when the carrier is shifted relative to the sidebands by 90° the fundamental component of the envelope modulation disappears. This is true for all odd multiples of ±90° (i.e., ±90°, ±270°, ±450°, . . .). Thus, when an envelope detector is used, the distortion is infinite. With a synchronous demodulator, the distortion is indeterminent because while there is no distortion, there is also no signal. The phasor diagrams in Fig. A-1 illustrate the effects of carrier/sideband phase errors on conventional AM signal and the worst case condition of a form of ISB AM stereo.

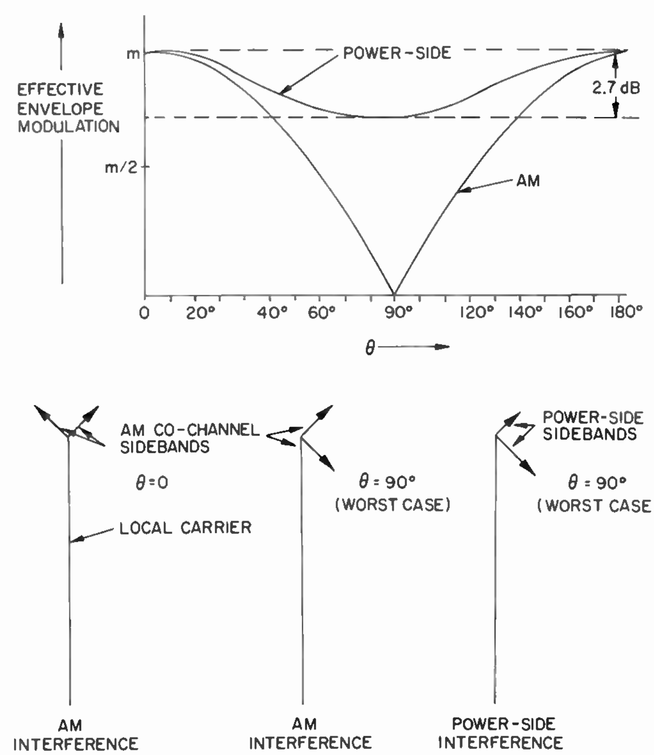


Figure A-1. Phasor diagrams for conventional AM wave and a form of ISB AM stereo wave, showing effect of carrier-to-sideband phase shift.

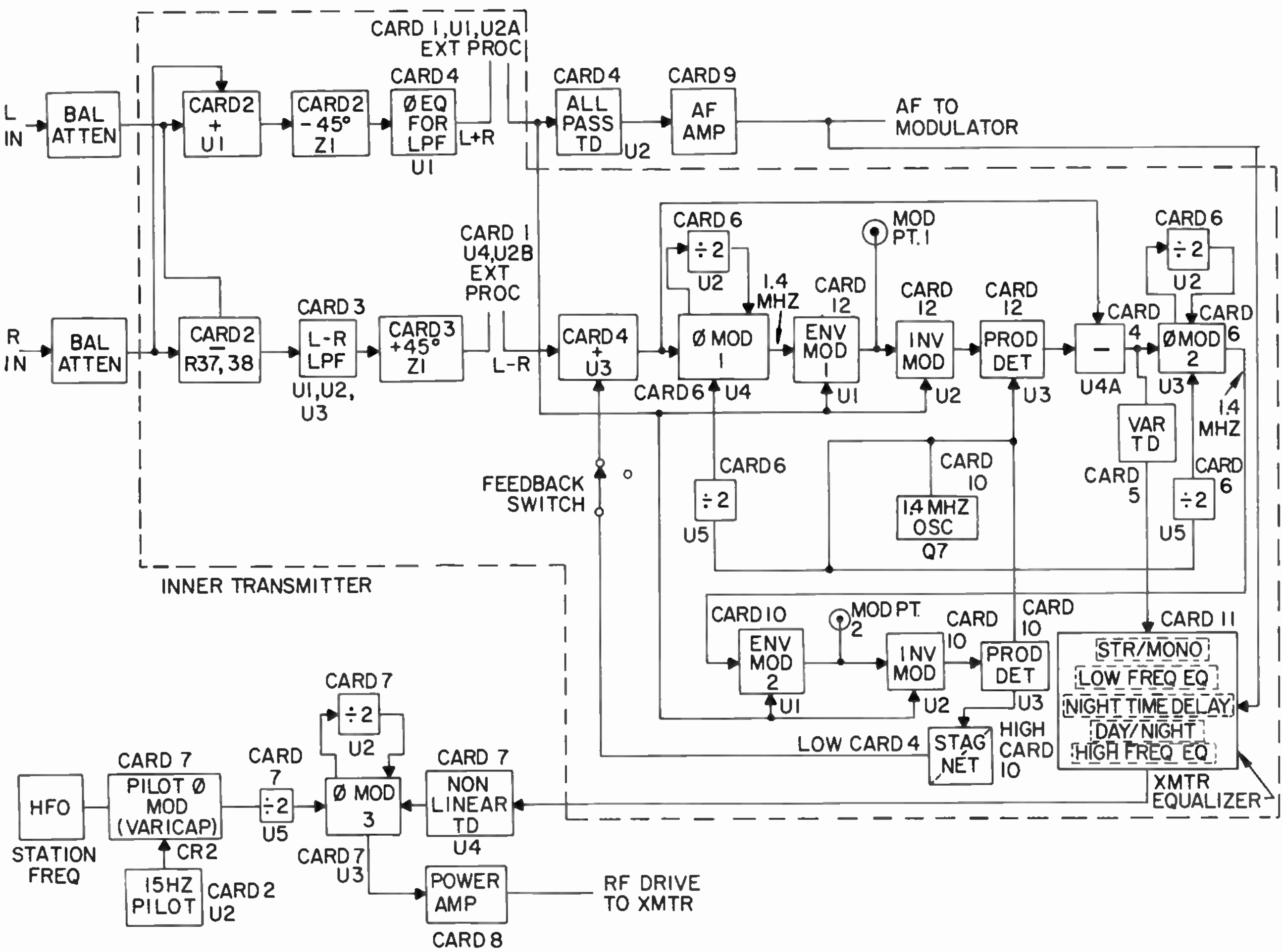


Figure A-2. Block diagram of Kahn/Hazeltine Model STR-84 Independent Sideband AM Stereo exciter.

DESCRIPTION OF ISB AM STEREO

Fig. A-2 is a block diagram of an ISB AM stereo exciter. Included in the following description is a brief analysis of the spectrum and distortion characteristics of this equipment.

Most of the blocks shown in Fig. A-2 are enclosed in dotted lines. These blocks comprise the "inner" two low-level transmitters. Both transmitters operate at 1.4 MHz, unless the station's operating frequency is too close to 1.4 MHz to avoid spurious problems. The output of the second inner transmitter is a low-power, low-distortion AM Stereo signal.

The circuitry outside the dotted lines serves three basic functions:

1. Allows the input signal to be adjusted to the proper level for the inner transmitters.
2. Preconditions the L + R and L − R signals fed to the station's transmitter so as to cause the station's transmitter to track the performance of the second inner transmitter, thereby producing a low distortion AM stereo signal.
3. Produces a 3-watt phase-modulated signal with the necessary L − R component and 15 Hz stereo pilot.

The L and R inputs are fed to attenuators, which in turn feed sum and difference networks. The output of the sum network, the L + R wave, feeds a phase equalizer. The phase equalizer matches the phase characteristic of the low-pass filter in the L − R channel, which ensures that under all conditions of modulation, adjacent channel interference is not significantly increased and all occupied bandwidth regulations are fully met.

The output of this equalizer feeds a phase-difference circuit. This all-pass network, in conjunction with the all-pass phase shift network in the L − R path, provides a constant phase difference of approximately 90° over the desired stereo frequency range.

The output of the phase-shift network feeds an output terminal which may, in turn, feed an external L + R audio processor.

If external audio processing is not utilized, the phase-shift network can be connected directly to a time-delay circuit, which in turn feeds the audio input circuit of the external transmitter. An advantage of providing system compatibility with an L + R audio processor is that any waveform characteristics imparted to the L + R wave (such as clipped peaks) can reach the transmitter without tilting the envelope of the transmitted wave. On the other hand, any clipping of the L and R waves fed to the input of the exciter will generally cause tilting of the audio wave. Therefore, audio clippers or asymmetrical modulation circuits should only be used in the designated external processor ports or *after* the exciter.

The output terminals for the composite L + R limiter feed a fixed time-delay circuit. This delay circuit is provided to cause the net delay in the overall L − R path (including the RF circuits of the associated transmitter) to be less than the delay in the overall L + R path (including the modulator circuits in the associated transmitter), by an amount that falls within the adjustable range of the variable time-delay circuit in the L − R path.

The output of the time-delay circuit feeds a DC coupled amplifier (including protective clipping circuits), which in turn feeds the associated transmitter.

The L − R path, is more complex than the L + R path. The L and R inputs feed the linear difference circuit. The output of the difference circuit feeds a low-pass filter, which provides a flat response (±0.5 dB) up to 6.24 kHz and allows some separation up to almost 7.5 kHz. As mentioned above, this filter is provided to limit the L − R frequency response to restrict adjacent channel interference which would be caused by the second order sideband component necessary for envelope detector compatibility.

The output of the LPF is fed to a phase network which, as mentioned above, provides, in cooperation with the phase network in L + R path, a quadrature relationship between the L − R and the L + R audio waves. The output of the phase network may be fed to an outboard composite processor which functions in cooperation with the composite limiter in the L + R path. The output of the external L − R processor limiter in the L + R path. The output of the external L − R processor feeds a summation circuit.

Assume, for example, that negative feedback is absent (i.e., the feedback switch is in the open position). The L − R, without any feedback term at the output of the summation circuit, feeds the phase modulator. Also feeding the phase modulator is a divide-by-two frequency divider, which in turn is fed by a 1.4 MHz crystal oscillator. The phase-modulated output of the phase modulator is a 1.4 MHz because this phase-locked-loop (PLL) circuit also incorporates a divide-by-two frequency divider, resulting in a frequency multiplication of two. The output of the PLL phase modulator is amplitude modulated by the L + R audio wave. When the proper amounts of phase modulation and amplitude modulation are provided, a wave having an ISB structure is produced. For a 1 KHz L-only signal, the upper sideband will be approximately 25 dB stronger than the lower sideband.

Fig. A-3 shows the spectrum calculation of *MOD PT. 1*, which monitors the output of the first inner loop transmitter. This calculation assumes the feedback path is open, the envelope modulation is 50%, and 0.5 radian of phase modulation is present. It is also assumed that the timing of the two components is perfect (i.e., no relative timing error). It is noteworthy that for L-only modulation, the upper sideband is present and the sidebands are reversed. As will be seen below, this situation is corrected at *MOD PT. 2*. Also, note that while the second order undesired sideband is only 29.7 dB below the modulated carrier (30.2 dB below the unmodulated carrier), theoretically the first order undesired sideband is fully suppressed. The fact that there can be complete suppression of the first order undesired sideband can be recognized by considering the following relationship between J0, J1, and J2, which are Bessel functions of the first kind.

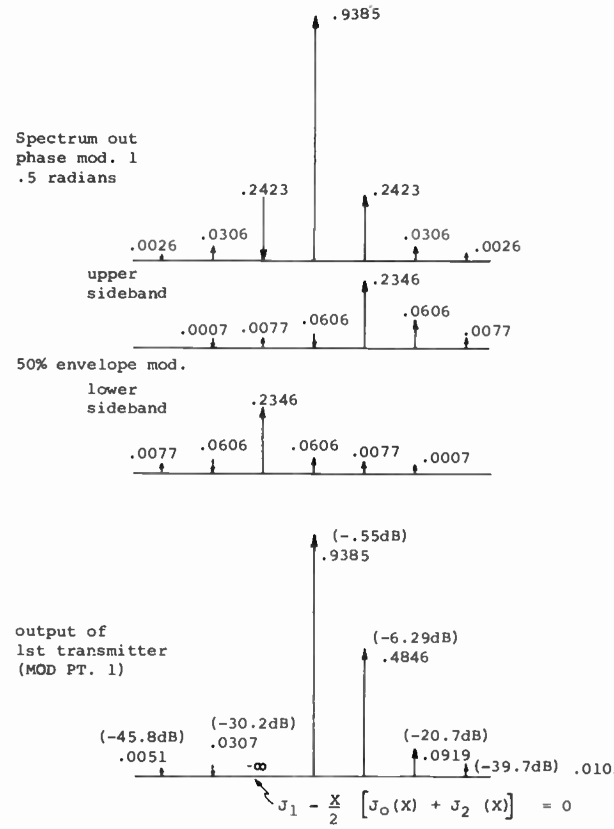


Figure A-3. Spectrum calculation of *MOD PT. 1*, which monitors the output of the first inner loop transmitter.

$$\frac{X}{2}\left[J_0(X) + J_2(X)\right] = J_1(X)$$

The above equation may be verified by examining the power series that define these Bessel functions. (It is noteworthy that this relationship can be used accurately to measure phase or frequency modulation where the value of X is less than one radian. If envelope modulation can be accurately measured to 100% and a spectrum analyzer is available, or some other means for measuring the first-order undesired sideband, it is only necessary to envelope modulate the phase or frequency-modulated wave and measure the amount of envelope modulation that causes the first order sideband to disappear.)

The output of the first transmitter is fed to the first simulated receiver. The simulated receiver incorporates a decoder and, accordingly, the wave from the amplitude modulator is inversely modulated. The output of the inverse modulator feeds a quadrature demodulator comprising a 90° phase shifter, which shifts the phase of the 1.4 MHz carrier wave, and a product demodulator, which multiplies the inverse modulated output with 90° phase shifted IF carrier. The audio output component of the product demodulator has approximately 12% of harmonic distortion. If this L − R component is combined with a distortion free L + R component, the total distortion would be about 6%. This distortion is reduced later in the system to substantially below 1%. The audio signal is then fed to a linear difference circuit which combines the audio out of the simulated receiver with a sample of the L − R wave which was fed to the first inner transmitter.

This L − R wave is 180° out of phase with the fundamental component of the audio wave at the output of the simulated receiver. The level of this undistorted L − R wave is made equal to twice the level of the fundamental component from the simulated receiver. This puts the distortion out of phase with the original wave from the simulated receiver and maintains the original distortion percentage. This operation also inverts the sidebands so that a left stereo signal is transmitted as a lower sideband.

The audio wave is then fed to a second inner transmitter comprising a phase modulator and an amplitude modulator. This second inner transmitter also operates at a frequency of 1.4 MHz.

The output of the second inner transmitter (available at *MOD. PT. 2*) has a much cleaner overall spectrum than the output of the first inner transmitter.

Fig. A-4 shows calculations for the spectrum characteristics at *MOD. PT. 2*, the output of the second inner transmitter. The output of the station's transmitter should match this spectrum, assuming the use of a low distortion transmitter with proper equalization. The spectrum output of the second phase modulator (PM_2) was determined by performing a Fourier analysis of the following equation:

$$PM_2 = \sin[w_c t + 0.5(-\cos w_A t - 0.023\cos 3w_A t)]$$

The spectral output components of PM_2 are shown in Fig. A-4 where the wave is amplitude modulated (lines 2 & 3) and the summation is shown (line 4) for each spectral component.

The actual amount of distortion can be determined by referring to Fig. A-5. The initial line is copied from the bottom line of Fig. A-4 and is the output spectrum of the second inner transmitter. This wave is inverse modulated by slightly over 26%. The resultant wave at the output of the inverse modulator is shown on the

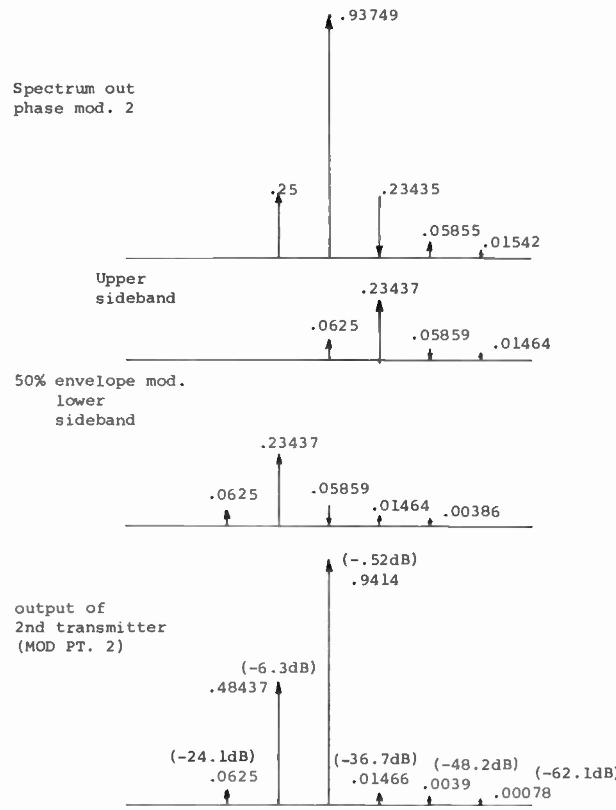


Figure A-4. Calculations for the specturm characteristics at *MOD PT. 2*, the output of the second inner transmitter.

bottom line. This wave is quadrature demodulated to derive the L − R signal. This calculation is shown at the bottom of Fig. A-5 indicating a total RMS second and third harmonic distortion of 2.01%. However, to determine the overall L- or R-only distortion, the figure can be reduced by a factor of 2 if we assume a perfect envelope detector is used to demodulate the envelope and obtain the L + R component. (Envelope detectors are available with distortion figures of less than 0.1%)

Accordingly, a total distortion figure of 1% can be achieved, without the negative feedback circuit in operation. With 10 dB to 12 dB of negative feedback in the L − R path, the total distortion is in the range of 0.25 to 0.3%. In practice 0.6% distortion figures have been obtained with inexpensive, though good quality, ISB AM stereo receivers.

Returning to the description of the overall block diagram of Fig. 2, the audio output of the second simulated receiver is used in a negative feedback circuit and is connected to the feedback switch. When the feedback switch is closed, the audio feedback wave is fed to the combining circuit which, in turn, feeds the first phase modulator. The result is that the proper L − R component is generated to produce a low distortion signal when demodulated by an appropriate AM stereo receiver.

The special audio that is fed to the second phase

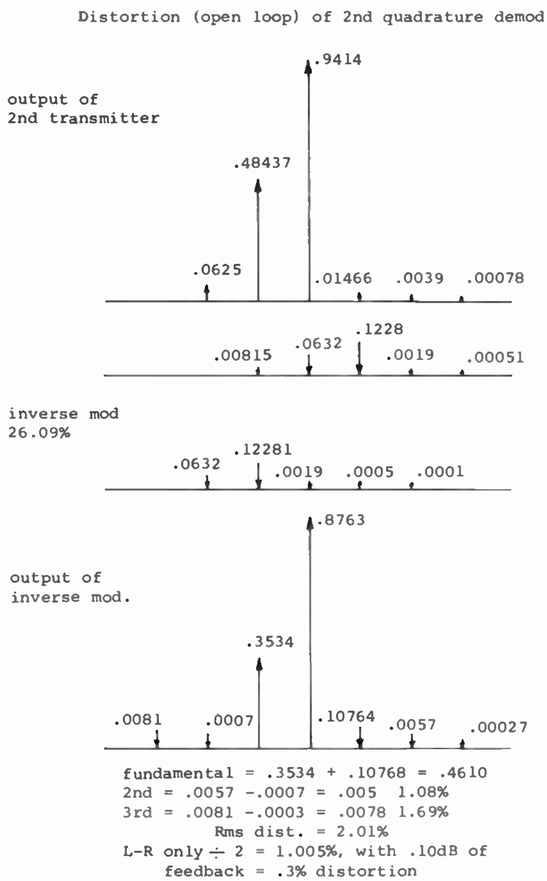


Figure A-5. Resulting distortion in the output of the inverse modulator.

modulator is also fed to the third phase modulator which operates at one-half the carrier frequency.

A divide-by-two circuit is inserted after the phase modulator of the PLL, causing the voltage-controlled oscillator to operate at the desired carrier frequency. The L − R audio wave is also passed through a day/night equalizer circuit. The L − R audio wave is passed through a variable time delay, and this time-delay circuit can be adjusted so that the overall time delay in the L − R path equals that in the L + R path including, of course, the delays in the associated transmitter. The audio is then passed through an adjustable, nonlinear time-delay circuit which, in turn, feeds the third modulator. The resulting phase-modulated wave is fed to a tuned RF amplifier, which produces three to four watts at the station's carrier frequency. It is noteworthy that no mixing circuits are used to produce the output phase-modulated wave. This significantly reduces the requirements of the filtering circuits normally provided to maintain low spurious heterodyne products when conventional frequency translation circuits are needed.

The output of the tune RF amplifier is used to take the place of the crystal oscillator in the station's AM transmitter. When this phase modulated wave is properly combined with the L + R modulations at the normal audio input of the transmitter, a specially contained, low distortion ISB wave is produced, similar to the wave observed at *MOD PT. 2*.

APPENDIX B

MOTOROLA C-QUAM® AM STEREO SYSTEM

Greg Buchwald
Motorola Inc., Schaumburg, Illinois

C-QUAM SYSTEM EQUATIONS

Any broadcast signal can be broken into three major components: amplitude, frequency, and phase. The equation for a monophonic transmission can be described by:

$$E_R = (1 + L + R)\cos(\omega_c t + \phi)$$

where 1 represents the carrier, L + R represents the monophonic (or Left and Right) information to be sent, $\omega_c t$ represents the carrier frequency, and ϕ represents phase modulation information which is, ideally, zero.

In fact as one uses ±100% modulation of L + R as an example and substitutes ±1 into the equation, it is obvious that at L + R = +1, the carrier level is instantaneously twice as high and at L + R = −1, the carrier level is instantaneously 0.0, or cutoff has occurred. From this equation it is also obvious why negative modulation is limited to 100% while positive modulation can exceed 100%. In the United States, the positive limit is 125% or 1.25 in the equation.

To insert stereophonic information, one could alter the amplitude, frequency, or phase of the transmitted signal. Altering the amplitude is to be avoided since the envelope would no longer represent 1 + L + R, but instead, a distorted component containing L + R. Substantial alteration of the frequency is also to be avoided. This leaves only the phase component available for modulation.

One method of adding a second channel of information to an existing carrier and utilizing the same spectral assignment is to make use of *quadrature modulation* (QUAM) techniques. Linear QUAM, not unlike that used to convey the chroma information in NTSC color transmission, can be used to convey the second channel of information. Advantages of QUAM are:

- No increase in occupied bandwidth
- Little signal-to-noise (SNR) degradation
- No loss of existing coverage
- The potential for synchronous reception techniques

However, the major disadvantage is that the envelope term is not 1 + L + R, but, instead is:

$$\sqrt{(1 + L + R)^2 + (L - R)^2}$$

leading to high levels of distortion in current monaural, envelope detector receivers and some difficulty in the conversion of existing broadcast transmitters due to additional requirements placed on the modulator stage.

The C-QUAM system was derived from QUAM, therefore it retains, to a large extent, the advantages of QUAM. In fact, C-QUAM can also use synchronous detection techniques particularly when conditions warrant. One difference between the generation of the C-QUAM signal versus the QUAM signal is found in the envelope audio term. By substituting the distorted term required by QUAM with the simple summation of the left and right channels, the envelope is made compatible with existing receivers. Mathematically, the system has been designed as follows.

The equation for the QUAM signal is:

$$E_R = \sqrt{(1 + L + R)^2 + (L - R)^2}\cos(\omega_c t + \phi)$$

where: $\phi = \tan^{-1}\left[\dfrac{L - R}{1 + L + R}\right]$

The desired envelope component is:

$$E = (1 + L + R)$$

Motorola engineers found that the cosine of the instantaneous phase modulation is:

$$\cos\phi = \frac{1 + L + R}{\sqrt{(1 + L + R)^2 + (L - R)^2}}$$

If QUAM is multiplied by cos ϕ:

$$\frac{1 + L + R}{\sqrt{(1 + L + R)^2 + (L - R)^2}} \times$$
$$\sqrt{(1+L+R)^2 + (L-R)^2}\cos(\omega_c t + \phi)$$

Then the resultant is:

$$E_R = (1 + L + R)\cos(\omega_c t = \phi)$$

where: $\phi = \tan^{-1}\left[\dfrac{L - R}{1 + L + R}\right]$

In the process, the envelope term is made compatible by sending $1 + L + R$ while the inphase (I) and quadrature (Q) components are multiplied by $\cos\phi$. Therefore the broadcast C-QUAM signal has the following characteristics:

Envelope (E) =	1 + L + R
Inphase (I) =	(1 + L + R)cosϕ
Quadrature (Q) =	(L − R)cosϕ

It can be seen that the mono, $1 + L + R$, signal may be directly derived from the envelope detector output, while a quadratine detector combined with division by $\cos\phi$ may be used to demodulate the $L - R$. The art of demodulation of C-QUAM signal will be discussed later.

THE C-QUAM ENCODING SYSTEM

There are several forms by which the C-QUAM encoder may be implemented. In the first form, used at the time of this writing by Motorola and Delta Electronics, a series of linear, balanced multipliers is employed to generate a quadrature modulated signal. The signal is then amplified to the point of limiting in order to remove the QUAM envelope term. The

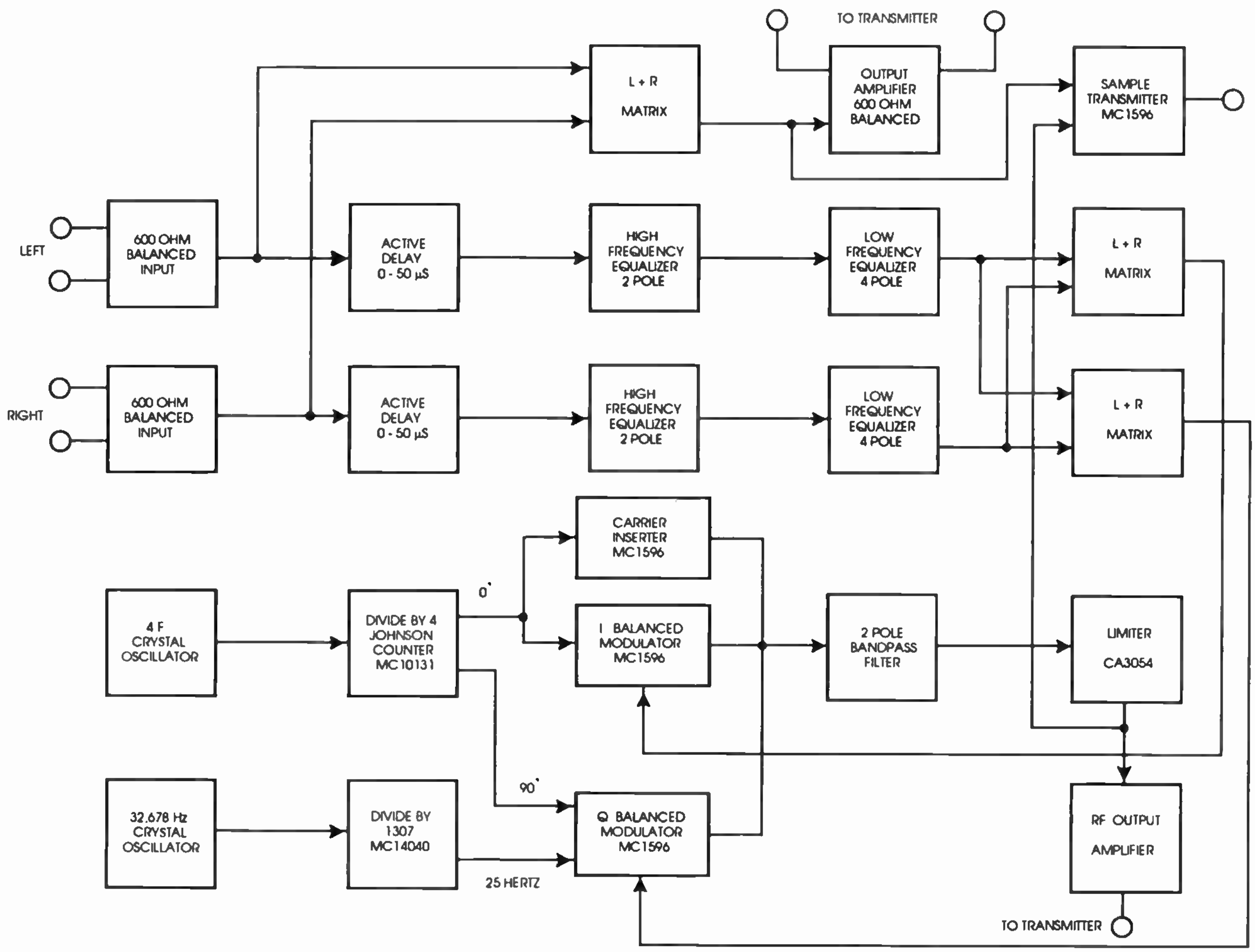


Figure B-1. Block diagram of C-QUAM exciter.

summation of the left and right channels provides a distortion-free monophonic audio term which modulates the transmitter. The resultant phase modulated signal is used to replace the crystal oscillator stage of the transmitter. The existing broadcast transmitter conveys the $L-R$ information in the form of complex phase modulation, and the envelope conveys the L + R information to both the existing monophonic as well as newer stereophonic receivers.

Other modulation techniques include the matrix switching method used by Harris Corporation and the Time Division Multiplex method used by Broadcast Electronics. The latter eliminates the need for audio matrices in the stereo modulator path.

To help visualize the generation of the signal, refer to Fig. B-1, the block diagram for the exciter. Since the linear, balanced modulator technique is the easiest to understand, it will be described herein. Interested readers are encouraged to study literature produced by other C-QUAM broadcast equipment manufacturers to obtain a full understanding of the various modulation techniques.

Beginning with the RF stages, a crystal at four times the carrier frequency ($4F_c$) is fed to a digital divider circuit which results in an output that is an on-frequency carrier signal at 0°, 90°, 180°, and 270°. These RF signals are utilized in pairs as differential references for three balanced modulators. The purpose of each balanced modulator is to supress the carrier feed-through and produce sideband information only when audio is present. It can be seen that the first two modulators are fed with the 0° and 180° reference signals. These modulators, therefore, produce an output referenced in-phase or synchronous to the carrier frequency. The third modulator is, conversely, fed with the 90° and 270° outputs from the RF divider and therefore forms a quadrature modulator stage. The first modulator is fed with a DC voltage which causes a precise offset from the null of the 0° carrier reference, thereby producing a precise DC carrier term at its output. The second modulator is fed with $L+R$ and therefore produces in-phase, $L+R$ modulation sidebands. The final modulator is fed with $L-R$ audio information and produces quadrature sidebands relating to the stereophonic information. It is important to note that the carrier-producing modulator is used in preference to simply unbalancing the in-phase ($L+R$) modulator for lower distortion.

The three balanced modulator outputs are summed and bandpass filtered to remove odd-ordered harmonics of the desired carrier frequency. Removal of odd-order harmonics prior to limiting is important since odd-ordered products represent a DC term which can unbalance the subsequent limiting stage, thereby introducing distortion. After bandpass filtering, the RF signal is amplitude limited by multiple stages which exhibit approximately 50 dB of limiting gain. The envelope component is therefore removed, leaving only the phase modulated carrier component. This constant amplitude RF signal is then further amplified to a level adjustable from zero to 30 volts, peak to peak. It is this amplified, constant amplitude carrier signal which is interfaced to the transmitter to create the composite stereo transmission.

Unlike FM stereo where a composite signal is fed into a wideband modulator input, the AM stereo signal is constructed at the power amplifier stage of the converted broadcast transmitter. A portion of the signal from the exciter enters the transmitter through an RF interface and follows one path to the RF power amplifier, while the remaining element, L + R, enters through the audio input terminals and follows a different path to the modulator where it is finally combined with the RF phase modulated information. Since two different circuit paths are used, the time delay along each path may be different, often by 40 μs or greater. This level of delay can be understood when one considers the phase shift through the low-pass filter in a conventional PDM transmitter which may exhibit a fifth or seventh order elliptical function. If the delay between the paths is not matched, both reduced separation and increased distortion will result. Loss of separation is easily understood since proper dematrixing of the audio signals in the receiver will occur only if the phase and amplitude of L + R and L − R are closely matched. The increase in distortion is a concept that can only be clearly understood when one examines the phase modulated component of the signal and finds that an L + R term is found in the PM component. If this L + R term does not match the envelope term, an increase in distortion would result. Therefore, the audio equalizer is a very important section of the exciter.

In the simplest form of equalization, a delay circuit in the audio path to the quadrature modulator section of the exciter would be utilized, thereby introducing time delay into the L + R and L − R audio fed into the QUAM generator so that it matches the L + R audio delay through the broadcast transmitter. Although adequate results will occur with this approach, a modification of this approach allows for additional correction of the signal to compensate for bandwidth or sideband symmetry problems commonly associated with broadcast antenna systems, particularly directional arrays.

A once-common form of single sideband generation particularly popular in the 1950s, utilized a scheme known as the "phasing method." The process was simple because one could simultaneously amplitude and phase modulate a signal, essentially generating a QUAM signal. If the processs is taken one step further so that both the RF terms and the audio terms fed to each modulator are shifted 90°, a single sideband transmission results without the use of expensive, sharp cutoff RF filters. The same holds true for AM stereo signal generation. If a small amount of phase shift is introduced into the audio, driving the phase modulated path of the transmitter with reference to the envelope audio path, the signal will take on an unsymmetrical sideband structure. If the antenna system exhibits an overall tilt towards the upper sidebands, for example, one would only need to predistort the signal so the the lower sidebands were favored in the transmitter, thereby restoring symmetry in the radiated signal.

The exciter can perform this task by inserting the delay sections into the left and right channels prior to matrixing. It is clear from Fig. B-1 that this is the case for most C-QUAM exciters. The audio is delayed independently in the left and right channels. It is then matrixed to produce L + R and L − R which feed the QUAM generator. An uncompensated summation of L + R is used to feed the external broadcast transmitter. Such equalization, known as differential equalization, can be quite powerful in correcting situations where asymmetrical antenna sideband radiation would otherwise limit stereo performance.

The audio equalizer is broken into several components. The first is the constant time delay circuitry which is used to compensate for the bulk of differential transit delays through the transmitter. The second section is the high-frequency equalizer which is used to compensate for additional phase shift (nonlinear group delay) characteristics in the broadcast transmitter audio path. These nonlinear delay characteristics are introduced by rolloff in the modulator stages, phase shift in PDM filters, and reflected antenna impedances. The final section is the low frequency equalizer which, not unlike tilt correction commonly found in audio processors and limiters, anticipates the phase shifting and amplitude rolloff action of modulation transformers, reactors, and coupling capacitors in older, plate modulated transmitters. By introducing a similar rolloff and phase shift into the phase modulated audio path, separation may be maintained at frequencies below 50 Hz.

The last section of the exciter to be discussed is the pilot generator. The pilot tone, a 25 Hz sine wave audio component, is injected into the L − R modulator at a 5% modulation level in the L − R channel. The purpose of the pilot tone is to indicate the presence of stereophonic information. This is different from the pilot tone signal in the FM stereo system where the 19 kHz signal is used as a synchronizing term in the demodulator. Indeed, AM stereo receivers could be built without the pilot tone detector, however, the consumer has grown accustomed to seeing a stereo indicator on the radio, hence it is incorporated into the system.

C-QUAM DECODING/RECEIVING TECHNIQUES

There are over a dozen ways to decode a C-QUAM transmission. The most common approach is the "feedback decoder" technique (Fig. B-2). In this drawing, there are three detectors: envelope, inphase, and quadrature.

The envelope detector demodulates the monophonic, L + R information. It may be a simple diode detector, however, most stereo demodulator integrated circuits utilize a limiter/multiplier approach which offers superior performance. Distortion mesaurements in the 0.1% to 0.3% region are commonly found at 99% negative modulation when this technique is used.

The inphase and quadrature demodulators are identical in action to the balanced modulators found in the exciter. Each has a reference signal, either 0° for the I detector, or 90° for the Q detector, and is phase locked to the incoming, received signal. However, referring back to the system equation, the actual received I signal is not $1 + L + R$, but instead, $(1 + L + R)\cos\phi$, and the received quadrature signal is $(L - R)\cos\phi$ rather than $L - R$ as desired. In a proper C-QUAM decoder, the incoming RF/IF signal must be divided by the term $\cos\phi$, generally prior to quadrature demodulation. In the feedback decoder, the divider stage accomplishes this task. The derivation of the $\cos\phi$ term conveyed in the phase modulated component is, however, an interesting process.

As indicated earlier, the I detector output would, since it is both phase and amplitutde sensitive, be $(1 + L + R)\cos\phi$ assuming no divider action. If the $\cos\phi$ component is eliminated from the I detector output, the I detector would produce a signal identical in theory to the envelope detector. Asssuming further that the I and Q detectors are fed from the same IF signal path, as $\cos\phi$ is removed from the I detector output, it is also removed from the Q detector output, leaving L − R as the result. Since it is known that the IF signal must be divided by $\cos\phi$, a divider stage is placed in the IF path feeding the I and Q detectors. If the output of the I detector is then analyzed against the envelope detector output in a high gain comparator, the resultant is $\cos\phi$. By connecting the output of the comparator to the control port of the divider, a feedback loop results which will cause the $\cos\phi$ signal normally found at the output of the I detector to be effectively transferred to the input of the divider. Since the IF path is therefore divided by $\cos\phi$, the Q detector performs the task of directly demodulating L − R.

The remaining circuitry in the decoder detects the pilot tone and dematrixes the $L + R$ and $L - R$ audio terms into the original left and right components.

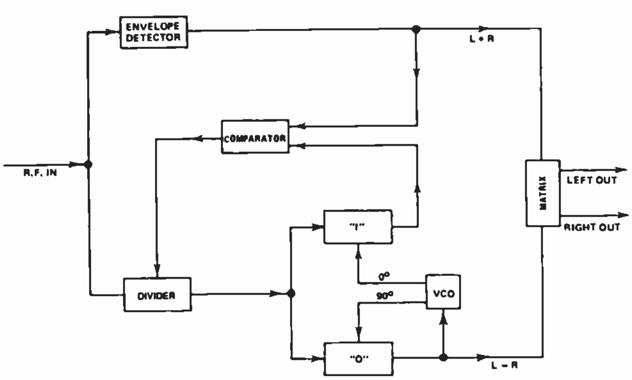


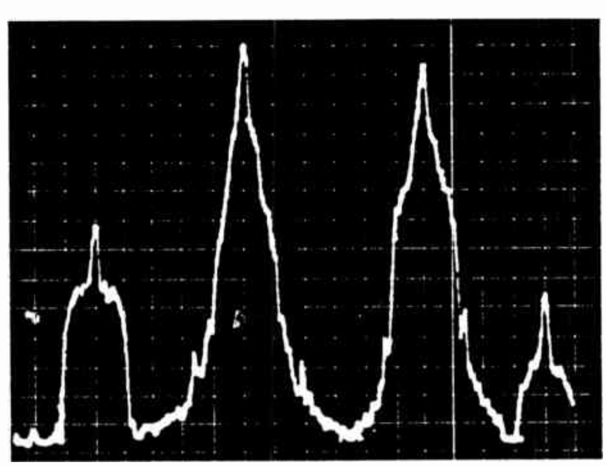
Figure B-2. Block diagram of feedback decorder.

PERFORMANCE CONSIDERATIONS

The C-QUAM system is capable of on-air performance figures in excess of 40 dB separation and under 1% distortion from 100 to 5000 Hz, particularly on newer solid-state broadcast transmitters. It is not uncommon to obtain separation figures in excess of 30 dB from 50 Hz to the 10 kHz NRSC limit.

RF spectra is also well controlled within the NRSC limits of occupied bandwidth now adopted by the FCC. A spectral photograph of four C-QUAM stations, three of which operate at a power level of 50 kW, depicts the spectral signal of C-QUAM broadcasts. It is virtually indistinguishable from monophonic broadcasts, see Fig. B-3.

The broadcaster is free to choose the type of audio limiting desired for use with the C-QUAM signal with the exception that, if matrix type limiters are utilized, they must contain single channel limiting circuitry, or the single channel limiter found in all C-QUAM exciters must be enabled. Since the L + R monophonic signal is summed and directly coupled to the broadcast



Frequency at 10 kHz / Division
Amplitude Response at 5 dB / Division

Figure B-3. Spectral photograph of four C-QUAM broadcasts.

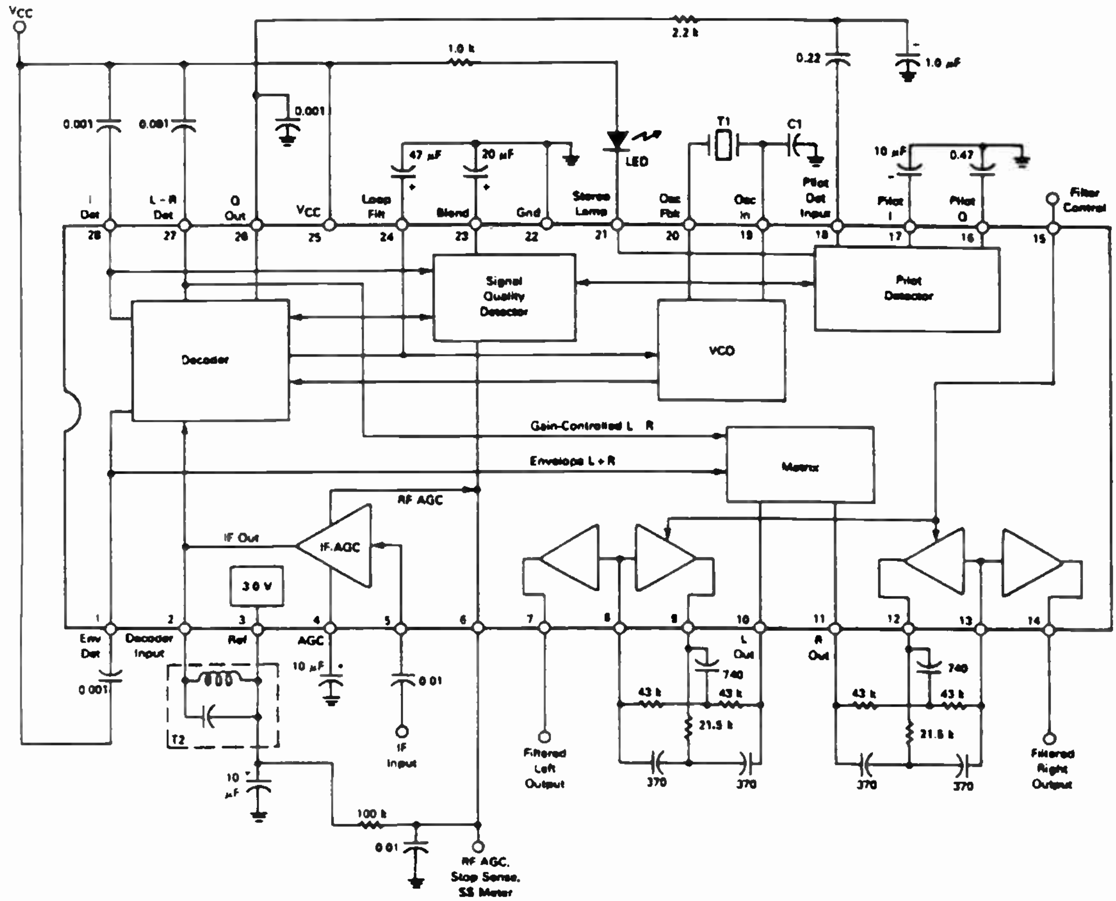


Figure B-4(A). Block diagram of the MC13022.

transmitter, no additional tilt or overshoot is placed upon the audio allowing full modulation of the envelope component. The modulation constraints of the system are as follows:

L + R (monaural modulation)	+ 125, − 100% (FCC limits—higher mod is possible)
L − R (stereo difference channel only)	± 100% (±45° phase modulation)
L, R only (single channel, ref. to L + R)	+ 125, − 75% (limit of 71.56° phase modulation)
PILOT TONE (injected into L − R channel)	25 Hz ±0.1 Hz, 5% + 1/ − 0% injection level, sine wave.

In addition, audio preemphasis may be used. Motorola recommends use of the NRSC modified 75 μS standard, however, tests with up to 30 dB preemphasis yield operation within FCC occupied bandwidth limits. The C-QUAM system has been used at audio bandwidths up to 15 kHz with excellent results, however, Motorola must emphasize the current FCC rules limiting audio bandwidth and subsequent RF emissions to 10 kHz as dictated by the current NRSC guidelines and adopted by the FCC.

Use of the C-QUAM system will not reduce the monophonic coverage of the broadcast station. Since the envelope is sent as a compatible signal (i.e., the amplitude modulation of the station continues as 1 + L + R), the stereo component should pass undetected in existing mono receivers and full modulation of the envelope may be contemplated. The stereo receiver will experience a slight S/N degradation due to the fact that a second channel of information has been opened to allow passing of the stereo information. The degradation is usually under 3 dB and is typically 1.5 dB, which is nearly imperceptible under normal stereo programming conditions. All C-QUAM receivers contain circuitry that restores receiver to mono reception when conditions warrant, thereby assuring no loss of coverage to the listener.

Receiver technology continues to expand with the advent of newer generation C-QUAM decoders. Original decoder ICs demodulated the incoming signal but performed very little conditioning of the signal to extend the performance of the receiver. Later generation decoders contain circuits which allow new receivers to be designed with adaptive bandwidth and synchronous detection features which improve reception under weak signal conditions. The MC13022 decoder IC, see Fig. B-4 A and B, contains such circuitry which allows receiver designers to produce radios that

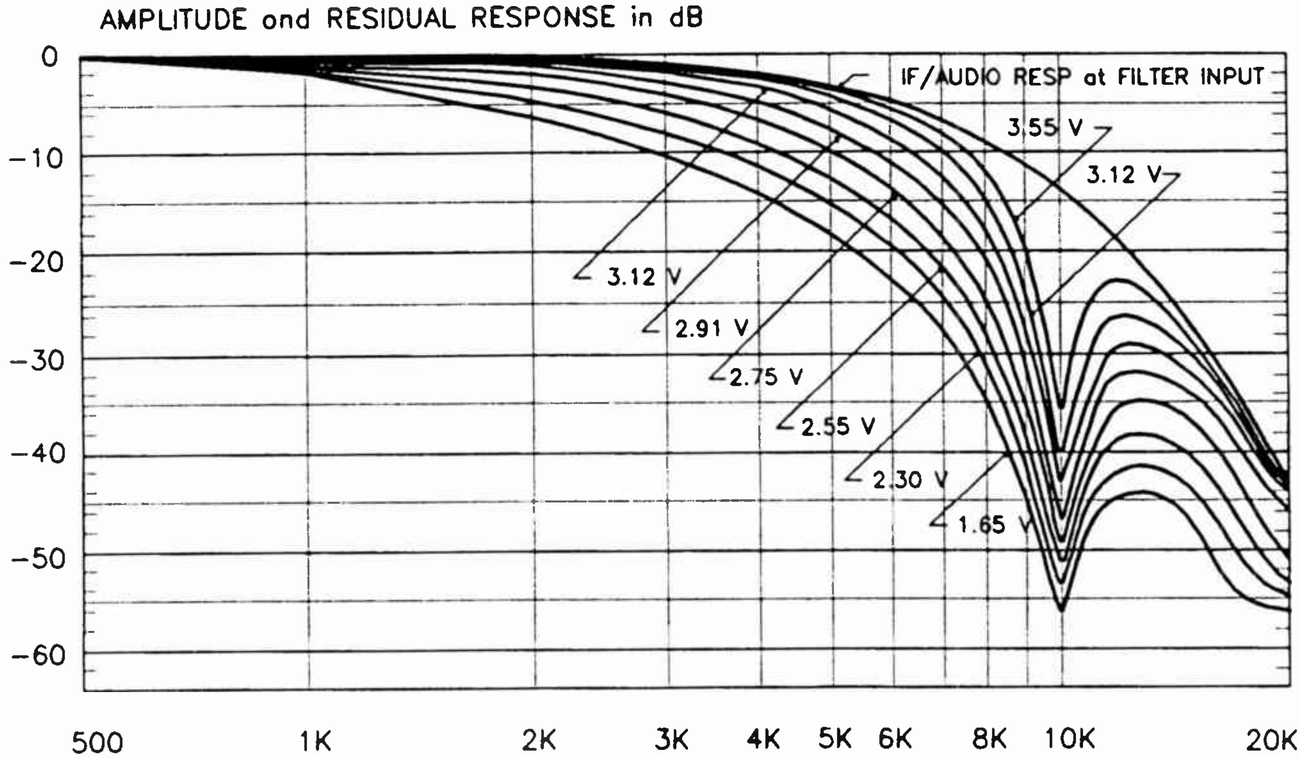


Figure B-4(B) MC13022 Variable Q notch filter response curves.

operate at NRSC bandwidths under stronger signal conditions and automatically variable bandwidths sliding to three to four kHz under weak incoming signal conditions. The merits of such a system are obvious. The listener can enjoy wideband AM reception when signal conditions warrant while the receiver will automatically reduce the bandwidth under poor conditions, thereby retaining the listener. In addition, NRSC de-emphasis is easily implemented with the MC13022. Future generation ICs are expected to further reduce the cost of adding AM stereo to existing designs, perhaps to a point of zero additional cost.

CONCLUSION

The C-QUAM AM stereo system has been designed to fulfill the demands of the AM broadcast service. It is easy to install and maintain, it occupies essentially no additional spectra, it is capable of excellent performance from 50 Hz up to the NRSC limit of 10 kHz, and performs well with matrix audio limiters which insure monophonic coverage and loudness. The C-QUAM system was designed with the receiver manufacturer in mind as well. Low cost implementations are possible with the ever-growing family of ICs available and performance rivaling FM stereo is available, particularly in the automotive and portable receiver markets. C-QUAM has repeatedly been selected as the standard for stereophonic broadcasting due to the fact that it addresses both broadcast and receiver industry related issues and exhibits superior performance.

3.3 FM Broadcast Transmitters

Geoffrey N. Mendenhall, P.E.
Mukunda B. Shrestha
Edward J. Anthony
Broadcast Electronics Inc., Quincy, Illinois

With contributions by
Warren B. Bruene, P.E., Dallas, Texas

INTRODUCTION

History of FM Broadcasting

Although the mathematical principles explaining frequency modulation (FM) had been known for many years, the advantages and practical application to radio broadcasting were not realized until the 1930's, when Major Edwin H. Armstrong did extensive developmental work proving that FM radio transmissions were possible. Many theoreticians claimed to have proof that Armstrong's experiments were impossible, based on mathematical models, claiming that an infinite transmission bandwidth would be required. He never received proper credit for his many contributions to the radio communications industry during his lifetime. (See Ref. 1 at the end of this chapter for more information about the career of Major Armstrong.)

Among the advantages of FM are freedom from static, wide audio bandwidth, and the ability of an FM receiver to capture the stronger of two signals transmitted on the same carrier frequency.

In 1940, following extensive public hearings, the Federal Communications Commission established the FM Broadcast Service and set aside 40 channels in the 42 to 50 MHz band with commercial operation scheduled to begin January 1, 1941. Although World War II stopped all nonmilitary radio construction, more than 40 FM stations continued to serve over 400,000 receivers. To eliminate the interference problems resulting from skywave reflection in the prewar FM band, the Commission moved the FM Broadcast Service to the 88 to 108 MHz band in 1945, thereby increasing the number of available channels to 100. However, the expected growth of FM broadcasting did not materialize. Since conversion of prewar FM receivers to the new band was not practical, purchase of a new receiver was the only way to receive the new FM stations. Television appeared to offer much more to the consumer than FM radio, since most FM stations merely duplicated the programming of an affiliated AM station. Despite the potential for a higher quality broadcast service, there was little public demand for new FM receivers and virtually no public reaction when FM stations dropped popular programs from their schedule or were off the air due to equipment failures. It is not surprising, therefore, that in May 1950, there had been only 16 new FM license applications during the previous 15 months in which 259 FM stations ceased operations. It was not until after the introduction of stereo multiplex FM broadcasting that public awareness of the FM band increased and then skyrocketed in the 1970's to make FM the dominant medium for musical programming.

Characteristics Of FM Compared To AM

Reduced Noise

The 88 MHz to 108 MHz FM broadcast band is relatively free of atmospheric and other noise interference. Emission at these frequencies is not propagated great distances by the ionosphere as it is in the 550 kHz to 1600 kHz AM broadcast band. Therefore, noise from lightning discharges is limited to line-of-sight distances and is almost negligible. Manmade noise is a far greater source of noise, particularly in urban areas. The level of manmade noise falls off at increasing higher frequencies so that the microvolts-per-meter noise level is about one-tenth as great in the FM band as it is in the AM band.

In addition, FM has an improved noise threshold characteristic when compared to AM. Limiting circuitry symmetrically clips the RF waveform in an FM receiver to remove any amplitude variations produced by static or impulse noise before they reach the demodulator which responds only to phase or frequency changes in the signal. The FM improvement factor is illustrated in Fig. 1. Note the sharp knee in

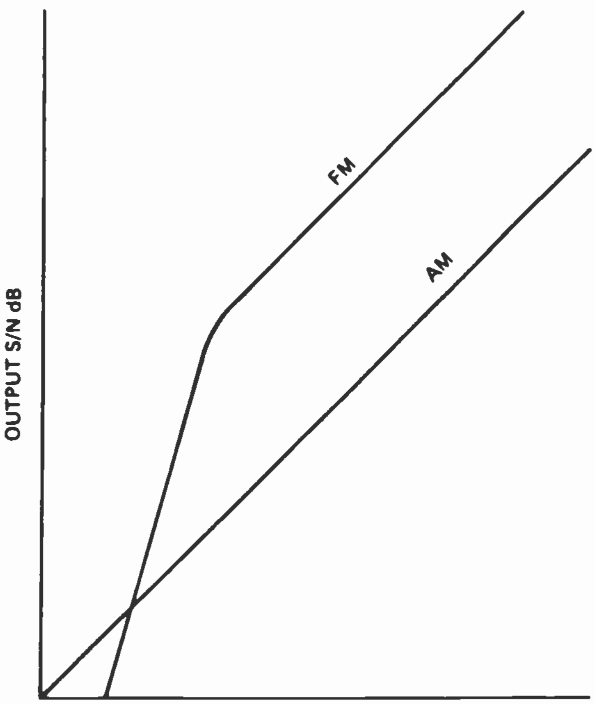


Figure 1. FM improvement factor.

the threshold curve above which noise and interference are suppressed, resulting in an improved signal to noise (S/N) ratio. This same capture effect causes a weaker FM signal on the same channel to be suppressed, resulting in greatly reduced co-channel interference.

For even greater noise reduction, preemphasis (75 micro-second time constant) is employed in the transmitter, whereby the audio frequency components above about 2.1 kHz are boosted in amplitude at the rate of 6 dB per octave before being applied to the modulator. Flat frequency response is restored in the receiver's deemphasis network by attenuating the higher frequencies the same amount they were boosted in the transmitter. At the same time, the high frequency noise components, which are characteristic to FM, are also attenuated, resulting in greatly reduced background noise. Preemphasis is discussed in more detail in the FM theory and FM exciter sections of this chapter.

The net result of the above factors is a much better signal-to-noise ratio on FM than on AM, and the ability to transmit much wider bandwidth information on FM. The lower background noise, along with wider frequency response, means that a signal of higher quality and wider dynamic range can be enjoyed.

Occupied Bandwidth

The advantages of the FM broadcast medium over the AM broadcast medium do not come free of compromise. A standard FM broadcast channel occupies more than ten times the bandwidth of an AM broadcast channel. This is because the more complicated nonlinear sideband structure of a frequency modulated carrier with wide deviation requires much more bandwidth than the simpler linear distribution of sidebands in an AM system. More details on the amplitude and spacing of FM sidebands are given in the section on FM modulation theory in this chapter. Fortunately, the FM broadcast band is located in the VHF portion of the frequency spectrum where more bandwidth is available than in the medium wave AM broadcast band.

High Fidelity

Uniform frequency response over the audible range of at least 50 Hz to 15 kHz, very low amplitude distortion (harmonic and intermodulation), very low noise level, and good transient response (uniform time delay versus frequency) are necessary for Hi-Fi performance. The FM channel authorizations provide for adequate audio frequency response and a low-noise radio link to the listener. The rest of the performance is a matter of equipment design.

Stereophonic Transmission

The wide channel allocations and ability of FM to multiplex compatibly several audio channels on one carrier permitted development of a practical stereo broadcasting system. This provides a means for the broadcast industry to provide the public with reproduction quality as good as or better than that which is available on stereo records or tapes. The advent of the *compact disc* and other digital audio source equipment will continue to challenge equipment manufacturers and station engineers further to improve the performance of the entire FM broadcast chain. Detailed technical information about stereophonic transmission theory and standards is provided in the chapter about FM Stereo and SCA.

Subsidiary Communications Authorization

The wide channel bandwidth authorized for FM broadcasting also makes it feasible to multiplex two subsidiary (SCA) audio or data channels together with the stereo transmission. SCA channels provide an important source of revenue to many stations as well as provide many useful audio and digital services to the community. Detailed technical information about SCA transmission theory and standards is provided in the chapter about FM Stereo and SCA.

FCC Transmission Standards

The Federal Communications Commission regulates and enforces the technical standards that apply to radio broadcasting in the United States. In theory, this will assure that the public is provided with a consistently high standard of transmission quality from station to station. The Rules and Regulations covering transmitter technical performance are set forth in Part 73 of the FCC Rules and Regulations available from the

U.S. Government Printing Office in Washington, D.C. The rules and regulations are changed from time to time to keep pace with new technology and changes within the broadcast industry. Every broadcast engineer should have access to a current copy of these Rules and Regulations so that the station's technical performance is maintained within the prescribed limits.

FREQUENCY MODULATION THEORY

Angular Modulation

Frequency modulation and phase modulation (PM) are both special cases of angular modulation. In any angular modulation system both the frequency and phase of the carrier vary with time as a function of the modulating signal.

The relationship between the frequency deviation of the carrier, the phase deviation of the carrier, and the sinusoidal modulating frequency is defined as the modulation index (m) where

$$m = \frac{\text{frequency deviation (peak-to-peak Hertz)}}{\text{modulating frequency (Hertz)}} \quad (1)$$

Since frequency modulation and phase modulation are both subsets of angular modulation, they are virtually indistinguishable from one another except in the modulator characteristics.

In a PM system, the modulating signal causes the phase of the carrier wave to vary according to the instantaneous amplitude of the modulating signal. A phase modulator generates a constant amount of phase deviation of the carrier with a constant amplitude modulating signal independent of the frequency of the modulating signal. The frequency deviation of the carrier produced by a phase modulator does increase as the modulating frequency is increased, even though the level of the modulating voltage is held constant. The net effect is that the phase modulator behaves as if it were a frequency modulator with a 6 dB/octave rising slope on the modulating signal input.

A frequency modulator generates a constant frequency deviation of the carrier, with a constant amplitude modulating signal independent of the frequency of the modulating signal. The phase deviation of the carrier produced by a frequency modulator decreases as the modulating frequency is increased, even though the level of the modulating voltage is held constant. The net effect is that the frequency modulator behaves as if it were a phase modulator with a 6 dB/octave falling slope on the modulating signal input.

In FM broadcasting, the signal should have frequency deviation that is proportional to the amplitude of the modulating signal, but independent of the frequency of the modulating signal as is generated by a frequency modulator.

The instantaneous frequency (rate of change of phase) of the RF output wave differs from the carrier frequency by an amount proportional to the instantaneous value of the modulating waveform. For example, consider a 100 MHz carrier wave frequency modulated by a 1000 Hz audio tone, and assume that a 1 volt input to the modulator causes ±20 kHz of frequency deviation on the positive and negative peaks of this tone. If the audio input amplitude is increased to 2 volts, the peak deviation will become ±40 kHz varying in sine-wave fashion from one peak deviation to the other and back again at the 1000 Hz rate. In FM broadcasting, the FCC has established that a peak frequency deviation of ±75 kHz shall constitute 100% modulation.

When preemphasis is used ahead of the frequency modulator, the system becomes a phase modulator at audio frequencies above the turnover point of the preemphasis network. This is because the frequency response of the preemphasis network rises at the rate of 6 dB/octave above this point. FM broadcasting with preemphasis really becomes a mixture of FM at low modulating frequencies and PM at high modulating frequencies.

FM Sideband Structure

The frequency modulated RF output spectrum contains many sideband frequency components, theoretically an infinite number. They consist of pairs of sideband components spaced from the carrier frequency by multiples of the modulating frequency. When the modulation index is small (m = 0.5) the amplitude of the second and higher order sidebands is small so that the output consists mainly of the carrier and the pair of first-order sidebands, as illustrated in Fig. 2A. The total transmitter RF output power remains constant with modulation, but the distribution of that power into the sidebands varies with the modulation index so that power at the carrier frequency is reduced by the amount of power added to the sidebands.

As the modulation index is increased, as in wide deviation FM broadcasting, the higher order sidebands become more prominent. The amplitude and phase of the carrier (J_0) as well as the sidebands (J_1 thru J_n) can be expressed mathematically by making the modulation index (m) the argument of a simplified Bessel function.

E(t) = total RF output voltage (2)

$A[J_0(m)\sin \omega c(t)]$ carrier amplitude

$+[J_1(m)\sin(\omega c+\omega m)t]$ first-order upper sideband

$-[J_1(m)\sin(\omega c-\omega m)t]$ first-order lower sideband

$+[J_2(m)\sin(\omega c+2\omega m)t]$ second-order upper sideband

$+[J_2(m)\sin(\omega c-2\omega m)t]$ second-order lower sideband

$+[J_3(m)\sin(\omega c+3\omega m)t]$ third-order upper sideband

$-[J_3(m)\sin(\omega c-3\omega m)t]$ third-order lower sideband

$\pm[J_n(m)\sin(\omega c\pm n\omega m)t]$ . . .higher order sidebands

where: A $=$ The unmodulated carrier amplitude constant
J_0 $=$ The modulated carrier amplitude
$J_1, J_2, J_3 \ldots J_n$ are the amplitudes of the nth order sidebands
m $=$ The modulation index
ωc $= 2\pi F_c$ (The carrier frequency)
ωm $= 2\pi F_m$ (The modulating frequency)

The numeric values of the Bessel functions (J_0 thru J_n), which express the amplitudes of the various frequency components, can be found in mathematical tables. Fig. 3 shows a graphical representation of how the Bessel function values for the carrier and the first eight pairs of sidebands vary with the modulation index.

In a monophonic FM broadcast transmitter, the modulation index can become very high at low modulating frequencies. With a 50 Hz audio input signal of sufficient amplitude to produce 75 kHz deviation (100% modulation), the modulation index is:

$$m = \frac{75000}{50} = 1500 \qquad (3)$$

With a 15000 Hz input at the same deviation (also 100% modulation), the modulation index is only:

$$m = \frac{75000}{15000} = 5 \qquad (4)$$

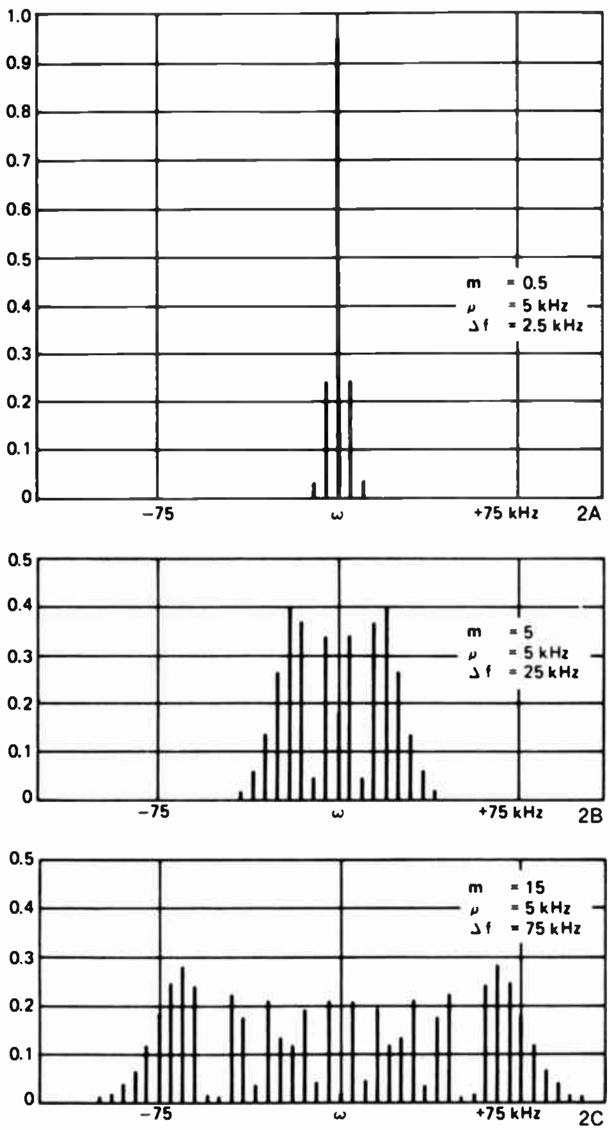


Figure 2. RF spectrum with modulation indexes of 0.5, 5.0, and 15.

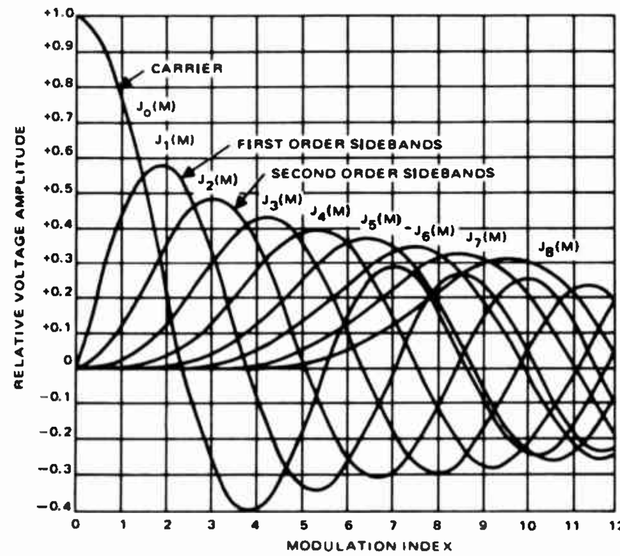
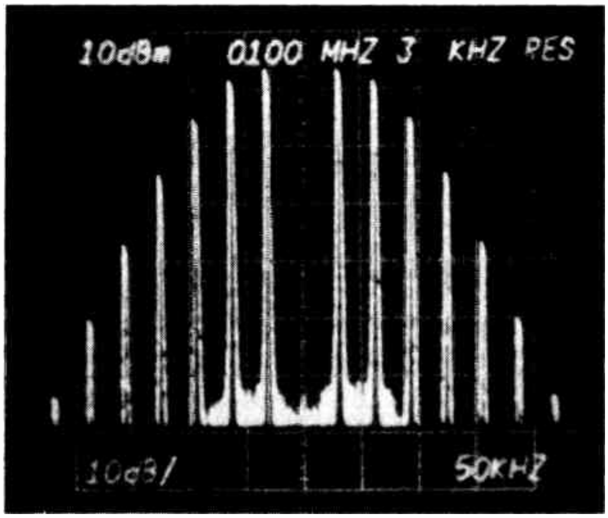


Figure 3. Relationship of carrier and sideband amplitudes to modulation index.

Figs. 2B and 2C illustrate the frequency components present for modulation indices of 5 and 15. Note that the number of significant sideband components becomes very large with a high modulation index. The total bandwidth occupied extends beyond ±75 kHz from the carrier depending upon the modulating frequency. This single tone modulating frequency analysis is useful in understanding the general nature of FM and for making tests and measurements. When program modulation is applied, there are many more sideband components present and they are varying so much that sideband energy becomes distributed over the entire occupied bandwidth rather than appearing at discrete frequencies.

Bessel Nulls

At certain modulation indices, the carrier amplitude goes to zero with all the transmitted power distributed at frequencies other than the carrier frequency. This carrier null phenomenon is useful as an extremely accurate method for measuring the frequency deviation and to check the calibration of modulation monitors. Referring again to Fig. 3, note that the carrier amplitude goes to zero and reverses sign at several values of modulation index including: 2.405, 5.520, and 8.654.



FOR M = 2.405, F_M = 31,185Hz, F_C = 100.00MHz

Figure 4. FM RF spectrum showing occupied bandwidth and Bessel Null single tone @ ± 75 kHz deviation.

Fig. 4 is a photograph taken from an RF spectrum analyzer showing the first Bessel null (m = 2.405) of a carrier at a frequency of 100 MHz.

To determine the audio input level required to achieve 75 kHz deviation, apply an audio tone of exactly 8667 Hz (75,000 divided by 8.654). Increase the audio level until the carrier disappears for the third time. At this audio level the deviation is exactly 75 kHz. The carrier amplitude (null) detector must have sufficient selectivity to separate the carrier from the sidebands and could be a spectrum analyzer or a receiver with a narrow IF bandwidth. The FM signal can be heterodyned down to a convenient frequency for measurement. Heterodyning does not alter the modulation index. When a frequency (or phase) modulated wave is multiplied or divided, this also multiplies or divides the frequency deviation and the modulation index by the same amount.

A listing of useful carrier and first order sideband nulls as function of the modulation index (m) and the modulating frequency (F_m) is given below:

NULL	MODULATION INDEX (m)		(F_m) FOR 75 kHz DEVIATION	
	CARRIER	1st SIDE-BANDS	CARRIER	1st SIDE-BANDS
1st	2.405	3.832	31,187	19,574
2nd	5.520	7.016	13,587*	10,690
3rd	8.654	10.174	8,667	7,372
4th	11.792	13.324	6,361	5,629
5th	14.943	16.471	5,023	4,554
6th	18.071	19.616	4,150	3,823
7th	21.212	22.759	3,536	3,295

* This tone is recorded on track 38 of the NAB Broadcast Audio System Test CD.

Occupied Bandwidth

After examining the Bessel function and the resulting spectra, it becomes clear that the occupied bandwidth of an FM signal is far greater than the amount of deviation from the carrier that one might incorrectly assume as the bandwidth. In fact, the occupied bandwidth is infinite if all the sidebands are taken into account, so it is now clear that a frequency modulation system would require the transmission of an infinite number of sidebands for perfect demodulation of information. In practice, a signal of acceptable quality can be transmitted in the limited bandwidth assigned to an FM channel.

Effects of Bandwidth Limitation

Practical considerations in the transmitter RF circuitry make it necessary to restrict the RF bandwidth to less than infinity. As a result, the higher order sidebands will be altered in amplitude and group delay (time). Bandwidth limitation will cause distortion in any FM system.

Consider the block diagram shown in Fig. 5A, where a perfect FM modulator is connected to a perfect demodulator via an RF path of infinite bandwidth. The demodulated audio shown in Fig. 5B contains no distortion components.

In Fig. 6A, a passive LC bandpass filter is inserted between the modulator and demodulator in order to restrict the bandwidth. Audio distortion products now appear at the output of our perfect demodulator as shown in Fig. 6B. These distortion products are due solely to the bandwidth restriction (300 kHz BW_3) imposed by the passive bandpass filter.

Figs. 7A and 7B show the effects of a narrowband RF bandpass filter on the RF spectrum of a composite signal consisting of a stereophonic subcarrier modulated only on the left channel with 4.5 kHz, plus a 67 kHz unmodulated SCA subcarrier. The only distortion evident on the RF spectrogram is the loss of some sidebands greater than 150 kHz from the center frequency, and some amplitude differences between the upper and lower sideband pairs. Fig. 7B shows the corresponding effects observed on the demodulated baseband spectrum for the same signal. Note the creation of many undesired intermodulation terms which cause crosstalk into both the stereophonic and SCA subcarrier bands. The change in the RF spectrum is subtle, but the resulting spectrum after demodulation is clearly modified.

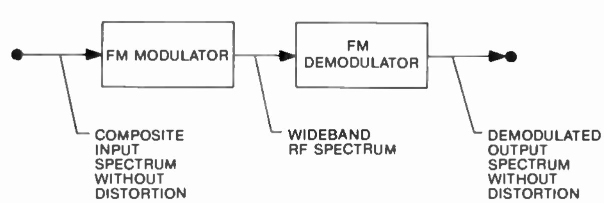
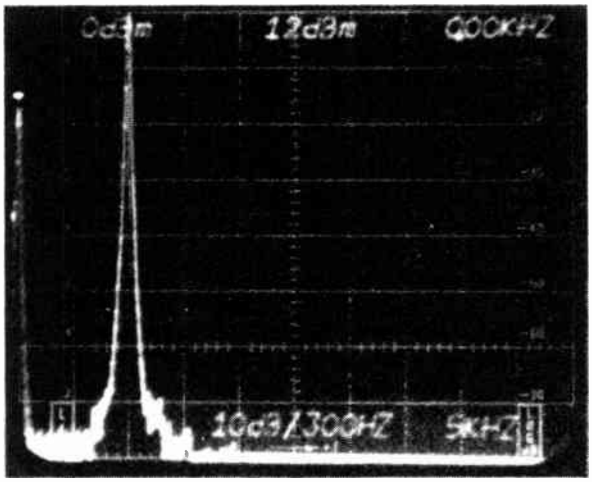
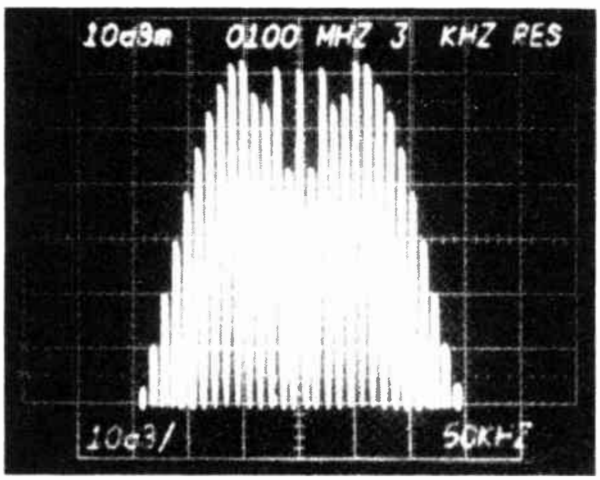


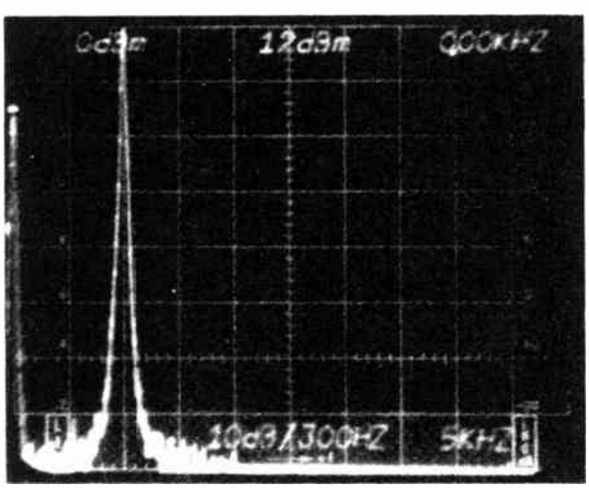
Figure 5A. Wideband RF path.



BASEBAND SPECTRUM TO FM MODULATOR



RF SPECTRUM TO DEMODULATOR



DEMODULATED BASEBAND SPECTRUM

Figure 5B. Single tone (10 kHz) modulation through wideband RF path.

As one can see, the distortion in any practical FM system will depend on the amount of bandwidth available versus the modulation index being transmitted.

Group Delay Response Symmetry versus Amplitude Response Symmetry

While both the amplitude response and group delay (time) response have an effect on the amount of distortion added to the FM signal, the symmetry of the group delay response is more important than the total group delay variation or the amplitude response symmetry. Best FM modulation performance is always obtained when the system is tuned for symmetrical group delay (time) response. Depending on the circuit topology, the tuning conditions for symmetrical group delay response may not coincide with the symmetrical amplitude response.

Limiting Factors within an FM Transmitter

Relating the specific quantitative effect of the bandwidth limitations imposed by a particular transmitter to the actual distortion of the demodulated composite baseband is a complicated problem indeed. Some of the factors involved are:

1. Total number of tuned circuits involved.
2. Amplitude and group delay response of the total combination of tuned circuits in the RF path.
3. Amount of drive (saturation effects) to each class "C" stage.
4. Nonlinear transfer function within each amplifier stage.

Improvement of the RF Path

The following design techniques can help improve the transmitter's bandwidth:

1. Maximize bandwidth by using a broadband exciter and a broadband IPA stage.
2. Use a single-tube design or a broadband, completely solid-state design where feasible.
3. Optimize both grid circuit and plate circuit of the tuned stage for the best possible amplitude response and symmetrical group delay response.
4. Minimize the number of interactive tuned networks.
5. Use a broadband antenna system that provides a low standing wave ratio on the transmission line.

For more detailed information about FM modulation theory, see Refs. 2, 3, and 4 at the end of this chapter.

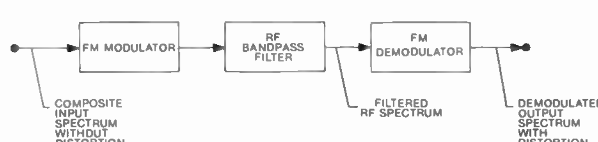
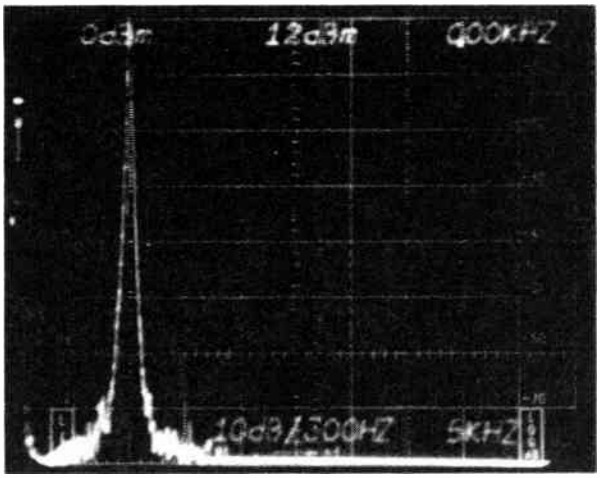
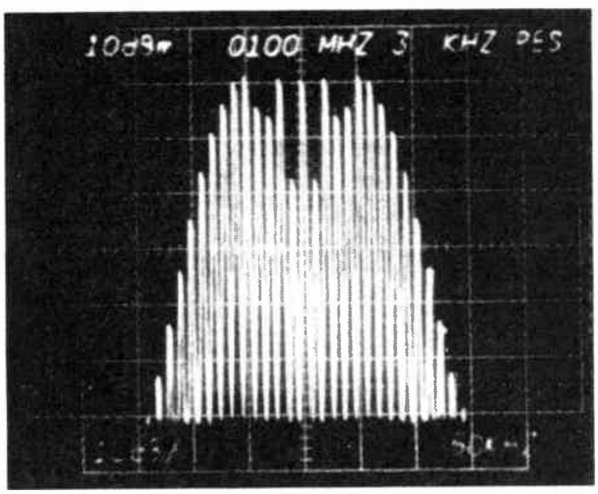


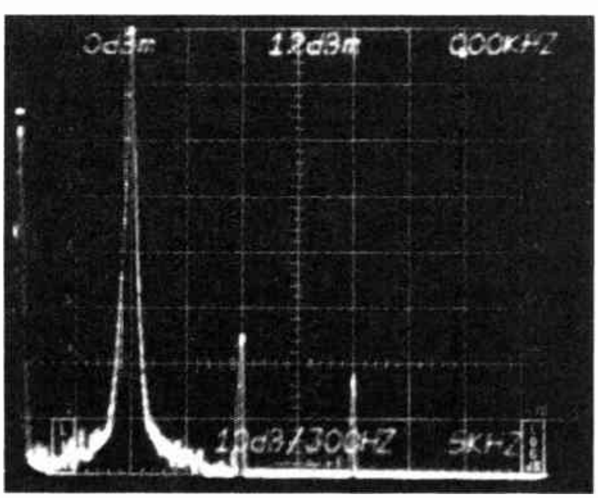
Figure 6A. Bandwidth limited RF path.



BASEBAND SPECTRUM TO FM MODULATOR

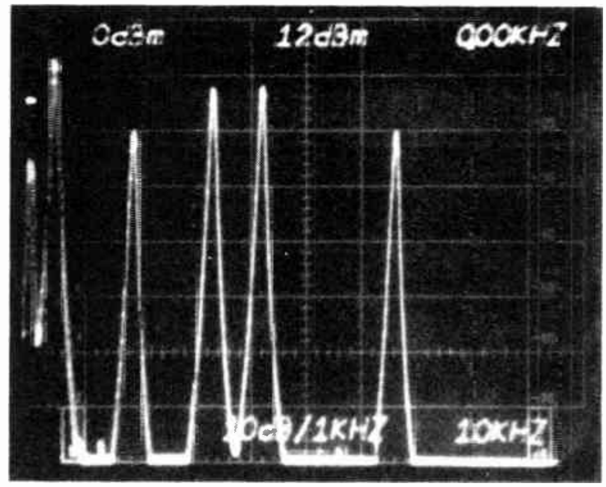


BANDWIDTH LIMITED RF SPECTRUM
TO DEMODULATOR

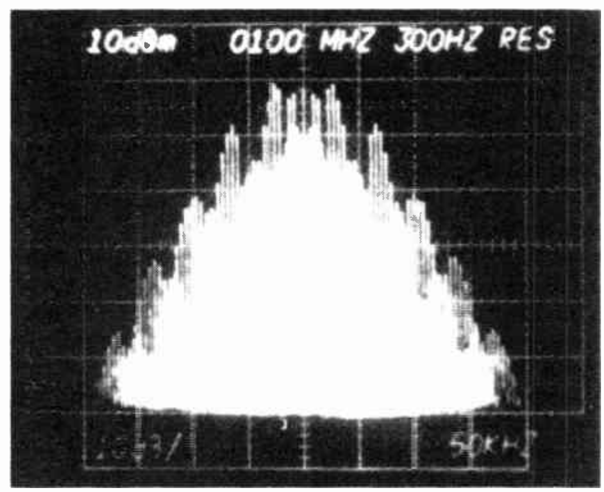


DEMODULATED BASEBAND SPECTRUM

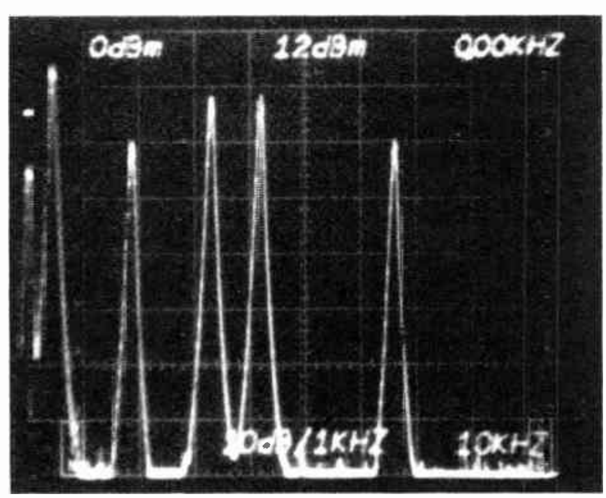
Figure 6B. Single tone (10 kHz) modulation through narrowband RF path.



BASEBAND SPECTRUM TO FM MODULATOR

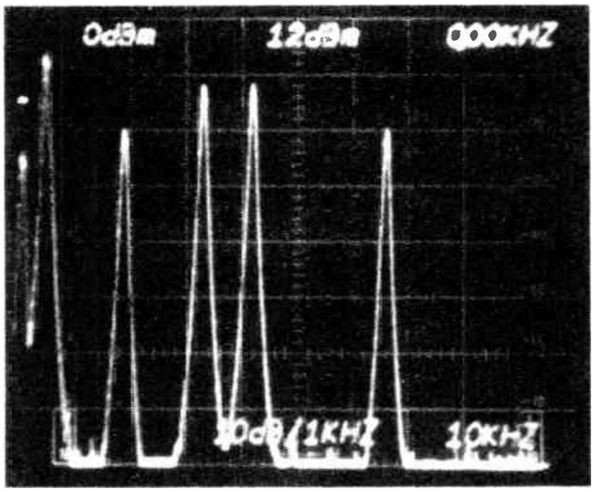


RF SPECTRUM TO DEMODULATOR

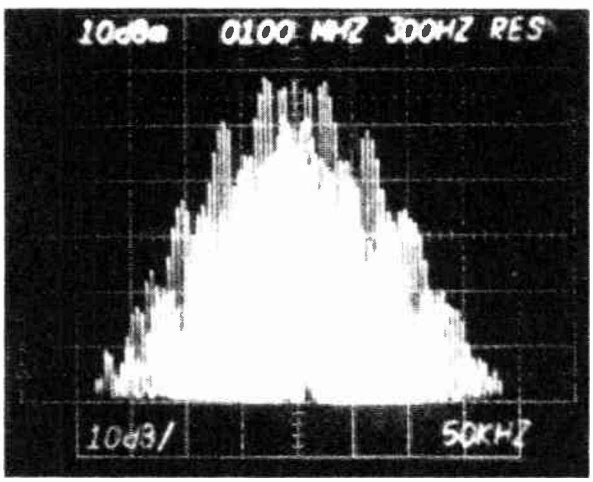


DEMODULATED BASEBAND SPECTRUM

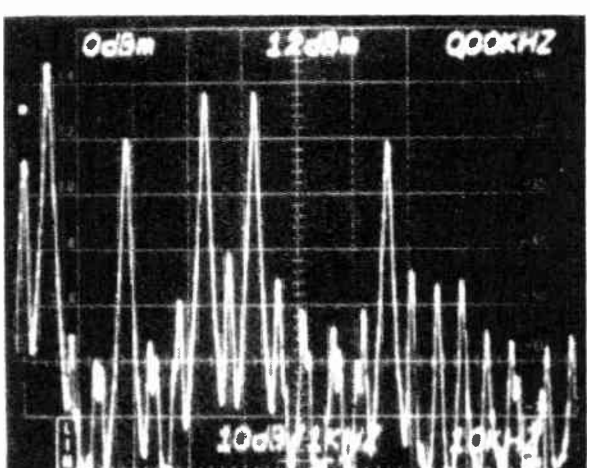
Figure 7A. Stereo (L or R = 4.5 kHz) plus SCA (unmod.) modulation through wideband RF path.



BASEBAND SPECTRUM TO FM MODULATOR



BANDWIDTH LIMITED RF SPECTRUM
TO SPECTRUM



DEMODULATED BASEBAND SPECTRUM

Figure 7B. Stereo (L or R = 4.5 kHz) plus SCA (unmod.) modulation through narrowband RF path.

PREEMPHASIS

The standards adopted for FM broadcasting require the use of preemphasis. The standard preemphasis curve is defined as an ideal RC network with a time constant of 75 microseconds. The 3 dB point is at a frequency of:

$$f = \frac{1}{2\pi RC} = \frac{1}{2\pi 75 \times 10^{-6}} = 2122\ \text{Hz} \qquad (5)$$

The 75 microsecond curve and the tolerance allowed by the FCC is shown in Fig. 8.

The reduction in receiver output noise due to the use of preemphasis in monophonic transmission is illustrated in Fig. 9. The noise voltage in a narrow bandwidth (for example, 1 Hz) increases directly with frequency, therefore, the power spectral density increases as the square of frequency as shown. When preemphasis is used, the noise voltage is attenuated above 2.1 kHz so that it remains constant with frequency. The power spectral density is also constant above 2.1 kHz. The area between these curves represents the noise power that is removed by the use of preemphasis. This diagram indicates the importance of preemphasis for high-fidelity transmission because the high-frequency noise at the receiver would be very much greater without deemphasis.

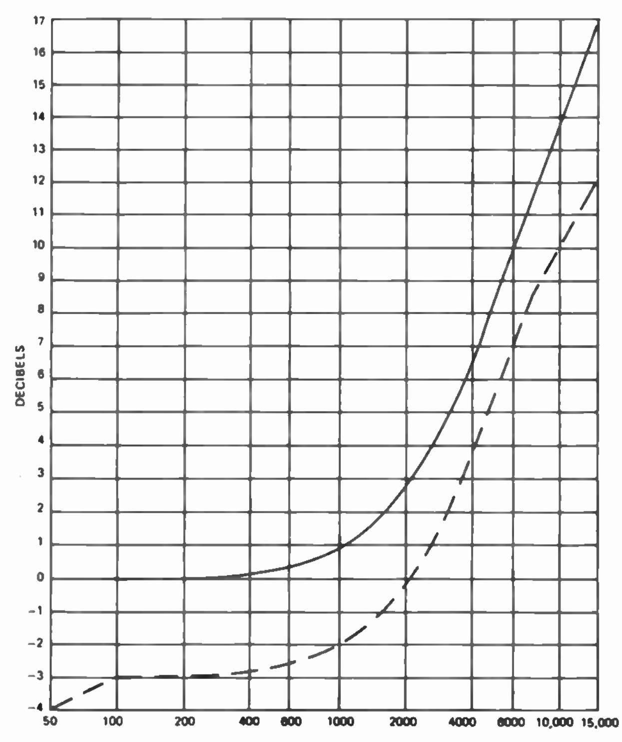


Figure 8A. FCC Standard 75 microsecond preemphasis curve (solid line) and tolerance limits (solid and dashed lines).

FREQUENCY IN HERTZ	DECIBELS
400 HZ	0.16 dB
1,000	0.87
2,000	2.76
3,000	4.77
4,000	6.58
5,000	8.16
6,000	9.54
7,000	10.75
8,000	11.82
9,000	12.79
10,000	13.66
11,000	14.45
12,000	15.18
13,000	15.86
14,000	16.49
15,000	17.07

Figure 8B. 75 microsecond preemphasis response.

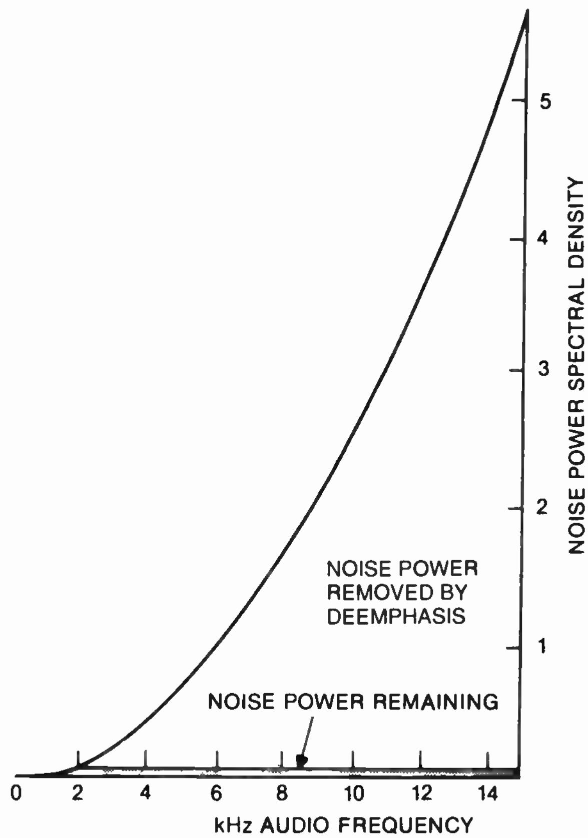


Figure 9. Noise power spectral density before and after deemphasis in receiver.

Preemphasis is practical because program energy tends to peak at several kilohertz and then falls off fairly rapidly at the higher frequencies. For this reason, the higher frequencies can be boosted in amplitude without causing much increase in modulation level. There is some increase, however, so the net improvement due to preemphasis is the ratio of the areas under the two curves of the diagram less this reduction in audio input level required to keep within the 100% modulation limit. Modern audio processing equipment takes the preemphasis curve into account when controlling peak modulation levels.

In the section on FM modulation theory, it was mentioned that the use of preemphasis ahead of an FM modulator actually causes the system to behave like FM at low modulating frequencies and like PM at high modulating frequencies.

The location of the preemphasis network in the system depends on whether the station is operating in the monaural or stereo mode. In the case of monaural transmission, the preemphasis network is usually located in the FM exciter just ahead of the modulator stage. Stereo transmission requires that the FM modulator have a flat response to the composite baseband signal from the stereo generator, so the individual preemphasis networks for the left and right channels are located in the stereo generator before the left and right audio channels are multiplexed into the composite baseband signal. The preemphasis time constant may be reduced to 25 microseconds if Dolby type "B" noise reduction is being used by the station. This will assure reasonable compatibility at low modulation levels, with the standard 75 microsecond receiver deemphasis.

FM TRANSMITTER POWER OUTPUT REQUIREMENTS

The transmitter power output (TPO) required is determined by the class of station, antenna height above average terrain, transmission line losses, and the antenna array gain. The required effective radiated power (ERP) is first determined, then the transmitter power output (TPO) can be calculated. Depending on the particular situation, the TPO could vary from as little as 50 watts to as much as 70 kilowatts.

THE FM TRANSMITTER

The purpose of the FM transmitter is to convert one or more audio frequency (composite baseband) input signals into a frequency modulated, radio frequency signal at the desired power output level to feed into the radiating antenna system. In its simplest form, it can be considered to be an FM modulator and an RF power amplifier packaged into one unit.

Actually the FM transmitter consists of a series of individual subsystems each having a specific function:

1. The FM exciter converts the audio baseband into frequency modulated RF and determines the key qualities of the signal.
2. The intermediate power amplifier (IPA) is required in some transmitters to boost the RF power level up to a level sufficient to drive the final stage.
3. The final power amplifier further increases the signal level to the final value required to drive the antenna system.
4. The power supplies convert the input power from the AC line into the various DC or AC voltages and currents needed by each of these subsystems.
5. The transmitter control system monitors, protects, and provides commands to each of these subsystems so that they work together to provide the desired result.
6. The RF lowpass filter removes undesired harmonic frequencies from the transmitter's output, leaving only the fundamental output frequency.
7. The directional coupler provides an indication of the power being delivered to and reflected from the antenna system.

Fig. 10 shows a simplified block diagram of a typical FM transmitter.

FM EXCITERS

The heart of an FM broadcast transmitter is its exciter. The function of the exciter is to generate and modulate the carrier wave with one or more inputs (mono, stereo, SCA) in accordance with the FCC standards. The FM modulated carrier is then amplified by a wideband amplifier to the level required by the transmitter's following stage.

Stereo transmission places the most stringent performance requirements upon the exciter. Since the exciter is the origin of the transmitter's signal, it determines most of the signal's technical characteristics; including signal-to-noise ratio, distortion, amplitude response, phase response, and frequency stability. Waveform linearity, amplitude bandwidth, and phase linearity must be maintained within acceptable limits throughout the baseband chain from the stereo and subcarrier generators to the FM exciter's modulated oscillator. From here, the FM carrier is usually amplified in a series of class "C" nonlinear power amplifiers, where any amplitude variation is removed. The amplitude and phase responses of all the RF networks which follow the exciter must also be controlled to minimize degradation of the baseband.

Before the advent of stereo broadcasting, most of the FM exciters employed phase modulation techniques. Some of these were adapted to stereo but it was difficult to achieve and maintain the performance requirements for stereo transmission.

In the FM modulation theory section, the important relationship between PM and FM was discussed. If the audio frequency response is made to fall off at the rate of 6 dB/octave across the entire audio band at the input of a phase modulator, the resulting modulated output will be identical to that of a frequency modulated carrier. In 1948, James R. Day, an associate of Edwin

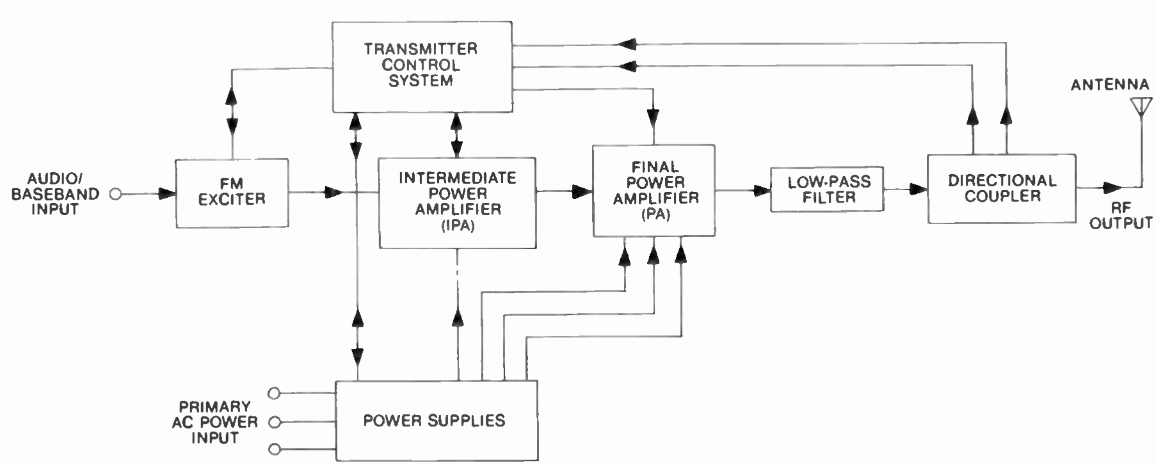


Figure 10. Simplified block diagram of an FM broadcast transmitter.

H. Armstrong, demonstrated a "Serrasoid" phase modulator which used pulse circuits to achieve large phase shifts with low distortion. The "Serrasoid" modulation technique, in conjunction with the audio response shaping mentioned above, became the standard method of generating wide deviation FM with low distortion.

This principle called "indirect FM" was used in many FM broadcast transmitters prior to the advent of stereo broadcasting. The advantage was that the carrier frequency could be generated by a stable crystal oscillator with the phase modulation occurring in later stages. The amount of phase deviation with low distortion was limited in most systems, so it was necessary to start with a low frequency oscillator and multiply its frequency and modulation index many times to achieve 75 kHz deviation at low modulating frequencies. This technique has been abandoned in favor of direct FM systems.

Direct FM

Direct FM is a modulation technique where the frequency of an oscillator can be made to change in proportion to an applied voltage. Such an oscillator, called a voltage tuned oscillator (VTO), was made possible by the development of varactor tuning diodes which change capacitance as their reverse bias voltage is varied (also known as a voltage controlled oscillator or VCO).

If the composite baseband signal is applied to the tuning terminal of a VTO, the result is a *direct* FM modulated oscillator. Fig. 11 is a block diagram that fits most of the modern direct FM exciters on the market.

The signal-to-noise ratio of an FM exciter is dependent on the short-term stability of the modulated oscillator. That stability is determined by factors such as operating level, noise figure of the oscillator transistor, circuit configuration, method of amplitude limiting, loaded "Q" of the oscillator tank circuit, and the mechanical stability of components. Optimization of these factors has resulted in a signal-to-noise ratio of better than 90 dB below 100% modulation in the current generation of FM exciters.

FM Modulator Linearity

Nonlinearities in the FM oscillator can, by altering the waveform of the baseband signal, create distortion in the demodulated output at the receiver. A secondary effect of this distortion may include stereo crosstalk into the secondary communications authorization (SCA).

The composite baseband signal is translated to a frequency modulated carrier frequency by the modulated oscillator. Frequency modulation is produced by applying the composite baseband signal to a voltage tunable RF oscillator. The modulated oscillator usually

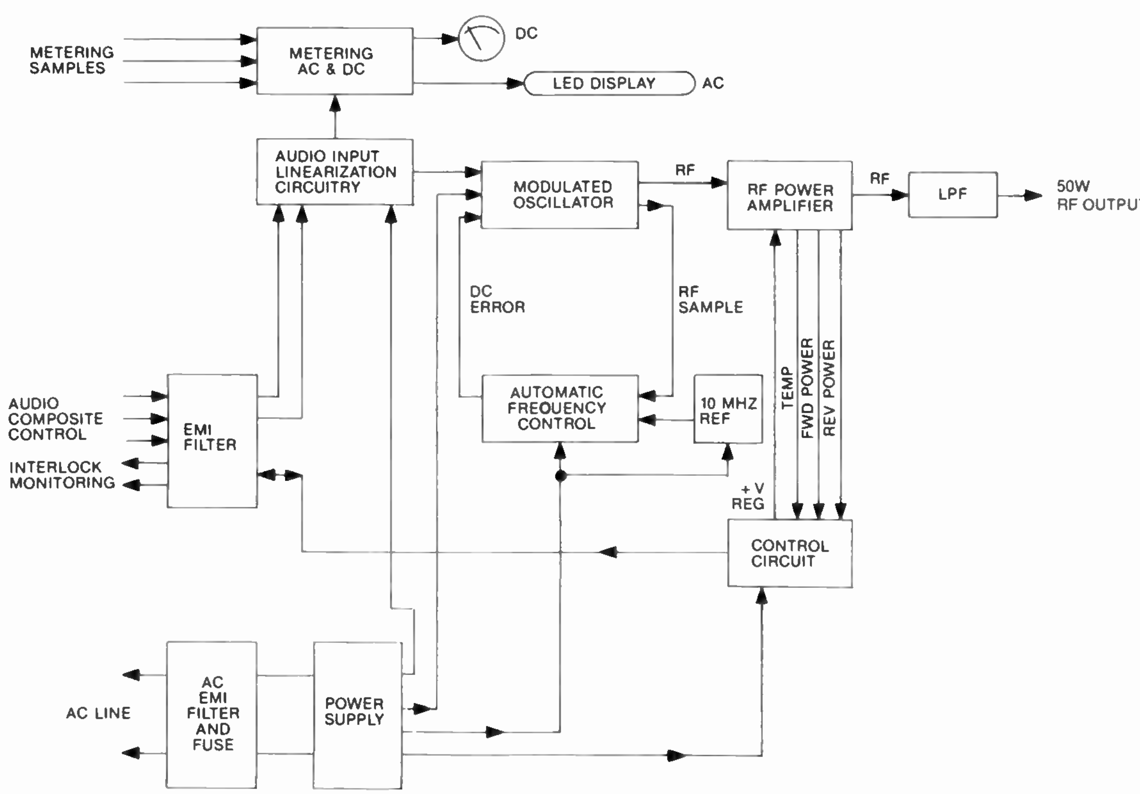


Figure 11. FM exciter block diagram.

operates at the carrier frequency and is voltage tuned by varactor diodes, operating in a parallel LC circuit.

To have perfect modulation linearity, the RF output frequency (F_c) must change in direct proportion to the composite modulating voltage (V_m) applied to the varactor diodes (C_v). This requirement implies that the capacitance of the varactor diodes must change as nearly the square of the modulating voltage as shown in following relationships:

(F_c) is proportional to (V_m) (Desired linear voltage to frequency translation) if:

$$F_c = \frac{1}{2\pi\,[(L)(C_t)]^{1/2}}$$

and: $$C_v = \frac{K}{(V^2_m)} \quad (\text{if } C_{fixed} = 0) \qquad (6)$$

where: F_c = Instantaneous carrier frequency
L = Inductance of resonant circuit
C_t = Total capacitance across L (C_{fixed} + $C_{varactors}$)
C_v = Capacitance of varactor tuning diodes
K = Varactor constant
V_m = Baseband modulating voltage

Unfortunately, the voltage versus capacitance characteristic of practical varactor diodes is not the desired square law relationship. All varactor-tuned oscillators have an inherently nonlinear modulating characteristic. This nonlinearity is very predictable and repeatable for a given circuit configuration, making correction by complementary predistortion of the modulating signal feasible. Suitable predistortion can be applied to the composite baseband signal by using a piece-wise linear approximation to produce the desired complementary transfer function. Fig. 12 shows a typical network of switching diodes and resistors used for complementary predistortion of the composite baseband. Fig. 13 shows how the predistortion network is cascaded with a nonlinear voltage-tuned oscillator to produce a linearized FM modulator.

It is also possible to improve both the linearity and signal-to-noise ratio of the modulated oscillator by demodulating its RF output to baseband and then

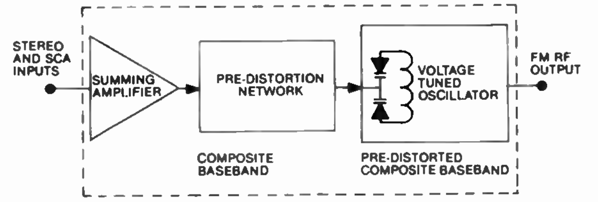


Figure 13. Linearized FM modulator block diagram.

feeding some of this baseband with the proper phase relationship back into the composite input of the modulator. This type of configuration places the entire modulated oscillator within a negative feedback loop and transfers the responsibility for maintaining linearity to the demodulator. Digital demodulation schemes can be made very linear, but the additional complexity and the potential problems with loop stability have limited the applications of this approach to linearization.

Modulator linearization has reduced harmonic and intermodulation distortion to less than 0.01% in the current generation of equipment. Any distortion of the baseband signal caused by the modulated oscillator will have secondary effects on stereo and SCA crosstalk, which are quite noticeable at the receiver in spite of the rather small amounts of distortion to the baseband. For example, if the harmonic distortion to the baseband is increased from 0.05% to 1.0%, as much as 26 dB additional crosstalk into the SCA can be expected.

For illustrative purposes, Figs. 14A, 14B, and 14C give representations of the fundamental and second order terms in the composite baseband spectrum with increasing amounts of harmonic distortion in the modulated oscillator. Fig. 14B shows this spectrum after 0.05% harmonic distortion has been added to each component. Note that the second order stereo (L − R) sidebands are 78 dB below 100% modulation, or about 58 dB below a 67 kHz SCA with a 10% injection. With normal energy distribution in L − R and the SCA, crosstalk from stereo into the SCA will be more than 60 dB below the SCA subcarrier. Fig. 14C shows the same baseband spectrum with 1.0% harmonic distortion. The second order stereo sidebands are only 32 dB below the SCA. Crosstalk may now increase as

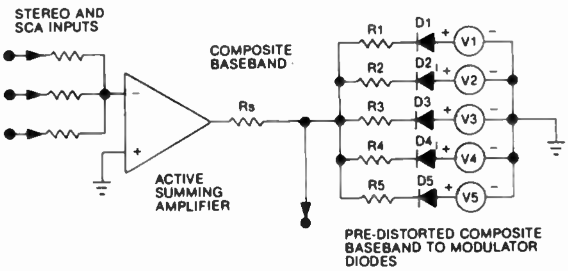


Figure 12. Predistortion network.

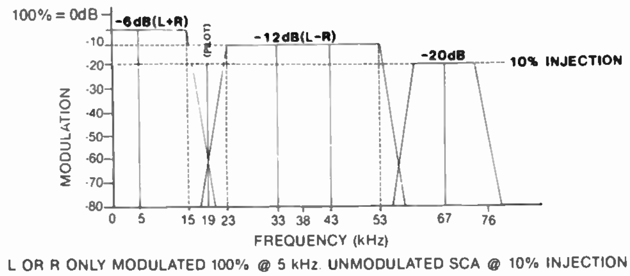
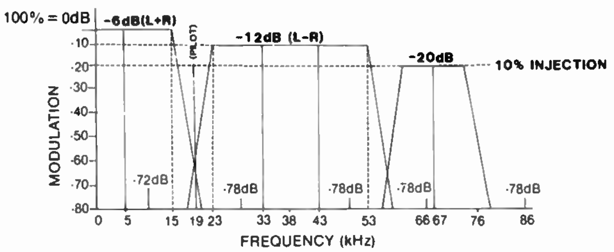


Figure 14A. Ideal demodulated composite baseband spectrum with no modulator distortion.



L OR R ONLY MODULATED 100% @ 5 kHz. UNMODULATED SCA @ 10% INJECTION. INTERFERING SECOND HARMONIC STEREO SIDEBANDS ARE 58 dB BELOW SCA. ONLY FUNDAMENTAL AND SECOND HARMONIC TERMS ARE SHOWN.

Figure 14B. Demodulated composite baseband spectrum with 0.05% harmonic distortion in modulator.

much as 26 dB, depending on the respective energy distributions in (L − R) and the SCA.

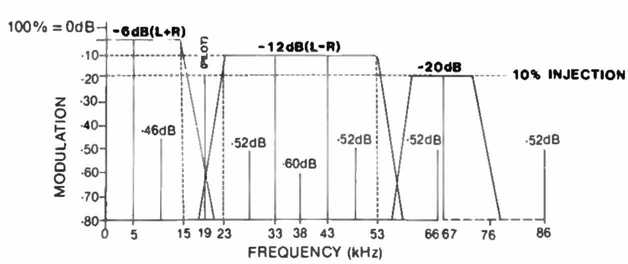
Transient intermodulation (TIM) distortion is usually not a factor in varactor-tuned modulated oscillators. The modulation bandwidth capability is generally more than ten times the composite bandwidth, and no negative feedback is used to maintain linearity.

Assuring that the composite baseband signal undergoes minimal distortion in the modulation process will suppress undesired harmonic and intermodulation products in the baseband, making the FM exciter transparent to the signals coupled into it. All exciter stages after the modulated oscillator operate as broadband amplifiers with minimal bandwidth limitations. The best FM exciter technology presently available is capable of transmitting compact disk quality with less than 0.01% distortion and a signal-to-noise ratio of better than 90 dB.

Automatic Frequency Control

The frequency stability of direct FM oscillators is not good enough to meet the FCC frequency tolerance of ±2000 Hz. This requires an automatic frequency control system (AFC) that uses a stable crystal oscillator as the reference frequency.

The modulated oscillator need not have good long-term stability since the AFC feedback loop will correct for long-term drift to keep the average carrier frequency within limits. The modulated oscillator does need excellent short-term stability (less than 1 second) because the control loop time constant must be long enough so that the AFC circuit does not try to remove desired low frequency audio modulation. This means that the oscillator is essentially running open-loop at frequencies above a few hertz so that the noise performance of the modulator will also be determined by the short term stability characteristics of the oscillator.



L OR R ONLY MODULATED 100% @ 5 kHz. UNMODULATED SCA @ 10% INJECTION. INTERFERING SECOND HARMONIC STEREO SIDEBANDS ARE 32 dB BELOW SCA. ONLY FUNDAMENTAL AND SECOND HARMONIC TERMS ARE SHOWN.

Figure 14C. Demodulated composite baseband spectrum with 1.0% harmonic distortion in modulator.

Phase-Locked Loop Automatic Frequency Control

Phase-locked loop (PLL) technology has provided a means of precisely controlling the carrier's average frequency while permitting wide deviation of the carrier frequency at baseband modulating frequencies. This implies that a PLL system behaves like an audio high-pass filter with higher modulating frequencies being ignored by the control loop while lower frequencies are considered to be errors in the average frequency and are tracked out by the loop. An added advantage of the PLL is the ability to synthesize the desired frequency from a single reference oscillator, thereby eliminating the need to change crystals when changing the frequency of the exciter.

The block diagram shown in Fig. 15 includes the key elements in the PLL. The output of the modulated oscillator operating at the carrier frequency is digitally divided down to a frequency of a few kilohertz or even less, called the comparison frequency. Likewise, the reference crystal oscillator is also digitally divided down to the reference frequency. The two frequencies are compared in a digital phase/frequency detector to develop an error voltage which corrects the carrier frequency of the modulated oscillator. The reason for dividing the modulated oscillator frequency so many times is to reduce the modulation index enough to limit the peak phase deviation at the reference frequency to a value that will not exceed the linear range of the phase/frequency detector. If the linear range is exceeded, the loop will lose lock. This is why some exciters may lose AFC lock in the presence of low frequency modulation components.

The phase detector output is integrated and low-pass filtered to remove the comparison frequency and all other frequency components above a few Hertz so that the AFC circuit does not try to track-out low frequency modulation. Some FM exciters use a dual-speed PLL in order to keep the loop turn-over frequency low enough to maintain good amplitude and phase response at 30 Hz, while also providing quick lock-up time. The PLL error correction circuitry must respond quickly during the initial frequency scan of the FM band to achieve lock-up to the precision reference oscillator in a few seconds. The loop bandwidth is wide during acquisition and lock-up. After lock is achieved, the bandwidth is reduced to provide the optimum modulation characteristic.

The reference oscillator is usually temperature compensated and requires no warm-up to maintain ±3

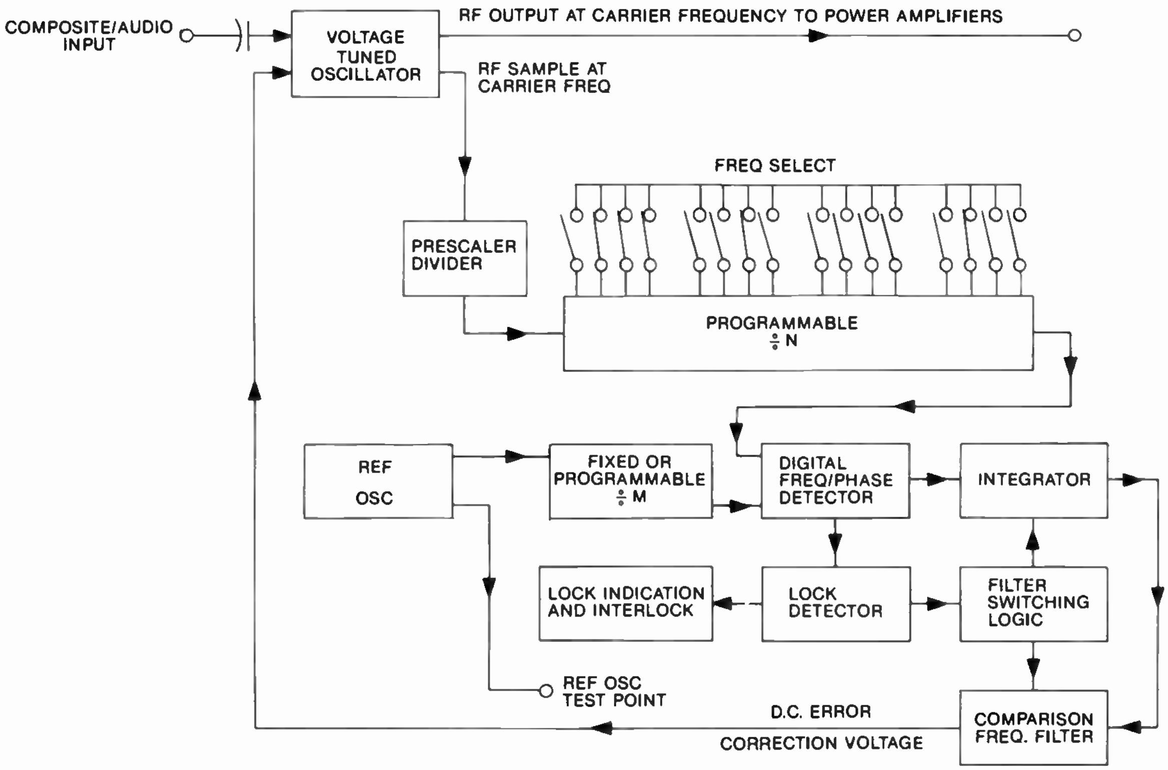


Figure 15. Phase-locked-loop frequency synthesizer.

PPM or better accuracy over the operating temperature range. 10 MHz is often selected as the reference frequency for convenient comparison to international frequency standards. (For more information about PLL frequency synthesizers see Ref. 5 at the end of this chapter.)

Direct Digital Synthesis

A new technology called *direct digital synthesis* (DDS) eliminates the need for a phase-locked loop (PLL) in the FM modulation process by directly synthesizing the carrier frequency, including FM modulation, from a look-up table in a programmable read-only memory (PROM) operating in conjunction with a fast digital-to-analog converter. When this technique is combined with digital signal processing (DSP) technology, the entire process of generating stereo baseband with SCAs, and FM modulating this baseband information onto the RF carrier, can be done entirely in the digital domain. Within the next few years, the cost to performance ratio of DDS/DSP technology will make it competitive with the current analog/digital technology used in present exciters. The full benefit of DDS/DSP technology will require digital mixing and digital transmission of audio information as a digital bit-stream all the way from the digital audio source thru a digital console, digital audio processing, and digital STL to the digital input port of the DSP/DDS exciter. This same technology is used fully in the digital audio broadcast (DAB) service, with technical standards presently being implemented worldwide.

Exciter Metering

Metering of important operating parameters can be provided by a combination of analog metering and a digital LED display. Steady-state parameters are usually selected by multiposition switch and displayed on a conventional analog multimeter. Typical steady-state functions include regulated, preregulated, and unregulated supply voltages; the AFC control voltage; RF power amplifier collector voltage and current; forward output power; and reflected power.

Either a color-coded LED display or a peak reading analog meter are usually provided to constantly monitor the time varying composite signal applied to the modulated oscillator. In either case, a high-speed peak detector gives accurate peak readings on signals from DC to 100 kHz. A one-shot multivibrator circuit provides a clear indication of short transient peaks exceeding 100% modulation.

Exciter Packaging

Protection of sensitive circuits within an FM exciter from external electromagnetic interference is important because the unit is often located in the near field of multiple broadcast antennas operating over a broad range of frequencies. The exciter should be protected from conducted eletromagnetic interference (EMI) by use of RC and/or LC filters on all leads entering the cabinet, including the AC line. Additionally, the power transformer may have an electrostatic shield between the primary and secondary windings. The modulated oscillator is often shock mounted to prevent the transmission of mechanical vibrations from the transmitter's blower. This avoids microphonic pick-up by the modulator that would degrade the FM signal-to-noise ratio. Magnetic shielding of the modulator is also used to prevent hum pick-up from nearby transformers. In some cases, a hum-bucking circuit is provided to help cancel hum induced into the modulator.

The mechanical construction of most present day exciters is designed around a plug-in modular or semi-modular approach, which allows easy removal of sub-assemblies for repair or replacement.

The exciter chassis may be mounted on pull-out slides so that all sub-assemblies are accessible while the unit continues to operate. Front-panel test jacks are often provided to allow measurement of the composite signal without removing or opening the unit.

Exciter Output Stage

The broadband RF amplifier in the exciter amplifies the output of the modulated oscillator from a power level of a few milliwatts up to an output level in the range of 5 to 50 watts. The output stage should be protected against damage by an infinite voltage standing wave ratio (VSWR) at any phase angle.

The typical RF amplifier is designed to have a bandwidth of at least 20 MHz, using successive broadband impedance matching sections for each stage. Each group of matching sections consists of micro-strip or lumped elements.

The broadband performance of the RF amplifier eliminates the need for adjustments to any particular frequency within the FM band. The exciter output is transparent to the signal generated by the modulated oscillator and the amplifier stability is enhanced under varying load conditions.

A micro-strip directional coupler is often incorporated in the RF amplifier output network. This coupler supplies information to the exciter control circuitry which provides automatic control of power output level and provides protection against operation under high VSWR conditions.

All standard FM exciters will produce at least 10 watts output so they can be used as a complete transmitter for educational stations with the addition of a harmonic filter to the output. For higher power levels, the exciter is used to drive an external power amplifier.

SYNCHRONOUS FM BOOSTERS

On July 16, 1987, the FCC approved Docket MM 87-13 which authorized FM stations to increase the power of their on frequency booster facilities from 10 watts maximum to "20% of the maximum permissible ERP for the class of primary station they rebroadcast."[1] The station may not, however, transmit beyond the predicted 1 mV/m contour of the main transmitter for class A and class C stations, the 0.5 mV/m contour for class B, or the 0.7 mV/m contour for class B1 stations.

Effects of Adding a Booster

For the purpose of analyzing the effects of a second carrier, the addition of a booster signal can be treated as interference. A second interfering carrier will both amplitude and phase (frequency) modulate an existing, desired carrier. The characteristics of this apparent modulation are given below.

$$Fm = |fc - fi| \qquad (7)$$
$$B = Ai/Ac$$

where: fc = Main carrier frequency
fi = Booster (interfering) carrier frequency
Ai = Booster (interfering) carrier amplitude
Ac = Main carrier amplitude
B = Percentage of amplitude modulation

Or, in words, an FM receiver detecting two carriers (unmodulated for simplicity), decodes a modulation tone equal in frequency to the absolute value of the frequency separation between the carriers. Moreover, the modulation index (both AM and FM) is simply the ratio of the carrier amplitudes. Notice, however, that the modulation index is never more than one, as increasing the amplitude of the interfering signal over that of the original carrier simply makes the carrier the interfering signal to the booster. For FM, B is measured in radians (see Eq. 7), while for AM, B is the percentage of amplitude modulation produced.

Synchronous Carriers

This gives rise to the need for synchronizing the carrier frequencies. An analysis of the equation for B (FM) shows, given a fixed carrier ratio (fixed modulation index), an increase in carrier frequency separation is equivalent to an increase in Δf, which for FM is equivalent to an increase in detected signal amplitude (see Eq. 8).

If, given $B = \Delta f/Fm$, B is held constant, and Fm is the frequency separation of the carriers, then:

$$\Delta f = B \times Fm \qquad (8)$$

Or, graphically, this phenomenon is shown in Fig. 16.

Knowing this, the advantages of frequency locking the carriers becomes obvious. By taking the limit as the difference in carrier frequency approaches zero, two things happen. First, the frequency of the detected tone approaches zero, and the equivalent FM deviation produced by that tone approaches zero. In other words, the interference disappears.

The Resultant Carrier

We are left with a single frequency carrier whose amplitude depends on the relative phase relationship between the two signals at any given reception point. If the signals are in phase and the carrier ratio is one

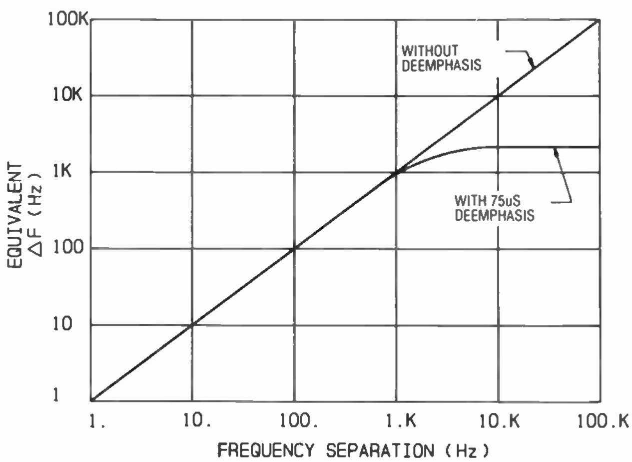


Figure 16. Relationship of carrier frequency separation to equivalent carrier frequency deviation (B = 1).

(0 dB), then the resultant amplitude is twice that of either carrier. On the other hand, if the carriers are 180° (out of phase), the net result is zero. Between these two extremes, the resultant amplitude and phase can be derived by Eq. 9.

$$Ar = \sqrt{[(Ac + A_1 \cos \omega_1(t))^2 + (A_1 \sin \omega_1(t))^2]} \quad (9)$$

where: Ar = Resultant carrier amplitude
Ac = Main carrier amplitude
A_1 = Booster carrier amplitude
$\omega_1(t)$ = Angle between A_1 and Ac

and: $$\phi r = \operatorname{Arctan} \frac{A_1 \sin \omega_1(t)}{Ar + A_1 \cos \omega_1(t)}$$

where: ϕr = Resultant carrier phase angle

Figure 17 shows this vector summation.

The Resultant Effect on Receivers

If we were in a reception area of equal, in-phase and frequency-locked carriers, located on a straight line between the two transmitting antennas, and provided

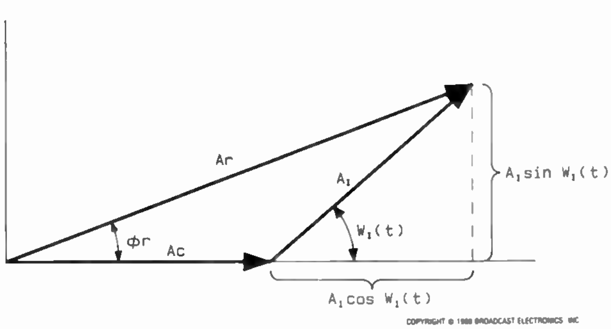


Figure 17. Vector summation of two carriers.

that the information on each carrier is occurring at the same time, all the conditions for adequate reception have been met. (See additional information under modulation delay equalization discussion.)

However, if we were to move 2.5 feet toward either transmitter (roughly equivalent to one-quarter wavelength at 100 MHz), we would now be in an area where the two signals are 180° out of phase, and there would no longer be a signal to detect. Notice that we have only moved one-quarter wavelength, yet have actually moved into a null! This is because we have moved one-quarter wavelength away from one transmitter and one-quarter wavelength closer to the other, giving a total one-half wavelength change. In a mobile receiver, this phenomenon is virtually identical to the "picket fencing" of multipath, except that the interfering signal is not reflected, but rather is a duplicate transmission from a booster site.

For the stationary receiver in an additive state, this is not a problem, and in fact, gross modulation problems can be effectively covered up by virtue of the capture effect of an FM receiver, provided the carrier ratios are adequate to accomplish this. This ratio depends on the particular receiver, but suffice to say it is a smaller ratio than is required to maintain adequate reception with a mobile receiver.

Even relatively small fluctuations in signal strength due to the adding and subtracting of carriers will cause a noticeable fluctuation in receiver signal to noise, unless the residual signal strength in the null area is adequate to keep the receiver well into quieting.

This effect, more than any other, makes the areas of "inadequate" carrier ratios unlistenable. It is important to note that for the mobile receiver, the "picket fencing" problem is present whether or not the carriers are locked.

MODULATION EFFECTS

So far, we have been dealt primarily with the effect of adding a second unmodulated carrier onto a main, unmodulated carrier. At best, understanding the consequences of imperfectly correlated modulation superimposed on two carriers is difficult, if not futile. There are, however, two areas of modulation equalization which are fairly straightforward.

Time Delay Equalization

Consider what happens to two identical signals which have a constant time delay between them. At any frequency where the time delay is equivalent to 180° or N multiples of 180°, there will be complete cancellation, producing a combing effect. This is especially destructive to composite FM stereo and subcarrier performance. In fact, a 7.5 μS group delay addition to one composite signal will cancel a 67 KHz subcarrier. Likewise, a 26.3 μS delay will cancel the stereo pilot tone.

Unfortunately, time delay equalization cannot eliminate cancellation of modulation components in equal carrier areas, since differences in propagation times between the main and the booster will be adequate in some cases to cause the cancellation. At best, the use of group delay can move the location of the nulls relative to the transmitters. Care must be taken to insure the network exhibits a constant group delay, as nonlinear delay can seriously degrade stereo performance, especially stereo separation.

Deviation Calibration

If the frequency deviation of both modulators is not precisely calibrated, a condition of dynamic interference will occur during modulation. For example, two separate modulators are fed identical amplitude, delay equalized sine waves. The level is adjusted to produce a nominal 100% (±75 kHz) modulation. The first modulator swings the carrier exactly ±75 kHz, as predicted, but the second modulator only modulates its carrier ±74 kHz (98.67%).

Careful examination shows that the second carrier will interfere with the first in the following manner. Assuming equal carriers in an additive RF location, and starting at time zero, we have two carriers of exact frequency, producing a single carrier whose amplitude is derived from Eq. 9. As we move positively in frequency with the modulation, the carrier frequencies diverge until, at the peak of modulation, the carriers are 1 kHz apart. From Eq. 7, this produces a 1 kHz FM modulation at a B of 1 (1 kHz deviation) and an AM modulation equivalent to 100% at 1 kHz.

Actually, the detected interference is a frequency sweep from DC to 1 kHz to DC and back to 1 kHz for each complete cycle of modulation applied.

Subjectively, this type of interference sounds similar to white noise, but it is only present during modulation. It is also most prevalent during the maximum modulation peaks, as this is the point of maximum carrier divergence with maximum detected loudness. This relationship of maximum interference during the periods of maximum modulation tends to mask the noise.

Correcting the Interference

Both forms of modulation related interference are most prevalent in areas of nearly equal carriers. With "adequate" carrier ratios, both types are effectively eliminated by the capture effect in the FM receiver. It is also important to keep the proper perspective in assessing the importance of these forms of interference. Remember, these are most prevalent in areas where, even without modulation, the carriers are adding and subtracting, tending to make the signal unlistenable as the receiver location is moved.

BOOSTER SYSTEMS

Two main components must be present at the booster station. A way must be found to transmit the station program material, either in the form of composite stereo, or possibly discrete left and right channels (or mono, if necessary). Composite stereo is preferred, otherwise a second stereo generator would be necessary at the booster. Some form of frequency locking information must also be present.

Using a Radio Link

By far, the most flexible method of interconnection is by the use of a radio link, such as a composite STL between the main transmitter and the booster. This method has several advantages, including high quality transmission, reliability, total signal control, and economy. It is also capable of transmitting the frequency reference signal with the composite stereo via subcarrier. In this way, one radio link supplies both the station programming and reference information.

An Alternate Approach—Leased Phone Lines

In some areas, frequency allocations for a radio link are not available. In such instances, dedicated phone lines may be the only alternative. A voice quality line can be used, provided that the frequency locking reference tone is in the 300 Hz to 3 kHz range. The reference tone could also be sent to the booster site on a composite FM subcarrier of the main transmitter.

RF POWER AMPLIFIERS

The remainder of the FM transmitter consists of a chain of power amplifiers, each having from 6 dB to 20 dB of power gain. Ideally, the transmitter should have as wide a bandwidth as practical with a minimum of tuned stages. Broadband solid-state amplifiers are preferred to eliminate tuned networks in the RF path. Higher powered transmitters in the multi-kilowatt range may use multiple tube stages each with fairly low gain, such as in the grounded grid configuration or a single grid driven power amplifier (PA) stage with high gain and efficiency. The dollars/watt economics of single-tube transmitters outweigh the bandwidth benefits of solid-state transmitters at the higher power levels with present technology. Design improvements in tube-type power amplifiers have concentrated on improving bandwidth, reliability, and cost effectiveness.

TRANSMITTER POWER OUTPUT REQUIREMENTS

The FCC regulates the power of FM broadcast stations in terms of effective radiated power (ERP). The authorized ERP applies only to the horizontally polarized component of radiation. Elliptical or circular polarization is also permitted where the ERP of the vertically polarized component may be as great as the authorized horizontal component. This means that twice as much total power may be radiated and twice as much transmitter power will be required.

The transmitter power requirement can be reduced by increasing the gain of the antenna. There is, of course, an economic trade off between the cost of a higher gain antenna versus the cost of a larger transmitter and the added primary power costs. For a high ERP, it is common to use antennas with up to 12 elements which provide a power gain of about 12.6 (or 6.3 in each polarization).

The long transmission lines associated with the tall towers commonly used are a source of considerable power loss. For example, the efficiency of 2,000 ft of 3⅛ in. rigid coax at 100 MHz is only about 62%.

FM transmitters are designed to operate over a range of power outputs so that with a few basic sizes any required power output can be furnished. Popular maximum ratings range from 100 watts to 70 kilowatts. Most installations use a maximum transmitter power output (TPO) of thirty kilowatts because it is more economical to achieve the maximum 100 kilowatts ERP with circular polarization by means of sufficient antenna gain.

RF Power Amplifier Performance Requirements

The basic function of the power amplifier is to amplify the power of the exciter output to the authorized transmitter power output level. Most of the overall transmitter performance characteristics are determined by the exciter, but a few are established or affected by the power amplifier characteristics:

1. The output at harmonics of the carrier frequency is almost completely a function of the attenuation provided by the output tank circuit and output low-pass/notch filters. The limit in decibels is [43 dB + 10(LOG watts) dB] or 80 dB whichever is less. (73 dB for 1 kW output or 80 dB for 5 kW and higher.)
2. The major source of AM noise usually originates in the last power amplifier stage. The FCC limit is 50 dB below 100% modulation.
3. The RF power output control system, which must keep the output within +5% and −10% of authorized output, is usually achieved in the final power amplifier.
4. Inadequate passband, particularly with respect to phase linearity across the signal bandwidth, can reduce stereo separation and cause SCA crosstalk.
5. The presence of standing waves on the transmission line to the antenna may also react with the power amplifier to cause degraded stereo separation and SCA crosstalk.

The power amplifiers should provide trouble-free service and be easy to maintain and repair. Good overall efficiency is also desirable to reduce the primary power consumption.

Power Amplifier Bandwidth Considerations

As mentioned earlier, the FM signal theoretically occupies infinite bandwidth. In practice, however, truncation of the insignificant sidebands (typically less than 1% of the carrier) makes the system practical by accepting a certain degree of signal degradation. The transmitter power amplifier bandwidth affects the modulation performance. Available bandwidth determines the amplitude response and group delay response. There is a trade-off involved between the bandwidth, gain, and efficiency in the design of a power amplifier.

The bandwidth of an amplifier is determined by the load resistance across the tuned circuit and the output or input capacitance of the amplifier. For a single-tuned circuit, the bandwidth is proportional to the ratio of capacitive reactance to resistance:

$$BW \propto \frac{K}{2\pi f R_L(C)} = \frac{K(X_c)}{R_L} \qquad (10)$$

where: BW = Bandwidth between half-power points (BW_3)
K = Proportionality constant
R_L = Load resistance (appearing across tuned circuit)
C = Total capacitance of tuned circuit (includes stray capacitances and output or input capacitances of the tube)
X_c = Capacitive reactance of C
f = Carrier frequency

The load resistance is directly related to the RF voltage swing on the tube element. For the same power and efficiency, the bandwidth can be increased if the capacitance is reduced.

INTERMEDIATE POWER AMPLIFIERS

The intermediate power amplifier (IPA) is located between the exciter and the final amplifier in higher power transmitters that require more than about 30 watts of drive to the final amplifier. The IPA may consist of one or more tubes or solid-state amplifier modules.

Interstage Coupling Circuits

The separate IPA output circuit and the final amplifier input circuit are often coupled together by a coaxial transmission line. Impedance matching is usually accomplished at either end by one of the configurations shown in Figs. 18A, 18B, 18C, and 18D.

All of the classical circuits shown in Fig. 18, except "D", require some interactive adjustment of the tuning and loading elements to provide a satisfactory impedance match for each operating frequency and RF drive level. The circuit in Fig. 18D utilizes multiple LC sections, with each section providing a small step in the total impedance transformation. This technique provides a broadband impedance match without adjustment; thereby improving the transmitter's stability, ease of operation and maintainability. A single grid resonating control is sufficient to tune and match the 50 ohm driver impedance to the high input impedance of the grid over the entire 88–108 MHz FM broadcast band with a 4:1 range of RF power levels.

The interconnecting transmission line between the coupling circuits should be properly matched to avoid

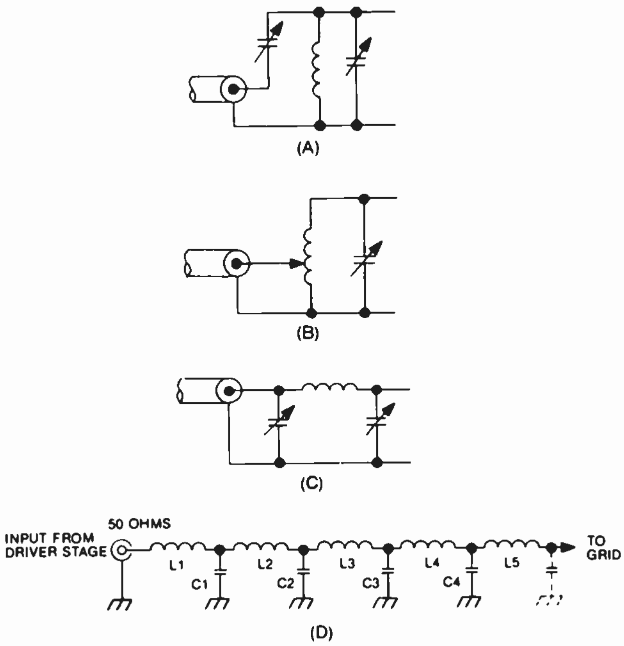


Figure 18. Interstage RF coupling circuits.

a high VSWR. Directional wattmeters are normally placed in the line to measure forward and reflected power from which standing wave ratio can be established. The VSWR is established by the match at the load end of the transmission line.

The transmission line matching problem is eliminated in some transmitters by integrating an IPA stage utilizing a tube(s) into the grid circuit of the final amplifier stage by having the plate of the IPA and the grid of the final tube share a common tuned circuit. This technique has the advantage of simplicity by not transforming the impedance down to 50 ohms and then back up to the grid impedance level, but does not allow the IPA to be connected directly to the antenna as a low-power back-up system.

Solid-state RF power devices possess a very low load impedance at the device output terminal, so that an impedance transformation that goes through the 50 ohm intermediate impedance level is required to couple these devices into the relatively high impedance of the final amplifier grid circuit. Therefore, virtually all solid-state IPA systems have a 50 ohm impedance point within the system that can be used to feed the antenna in an emergency.

High power transmitters utilizing a grounded-grid amplifier configuration in the output stage, require large amounts of drive power (typically greater than 1,500 watts). The IPA may be a standard 3 kW transmitter that can also be used as a stand-by transmitter.

Most of the newer design high-power transmitters only require between 150 and 600 watts of drive into a high gain final amplifier. This permits the use of a system of solid-state, wideband modules to boost the exciter's power up to the level required to drive the grid of the final tube.

Solid-State IPA Systems

A solid-state IPA almost always consists of a system of individual amplifier modules that are combined to provide the desired power output. The advantages of using several lower power modules instead of a single high-power amplifier are:

1. Redundancy is provided by isolating the input and output of each module, permitting uninterrupted operation at reduced power if one or more of the modules fails.
2. The ability to repair or replace failed modules without having to go off-the-air.
3. More effective cooling of each power device junction by splitting the concentration of heat to be dissipated into several areas instead of one small area.
4. Better isolation between the amplifier modules and the input circuit of the final power amplifier is provided by the combiner/isolator.
5. Redundant power supplies and air cooling systems for each module improve overall reliability.

Each RF power amplifier module consists of one or more solid-state devices with broadband impedance transformation networks for input and output matching. A new generation of class "C" BIPOLAR and MOSFET devices permit the design of broadband amplifier stages that exhibit both high efficiency and the wide bandwidth necessary to cover the FM broadcast band.

Regardless of which type of solid-state device is used, the input impedance is always lower than the desired 50 ohm input impedance, so a broadband impedance transformation scheme is required. This is usually accomplished by a combination of coaxial baluns and push-pull coaxial line sections that are cross-coupled to provide 4:1 or higher transformation ratios over the FM band.

By operating two devices in push-pull, the input impedance (differential) is double that of a single ended circuit and the suppression of even order harmonics is enhanced. Two devices fed in this manner also provide some degree of redundancy within the module itself since partial RF output can be obtained with one device failed. In a similar manner, the low output impedance of these solid-state devices can be transformed up to the desired 50 ohm module output impedance where combining occurs. Fig. 19 illustrates a simplified schematic of a broadband IPA module utilizing the push-pull configuration.

IPA Splitting and Combining

There are two types of splitting/combining schemes used:

1. 90° hybrid splitter or combiner ("N-1" hybrids required to split or combine "N" inputs). (See section on transmitter output combining.)

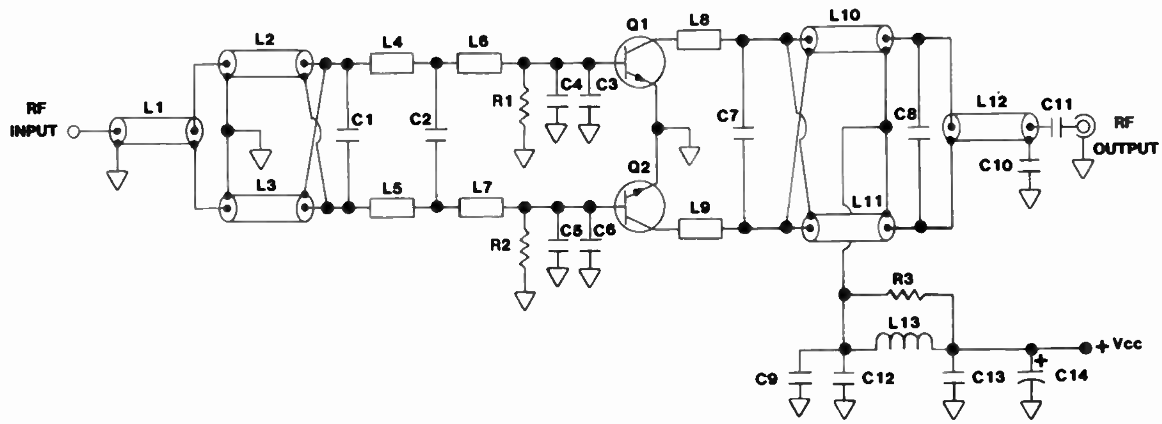


Figure 19. Simplified schematic of a broadband intermediate power amplifier module.

2. Wilkinson "N-way" in-phase splitter or combiner.

Either type of splitter/combiner must provide isolation between the individual power amplifier modules and low loss splitting or combining of the total power. By choosing the proper lengths for the coaxial interconnecting cables, either of the above methods can be configured to operate the individual modules "in-phase" so that the loading on each of the modules tracks the other modules when the impedance at the output of the combiner is varied.

The cascaded 90° hybrid system shown in Fig. 20 provides double isolation between the IPA and the grid circuit by first combining the two pairs of amplifiers and then combining the outputs of the first two combiners. A portion of the reflected power, caused by a mismatch at the output, will be dissipated in the reject loads so that the IPA modules will operate into a lower VSWR than exists at the output. The unbalanced 50 ohm reject loads are accessible for monitoring of reject load power, which is useful in determining the balance of the system. The coaxial interconnecting cables between module pairs must be offset in electrical length by one-quarter wavelength (90°) at the operating frequency so that the modules operate in phase while the hybrids operate in quadrature.

The Wilkinson system shown in Fig. 21 is a simple and effective way to split and combine modules operating in phase, but usually requires a balanced reject load making reject power measurements more difficult. By adding additional coaxial balun sections to the Wilkinson, it is possible to use unbalanced reject loads. Since the Wilkinson operates in phase, all of the coaxial interconnecting cables should be equal in length.

Since most IPA splitter/combiner systems are designed around a 50 ohm input and output impedance level, these systems can be easily used as a low power stand-by transmitter by routing the output to the

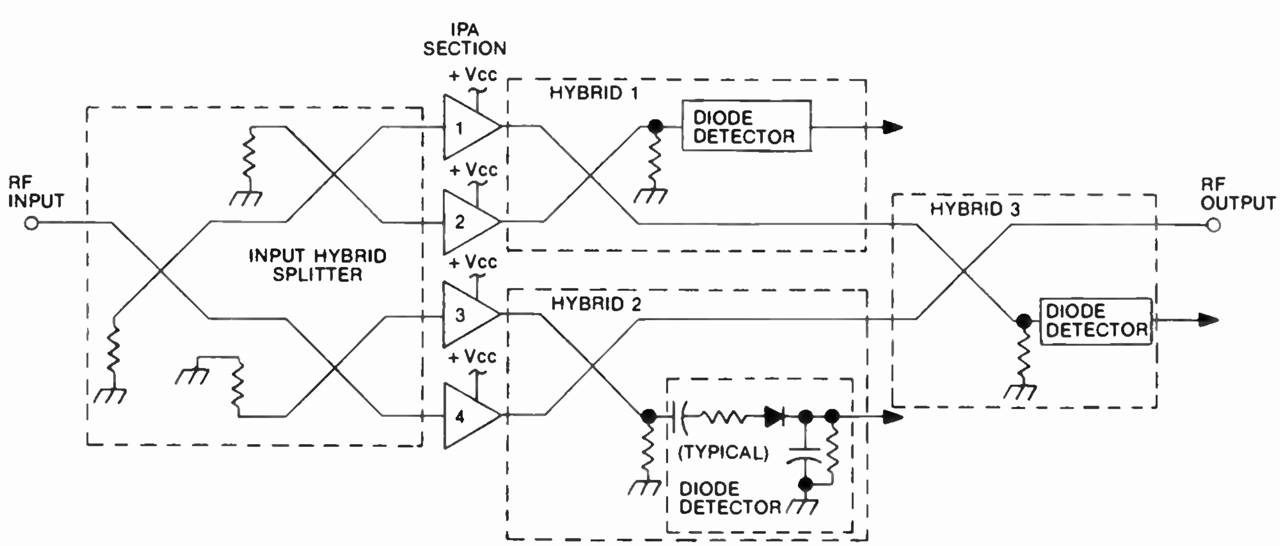


Figure 20. Cascaded 90° hybrid splitting/combining system.

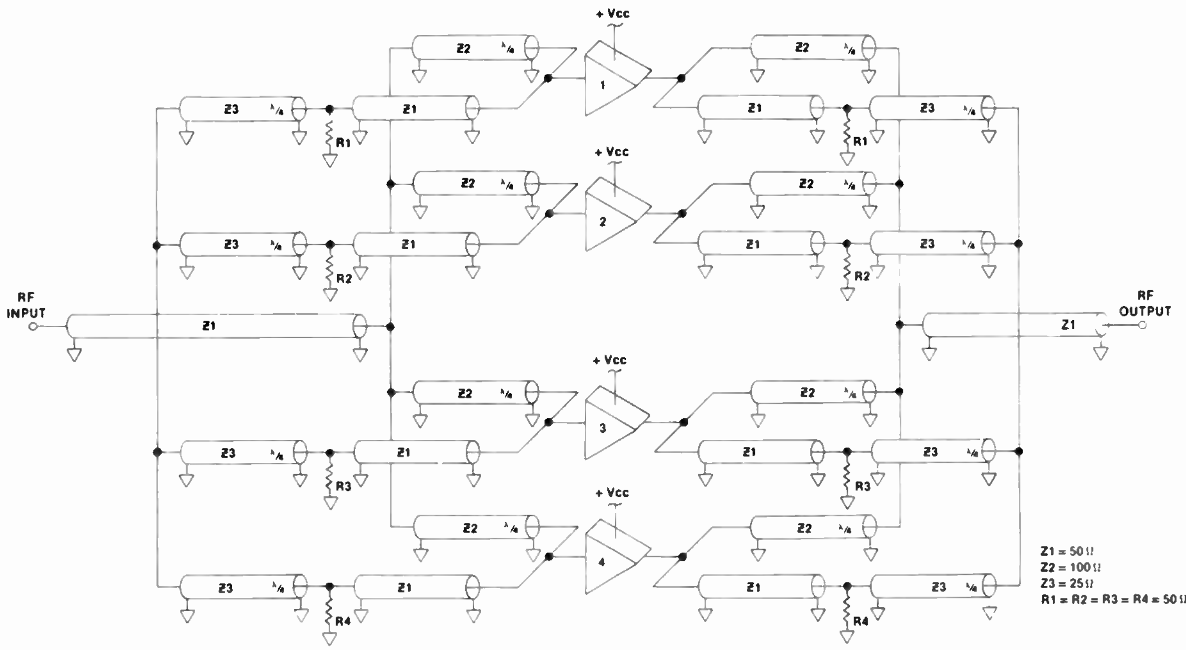


Figure 21. Wilkinson/Gysel in-phase splitting/combining system with unbalanced reject loads.

antenna system. An RF low-pass filter (LPF) is required only when directly feeding the antenna system. The harmonic suppression of the IPA is not as critical when driving a nonlinear power amplifier that also generates harmonics, because this stage will have its own LPF.

SOLID-STATE FM BROADCAST TRANSMITTERS

The techniques used to construct IPA systems can also be used to construct a completely solid-state transmitter using arrays of combined modules for the final output stage. An additional RF low-pass filter is usually required to meet FCC emission requirements. Several manufacturers offer solid-state FM broadcast transmitters with power outputs ranging from 100 watts up to 10 kilowatts, but present economic factors still favor the single tube FM transmitter for power levels above 5 kilowatts. In order for a solid-state transmitter to be cost and power consumption competitive with a single-tube transmitter, the efficiency of the solid-state RF power amplifiers and combining system have to be at least 80% efficient. This high efficiency is difficult to achieve with presently available solid-state devices at VHF frequencies.

Advantages of Solid-State Transmitters

The primary advantages of a solid-state transmitter are the built-in amplifier redundancy, the ability to cover the entire FM band without the need for retuning, and elimination of tube replacement costs. Even a tubeless transmitter is not entirely maintenance free, because high-power solid-state RF devices do age and wear out like tubes, depending on the junction operating temperature and internal current density. The typical life of an RF transistor conservatively operated, is typically at least ten years.

Output Filtering Requirements

Broadband, solid-state transmitters are much more likely to generate RF intermodulation products than single-tube transmitters which have selectivity in the output stage. Broadband, solid state, amplifiers may require tuned output band-pass filters when operating in a dense RF environment in order to prevent RF intermodulation products from being generated in the PA modules.

For more information about solid state amplifiers and hybrid splitter/combiners, see Refs. 6, 7, 8, and 9 at the end of this chapter.

VACUUM TUBE POWER AMPLIFIER CIRCUITS

The amplitude of an FM signal remains constant with modulation so that efficient class B and C amplifiers can be used. Most exciters being manufactured at this time, provide 10 watts to 50 watts of output power. It is technically feasible to develop transistor amplifiers for any required power, but they are not yet economically competitive at high power levels. Additional circuitry is involved because it takes the combined output of many transistors to produce a few kilowatts of power output. For this reason, the following discussion will relate to vacuum tube amplifier circuits.

FM broadcast power amplifier circuits have evolved into two basic types. One type uses a tetrode or pentode tube in a grid-driven circuit, while the other uses a high-mu triode in a cathode-driven (grounded-grid) circuit.

Cathode-Driven Triode Amplifiers

The high-mu triodes being used in cathode-driven (grounded-grid) FM amplifiers were originally developed for linear single sideband (SSB) amplifiers. Their characteristics are well adapted to FM broadcast use because the circuit is very simple and no screen or grid bias power supplies are required. Fig. 22 shows the basic circuit configuration. In this case, the grid is connected directly to chassis ground. DC grid current is the difference between DC cathode current and DC plate current. The output tank circuit is a shorted coaxial cavity which is capacitively loaded by the tube output and stray circuit capacitance. A small capacitor is used for trimming the tuning, and another small variable capacitor is used for adjusting the loading. A pi-network matches the 50 ohm input to the tube cathode.

The triodes are usually operated in the class B mode in order to achieve maximum power gain, which is on the order of 20 (13 dB). They can be driven into class C operation by providing grid bias. This increases the plate efficiency, but also requires increased drive power.

Most of the drive power into a grounded-grid amplifier is fed through the tube and appears in the stage's output. This increases the apparent efficiency so that the efficiency factor given by the manufacturer may be higher than the actual plate efficiency of the tube. The true plate efficiency is determined by dividing the output power by the total input power, which includes both the DC plate input power (Ip × Ep) and the RF drive power. Since most of the drive power is fed through the tube, any changes in loading of the output circuit will also affect the input tuning and driver stage.

There is RF drive voltage on the cathode (filament) of the tube, so some means of decoupling must be used to block it from the filament transformer. One method employs high current RF chokes since the inductance can be very low at this frequency range. The other commonly used method feeds the filament power through the input tank circuit inductor.

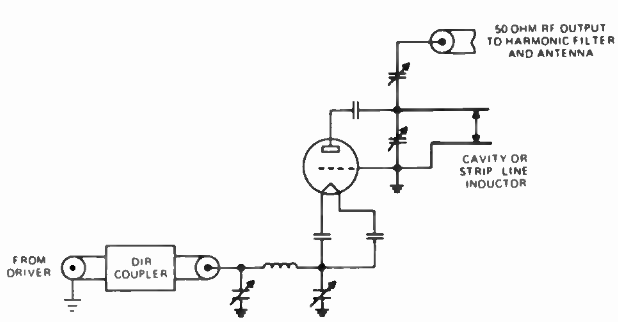


Figure 22. Cathode-driven triode power amplifier.

Cathode-driven stages are normally used only for the higher power stages. The first stage in a multi-tube transmitter is nearly always a tetrode because of its higher power gain.

Grounded-Grid versus Grid-Driven Tetrode Operation

Tetrodes may also be operated in the grounded-grid configuration by placing both the control grid and the screen grid at RF ground. Higher efficiency and gain can be achieved by placing negative bias on the control grid while placing a positive voltage on the screen grid of a cathode-driven tetrode.

The input capacitance of a tetrode in a grounded-grid configuration is much less than a grid-driven configuration and the input impedance is lower, providing better bandwidth.

Approximate input capacitances of some typical tubes are:

TUBE TYPE	Cin (PF) Grounded Grid	Cin (PF) Grounded Cathode
4CX 3,000A	67	140
4CX 3,500A	59	111
4CX 5,000A	53	115
4CX 15,000A	67	161
4CX 20,000A/8990	83	190

The typical drive power requirements, as a function of plate voltage, for a 5 kW power amplifier are:

Configuration	Plate Voltage/Drive Power 4500 Volts	5200 Volts
Grounded-Grid	340 Watts	280 Watts
Grid-Driven	190 Watts	140 Watts

Television transmitters sometimes use a grounded-grid tetrode configuration to increase the input bandwidth of the amplifier. There are several trade-offs between the performance of grounded-grid and grid-driven operation of a tetrode PA with respect to gain, efficiency, amplitude bandwidth, phase bandwidth, and synchronous AM under equivalent operating conditions:

1. When driving the PA into saturation, the bandwidth of the PA is limited by the output cavity bandwidth in the grounded-grid amplifier. The PA bandwidth in the grid-driven amplifier is limited by the input circuit "Q", which is basically determined by the amount of swamping resistance.
2. Output bandwidth under saturation can be improved in either configuration by reducing the plate voltage. This involves a trade-off in efficiency with a smaller voltage swing. The bandwidth improve-

ment can be obtained with a loss of PA gain and efficiency.

3. A grounded-grid saturated PA improves bandwidth over a grid-driven saturated PA at the expense of amplifier gain. Best performance for FM operation is obtained when the amplifier is driven into saturation where little change in output power occurs with increasing drive power. Maximum efficiency also occurs at this point.
4. The phase linearity in the 0.5 dB bandwidth is better in a grid-driven configuration. The class C, grounded-grid, PA exhibits a more nonlinear phase slope within the passband, yet has a wider amplitude bandwidth. This phenomenon is due to interaction of the input and output circuits because they are effectively connected in series in the grounded-grid configuration. The neutralized grid-driven PA provides more isolation between these networks, so they behave more like independent filters.

Grid-Driven Tetrode and Pentode Amplifiers

A small tetrode tube, such as the 4CX250B, is commonly used as the only amplifier stage in 250 watt transmitters and as the driver for higher power stages. The largest single stage transmitter presently available uses a high gain 5CX1500A pentode to deliver 2.5 kW. Higher power levels require two stages (an IPA and the final PA).

Transmitters using tetrode amplifiers throughout usually have one less stage than those using triodes. Since tetrodes have higher power gain, they are driven into class C operation for high plate efficiency. Against these advantages is the requirement for neutralization, along with screen and bias power supplies.

Fig. 23 shows a schematic of a grid-driven tetrode amplifier. In this example, the screen is operated at DC ground potential, and the cathode (filament) is operated below ground by the amount of screen voltage required. This is called grounded-screen operation. It has the advantage that stability problems, due to undesired resonances in the screen bypass capacitors, are eliminated. With directly heated tubes, it is necessary to use filament bypass capacitors. During grounded-screen operation, these bypass capacitors will need to have a higher breakdown voltage rating since they will have the DC screen voltage across them. The filament transformer must have additional insulation to withstand the DC screen voltage. The screen power supply provides a negative voltage in series with the cathode-to-ground, and must have the additional capacity to handle the sum of the plate and screen currents. A coaxial cavity is used in the output circuit so that the circulating current is spread over large surfaces keeping the losses very low. This cavity is a shorted quarter-wavelength transmission line section which resonates the tube's output capacitance. The length is preset to the desired carrier frequency and then a small value variable capacitor is used to trim

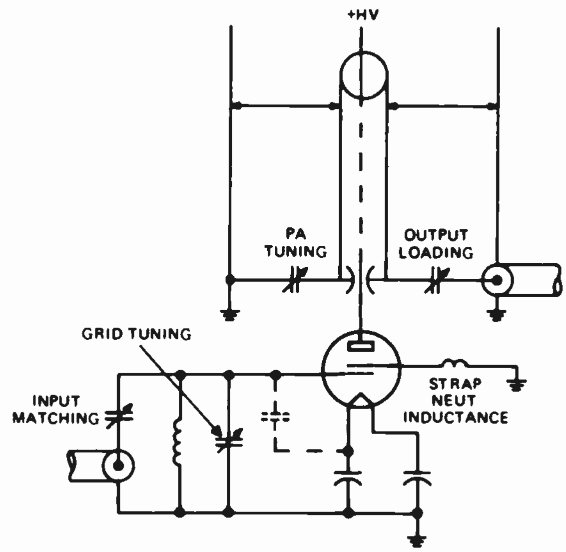


Figure 23. Grid-driven grounded screen tetrode power amplifiers.

the system to resonance. Capacitive output coupling is used from the high RF voltage point to the 50 ohm transmission line. The 50 ohm input is capacitively coupled into the grid circuit inductor to provide the correct impedance match.

Pentode amplifiers have even higher gain than their tetrode counterparts. The circuit configuration and bias supply requirements for the pentode are similar to the tetrode since the third (suppressor) grid is tied directly to ground. The additional isolating effect of the (suppressor) grid eliminates the need for neutralization in the pentode amplifier.

Impedance Matching Into The Grid

The grid circuit is usually loaded (swamped) with added resistance. The purpose of this resistance is to broaden the bandwidth of the circuit by lowering the circuit "Q" and to provide a more constant load to the driver. It also makes neutralizing less critical so that the amplifier is less likely to become unstable with varying output circuit loading.

Cathode or filament lead inductance from inside the tube, through the socket and filament capacitors to ground, can heavily load the input circuit. This is caused by RF current flowing from grid to filament through the tube capacitance and then through the filament lead inductance to ground. An RF voltage is developed on the filament which in effect causes the tube to be partly cathode driven. This undesirable extra drive power requirement can be minimized by series resonating the cathode return path with the filament bypass capacitors or by minimizing the cathode-to-ground inductance by using a specially designed tube socket using thin-film dielectric "sandwich" capacitors for coupling and bypassing.

High power, grid driven, class C, amplifiers require a swing of several hundred RF volts on the grid. To develop this high voltage swing, the input impedance of the grid must be increased by the grid input matching circuit. Since the capacitance between the grid and the other tube elements may be 100 picofarads or more, the capacitive reactance at 100 MHz will be very low unless the input capacitance is parallel resonated with an inductor. Bandwidth can be maximized by minimizing any additional circuit capacitance and utilizing a portion of the tube input capacitance as part of the impedance transformation network. Figs. 24A and 24B show two popular methods of resonating and matching into the grid of a high power tube. Both methods can be analyzed by recognizing that the desired impedance transformation is produced by an equivalent "L" network.

In Fig. 24A, a variable inductor (L_{in}) is used to raise the input reactance of the tube by bringing the tube input capacitance (C_{in}) almost to parallel resonance. Parallel resonance is not reached since a small amount of parallel capacitance (C_p) is required by the equivalent "L" network to transform the high impedance (Z_{in}) of the tube down to a lower value through the series matching inductor (L_s). This configuration has the advantage of providing a low-pass filter by using part of the tube's input capacitance to form (C_p), instead of an external bandwidth-reducing variable capacitor.

Fig. 24B uses variable inductor (L_{in}) to take the input capacitance (C_{in}) past parallel resonance so that the tube's input impedance becomes slightly inductive. The variable series matching capacitor (C_s) forms the rest of the equivalent "L" network. This configuration is a high-pass filter.

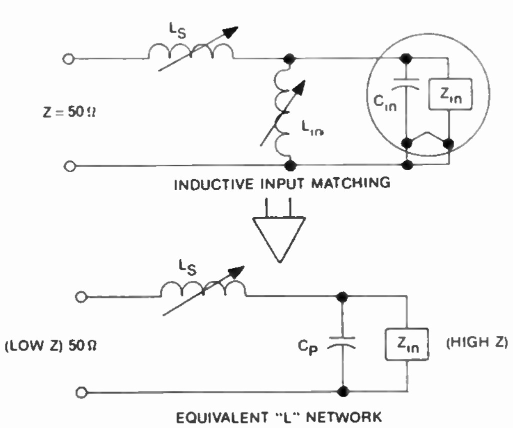


Figure 24A. Inductive input matching.

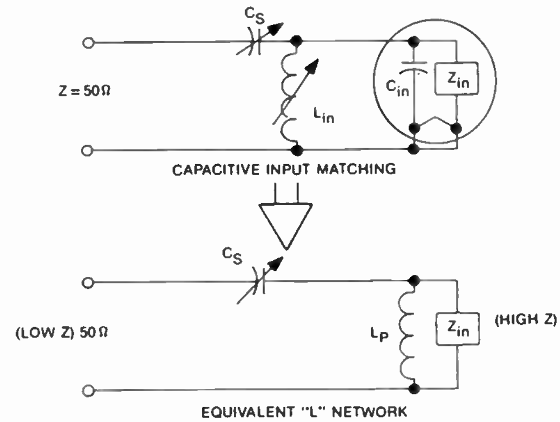


Figure 24B. Capacitive input matching.

Neutralization

Apparently none of the cathode-driven amplifiers utilizing triodes require neutralization. It is necessary that the grid-to-ground inductance, both internal and external to the tube, be kept very low to maintain this advantage. Omission of neutralization will allow a small amount of interaction between the output circuit and the input circuit through the plate-to-filament capacitance. This effect is not very noticeable because of the large coupling between the input and output circuits through the electron beam of the tube. Cathode-driven tetrodes have higher gain and therefore require some form of neutralization.

Grid-driven, high gain tetrodes need accurate neutralization for best stability and performance. Self-neutralization can be accomplished very simply by placing a small amount of inductance between the tube screen grid and ground. This inductance is usually in the form of several short, adjustable-length straps. The RF current flowing from plate to screen in the tube also flows through this screen lead inductance. This develops a small RF voltage on the screen, of the opposite phase, which cancels the voltage fed-back through the plate-to-grid capacitance. This method of lowering the self-neutralizing frequency of the tube works only if the self-neutralizing frequency of the tube/socket combination is above the desired operating frequency before the inductance is added. (Ref. 10 at the end of this chapter has more information on the theory of self neutralization.) Special attention must be given to minimizing the inductances in the tube socket by integrating distributed bypass capacitors into the socket and cavity deck assembly. Pentodes normally do not require neutralization because the suppressor grid effectively isolates the plate from the grid.

POWER AMPLIFIER OUTPUT CIRCUITS

Usually, the output circuit consists of a "high-Q" (low loss) transmission line cavity, strip line, or a lumped inductor that resonates the tube output capacitance. A means of trimming the tuning and a means of adjusting the coupling to the output transmission line must also be provided by the output circuit. The tank circuit loaded "Q" is kept as low as practical to minimize circuit loss and to maintain as wide an RF bandwidth as possible.

The Power Amplifier Cavity

The vacuum tube power amplifier is constructed in an enclosure containing distributed tank circuit elements for minimum loss. The efficiency of the PA depends on the RF plate voltage swing, the plate current conduction angle, and the cavity efficiency. The cavity efficiency is related to the ratio of loaded to unloaded "Q" as follows:

$$N = 1 - \left(\frac{Q_l}{Q_u}\right) \times 100$$

where: N = Efficiency in percent
Q_L = Loaded "Q" of cavity
Q_u = Unloaded "Q" of cavity

The loaded "Q" depends on the plate load impedance and output circuit capacitance. Unloaded "Q" depends on the cavity volume and the RF resistivity of the conductors due to skin effects. A high unloaded "Q" is desirable, as is a low-loaded "Q", for best efficiency. As the loaded "Q" goes up, the bandwidth decreases. For a given tube output capacitance and power level, loaded "Q" decreases with decreasing plate voltage or with increasing plate current. The increase in bandwidth at reduced plate voltage occurs because the load resistance is directly related to the RF voltage swing on the tube element. For the same power and efficiency, the bandwidth can also be increased if the output capacitance is reduced. Power tube selection and minimization of stray capacitance are areas of prime concern when designing for maximum bandwidth.

The methods used to improve the bandwidth of PA output circuits include minimizing added tuning capacitance. The ideal case would be to resonate the plate capacitance alone with a "perfect" inductor, but practical quarter-wave cavities require either the addition of a variable capacitor or a variable inductor using sliding contacts for tuning. The inherent mechanical and electrical compromises are the requirement for a plate DC blocking capacitor and the presence of maximum RF current at the grounded end of the line where the conductor may be nonhomogeneous, with accompanying losses in the contact resistance.

The Quarter-Wavelength Cavity

The quarter-wavelength coaxial cavity is the compact and popular output circuit illustrated in Fig. 25. The tube anode is coupled through a DC blocking capacitor to a shortened quarter-wavelength transmission line. The tube's output capacitance is brought to resonance by the inductive component of the transmission line that is physically less than a quarter-wavelength long. Plate tuning can be accomplished either by adding end-loading capacitance at the high-impedance end of the line with a variable capacitor or by sliding the ground plane at the low-impedance end of the line. The plate tuning capacitor may be a sliding or rotating plate near the anode of the tube. The center conductor of the transmission line (chimney) is at DC ground while the anode of the tube operates at a high RF and

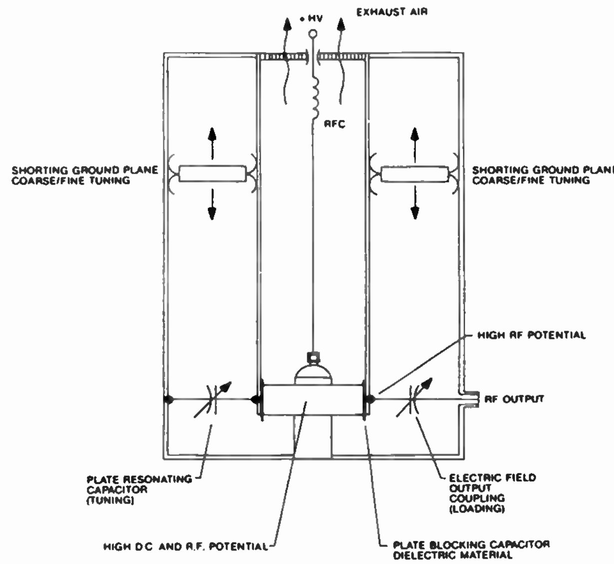


Figure 25. The quarter-wavelength cavity.

DC potential. DC voltage is fed through an isolated quarter-wavelength decoupling network inside the chimney to the anode of the tube, while the plate blocking capacitor prevents DC current flow from the anode into the chimney.

The Folded-Half-Wavelength Cavity

Another approach to VHF power amplification uses the re-entrant, folded half-wave cavity design illustrated in Fig. 26. The DC anode voltage is applied to the lower portion of the plate line through a choke at the RF voltage null point. This RF voltage null point is reasonably consistent through the entire commercial FM broadcast band. The second harmonic voltage peak, which is located at this same point, exhibits a high source impedance and high voltage, allowing suppression with a simple series LC arrangement. This method provides about 50 dB of suppression to the second harmonic with no power loss at the fundamental frequency. The half-wave line is tuned by mechanically expanding or contracting the physical length of a flexible extension (bellows) on the end of the secondary transmission line stub, which is located concentrically within the primary transmission line (chimney). The bellows assembly is constructed of beryllium copper, providing high conductivity and long flex life.

Coarse frequency adjustment is accomplished by presetting the depth of the top secondary section of plate line into the tank cavity. A plate blocking capacitor is unnecessary since an air gap is provided between the primary line carrying the DC plate voltage and the secondary line at DC ground potential. Ohmic losses are minimized because the high RF current point is located in the central area of the homogeneous primary line where no joints, fasteners, or obstructions occur. Due to the folded nature, this configuration requires

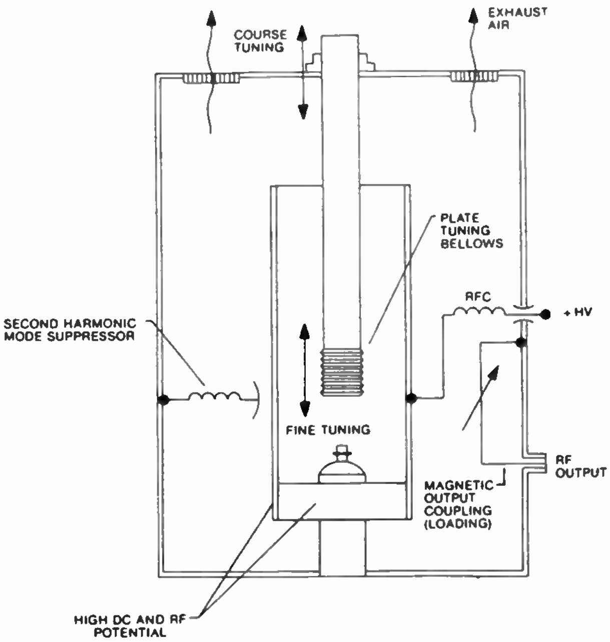


Figure 26. The folded-half-wavelength cavity.

only slightly more physical space than the quarter-wave design.

Other power amplifier configurations may use lumped components or hybrid combinations with distributed transmission line elements to achieve similar results. The discrete circuit elements are chosen for their individual inductance or capacitance, instead of being operated in a purely quarter-wave or half-wave mode. Stray inductance and capacitance add to the component values resulting in the hybrid nature of these circuits.

The RF voltage and current distributions for the quarter-wavelength and the folded-half-wavelength cavities are shown in Fig. 27.

Regardless of the specific configuration, the output circuit must transform the high resonant plate impedance down to the output transmission line impedance of 50 ohms.

The bandwidth of either a quarter-wave or half-wave transmission line cavity is optimized by choosing the highest characteristic impedance mechanically allowable. The center conductor is sized for minimum impedance discontinuity and is clamped to the outer surface of the anode fins for better heat transfer. In the folded half-wave cavity, the secondary tuning line (with adjustable bellows) is sized to maintain a similar characteristic impedance without appreciable end-loading capacitance.

(Refs. 10 and 16 at the end of this chapter give detailed information about the design of tube type RF power amplifiers.)

Output Coupling

Power may be coupled from a quarter-wavelength cavity to the transmission line by a capacitive probe located at the high RF voltage point located at the anode end of the quarter-wave line as shown in Fig. 25. The loaded "Q" of this circuit varies with the degree of capacitive coupling. Another method of coupling power from the quarter-wavelength cavity uses a tuned loop located near the grounded (high current) end of the line. In this case, the tuned loop operates both as an inductive and a capacitive pick-up device.

Power may be coupled from the half-wave line by an inductive loop located in the strong fundamental magnetic field near the center of the cavity as shown in Fig. 26. One end of the output loop that couples energy to the transmission line is electrically grounded to the cavity wall. Output loading is mechanically controlled by changing the position of the loop in the magnetic field generated at the null point of the primary line structure. Multiple phosphor-bronze leaves pro-

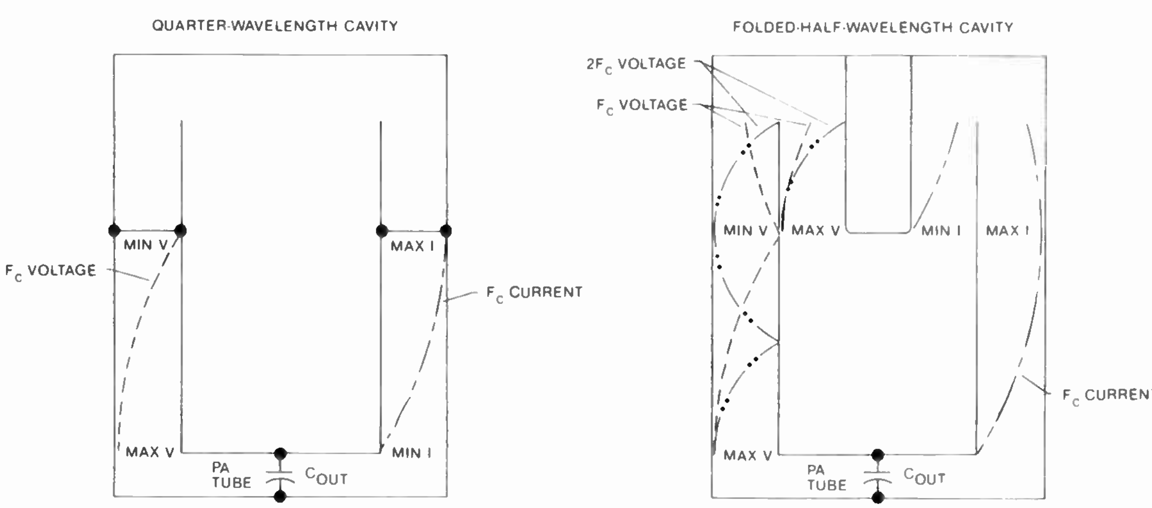


Figure 27. Cavity RF voltage and current distributions.

vide the connections to the loop allowing mechanical movement without sliding contacts. The loaded "Q" of the cavity varies as the square of the effective loop area and inversely as the square of the distance of the loop center from the cavity center axis. As this loop is positioned so that it links more or less magnetic field, it determines the output loading of the transmitter. Magnetic coupling at the voltage null point also eliminates interaction between the tuning and loading controls.

Power Amplifier Source Impedance

At the milliwatt levels used in RF test equipment, it is customary to provide both 50 ohm source and load impedances at both ends of a coaxial transmission line. This approach minimizes any reflections on the line since both the transmitter (source) and the termination (load) will absorb reflected energy. A 50 source impedance is usually provided by placing a 50 ohm build-out resistor in series with a low impedance voltage source (Thevenin equivalent). The closed circuit voltage with this configuration is exactly one-half of the open circuit voltage, meaning that half of the total available RF power is dissipated in the source resistance. The best possible efficiency for this system is 50% assuming the voltage source is 100% efficient without the source resistance.

It becomes obvious that while an FM transmitter is designed to drive a 50 ohm load, it does not itself have an output source impedance of 50 ohms. In order to achieve high efficiency, the transmitter must have a very low output source impedance so that nearly all of the power is delivered to the load. The plate dissipation indirectly represents some of the power lost within the low source resistance. Since the low source impedance of the transmitter provides a mismatch to reflected power from the load, this power is almost totally reflected back from the transmitter output stage toward the load again.

EFFECTS OF CIRCUIT TOPOLOGY AND TUNING ON FM MODULATION PERFORMANCE

FM broadcast transmitter RF power amplifiers are typically adjusted for minimum synchronous AM (incidental amplitude modulation) which results in symmetrical amplitude response. This will assure that the transmitter's amplitude passband is properly centered on the FM channel. The upper and lower sidebands will be attenuated equally or symmetrically which is *assumed* to result in optimum FM modulation performance. This will be true if the RF power amplifier circuit topology results in simultaneous symmetry of amplitude and group delay responses.

Actually, symmetry of the group delay response has a much greater effect on FM modulation distortion than the amplitude response. Tuning for symmetrical group delay will cause the phase/time delay errors to affect the upper and lower sidebands equally or symmetrically. The group delay response is constant if the phase shift versus frequency is linear. All components of the signal are delayed in time, but no phase distortion occurs.

The tuning points for symmetrical amplitude response and symmetrical group delay response usually do *not* coincide, depending on the circuit topology. Therefore, simply tuning for minimum synchronous AM (symmetrical amplitude response) does not necessarily result in best FM modulation performance.

Measurements taken on a typical FM transmitter as well as computer simulations[23] showed that tuning the RF power amplifier for symmetrical group delay response resulted in minimum distortion and crosstalk. It confirms that group delay response asymmetry causes higher FM modulation distortion and crosstalk than amplitude response asymmetry. Therefore, RF power amplifier circuit topologies that exhibit coincidence of symmetrical amplitude and group delay responses will result in a better overall FM modulation performance. The transmitter should be tuned for symmetrical group delay response, which results in best FM modulation performance rather than symmetrical amplitude response, which results in minimum synchronous AM.

POWER SUPPLIES

Power supplies provide the appropriate DC or AC voltages to the various subsystems with the transmitter. The range of voltages and currents provided can be from less than five volts at a few milliamperes, to over 10,000 volts at several amperes in a typical FM transmitter. Safety must be a prime consideration when working around potentially lethal power supplies. Power supplies must be designed with adequate bleeder resistors and interlocks to discharge high voltages before an operator can come in contact with these circuits. The degree to which the AC components are filtered out of the DC outputs of the power supplies will, in large part, determine the "asynchronous" (without FM modulation) AM noise of the FM transmitter.

FM transmitters usually contain multiple power supplies for each of the functional blocks within the system. These power supplies fall into two general categories:

1. Single-phase supplies (single input winding on the transformer).
2. Polyphase supplies (three or more input windings on the transformer).

Single-Phase Power Supplies

Single-phase power supplies with conventional full-wave rectification and filtering are most often used for the FM exciter, the control system, bias supplies, and the intermediate power amplifier (IPA). The filament transformer is also a single-phase transformer. A single-phase supply requires a larger filter choke to achieve the "critical inductance" requirement and a greater value of filter capacitance to maintain acceptably low ripple content compared to a polyphase supply. Large value filter components also mean that the greater stored energy in these components can have a more destructive effect if an arc-over occurs. Choke-input filter sections are normally used to help limit the in-rush current while the shunt capacitor is charging during turn-on. This also maximizes utilization of the transformer and rectifiers by keeping the charging current nearly constant, providing the best filtering action. Choke-input filters have the undesirable characteristic of poor voltage regulation over a wide range of loads. The output voltage will "soar" above the nominal value with no load unless there is enough current through the "bleeder" resistor to keep the choke in the constant current range. Fortunately, in a FM transmitter application, the load on the power supply is fairly constant since the power output of the transmitter does not vary significantly with FM modulation. In higher power transmitters where the main power source is three-phase, care must be taken to balance each of the individual single-phase loads among the three phases so that the total load on each of the individual phases is equal.

Polyphase Power Supplies

Polyphase power supplies are used for the final power amplifier high voltage supply in high power transmitters. Sometimes they are also used for tube or solid state IPA supplies. Large blowers used to cool transmitters are usually operated from a three-phase power source. Care must be taken to make the three line connections to the blower motor in the proper sequence so that the motor will turn in proper direction.

The most common type of polyphase supply is three-phase with full-wave rectification and LC filtering. Other polyphase systems encountered in broadcast equipment are usually multiples of the three phases with twelve phase rectifiers becoming more popular. The main advantages of a polyphase power supply are:

1. Division of the load current between three or more lines to reduce line losses and the size of each of the lines.
2. Greatly reduced filtering requirements after rectification due to the low ripple at the output of a polyphase full-wave rectifier.
3. Better voltage regulation with a choke input filter with typically 6% or less "soaring" from no load to full load.
4. Greater choice of output voltages from a given transformer by selection of either a DELTA or WYE configuration.

The main disadvantage of polyphase systems is their susceptibility to phase imbalance, which causes degraded performance of the power supply. If significant imbalance exists in a polyphase system, ripple rejection will be reduced in the polyphase rectifier with a resulting increase in AM noise.

The broadcast engineer should be particularly careful to be sure that the local utility does in fact, provide true three-phase power to the transmitter site. This can usually be verified by making sure that there are three transformers on the utility pole feeding the transmitter site. In many rural areas, the utilities are still synthesizing pseudo-three-phase service by providing the so called "open-delta"(V-V) or "Scott"(T-T) connection with two transformers instead of true three-phase service. Operation on an "open-delta" service will degrade the transmitter's performance and increase the susceptibility of the transmitter to damage from transients on the line. Most transmitter manufacturers state that their warranty is void if the transmitter is connected to an "open-delta" system.

Special Power Supplies

Trends toward completely solid-state power amplification demand lower voltages at much higher currents. Voltage regulation of these high current supplies is necessary to suppress ripple, but the design of these specialized regulators is different from the typical high-voltage power supply. Linear regulators are used at the lower power levels because they are low in efficiency, but they are simple and provide excellent ripple rejection without the need for suppression of switching transients. The linear regulators use series or shunt devices which change resistance dynamically to provide regulation with changes in load and therefore dissipate some of the power within the dynamic resistance.

Switch-mode regulators are used at higher power levels because they are high in efficiency, but they are more complicated and require additional suppression of the switching transients. The high efficiency comes from the digital "on" or "off" nature of the switching regulator which reduces resistive losses by using low loss reactive components to store energy during switching.

In some cases, phase control switching regulation is applied to the high voltage power supply feeding the final output tube. The regulation is accomplished by switching thyristors in the AC mains ahead of the primary winding of the transformer. As the switching duty cycle is reduced, the plate voltage is also reduced. Care must be taken to protect solid-state devices connected to the main power line from lightning transients. Heavy duty transient suppressors are available for this purpose.

Low-voltage/high-current power supplies contain extremely large amounts of stored energy that can be dangerous due to the high peak currents that can occur during a short circuit across a component with high stored energy. For this reason, special attention must be paid to methods of safely discharging these circuits without damage to components or injury to the operator. The voltage regulator should provide short-circuit protection with some type of current limiting. The main danger to the operator from this type of power supply is burns due to the nearly instantaneous heating of metallic tools or other conductors that accidentally get into a high current path such as a short across the filter capacitor.

(For more detailed information on power supply theory, see Refs. 2, 12, and 13 at the end of this chapter.)

Step-Start

Step-start is often used in large power supplies where peak in-rush currents become high when the power supply is first turned on. These peak currents are caused by the need to overcome the hysteresis effect to initially magnetize the core of the transformer when AC power is first applied and to charge the filter capacitor. Step-start systems temporarily insert a resistance or reactance in series with the power lines to limit the current to a reasonable value until initial magnetization of the core and filter charging are completed. The system should be designed so that failure of the step-start mechanism will not prevent getting "on-the-air" without step-start. This is accomplished by using "fail-safe" systems where the current limiting components are only in series with the line during the starting interval until the main circuit is closed in parallel with the step-start circuit, then the connection to the current limiting components is opened.

TRANSMITTER CONTROL SYSTEMS

Transmitter control systems are often overlooked or given little priority in the selection and maintenance of a broadcast transmitter. The transmitter's control system does however, serve several important purposes:

1. Must provide basic on/off control of the transmitter.
2. Must provide overload protection to protect the transmitter from damage.
3. Must provide safety interlock protection to prevent injury to people and accessory equipment such as RF switching equipment or RF loads.
4. Must provide a means for controlling the transmitter output power.
5. Must provide remote control capability and interfacing at installations where the transmitter is not at the same location as the control operator.
6. Must provide warm-up and cool-down timing sequences of filaments or other time sensitive operations.
7. May provide status indications of overloads or other critical parameters.
8. May provide automatic regulation of the transmitter output power.
9. May provide diagnostic indications to aid in adjustment and maintenance.
10. May provide totally automatic operation of the transmitter plant (ATS).
11. May provide integrated remote control capability.

The transmitter's ability to stay "on-the-air" will only be as good as the reliability of the control system, so it is easy to recognize the importance of the selection and correct operation of the transmitter control system.

Early transmitter technology relied on simple relay logic combined with electromechanical contactors to control the transmitter. The speed and variety of overload protection was limited and diagnostics were not available. While these systems had the virtue of being simple and immune to radio frequency interference (RFI), the relay logic required continuing maintenance in order to keep the contacts clean and adjusted.

Later, the transmitter control logic relays were replaced by solid-state digital logic. This provided much more flexibility and reliability, but also raised concerns about the ruggedness and ability of solid-state logic to survive lightning strikes. Although this concern is valid, modern solid-state logic systems are well protected against damage by optical isolation, shunt protection techniques, and RFI filtering. Operating experience with the current generation of transmitters has proven that a properly designed solid-state control system is far more reliable than its relay predecessors.

Automatic Power Control

Many transmitters also provide automatic power control (APC) circuitry to maintain the transmitter's power output within preset limits by correcting for changes in line voltage, component aging, or small amounts of drift in operating parameters. The APC circuitry compares a sample of the transmitter output power against a reference, and then adjusts the RF drive or other voltages within the transmitter to bring the output power within tolerance.

Some of the more sophisticated APC circuits also provide proportional VSWR foldback of the transmitter output power. If a sample of the reflected power on the transmission line exceeds a safe limit, the transmitter output power is proportionally reduced to a safe level until the problem is resolved. This feature prevents lost air time during antenna icing or other limited VSWR situations. All standard APC circuits should provide fast-acting shutdown of the transmitter during a catastrophic failure of the antenna system such as a short or open circuit. A well designed APC system should operate like a damped analog feedback system with fast response and low overshoot.

Microprocessor Control Systems

Recently, microprocessor technology has been applied to broadcast transmitter control systems. Microprocessor based control systems can expand the functions of the transmitter controller from basic housekeeping duties into powerful self diagnostics, controller redundancy, integrated remote control, user customization, and even self-correction of operational problems.

Microprocessor technology lends itself well to industrial control applications like broadcast transmitters because the hardware can be made just as reliable as hard-wired digital logic, but has the advantage of allowing changes and growth in the operational features by simple changes in software instructions without requiring a complete redesign of the hardware.

Several transmitter manufacturers are marketing transmitters with microprocessor based control systems. Some of the typical features that distinguish these control systems from nonmicroprocessor systems are:

1. Built-in trouble tree with fault location and diagnostic read-outs often using plain English messages (user-friendly).
2. Simultaneous read-outs of all operating parameters.
3. Real-time calculation of efficiency, dissipation, VSWR, and other parameters requiring mathematical operations.
4. Built-in clock/calendar for logging changes in operating status, power failures, and overloads.
5. Tolerance flagging on key operating parameters as warnings for logging and for preventive maintenance.
6. The ability to communicate with the outside world for remote control or logging purposes through a standard serial interface.
7. Integrated remote control capability without external remote control equipment and interfacing.
8. Provision to customize the system features to the individual requirements of the station through the use of software menus.
9. Tuning aids which will allow the operator to adjust the system for peak efficiency, minimum dissipation, and minimum VSWR by means of a real time display of these calculated parameters.
10. The method of communicating information to the operator varies from one system to another, but most use LED or LCD read-outs with codes or alphanumeric messages. Some microprocessor controllers use CRT displays so that a large amount of information can be displayed in several different formats.

Controller Back-Up Systems

A certain degree of redundancy is desirable in the transmitter control system so that the transmitter can stay on-the-air even if a portion of the system fails. There are several approaches being used in present equipment to provide back-up systems. A multi-level hierarchy can be used which automatically hands over basic control functions to a primary controller in the event of a problem in the microprocessor hardware or software. Good system design separates diagnostic and supervisory functions from basic control functions so that a failure in a higher level function will not affect the ability of the system to remain on-the-air without interruption. Watchdog circuits and software are embedded within the control system to detect failures and initiate corrective action before an interruption in service occurs.

It is also possible to have two entirely independent microprocessor systems operating in parallel with only one system in charge at one time. Sophisticated software is required to determine if the two systems agree and, if they don't agree, an algorithm must be used to decide which one is right so that control can be delegated to the properly functioning system.

The ability to quickly substitute another replacement control system or a simplified controller bypass unit is useful in the event of total failure of the control system.

Remote Control Interfacing

Regardless of the type of control system used, the ability to interface easily with standard remote control systems is very important. Most modern systems interface with parallel control lines for each individual function requiring a momentary contact closure of 24 volts or less at a current of 50 milliamperes or less. These levels are compatible with relay logic or optically isolated solid state logic. Analog levels output from the transmitter for remote meter readings generally are fully buffered and fall into the range of 0–10 volts DC for a full scale reading at an impedance level of less than 10K ohms.

With the advent of microprocessor-based control systems, there is a trend toward using standard computer asynchronous serial interfacing instead of parallel interfacing. Serial interfacing reduces the number of connections to the transmitter and can carry both control functions and digitized meter readings through the same interface. By converting analog information into digital information before transmission to the remote control point, the need for calibration and recalibration of the remote metering point is reduced. Microprocessor based control systems also allow the remote control system functions to be integrated directly into the transmitter itself, giving the remote control point access to more in-depth information about the transmitter than is possible by interfacing with an external remote control system. A low-cost personal computer or laptop computer can be used to control the transmitter through an ordinary dial-up phone line or through a radio link.

For more specific information, see the chapter on "Transmitter Remote Control Systems."

RF OUTPUT LOW-PASS FILTERS

The high efficiency, nonlinear RF power amplifiers used in FM broadcast transmitters generate significant amounts of energy on frequencies that are integer multiples (harmonics) of the desired fundamental frequency. The output circuit alone does not provide enough harmonic attenuation to meet FCC regulations. To comply with Part 73 of the FCC Rules and Regulations, and to prevent interference to other services, a low-pass filter must be installed in the transmission line at the output of the transmitter. The FM band is narrow enough that one low-pass filter design can be used for any FM channel carrier frequency. These filters usually consist of multiple LC sections arranged so that frequencies within the FM band are passed with little attenuation (typically 0.1 dB or less) while frequencies above the FM band are highly attenuated (60 dB or more).

The most common type of filter used in this application is called a "reflective" filter, meaning that the frequency components outside the passband are reflected back out of the filter toward the source because it provides a mismatch at these undesired frequencies. The filter can be constructed using either "lumped" inductors and capacitors, or by using a section of nonconstant impedance transmission line to form *distributed* inductors and capacitors. The filters designed for low power transmitters often employ *lumped* elements (coils and capacitors), because they are compact and can be integrated into the transmitter cabinet. The distributed type of filter is most often used with high power FM broadcast transmitters because of its simplicity, extreme ruggedness, and ability to handle higher power levels. The distributed filter does have the disadvantage of having larger physical dimensions than a similar lumped filter, which may necessitate mounting the filter external to the transmitter cabinet. Fig. 28 shows a cut-away view of a typical distributed low-pass filter. Note that the areas where the center conductor of the transmission line is smaller than that required for the input Z_o are inductive, while the areas where the center conductor is larger in diameter are capacitive.

When two filters (such as the output cavity and the harmonic filter) are connected together by a transmission line, the total harmonic attenuation will vary with interconnecting line length. The attenuation characteristics of the harmonic filter are specified for the condition where both the source and load impedances are equal to the desired transmission line impedance.

In actual use, the source impedance at the output of the tank circuit is much less than the 50 ohm load impedance presented by a properly terminated filter. If an unfortunate length of line is selected, the harmonic attenuation may be insufficient and the transmitter tuning may be affected. This undesirable condition can be corrected by changing the line length by approximately one-quarter wavelength. The line length between the tank circuit and harmonic filter is usually supplied pre-cut to a value known to be satisfactory by the transmitter manufacturer.

Harmonic Notch Filters

In some cases, a second harmonic notch filter is required in addition to the low-pass filter because the second harmonic component from amplifier is high and the cut-off slope of the low-pass filter is not steep enough to provide sufficient second harmonic attenuation. The additional attenuation required (typically 30 dB) can be provided by an absorptive notch filter, which places a short circuit across the transmission line at the second harmonic while providing a high impedance at the fundamental. A one-quarter wavelength (at the fundamental frequency) shorted coaxial stub is often used for this function. The second harmonic energy is dissipated in the equivalent series resistance of the series tuned circuit formed by the stub. Some types of transmitters, including those using the folded half-wavelength cavity, have internal second harmonic mode suppressors which eliminate the need for an external notch filter.

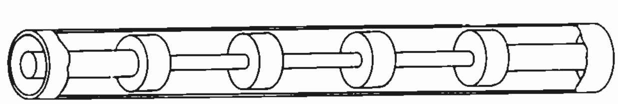


Figure 28. Cut-away view of a distributed low-pass filter.

TRANSMISSION LINE POWER AND SWR MEASUREMENTS

Directional wattmeters are instruments that measure the forward power (P_f) and reflected power (P_r) in a transmission line. The net power delivered to the load (antenna) is ($P_f - P_r$). If the transmission line is perfectly matched all the forward power will be absorbed by the load and there will be no reflected power. The peak voltage at each point along the line will be the same value and similarly the current at each point along the line will also have a uniform value. If the transmission line is mismatched, there will be reflected power with a resulting "standing wave" on the line. This means that the voltage and current distributions along the line will no longer be uniform with high values at certain points on the line and low values at points on the line that are one quarter wavelength away. The ratio of high value to the low value is called the standing wave ratio (SWR).

Standing Wave Ratio

The standing wave ratio (SWR) on the transmission line is related to the ratio of the forward to reflected power by the following formula:

$$SWR = \frac{1 + \sqrt{(P_r/P_f)}}{1 - \sqrt{(P_r/P_f)}} \qquad (12)$$

This relationship is shown graphically in Fig. 29 so the approximate SWR can be obtained without computation.

The standing wave is due to the presence of two components of power: one traveling toward the load; and the other reflected by the load mismatch, traveling back toward the generator.

These components are defined as:

$$P_f = \frac{E_f^2}{Z_o} = I_f^2(Z_o) \qquad (13)$$

$$P_r = \frac{E_r^2}{Z_o} = I_r^2(Z_o) \qquad (14)$$

$$P_n = (P_f - P_r) \qquad (15)$$

The subscripts (f) and (r) are used to denote the forward and reflected values of power, voltage, and current; while Z_o is the characteristic impedance of the transmission line. (P_n) is the net power absorbed by

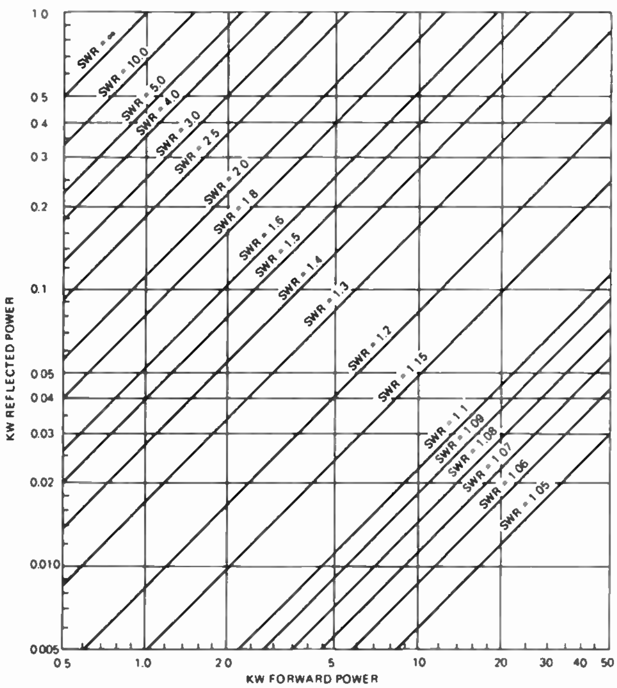


Figure 29. Chart of SWR versus forward and reflected power.

the load (transmission line loss and antenna radiation).

Since the forward and reflected voltage and currents are traveling in opposite directions, they will add in phase at some point along the line length to produce a voltage maximum. One-quarter wavelength along the line in either direction from this maximum, the forward and reflected components are out of phase and produce a voltage minimum. The forward and reflected components of current also add vectorially to produce a current standing wave. The magnitude of the standing wave is defined as:

$$SWR = \frac{E_{max}}{E_{min}} = \frac{I_{max}}{I_{min}} \qquad (16)$$

The ratio of the highest voltage point on the line to the lowest voltage point on the line is a commonly used measure of system performance defined as the voltage standing wave ratio (VSWR). A VSWR of 1.0:1.0 means that a perfect match has been achieved while a VSWR of 2.0:1.0 means that a mismatch is causing approximately 11% of the power to be reflected.

At the point of reflection (the load mismatch), the phase of the reflected current is reversed 180° from the forward current. The reflected voltage does not have this phase reversal. This displaces the voltage and current standing waves by 90° along the line so that the (E_{max}) and (I_{min}) occur at the same points, while (E_{min}) and (I_{max}) occur 90° (one-quarter wavelength) away, in either direction from (E_{max}) and (I_{min}).

The fact that the reflected current is reversed in phase makes it possible to measure forward and reflected power separately using a device called a directional coupler. A small voltage is obtained by inductive coupling which represents the current in the transmission line. To this is added a sample of the voltage across the line that is simultaneously obtained by capacitive coupling. These two samples are adjusted to be exactly equal when the line is terminated with its characteristic impedance (no standing waves and no reflected components). The two RF samples are added, which gives a resultant RF voltage proportional to the forward components of voltage and current as illustrated in Fig. 30A.

The forward components of the samples are equal and in phase, but the reflected components of voltage and current balance out. By having a second coupling section physically turned around in the opposite direction, the phase of the current sample is reversed and the reflected components add while the forward components balance out as illustrated by the vector diagram in Fig. 30B. These voltages output by the directional coupler, representing the forward and reflected powers, are usually rectified and buffered to feed the automatic power controller and a power indicating meter. Since power is proportional to the square of the voltage on the transmission line, the meter scale is calibrated to read the square of its input along with a diode correction factor so that forward and reflected power are read out directly.

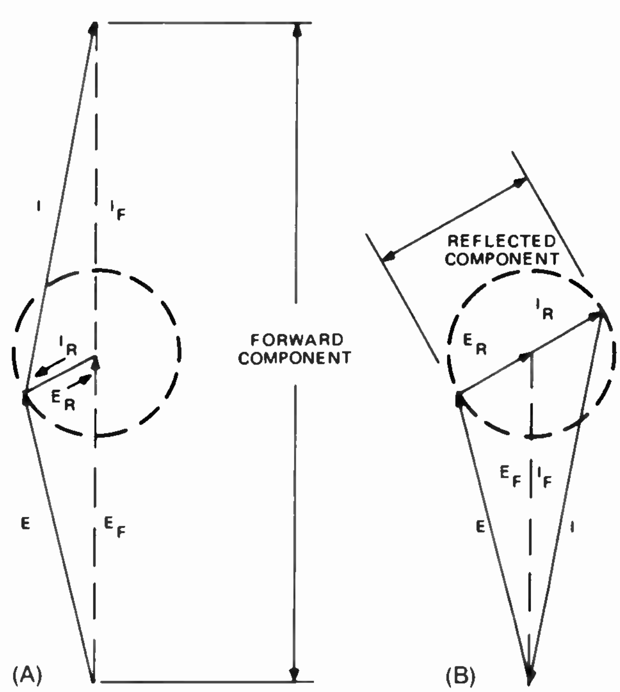


Figure 30. Phasor addition of voltage and current samples to separate forward and reflected components.

VSWR Measurement

Although some FM transmitters can operate into a VSWR of greater than 1.8:1.0, the VSWR on an FM antenna transmission line should normally be kept to a value of 1.1:1.0 or less for good stereo performance. It takes very little reflected power to produce substantial VSWR as shown in Fig. 29. For this reason, the reflected power is usually read on a more sensitive meter position. Problems in the antenna system such as loose connections or icing may cause excessive VSWR. Instruments external to the transmitter are available that monitor reflected power and energize an alarm if it becomes excessive. As long as the transmitter power output is fairly constant, the use of reflected power to indicate excessive VSWR is simple and adequate.

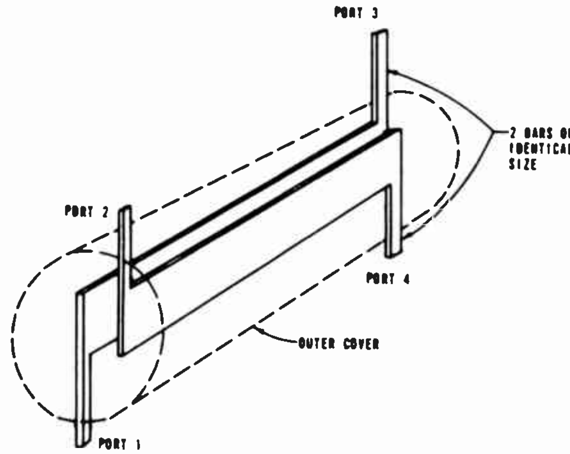


Figure 31. Physical model of 90° hybrid coupler.

COMBINED TRANSMITTERS

It is possible to combine the output of two RF power amplifiers for higher power levels. The important advantage is that the broadcast transmission is not interrupted if one amplifier fails. The radiated signal strength merely drops 6 dB until the failed amplifier is repaired and put back on the air. A dual amplifier system costs more than a single amplifier for a given total power output, but there are the economic advantages of reducing lost air time and eliminating the need for a separate standby transmitter. Automatic or manual output switching can be used to route the full power of the remaining amplifier directly to the antenna. Some stations go one step further and also install dual exciters with automatic switching so that if one exciter fails, the other unit is quickly switched into service.

Hybrid Couplers

Hybrid couplers are reciprocal, 4-port devices that can be used either for splitting or combining RF sources over a wide frequency range. Fig. 31 shows an exploded view of a typical 3 dB, 90° hybrid coupler. The coupler consists of two identical parallel transmission lines that are coupled over a distance of approximately one-quarter wavelength and are enclosed within a single outer conductor. Ports at the same end of the coupler are in-phase while ports at opposite ends of the coupler are in quadrature (90° phase shift) with respect to each other.

The phase-shift between the two inputs or outputs is always 90° and is almost independent of frequency. If the coupler is being used to combine two signals into one output, these two signals must be fed to the hybrid coupler in phase quadrature. The reason this type of coupler is also called a 3 dB coupler is that when used as a power splitter, the split is equal or half-power (3 dB) between the two outputs.

Hybrid Combiners

The output hybrid combiner effectively isolates the two amplifiers from each other. Tuning adjustments can be made on one amplifier including turning it on and off without appreciably affecting the operation of the other amplifier. Good isolation is necessary so that if one transmitter fails, the other will continue to operate normally instead of in a mistuned condition.

Two of the ports on the hybrid coupler are the inputs from the power amplifiers. The *sum port* is the antenna output terminal; and the *difference port* goes to a resistive dummy load called the "reject load," since only the rejected power due to imbalance appears here. When the power fed to each of the two inputs is equal in amplitude with a phase difference of 90°, the total power is delivered to the sum port (antenna). Very little of the power appears at the reject load if the phase relationship and power balance is correct. If the phase relationship is reversed between the two amplifiers, all the power is delivered to the reject load, so care must be taken to be sure that the proper one of the two possible 90° phase relationships is used. When all the ports on the hybrid combiner are properly terminated, isolation of 30 dB or more can be achieved between the power amplifiers. For perfect isolation between the amplifiers, the load impedance on the sum and difference ports must be exactly the same. This is approached in practice by providing a 1.0:1.0 VSWR with a resistive 50 ohm load for the termination (reject load) on the difference port and then reducing the VSWR on the antenna transmission line as low as possible by trimming the antenna match. This will keep the input port impedances from changing very much when one amplifier is not operating.

The input ports will present a load to each transmitter with a VSWR that is lower than the VSWR on the output transmission line. This is because part of the reflected power coming into the output port will be directed to the reject load and only a portion will be fed back into the transmitters. Fig. 32 shows the effect of output port VSWR on the input port VSWR, and on the isolation between ports.

If the two inputs from the separate amplifiers are not equal in amplitude or exactly in phase quadrature,

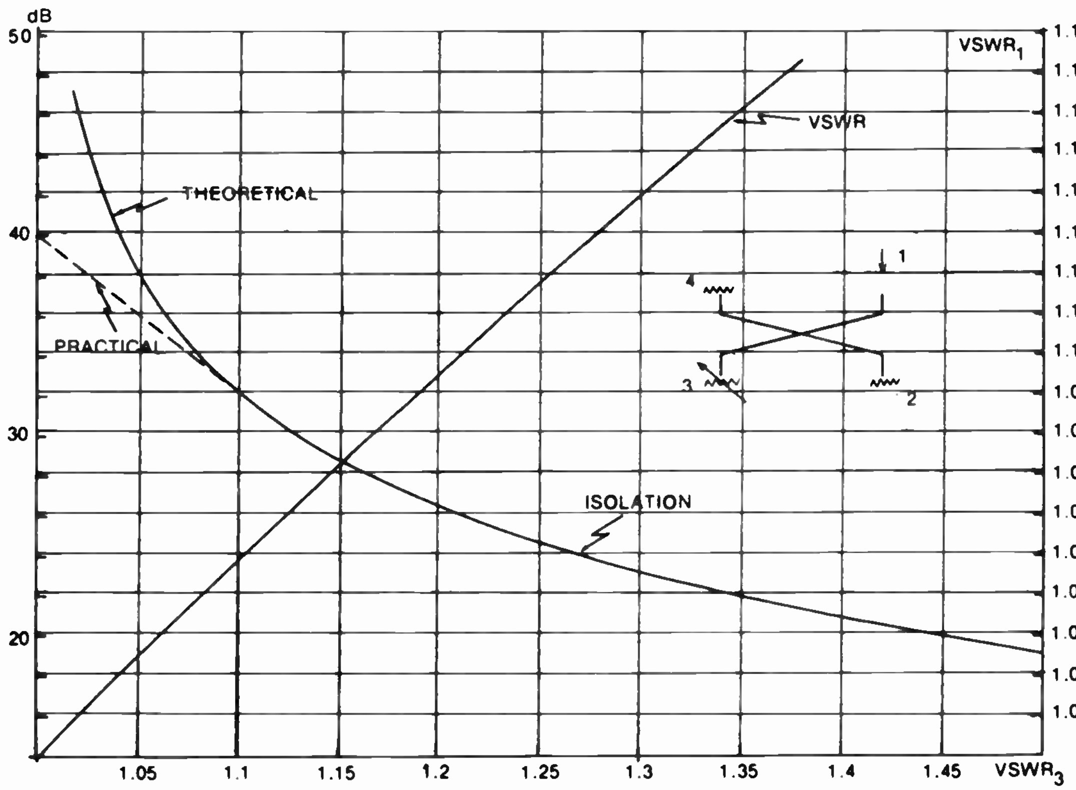


Figure 32. Isolation of hybrid coupler.

some of the power will be dissipated in the difference port reject load. The match in input power and phase is not extremely critical as shown in Fig. 33 and Fig. 34. The power lost in the difference port reject load can be easily reduced to a negligible value by touching up the amplifier tuning and by adjusting the phase shift. For example, if one amplifier is delivering only half the power of the other amplifier, only about 3% of the total available power will be dissipated in the reject load and 97% is still fed to the output transmission line.

If one transmitter fails completely, half of the working amplifier's output goes to the antenna and the other half is dissipated in the difference port reject load. This is why the radiated output drops by 6 dB or to one-fourth of the original combined power. The reject load must be rated to handle a minimum of one-fourth of the total combined power, but often the reject load is rated to handle one-half the total power so that it can also be used as a test load for one of the transmitters.

Hybrid Splitting of Exciter Power

Fig. 35 shows a block diagram of a pair of combined amplifiers with dual exciters. The exciters cannot be

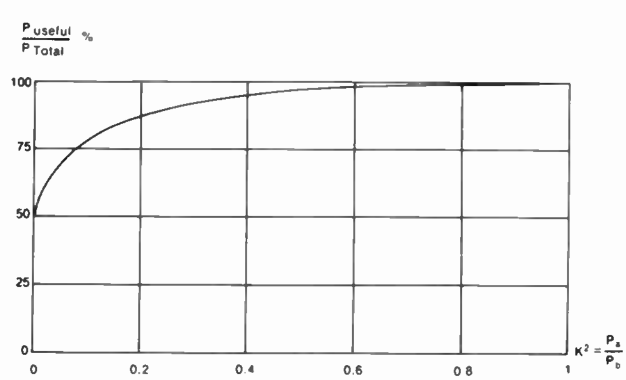


Figure 33. Power imbalance in hybrid couplers.

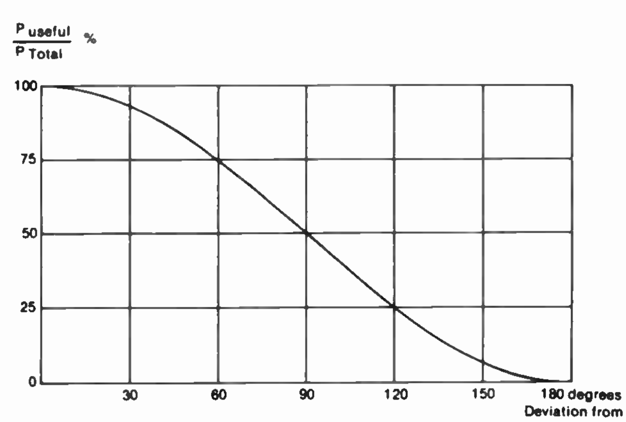


Figure 34. Phase sensitivity, hybrid coupler.

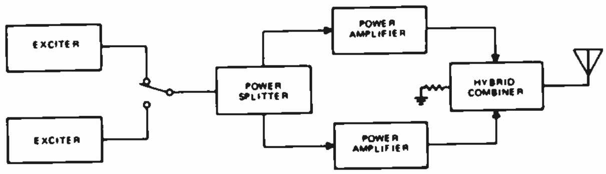


Figure 35. Block diagram of transmitter with two power amplifiers, a hybrid combiner and dual exciters.

operated in parallel like the amplifiers because their RF outputs would have to be on exactly the same carrier frequency and almost exactly in phase under all modulation conditions. An automatic or manual exciter switcher is used to direct the output of the desired exciter to the combined transmitter while the other stand-by exciter is routed to a dummy load. The one exciter in use, feeds a hybrid splitter/phase shifter which transforms one 50 ohm input into two isolated 50 ohm outputs that have a 90° phase shift between them with half the power going to each output. The operation of this hybrid splitter is the reciprocal of the hybrid combiner described above. The exciter must have enough power output capability to drive both power amplifiers. In some cases an additional IPA is required between the exciter and the splitter to boost the drive level. The length of coax from the power splitter to each amplifier input must be cut to a precise length so that the amplifiers will be fed in the proper phase relationship.

Each of the power amplifiers is assumed to have equal gain and phase shift. In practice, it may be difficult to get the amplifiers tuned so that their gains and phase shifts are equal at the same time. For this reason, a line stretcher or variable phase shift network is usually included with the exciter splitter so that the station engineer will have the ability to adjust phasing independent of amplifier tuning.

For more detailed information on the theory of hybrid couplers, see the chapter on television transmitters or Ref. 7 at the end of this chapter.

FILTERPLEXING

The practice of having several FM stations share a single broadband antenna system is becoming more popular in recent years. In order to connect several transmitters on different frequencies together onto one antenna system, a special device called a filterplexer is required. The purpose of the filterplexer is to provide isolation between the various transmitters while efficiently combining their power into a single transmission line. This is usually accomplished by a system of band-pass filters, band-reject filters, and hybrid combiners. The isolation is required to prevent power from one transmitter from entering another transmitter with resulting spurious emissions, as well as to keep the rest of the system running in the event of the failure of one or more transmitters.

An important consideration in the design of a filterplexing system is the effect on the phase response (group delay characteristic in the passband) of each of the signals passing through the system due to individual bandwidth limitations on each of the inputs.

RF Intermodulation Between FM Broadcast Transmitters

Interference to other stations within the FM broadcast band, as well as to other services outside the broadcast band, can be caused by RF intermodulation between two or more FM broadcast transmitters. Transmitter manufacturers have begun to characterize the susceptibility of their equipment to RF intermodulation, so this information will be available to the designers of filterplexing equipment.

The degree of intermodulation interference generated within a given system can be accurately predicted before the system is built if the actual mixing loss of the transmitters is available when the system is designed. Accurate data on *mixing loss* or *turn-around-loss* not only speeds the design of filterplexing equipment; but also results in higher performance and more cost effective designs, because the exact degree of isolation required is known before the system is designed. Filterplexer characteristics, as well as antenna isolation requirements, can be tailored to the specific requirements of the transmitters being used. The end user is assured in advance of construction that the system will perform to specification without fear of overdesign or underdesign of the components within the system.

Mechanisms Which Generate RF Intermodulation Products

When two or more transmitters are coupled to each other, new spectral components are produced by mixing of the fundamental and harmonic terms of each of the desired output frequencies. For example, if only two transmitters are involved, the third-order intermodulation (IM_3) terms could be generated in the following way. The output of the first transmitter (f_1) is coupled into the nonlinear output stage of the second transmitter (f_2), because there is not complete isolation between the two output stages. (f_1) will mix with the second harmonic of (f_2) producing an in-band third-order term with a frequency of $[2(f_2) - (f_1)]$. In a similar fashion, the other third-order term will be produced at a frequency of $[2(f_1) - (f_2)]$. This implies that the second harmonic content within each transmitter's output stage, along with the specific nonlinear characteristics of the output stage, will have an effect on the value of the mixing loss.

It is possible, however, to generate these same third-order terms in another way. If the difference frequency between the two transmitters $[(f_2) - (f_1)]$, which is an out-of-band frequency, remixes with either (f_1) or (f_2), the same third-order intermodulation frequencies are produced.

$$f_1 = 100.3 \text{ MHz} \qquad f_2 = 101.1 \text{ MHz}$$
$$2f_1 - f_2 = [2(100.3) - (101.1)] = [200.6 - 101.1] = 99.5 \text{ MHz}$$
$$2f_2 - f_1 = [2(101.1) - (100.3)] = [202.2 - 100.3] = 101.9 \text{ MHz}$$
OR
$$[f_1 - (f_2 - f_1)] = [100.3 - (101.1 - 100.3)] = [100.3 - 0.8] = 99.5 \text{ MHz}$$
$$[f_2 + (f_2 - f_1)] = [101.1 + (101.1 - 100.3)] = [101.1 + 0.8] = 101.9 \text{ MHz}$$

Figure 36. Third order intermodulation products.

Empirical measurements indicate that the $[2(f_2) - (f_1)]$ type of mechanism is the dominant mode generating third-order IM products in modern transmitters using a tuned cavity for the output network.

Fig. 36 shows an example of how the intermodulation product frequencies may be calculated. Fig. 37 and Fig. 38 show the resulting frequency spectra.

Intermodulation as a Function of Turn-Around-Loss

Turn-around-loss or *mixing loss* describes the phenomenon whereby the interfering signal mixes with the fundamental and its harmonics within the nonlinear output device. This mixing occurs with a net conversion loss, hence the term turn-around-loss has become widely used to quantify the ratio of the interfering level to the resulting IM_3 level. A turn-around-loss of 10 dB means that the IM_3 product fed back to the antenna system will be 10 dB below the interfering signal fed into the transmitter's output stage.

Turn-around-loss will increase if the interfering signal falls outside the passband of the transmitter's output circuit, varying with the frequency separation of the desired signal and the interfering signal. This is

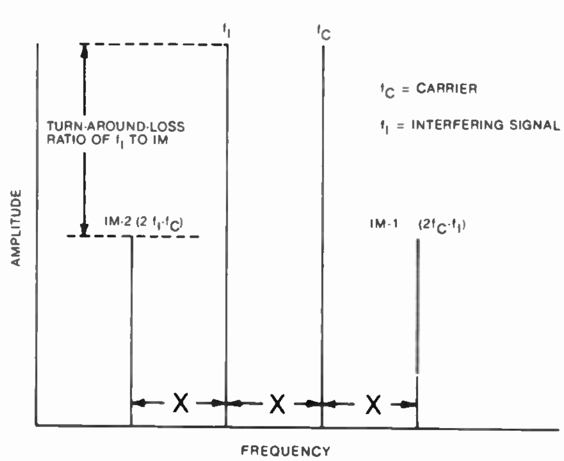


Figure 37. Frequency spectrum of third order IM with the interfering level equal to the carrier level.

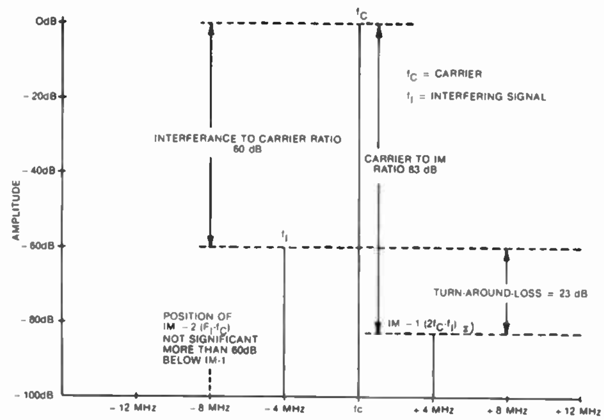


Figure 38. Typical frequency spectrum of third order IM of a broadcast FM transmitter.

because the interfering signal is first attenuated by the selectivity going into the nonlinear device, and then the IM_3 product is further attenuated as it comes back out through the frequency selective circuit.

Turn-around-loss can actually be broken down into three individual parts:

1. The basic in-band conversion loss of the nonlinear device.
2. The attenuation of the out-of-band interfering signal due to the selectivity of the output stage.
3. The attenuation of the resulting out-of-band IM_3 products due to the selectivity of the output stage.

As the turn-around-loss increases, the level of undesirable intermodulation products is reduced and the amount of isolation required between transmitters is also reduced.

The transmitter output circuit loading control directly affects the source impedance, and therefore affects the efficiency of coupling the interfering signal into the output circuit where it mixes with the other frequencies present to produce IM_3 products. Light loading reduces the amount of interference that enters the output circuit with a resulting increase in turn-around-loss. In addition, the output loading control setting will change the output circuit bandwidth (loaded "Q"), and therefore also affect the amount of attenuation that out-of-band signals will encounter passing into and out of the output circuit.

Second harmonic traps or low-pass filters in the transmission line of either transmitter have little effect on the generation of intermodulation products. This is because the harmonic content of the interfering signal entering the output circuit of the transmitter has much less effect on IM_3 generation than the harmonic content within the nonlinear device itself. The resulting IM_3 products fall within the passband of the low-pass filters and outside the reject band of the second harmonic traps, so these devices offer no attenuation to RF intermodulation products.

Special Requirements of Solid-State Transmitters

Wideband, solid-state, transmitters pose some special problems for filterplexing. Solid-state, nonlinear, power amplifiers are typically very efficient mixing devices and, because there is no output selectivity to attenuate out-of-band interference or out-of-band RF intermodulation products, the turn-around-loss can be as little as 6 dB regardless of frequency separation. Additional external bandpass filtering is usually required at the output of each solid-state transmitter to reduce RF intermodulation to acceptable levels.

For more detailed information about RF intermodulation between transmitters, see the chapters on FM and TV antenna systems and Ref. 14 at the end of this chapter.

OPERATIONAL MEASUREMENTS

Certain parameters are considered by the FCC to be important enough to justify almost continuous observation. Especially important are the modulation level, carrier frequency, and output power level. The FCC requires that some of the technical operating parameters be entered at regular intervals in the operation and maintenance logs. These logs must be kept available for inspection a minimum of two years from the date of entry. It is always important to consult the current revision of Part 73 of the FCC Rules and Regulations because they may have changed since this chapter was printed.

FM Modulation Measurement

The measurement of FM modulation can be accomplished with a broadcast type modulation monitor or with one of the newer modulation analyzers. Some FM exciters have built-in peak modulation displays for convenience in set-up and adjustment. Once the initial levels are correctly set, modern audio processing equipment will usually hold the modulation levels within the desired window.

Much interest and concern has developed throughout the broadcasting industry as to the best method for the determination of modulation percentage for complex program material. The inability of mechanical meter movements to follow short-duration, nonrepetitive peaks accurately has received special attention. Standard modulation meter movements cannot follow modulation peaks with the required accuracy. For this reason, modulation monitors have a peak-indicating device that can be preset to flash at the particular level of interest. This device should be used instead of the meter to determine peak modulation conditions of the transmitter.

The reason for setting a peak deviation limit is so that the related occupied bandwidth does not increase to the point of interfering with stations on adjacent channels. The FCC presently enforces the modulation limit by monitoring the instantaneous peak deviation of the station as displayed on an oscilloscope. This method of measurement does not exactly correlate with the station's occupied bandwidth because the duty-cycle of the modulation peaks is not taken into account. As a result, many sophisticated "peak limiting" and "overshoot control" devices have appeared on the market to maximize loudness without exceeding the peak deviation limit, by removing the low energy peaks that would extend beyond 100% modulation. The use of these devices does cause some degradation of the audio quality and they would no longer be used if the method of modulation measurement were changed to one based on occupied bandwidth. Recently introduced modulation measurement devices ignore short duration overshoots and provide modulation level indications that more accurately reflect the resulting occupied bandwidth. The FCC is considering rule changes that would allow enforcement of FM modulation by occupied bandwidth after the technical details on the exact method of measurement are worked out.

Carrier Frequency Measurement

The average carrier frequency must be measured and maintained to within ±2 kHz of the assigned channel with an accurate frequency monitor. These monitors fall into two categories; analog display of the frequency error from the nominal carrier frequency and digital display of the absolute carrier frequency. The trend is toward the digital counter because of its inherent accuracy and ease of use.

The requirement that a station utilize type-approved modulation and frequency monitors has been eliminated by the FCC. Each station is still required to maintain its frequency, modulation, and audio performance within the FCC limits defined in Part 73 of the Rules and Regulations, but the responsibility for selecting the method of measurement and type of measuring equipment is now up to the station operator. Every quality conscious station will want to have the necessary equipment to accurately evaluate the signal being broadcast. There are new options available in high performance yet general purpose test equipment, now that monitor type-approval is no longer required. For instance, modern modulation analyzers provide frequency agility as well as greater measurement capability than the more specialized modulation monitors. General purpose frequency counters are now available with sufficient accuracy to measure the carrier, subcarrier, and stereo pilot frequencies directly. Spectrum analyzers provide a wide range of capability including the measurement of harmonic and spurious frequencies at the carrier frequency, composite baseband, Bessel nulls, occupied bandwidth, stereophonic and SCA crosstalk, and synchronous AM.

Measurement of RF Power Output

The methods for determining RF output power are specified in Part 73 of the FCC Rules and Regulations. An accurately calibrated directional wattmeter provides an excellent means of making a direct measurement of RF output power. The directional wattmeter is seldom used as the primary RF power determining method because of the requirement for recalibration to a traceable standard at regular intervals. Use of the indirect method of power measurement avoids this requirement. The FCC is considering a change in the Rules that would permit the use of the transmitter power output meter directly, if it is periodically calibrated by comparison with the indirect method, instead of with a dummy load and standard wattmeter.

Using the indirect method, the output power is calculated from a measurement of the DC input power multiplied by the efficiency factor of the final amplifier stage. The efficiency factor is provided by the transmitter manufacturer on the final test data sheet or in the instruction manual and must be applicable to the particular frequency and power level in use. The power input to the final amplifier stage is normally defined as the product of plate voltage and plate current to this stage. Multiple output stages which are combined for the total power, must have their individual DC power inputs arithmetically summed to obtain the total power input.

The directional wattmeter can be used as a check when compared to the power output calculated by the indirect method to determine if the efficiency factor has changed due to incorrect tuning, changing antenna conditions, or a weak output device.

Logging of Transmitter Operating Parameters

Although the FCC has deleted specific logging requirements, FM stations may still wish to log certain transmitter parameters at three-hour intervals, in order to track equipment performance trends. The minimum suggested entries are:

- Final amplifier plate voltage
- Final amplifier plate current
- RF transmission line current, voltage, or power

The carrier frequency, stereo pilot frequency, and SCA subcarrier frequency require less frequent measurement and logging. Microprocessor based control systems often provide automatic logging of all transmitter parameters; including overloads and tolerance flagging.

Proof-of-Performance Measurements

While the FCC no longer requires proof of performance tests, it is advisable to conduct such tests periodically to ensure the equipment is operating properly. These tests should include:

- Audio frequency response
- Audio frequency harmonic distortion
- FM signal-to-noise ratio
- AM noise level

Fig. 39 is a block diagram of a typical test set-up for proof-of-performance measurements.

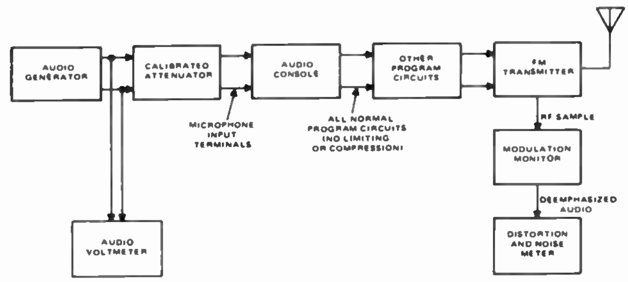


Figure 39. Block diagram of test set-up for audio performance measurements.

Measurement of Audio Frequency Response

The audio frequency response of the system is measured in reverse, that is, a constant percentage of modulation is maintained for all modulating frequencies by adjusting the amount of attenuation between the audio generator and the microphone input terminals. This is necessary because of the rising response due to preemphasis. Frequency response data is taken at three levels of modulation: 25%, 50%, and 100%. The audio voltmeter which measures the audio generator output voltage is used to maintain a constant voltage level versus frequency at the generator output terminals. The precision attenuator dials are adjusted for each modulating frequency to maintain the desired modulating level and the attenuator readings, in decibels, are recorded. Readings should be taken for the following modulating frequencies: 50, 100, 400, 1000, 5000, 7500, 10000, and 15000 Hz. Measurements at 7500 Hz are not required by the FCC Rules and Regulations, but are helpful when the data is plotted within the preemphasis curve limits.

When the attenuation, in decibels, is plotted versus frequency, the 75 microsecond preemphasis curve is obtained if the system frequency response is perfect. Deviations from the ideal response are permitted to an extent which allows the measured curve to be fitted between the upper and lower limit curves shown in Fig. 8. The procedure for doing this is to offset the measured curve by subtracting or adding the same number of decibels from each of the measured values. This process may be repeated until a fit is obtained. If it is impossible to obtain a fit within the limits by subtracting or adding the same value of attenuation from all measured values, the system frequency response is inadequate and corrections must be made.

Measurement of Audio Harmonic Distortion

Total harmonic distortion (THD) of the system from microphone input terminals to transmitter output is measured by modulating the transmitter with sinusoidal modulating signals having low distortion, and observing the harmonic content at the output of the modulation monitor. For this measurement, preemphasis is used in the transmitter and deemphasis is used in the monitor. The distortion analyzer must respond to de-emphasized harmonics through 30 kHz.

The type of distortion meter normally used in this test reads not only harmonic distortion but also noise in the audio passband. For this reason, THD measurements above 5 kHz may be noise limited due to the effect of deemphasis. A more accurate method of distortion measurement in the presence of noise is to use an audio frequency spectrum analyzer to determine the total root mean square (RMS) value of the individual distortion products.

The THD is often measured under the following conditions:

- For modulating frequencies of 50, 100, 400, 1000 and 5000 Hz at modulating levels of 25, 50, and 100% modulation
- For modulating frequencies of 10 kHz and 15 kHz at a modulating level of 100%

The preceding measurements must show that the system THD is less than 3.5% for modulating frequencies between 50 and 100 Hz, less than 2.5% for modulating frequencies between 100 and 7500 Hz, and less than 3.0% for modulating frequencies between 7500 and 15000 Hz. Most modern equipment will pass these tests by a wide margin. If distortion levels greater than these are measured, the system requires adjustment or repair. Tuning the transmitter for minimum even order harmonic distortion will result in a symmetrical group delay response and optimum FM modulation performance. This can be accomplished either by observing the even order harmonics in the demodulated baseband, or by placing an audio bandpass filter tuned to the second harmonic or other even order harmonic of the modulating frequency ahead of the audio distortion analyzer.

Measurement of Two-Tone Intermodulation Distortion

Audio intermodulation distortion (IMD) measurements are a quick and accurate way to evaluate the system performance before a complete set of THD measurements are made. If the system will pass a single 60 Hz/7 kHz, 4:1 (SMPTE-IMD) measurement at 100% modulation, it will probably pass all of the THD measurements up to 10 kHz. (This dual frequency test signal (SMPTE-IMD) is on track 30 of the *NAB Broadcast and Audio System Test CD*.)

Other types of difference-tone, swept two-tone, and sinewave/squarewave IMD measurements can reveal more subtle problems in the system such as transient IMD due to insufficient audio feedback bandwidth in the audio amplifier stages.

Measurement of FM Signal-To-Noise Ratio

The FM signal-to-noise ratio of the system is also measured from the microphone input terminals to the transmitter output. The residual noise level at the monitor output is measured with an audio voltmeter. For this measurement, preemphasis is employed in the transmitter and deemphasis in the monitor. The residual audio noise level is referenced to the signal level produced by 400 Hz (L + R, mono) modulation at the 100% level (75 kHz deviation).

The procedure for making the FM signal-to-noise ratio measurement is as follows:

1. Modulate the transmitter with a 400 Hz sine wave applied at the microphone input terminals of the console and set the level for 100% modulation.
2. Read and record the audio signal level appearing at the modulation monitor output terminals. If the monitor has audio metering capability, the meter gain should be set for a 0.0 dB reference level according to the manufacturer's instructions.
3. Remove the modulation and terminate the console audio input terminals with a resistor equal to the normal microphone output impedance.
4. Read and record the residual audio noise voltage in decibels below the 400 Hz reference signal level. The measured signal-to-noise ratio must be at least 60 dB. State of the art equipment can provide a signal-to-noise ratio of better than 90 dB, so it is relatively easy to meet the FCC requirement of 60 dB.

Measurement of AM Signal-To-Noise Ratios

The perfect FM transmitter will have an absolutely constant output, regardless of FM modulation or power supply variations. In practice, there is always some residual amplitude modulation of the FM transmitter. There are two types of AM signal-to-noise ratios that are of interest to the FM broadcast engineer:

1. *Asynchronous AM signal-to-noise ratio* measured without FM modulation is required by the FCC Rules and Regulations, and is primarily related to power supply ripple.
2. *Synchronous AM signal-to-noise ratio*, or ICAM (incidental carrier AM), measured with FM modulation is *not* required by the FCC Rules and Regulations and *is* related to the tuning and overall bandwidth of the system.

Asynchronous AM

Residual amplitude modulation of the transmitter output without FM modulation, due primarily to power supply ripple, is measured with an AM envelope detector. Most FM modulation monitors include an AM detector for this purpose. The detector must include 75 microsecond deemphasis of its output. AM noise measurements must be made directly at the transmitter output (or an accurate sample of its output). No amplifying or limiting equipment may be used between the transmitter output and the AM detector, since this equipment would modify the residual AM noise level present. The FCC Rules and Regulations require residual AM noise to be 50 dB below the level which would represent 100% amplitude modulation of the carrier. Since the transmitter cannot be amplitude modulated, this reference must be established indirectly by a measurement of the RF carrier voltage. Refer to the instructions of the detector manufacturer to determine the reference level. Generally, the reference level is determined by setting a carrier level meter to a specified reading. If the transmitter is unable to meet the 50 dB requirement, the problem can usually be traced to a power supply component or to line imbalance in a three-phase system.

Synchronous AM

Synchronous AM is a measure of the amount of incidental amplitude modulation introduced onto the carrier by the presence of FM modulation. Although this measurement is not required by the FCC Rules and Regulations, it provides information about the amplitude response and tuning of the transmitter. Measurement of synchronous AM also gives the station engineer an idea of the overall system bandwidth and whether the passband is positioned correctly. Since all transmitters have limited bandwidth, there will be a slight drop-off in power output as the carrier frequency is swept to either side of the center frequency. This slight change in RF output level follows the waveform of the signal being applied to the FM modulator, causing AM modulation in synchronization with the FM modulation. The concept is similar to the slope detection of FM by an AM detector used in conjunction with a tuned circuit.

Synchronous AM measurements are made directly at the transmitter output (or an accurate sample of its output). No amplifying or limiting equipment may be used between the transmitter output and the AM detector since nonlinearities in this equipment could modify the synchronous AM level present. Since the transmitter cannot be fully amplitude modulated, an equivalent reference level must be established indirectly by a measurement of the RF carrier voltage. Refer to the instructions of the detector manufacturer to determine this reference level. Generally, the reference level is determined by setting a carrier level meter to a specified reading, or to obtain a specific DC voltage level at the output of the detector diode without modulation. Care must be taken when making these measurements that the test setup does not introduce synchronous AM and give erroneous readings, which would cause the operator to mistune the transmitter to compensate for errors in the measuring equipment.

The input impedance of the envelope detector must provide a nearly perfect match so that there is a very low VSWR on the sampling line. Any significant VSWR on the sampling line will produce synchronous AM at the detector because the position of the voltage peak caused by the standing wave moves along this line with FM modulation. A thru-line type of directional coupler, normally used to drive the wattmeter movement, has the envelope detector diode built into the sampling element; and provides a DC component that the meter movement responds to, plus the demodulated synchronous AM component that the meter movement does not respond to. If the thru-line element output is fed to an oscilloscope instead of the wattmeter movement, the synchronous AM waveform can be accurately measured. This approach eliminates the errors due to VSWR on the sampling line, since the detector is located at the sampling point. The manufacturer of the thru-line coupler can supply the special connectors and/or cables to connect its output to the oscilloscope. Care must be taken to avoid hum pick-up from AC ground loops while making these low level measurements. Both the thru-line element detector and the precision envelope detectors have some residual RF on their DC output, so an RF filter network may be required between the detector and the input of the oscilloscope.

Most FM demodulators cannot be relied upon to make accurate synchronous AM noise measurements, so it is a good idea to cross-check the demodulator reading directly against the demodulated output of a precision envelope detector. This can be done by first measuring the DC component of the waveform with a voltmeter, or by DC coupling the scope input. The scope is then AC coupled and the input sensitivity is increased until an accurate peak-to-peak measurement of the AC modulation component can be made. The peak-to-peak AC voltage is then divided by twice the DC component to obtain the *voltage ratio*. Twenty times the LOG (base 10) of the voltage ratio is the actual synchronous AM noise level in dB below equivalent 100% AM modulation. Multiplying the voltage ratio by 100 yields the percent of AM modulation. Note that the *peak* detected value of the carrier must be doubled to convert it to the *peak-to-peak* value of the carrier. The ratio of the peak-to-peak modulation component to the peak-to-peak carrier is then used to calculate the percentage of synchronous AM modulation.

Built-In AM Noise Measurement Capability

Some FM transmitters feature a built-in precision envelope detector incorporated into the automatic power control (APC) system. A calibrated front panel AM noise test jack allows observation of the synchronous AM waveforms or direct measurement of the synchronous AM noise level on a standard audio voltmeter.

How Good Should Synchronous AM Be?

Synchronous AM of 35 dB or more below equivalent 100% AM, is considered to be acceptable since the limited bandwidth of the IF filter in the receiver will reintroduce higher levels of synchronous AM to the FM signal before demodulation. Higher levels of synchronous AM can cause increased "chopping" of the signal at the receiver near limiting threshold under weak signal "fringe area" conditions and can exacerbate multipath problems. Excessive synchronous AM is also an indirect indication of passband induced distortion problems that degrade stereo performance and SCA crosstalk.

Many of the older multi-tube transmitter designs presently in use will have as much as 6% (−30 dB)

TABLE 1
Synchronous AM versus bandwith.

SYNCHRONOUS AM (below 100% AM) (±75 kHz @ 1 kHz FM)	APPROXIMATE BANDWIDTH OF TRANSMITTER (−3 dB)	RF LEVEL VARIATION AT RECEIVER LIMITER (%)	(dB)
−30 dB	410 kHz	6.32%	0.57 dB
−35 dB	550 kHz	3.54%	0.31 dB
−40 dB	730 kHz	2.00%	0.18 dB
−45 dB	1.00 MHz	1.12%	0.10 dB
−50 dB	1.34 MHz	0.64%	0.06 dB
−55 dB	1.82 MHz	0.36%	0.03 dB
−60 dB	2.46 MHz	0.20%	0.02 dB

synchronous AM when simply tuned for best power output and efficiency, even though the asynchronous AM (without modulation) may be better than −50 dB. Some of the newer single-tube transmitters can be adjusted for 50 dB or more suppression of synchronous AM. The synchronous AM level of virtually any FM transmitter can be improved by proper tuning techniques. An approximation to the overall system bandwidth can be related to the synchronous AM as shown in Table 1.

Limitations of Synchronous AM Measurements

Synchronous AM measurements are an indirect way of evaluating and optimizing FM performance. Even though synchronous AM measurements are a helpful aid to correctly tune an FM transmitter, these measurements tell only (the amplitude response) half of the total story. Transmitter tuning also affects the group delay (time) response which in turn affects the relative time delays of the higher order FM sidebands. Even though the amplitude response appears flattened when the grid is heavily driven, the group delay (time) response still has a serious effect on the higher order FM sidebands.

Synchronous AM versus Symmetrical Group Delay Response

Recent research using computer simulations,[23] as well as empirical measurements made on FM transmitters, showed that group delay asymmetry results in much more distortion than asymmetrical amplitude response. As long as the group delay response is symmetrical, the amount of synchronous AM will have little effect on the FM modulation performance and distortion. Most FM transmitters will exhibit a significant increase in synchronous AM when tuned for symmetrical group delay response, even though this results in best FM modulation performance. Tuning for minimum synchronous AM is a good starting point, but it is more desirable to finish tuning at the symmetrical group delay point.

Fine tuning the input and output for minimum even order harmonic distortion will optimize the group delay (time) response. Transmitters that utilize wideband solid-state intermediate power amplifiers (IPA) will add less distortion to the FM signal because both the amplitude and group delay (time) response will be better than systems utilizing several tuned stages.

TUNING THE TRANSMITTER FOR BEST MODULATION PERFORMANCE

The modern power amplifiers which have been discussed in the preceding sections, can operate with high reliability and power efficiency without compromising subcarrier performance, if they are properly adjusted. All optimization should be done with any automatic power control (APC) system disabled so that the APC will not chase the adjustment in an attempt to keep the output power constant. The transmitter should be connected to the normal antenna system rather than to a dummy load. This is because the resistance and reactance of the antenna will be different from the dummy load and the optimum tuning point of the transmitter will shift between the two different loads. The tuning sequence is:

Initial Tuning and Loading

The transmitter is first tuned for normal output power and proper efficiency according to the manufacturer's instruction manual. The meter readings should closely agree with those listed on the manufacturer's final test data sheet if the transmitter is being operated at the same frequency and power level into an acceptable load.

Input Tuning and Matching

The input tuning control should first be adjusted for maximum grid current and then fine tuned interactively with the input matching control for minimum reflected power to the driver stage. Note that the point of maximum grid current may not coincide with the minimum reflected power to a solid-state driver. This is because a solid state driver may actually output more power at certain complex load impedances than into a 50 ohm resistive load. The main objective during input tuning is to obtain adequate grid current while providing a good match (minimum reflected power) to the coaxial transmission line from the driver. In the case of an older transmitter with a tube driver integrated into the grid circuit of the final amplifier, the driver plate tuning and the final grid tuning will be combined into one control; which is adjusted for maximum grid current.

Output Tuning

The output tuning control adjusts the resonant frequency of the output circuit to match the carrier frequency. As resonance is reached, the plate current will drop while both the output power and screen current rise together. Under heavily loaded conditions this "dip" in plate current is not very pronounced, so tuning for a peak in screen current is often a more sensitive indicator of resonance.

Amplifiers utilizing a folded halfwave cavity will display little interaction between output tuning and output loading, because the output coupling loop is located at the RF voltage null point on the resonant line. Quarter-wave cavities will require interactive adjustment of output tuning and output loading controls, since changes in loading will also affect the frequency of the resonant line.

Output Loading

There is a delicate balance between screen voltage and output loading for amplifiers utilizing a tetrode tube. Generally there is one combination of screen voltage and output loading where peak efficiency occurs. At a given screen voltage, increasing the amplifier loading will result in a decrease in screen current, while a decrease in loading will result in an increase in screen current. As the screen voltage is increased to get more output power, the loading must also be increased to prevent the screen current from reaching excessive levels. Further increases in screen voltage without increased loading will result in a screen overload without an increase in output power.

Automatic Power Control Headroom

Automatic power control (APC) feedback systems are utilized in many transmitters to regulate the power output around a predetermined setpoint with variations in AC line voltage or changes in other operating parameters. Most modern FM broadcast transmitters utilize a high gain tetrode as the final amplifier stage with adjustment of the screen voltage providing fine adjustment of the output power.

For each power output level there is one unique combination of screen voltage and output loading that will provide peak operating efficiency. If the screen voltage is raised above this point without a corresponding increase in loading, there will be no further increase in power output with rising screen voltage and screen current. If the screen voltage is raised without sufficient loading, a screen current overload will occur before the upward adjustment in power output is obtained. To avoid this problem, it is a good idea to tune the transmitter with slightly heavier loading than necessary to achieve the desired power output level in order to allow for about 5% headroom in adjustment range. The output loading can be adjusted for a peak in output power of 5% over the desired level and then the screen voltage can be reduced enough to return to the desired level. This procedure will allow headroom for an APC system controlling screen voltage and will result in about a 1% compromise in efficiency, but it will assure the ability to increase power output up to 5% without encountering a screen overload.

Centering the Passband

A simple method for centering the transmitter passband on the carrier frequency involves adjustment for minimum synchronous AM. If the bandpass is narrow or skewed, increasing synchronous amplitude modulation of the carrier will result. A typical adjustment

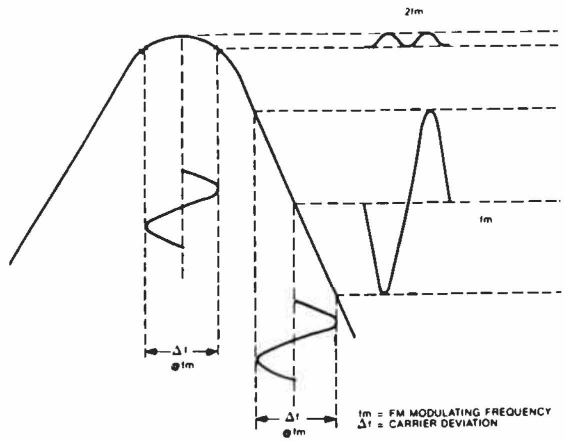


Figure 40. Synchronous AM waveforms.

procedure is to FM modulate 100% at 1 kHz, and fine-adjust the transmitter's grid tuning and output tuning controls for minimum 1 kHz AM modulation as detected by a wideband envelope detector (diode and line probe). 1 kHz is used as the FM modulating frequency rather than 400 Hz so that the audio high-pass filter in the audio analyzer can be used to eliminate the AC line frequency related asynchronous component from the synchronous AM component. It is helpful to display the demodulated output from the AM detector on an oscilloscope while making this adjustment. Note that as the minimum point of synchronous AM is reached, the demodulated output from the AM detector will double in frequency to 2 kHz, because the fall-off in output power is symmetrical about the center frequency causing the amplitude variations to go through two complete cycles for every one FM sweep cycle. This effect is illustrated in Fig. 40. It should be possible to minimize synchronous AM while maintaining output power and efficiency in a properly designed power amplifier.

Effect of Transmitter Tuning on the FM Sidebands

The higher order FM sidebands will be slightly attenuated in amplitude and shifted in time (group delay) as they pass through the final amplifier stage. These alterations in the sideband structure that are introduced by the amplifier passband, result in distortion after FM demodulation at the receiver. The amount of distortion is dependent on the available bandwidth versus the modulation index being transmitted. For a given bandwidth limitation, the distortion can be minimized by centering the passband of the amplifier around the signal being transmitted. This will cause the amplitude and group delay errors to affect both the upper and lower sidebands equally or symmetrically. Tuning an amplifier for minimum plate current or for best efficiency does not necessarily result in a centered passband. One way to center the amplitude passband is to tune the amplifier for minimum synchronous AM modulation while applying FM modulation to the transmitter. Since the circuit topology of most transmitters exhibit a difference in tuning between the symmetrical amplitude response and the symmetrical group delay response, FM modulation performance can be further improved by tuning for symmetrical group delay rather than for minimum synchronous AM. The symmetrical group delay tuning point usually does not coincide exactly with the symmetrical amplitude tuning point, and falls between the point of minimum synchronous AM and the point of maximum efficiency.

The transmitter may be tuned for minimum intermodulation distortion in left-only or right-only stereo transmissions. Stereo separation will also vary with tuning. For stations employing a 67 kHz SCA, transmitter tuning becomes very critical to minimizing crosstalk into the SCA. Modulate one channel only on the stereo generator to 100% with a 4.5 kHz tone. This will place the lower second harmonic (L-R) stereo sideband on top of 67 kHz SCA. Activate the SCA at normal injection level without modulation on the SCA. Tune the transmitter for minimum output from the SCA demodulator. This adjustment can also be made by listening to the residual SCA audio while normal stereo programming is being broadcast.

A more sensitive test is to tune for minimum even order harmonic distortion; which will result in a symmetrical group delay response and will optimize distortion, separation, and crosstalk.

The latest generation of power amplifiers has been designed to operate without compromising subcarrier performance. By providing broadband matching circuits, adjustment of these transmitters for optimum FM modulation performance (minimum distortion, minimum crosstalk, maximum separation, etc.) is very repeatable and stable.

The field adjustment techniques are listed below in ascending order of sensitivity:

1. Tune for minimum synchronous AM noise.
2. Tune for minimum IMD in the left or right channel only.
3. Tune for minimum crosstalk into the unmodulated SCA subcarrier.
4. Tune for minimum even order harmonic distortion (symmetrical group delay).

In any of these tests, the grid tuning is frequently more critical than the plate tuning. This is because the impedance match into the input capacitance of the grid becomes the bandwidth limiting factor. Even though the amplitude response appears flattened when the grid is heavily driven, the group delay (time) response has a serious effect on the higher order FM sidebands.

For more information about tuning an FM transmitter for optimum performance, see Refs. 14, 15, 20, and 22 at the end of this chapter.

OPTIMUM TUNING VERSUS EFFICIENCY

VHF amplifiers often exhibit a somewhat unusual characteristic when tuning for maximum efficiency. The highest efficiency operating point does not exactly coincide with the lowest plate current, because the power output continues to rise on the inductive side of resonance coming out of the dip in plate current. If the amplifier is tuned exactly to resonance, the plate load impedance will be purely resistive and the load line will be linear. As the output circuit is tuned to the inductive side of resonance, the plate load impedance becomes complex and the load line becomes elliptic instead of linear since the plate current and plate voltage are no longer in phase. Best efficiency occurs when the phase of the instantaneous plate voltage slightly leads the plate current. This effect is believed to be caused by the nonlinear gain characteristics of the power amplifier tube operating on an elliptic load line.

INSTALLATION CONSIDERATIONS

Adequate planning and care in the installation of an FM broadcast transmitter and associated equipment will help avoid many problems that may be difficult and expensive to correct later. For example, poor grounds and ground loops may cause high noise levels.

Wiring the Transmitter Plant

Separate metallic shielded conduits or troughs should be provided for the audio and the AC wiring. A third conduit should be used if computer logic levels are employed for equipment control. These conduits or wiring troughs may be either overhead or below the cabinets. The AC wiring should be well separated from the audio pairs to prevent the induction of unwanted hum and noise into the audio circuits.

Audio shields should be grounded at only one point to prevent ground loops in the shields. This point may have to be found experimentally to give the lowest noise pickup. The equipment racks and transmitter should be connected together by copper straps at least two inches wide, tied to a good earth ground at one point. If a good ground screen is not available, a satisfactory ground can be provided by driving four or five copper ground rods eight to ten feet long into the ground with a spacing of about three feet. These ground rods should be tied together with a wide copper strap. The straps connecting the equipment to the earth ground should be as short and direct as practical.

It is often difficult to remove VHF-RF from the equipment by grounding, because at FM carrier frequencies, nearly any connection to an earth ground has an appreciable impedance. The best way to keep RF out of sensitive low-level circuits is by keeping them enclosed within an RF shield, and by filtering leads that enter the shielded unit when necessary. Filters in the audio lines may be made up of small bifilar RF chokes and disc capacitors.

For stereo transmission, it is necessary to keep the L and R audio lines phased properly. To insure proper monaural compatibility, correct audio polarity must be maintained throughout the station from the microphones, tape machines, and turntables through all of the audio equipment to the stereo generator audio input terminals. Stereo phone line pairs or separate RF studio-to-transmitter links (STL) should also be checked for proper polarity and equal phase delay.

The transmitter equipment should be located and arranged to provide sufficient room around the front, sides, and rear for easy access during servicing and maintenance. Servicing of certain components may be easier by removing a side panel from the transmitter.

Transmitter Cooling

Almost all FM broadcast transmitters require forced air cooling, to remove heat from the output stage and other assemblies within the cabinet. A very important consideration in locating the transmitter is the provision for adequate cooling air. As a rough approximation, it can be assumed that the overall efficiency of the transmitter is about 50%. In other words, it will generate about the same number of kilowatts of heat as it does RF power output.

Fig. 41A shows a transmitter located in an air-conditioned room. This type of closed-loop system requires no special ducting and has the advantage that the transmitter intake air is usually much cleaner than outside air. The transmitter exhaust air places a substantial heat load on the air-conditioner during the summer, but it becomes a source of heat in the winter. The transmitter manufacturer can usually supply data on the number of cooling BTU's required, so that the proper size air-conditioner can be selected. This method is used frequently with the lower power transmitters. A protective system should be provided to prevent over-heating of the transmitter if the air conditioner fails.

Fig. 41B shows a transmitter located in a wall separating an air conditioned room and a ventilated, but not air-conditioned room. A large exhaust fan is provided in the ceiling to remove the rising hot air while an adequate cool air intake is provided in the lower portion of an outside wall. Adequate air filtering is required to keep the transmitter interior clean.

Fig. 41C shows a transmitter located in an air-conditioned room with intake and exhaust air ducts to the outside. An auxiliary blower or fan is normally required to overcome pressure drop in the ducting. This type of system requires careful design to make sure that the air flow through the transmitter is not impeded by the duct work. Additional air interlocks may be required to protect the transmitter from a failure of the external fan. The air intake and exhaust openings to the outside should be provided with rain shields, insect screens, and dust filters as dictated by the environment. The location of the air intake and exhaust openings should be arranged so that wind pressure will not impede the air flow.

Air filters should be periodically cleaned or replaced according to the transmitter manufacturer's instructions. This is very important because dust or insect clogged air filters may reduce the cooling air flow enough to cause overheating of some of the components. The probability of component failure increases very rapidly when cooling is insufficient. Particular attention should be paid to removing dirt and dust from high voltage components during regular maintenance after all power is removed and all components are discharged.

Dust should be cleaned from the transmitter by means of a suitable brush and vacuum cleaner or as otherwise recommended by the transmitter manufacturer. Usually weekly cleaning is sufficient.

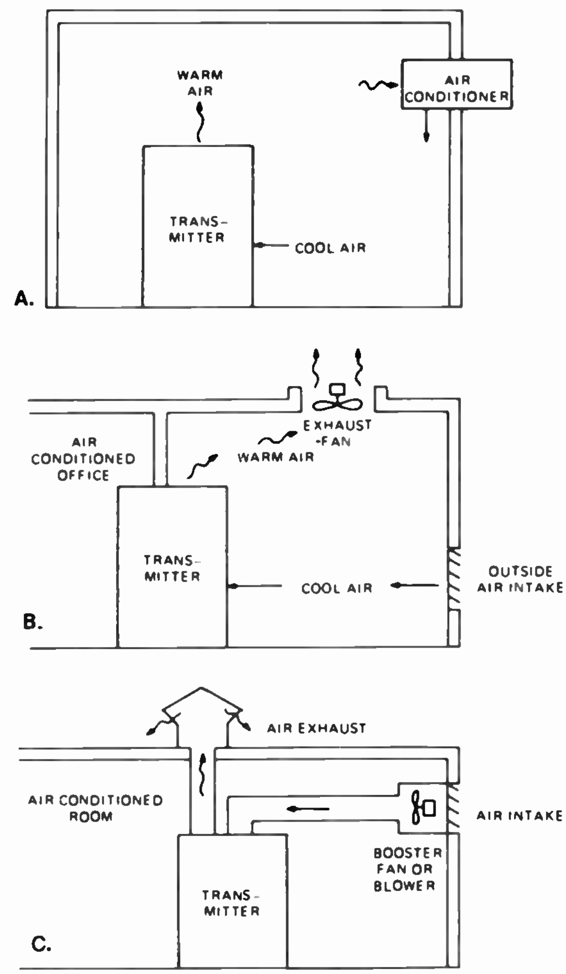


Figure 41. Three methods of providing cooling air for the transmitter.

FM TRANSMITTER PLANT MAINTENANCE

Care of Power Tubes

The operating life of high-power vacuum tubes can be extended by proper care. Most high power tubes utilize a directly heated cathode composed of a thoriated tungsten filament structure.

The key points to extending the life of RF power tubes are:

1. Store tubes upright (upside-down or right-side-up) along the axis of symmetry, not on their side. This will help to keep the internal elements concentrically aligned.
2. Use care when handling tubes to prevent mechanical shocks to the delicate internal structure. Do not set a tube on a hard surface without padding.
3. Keep the tube seals and anode cooler free of dust and dirt by cleaning once each week even in a clean environment.
4. Keep a spare tube on hand and rotate the tubes every few months to help keep the chemical "gas-getter" active so that the tubes remain gas free.
5. Keep a regular record of all tube operating parameters so that any trend of changes will be noticeable. In the event of a tube failure during the warranty period, this data will be essential to receiving credit on a replacement tube.
6. Monitor the filament voltage on a true RMS responding instrument and log any changes for future reference. The sampling point for this voltage measurement should be located as close to the tube's filament contacts as possible to minimize errors due to voltage drops in the filament wiring.

A tube that has been properly operated should gradually lose emission from the cathode until it is no longer useful, because the emissive material is gradually consumed. The carcass of the tube can then be rebuilt with a new cathode and recycled back into service. Tube life is not directly related to plate dissipation (within the ratings), but is related to the filament operating temperature (i.e., filament voltage) and the current density (milliamperes per heater watt) emitted by a given size filament. This means that operating a given tube type at a lower filament voltage and plate current will proportionately increase the life of the tube. For directly heated, thoriated tungsten filaments the plate current should be less than four milliamperes per watt of heater power for extended life.

Normally a new tube will deliver full output at a reduced filament voltage. By operating the tube at the optimum filament voltage, the filament life can be significantly extended. The optimum value may be found by slowly reducing the filament voltage from the manufacturer's rated value until the RF power output drops about 2% and then increasing the filament voltage until the RF power increases back up 1%. Recent informational bulletins[26] on extending tube life from tube manufacturers permit the filament voltage to be reduced more than 5% below the manufacturer's rating, as long as operation is closely monitored to stay above the point where there is a 1% drop in power output. A brand new tube should be operated at the full rated filament voltage for the first 200 hours before the voltage is reduced to the optimum value for long life. This will assure that the "gas-getter" is properly activated. As the tube ages, the filament voltage will have to be increased to stay at the optimum value, until RF output power cannot be maintained at or above the rated value of filament voltage. At this point, the tube's useful life comes to an end. Check the manufacturer's data sheets and most recent application notes which are often enclosed with the tube for detailed information. An excellent guide to the proper care of power tubes is listed in Refs. 10 and 26 at the end of this chapter. Ref. 17 gives detailed information about how to specify a proper forced-air cooling system for power tubes.

Preventive Maintenance

Preventive maintenance is equipment inspection and maintenance performed at regular intervals before an operational problem develops. The long term benefits are great because potential problems are discovered and solved while they are still easily manageable. A check-list of a few typical preventive maintenance items for an FM transmitter plant might include:

1. Weekly overall internal and external cleaning and inspection for damage or excessive wear.
2. Lubricate motors, tuning gears, and other moving parts at intervals recommended by the manufacturer.
3. Check and log all meter readings, including filament voltage, daily. Then compare these readings with the previous set of readings as an aid to diagnosing a developing problem.
4. Regularly exercise the automatic power control and any other servo systems.
5. Check the antenna lighting and deicer systems.
6. Check the transmission line pressurization and VSWR.
7. Check all air filters in the transmitter plant and clean or replace as required.
8. Check the proper operation of all monitoring and remote control equipment.
9. Good overall housekeeping will pay big dividends in the long run by keeping the equipment clean and free of problems that would otherwise be caused by dirt build-up.

Maintenance Systems

The key to making any maintenance program work is to set up formal check-list, logging, parts inventory management, and repair scheduling systems. These systems provide the conscience and discipline required to keep the maintenance routine accurate and complete. Each station should develop a system suited for the particular physical plant involved. Often there will be more maintenance and repair work needing attention than there is time to do it all. In this case, the maintenance system should set the priorities for completing each item, and assure that no item gets forgotten. Accurate notebooks describing all installation and maintenance work are a very helpful part of any maintenance system when working on the equipment years after the human memory has faded.

REFERENCES

1. Lessing, Lawrence, "Man of High Fidelity: Edwin Howard Armstrong", Bantam Publishing, 1959, The Armstrong Foundation, Columbia University, New York, N.Y.
2. Reference Data for Radio Engineers, ITT—Howard W. Sams and Co., Inc., New York, N.Y., Fifth Edition, 1970.
3. Terman, Frederick E., "Electronic and Radio Engineering", McGraw-Hill Book Company, New York, N.Y., Fourth Edition, 1955.
4. Clarke, Kenneth K. and Hess, Donald T., "Communications Circuits: Analysis and Design", Addison-Wesley Publishing Company, Reading, Massachusetts, Second Printing, 1978.
5. Rohde, Ulrich L., "Digital PLL Synthesizers-Theory and Design", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.
6. Krauss, Herbert L., Bostian, Charles W., and Raab, Frederick H., "Solid State Radio Engineering", John Wiley and Sons, New York, N.Y., 1980.
7. Uyttendaele, Antoon G., "Design Requirements and Operational Features of the Gates Dualtran RF Switching System", Broadcast Products Division of Harris Corporation, Quincy, Illinois.
8. Johnson, Joseph, "Solid Circuits", Acrian/Communications Transistor Company, Cupertino, California, 1973.
9. Howe, Harlan Jr., "Simplified Design of High Power, N-Way, In-Phase Power Divider/Combiners", Microwave Journal, December, 1979.
10. Eimac, "Care and Feeding of Power Grid Tubes", Eimac Division of Varian Corporation, San Carlos, California, 1967.
11. Hershberger, David and Weirather, Robert, "Amplitude Bandwidth, Phase Bandwidth, Incidental AM, and Saturation Characteristics of Power Tube Cavity Amplifiers for FM", Harris Corporation, Quincy, Illinois, 1982.
12. Motorola Silicon Rectifier Manual, Semiconductor Group-Motorola Inc., Phoenix, Arizona, 1980.
13. Hnatek, Eugene R., "Design of Solid-State Power Supplies", Van Nostrand Reinhold Company, New York, N.Y., Second Edition, 1981.
14. Mendenhall, Geoffrey N., "A Study of RF Intermodulation Between FM Broadcast Transmitters Sharing Filterplexed or Co-located Antenna Systems", Broadcast Electronics Inc., Quincy, Illinois, (c)1983.
15. Mendenhall, Geoffrey N., "The Composite Signal—Key to Quality FM Broadcasting", Broadcast Electronics Inc., Quincy, Illinois, (c)1981.
16. Lyles, John T.M. and Shrestha, Mukunda B., "Transmitter Performance Requirements for Subcarrier Operation", Broadcast Electronics Inc., Quincy, Illinois, (c)1984.
17. Steinkamp, Jeffrey H., "Air Flow Measurements Using a Variable Frequency AC Generator", Broadcast Electronics Inc., Quincy, Illinois, (c)1984.
18. Robinson, John T. and Whipple, L. Glenn, "FCC Regulations Governing the Audio Fidelity Characteristics of FM Broadcast Equipment", Audio Engineering Society Journal, Vol. 16, No. 3, July 1968.
19. Bruene, Warren B., "An Inside Picture of Directional Wattmeters", American Radio Relay League—QST magazine, April 1959.
20. Mendenhall, Geoffrey N., "Tecniques for Measuring Synchronous AM Noise in FM Transmitters", Broadcast Electronics, Inc., Quincy Illinois, (c) 1988.
21. Anthony, Edward J., "Practical Considerations for the Implementation of a Reliable Synchronous FM Booster", Broadcast Electronics, Inc., Quincy, Illinois, (c) 1988.
22. Shrestha, Mukunda B., "The Significance of RF Power Amplifier Circuit Topology on FM Modulation Performance", Broadcast Electronics, Inc., Quincy, Illinois, (c) 1990.
23. "FMSIM" (c) 1990—FM Stereo Simulation and Analysis Program by Quantics developed in cooperation with Broadcast Electronics, Inc., P.O. Box 3606, Quincy, Illinois 62305–3606, (217)224–9600.
24. Anthony, Edward J., "Optimum Bandwidth for FM Transmission", Broadcast Electronics, Inc., Quincy, Illinois, (c) 1989.
25. Shrestha, Mukunda B., "Personal Safety Considerations with Broadcast Transmitters", Broadcast Electronics, Inc., Quincy, Illinois, (c) 1989.
26. Artigo, Robert, "Extending Transmitter Tube Life", Broadcast Management/Engineering Magazine March 1982, Revised March 1990 and Reprinted as Eimac Application Bulletin AB-18, Eimac Division of Varian Corporation, San Carlos, California.

3.4 Subcarrier Transmissions and Stereophonic Broadcasting

John Kean
Moffet, Larson & Johnson, Inc., Falls Church, Virginia

Stereophonic sound FM broadcasting and Subsidiary Communications Authorization (SCA) services are a form of multiplexing that had its origins fifty years ago, when high-fidelity audio and facsimile messages were simultaneously transmitted from the Empire State Building in New York City to experimental receive sites in New Jersey.[1] Those historic multiplexing efforts demonstrated the value of wideband frequency *modulation*, allowing modern FM broadcasting to provide services to the public of both quality and variety.

Despite its illustrious start, FM broadcasting in the 1950s was a marginal business. The number of operating FM's declined annually from about 700 in 1947 to slightly over 500 in 1957. Many depended on additional income from their SCA services, which offered background music and other programs to stores and offices over medium-fidelity audio subcarriers. The FCC gave the then struggling industry a boost by granting something AM and TV didn't have: stereo. The FCC's April, 1961 Report and Order authorized transmission of a stereophonic sound system that combined the system proposals of Zenith Radio Corp., Chicago, IL, and General Electric Co., Syracuse, NY.

FM broadcasters did not show much excitement for their new capability, but the growing availability of stereophonic LP records and home stereo equipment created a natural market for FM stereo. Most agree that FM's resurgence and eventual dominance as an audio medium was due in large part to stereophonic transmission. SCA, as well, is still actively employed by FM licensees.

THE COMPOSITE BASEBAND

This section contains definitions of the terms commonly used in the FM stereo and SCA systems.[2] Since some of these terms are misused or ambiguous, the intention here is to establish meanings that will be used in the rest of this chapter.[3] The list was selected to clarify certain terms and is not a complete glossary.

Glossary of Terms For Stereo and SCA

Multiplexing. In its simplest sense, multiplexing implies that two or more independent sources of information are combined for carriage over a single medium (namely, the radio frequency carrier), and then are separated at the receiving end. In stereophonic broadcasting, for example, program information consisting of left and right audio signals are multiplexed onto an FM carrier for transmission to receivers, which subsequently recover the original audio signals.

Channel. A transmission path. The usage herein distinguishes between the concept of a channel (i.e., main channel, stereophonic subchannel, etc.), and left and right audio signals.

Composite Baseband Signal. A signal which is the sum of all signals that frequency modulate the main carrier. The signal includes the main channel signal, the modulated stereophonic subcarrier, the pilot subcarrier, and the SCA subcarrier(s).

FM Baseband. The frequency band from 0 hertz (Hz) to a specified upper frequency which contains the composite baseband signal.

Main Channel. The band of frequencies from 50 (or less) Hz to 15000 Hz on the FM baseband which contains the main channel signal.

Main Channel Signal. A specified combination of the monophonic, or left and right audio signals, which frequency modulates the main carrier.

Stereophonic Sound. The audio information carried by plurality of channels arranged to afford the listener a sense of the spatial distribution of sound sources. Stereophonic sound includes, but is not limited to, biphonic (two channel), triphonic (three channel), and quadraphonic (four channel) services.

Stereophonic Sound Subchannel. The band of frequencies from 23 kHz to 99 kHz (53 kHz for two channel transmission) containing sound subcarriers and their associated sidebands.

Subchannel. A transmission path specified by a subchannel signal occupying a specified band of frequencies.

Subchannel Signal. Subcarrier(s) and associated sideband(s) which frequency modulate the main carrier. It is synonymous with subcarrier, as in the stereophonic subcarrier or SCA subcarrier.

Frequency Deviation. The peak difference between the instantaneous frequency of the modulated wave and the average carrier frequency.

Percentage Modulation. The ratio of the actual frequency swing of the carrier to the frequency swing defined as 100% modulation, expressed in percentage. Although current FCC rules conditionally permit greater than 100% modulation when SCAs are transmitted, a frequency swing of ±75 kHz is still defined as 100% modulation.

Injection. The ratio of the frequency swing of the FM carrier by a subchannel signal to the frequency swing defined as 100% modulation, expressed in percentage. The total injection of more than one subchannel signal is the arithmetic sum of each subchannel injection.

Crosstalk. An undesired signal occurring in one channel caused by an electrical signal in another channel.

Linear Crosstalk. A form of crosstalk in which the undesired signal(s) is created by phase or gain inequalities in another channel or channels. Such crosstalk may be due to causes external to the stereophonic generator; consequently it is sometimes referred to as system crosstalk.

Nonlinear Crosstalk. A form of crosstalk in which the undesired signal(s) is created by harmonic distortion or intermodulation of electrical signal(s) in another channel or channels. Such crosstalk may be due to distortion within the stereophonic generator or FM transmitter; consequently it is sometimes referred to as transmitter crosstalk.

Frequency Spectrum and Modulation Limits

The FCC's stereophonic transmission standards are contained in Section 73.322 of the FCC Rules, and the SCA transmission standards are contained in Section 73.319. The characteristics and methods of generating these signals are discussed below, but readers are also referred to the Commission's Rules for a listing of its standards.

The composite baseband extends to 99 kHz and may be used in support of either stereophonic or SCA multiplex services.. Stereophonic broadcasting includes, but is not limited to, biphonic (two channel) service. Within the frequency range of 23 kHz to 99 kHz, any form of amplitude modulation (DSB, SSB, etc.), angle modulation (FM or PM), or frequency shift keying of a multiplex subchannel is permitted.

The term SCA is an acronym for a "Subsidiary Communications Authorization," once required in order for a station to begin broadcasting a multiplex service. Although the familiar term "SCA" continues, the authorization is no longer required. Broadcast licensees may begin transmitting multiplex services without prior notification of or authorization from the FCC.

Under certain conditions when SCA multiplex subcarriers are operated, a total modulation of up to 110% is legal. Fig. 1 shows the baseband frequency ranges and modulation limits for various modes, from monophonic to stereophonic, plus fully-loaded SCA operation.

Fig. 1A represents the basic monophonic program mode, where the baseband width is limited to approximately 15 kHz and no other signals are multiplexed. In this case all the modulating energy is contributed by the main channel. Not more than 100% modulation is permitted in this case.

Fig. 1B shows the baseband with SCA operation in addition to monophonic main channel service. Total SCA injection up to 30% is permitted within the band from 20 kHz to 99 kHz. This injection figure may be contributed by one or more SCA subcarriers. To insure that the bandwidth of the main carrier (and its interference to other stations on adjacent and alternate channels) is not significantly increased, the arithmetic sum of all modulation must not exceed 100% plus one-half of the SCA injection. SCA injection between 75 kHz and 99 kHz may not exceed 10% under any conditions.

Note that the Commission Rules permit transmission of multiplex subcarriers when *no* broadcast program service is carried on the main channel, provided that the above modulation rules are met.

Fig. 1C shows the basic stereophonic sound program

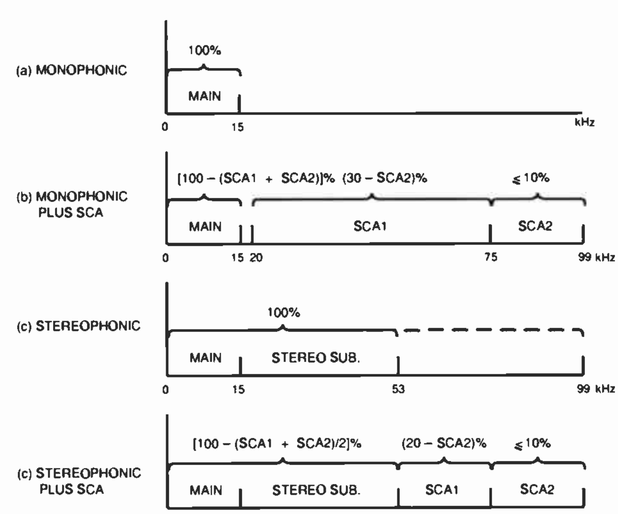


Figure 1. FM baseband scenarios with allowable modulation limits for monophonic, stereophonic, and SCA operation. Modulation percentages are referred to 75 kHz carrier deviation.

mode, without SCA operation. As is the case for monophonic program operation, the modulation must be limited to 100%. Frequencies up to 99 kHz are available for multichannel sound program transmission.

Fig. 1D adds a single band for SCA operation to the stereophonic mode. While total SCA injection may be up to 20%, no more than 20% total injection may be employed within the frequency bands from 53 kHz to 75 kHz, and 10% from 75 kHz to 99 kHz. (There is one exception to determining total SCA injection involving subcarriers that are multiples of and synchronous to the stereophonic pilot. This is discussed in the SCA multiplex section of the chapter.) The modulation contributed by the main channel and stereophonic subchannel signals must be no more than 100% minus one-half the total SCA injection. Since the total injection may be up to 20%, total modulation may be up to 110%.[4]

GENERATING THE STEREOPHONIC BASEBAND SIGNAL

Fig. 2 shows the composite baseband that modulates the FM carrier for biphonic broadcasting. (SCA multiplex subchannels are not part of this band and will be discussed later.)

The two-channel stereo baseband has a bandwidth of 53 kHz, and consists of:

1. A main channel which consists of the sum of left plus right audio signals $(L+R)$. This is the same signal broadcast by a monaural FM station, but is reduced by approximately 10% to allow for the stereo pilot injection.
2. A stereophonic sound subchannel $(L-R)$ is required, consisting of a double sideband amplitude modulated subcarrier with a 38 kHz center frequency. The modulating signal is equal to the instantaneous difference of the left and right audio signals. The subcarrier is suppressed to avoid wasting modulation capability. The pairs of AM sidebands have the same peak modulation potential as the main channel.
3. A 19 kHz subcarrier *pilot* which must be exactly one-half the frequency of the stereophonic subcarrier and very nearly in phase to it. It supplies the reference signal needed to synchronize the decoder circuitry in receivers. The frequency tolerance is ±2 Hz and the main carrier must be modulated between 8% and 10%.

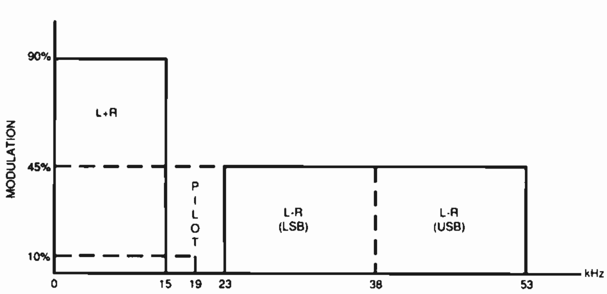


Figure 2. Biphonic (two-channel) stereo baseband.

In general two principles have been used to generate the stereophonic subchannel. The methods are called time division multiplex (TDM) or switching method and frequency division multiplexing (FDM) or matrix method.

Frequency Division Multiplexing

A basic method for generating the stereophonic baseband involves the direct generation of the double sideband suppressed $L-R$ subchannel along with the $L+R$ channel.

A simplified block diagram of the FDM system is shown in Fig. 3. Both left and right audio channels are preemphasized and low pass filtered. In the matrix the left and right audio signals are both added and subtracted. The audio signals are added to form the $L+R$ main channel which is also used as the monaural broadcast signal.

The subtracted signals are fed to a balanced modulator which generates the $L-R$ subchannel. Since a balanced modulator is used, the carrier at 38 kHz will be suppressed, leaving only the modulated sidebands.

The 38 kHz oscillator is divided by two to make the 19 kHz pilot tone. Finally, the main channel, stereophonic subchannel and pilot are combined in the proper proportions (45 + 45 + 10) to form the composite output.

An examination of the composite stereo waveform in the time domain, such as displayed by an oscilloscope, is helpful. First consider a 1 kHz sine wave applied equally to the L and R audio inputs. This is shown in Fig. 4A without a pilot signal. The only frequency present in the spectrum graph is 1 kHz, since the matrix produces no difference signal necessary to generate sidebands in the stereophonic subchannel.

Fig. 4B illustrates the ideal composite signal when the same 1 kHz tone is applied to the L and R inputs but exactly out of phase. With the pilot still off, two frequency components at 37 kHz and 39 kHz are generated. No $L+R$ signal appears from the matrix, thus only the sidebands of the modulated 38 kHz subchannel are present.

The symmetrical envelope shown represents a double-sideband suppressed carrier (DSBSC) AM signal. Note that the amplitude of each sideband is one-half that of the $L+R$ component in Fig. 4A. In the receiver's stereo decoder the sidebands are added together to produce an output equal to the full left signal.

Finally, consider the waveform in Fig. 4C, when the composite signal (still without pilot) is generated by applying a 1 kHz tone to the L input alone. The baseline of the waveform envelope will be a straight line if there is no amplitude or phase difference between the main channel and subchannel. Three frequency components

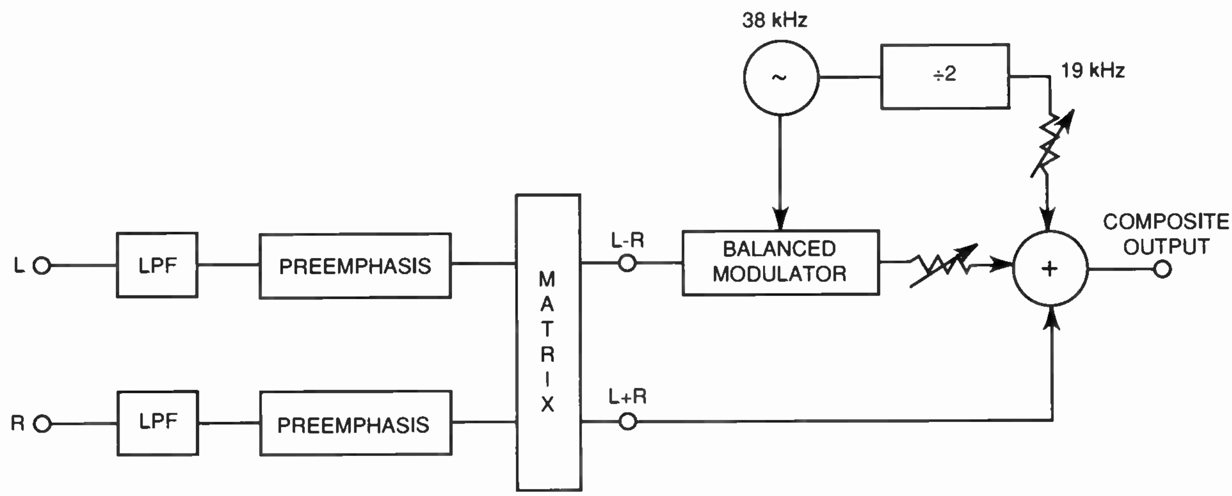


Figure 3. Functional blocks of a frequency division multiplex stereo generator.

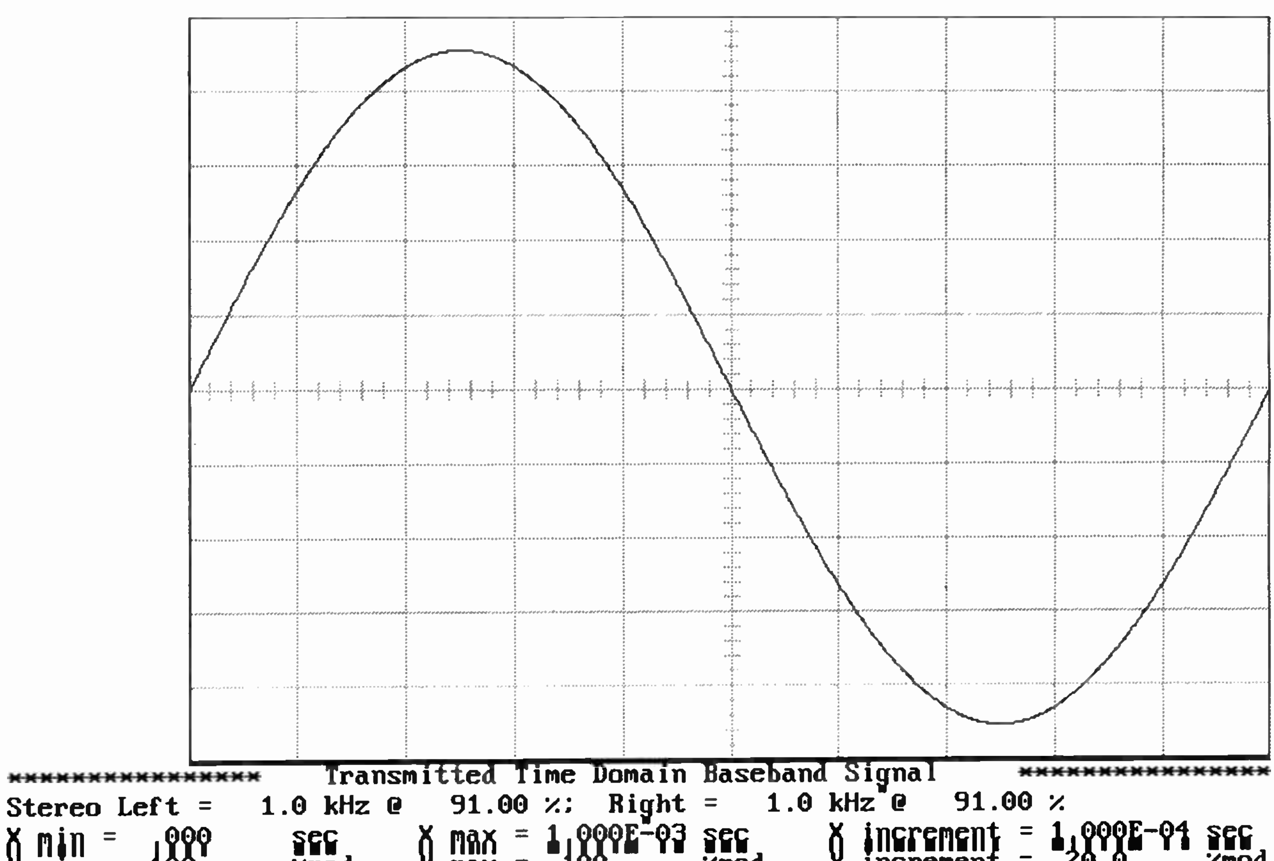


Figure 4A(1). 1 kHz Left at 91%, 1 kHz Right at 91%, 19 kHz at 0%. Identical sinusoidal L and R inputs. (Courtesy of Quantics.)

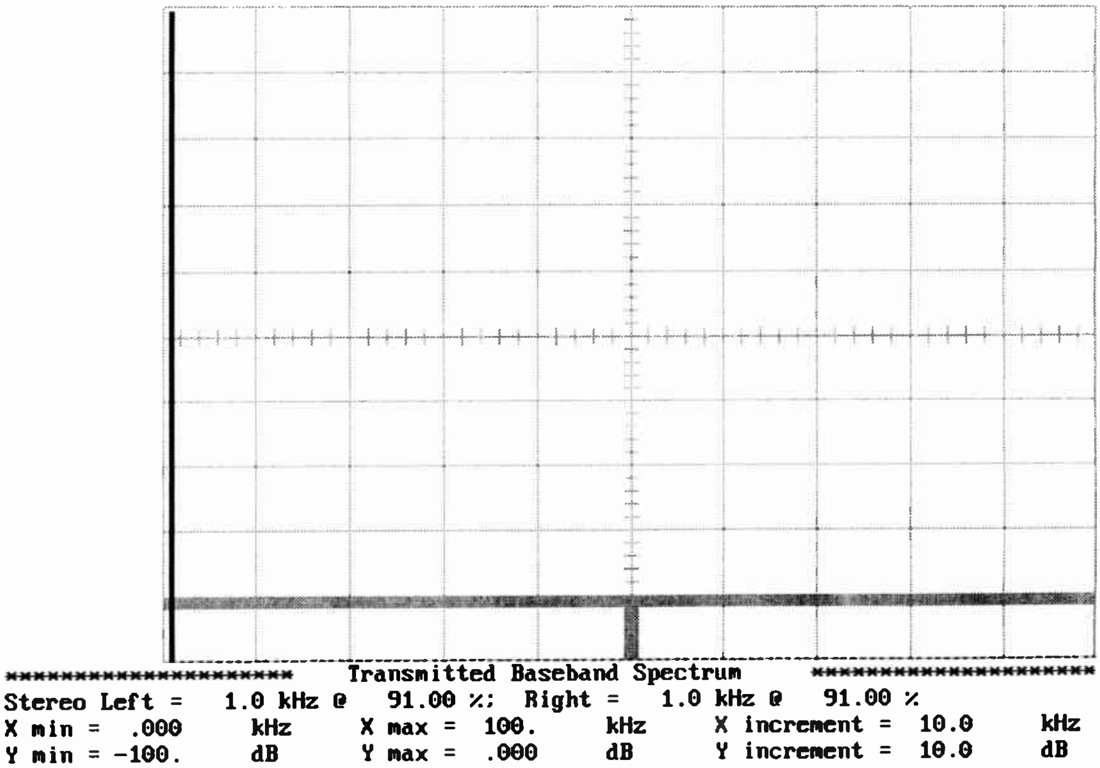


Figure 4A(2). Same modulation conditions; only 1 kHz fundamental at 90% modulation (−0.92 dB) is present. (Courtesy of Quantics.)

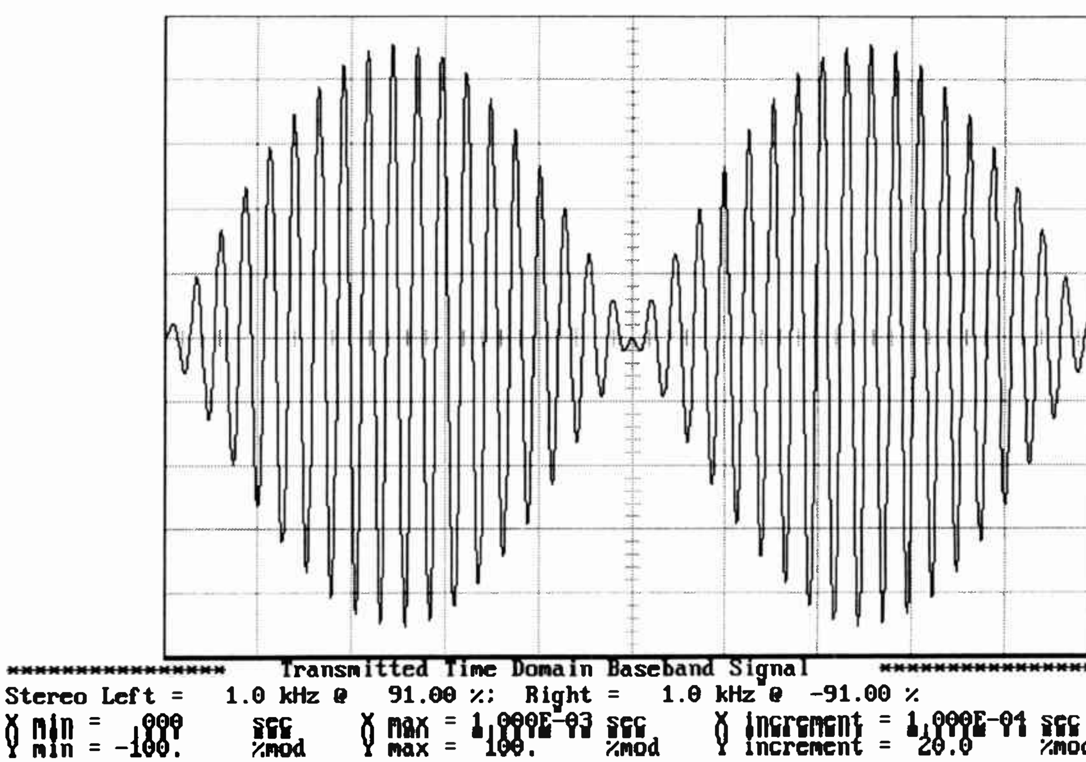


Figure 4B(1). kHz Left at 91%, 1 kHz Right at 91%, 19 kHz at 0%. Identical but out-of-phase sinusoidal L and R inputs. (Courtesy of Quantics.)

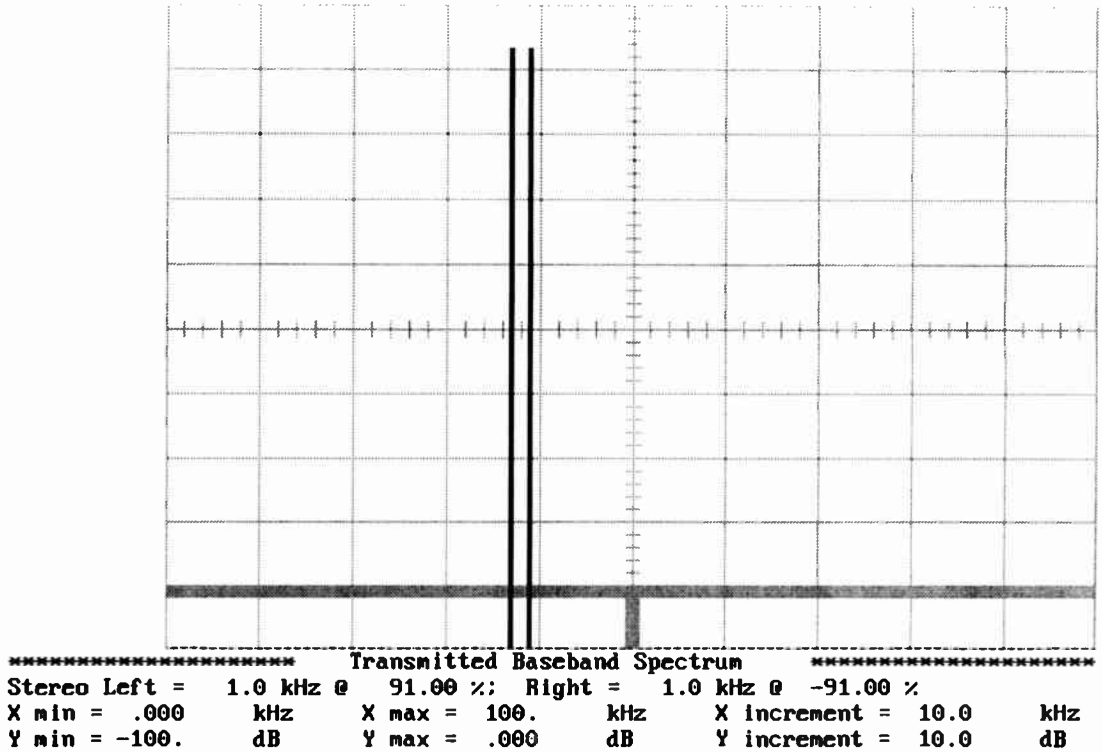


Figure 4B(2). 0-100 baseband spectrum; 1 kHz difference signal appears as upper and lower sidebands of 38 kHz subcarrier each at 45% modulation (−6.94 dB) for 90% total modulation. (Courtesy of Quantics.)

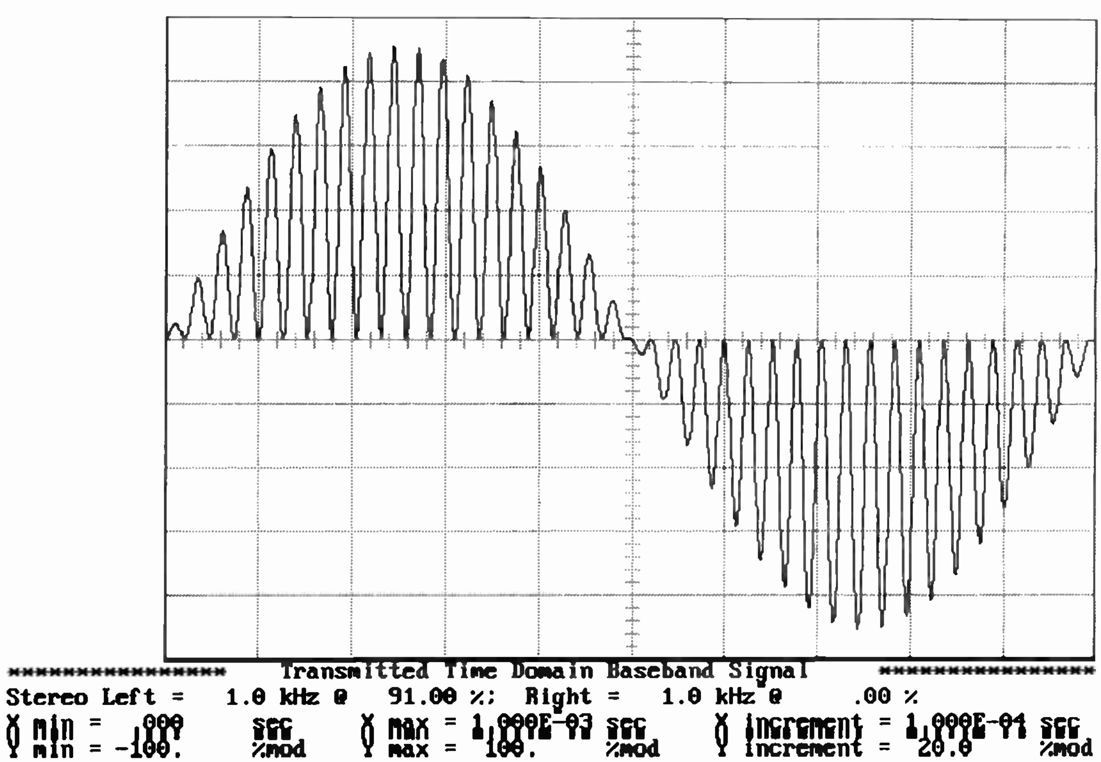


Figure 4C(1). kHz Left at 91%, 1 kHz Right at 91%, 19 kHz at 9%. Sinusoidal L input at 90% modulation. (Courtesy of Quantics.)

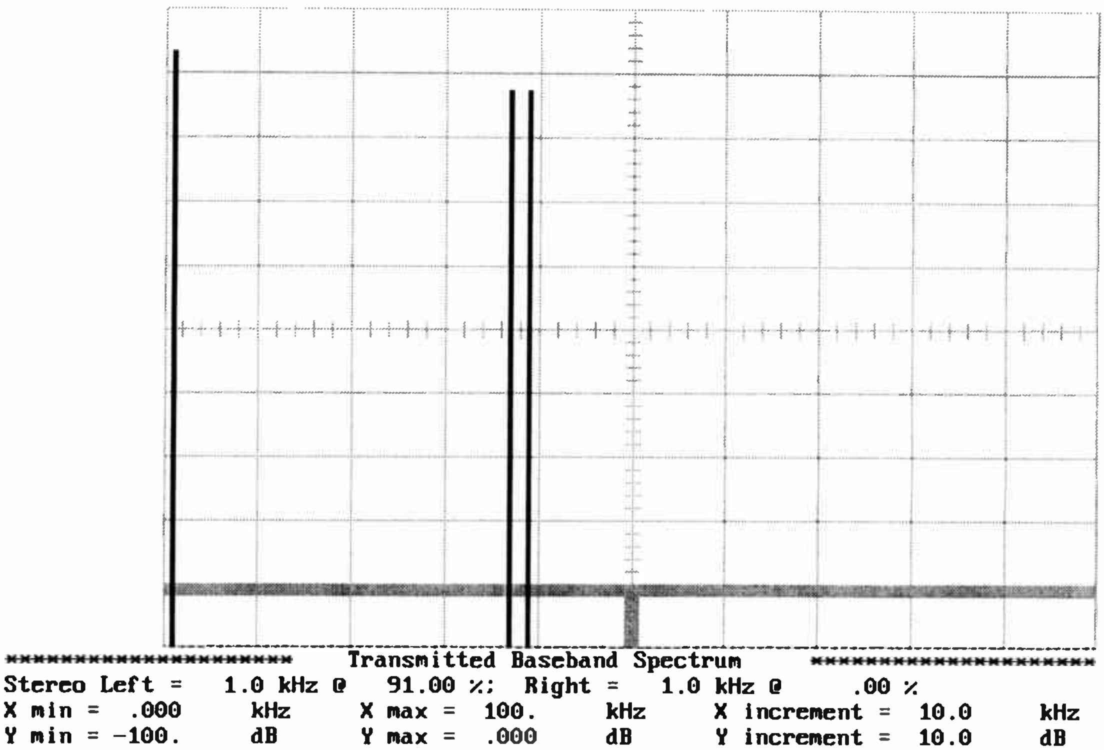


Figure 4C(2). Baseband spectrum of 4C; modulation of 1 kHz signal appears a fundamental frequency at 45% level and as upper and lower sidebands of 38 kHz subcarrier at 22.5% each (−13.0 dB). (Courtesy of Quantics.)

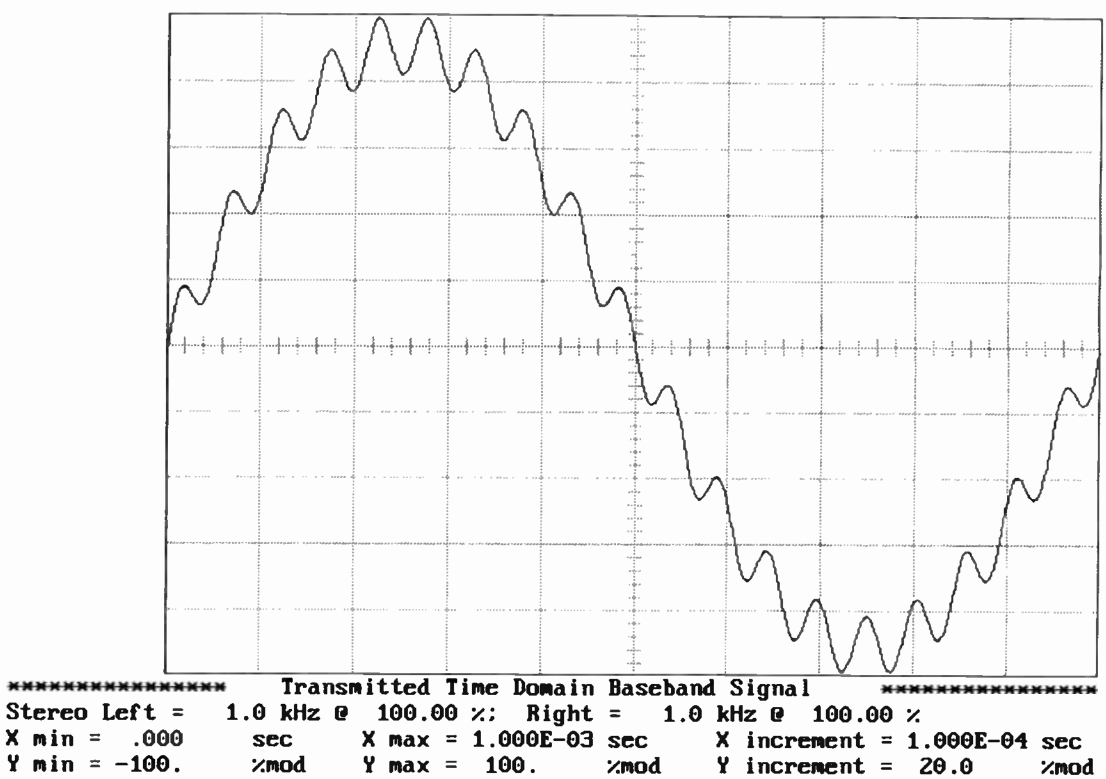


Figure 4D(1). Same modulation conditions as A, with 19 kHz pilot at 9% injection (−20.9 dB). (Courtesy of Quantics.)

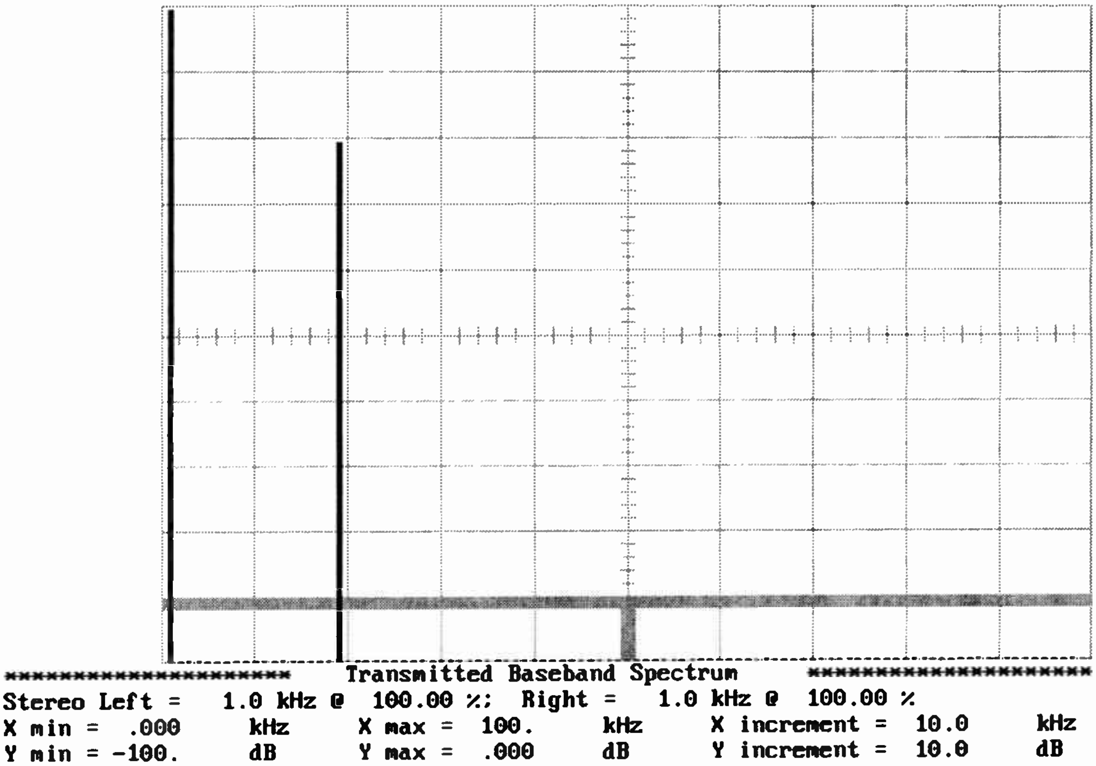


Figure 4D(2). Same modulation conditions as A, but with 19 kHz pilot at 9%. (−20.9 dB). (Courtesy of Quantics.)

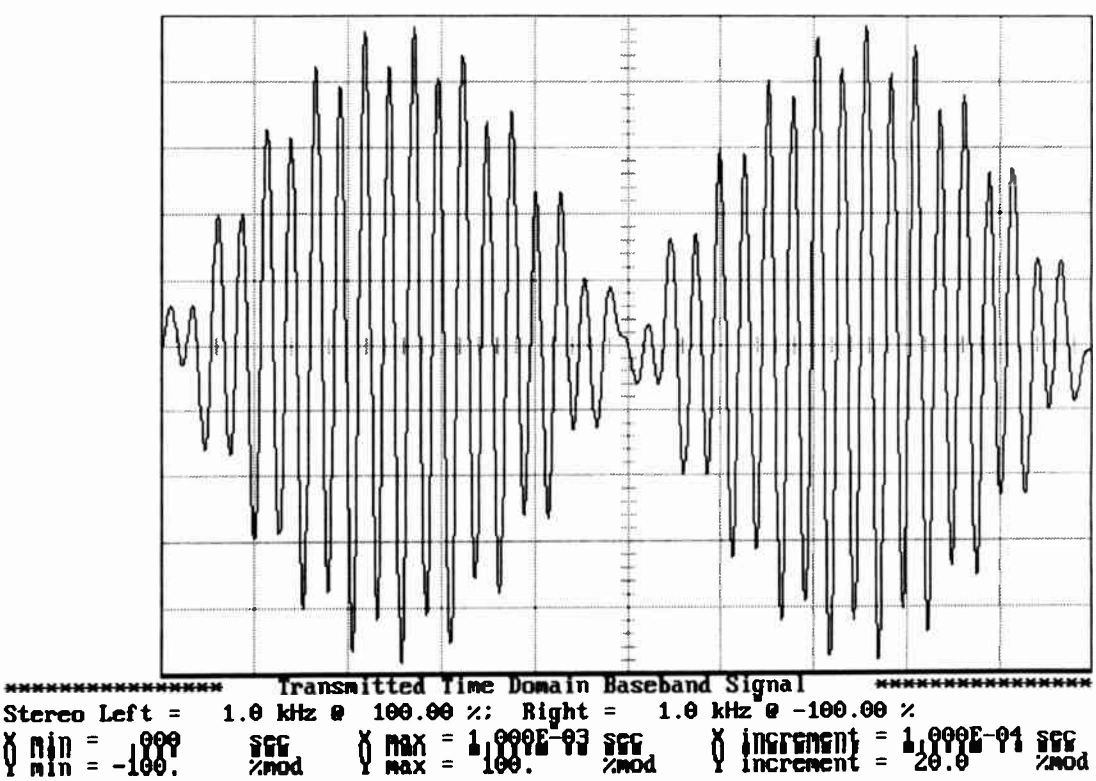


Figure 4E(1). Same modulation conditions as B, but with 19 kHz pilot at 9%. (Courtesy of Quantics.)

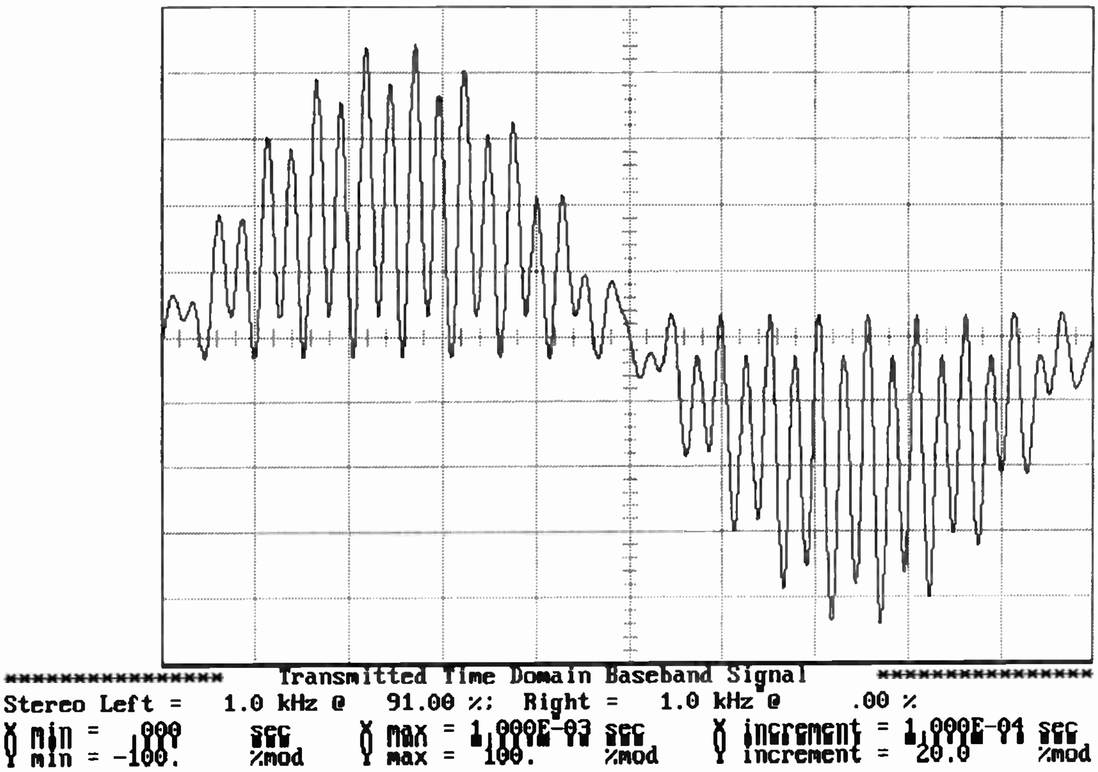


Figure 4E(2). Same modulation conditions as C, but with 19 kHz pilot at 9%. (Courtesy of Quantics.)

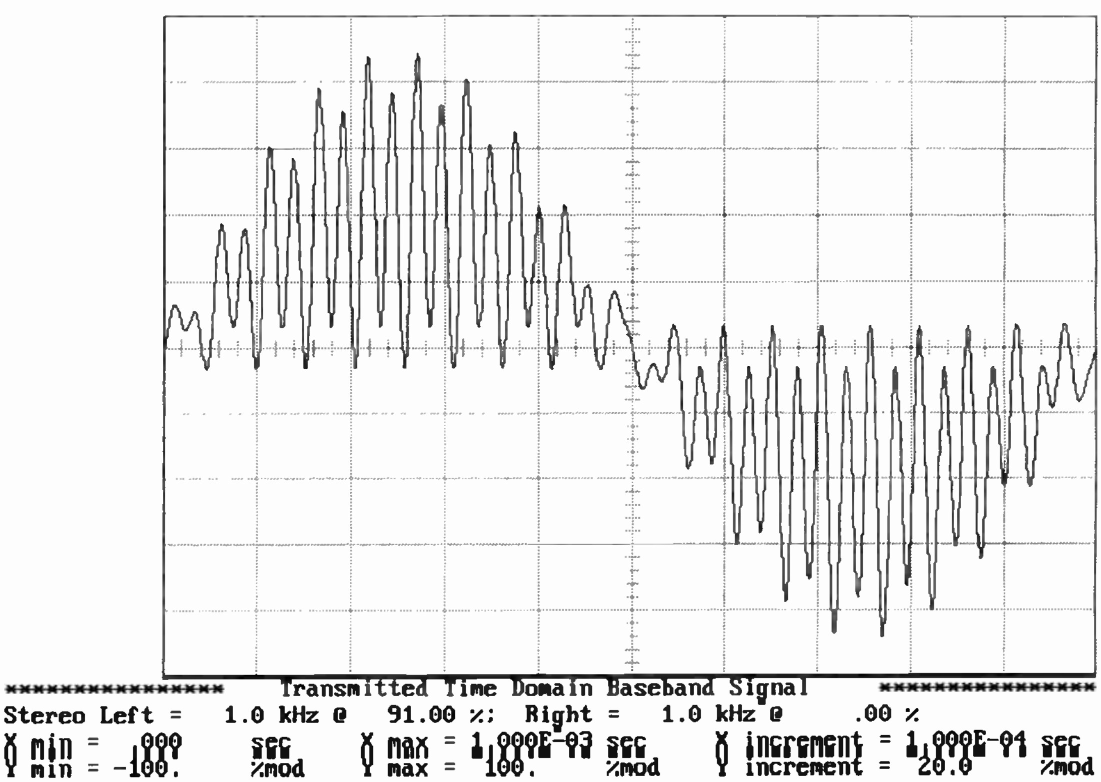


Figure 4F. Same modulation conditions as C, with 19 kHz pilot at 9%. (Courtesy of Quantics.)

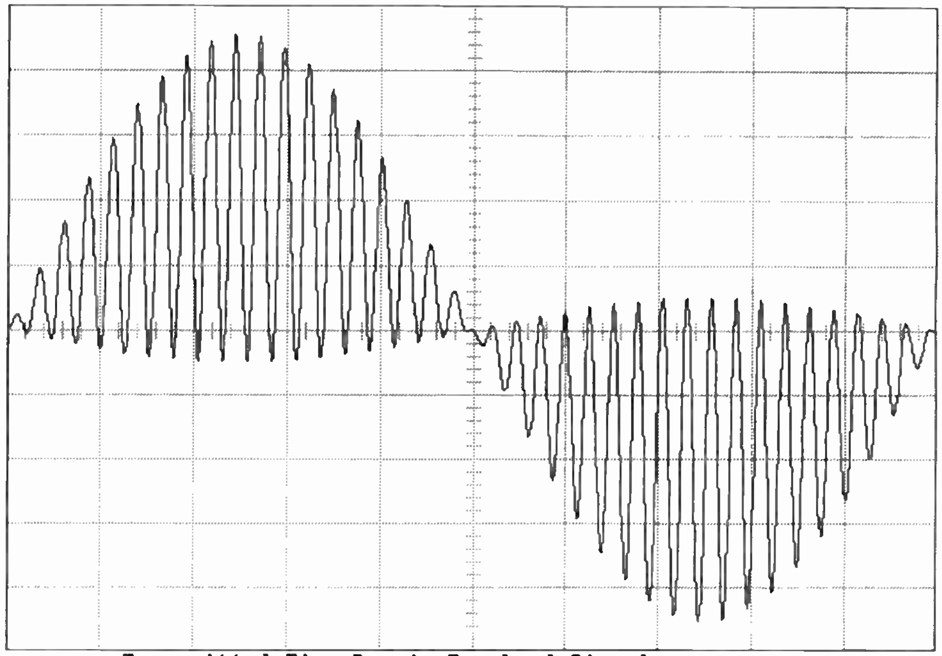


Figure 4G. Same modulation conditions as C, with mis-matched gain between Main Channel (L+R) and Subchannel (L−R). (Courtesy of Quantics.)

are present: 1 kHz, 37 kHz, and 39 kHz. These sidebands are each one-half the voltage amplitude of the 1 kHz signal in the main channel; together they equal the energy of the main channel in this instance. Figs. 4D-F show the same conditions as 4A-C with the pilot turned on. Fig 4G is an example of a possible degradation that can be caused by imbalance between the main channel and the subcarrier.

The last diagram looks the same when an *R*-only 1 kHz tone is applied, but the phase of the two sidebands would be reversed with respect to the 38 kHz subcarrier (and the pilot). Adding the pilot at 8% to 10% produces similar waveforms, but the oscilloscope display of the waveform baseline is fuzzier. For this reason, most stereo generators allow the pilot to be turned off for baseline measurements.

Time Division Multiplex

A different type of stereo generator is in use which produces a result similar to frequency division multiplexing by using a switching technique.

Generation of both the $L+R$ and $L-R$ channels is accomplished by an electronic switch that is toggled by a 38 kHz signal. The switch alternately samples one audio channel and then the other, as shown in Fig. 5. According to Nyquist criteria, the original signal can be reconstructed from periodic samples, provided that the samples are taken at a rate at least twice the frequency of the highest audio frequency component (approximately 15 kHz in broadcast FM).

The output waveform for the TDM generator shown in Fig. 6 is the time domain (as an oscilloscope would display the signal) for a sequence of input signals. The diagrams at the right of the waveform show the same signal in the frequency domain (as would be displayed on a spectrum analyzer).

In Fig. 6A, there are no input signals present. Ideally, no output signals are possible, and in practice only a small amount of leakage of the switching transients are present. Since the transfer time of the switching signal is extremely quick, harmonics of the fundamental 38 kHz are possible.

A 9 kHz audio tone is applied to the *L* and *R* inputs in Fig. 6B. The 9 kHz input signals are combined at full amplitude (90% modulation) and no subchannel sidebands are generated.

In Fig. 6C, only the left channel has a signal present. As the switch selects the *L* audio line, samples are passed along to the composite output. Therefore, the output waveform shows the same signal, chopped into segments of 1/38,000 of a second. Since the total area under the waveform has been divided in half, it should be apparent that the energy of the 9 kHz signal in the $L+R$ channel is only half the amplitude that it would be if an equal 9 kHz signal were also present at the right channel. The equation for the output signal (e) for an input signal (σ) at any instant (t) is

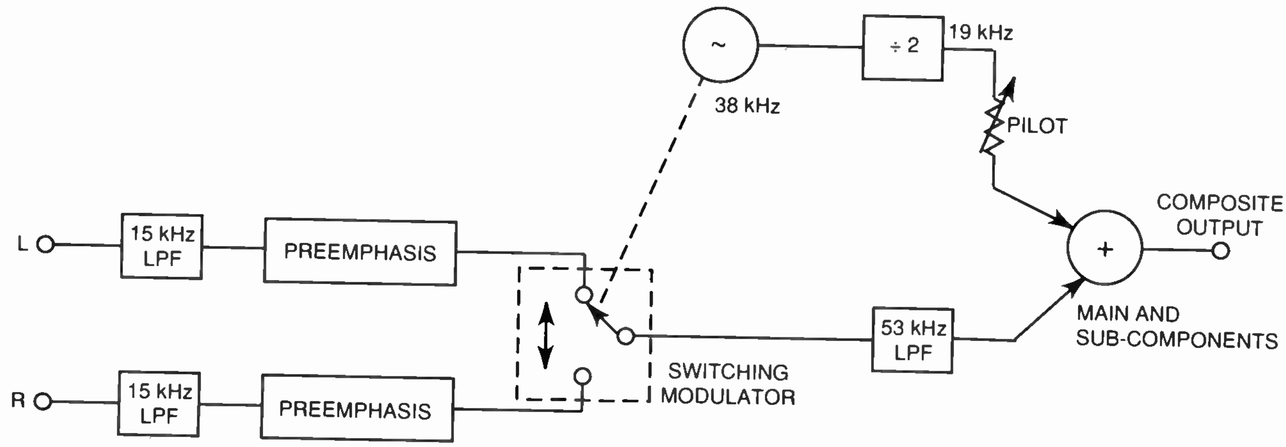


Figure 5. Functional blocks of a time division multiplex stereo generator.

$$
\begin{aligned}
e = {} & \tfrac{1}{2}\sin \sigma t + \\
& \text{(main ch. audio)} \\
& 1/\pi\,[\sin(\phi + \sigma) + \sin(\phi - \sigma)t] - \\
& \text{(DSBSC)} \qquad\qquad (1) \\
& \tfrac{1}{3}\pi\,[\sin(3\phi + \sigma)t + \sin(3\phi - \sigma)t] \\
& \text{(3rd harmonic)} \\
& \ldots \text{etc.} \quad \text{(higher harmonics)}
\end{aligned}
$$

Fig. 6C shows the original 9 kHz signal (at half amplitude), and a pair of sidebands centered about the 38 kHz switching frequency. No 38 kHz signal is generated if the switching waveform has perfect symmetry, that is, if the switch is connected to the left and right channels for precisely equal periods. Note that a harmonic of the stereophonic subcarrier is shown, centered around 114 kHz which is three times the switching frequency. Only one extra term was shown in the equation; however, other terms at the fifth, seventh, etc. are present.

In addition to the odd-order harmonics of the 38 kHz subchannel, asymmetry in the switching signal, or other circuit imbalances, can create some sidebands centered about the second harmonic at 76 kHz. All these harmonics must be removed by filtering, as shown in the diagram.

When the odd harmonics are filtered out, the proper DSBSC waveform results. However, it is slightly greater in amplitude than the $L+R$ signal because the fundamental component of the square wave is $4/\pi$ or 1.27 times larger than the square wave amplitude. This is easily corrected by adding enough of the L and R audio to the output to equalize the amplitude.

In Fig. 6D, the TDM signal is shown when the L and R signals are equal in amplitude and exactly reversed in phase. This waveform matches the composite stereo signal shown in Fig. 4C.

The composite lowpass filter must have very steep cutoff characteristics but should have flat amplitude response and linear phase shift with frequency (equal time delay at all frequencies) below 53 kHz. While this approach to stereophonic generation is simple and stable, the filter can degrade stereo separation, especially at higher audio frequencies.

The 19 kHz pilot square wave from the ÷2 digital divider must also be filtered to remove harmonics. This additional time delay (phase shift) of the pilot with respect to the 38 kHz information must be compensated to have optimum channel separation.

A significant improvement on the original switching concept is shown in Fig. 7. As mentioned earlier, the higher order terms of the square wave-driven switch are responsible for generating the harmonics of the 38 kHz subchannel which must be removed by filtering. By using a "soft switch" to connect back and forth between the L and R channels it is possible to eliminate the lowpass filter and its side-effects.

This is accomplished by using the electrical equivalent of a variable attenuator, shown in the diagram by a potentiometer. The slider is driven from end to end

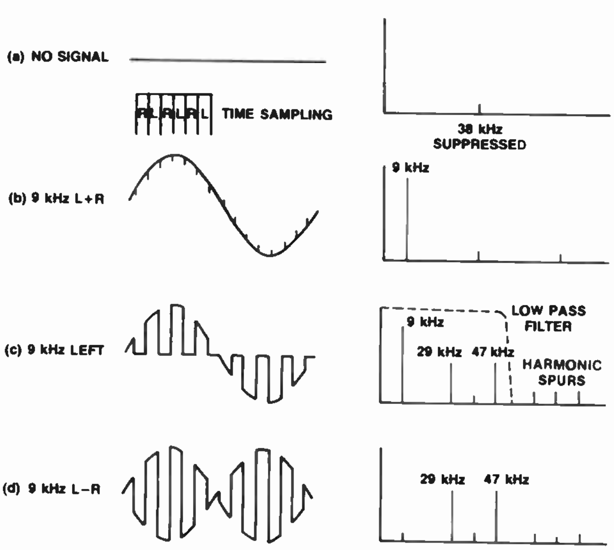


Figure 6. Time domain and frequency domain diagrams of stereo baseband signals.

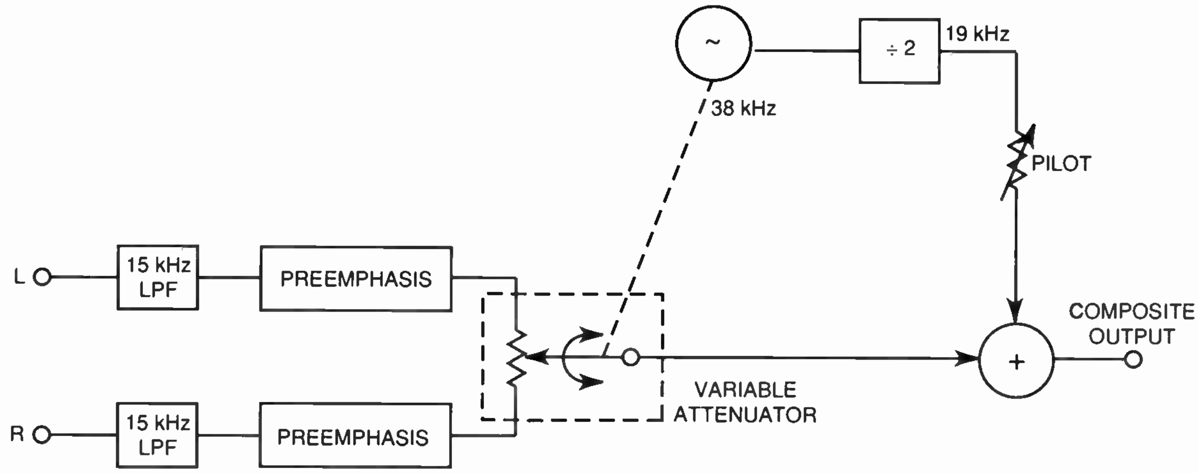


Figure 7. Functional blocks of a time division multiplex stereo generator using a variable attenuator.

of the potentiometer by a sine wave. Since a sine wave is represented only by a single, fundamental frequency, the signal output at the slider has the proper DSBSC characteristics without the harmonics. The equation for the composite signal generated in this way is

$$e = \tfrac{1}{2} \sin \sigma t \, (L + R \text{ audio}) +$$

$$1/\pi \, [\sin(\phi + \sigma)t + \sin(\phi - \sigma)t] \quad (2)$$
(38 kHz DSBSC)

As the equation shows, only the fundamental sidebands of 38 kHz are present in the sampled signal, along with the main channel component. Like the fast-switching TDM system, the $L+R$ and $L-R$ channels are generated in one operation so that the circuit remains relatively simple. No filter of the output is required, provided that the 38 kHz sinewave is free from harmonics and the variable attenuator has good linearity.

FMX™ Stereo Noise Reduction

The FMX™ system was developed to reduce the noise in stereo reception by compressing the stereo $L-R$ difference signal before transmission and expanding it to its original dynamic range at the receiver. The system utilizes the standard FM stereo composite format with the addition of a new subcarrier for the compressed signal. The added DSBSC subcarrier is in phase quadrature with the regular stereo subcarrier at 38 kHz, as shown in Figure 8A.

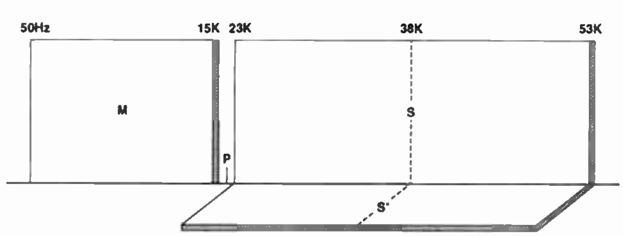


Figure 8A. Baseband spectrum.

The compression characteristic shown in Figure 8B provides linear operation at a fixed gain of 14 dB over much of the program dynamic range, while at the highest modulation levels the FMX program signals are attenuated to permit delivery of full modulation by the standard stereo composite signal. The FMX receiver has a quadrature channel demodulator, expander, and matrix that restores the compressed $L-R$ program to its original dynamic range and combines it with the $L+R$ signal to provide stereo with reduced noise.

Digital Stereo Generation

The major manufacturers of stereo generators have recently introduced all-digital composite generator designs. These units carry out the same functions as analog stereo generators, but with the flexibility and consistency commonly expected of digital audio sys-

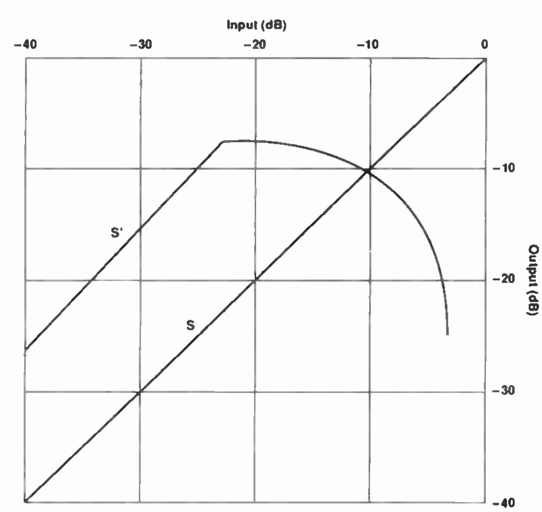


Figure 8B. Compression characteristic.

tems. Some generators also perform all the audio compression and peak limiting functions in the digital domain. This chapter is confined to a discussion of digital stereo generation techniques.

In a digital implementation it is made easier if the output sampling rate is a binary multiple of either of the two steady state vectors in the composite signal: the stereo pilot or the $L-R$ subcarrier. For example, an oversampling rate of 304 kHz is sometimes employed; this is eight times the subcarrier frequency (16 times the pilot).

The input sampling rate is also an important consideration. The reader will note that switching-type stereo generators, discussed earlier, operate at only 38 kHz. While higher rates do not yield any additional improvement in fidelity, it is desirable to use a higher rate throughout the generator system.

While Digital Compact Discs© are fixed at 16-bit quantization precision, most digital stereo generator designs use more bits per sample. This is because analog-to-digital (A/D) conversion devices do not perform to the theoretical limits of 16 bits, and because the mathematical precision of the processing architecture can produce rounding errors that contribute noise and distortion to the processed signal. Eighteen to 24 bit data paths are generally chosen.

Analog audio for each channel is passed through an anti-aliasing lowpass filter and is then A/D converted into digital data streams. At this point, generators with audio processing carry out algorithms that gain-control, preemphasize, limit and filter the digital audio stream. The dynamic effects are very similar to analog processors even though the treatment is entirely digital.

The processed L and R signals are then numerically matrixed using simple addition into $L+R$ and $L-R$ channels. The digitized 38 kHz sinewave is derived from a look-up table and digitally multiplied by the $L-R$ channel to produce the 38 kHz double-sideband suppressed carrier (DSBSC) subchannel. The 19 kHz pilot is generated by another look-up table that is locked via software control so that the phasing relative to the 38 kHz subchannel is perfect. The pilot and 38 kHz subchannel are finally summed and applied to a digital-to-analog converter to form the complete composite stereo signal. Some manufacturers provide a separate digital output port for future digital FM exciters.

Stereo Decoder Circuits

Stereo FM receivers include a circuit to convert the multiplexed composite signal at the FM detector into the original left and right audio channels transmitted by the FM station. There are at least as many ways to decode the stereophonic signal as there are ways to encode (generate) the composite signal. In practice, only one type of decoder is commonly used: the so-called phase-locked loop (PLL) integrated circuit.

The circuit in Fig. 9A is seldom used, but is shown for comparison. It is the closest complement to stereo generators using frequency division multiplexing. At the input, the composite signal is split with three equal time-delay filters into the main $(L+R)$ channel, pilot signal, and stereo $(L-R)$ subchannel. Next, the pilot is doubled to 38 kHz. This regenerated carrier is reinserted into the double-sideband AM signal from the subchannel filter. The AM signal is demodulated to yield the $L-R$ (difference) audio. Finally, the $L+R$ and $L-R$ signals are combined in a sum and difference matrix to produce the original left and right audio channels.

Because of the costly filters needed to separate the composite spectrum, the frequency division multiplex circuit is not used in consumer equipment. Similar circuits have been used in broadcast modulation monitors, where metering of the separate channels is required.

The circuit shown in Fig. 9B is universally used due to its simplicity, high performance, and low cost. While this stereo decoder is commonly referred to by its internal phase-locked loop (PLL), its performance is really distinguished by its time division demultiplexer (shown in the dashed box as a toggle switch).

Following a buffer amplifier, the composite baseband signal is sampled by a phase-locked loop within the integrated circuit (IC). A voltage controlled oscillator, usually running at 76 kHz (four times the pilot frequency), is held in phase with the pilot by a reference signal from the phase comparator and loop filter. When divided by two, the result is a square wave at 38 kHz having nearly perfect duty cycle (high and low states have equal timing) and very fast rise and fall times. This signal is ideal for driving the output audio switcher (demultiplexer). This stage is a transistor matrix designed to rapidly transfer the composite baseband to the left and right audio output in time with the switch in the station's stereo generator. Fast, clean audio switches are relatively easy to make and do not drift.

Because PLL stereo decoders normally use square wave switching, the circuit is able to demodulate baseband signals that are odd harmonics of 38 kHz. The third harmonic (114 kHz) is most troublesome, since noise and spurious signals near this frequency are shifted to the audio baseband, as is the 38 kHz stereophonic subchannel. Engineers should be watchful of the frequency band centered on 114 kHz in their transmitted signal since audible noise may occur in consumer receivers.

Recent stereo decoder designs have reduced sensitivity to energy outside the composite baseband. One approach utilizes a second composite audio toggle switch operated at 114 kHz. The demodulated product is inverted and mixed equally with the 38 kHz switching outputs, canceling the response to signals in the 114 kHz range. The other approach, called a Walsh demodulator, applies a properly timed stair step imitation of a sinusoid to a digital multiplier in the output signal path.[5] This reduces sensitivity to third and fifth harmonics, as well as adjacent channel noise and interference, by up to 20 dB.

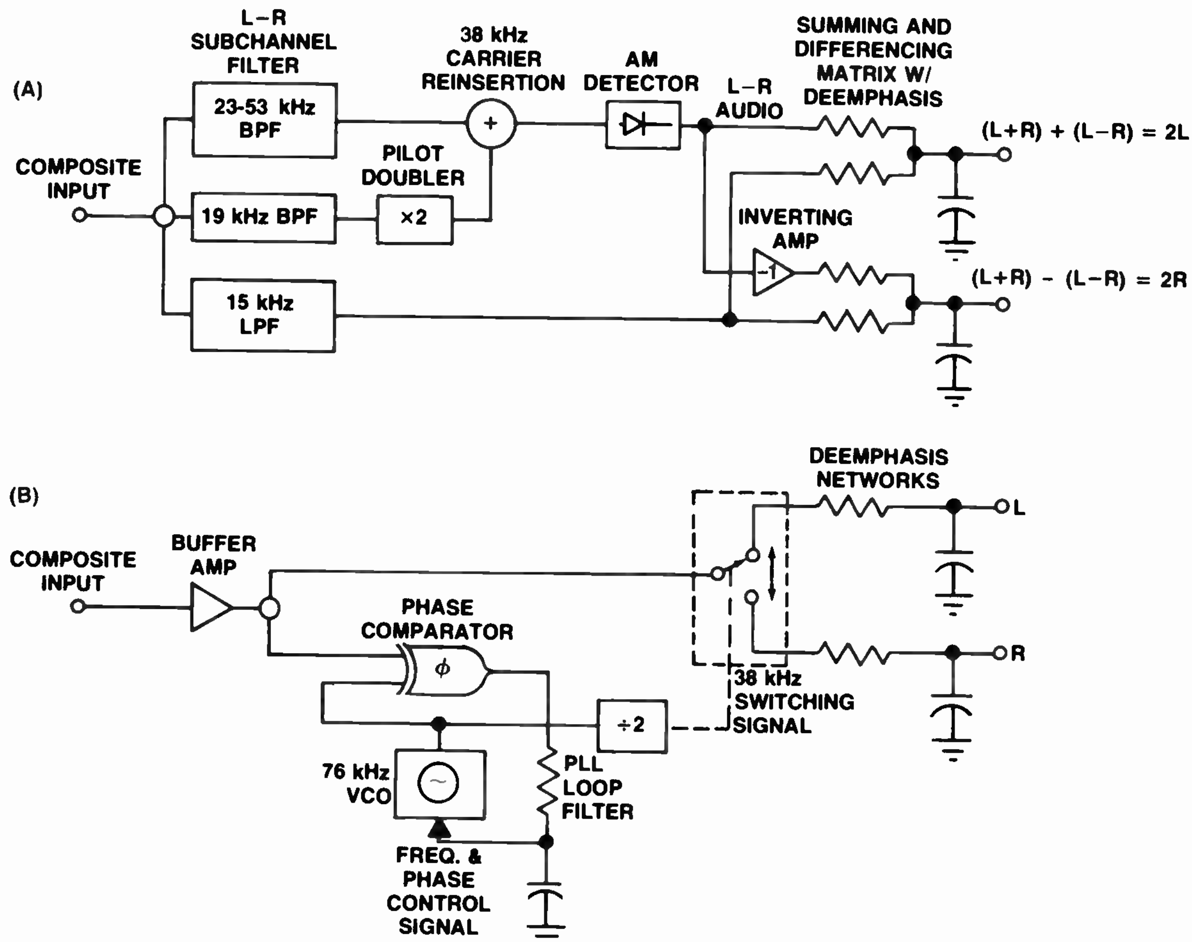


Figure 9. Functional blocks of stereo decoders using (A) L+R and L−R matrixing and (B) phase-locked time-division multiplexing.

FM SCA TRANSMISSION

Background

From its beginning as a broadcast service, people have recognized the potential of FM for multiplexed services. As early as 1940, the FCC permitted multiplex facsimile transmission on FM stations, but not until much later did auxiliary FM services attain a wide acceptance among FM broadcasters.

In 1955 the FCC established the Subsidiary Communications Authorization and created an entirely new industry. The original intent of the authorization was to permit programming of background music to offices, stores, restaurants, etc., where it was uneconomical to provide this service via telephone lines. For many commercial FM stations, the SCA operation became a major source of revenue which enabled them to survive in the 1950s.

By the early 1980s, improvements in transmitter and receiver technology and the desire for new revenue prompted commercial and noncommercial broadcasters to seek changes in the SCA rules. In a series of rulemakings in 1982 and 1983 the Commission made numerous changes to expand technical opportunities and reduce legal regulation.[6] These changes extended the baseband frequency limit from 75 kHz to 99 kHz, allowed any type of subcarrier modulation to be used, changed the subcarrier injection requirements to permit multiple services, and increased limits for the total modulation during SCA multiplex operation to reduce main channel modulation loss.

FCC Rule Requirements For SCA Operation

Section 73.319 of the FCC Rules sets forth the technical standards for FM multiplex subcarriers. Users should note, however, that the Commission does not set standards for minimum SCA subcarrier performance. This is left for the broadcaster or lessee of the service to determine. In its rules, the Commission defines the transmission conditions under which subcarriers may be operated, to minimize interference to

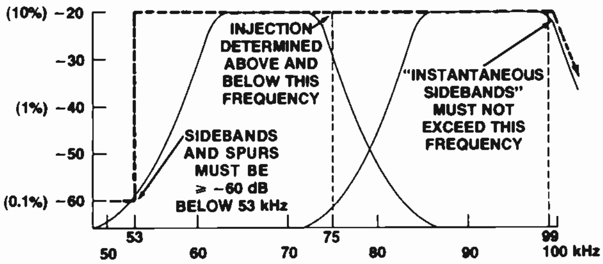


Figure 10. Injection, channel bandwidths and spurs limits for SCA operation when stereo is transmitted. Two possible subcarriers (at 67 and 92 kHz) are shown.

the main channel, the stereophonic subchannel, and to other FM stations.

Subcarrier injection and bandwidth are considered here only with stereophonic operation, since this is by far the most common FM mode. Fig. 10 shows the upper portion of the composite baseband, from approximately 50 kHz to 100 kHz. This is the same as the upper portion of Fig. 2 with the addition of two hypothetical subcarriers centered at 67 kHz and 92 kHz. Any number of SCA subcarriers may be operated in this frequency range provided that total bandwidth and injection limits are met.

On the baseline are several frequency markers at the upper and lower frequency limits of the two subcarrier ranges. The actual frequencies of the subcarriers are not specified, but are referred to as "*fc1*" and "*fc2*".

Alongside each subcarrier are arrows marking the level in dB below 100% modulation of the main carrier. For example, −20 dB marks the injection at the center frequency of both subcarriers. Since 10% is the maximum injection permitted under the Commission Rules within each SCA subchannel,

$$\text{Injection} = 20 \times \log_{10}(0.1) = 20(-1) = -20\,\text{dB} \qquad (3)$$

At 53 kHz, an arrow marks a level of −60 dB. The FCC requires that any frequency modulation of the main carrier due to the SCA operation shall be at least 60 dB below 100% modulation in the frequency range of 50 Hz to 53 kHz when stereo is transmitted. This figure must include spurious and intermodulation products as well as subcarrier sideband energy.

At 99 kHz, the level of −20 dB is marked, denoting the FCC requirement that "instantaneous sidebands" be restricted within this frequency limit. The Commission has not officially defined instantaneous sidebands, but it is normally considered to be the instantaneous frequency of the subcarrier at its peak deviation (for frequency modulated subcarriers) or the highest sideband frequency (for amplitude modulated subcarriers).

Interference Between FM SCA Subcarriers

In practice, two subcarriers should be separated as far apart in frequency as possible, while observing the limit of spurious energy below 53 kHz and the instantaneous sidebands at 99 kHz. While the FCC Rules are silent on the choice of SCA frequencies, 67 kHz and 92 kHz have become the defacto standards for FM subcarriers. The first frequency was adopted when the original stereophonic standards restricted the subcarrier spectrum between 53 kHz and 75 kHz. The second frequency was recommended to situate the instantaneous sidebands below 99 kHz while maintaining a safe separation from a 67 kHz subcarrier.[7]

Some overlap of the subcarrier sidebands does occur, as depicted in Fig. 11. This does not cause significant interference between the two SCA subchannels when the systems use frequency modulation. The following is a list of bench test results of a standard table model SCA receiver:

1. Unweighted 67 kHz SCA SNR at 65 dBf, referred to 5 kHz deviation: 57.5 dB.
2. Same, with 4 kHz tone modulation of 92 kHz subcarrier at 7 kHz deviation: 56.5 dB.
3. Same as first case with tone modulation of main channel at 67.5 deviation: 52.5 dB.

The above data show that crosstalk is greater from main-to-subchannel than from a 92 kHz subcarrier into the 67 kHz subchannel demodulator.

Other Types of Analog Subcarrier Modulation

Other frequencies and total number of subchannels are permissible, according to the occupied bandwidths and interference margins required. A hypothetical SCA baseband spectrum combining five amplitude-modulated subcarriers is shown in Fig. 10. Here, a variety of bandwidths and injection levels are used.

In Fig. 12 note that single-sideband suppressed carrier (SSBSC) AM systems are used for the four

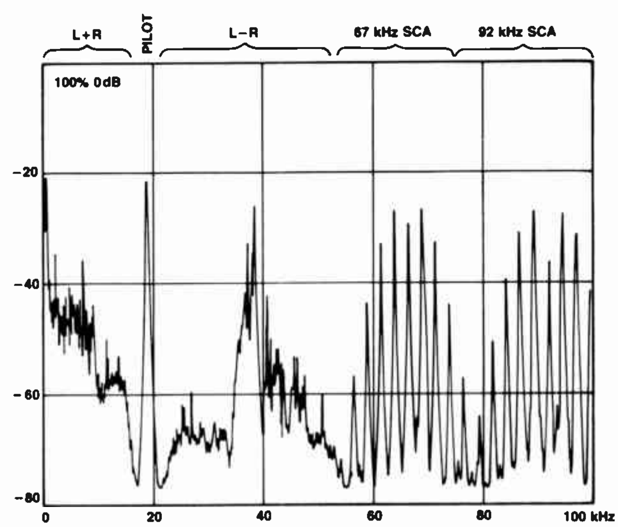


Figure 11. Composite baseband of FM station with 67 kHz and 92 kHz FM-SCAs, modulated with 2.5 kHz tone at 5 kHz and 7 kHz peak deviation, respectively. Note overlap of sidebands between subcarriers. Station is carrying stereo programming.

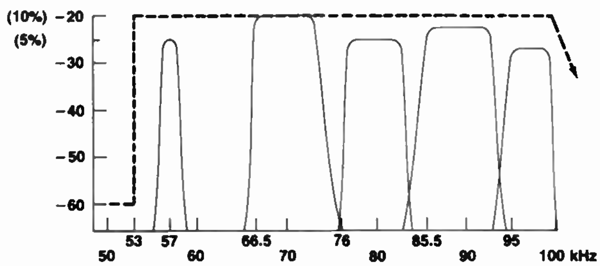


Figure 12. Carrier envelope spectrum, DSB-AM and SSB-AM SCA subcarriers. Vertical markings are dB referred to 75 kHz deviation.

other subchannels. On the first (at 66.5 kHz), 10% peak envelope injection is being used, while the other three use levels which bring the total to 20%.

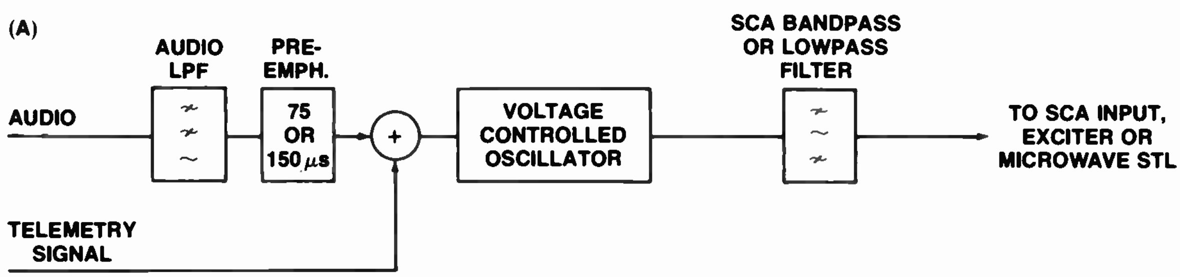
At 57 kHz, the envelope of a narrow subcarrier is shown. Actually, standards exist to interleave two subcarriers within the same channel. One is called the Radio Data System (RDS) and is in use by a large number of Western European countries for personal paging, station and program identification, emergency alerting, etc. This 57 kHz system is currently used for subcarrier paging in the U.S. The modulated carrier spectrum resembles suppressed-carrier double sideband AM that is locked to the third harmonic of the pilot. Digital data rate is 1187.5 bits/second and nominal deviation of the carrier is 2.67% (±2.0 kHz).

Another 57 kHz system was also developed in Europe for Automotive Road Information (ARI). It is used to identify stations that broadcast traffic information and to activate specially-equipped FM car radios during traffic announcements. ARI is a slow-rate signaling system with very narrow bandwidth. The 57 kHz ARI carrier is centered within the suppressed-carrier region of the RDS signal (if RDS is transmitted).

Examples of Analog SCA Operation

Fig. 13A shows the functional blocks of a standard aural SCA generator using frequency modulation. At the input, audio is lowpass filtered to limit the modulation sidebands that could interfere with the stereophonic subchannel or an upper SCA subchannel.

A 5 kHz audio cutoff is frequently used so that fidelity compares favorably to unequalized telephone service. Using a Bessel function analysis one can determine the peak deviation which may be used to maintain FM sidebands below 53 kHz within FCC limits.



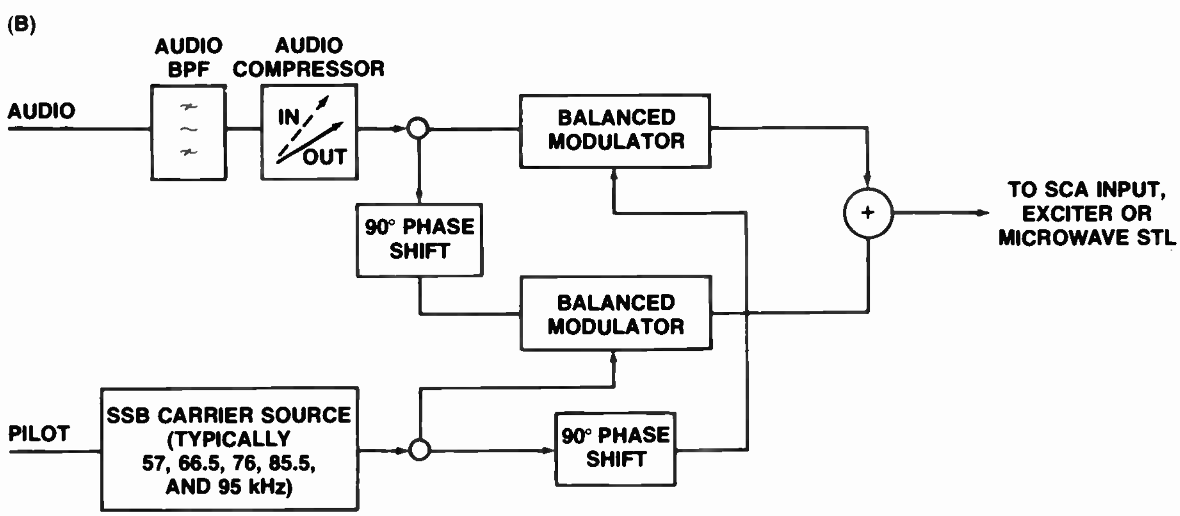


Figure 13. Functional blocks of (A) a frequency-modulated SCA subcarrier generator, and (B) a single-sideband AM subcarrier generator.

TABLE 1

Sinusoidal SCA Modulating Frequency	Maximum SCA Peak Deviation
3.5 kHz	5.0 kHz
5.0 kHz	3.5 kHz
7.5 kHz	2.0 kHz

For the above protection requirements, equipment manufacturers often recommend 3.5 kHz or 4.0 kHz peak deviation with a 5 kHz low-pass filter when stereo is transmitted. In practice, slightly higher deviation may be used since the calculations assume single tone modulation which has greater modulation power than program audio.

Preemphasis follows the low-pass filter in the sample FM SCA generator. A value of 150 μSec is commonly used to combat the rising high-frequency noise characteristic in the FM SCA channel. Unfortunately, this extreme boost (almost 17 dB at 5 kHz, compared to 400 Hz) requires substantial amounts of peak limiting to allow full modulation at low and middle audio frequencies.

The processed audio signal is fed to the FM modulator, usually with a provision for subaudible (less than 30 Hz) channel for telemetering return. The telemetry channel takes advantage of the extended low frequency response capability of FM subcarriers by sending tone signals in the 10 Hz to 30 Hz range at a level of 15 dB to 25 dB below reference peak deviation. A high-pass filter is inserted ahead of the audio input when subaudible telemetry is used to avoid interference to the low frequency channel.[8]

Ideally, no filtering of the frequency-modulated subchannel is required. However, some of the FM generator circuits produce spurious or harmonic energy. For this reason, a bandpass or low-pass filter is used which passes all the significant sidebands but suppresses spurious signals to acceptable levels. The modulated subchannel signal is amplified to a level required by the subcarrier input of the FM exciter or studio transmitter link equipment.

A simplified block diagram of a SSBSC AM subchannel generator is shown in Fig. 13B. The input audio is bandpass filtered to control subchannel bandwidth (to 5 kHz, for example), and may be compressed according to specific amplitude and spectral standards. A complementary expander circuit in the receiver restores the original program dynamics and reduces noise in the channel. Large amounts of preemphasis are not as important since the spectral characteristic of the noise is flat with a linear modulation system, not rising with frequency as in angular modulation systems.

The processed audio signal is fed to a pair of balanced modulators, one path going through an allpass filter which shifts the audio phase 90° throughout the required audio range. The balanced modulators are driven by a subcarrier frequency source which may be frequency-locked to a harmonic of the stereophonic pilot (and its half-frequency of 9.5 kHz, as shown). One of the modulators is driven by a carrier signal which is shifted 90°. When the two modulated carriers are summed, cancellation of the lower (or upper) sidebands creates the desired single-sideband signal.

Since SSBSC AM requires a carrier for detection, the pilot provides an excellent source for reinsertion at the receiver. With monophonic audio, no pilot is transmitted, so another carrier reference must be used. One approach is to carry the 19 kHz reference at a reduced level which will not trigger the stereophonic decoders in FM receivers.

SCA DATA SYSTEMS

SCA subchannels have characteristics that are quite favorable for transmitting data: while the channels are noisy for high fidelity audio services, they can provide good to excellent data error performance; and moderately fast data rates are possible without complicated or expensive receivers. Combined with the high power and elevation of FM broadcast stations, SCA data systems may be operated over large service areas.

There are many methods of putting data on an SCA subchannel, the simplest being connection of a telephone-type data MODEM to the audio terminals of the SCA generator. This method is commonly called *audio-frequency-shift keying* (AFSK).

The cost and complexity of this form of SCA data broadcasting is low. However, this technique occupies the bandwidth of a standard audio subcarrier and the highest practical speed is about 2,400 bits/second. This may be acceptable where the hardware must be "off the shelf" and data capacity of baseband spectrum efficiency is not a consideration.

Direct Data Modulation

Direct data modulation eliminates the intermediate steps that convert the data into audio signals, allowing the subcarrier to be directly modulated by the binary data. This efficiency allows a greater amount of data to transmitted within the available SCA subchannel bandwidth.

In *non-return-to-zero frequency-shift keying* (NRZ/FSK), the two binary states are represented by two different subcarrier frequencies. For example, a "0" could be 65 kHz and a "1" 69 kHz. FSK is the most popular form of direct data modulation at this writing, probably because of its robust performance, simplicity and low cost. Using careful design, a modulated bandwidth similar to that of background music services (i.e., about ±16 kHz at −25 dB below subcarrier injection) maximum speed 19 Kbits/second are currently possible.[9]

Data FSK systems must use filtering methods to control occupied bandwidth, since the rapid frequency transitions generate high-order sidebands that could interfere with adjacent subchannels. Filtering is accomplished by either or both of the following ways, shown in Fig. 14:

1. Where a pair of oscillators are alternately toggled by the data states, filtering must be done to

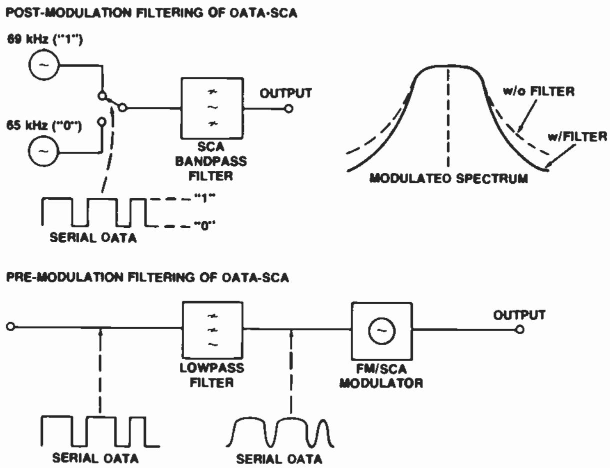


Figure 14. Two basic methods for generating and controlling the bandwidth of data subcarriers used for SCA.

the output spectrum to directly reduce occupied bandwidth.

2. Where analog FM modulators are used, filtering the slew rate of the switching signal tends to reduce high-order sidebands indirectly.

The spectrum of the composite baseband of a Washington, D.C. FM station is shown in Fig. 15, carrying a 4800 baud data signal. The spectrum was generated by a specially designed SCA generator employing both pre- and post-modulation filtering.

Phase-shift keying (PSK) is another technique for transmitting high-speed data. With PSK, one phase of the carrier represents one binary state and a second phase (usually 180° apart) is used for the second state. High speed data transmission is possible, up to 56,000 bits/second, although most of the 53 kHz to 99 kHz bandwidth might be required and receiver/decoders would be much more expensive than those for moderate speed FSK systems.

MONOPHONIC, STEREOPHONIC, AND SCA PERFORMANCE

While monophonic, stereophonic and various SCA services are conveyed by the same FM carrier, performance varies considerably due to the relative carrier deviation of each and the particular frequency span occupied within the composite baseband. The performance of a channel may be expressed in terms of demodulated signal to noise ratio for a given radio frequency carrier level, or, conversely, coverage range for a given signal-to-noise ratio. The purpose of this section is to summarize and illustrate these relations but not present the lengthy derivations for the equations.

The performance of a conventional FM receiver in the presence of random (thermal) noise is commonly judged on the basis of the variation of the output signal-to-noise power ratio as a function of the carrier-to-noise power ratio contained within the receiver bandwidth, usually determined by the IF bandwidth. An estimate of the IF bandwidth required for transmission of (monophonic) broadcast FM is given by[10]

$$\beta_{IF} = 2(F_{dev} + 2f_m) \tag{4}$$

where: F_{dev} = The peak frequency deviation (75 kHz)
f_m = The highest baseband modulating frequency (15 kHz)

This β_{IF} figure is rounded to 200 kHz for later use.

The equivalent input noise power in a bandwidth β launched into a noise-free (cool) receiver is

$$rB = 4 \times 10^{-21}\,\beta f_N \tag{5}$$
$$= 8 \times 10^{-16} f_N \quad \text{(Watt)}$$

where: f_N = The receiver noise figure
β = Nominal noise bandwidth in Hz

The carrier power is

$$C = (V_{\mu V^2}/R) \times 10^{-2} \quad \text{(Watt)} \tag{6}$$

where: $V_{\mu V^2}$ = the RMS input voltage in microvolts
R = the receiver r.f. input resistance

For R = 75 ohms the carrier to noise ratio in decibels is

$$CNR = 10\log_{10}(C/rB) \tag{7}$$
$$= 12.2 + 20\log_{10} V_{\mu V} - 10\log_{10} f_N$$

For example, for a 10 dB noise figure the carrier-to-noise ratio (CNR) is 0 dB for an input voltage of 0.77 μV over 75 ohms (1.55 μV over 300 ohms).

For a half-wave dipole antenna matched to a 75 ohm receiver, the input voltage is related to the field strength as

$$V_{\mu V} = 48.5\, E_{\mu V}/f_{MHz} \tag{8}$$

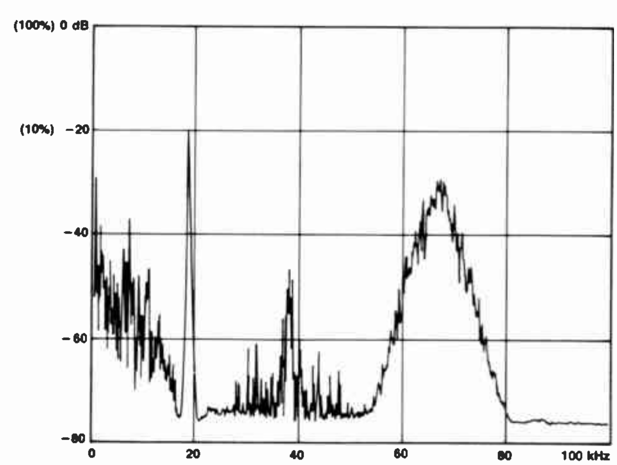


Figure 15. Composite baseband of an FM station carrying a 4800 baud frequency-shift-keyed data subcarrier. Center frequency of the subcarrier is 67 kHz, with an injection of 9%. Station is carrying stereo programming.

where: $E_{\mu V}$ = The field strength in microvolts
f_{MHz} = The carrier frequency in MHz[11]

Thus, for 98 MHz (middle of the FM band)

$$V_{\mu V} = \frac{E_{\mu V}}{2} \quad (9)$$

With the FCC F(50,50) field strength curves, it is possible to relate $E_{\mu VV}$ to the distance to the transmitter, effective radiated power, and antenna heights. Ultimately, it is possible to predict the carrier-to-noise ratio for a given broadcast system. For a 200 kHz nominal bandwidth, the CNR in dB can be shown to be

$$CNR = E_o + P + G - F + 20\log_{10}(h/30 \times 98/f) + 6 \quad (10)$$

where: E_o = Field strength for 1 kW ERP at 1000′ HAAT, in dBμ V/m
P = ERP in dBkW
G = Receiver antenna gain relative to dipole, in dB
h = Receiving antenna height above ground, in feet
F = Receiver noise figure in dB
f = Frequency in MHz

For example, if $F = 10$, $P = 10$, $G = 0$, $h = 30$, and $f = 98$, then $CNR = E_o + 6$.

Next, the signal-to-noise ratio is established by the carrier-to-noise ratio

$$SNR = 40 + 10\log_{10} CNR\ dB \quad (11)$$

This expression assumes a peak deviation of 75 kHz, a deemphasis corner frequency of 2.1 kHz (75 μS) and a nominal intermediate frequency noise bandwidth of 200 kHz.[12]

From the above, a carrier-to-noise ratio of 0 dB would result in a signal-to-noise ratio of 40 dB. However, standard FM demodulators typically exhibit a "threshold" at about 12 dB CNR, below which the noise floor changes from a "hiss" to a raspy kind of noise with noticeable "clicks" jutting above the noise. For this reason, the monophonic channel actually deteriorates below a signal-to-noise ratio of about 52 dB.

A well-known effect of FM demodulation is that the noise voltage in a narrow band (for example, 1 Hz) increases with frequency. This has important implications for systems which employ subchannels, since both the bandwidth and center frequency affect the signal-to-noise ratios. For example, the stereophonic $(L-R)$ subchannel shown in Fig. 2 extends nominally from 23 kHz to 53 kHz and occupies twice the bandwidth of the main $(L+R)$ channel. When the noise of the $L-R$ subchannel is combined during stereo decoding, the unweighted noise floor is 23.1 dB higher than the monophonic system under the same conditions.[13]

The signal-to-noise ratio of an SCA subcarrier, relative to the main channel, is a function of a number of variables and is

$$N_s/N_o = (2f_c^2/m^2D^2)\frac{f_1^3(b_1 - \arctan b_1)}{(f_o^3(b_o - \arctan b_o)} \quad (12)$$

where: f_c = Subcarrier frequency
m = Injection level as a percent of 75 kHz
D = Deviation of FM subcarrier
f_o = −3 dB frequency of deemphasis in main audio channel (2.12 kHz)
f_1 = −3 dB frequency of deemphasis in SCA audio channel (1.06 kHz)
b_o = Main channel audio output bandwidth (15 kHz)/f_o
b_1 = SCA channel audio output bandwidth (5 kHz)/f_1

An example of the excess noise in the 5 kHz SCA audio channel compared to the 15 kHz main channel is given below, if $m = 0.1$ (10%). Note that the noise level of a 67 kHz subcarrier with 5 kHz deviation and 92 kHz subcarrier with 7 kHz deviation are nearly equal. Comparative tests by the author have indicated that susceptibility to impulse noise, e.g., gasoline engine ignition radiation, is greater at the higher subcarrier frequency even though gaussian noise levels are approximately the same. This may be due to a higher noise threshold in the 92 kHz demodulator (due to a wider predetection bandwidth than the 67 kHz demodulator) or to differences in the spectral distribution of impulse noise compared to gaussian noise in the IF (composite) demodulator.[14]

f_c	(SCA freq., kHz)	67	92	92
D	(deviation, kHz)	5	5	7
N_s/N_o	(above main, dB)	36.4	39.2	36.2

Fig. 16 graphs the relative subcarrier noise level against increasing frequency of a subcarrier, compared

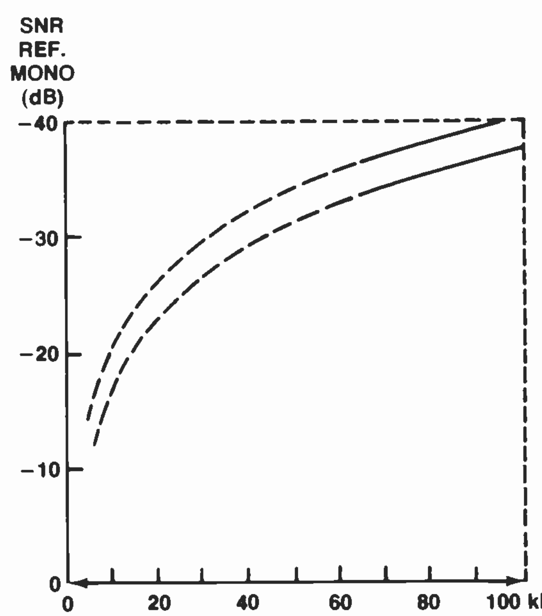


Figure 16. SNR of FM/SCA audio (5 kHz bandwidth) versus subcarrier frequency, compared to monophonic reception, referred to 5 kHz and 7 kHz deviation.

to the main (monophonic) channel for the same receive conditions. Shown are two peak deviations (5 kHz and 7 kHz) for FM SCA, and SSBSC AM SCA with 10% peak envelope injection. The graph segment representing 7 kHz deviation is dashed below 70 kHz since excessive interference to the $L-R$ subchannel is possible in this frequency range when stereo is transmitted.

Single Sideband AM Subcarriers

The conditions under which standard SCA subcarriers are operated are quite close to those existing with narrow band frequency modulation. For example, when a maximum audio bandwidth f_m of 5 kHz and peak deviation f_d of 5 kHz are used, the modulation index is 1.

$$i = f_d/f_m = 5/5 = 1 \qquad (13)$$

After frequency demodulation, the signal-to-noise power ratio (SNR) will be[15]

$$SNR = \frac{3 \cdot i^2}{CNR} = \frac{3 \cdot 1^2}{CNR} \qquad (14)$$

where CNR is the carrier-to-noise power ratio in the bandwidth occupied by the modulated signal. The noise power bandwidth B_n is difficult to predict for two reasons. First, the (usually) inexpensive SCA predetection filter has sloping skirt selectivity, thus departing from the ideal rectangular shape assumed in the calculation. Second, the actual bandwidth may vary, depending on the amount of audio distortion to be tolerated. The extremes of choice may be:

$$i \le 1, \quad B_n = 2f_m = 10\,\text{kHz} \qquad (15)$$

$$i > 1, \quad B_n = 2(f_d + f_m) = 20\,\text{kHz} \qquad (16)$$

A noise power bandwidth of 15 kHz is a compromise value that happens to represent the measured performance in some common FM SCA receivers.

In single sideband systems, one is simply frequency-shifting the modulation passband to a particular frequency in the composite baseband. There is no inherent noise reduction effect; the recovered audio signal-to-noise ratio is approximately equivalent to the noise contained within the subchannel bandwidth.

While the FM subcarrier system improves SNR, its greater bandwidth admits about three times the noise power. At a modulation index of one, FM improvement is approximately 10 log(1) or 0 dB for an ideal detector. The use of preemphasis and deemphasis improves the FM system SNR without significantly affecting the subcarrier's occupied bandwidth. The use of 150 µS deemphasis, common in FM subcarrier demodulators, reduces RMS noise in a 5 kHz channel by 8.8 dB. The practical improvement in signal-to-noise ratio measured by the author is more typically 4 dB. This is because the extreme high-frequency boost is not a good fit to the program material.[16]

Since single sideband modulation is spectrally efficient, even when compared to narrow band FM, it is practical to divide the 53–99 kHz range into a number of independent subchannels. In the one system, five subchannels each having a 5 kHz baseband are operated at 57 kHz, 66.5 kHz, 76 kHz, 85.5 kHz and 95 kHz. These are shown in Fig. 12.

20% peak injection is the maximum permitted by the FCC Rules when stereo is transmitted, therefore, each subchannel would receive 4% injection (20/5 = 4). It is reasonable to assume that slightly higher injection could be permitted for each subchannel due to the random interleaving of the total subchannel injection. If 5% were employed, then, the SNR would be lower than that for a channel using 10% peak injection.

$$\text{change in } SNR = 20\log(10/5) = -6\,\text{dB} \qquad (17)$$

Assuming an approximated improvement of 4 dB for a standard FM subchannel, and a reduction of 6 dB for one of five single sideband subchannels, there is an estimated difference in SNR, assuming the same center frequency of both subchannels and flat noise density over this spectrum.

$$SNR_{FM} - SNR_{SSB} = 4 - (-6) \approx 10\,\text{dB} \qquad (18)$$

The power density from an AM demodulator remains uniform across the baseband (unlike FM demodulators, in which noise density rises in proportion to frequency). Therefore, preemphasis/deemphasis is not as necessary to combat noise. The use of a complementary noise reduction system (compandor) can significantly improve the signal-to-noise ratio of either system.

Using the above estimates, the performance of the (uncompandored) single sideband and (deemphasized) FM subchannels is compared to main channel and stereo in Fig. 17. The coverage radii are based on the FCC F(50,50) field strength predictions for an FM station having 10 kW ERP, 1,000 feet height above average terrain, and a dipole receiving antenna at 30 feet above ground level into a receiver having a 10 dB noise figure.

SCA-To-Stereo Interference Considerations

Introduction of new signals into the modulating baseband of an FM station requires attention to their possible interactions with existing baseband signals, as well as its chosen performance characteristics. Several findings regarding the potential for perceptible interference (crosstalk) are summarized below.

FM stereo receivers having stereo decoders with diode switching bridges driven by a high-level 38 kHz sinusoid (usually identified by discrete component construction and tuned transformer stages) were responsible for the original "birdie" or beat-note interference in early SCA/Stereo broadcast systems. For historical perspective, the sources of this interference were:

- Nonlinearity in the diode bridge due to the conduction characteristics of the diodes and the sinusoidal drive signal
- Imbalance in the center-tapped transformer driving

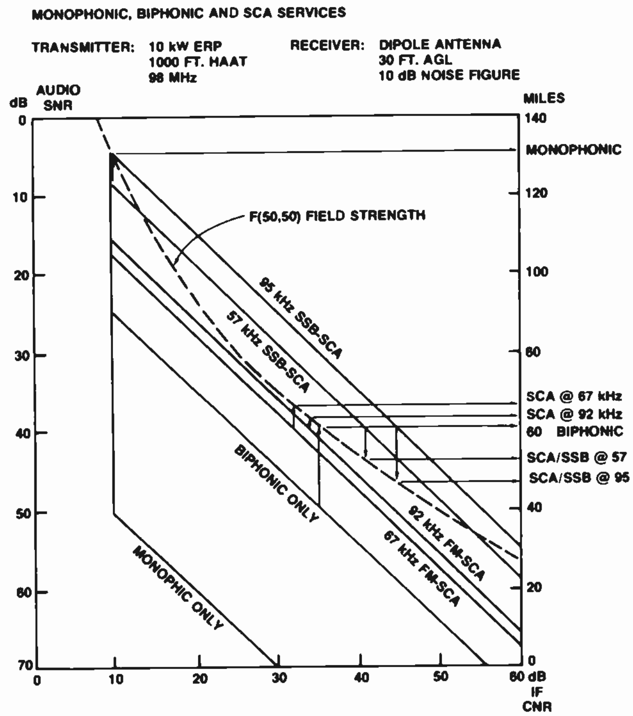


Figure 17. Predicted service distances for 50 dB monophonic and biphonic SNR, and 40 dB SCA SNR. To use, find intersection of desired SNR with service (diagonal) line; next move up (or down) to F(50,50) dashed line; then read across to mileage scale.

the diode bridge and second harmonics of the 38 kHz sinusoid at this point

The result caused a second-order intermodulation (mixing) of twice the 38 kHz switching frequency minus the instantaneous frequency of the SCA subcarrier as

$$9\,\text{kHz} = 76\,\text{kHz}\,[38\,\text{kHz} \times 2] - 67\,\text{kHz} \qquad (19)$$

The product is not 10 kHz as was often believed. Today, the small amount of beat-note interference that results from a 67 kHz subcarrier is 10 kHz, but due to an entirely different mechanism.

Over the past decade these receivers have been entirely replaced by sets with integrated circuit decoders, which have virtually no internally-generated beat-note problem. As a result, the interference to stereo service from SCA operation has dropped significantly.

Contemporary receivers create very little beat-note interference from SCA operation because of the type of stereo decoding done within the integrated circuit. The fact that the circuit contains a phase-locked loop (PLL) has little to do with this improvement. Modern decoders derive their 38 kHz signal by a digital technique, which is well-suited to circuit integration. This results in a 38 kHz square wave of fast rise time and balanced duty cycle (equal time in high and low states). The output circuit is usually a differential transistor pair configured as a balanced demodulator. When driven by the 38 kHz square wave (which has virtually no even-order harmonics), very little mixing is possible which can result in the product shown in Eq. 19.

If SCA beat-notes created within stereo receivers were the only source of interference, there would be virtually no problem with SCA operation. However, experience shows that a minor source of interference remains when certain SCA frequencies are used. The source of this interference is *external* to the receiver: multipath distortion of the radio frequency signal itself.

Large obstructions such as mountains, hills, and tall buildings can reflect VHF broadcast waves well enough to create the simultaneous reception of a direct and one or more reflected-path signals. Upon demodulation the baseband signals will include a combination of linear, second- and third-order distortion products, as shown in Fig. 17. Higher order distortions are negligible when the distortion is small.[17] Hence, these products take the form

$$2A,\ 2B,\ 2C,\ \text{etc.}$$
(where A, B, and C are the fundamentals)
$$A+B,\ A-B,\ A+C,\ A-C,\ B+C,\ B-C,\ \ldots\text{etc.}$$
$$2A+B,\ 2A-B,\ 2B+A,\ 2B-A,\ \ldots\text{etc.}$$

Second-order distortion (frequency sum and difference) is the most troublesome, because new product frequencies are created which may fall within the main and stereophonic channels where they can become audible. This is illustrated in Fig. 18, which shows the baseband of an actual FM station. The only modulation present in the transmitter is the stereo pilot and unmodulated SCA subcarriers at 67 kHz and 92 kHz, plus a small amount of noise from the audio chain which can be seen from zero to 15 kHz and centered around 38 kHz in the $L-R$ subchannel.

Several second-order products are visible in the spectrum graph. This distortion takes the form:

$$f_{SCA1} \pm f_{SCA2} = f_{IM}\ product \qquad (20)$$

OR

$$f_{SCA} \pm pilot = f_{IM}\ product$$

In the case of 67 kHz and 92 kHz subcarriers, significant products are

$$92 - 67 = 25\,\text{kHz}$$
$$67 - 19 = 48\,\text{kHz}$$
$$92 - 19 = 73\,\text{kHz}$$
$$67 + 19 = 86\,\text{kHz}$$
$$92 + 19 = 111\,\text{kHz}$$

While these IM products are not directly audible, some may be once they are demodulated by a stereo decoder:

$$38 - 25 = 13\,\text{kHz}$$
$$48 - 38 = 10\,\text{kHz}$$
$$114 - 111 = 3\,\text{kHz}$$

(The last product could be demodulated by IC stereo decoders using a 38 kHz square wave switching signal. However, sensitivity to a 111 kHz product is rare.)

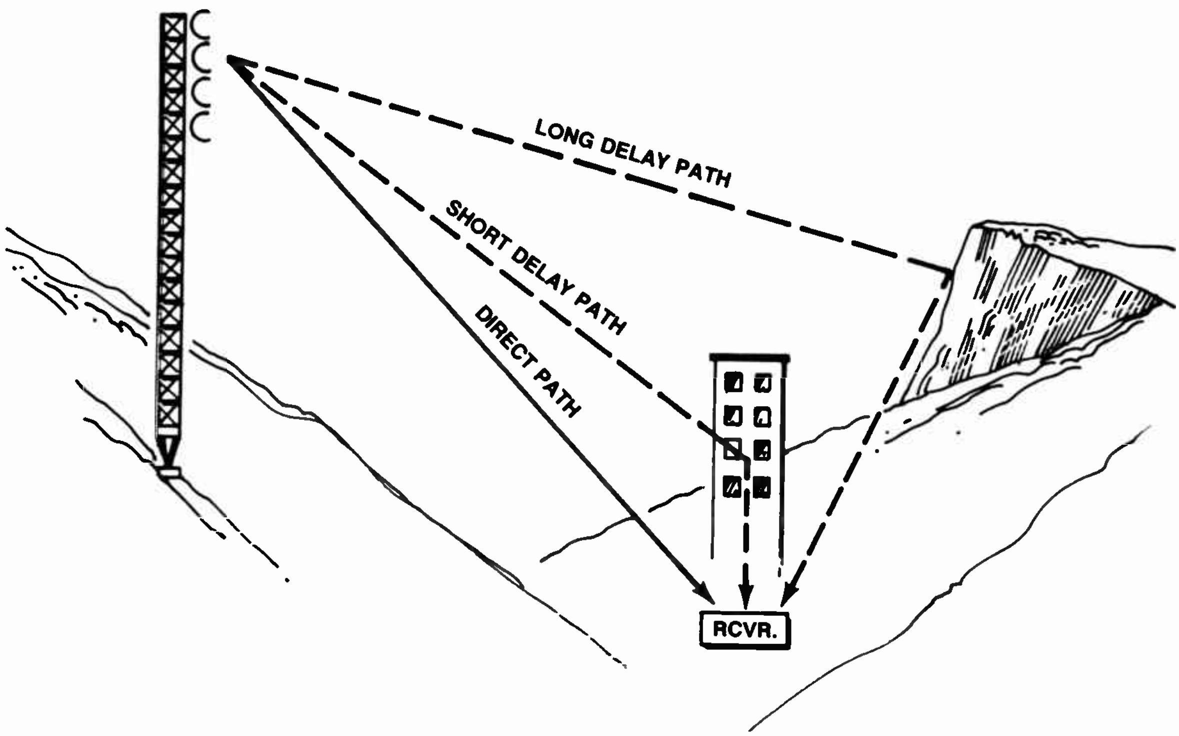


Figure 18. Examples of reflected signals arriving at an FM receiver, along with a direct path RF signal. Distortion of the demodulated FM signal is related to both the direct/reflected signal ratio and delay time of the reflected signal; distortion increases as the signal ratio approaches 1:1 and as the secondary path delay time rises.

Some other general findings about multipath effects are[18]

1. Multipath distortion is almost inversely proportional to the DU (desired-to-undesired path signal ratio) if the ratio is greater than about 10 dB.
2. Baseband signal distortion increases almost proportionally to the delay time up to about 10 μS, then increases with delay time at a slower rate.
3. A high DU ratio is required to suppress the beat-note interference to stereo program reception from an SCA subcarrier (about 20 dB at a delay time of around 5 μS and about 30 dB at around 20 μS).

It should be emphasized that the perceptibility of this form of beat-note interference depends on the loudness and consistency of main channel programming, (i.e., program audio in most broadcast FM stations is capable of masking the beat-note). Generally, only stations which do light amounts of audio processing and broadcast programming with wide dynamic range are even aware of an occasional case of interference.

The *type* of SCA modulation or the subcarrier frequency may significantly affect the perceptibility of any beat-note. For example, properly encoded high-speed data subcarriers using FSK or PSK spread the carrier energy in such a way that it sounds like faint, band-limited noise which is most easily masked. Furthermore, higher subcarrier frequencies will shift the beat-note image out of the audible range: experience has shown that 92 kHz to 95 kHz subcarriers have little problem in this regard.

Crosstalk From Main Channel Program Into SCA Subchannels

Crosstalk remains a problem for audio SCA services, despite the development of low-distortion receivers for

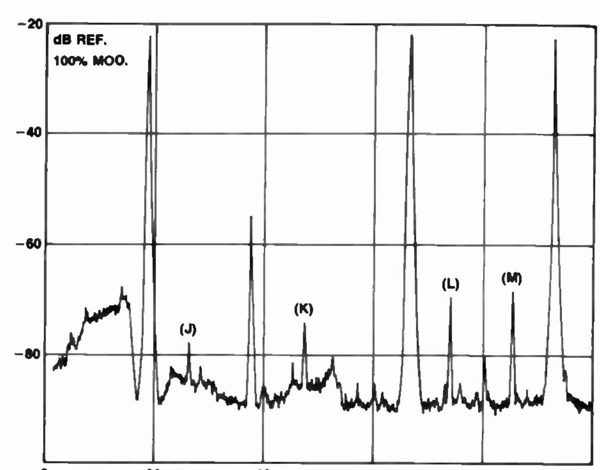


Figure 19. Composite baseband of an FM transmitter system with pilot, 67 kHz, and 92 kHz SCAs at 9 percent injection. IM products are identified by letters and are described in the text.

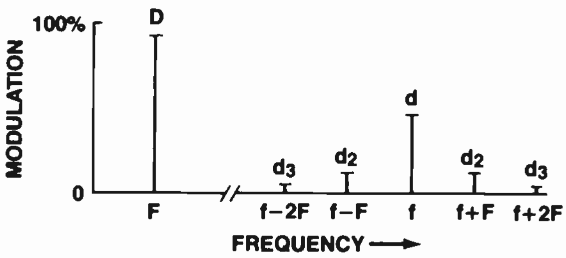


Figure 20. Spectrum of a subcarrier "f" shown with second and third order distortion products resulting from low frequency modulation "f".

SCA use. The cause is the same as SCA-to-stereo beat-notes: multipath distortion.

In the reference case illustrated in Fig. 20, the baseband signal consists of one low-frequency sine wave of frequency F, i.e., 1 kHz, which modulates the main carrier by a deviation D (95% or 71.25 kHz for broadcast FM), and a sine wave subcarrier at a variable frequency f Hz which modulates the main carrier with a very small deviation d, having a modulation index less than 0.3.

Second-order distortion will cause sidebands at frequencies of $f-F$ and $f+F$ with an output injection $d_2(f-F)$ and $d_2(f+F)$. The sum of the sideband amplitudes is

$$d_2(f) = d_2(f-F) + d_2(f+F) \tag{21}$$

The distortion is defined as the ratio of the sum of the sidebands to the amplitude of the subcarrier d_1 is

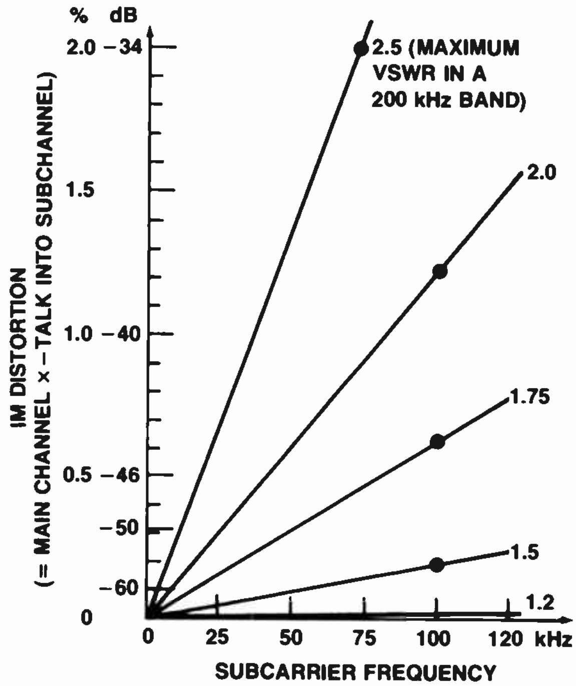
$$\delta = \frac{d_2(f)}{d_1(f)} \tag{22}$$

This distortion is easily measured as a function of the subcarrier frequency f with a spectrum analyzer. If the sidebands are equal, they may, in the extreme, either amplitude-modulate or phase-modulate the subcarrier. In general there is a combination of amplitude- and phase-modulation.

Third-order distortion is similarly defined by the sidebands created at frequencies $f-2F$ and $f+2F$.

While FM SCA detectors usually include amplitude limiting, they are intended to convert any angular modulation, whether program audio or not, into an output signal. Thus any phase nonlinearity in the RF system or multipath will generate second-and higher-order sidebands around the subcarrier, causing main-to-SCA crosstalk at audio modulating frequencies.

Fig. 21 shows the relationship between main-to-SCA crosstalk, subcarrier frequency, and maximum antenna VSWR in a 200 kHz band. It is evident that higher frequency subcarriers require somewhat larger, but not very much larger bandwidth. It has also been determined that antenna matching must be improved as the transmission line becomes longer. This is especially important for higher frequency subcarriers since the phase error is compounded rapidly with an increase in



NOTE: ANTENNA IS MATCHED TO TRANSMISSION LINE AT THE CENTER FREQUENCY.

Figure 21. Distortion due to limited antenna bandwidth vs frequency of subcarrier with max VSWR in 200 kHz band as parameter.

subcarrier frequency. Distortion is proportional to the subcarrier frequency. Thus, for equal distortion, doubling the subcarrier frequency requires halving the reflection coefficient.[19]

ENDNOTES AND REFERENCES

1. Edwin H. Armstrong, "A Method Of Reducing Disturbances In Radio Signalling By a System of Frequency Modulation", *Proceedings of the I.R.E.*, May 1936, Vol. 24. No. 5, pp. 689–740.
2. Electronic Industries Association, "Report on the National Quadraphonic Radio Committee to the Federal Communications Commission", Vol. II, Chapter 1, pp. 21–29.
3. Federal Communications Commission, "Rules and Regulations", 47 C.F.R. Section 73.310.
4. National Association of Broadcasters/Westinghouse Broadcasting & Cable, Inc./National Public Radio, "Increased FM Deviation, Additional Subcarriers and FM Broadcasting: A Technical Report", Aug. 30, 1983, pp. 5–6.
5. S. Takahashi, H. Iida, "Application of Walsh Functions to an FM Stereo Demodulator", J. Audio Eng. Soc., Vol. 33, No. 9, September 1985.

6. FCC BC Docket No. 82–536, *First Report and Order* (Apr. 7, 1983), *Second Report and Order* (Mar. 29, 1984).
7. John C. Kean, "Laboratory and Field Tests of Several FM/SCA Frequencies", National Public Radio *Engineering Report*, Oct. 15, 1981, pp 13–14.
8. Moseley Associates, Technical Manual for Model SCG-8 FM Subcarrier Generator.
9. Eric Small and Eric Stoll, "Data SCA: Some Real World Experiences", Modulation Sciences pamphlet, Aug. 1984.
10. Howard W. Sams, Inc., "Reference Data for Radio Engineers, 6th Edition", 1975, pp. 23–9—23–12.
11. Ibid., NQRC, Vol. II-1, Chapter 1, pp. 31–40.
12. Ibid., NQRC, Supplement to Volume II, Chapter 1, Section 1.4, p. xviii.
13. Ibid., pg. 36.
14. Ibid., Kean, pg. 24.
15. Ibid., Sams, pp. 23–4—23–12.
16. Ibid., Kean, pg. 24.
17. Ibid. NQRC, papers of James Gibson, RCA, Chapter III, pg. 3/7.
18. Mitsuo Ohara, "Distortion and Crosstalk Caused by Multipath Propagation in Frequency-Modulated Sound Broadcasting", *IEEE Transactions on Broadcasting*, Vol. BC-26, No. 3, Sept. 1980, pp. 76–81.
19. Ibid., NQRC/Gibson, pg. 3/11.

3.5 TV Transmitters

Robert R. Weirather, Gregory L. Best, Robert Plonka
Harris Corporation, Quincy, Illinois

INTRODUCTION

Significant advances have been made in TV transmitter technology in recent years. New technology and ideas have been introduced in order to continue to provide high quality TV signal transmission while improving reliability, reducing maintenance, and lowering overall cost of ownership. These new technologies fall primarily into the following areas: solid-state high-power amplifiers, efficiency improvements for UHF transmitters, and transmitter combiners which eliminate off-air switching. Each will be covered in this chapter.

Further, the FCC continued its policy of technical deregulation which has served to allow more flexibility in transmitter design and system operation.

VISUAL EXCITER

The visual exciter has the purpose of receiving a video baseband signal, processing it, and converting it to a fully modulated vestigial sideband on-channel signal. Since intermediate frequency (IF) modulation has become a standard within the industry, most of the signal processing occurs either in the video stages or in the IF stages. The basic block diagram of an IF modulated transmitter is shown in Fig. 1. We will discuss each basic block diagram in greater detail.

Video Processing

For the TV transmitter's goal of becoming transparent to the incoming signal, it is of utmost importance to make sure that the incoming signal is optimized. It is often difficult to define exactly where the station video processing of a signal ends and the transmitter video processing begins. Typically some video processing is done external to the transmitter at the station plant by a *proc amp*. Different manufacturers of transmitters employ various means of accomplishing the same purpose and therefore, may use different approaches to solve the same problem. The main functions of the exciter video processing circuitry are:

- Obtain proper sync to video ratio
- Remove any common mode signal
- Provide overall video level control
- DC restoration
- Overmodulation prevention
- Frequency response correction for the signal applied to the transmitter

Quite often precorrection, used to compensate for some transmitter distortions, is included in the video processing section. Types of transmitter precorrection which are sometimes included are differential gain,

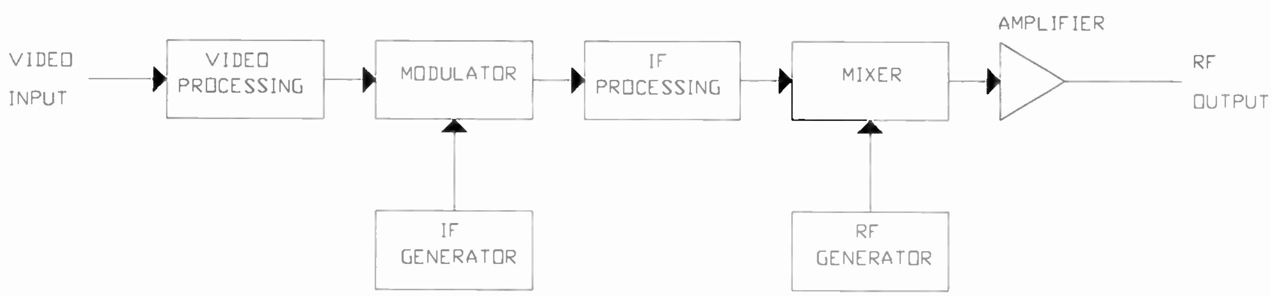


Figure 1. IF modulated transmitter.

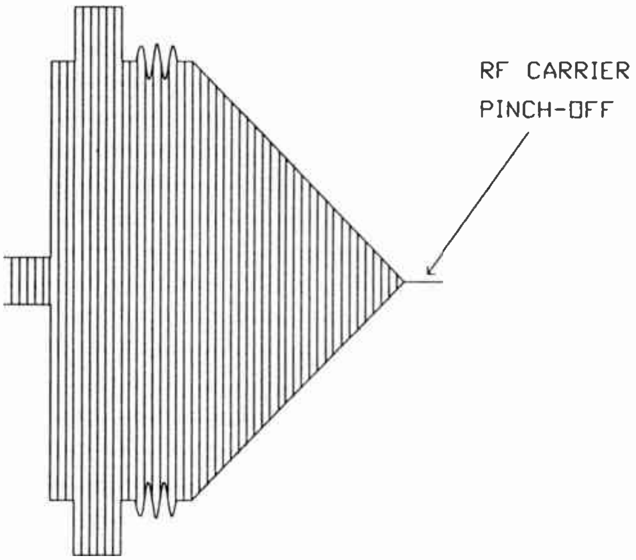


Figure 2. RF carrier pinch-off.

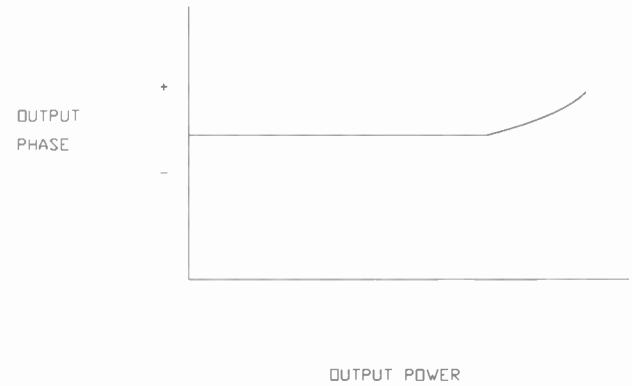


Figure 4. Phase transfer curve.

differential phase, and group delay compensation (for transmitter and/or for the notch diplexer, if used).

DC restoration, (clamping) is important because picture brightness information is contained in the DC component of the video signal. If AC coupling is used anywhere in the video circuitry, the DC level will tend to vary as the capacitors in the coupling circuits charge and discharge. AC coupling is convenient since differences in AC and DC grounds can be allowed without introducing waveform distortion. By clamping a consistent part (such as sync tip or "back porch") of a TV waveform during each TV line to a fixed voltage which does not vary, the correct DC level is applied to the modulator.

Common mode signals such as noise or AC hum can be removed by using a differential input mode for the video signal input stage. It is desireable to use an RF choke at the video input stages also to prevent rectification of ambient RF and remodulation of the main signal from that rectified RF.

White peak limiting is used to prevent modulation from reaching 0% or in other words, to keep the carrier from being "pinched off" as shown in Fig. 2. When an intercarrier TV receiver encounters a carrier that has been pinched off, it temporarily has no signal to receive, its automatic gain control circuits increase to maximum gain, and since there is nothing but noise available when the carrier is pinched off, that noise is greatly amplified, transferred to the aural intercarrier and becomes audible as a buzz in the receiver.

High frequency peaking circuits in a transmitter's video processing circuitry are often used to compensate for long video signal runs in a transmitter plant.

In UHF exciters, often the horizontal and vertical sync portions of the TV signal are detected and sent to klystron pulsers.

When a transmitter is adjusted for maximum efficiency, its transfer characteristic is not ideal. Typical transmitter amplitude and phase transfer curves are shown in Figs. 3 and 4. Often the video signal is predistorted in the opposite direction of the errors in amplitude and phase produced by the nonlinear transfer curve of the transmitter. An example of this predistortion is given by Fig. 5.

Nonlinear chroma gain at different luminance levels that exhibits a change in the saturation of the colors, is termed *differential gain*. Nonlinear chroma phase at

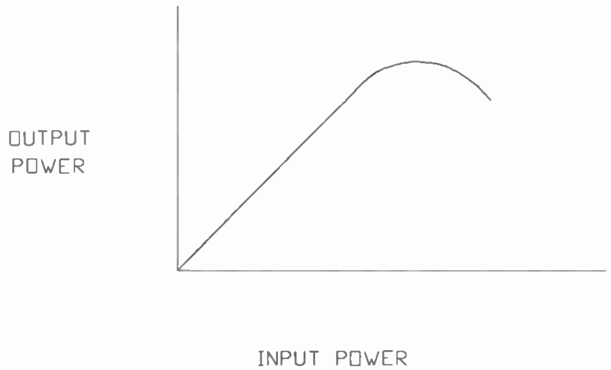


Figure 3. Amplitude transfer curve.

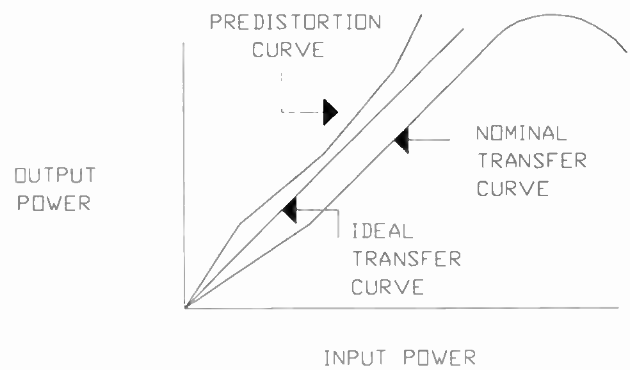


Figure 5. Sample predistortion curve.

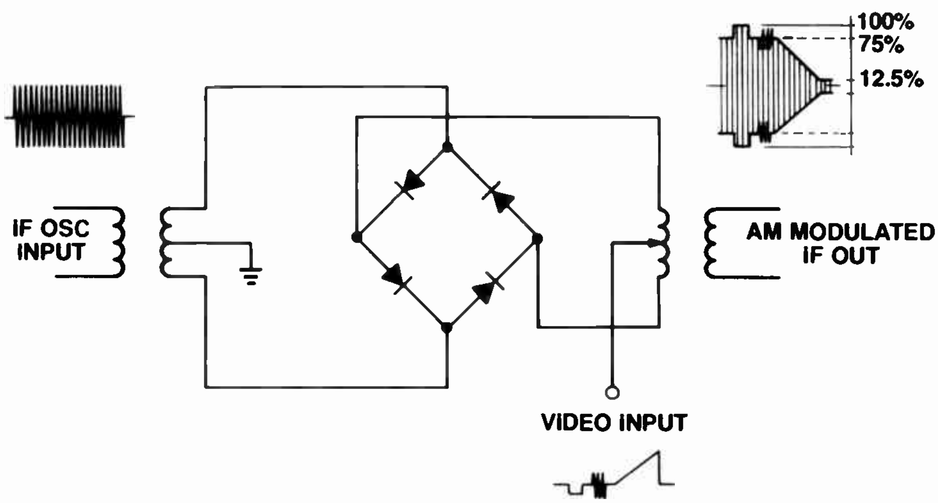


Figure 6. Balanced mixer.

different levels of luminance that exhibits a change in hue of the colors is termed *differential phase*. Differential gain and phase predistortion is often done in the exciter to compensate for nonlinearities in the transmitter.

Luminance nonlinearity is present when there is a change in luminance gain with different brightness levels.

Both luminance nonlinearity and differential gain are amplitude distortions. Precorrection can be accomplished by changing the gain of a video amplifier over a portion of the total video signal. The gain change is usually keyed by the video signal reaching a certain level. Differential phase can be corrected by splitting the video signal into two paths which are 90° different at the color subcarrier. By changing the gain of one path over a portion of the video waveform while not changing the other path phase changes can be obtained. Other methods can also produce the same result.

Modulator

With nearly all IF modulated transmitters, the modulator is a broadband, balanced, diode mixer. It is configured for maximum rejection of the local oscillator signal and biased so as to provide excellent linearity, low noise, and capability to achieve carrier cutoff. A schematic of a typical balanced mixer is shown in Fig. 6. The video signal is DC offset to provide the proper modulation level. The video signal is used to control the attenuation of the diodes. Peak of sync corresponds to maximum IF envelope output and white corresponds to minimum IF output. The output signal of the modulator is a double sideband AM signal having the proper depth of modulation (12.5%).

Group Delay Compensation

Waveform distortion can be caused by group delay inherent in RF amplifiers and tuned RF filters, combiners, and other output systems. Group delay distortion is the nonuniform delay of different frequencies over the bandwidth of the TV signal. Group delay distortion occurs in tube cavity amplifiers and notch diplexer combiners. In general, the closer the amplitude roll-off frequency is to the visual passband, the higher the group delay distortion.

IF and Video Delay Compensators

Group delay impairments show up in a picture as color smear and halo effects on edges. On a waveform monitor the effects may be seen using the 2T and modulated 12.5T contained in the composite test signal shown in Fig. 7. Pulse responses with unacceptable group delay may include exaggerated pre- or post-ringing on 2T and modulated 12.5T base line disturbance.

The multipulse signal shown in Fig. 8 may also be used to analyze group delay. The multipulse signal consists of a gray flag (80 IRE), 2T pulse, and five sine-squared pulses modulated with five discrete frequencies (consisting of one 25T pulse and four 12.5T

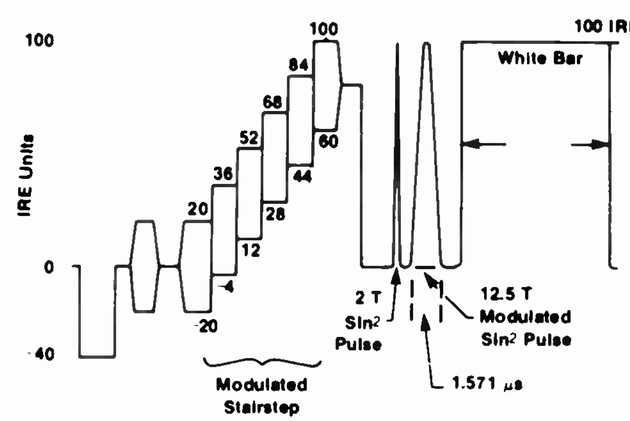


Figure 7. Composite test signal.

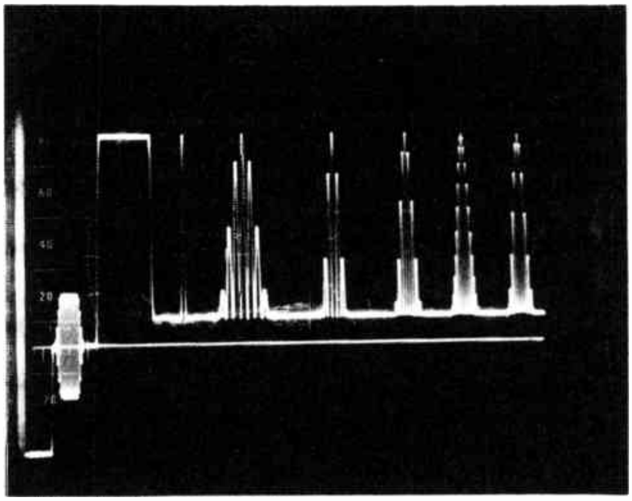


Figure 8. Multipulse.

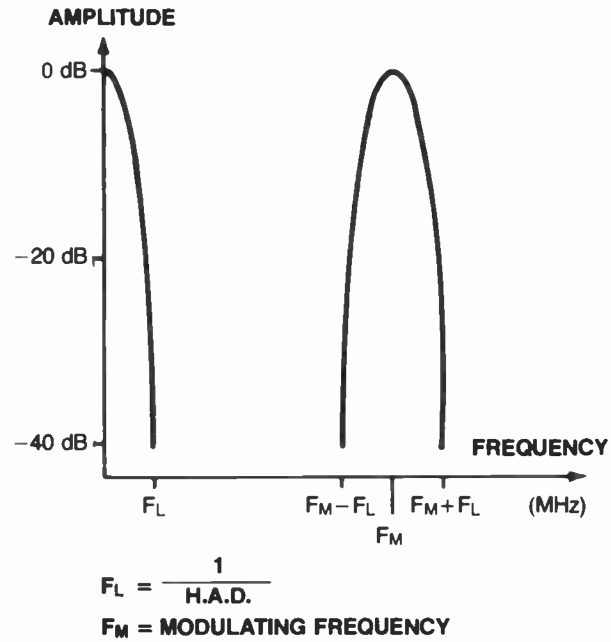


Figure 9. Energy spectrum for modulated pulses.

pulses). This signal may be used to measure linear distortion in TV systems such as amplitude versus frequency response and group delay distortion.

The modulated pulses contain two spectra of information, low frequency and high frequency, as illustrated in Fig. 9. These modulated pulses are useful for measuring group delay errors. If the low-frequency spectra and the high-frequency spectra are delayed equally, the results will be a symmetrical modulated pulse with the same shape as the input. The gain versus frequency distortion will alter the base line flatness but will not change pulse symmetry. Delay errors will result in an asymmetrical pulse baseline. Combinations of group delay and gain errors are shown in Fig. 10.

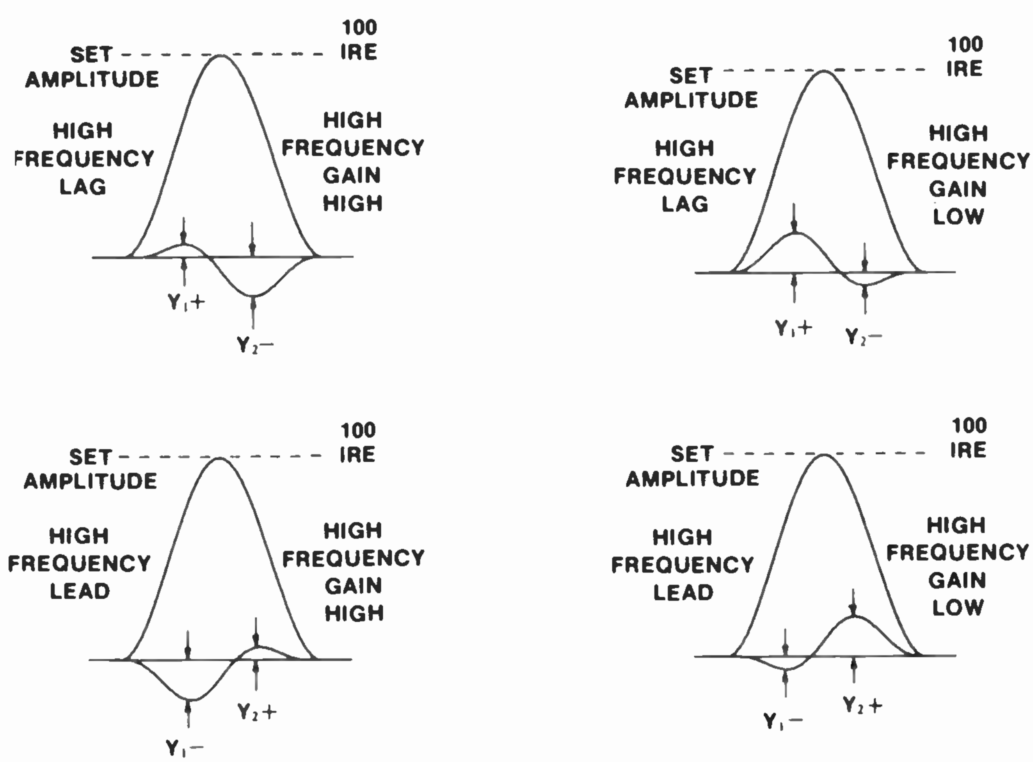


Figure 10. Modulated sine-squared pulses with gain and phase errors.

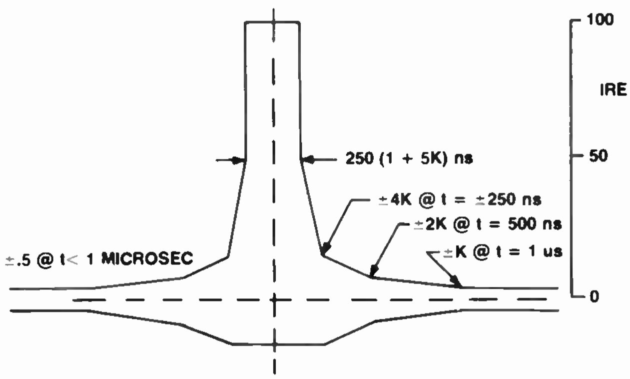


Figure 11. 2T sine-squared pulse graticule.

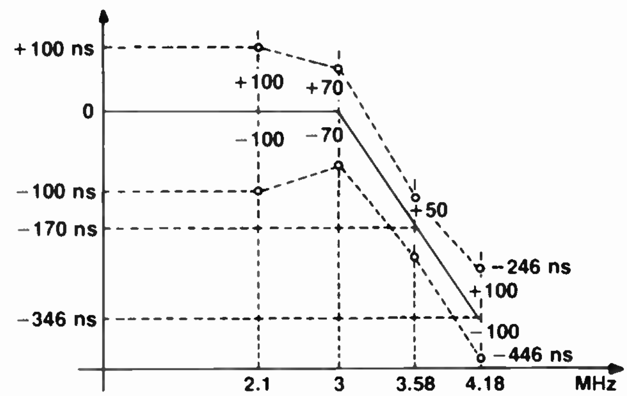


Figure 13. Predistorted group delay curve.

A method of quantifying these distortions uses the waveform graticule shown in Fig. 11. This graticule was arrived at empirically and represents constant perceptible distortion levels. This shows that overshoots closer to the desired pulse are not as perceptible as ringing further away.

The desired K factor graticule, i.e., 2%, is overlapped on a waveform monitor and the group delay corrector is adjusted until the 2T waveform lies entirely within the graticule. This technique is often preferred to swept group delay measurements because the results are in terms of perceptibility.

Low-frequency group delay and amplitude errors are referred to as short time waveform distortions. Fig. 12 shows typical distortions of a 2T pulse.

If the group delay error affects only the high frequency side of the passband, precorrection can be accomplished at video. This is the case for most visual-aural RF notch type diplexers and the FCC receiver equalizer curve. In the case of the notch diplexer, a pair of cavities resonant at the aural frequency are inserted between two 90° hybrid couplers. (Refer to Fig. 87.) The presence of the tuned circuits near the upper edge of the visual passband can cause a significant group delay error.

Another source of group delay is the tuned circuits of the high power tetrode or klystron amplifiers.

Finally the FCC requires the TV transmitter to predistort group delay according to the curve shown in Fig. 13.

The delay shown in Fig. 13 is a characteristic delay curve the FCC requires all visual transmitters to have. The purpose of it is to compensate for the delay error caused by discrete LC aural notch filters in early TV receivers. The reason this was done was that it would be far cheaper to group delay compensate one transmitter than every TV set. The notch filter in TV receivers is necessary to prevent the aural carrier from mixing with the detected video and chroma subcarrier signals and producing visible spurious beats.

Group delay predistortion may be accomplished at video baseband or IF. The techniques used are similar in concept. Both active and passive equalizers may be employed. IF group delay correction is necessary to correct group delay errors below visual carrier while not affecting the signal above visual carrier. Above visual carrier both IF and video correction is effective.

Passive Group Delay Equalizer

A common form of a passive group delay equalizer is shown in Fig. 14. It provides a flat frequency response and a nonlinear group delay which peaks at the resonant frequency of the circuit. This type of circuit is referred to as a passive all-pass network.

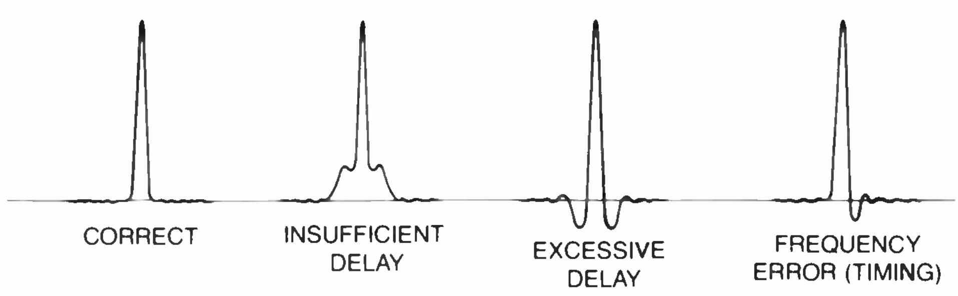


Figure 12. Typical distortion of a 2T pulse.

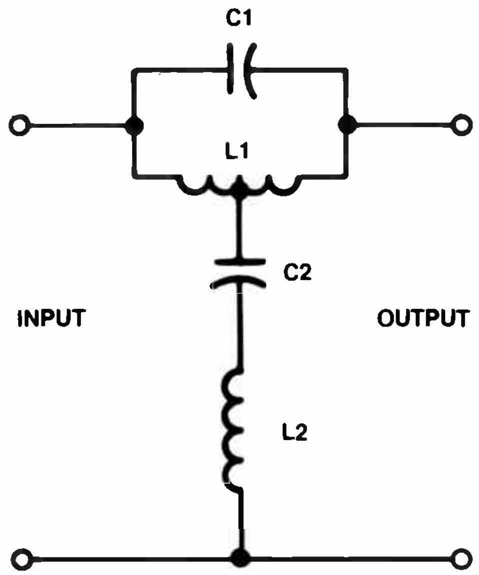


Figure 14. All-pass network.

Without going through a detailed analysis of the circuit, much can be understood by examining the circuit at frequencies well below and above resonance. At low frequency the network can be approximated by a low series inductive reactance due to L1. Here the output voltage leads in phase. At high frequency well above resonance the circuit can be approximated by a series capacitive reactance due to C1. At high frequencies the output voltage lags in phase. The output amplitude, however, is constant across the band. The phase of the network is plotted in Fig. 15. The slope of the phase is defined as group delay and is expressed as

$$\tau = (d\phi/d\omega)$$

where: ϕ = phase angle
ω = angular frequency

The steeper the phase slope the higher the maximum group delay. A plot of the group delay is shown in Fig. 16.

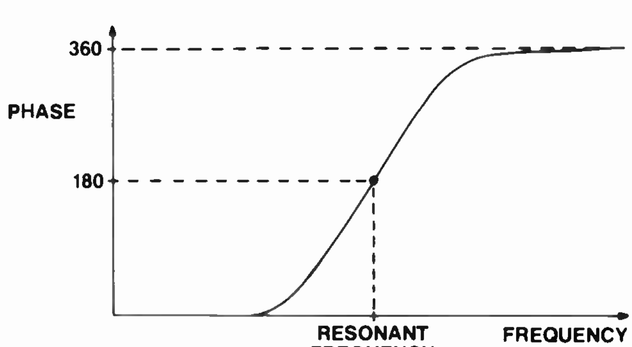


Figure 15. All-pass phase.

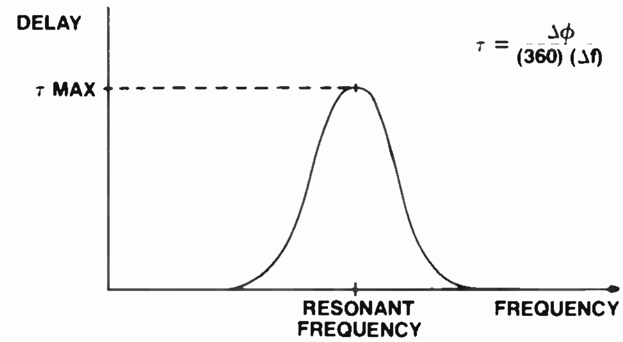


Figure 16. All-pass group delay.

Active Group Delay Equalizer

As in the passive type the network has a constant amplitude frequency response and a nonlinear phase response. The action of an active all-pass network can be best explained using the simplified schematic of Fig. 17. If the phase of voltage e_1 were plotted as a function of frequency, it would trace out a circle with maximum amplitude and zero phase at resonance. Voltage e_2 on the other hand has a constant amplitude and phase.

For the case when e_1 voltage is equal to twice e_2, the output voltage of the summing amplifier, e_3, has the characteristics of an ideal all-pass network. The output amplitude is constant and the resonator tuning determines the frequency of maximum group delay and the resonator Q determines the magnitude of delay.

For the case where voltage e_1 is larger than twice e_2, the output will have an amplitude peak at resonance and conversely if e1 is smaller, the output will have a dip at resonance.

Vestigial Sideband Filter

The FCC requires that the radiated TV signal have a major portion of the lower sideband (vestigial sideband) suppressed. In addition, the upper sideband signal must be contained within 4.75 MHz of the visual carrier. With the advent of IF modulation the filtering is at low-power stages and transmitter manufacturers have selected solid state filters using Surface Acoustic Wave technology to accomplish the stringent filtering requirements.

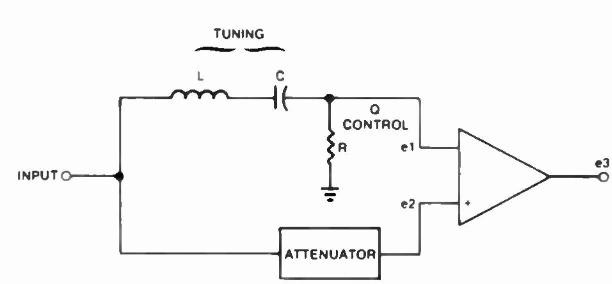


Figure 17. Active all-pass network.

The term *surface acoustic wave* (SAW) refers to the propagation of elastic waves on the surface of a piezoelectric crystal. The wave propagation is roughly the speed of sound, and therefore called acoustic. A time varying voltage on a metalized transducer is used to induce a deformation on the surface and produces a electromechanical wave. The deformations produce local electric fields which travel along with the mechanical wave, and interact with metal electrodes, which convert the mechanical wave back to a time varying voltage. The length, spacing, and the number of the electrodes determine the wave shaping properties of the filter.

The electrodes act like tapped delay lines, as illustrated in Fig. 18. The electrodes are designed to amplitude scale and delay the signal. The output is selectively attenuated depending on the time delay and signal frequency relationship. This type of filter is called transversal because the attenuation is controlled by delay lines rather than resonators. The transversal filter was invented in the early 1940s and used coax cables as delay lines. Today, surface acoustic wave transducers provide the same function. The wavelength of an IF acoustic signal is approximately 0.003".

This small wavelength allows the filter to be very small and compact. Because the transducers are only on the surface, photographic masks can be used to accurately control the physical dimensions of the transducers. The photographic mask lends itself to modern manufacturing techniques and insures a reproducible filter with a permanent amplitude versus frequency response and group delay.

A characteristic of a transversal filter is that group delay can be set independently of the amplitude characteristic. This is not true of discrete LC filters where basic physical laws couple group delay and amplitude.

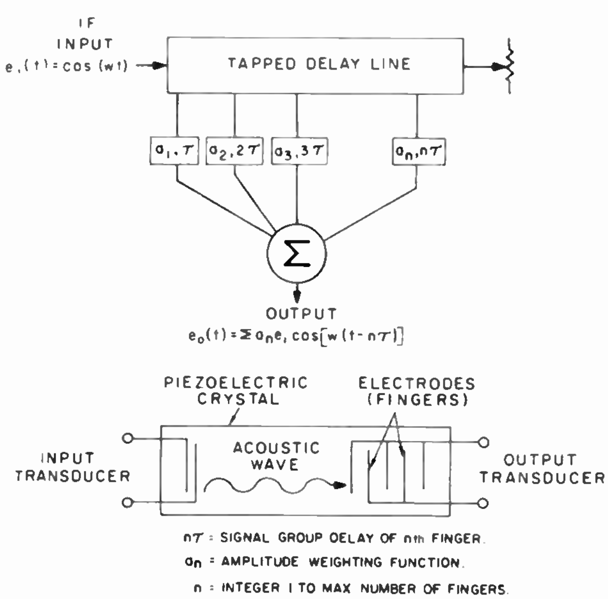


Figure 18. Saw filter.

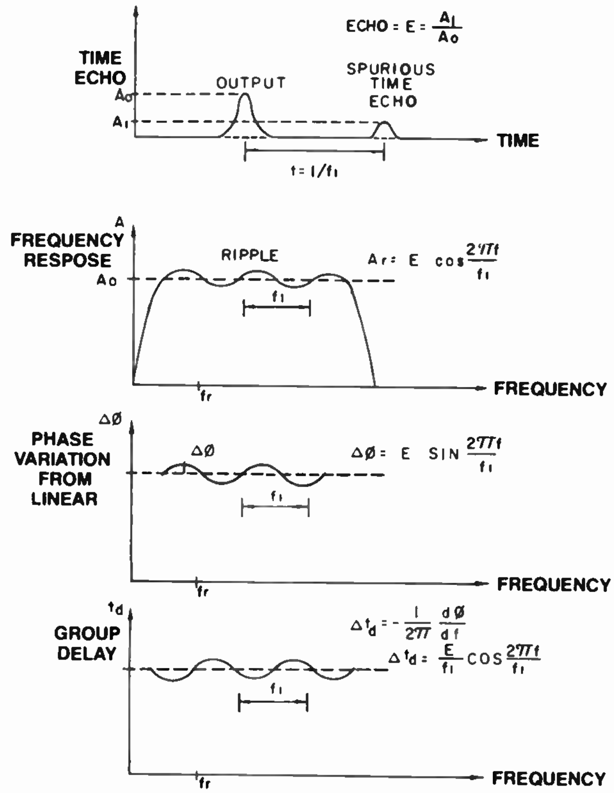


Figure 19. Response, phase, and delay ripples caused by time echoes.

The independent nature of group delay and amplitude allows a manufacturer to provide almost any type of group delay curve (such as a constant delay or the TV receiver group delay precorrection curve) across the passband.

Although the FCC does not require attenuation of visual signals in the aural passband, video signal components can cause interference called visual-to-aural crosstalk. Reduction of visual-to-aural crosstalk is essential to the proper transmission of stereo sound and other subcarrier broadcast services. High resolution cameras and character generators can produce spectral components at 4.5 MHz and beyond. Without sufficient attenuation these signals can cause distortion, especially in aural subcarriers which are more sensitive to visual crosstalk than the main aural signal. SAW filters thus are able to perform many functions simultaneously, i.e., bandwidth shaping, group delay equalization, and video to aural crosstalk reduction.

Distortion products that can be created by SAW filters are due to signal reflections (triple travel) and direct feedthrough. These distortions cause time echoes displaced either before or after the desired responses. In the frequency domain these echoes show up as ripples in the passband. (Fig. 19) A given time echo will contribute uniquely to the amplitude, phase and group delay characteristics by the addition of a ripple component of the waveform in the passband.

The amplitude of the ripple is proportional to the echo level, and its period is proportional to the reciprocal of the SAW time delay.

The SAW group delay characteristic will vary sinusoidally with the same period as the passband ripple. Its magnitude, however, is a function of both echo level and inversely proportional to time delay, and therefore not a sure test of signal distortion. The group delay ripple for long delayed echoes will give peak-to-peak values which have no correlation with conventional signal distortion estimates. For that reason, fast peak-to-peak group delay errors that occur closer than the reciprocal of the filter time delay can usually be ignored.

When pulse waveforms in Figs. 7 and 8 are used to measure group delay distortion, the 2T preshoot, overshoot, and asymmetry can be used to gauge low-frequency group delay error. Modulated pulse baseline "S" curve is a result of group delay and the peak excursions of the "S" curve can be used to quantify the amount of group delay distortion.

IF Linearity Precorrection

Amplitude and phase nonlinearities are sometimes corrected at video baseband. The alternative is to provide the correction at IF frequencies. There are advantages to correcting at IF in that, since most distortions are caused in the high power RF amplifiers after vestigial sideband filtering, a corrector placed after the vestigial sideband filter can more accurately predistort the modulated signal. Any precorrection spectra generated at IF after the VSB filter will produce energy components which can cancel intermodulation products generated in the final amplifier stage. This is particularly important in the case of a pulsed klystron transmitter.

Correction of distortion at IF is particularly helpful for chroma distortions. Chroma spectra at 3.58 MHz has only a single sideband information and thus has less energy than equivalent luminance signals. An ideal diode detector frequency response of the NTSC modulation signal is plotted in Fig. 20. Here it is seen

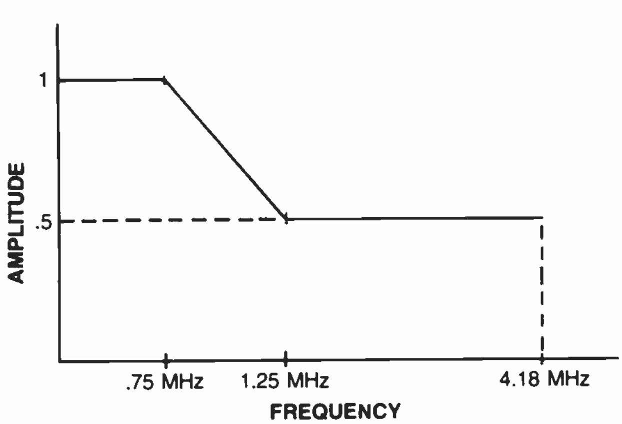


Figure 20. Frequency response using ideal diode detector.

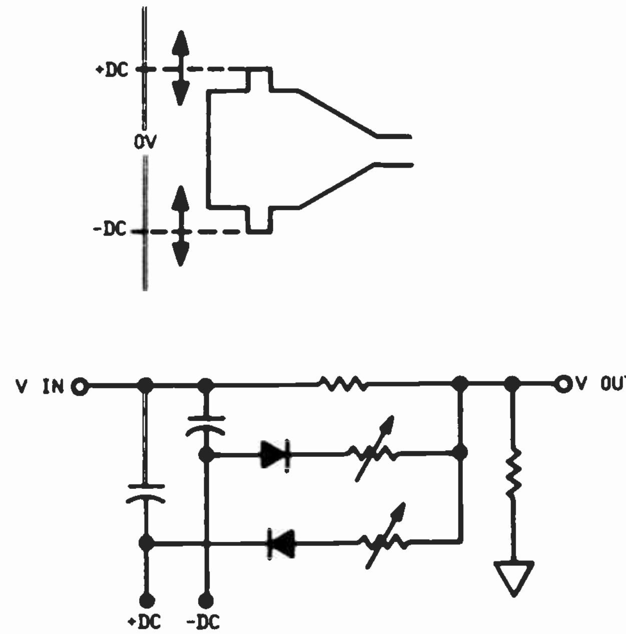


Figure 21. Basic gain expansion circuit.

that beginning at 0.75 MHz the video begins to fall to −6 dB. For video signals lower than 0.75 MHz, the RF spectrum is double sideband and has twice the peak RF voltage.

Intermodulation products are caused in high power amplifiers by nonlinear amplitude transfer characteristics. As the power output increases towards saturation, amplitude compression begins and in general the phase begins to lag as well. The nonlinear transfer characteristics give rise to mixing products which occur at sum and difference frequencies around the visual carrier. This process also creates the frequency spectra of what is called the lower sideband.

Correctors for amplitude distortion are usually similar in concept to video differential gain correctors. Linearity correctors generally use diodes which are keyed by the IF modulated signal. When the diodes conduct, the gain or attenuation is reduced as needed.

An example of a basic gain expansion circuit is shown in Fig. 21. The signal is normally attenuated a fixed amount by using a resistive L-pad. The diodes are normally reverse biased by equal, but opposite polarity, DC voltages. Reducing the DC voltage amplitude permits the diodes to conduct on the signal peaks. This inserts additional resistance in parallel with the series arm of the L-pad attenuator thereby decreasing the attenuation. Varying the resistance in series with the diodes provides for a variable gain expansion.

ICPM

Phase distortions in high power amplifiers produce incidental carrier phase modulation (ICPM), which are spectral components in quadrature with the modulation signal. Fast video amplitude changes such as a step or pulse will cause larger incidental phase spectral components than slow changes. Receivers make this condition worse by attenuating the lower sidebands below 0.75 MHz. The receiver then responds to the extra sidebands created by the phase modulation as if they were amplitude modulated single sidebands and produce spikes. The faster the rise time on the signal the more high frequency energy is present resulting in edge distortions in the displayed picture.

The picture impairment is similar to simultaneous group delay and differential phase errors in that edges are less sharp and color hue changes with brightness. On a waveform monitor, overshoots are visible as trailing edges and as rounding on leading edges. These overshoots vary in severity depending on how close the power amplifier is driven towards saturation.

Audio impairment is produced by ICPM in receivers employing intercarrier conversion. Intercarrier receivers use an AM or synchronous detector to produce a 4.5 MHz aural IF signal from the composite video IF. Any phase modulation present on the visual carrier is then transferred to the aural intercarrier. In the monaural baseband audio the increasing amplitude frequency effect of ICPM is nullified by de-emphasis to some degree. With multichannel sound, however, there is no de-emphasis applied to the baseband stereo signal, and thus the distortion is more pronounced at the stereo subchannel and pilot frequencies. To counteract the effects of ICPM and other noise sources on the stereo subchannel, an audio companding is employed that greatly reduces the potential interference. Although the audio companding process can reduce some of the effects of ICPM, ICPM correction is essential in delivering clear, low-noise audio to intercarrier receivers.

There is no defined level of ICPM for a given stereo performance level since the signal to buzz ratio is highly dependent on the picture spectral components.

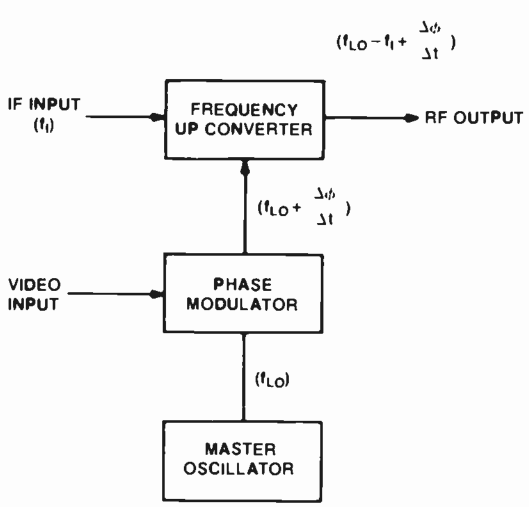


Figure 22. Master oscillator phase modulator block diagram.

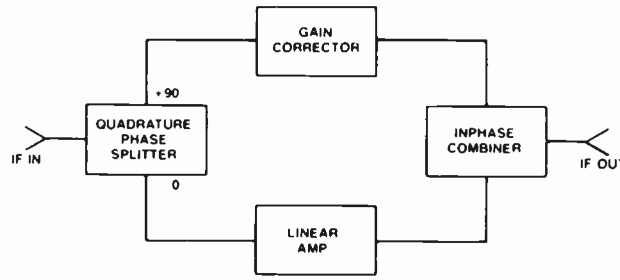


Figure 23. Direct IF ICPM corrector.

Refer to the EIA Recommended Practices for the current recommendations on ICPM limits.

ICPM precorrectors can be grouped into two types: ones using a phase modulator and the other operating on the signal directly. The phase modulator type uses video to vary the IF or master oscillator phase with the opposite phase characteristics of the nonlinear amplifiers. A phase modulator can also operate on the IF signal directly using a video signal to set the amount of modulation. A block diagram of a master oscillator phase modulator is shown in Fig. 22.

ICPM precorrectors operating directly on the IF signal can be implemented several ways. Direct precorrection at IF is similar in concept to baseband differential phase precorrection. See Fig. 23. In both cases, the visual signal is split into two paths which are in phase quadrature. In the IF corrector, the entire channel of frequencies is in quadrature, whereas in the video precorrector only the chroma band is in quadrature. One method of implementation is to modify the quadrature signal gain function with level dependent diode expansion or compression circuits. This can be done using the same techniques as in the linearity corrector.

RF Generator

Transmitters employing IF modulation generate the following frequencies; visual IF, modulated aural IF, and master oscillator signal(s) for translating visual and aural IF to the final carrier frequencies. These oscillators have been implemented with either digital synthesizer techniques or crystal oscillators. An advantage of the synthesizer is that only one crystal is needed and it operates at a single standard frequency for all TV channels. The crystal oscillator approach, however, may involve simpler circuitry.

The two commonly used IF frequencies are 37 MHz and 45.75 MHz. There are many reasons for selecting one IF frequency or the other. One advantage of 37 MHz is that the temperature drift sensitivity of most IF components such as the SAW filter is related directly to carrier frequency. Thus the lower IF has a 12% less drift sensitivity than components at 45.75 MHz. The second harmonic of 37 MHz falls in between channels 4 and 5 so as not to cause interference. On the other hand, 45.75 MHz is a common demodulator IF which can be useful for IF troubleshooting. Temperature drift may be minimized at either IF by maintaining the SAW filter at a stable temperature.

The important performance characteristics of an oscillator are: its phase noise, sensitivity to frequency drift with time and temperature, and susceptibility to mechanically induced phase and frequency shifts called microphonics.

In replacing a crystal, it is important to follow the recommendations of the oscillator manufacturer to insure proper operation. Synthesizer performance should be properly maintained to prevent inadvertent phase noise and spurious frequency generation.

OFFSET FREQUENCY CONTROL

The limited number of available channels for TV Broadcasting makes it necessary to assign the same carrier frequencies to many stations. To avoid interference between stations operating on the same frequency (co-channel interference) geographical separation and radiated powers are carefully selected.

Nevertheless, considerable co-channel interference was encountered in many locations. Investigation has shown that additional means were able to reduce co-channel interference by reducing optical or perceptive interference elements in the system. In other words, steps were taken not to alter the strength of the interfering carrier but to choose operating parameters to reduce the visibility of the interference.

Considerable investigative work was done to lay the scientific and physiological basis and to define the parameters which need to be controlled to reduce the visibility.

It turned out, with horizontal and vertical sweep ratios fixed, the visual carrier frequencies of the interfering stations was the parameter which needed to be controlled.

Co-channel interference between television stations appears to viewers as a horizontal pattern of alternating light and dark bars on the viewing screen—very much like the shadows cast by venetian blinds. It has been demonstrated for many years that the visibility of these bars varies cyclically as a function of the difference in frequency of the interfering carriers (Fig. 24). The interference is most visible when the carriers were offset by multiples of the line frequency (15734 kHz) and least visible when the carriers are offset by odd multiples of one-half the line frequency. In addition to the gross maxima and minima, fine grain maxima and minima occur when the frequency offset is a multiple of the frame frequency (29.97 Hz).

Ideally, stations would be offset by odd multiples of one-half the line frequency to provide minimum interference visibility. However, a third station in the same area would be offset from one of the other stations by an even multiple of the line frequency. Hence, maximum visibility of the interference would occur. Therefore, 10 kHz offsets currently used in the United States were chosen to provide approximately equal reduction of the interference patterns for any number of stations in geographical proximity. (See Fig. 25.)

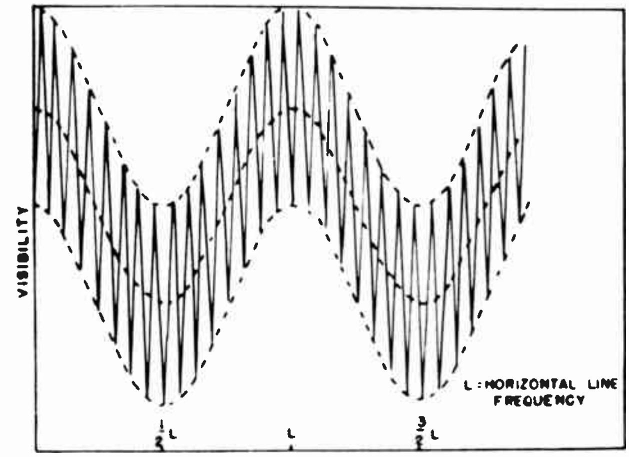


Figure 24. Co-channel interference.

Precise Frequency Offset

Although it is not practical to utilize the gross minima occurring at odd multiples of one-half the line frequency, it was determined experimentally that utilizing the fine grain minima, occurring at even multiples of the frame frequency, would be very advantageous in reducing the visibility of co-channel interference.

The nearest even multiple (334th) of the frame frequency to the 10 kHz offset is 10010 Hz. In a three station arrangement, one station will have zero offset, and the other two stations will be offset by ±10010 Hz. Experiments indicated that changes in the frequency differences of 5 Hz had a negligible effect on the reduction of the interference visibility.

To maintain the precision offset within 5 Hz requires maintaining each visual carrier frequency within 2 Hz or 3 Hz. Maintaining a television transmitter to such tight frequency tolerances requires some type of control system using an extremely stable frequency source.

In exciters using two independent oscillators (one for IF and one for the local oscillator), the visual carrier signal may be derived from mixing the oscillators together and comparing that signal to the reference signal in the comparator. That resultant error signal can be used to adjust one of the oscillators, preferably the local oscillator. For exciters using synthesizers, the sythesizer reference oscillator is compared to the precise frequency standard in the phase detector and the resultant error voltage is used to adjust the synthesizer oscillator.

By phase-locking the visual carrier to a stable reference oscillator, the master oscillator acquires the stability of the reference source. Sources which use an internal WWV Receiver/Comparator to self-correct or an Atomic Frequency Standard can easily provide the stable reference source required. These reference sources allow a transmitter to be maintained within a few hertz of a desired frequency indefinitely. Experimental results have indicated that frequency differences between transmitters can vary as much as 5 Hz

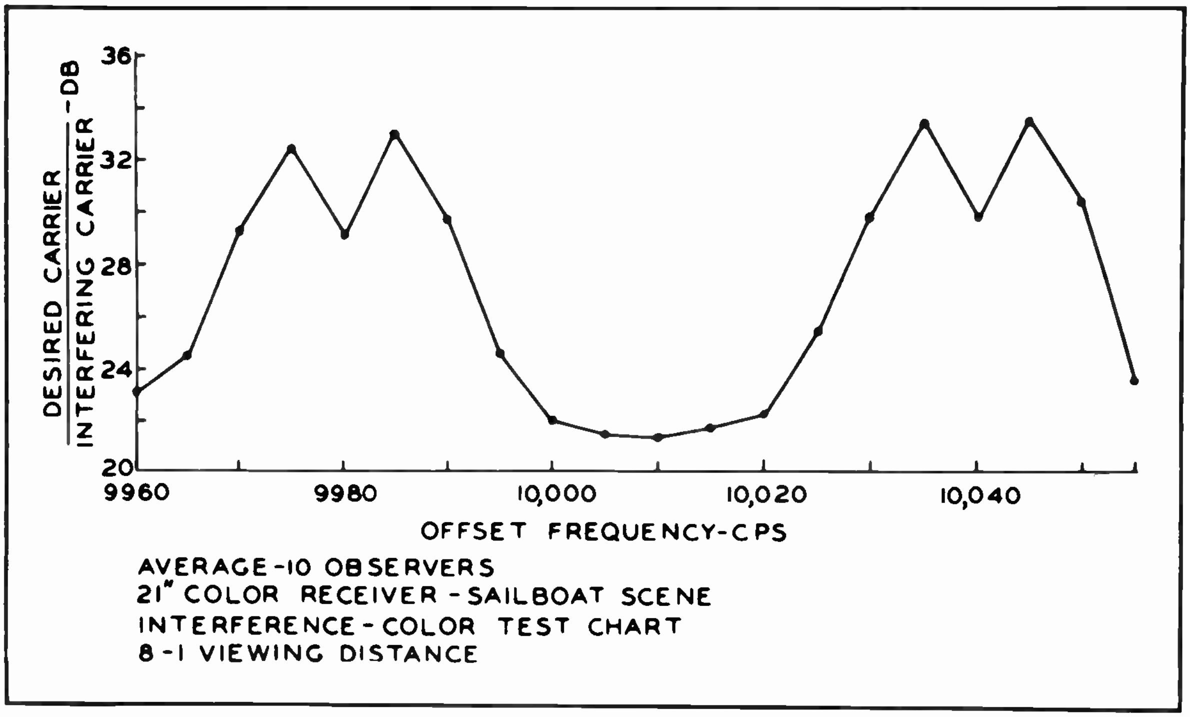


Figure 25. 10 kHz offset pattern.

from the precision offset before co-channel interference becomes noticeably worse. This means that two television stations can minimize their co-channel interference without the requirement to continuously adjust the transmitter frequency.

One word of caution when working with any frequency standard. If power is removed from the standard and no backup power supply is incorporated either internally or externally, ensure that the standard has stabilized before making any adjustments. Also ensure that after any adjustment has been made, a suitable time period has elapsed to allow the standard to settle before making any reading.

To ascertain the carrier frequency of a station operating under precision frequency control, assuming a measuring accuracy ten times better than the quantity to be measured, means measurement equipment must have the following accuracy requirements:

VHF Low Band	$+2.5 \times 10^{-9}$
VHF High Band	$+1 \times 10^{-9}$
UHF	$+3 \times 10^{-10}$

Frequency counters with the above accuracy and with National Institute for Standards and Technology (NIST) broadcasts, traceable calibration may be used for making measurements.

AURAL EXCITER

In its most basic form, the aural exciter consists of some audio processing, an FM modulated IF oscillator, and an up-converter to obtain the desired carrier signal as shown in Fig. 26.

Audio Processing Circuits

To ensure that the transmitter is not the limiting factor in audio (monaural and stereo reproduction), it is desired that the transmitter add as little distortion to the incoming signal as possible.

Baseband audio of the BTSC[1] MTS (Multi-channel Television Sound) system (including the mono, stereo, SAP, and profession channels) include frequency components to 105 kHz. Emphasis must be placed on phase linearity, low distortion, reduction of any amplitude ripples, or roll-off over the stereo (BTSC) passband to achieve good stereo separation and minimum crosstalk between the main stereo and the SAP channels.

While unbalanced coaxial inputs are used for the MTS input to the aural exciter, it is still necessary to use some form of common mode rejection to reduce the possibility of hum and noise from getting directly into the audio stages.

All errors in phase linearity and amplitude response within the audio circuitry contribute to stereo separation degradation. *As a general rule, amplitude roll-off*

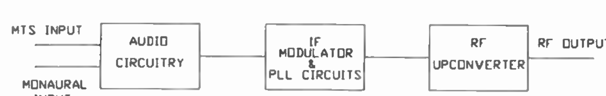


Figure 26. Aural exciter.

should be less than 0.1 dB and departure from phase linearity should be less than 1 degree for good quality stereo.

IF Modulation

All modern transmitters are IF modulated. To ensure low power stages do not contribute any group delay or amplitude roll-off, wideband amplifiers should be used. Modulated oscillator linearity requirements include flat modulation sensitivity versus frequency characteristics up to 47 kHz minimum, and typically out to 120 kHz.

With the advent of the MTS stereo system, intermodulation (IMD) and harmonic distortion (THD) products which, in monaural operation lie above 15 kHz (which would therefore be reduced by deemphasis in the receiver), now lie in the stereo channel or the SAP channel and will degrade stereo separation or will cause crosstalk into the SAP channel. In addition, IMD products generated in the stereo channel may now lie in the mono or SAP channel.

As the MTS signal is up-converted to the aural RF carrier frequency, the residual FM of the local oscillator signal becomes a determining factor. The level of FM produced by the local oscillator should be 10 dB lower than the modulated oscillator. Synthesized sources should be inspected for spurious frequencies which may show up as FM noise.

RF Amplifiers and Diplexers

All RF amplifiers are optimized (for stereo performance) when a flat, symmetric frequency response and a minimized variation in group delay across the modulation passband is achieved. Since FM modulation and demodulation is a nonlinear process there is not a one-to-one correspondence between RF amplitude/phase response and baseband stereo separation and crosstalk. Measurements indicate that a 1.5 MHz, 3 dB bandwidth will provide excellent stereo and SAP performance. RF amplifiers must be swept-frequency analyzed to insure the passband is sufficiently wide to pass the stereo signal and that its passband amplitude and group delay characteristics are symmetrical about the carrier frequency.

The diplexer is the last and possibly the most critical element in the aural chain with regard to stereo separation and crosstalk between stereo channel and SAP. However, hybrid diplexers, since they are broader bandwidth than notch diplexers, normally do not present any degradations to multichannel sound. If correct tuning of the notch diplexer itself and of any associated group delay compensation or amplitude correction is not properly maintained, stereo separation and crosstalk will be degraded. Low insertion loss at 4.18 MHz and proper aural bandwidth are both required for optimum visual and sound performance. Refer to the EIA Recommended Practices (Systems Bulletin #5) for other current recommendations.

Minimizing antenna reflections (VSWR) helps to keep the pilot signal from becoming distorted and the audio channel from losing separation or becoming noisy.

In summary, these characteristics have been incorporated into transmitters designed for MTS:

1. Wideband low distortion audio stages with excellent signal-to-noise ratio (SNR) are employed to insure negligible distortion in the baseband signal.
2. Low distortion wideband modulated oscillators with improved noise performance and phase locked loop (PLL) techniques are used so that the quality of TV multichannel sound will approach that of the FM broadcast service.
3. The bandwidth of IF and RF stages is wide. Power bandwidth tradeoffs may need to be made to achieve optimum stereo performance.

COMPENSATION OF AURAL PASSBAND FOR OPTIMUM STEREO PERFORMANCE

This section describes circuits that offer group delay equalization for a TV aural transmitter when operating through a notch diplexer. These circuits introduce group delay correction in the IF section of the aural exciter and after up-converting to the operating channel, the group delay precorrection effectively corrects the adverse group delay existing in the output diplexer circuit. The end result is improved TV stereo separation. In addition, precorrection of the FM bandpass can allow the use of lower cost, single-cavity notch diplexers. Stagger-tuned dual-cavity notch diplexers also have been utilized to provide a broad bandwidth desired for good stereo separation and negligible crosstalk between the different MTS components. However, dual-cavity notch diplexers introduce more group delay in the visual path, are more expensive, and should be tuned differently than single-cavity diplexers.

Many TV engineers have noticed a significant difference in the level of distortion when making performance measurements on an aural TV transmitter at points before the notch diplexer and after the notch diplexer. With the introduction of TV BTSC stereo broadcasting, the increase in THD distortion and stereo separation errors caused by the diplexer became a major problem which had to be solved.

The notch diplexer is a passive device and under first consideration it may seem strange that it can introduce nonlinear distortion and stereo separation errors. The basic problem is that the FM stereo signal is sensitive to the notch diplexer group delay and amplitude response over the occupied bandwidth of the FM signal.

The group delay and amplitude response of a single-cavity diplexer is shown in Fig. 27. The phrase, "single-cavity diplexer", means a single aural notch cavity is used in each branch of the diplexer. Dual-cavity notch diplexers use two aural notch cavities in each branch of the diplexer. The response curves shown in Fig. 27 are typical of a notch diplexer optimized for minimum aural reject power. The bandpass is somewhat narrow and the group delay is steep. Fortunately, the response

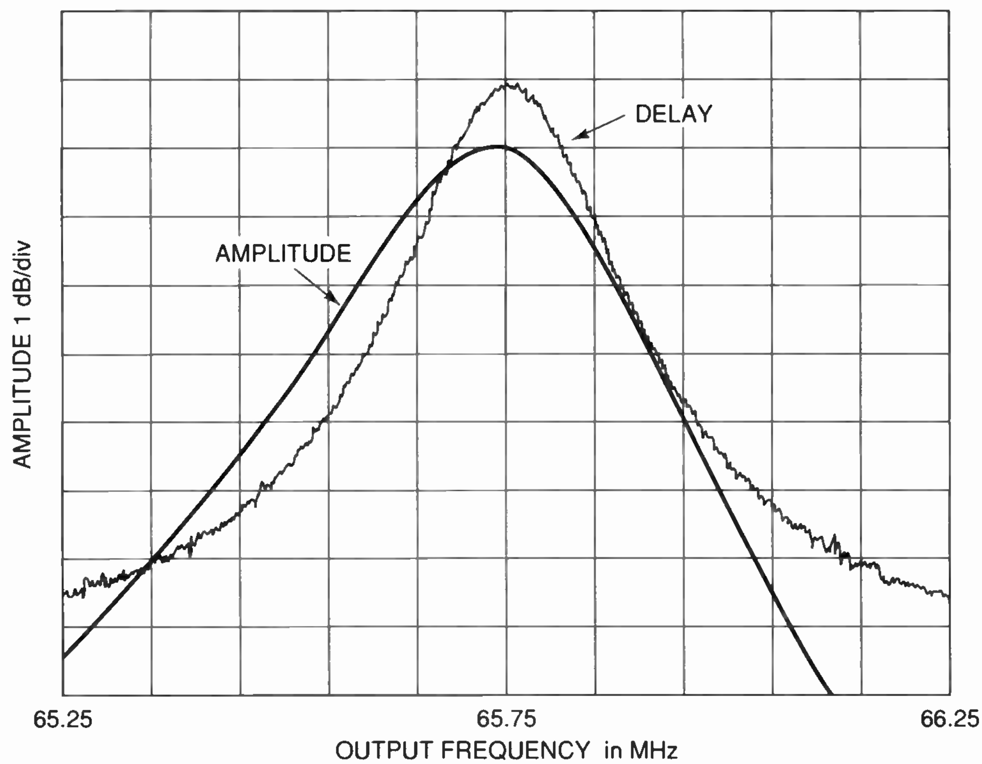


Figure 27. Typical single cavity diplexer amplitude and group delay.

curves in Fig. 27 show a high degree of symmetry which makes precorrection possible.

The group delay equalization concept as applied here to the FM modulated signal is essentially a feed-forward correction scheme. The phrase "feed-forward" is used here to describe the technique of generating correction signals early in the RF line up of a transmitter for the purpose of correcting distortions occurring downstream in the system.

Feed-forward correction signals are operated open loop (without feedback) and are manually adjusted for optimum operation and, once adjusted, left to operate in this manner.

IF Group Delay Correction

The equalization technique of adjusting the circuit to produce an inverse curve of that of the diplexer for correction is the same as that used on the visual transmitter.

There is however, a significant difference between group delay correction on a FM system versus an AM system. The aural transmitter must incorporate group delay correction at RF or IF to compensate for group delay distortions occurring in the output diplexer system. *Group delay correction before the FM modulation process is not effective for equalizing group delay occurring at RF.* The reason for this is that the occupied bandwidth of an FM signal increases or decreases as a function of baseband signal level. Fig. 28 shows the occupied bandwidth when the carrier is deviated 25 kHz and how it compares to a typical notch diplexer group delay curve. (Also note the precorrected group delay curve shown as an additional overlay.) When the baseband signal level is increased, the FM deviation will increase and a number of additional significant sidebands will be generated which will begin to extend beyond the acceptable group delay curvature region. The result is that distortion is observed in the demodulated FM signal.

In addition, as the baseband high frequency content increases, i.e., when switching from a mono to a stereo signal, significant sidebands extend even further outward increasing the demodulated distortion. Fig. 29 shows the spectral content of a TV stereo signal (with notch diplexer group delay curves, standard and precorrected, overlayed).

Simply put, there is not a direct relationship between the amount of group delay correction injected at baseband versus the amount of group delay correction achieved at the output of the RF system for a FM modulated signal.

To avoid the additional cost of extra cavities in the diplexer to achieve better stereo performance, an adjustable group delay circuit designed for the low power IF section of the aural exciter can provide the inverse group delay curve necessary to equalize the output. The corrector location in the aural transmitter system is shown in Fig. 30.

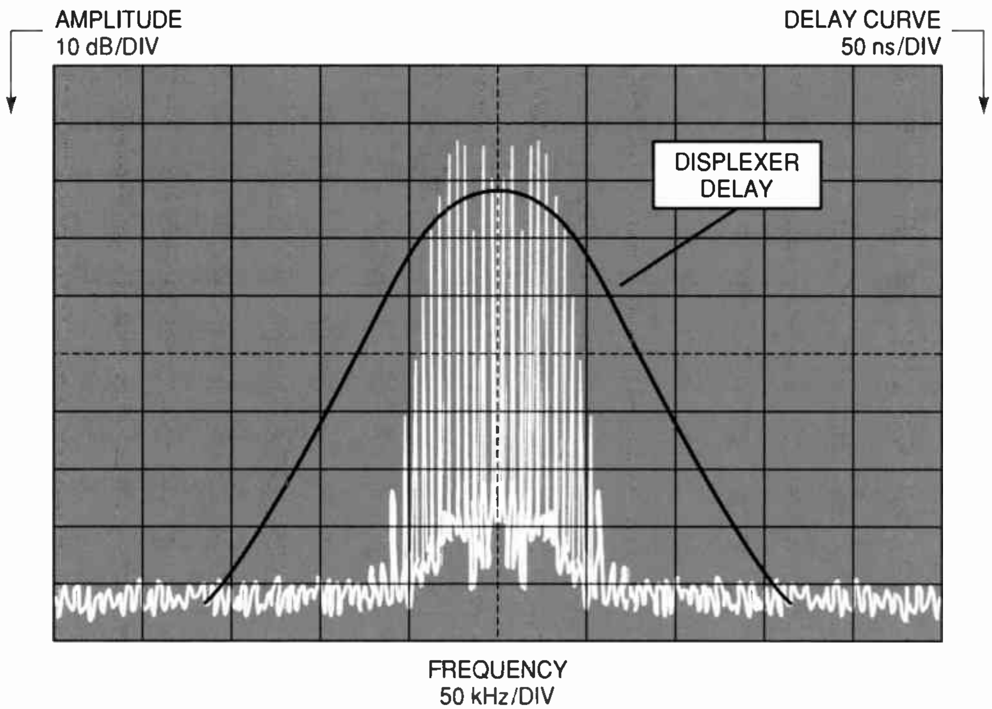


Figure 28. Occupied bandwidth with 25 kHz deviation with overlayed notch diplexer bandwidth.

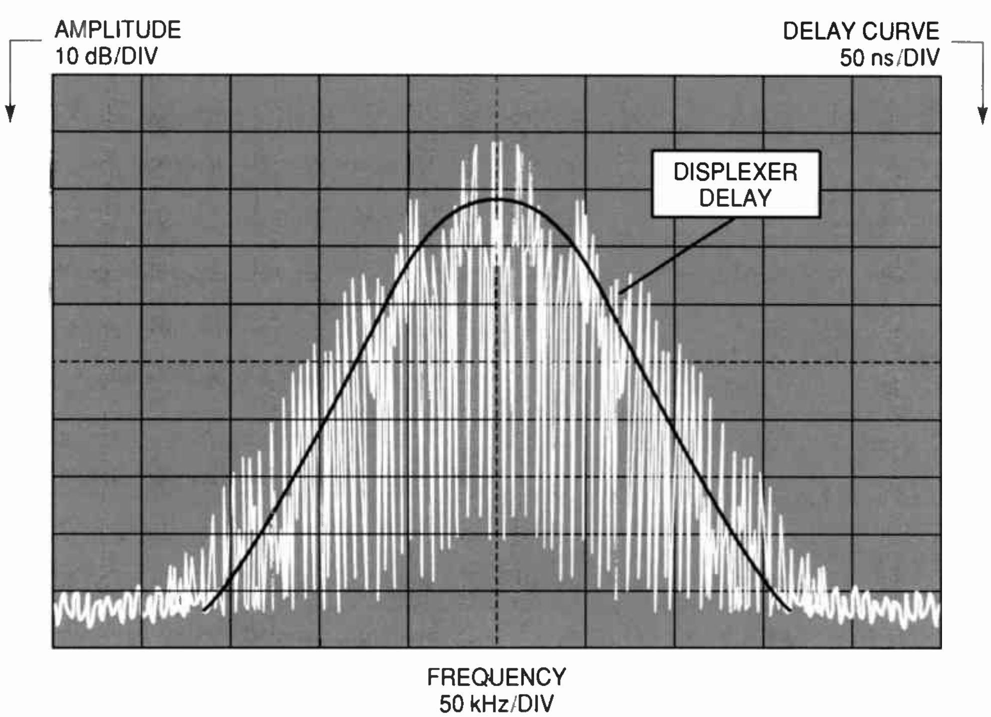


Figure 29. Occupied bandwidth of TV stereo signal with 55 kHz deviation and overlayed notch diplexer bandwidth.

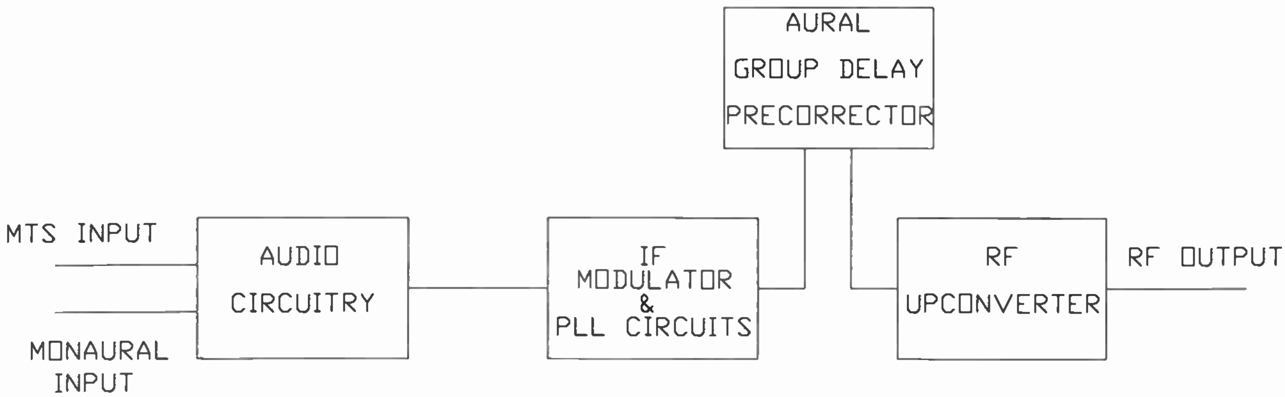


Figure 30. Aural exciter with group delay precorrector.

Equalization at IF frequency is effective because the up-conversion circuit uses a signal mixing process rather than a frequency multiplying process. In this manner, the precorrected sidebands are passed directly to the output for compensation.

Circuit Design

The circuit configuration is shown functionally in Fig. 31.

Results

The measured amplitude and group delay response of the group delay corrector is shown in Fig. 31. An ideal precorrection circuit without any circuit losses would have a flat response without a dip. The response dip shown in Fig. 32, however, is very useful because it also provides a first order correction to the notch diplexer amplitude response.

Fig. 33 shows the overall system equalization when the delay corrector is switched in and out of the circuit. The diplexer delay response contributes nearly all of the delay errors. The lower curve shows that a significant amount of equalization has been achieved over the occupied bandwidth of a stereo signal.

The effect on stereo separation is shown in Fig. 34 with the corrector switched in and out of the circuit. The curves generated in Fig. 34 are from measured data through the system with the response characteristics

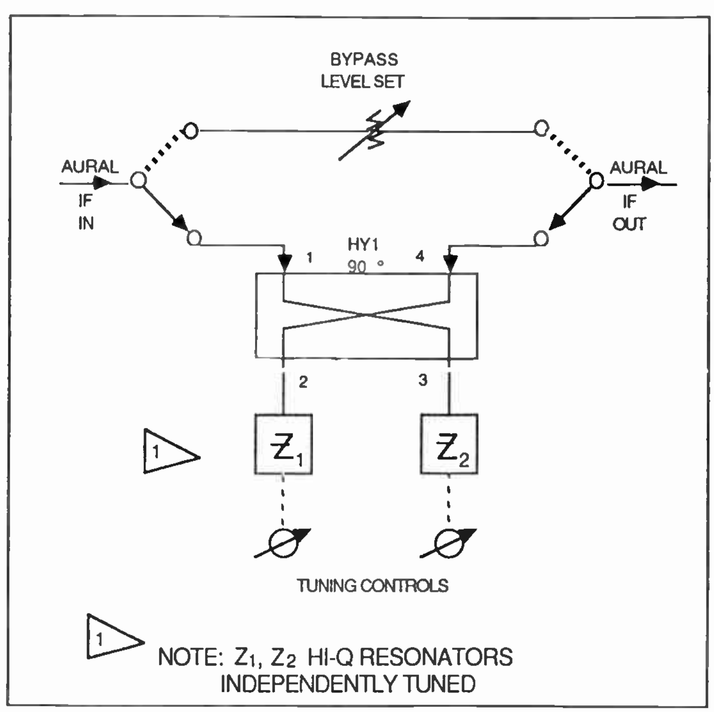


Figure 31. Aural group delay corrector functional block diagram.

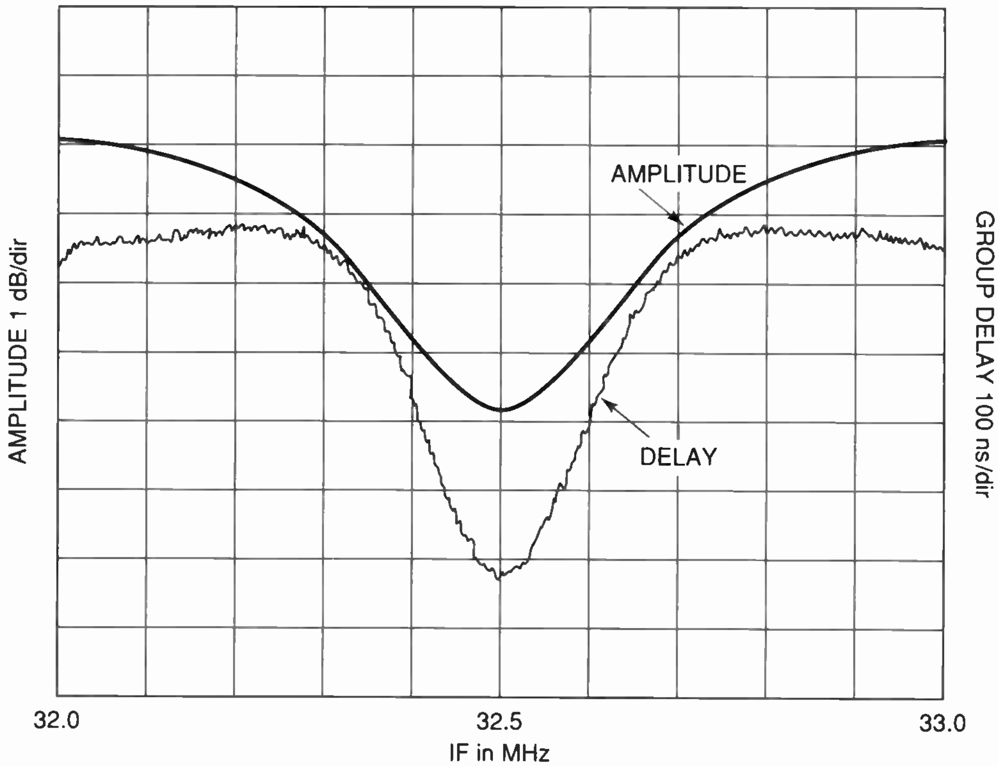


Figure 32. Aural group delay corrector measured results.

shown previously in Fig. 33. Fig. 34 shows that more than 10 dB stereo separation improvement can be obtained over midband audio frequencies.

SOLID-STATE TRANSMITTERS

More and more emphasis has been placed on maximizing the on-air time of TV transmitters. Recent technological advances in field effect transistors (FET) have made the development of solid state high power, linear amplifiers for TV applications both practical and cost effective.

Designing For Reliability

Because there are many factors which can affect the reliability of a TV transmitter, it must be understood that the application of solid-state high power technology in itself does not necessarily ensure a high reliability, low maintenance transmitter. Overall design philosophy, device technology, module design, control architecture, power supplies, cooling systems and cabinet design are some of the critical areas which also must be considered. By paying careful attention to the overall architecture and all of the subassemblies in the entire transmitter, a very reliable system results.

In a transmitter design which uses circuits in series with no system redundancy, it is clear that if one device fails then the entire transmitter will fail. Fig. 35 shows a system of three series devices with no redundancy.

If each device (a, b, c) has a probability of survival over a given period of time (P) of 0.5, then the overall probability of the system surviving P(s) over the same period of time is given by the formula:

$$\begin{aligned} P(s) &= P(a) \times P(b) \times P(c) \\ &= 0.5 \times 0.5 \times 0.5 \\ &= 0.125 \end{aligned}$$

If three identical devices are operated in parallel and only one is required for adequate operation of the system, the probability of the system surviving over the same time period is greatly enhanced. Fig. 36 shows this configuration.

The overall system reliability now becomes:

$$\begin{aligned} P(s) &= P(a) + P(b) + P(c) - P(a)P(b) - \\ &\quad P(a)P(c) - P(b)P(c) + P(a)P(b)P(c) \\ &= 1.5 - 0.25 - 0.25 - 0.25 + 0.125 \\ &= 0.875 \end{aligned}$$

System On-Air Availability

Related to reliability, but perhaps even more important to the broadcaster, is on-air availability. On-air availability is the percentage of time the transmitter

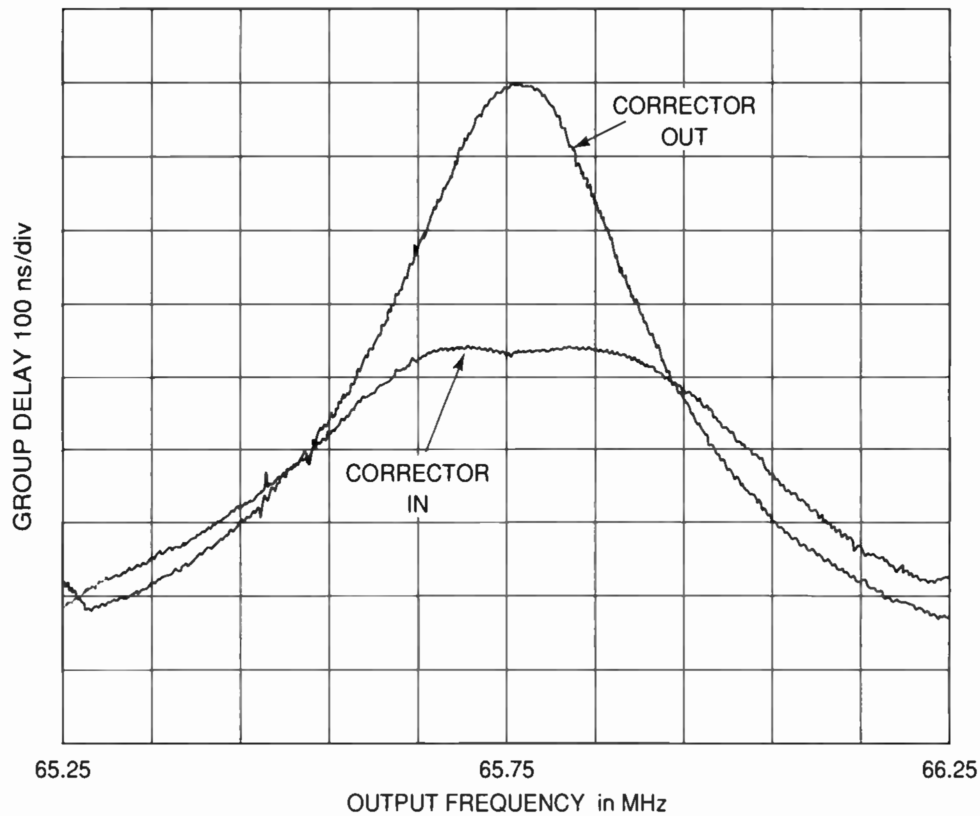


Figure 33. Overall transmitter delay with aural delay corrector switched in and out.

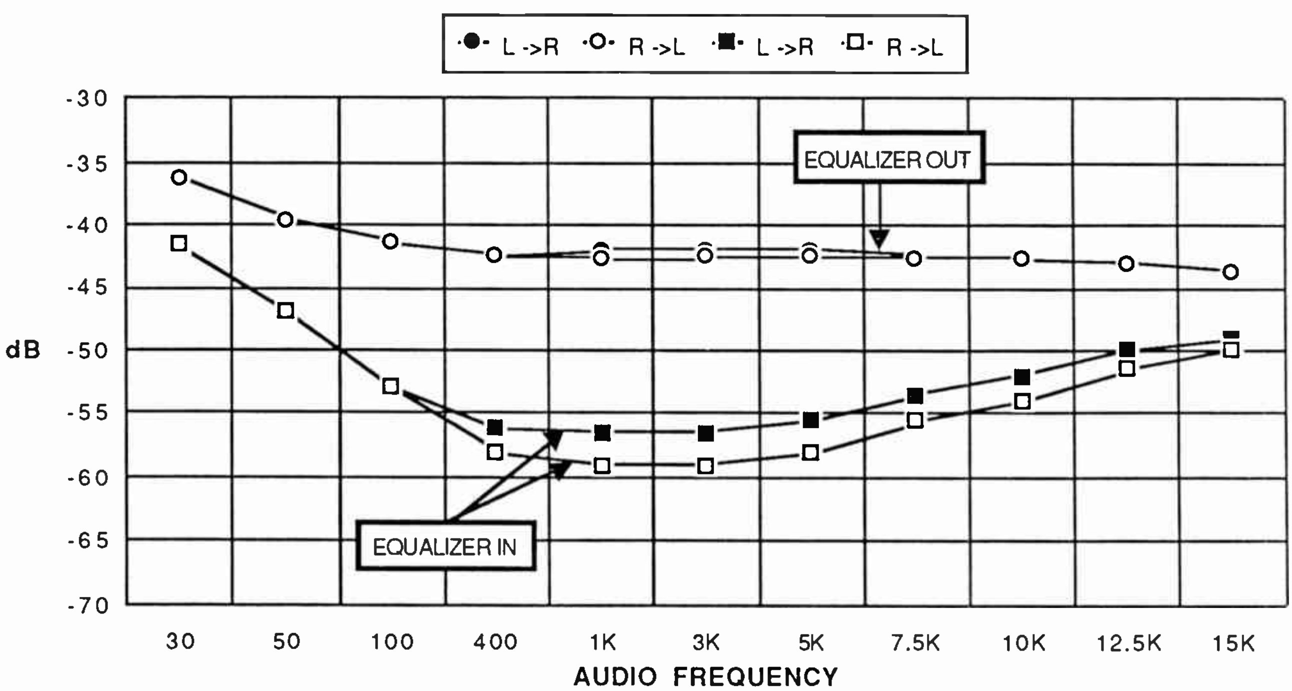


Figure 34. TV stereo separation with and without group delay equalizer.

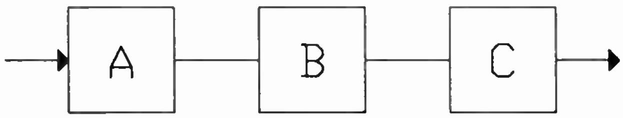


Figure 35. Circuits in series.

is in service, or could be in service. Availability is defined by the following equation:

$$\text{On-Air Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR} + \text{MPMT}} \times 100\%$$

where:

MTBF = Mean Time Between Failures (hours)
MTTR = Mean Time To Repair (hours)
MPMT = Mean Preventative Maintenance Time (hours)

From this equation, one can see that there is little point in designing a transmitter that has an extremely high MTBF figure if, due to poor design and mechanical packaging, it takes an inordinate length of time to make repairs, or the transmitter has to be shut down frequently for routine preventative maintenance.

Many stations have very short sign-off windows or operate 24 hours a day. This often results in a less-than-optimum maintenance schedule which can lead to premature failure or out-of-tolerance operation. One way to reduce the amount of off-air maintenance time is by making provisions for on-air maintenance or to have redundant transmissions. This significantly reduces the MPMT factor.

Several design factors should be considered for a transmitter for optimum on-air availability:

1. Very high reliability for the fundamental circuits.
2. Provision for fast and easy access to all modules and subassemblies.

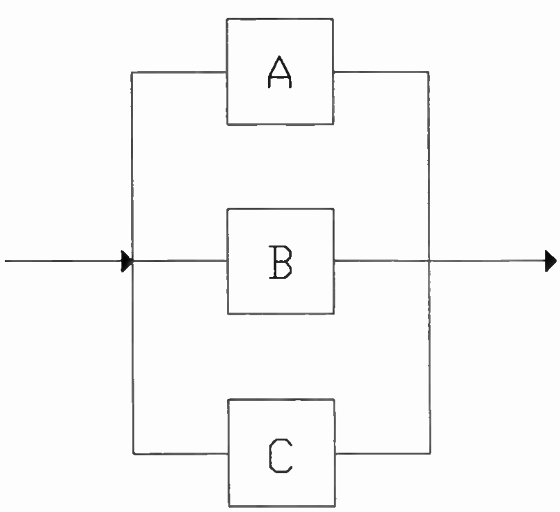


Figure 36. Circuits in parallel.

Figure 37. Modular VHF-TV transmitter (Model HT-30HS). (Photo courtesy of Harris corporation)

3. Maximum use of like part and subassemblies. Because only a few items would be needed, this allows most TV stations to maintain a full inventory of spares. If spares are on hand, it follows that the repair time will be much shorter.
4. Repair of transmitter at module or subassembly level. Modules which have been removed can then be repaired by station personnel or returned to the manufacturer for exchange.

Figs. 37 and 38 show two different modular solid state VHF-TV transmitters.

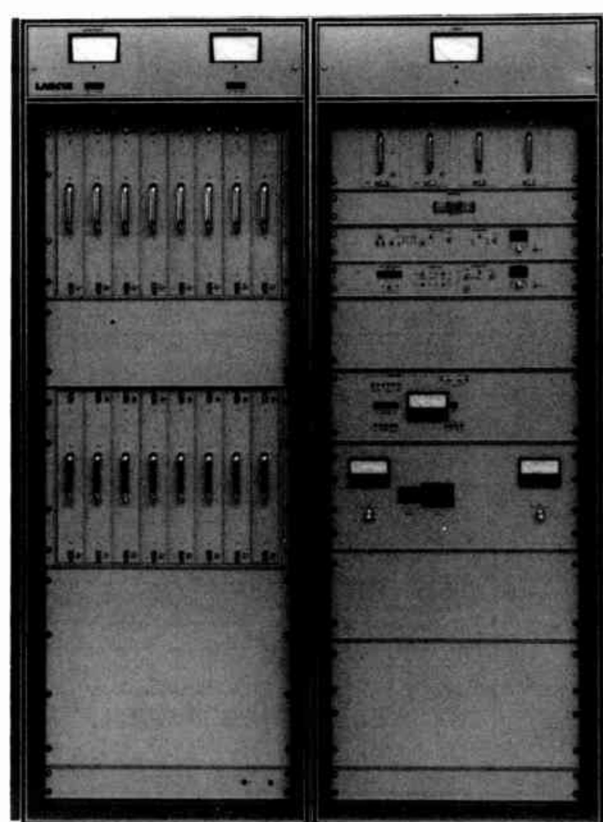


Figure 38. Modular VHF-TV transmitter (Model TTS-12M). (Photo courtesy of Larcan)

By combining a few or many RF modules, it is practical to create transmitters at any power range up to 60 kW. Solid state transmitters maintain their performance over extended periods of time due primarily to the fact that they have no tuning controls nor filament emission degradation with time. Another significant advantage is that no warm-up time is required. Solid-state transmitters can be producing full rated power within seconds of activation. Since solid state transmitters are of a different structure than tube transmitters, it is worthwhile to review some of the new transmitter architecture.

AGC

An automatic gain control (AGC) system is used to maintain constant power output from the transmitter. Ambient temperature changes will cause gain changes in a solid state amplifier or a faulty RF power module will cause a reduction in final output power. Then, RF drive power must be boosted to maintain constant power output. In most cases a detected RF sample of the PA output is fed to an input of a comparator. The exciter output sample or a voltage proportional to the exciter output is applied to the other input of a comparator. The DC output is then integrated and fed to an attenuator which varies the RF drive level at a low power level.

Combining Multiple Amplifier Cabinets

When combining RF power amplifiers, they must be matched in phase and gain for maximum power to the antenna and minimum power to the reject load. Electronic phase shifters and attenuators must have the capability for remembering their settings in case of AC power failure.

RF Amplifiers

Combining several RF power modules to achieve the desired transmitter output power increases the parallel redundancy and the on-air availability described earlier. Choosing the optimum power level for the PA modules of a solid-state transmitter is not an easy task. This point is illustrated in Fig. 39.

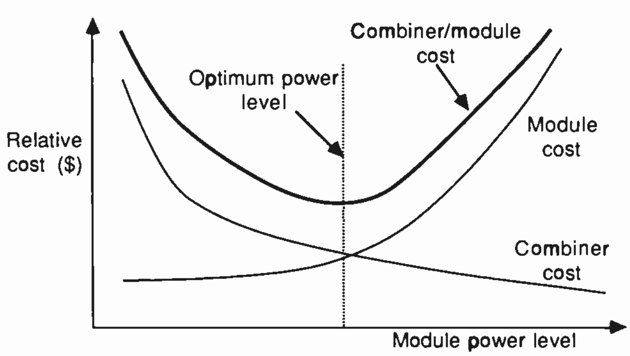


Figure 39. Module cost versus module power level.

A 1 kW output power module has been adopted in nearly all manufacturers' PA modules based on overall cost, practical weight, and size limitations.

Self-protection of each PA module against various fault conditions is good engineering practice. It is desirable that one sub-assembly failure not cause another sub-assembly failure. By using self-protecting modules, the cabinet control logic and overall transmitter control logic can be kept very simple, thus improving overall reliability. Self-diagnostics for the module aid in minimizing the time to repair. Protection from over-voltage, RF overdrive, VSWR, overtemperature, and ensuring proper load sharing among devices is essential to maintaining amplifiers for long life.

Modular amplifiers, which can be removed while "hot" can help in improving overall on-air availability. If an amplifier module fails, the transmitter can continue to function properly indefinitely without disrupting transmitter operation. If a spare PA amplifier is kept at the station, it can be used while the failed unit is repaired or returned to the manufacturer.

Temperature compensated, regulated bias supplies for the amplifiers are essential. Otherwise, performance and power output would vary roughly in direct proportion to temperature or supply voltage changes.

Amplifier level faults are most easily verified by swapping modules in different slots. If the fault follows the module it is an internal module problem but if the fault remains at the same slot after substituting the module, then the problem may be somewhere else in the system.

Solid-State Devices

Both bipolar and field effect transistor (FET) technology exist today as suitable RF amplification devices. Power amps are operated in class AB for the best trade-off of efficiency and linearity. The driver amplifiers usually contain class A operated RF amplifiers. Although both types have their merits, FETs have some advantages over bipolar devices.

FETs have a higher amplification factor than bipolar transistors, helping to reduce the number of devices required as driver stages. Higher supply voltages help to reduce the current capacity of the power supply. Simpler bias circuitry minimizes parts count.

Cooling System

Proper cooling of the solid-state modules is extremely important in obtaining a high MTBF. The MTBF of a FET or transistor essentially doubles for every 10°C drop in the actual FET junction temperature.

Distributed cooling systems employing more than one fan offer good redundancy. Current motor/fan technology has matured to the point where a few larger direct drive fans can be equally as reliable as, or more reliable than many smaller fans. Since many RF power amplifiers may be employed, a large volume of air is needed to adequately cool the heatsinks. Low pressure fans or blowers may be used if heatsink fin density is not high. This aids in reducing audible noise generation from the transmitters. The heat is distributed over a large volume of air and thus the temperature rise is relatively low; on the order of five to seven degrees Celsius.

Power Supplies

Power supply design is another area which is critical to the reliability of a solid-state transmitter. Since FET and bipolar devices are low voltage devices, the power supplies which serve them provide low voltage and high current. Therefore, high reliability connections must be guaranteed in the DC distribution. Since available power output from a FET or bipolar transistor varies roughly as

$$P_o = \frac{(Vcc)^2}{2R_L}$$

it is desirable that the supply remain very tightly controlled over incoming AC line variation.

Also, any anomaly present such as voltage or current transients or voltage sags at the AC input to the power supply should be significantly suppressed before reaching the amplifier transistor device. Transmitters should successfully pass the applicable portions of the ANSI/IEEE C62.41 transient testing standard (also referred to as the IEEE-587 standard).

Since the amplifier's current demand will vary with picture level, the power supply output voltage must remain stable from very little load (i.e., white picture) to very large loads (i.e., producing sync output power.)

Efficiency of the power supply is important since the lost power appears as waste heat energy.

Control Systems

If individual amplifier modules and power supplies are self-protecting, control and monitoring functions can be kept simple and straightforward.

One approach for the control system is to use a single controller to control and monitor all the functions of the transmitter.

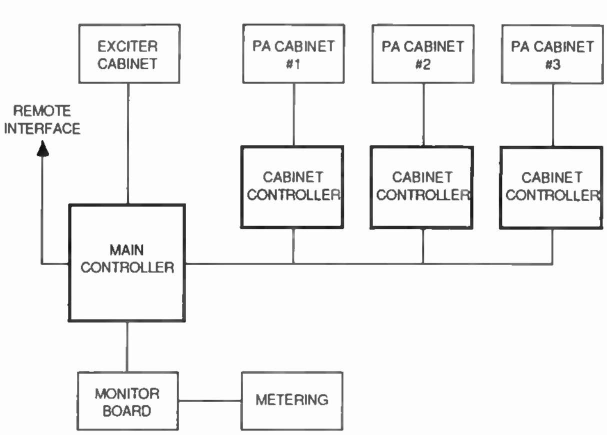


Figure 40. Distributed control and monitoring.

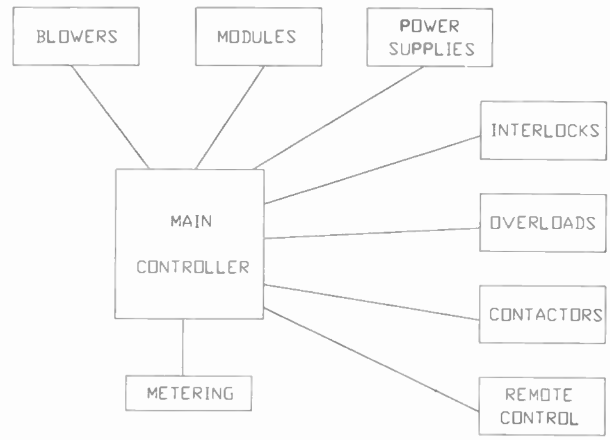


Figure 41. Centralized control system.

Another approach is to distribute the control system throughout the transmitter. The distributed control system can be designed so that the failure of any individual controller does not affect the operation of the others.

Both distributed control system architecture and centralized control systems are shown in Figs. 40 and 41.

After AC power failures, the controller should have back-up memory to restore the transmitter to the same operating condition as before power was interrupted. This would apply to all the control systems within the transmitter.

A system of indicators is essential to quick fault diagnosis. Some typical status conditions which may be displayed are: exciter fault, VSWR fault, VSWR foldback, power supply fault, controller fault, air loss, door open, failsafe interlock, phase loss, module fault, visual drive fault, aural drive fault, and external interlock(s).

VSWR foldback reduces power during high VSWR operation, such as antenna icing, and restores RF power back to normal when difficulties are removed. This is a technique used to keep transmitters on the air. Other options used to enhance the on-air capability of solid state transmitters may include dual exciters, 20% aural power, and redundant drive chains. An example of a VSWR foldback block diagram is shown in Fig. 42.

AC Distribution

Tube-type transmitters typically have one AC service connection to the transmitter. A more reliable method is to provide power to the modular RF amplifier cabinets through a distributed AC power feed system in which each cabinet is protected by a separate AC breaker that is external to the transmitter, as shown in Fig. 43. This concept also allows a cabinet to be safely serviced while the remaining cabinets are operational. Phase monitors guard against low voltage, loss of one phase, or reversal of the phase sequencing.

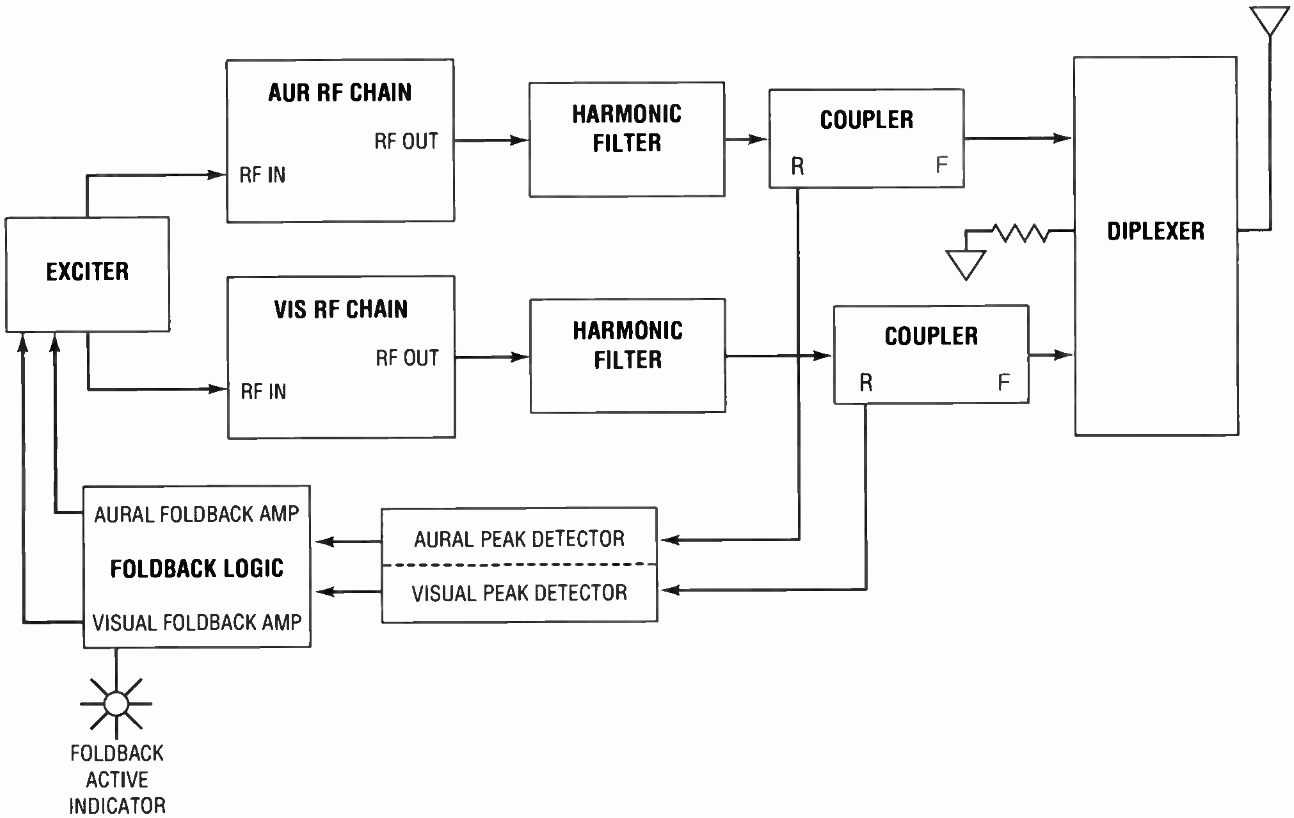


Figure 42. VSWR foldback block diagram.

Combiners and Dividers

There are several choices as to the method used for dividing and combining RF power for the solid-state visual and sound amplifier modules. The most predominant method used has been with in-phase N-way ring combiners. Two common examples are described below.

Microstrip Wilkinson Combiner

Fig. 44 shows an example of this type of combiner. Microstrip is used as the transmission line to carry the RF power. When all amplifiers are operating, equal voltages are presented to each side of the load resistor so that no power is dissipated. When an amplifier failure occurs, the power is distributed between the loads and the output. In this type of combiner, the impedance of the transmission lines and the length of the lines may be varied to achieve the desired impedance transformation.

Balanced reject loads are used to absorb RF power in case of an amplifier failure and to provide isolation for the other amplifiers.

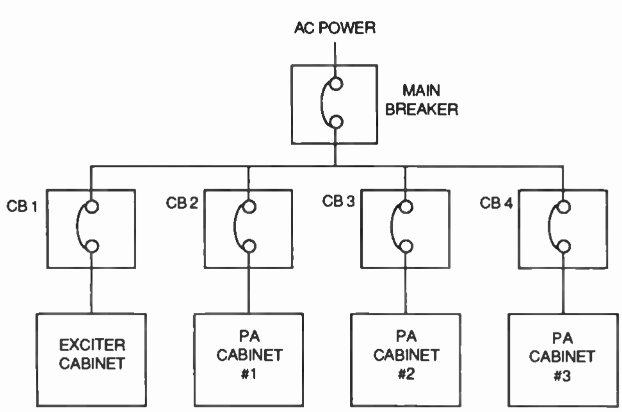


Figure 43. Distributed AC power system.

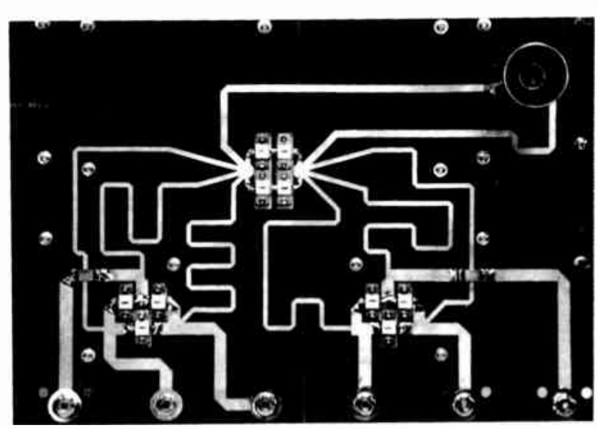


Figure 44. Microstrip Wilkinson combiner.

The combiner is housed in a shielded casing so that outside fields cannot alter the balanced configuration. In case of an amplifier failure, the reject loads conduct heat through the flange to the heatsink where the heat is exchanged to the moving air stream. This type of combiner is used generally for lower numbers of amplifiers (i.e., two to six).

Ring Combiner

Fig. 45 shows an example of this type of combiner. The higher power handling capability of the coax lines used in this combiner will also allow a large range of amplifiers to be combined, (i.e., two to 20). It also provides isolation from one amplifier to another using reject loads which are not in the direct RF path to the output. The operation of the multiport combiner is easily understood by first understanding a two-way combiner.

Refer to Fig. 46 for a simplified version of the combiner. Each of the transmission lines is a quarter wavelength. When equal voltages are applied to both input ports (both amplifiers operating) the combined signal arrives at the output. This is true for three reasons: (1) the distance from each input port to the output is electrically equal whether the signal follows the shorter path or the longer path, (2) the signal from one amplifier arrives at the load port out of phase with that from the other amplifiers so no power is dissipated, and (3) the signals from amplifier 1 arriving at input port 2 via the short and long paths are out of phase and vice versa. *Thus under normal conditions, all of the power appears at the output, none is absorbed in the loads, and there is complete isolation between amplifiers.*

Power is absorbed in the load resistors only when an amplifier is not operating. Assume that only amplifier number 1 is operating. The signal path is electrically equal not only for the long and short paths to the output but also for the right and left paths to either load resistor. The power from the operating amplifier

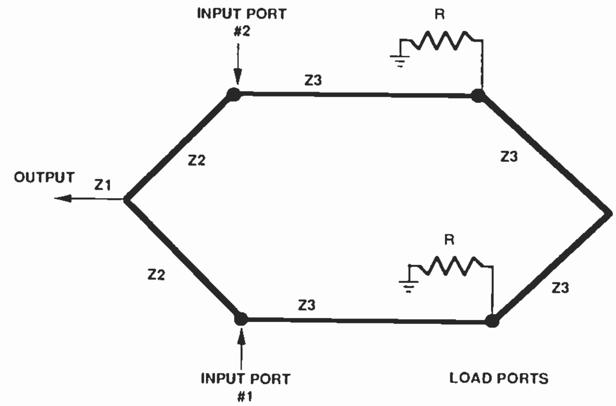


Figure 46. Two way ring combiner.

is split equally between the output and the isolation loads. Due to the isolation inherent in the network, the input ports remain matched even when one or more amplifiers are removed.

Since the transmission line that connects the amplifier port to the output port is a quarter wavelength long, and 75 ohm transmission line is used, the 50 ohm amplifier impedance is transformed to: $(75/50) \times 75 = 112.5$ ohms. For combining N amplifiers, the impedance at the output combiner junction is 112.5/N. This impedance is then matched to 50 ohms.

In this type of combiner, the reject loads may also be mounted to a grounded structure. Removing the heat from this structure should be done as quickly and efficiently as possible. One method to accomplish this is to use a heat pipe for the mounting structure. In case of an amplifier failure or removal, the reject load temperature rises, the fluid in the lower section of the heatpipe heats up until it vaporizes. The vapor rises to the finned area where the heat is exchanged to the moving air stream. The vapor condenses as it releases its heat and returns to the bottom of the pipe to absorb more heat.

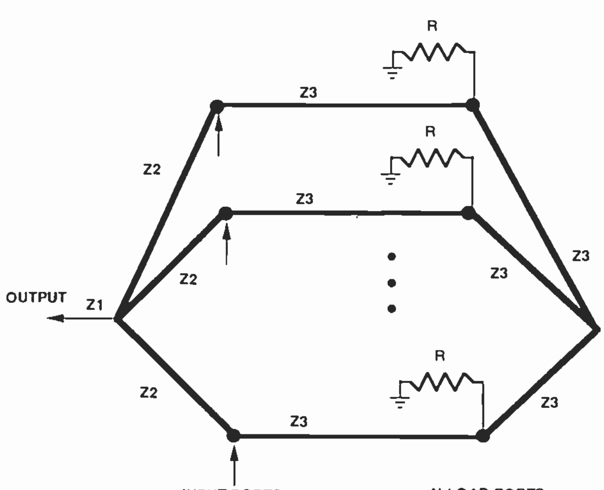


Figure 45. N-way ring combiner.

VHF TUBE TRANSMITTERS

Triodes and tetrodes are still used in the RF power stages of television transmitters. This section will highlight some of the facts on high power gridded transmitter tubes. The parts of the power tube will be reviewed.

General

A vacuum tube uses grids to control the flow of current through the tube. The cathode emits the electrons which travel to the plate. The plate is more positive than the cathode. The grids control this flow of electrons, thereby controlling the plate current. They also can modulate the electron stream causing the plate current to have the same waveform as the grid voltage, depending on class of operation.

A triode is a three element tube consisting of:

1. A heated cathode which emits electrons.
2. A control grid which modulates the electron stream in accordance with the DC and AC voltages impressed between the control grid and cathode.
3. A plate which accepts the electron stream. The plate is positive with respect to the cathode.

Input signals, DC and AC, are applied between control grid and cathode. Output signals, DC and AC, flow to the tube output load impedance by way of the plate and cathode.

A *tetrode* is a four-element tube which has a *screen grid* added between the control grid and plate. The screen grid acts as an electrostatic shield which helps to isolate the plate output signal from the control grid input signal, and is operated at a positive DC potential with respect to the cathode. This potential is much lower than the DC plate voltage and is usually operated at AC (RF) ground potential.

The tetrode has higher gain than the triode. The plate voltage has minimum effect on the plate current as long as the instantaneous plate voltage is greater than the screen grid DC voltage. The plate voltage is normally allowed to swing close to, but not below, the value of DC screen grid voltage. If the tetrode screen grid is not maintained at AC (RF) ground potential, the tube can take on the characteristics of a triode. The tetrode input bias and signal are applied between the control grid and cathode and are similar in performance to the triode. The output AC (RF) signal current path in the tetrode is from the plate, through the load impedance, and returns through the screen grid.

The Cathode

A power grid tube will have high peak and average current levels. Therefore, the cathode must be hot to emit many electrons. It normally requires large heater power at low voltage and high currents. DC current is often used to reduce AC noise levels. It is usually directly heated (the heater in the cathode) to provide quick warm up with lower heater power. Some type of cooling to remove excess cathode heat from the tube filament/cathode contacts and the tube socket may be required.

The Grids

The voltages (DC, AC, and RF) of all grids are measured with respect to the cathode. Grids can have current flow which is positive when the grid accepts electrons from the electron stream or negative when the grid emits electrons.

Grids can emit electrons and negative grid current can occur because they are located close to the hot cathode. The grids can become very hot and have small amounts of thermionic emission (primary emission) or electrons emitted by being bumped off by other electrons (secondary emission).

As tubes age, some of the cathode electron emitting coating may be deposited on the relatively cooler grids thereby increasing their tendency to emit electrons.

The control grid has a negative DC bias and an AC driving signal superimposed upon it. The instantaneous grid current can be negative when the grid voltage is swinging maximum negative, zero when the grid is only slightly negative, or positive when the grid is positive.

The screen grid has a positive DC bias and is at AC ground potential. The instantaneous screen grid current is negative when the plate voltage swings maximum positive, and positive when the plate voltage swings close to or below the screen grid DC voltage, and zero when the plate voltage is between these extremes.

Both grids have instantaneous positive, zero, or negative currents that follow the voltage swing. The average of these currents is the DC grid current. Since both DC voltage and current are present in both grids, the grids can dissipate power. The control grid also has AC (RF) voltages and currents present so that AC power dissipation in grids must also be considered.

The DC plate power that is applied to the tube is either converted to AC (RF) output power or dissipated as heat. This heat is developed in the plate itself and must be removed by air or water cooling.

Below 30 MHz, RF amplifiers use lumped components to implement the matching circuits for the tube input and output. At VHF and UHF frequencies, several problems make the use of lumped components (L&C in the same purchase) impractical.

The Transmission Line Cavity

As frequency increases, resonant circuits are smaller to reduce inductance and capacitance, larger in diameter to reduce skin effect, and closer to the tube to reduce the effects of stray inductance, and there is difficulty in predicting exactly what values of resistance, inductance, and capacitance a component or circuit may have.

These problems can be managed in low-power circuits, but with high power circuits, arcs and shorts due to high DC and RF voltages become a problem. Larger size and spacing of components are a good start towards arc and short prevention, but this is in opposition to the smaller size and spacing needs dictated by the high frequency operation. Also in high power circuits, the unpredictability of the circuit values of R, L, and C make it difficult to control the vitally important parameters of dissipation, efficiency, and reliability of operation.

One solution to the above problem is the resonant transmission line cavity amplifier. In this type of amplifier the tube becomes part of a resonant transmission line. The elements of these tubes themselves are arranged to look like concentric coaxial transmission lines. The design of these power tubes stresses low inter-electrode capacity and low distributed inductance. The stray inter-electrode and distributed capacity and inductance of the tube becomes part of the resonant transmission line. The resonant transmission line is physically larger than the equivalent lumped constant LRC resonant circuit operating in the same frequency. This larger physical size aids in solving the

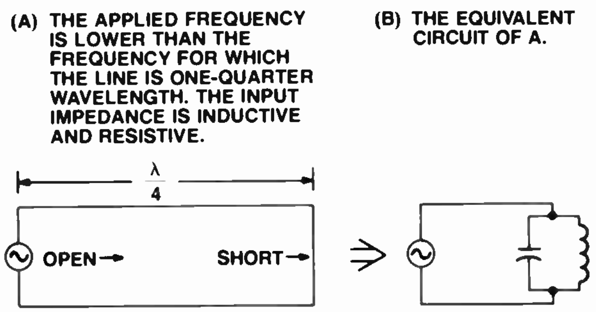


Figure 47. Shorted quarter wavelength line.

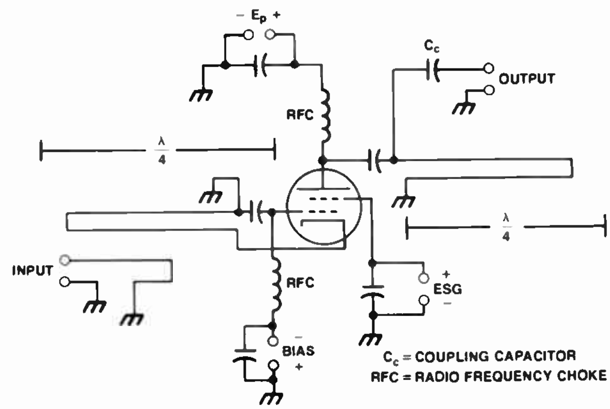


Figure 49. A shorted quarter wavelength transmission line amplifier.

high power operation problems of skin effect losses, prevention of arcs and shorts, and yields reliable and predictable operation. A commonly used transmission line cavity amplifier uses a quarter wavelength transmission line as its resonant element.

The Shorted Quarter Wavelength Line

A shorted quarter wavelength transmission line has a very high, purely resistive input impedance. Electrically, it looks like a parallel resonant circuit as shown in Fig. 47).

The Shorted Transmission Line Less Than a Quarter Wavelength Long

When the physical length of the line is less than one quarter wavelength, the impedance will be lower and the line will look inductive. Refer to Fig. 48. This inductance will be used to resonate with capacitive reactance in the tube and surrounding circuit.

Quarter Wavelength Cavity and Amplifier

In Fig. 49, shorted transmission lines are used to resonate the inputs and outputs of this amplifier. Notice that the length of the lines is less than a quarter wavelength but the tube's shunt input and output capacity and its series lead inductance will electrically lengthen and resonate the transmission lines. The input is shown inductively coupled, but it could just as easily have been capacitively coupled to the cathode. The input could also have a lumped constant resonant circuit or a transmission line resonant circuit since its power level is low. The output coupling is capacitive, but it also could have been inductive. The construction of the tube lends itself to transmission line tuning techniques.

Coaxial Construction of a Tetrode RF Power Amplifier Tube

The Anode (Plate)

The plate resembles a copper cup with the plate contact ring welded to the mouth and the cooling fins silver soldered or welded to the outside of the cup as shown in Figs. 50 and 51.

The contact ring is bonded to the base ceramic spacer through a strain isolation ring. This ceramic spacer is the same ceramic that is shown above the screen contact ring shown in Fig. 52.

The Screen Structure

The screen grid consists of many vertical supports fastened to a metal base cone. The other end of the metal base cone fastens to the screen contact ring. The inductance of the individual vertical supports is reduced by building the screen grid of many of vertical

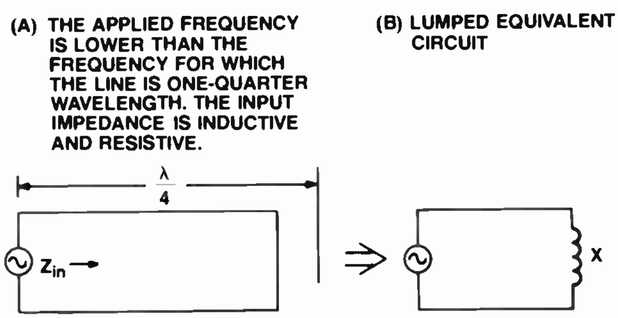


Figure 48. Shorted line less than one-quarter wavelength.

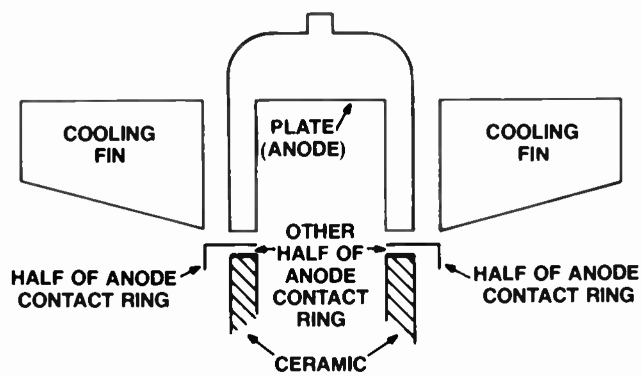


Figure 50. Cutaway view of the anode structure.

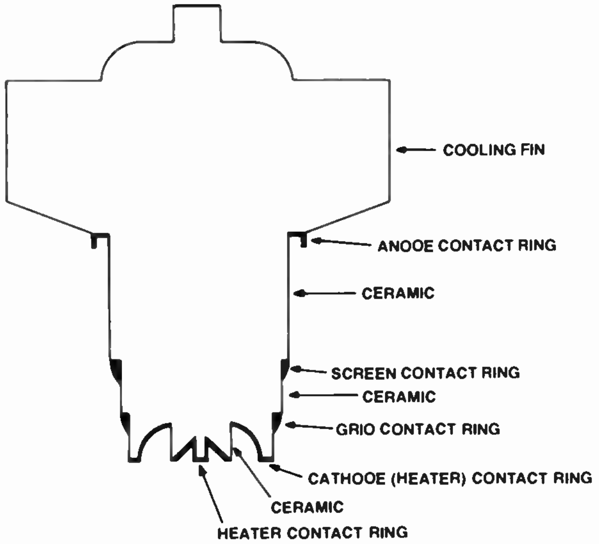


Figure 51. Cutaway view of the exterior of a RF tetrode.

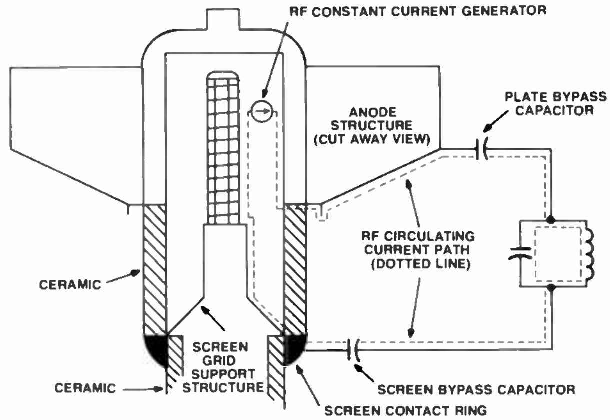


Figure 53. Showing the plate and screen assembly and RF circulating current path (dotted line).

supports in parallel. The vertical supports are held rigid by horizontal rings welded to them and a metal cap on the top of the assembly. The screen contact ring, metal base, and metal base cone also function to reduce lead inductance and RF resistance due to skin effect (refer to Fig. 52).

A cutaway view of the plate circuit and the screen circuit in Fig. 53 shows a concentric construction that resembles a coaxial transmission line.

Consider that the output RF current is generated by an hypothetical current generator between the plate and screen grid. The RF current travels along the inside of the plate structure on its surface (skin effect), through the ceramic at the bottom of the anode contact ring, around the anode contact ring, across the bottom of the fins, and to the band around the outside of the fins. From here it flows through the plate bypass capacitor to the RF tuned circuit and load, and returns to the screen grid. The return current travels through the screen contact ring, up the cone, and up the screen bypass capacitor, then through the screen grid to return to the hypothetical generator. The screen grid has RF current returning to it but due to its low impedance, the screen grid is at RF ground potential. The RF current generator appears to be feeding an open ended transmission line consisting of the anode (plate) assembly, and the screen assembly. The RF voltage developed by the anode is due to the plate impedance presented to the anode by the resonant circuit and its load.

The control grid assembly and the cathode assembly are also cylindrically constructed and concentric. The control grid assembly is constructed similarly to the screen grid but slightly smaller. Fig. 54 shows the screen grid, control grid, and the cathode assemblies as they are placed in the tube.

Fig. 54 also shows the current path of an RF generator feeding a signal into the grid/cathode assembly. It resembles a transmission line terminated by the RF resistance of the tube's electron stream. The outer contact ring for the cathode heater is the inner conductor of this transmission line.

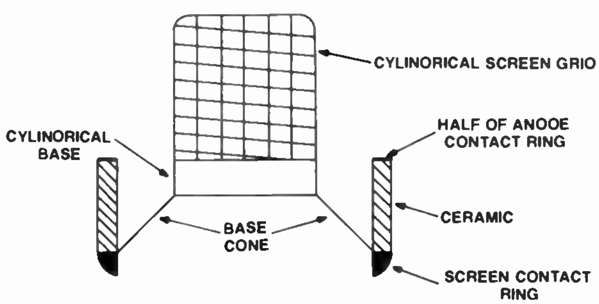


Figure 52. The screen grid assembly.

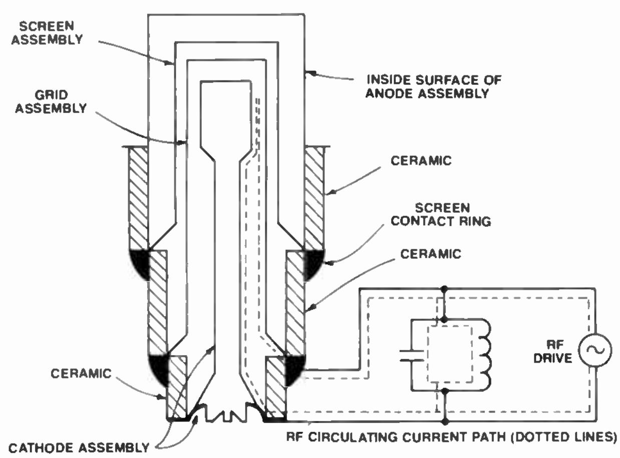


Figure 54. Showing details of assembly of grids and cathode components and the simplified RF input circuit (bias circuit not shown).

Double-Tuned Tube Power Amplifiers For TV Transmitter Applications

Television RF power amplifiers use tetrode tubes because of their high gain, good linearity, good efficiency, and high power advantages. A double-tuned output circuit is used to achieve proper bandwidth and efficiency.

As an example, the double-tuned overcoupled amplifier shown in Fig. 55 should not be thought of as an amplifier and a socket into which a tube is placed. The tube is an integral part of the cavity. The internal electrical properties of the tube will determine the amplifier gain and power handling capabilities. It will also dictate the dimensions and functions of the circuitry.

Effects of Tuning an Overcoupled P.A.

All double-tuned overcoupled visual power amplifiers have four controls to accomplish output tuning.

Plate tune (primary tune) resonates the plate circuit and tends to tilt the response and slide it up and down the bandpass (shown as A on Fig. 56, also Fig. 57 and Fig. 58).

Coupling sets the bandwidth of the PA. Increased coupling increases bandwidth and lowers the PA efficiency. When the coupling is adjusted, it can tilt the response and change the center of the bandpass necessitating the readjustment of the plate tune control (shown as B on Fig. 55, also Fig. 56 and Fig. 58).

Secondary tune resonates the secondary cavity and will tilt the response if adjusted. Generally it will not slide the response up and down the bandpass as will the primary tune (shown as C on Fig. 55, also Fig. 56 and Fig. 59).

Loading (secondary load or output load) determines the value of ripple in the response. Heavier loading (maximum C or minimum L) creates a haystacked response and light loading creates excessive ripple. Adjustment of the loading control usually tilts the response and changes bandwidth. This necessitates readjustment of secondary tune and coupling. In some cavities, primary tune may also have to be readjusted. The coupling control will also effect the value of ripple, but its greatest effect will be on the bandwidth (shown as D on Fig. 55, also Fig. 56 and Fig. 60).

Tuning For Power

Sweep the entire transmitter into the antenna or dummy load at 100% power. If tuning of any part or

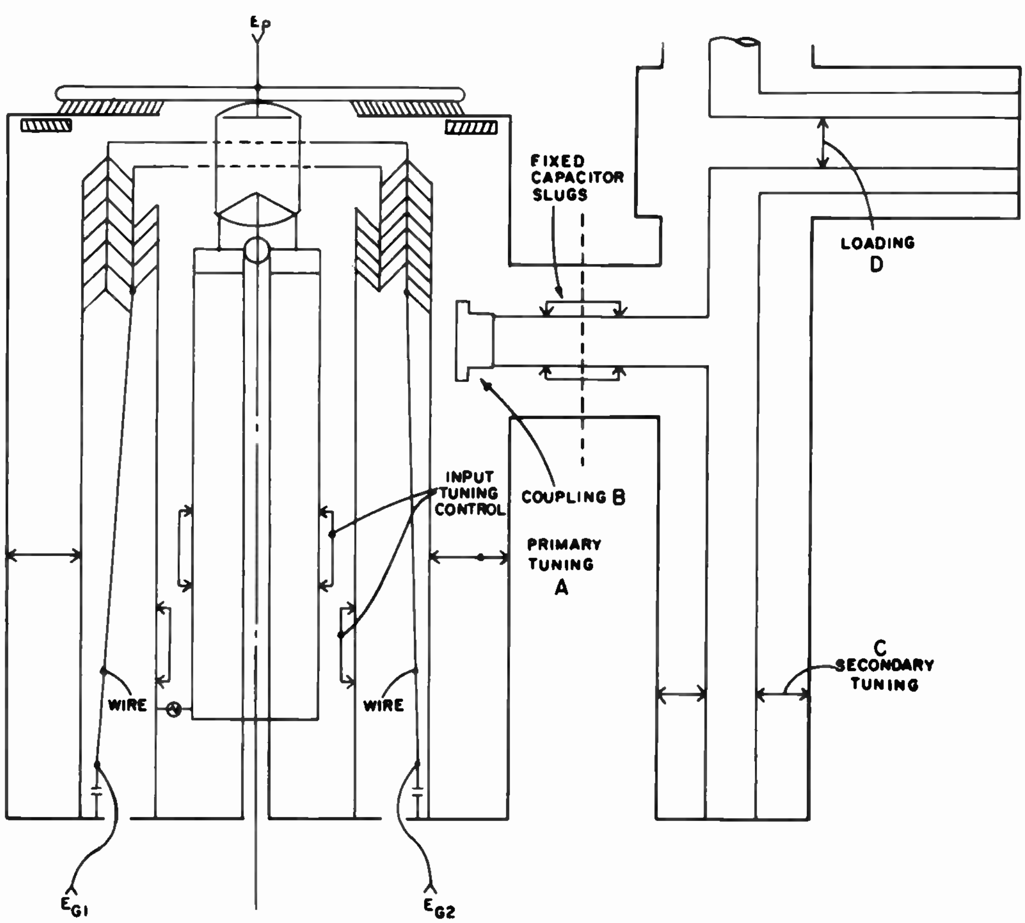


Figure 55. High band coaxial transmission line cavity amplifier for TV use.

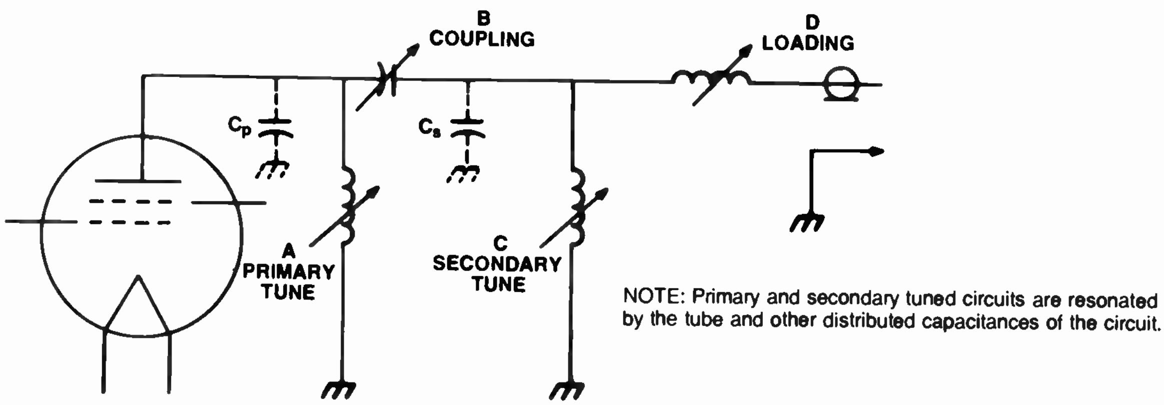


Figure 56. Double-tuned power amplifier.

all of the transmitter is in doubt, check the following. Observe the frequency response, reflected power (VSWR), plate current, screen grid current, and grid current. If the output response is not right, check each stage, starting at the exciter, to be sure each stage is flat and has the correct bandwidth. The final amplifier should be the narrowest of all stages.

After the transmitter sweep is completed and the overall response has the proper bandwidth, transmit a black picture, (video set at blanking, 75% modulation), operate into the station load and check for:

1. Proper sync level at the transmitter output. Adjust the visual exciter linearity corrector as necessary.
2. Excessive plate dissipation.

Plate dissipation = (Plate voltage × Plate current) − (Average RF output power)

3. Excessive screen current. The screen current will increase rapidly at higher power if the PA is lightly loaded. Very light loading can cause sync compression.
4. Excessive grid current, if PA is loaded too heavy, can cause sync compression.
5. Amplifier output efficiency. It can also be a guide to proper operation of the stage. The efficiency at

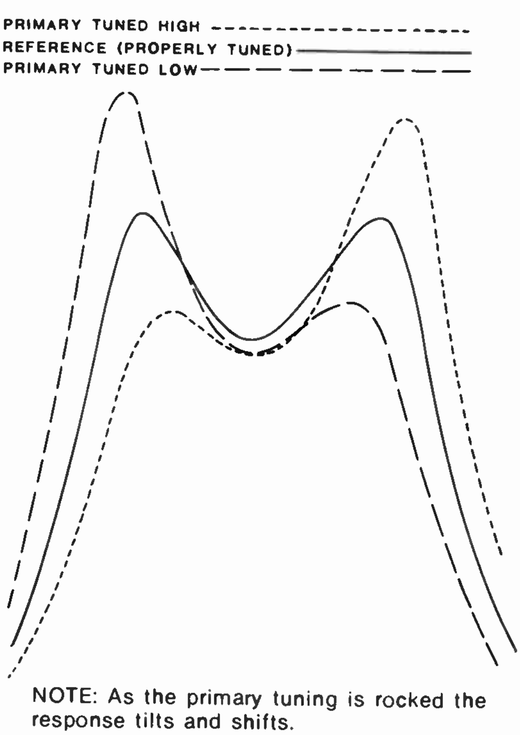


Figure 57. Primary tuning.

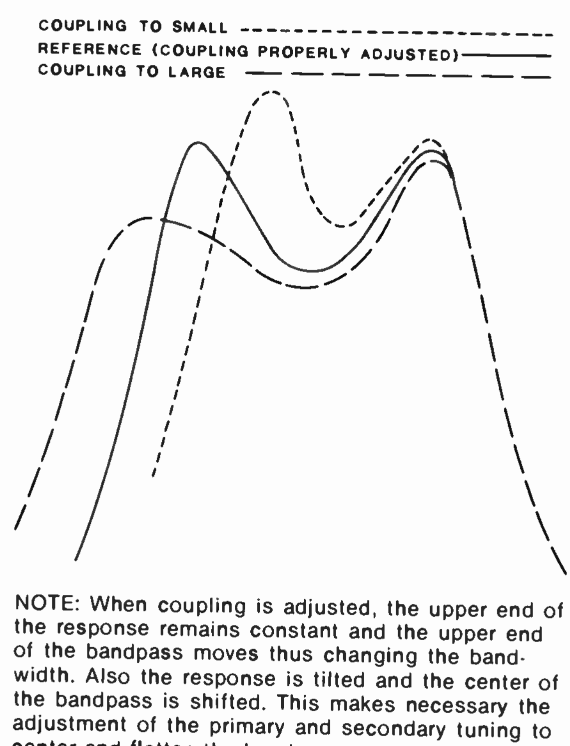
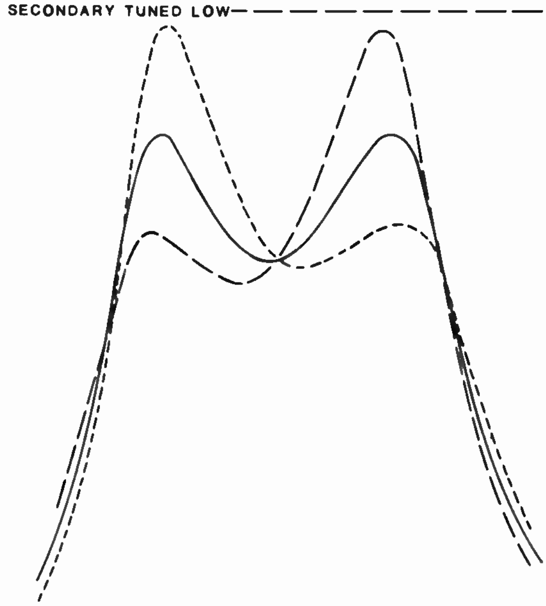


Figure 58. Coupling plate loading.



NOTE: As secondary tuning is rocked, the response tilts, but it tilts in the opposite direction to the primary. Thus, the secondary and primary tuning can be changed together to shift the response up or down in frequency.

Figure 59. Secondary tuning.

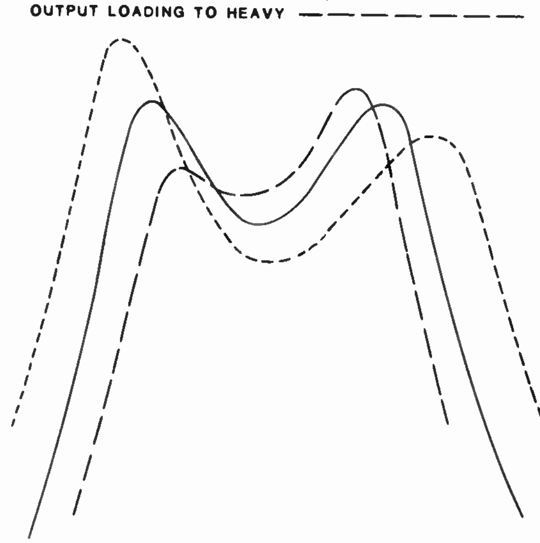


Figure 60. Loading adjustments.

VHF typically will range from 41% to 45%. At UHF the range is somewhat lower.

$$\text{Efficiency} = \frac{\text{Average power output}}{(\text{Plate voltage} \times \text{Plate current})}$$

Heavy Loading	Smaller Ripple	Lower Plate Impedance	Higher I_p and Positive I_g Lower Gain	Lower Positive I_{sg}
Light Loading	Larger Ripple	Higher Plate Impedance	Lower I_p and Positive I_g Higher Gain	Higher Positive I_{sg}

It is assumed that the RF output power is measured by an accurately calibrated RF wattmeter or by a calorimeter. If the wattmeter is not accurate, the amplifier dissipation and efficiency may be in doubt and the tube life may be shortened. If the wattmeter reads low, 100% output power will appear difficult to obtain. The amplifier might be trying to produce more than its full rated power. The symptoms would be high plate dissipation, and sync compression. These are the same symptoms that improper tuning could yield.

If the wattmeter reads high, power will appear easy to obtain and efficiency will appear high.

Plate dissipation might appear low and in reality be high. This will become apparent over a long time by a blackened anode.

A calorimeter is the most accurate way of measuring power but is not always practical. If the wattmeter is to be calibrated, send the wattmeter transmission line section, the sampling slugs, and the meter (with the meter cable) together with information giving your frequency and power to the manufacturer for calibration.

The power measured on a calorimeter or on the wattmeter is average power. To find peak power (assuming a black picture with no setup) use this formula:

$$\text{Peak power} = \frac{\text{Average power}}{0.595}$$

If a large value of plate current is required to make power, it indicates that the plate voltage is not swinging very far.

The low plate voltage swing is also indicated by the low screen current. Remember that positive screen current flows only during the time that the plate voltage swings close to the screen voltage.

The low swing of plate voltage along with the high swing of plate current indicates a low plate impedance (heavy loading).

$$\text{Plate impedance } (Z_p) = \frac{\text{The swing of plate voltage } (\delta e_p)}{\text{The swing of plate current } (\delta i_p)}$$

To make the amplifier more efficient and bring plate dissipation down, the plate impedance must be increased. To increase the plate impedance, the amplifier loading must be decreased by performing the following corrections:

Solid line —Transmitter overall response
Dotted line—The part of PA response outside of the transmitters bandpass.

Figure 61. Nonsymmetrical response caused by the amplifiers outside the transmitters bandpass. Solid line is transmitter overall response. Dotted line is the part of PA response outside of the transmitters bandpass.

1. Sweep the transmitter and decrease the PA loading.
2. This will cause the response to tilt. Correct this tilt by adjusting the secondary tuning.
3. Decreasing the loading will also cause the bandwidth (and ripple) to increase. This can be counteracted by decreasing coupling.
4. Changing coupling will cause the response to tilt. This can be corrected by adjusting the primary and/or secondary tuning.
5. The above procedure may have caused PA response to slide up or down out of the bandpass. It will show up as an asymmetrical bandpass (as shown in Fig. 61). It can be corrected by adjusting both primary and secondary tune simultaneously to center the response.

The transmitter's overall response will have the same bandpass but will now have slightly more ripple. The ripple content in the transmitter's overall response should still be within the 0.25 dB to 1.25 dB limits.

CAUTION: When changing loading, coupling, and primary and secondary tuning, it is possible to get the PA bandwidth too wide. The excessive bandwidth of the PA may be masked by narrow driver response. This excessive bandwidth will cause PA overdissipation. Correct bandwidth is shown in Fig. 62 and improper bandwidth is shown in Fig. 63.

Once again, transmit a black picture into the dummy station load. Switch the vestigial sideband filter and the linearity corrector in and check sync level at the transmitter's output. Plate current and grid current should be lower and screen grid current should be higher. The amplifier efficiency should fall between 41% to 45%.

This is a general tuning procedure given to illustrate tuning methods, control interactions and tube operating characteristics. It will work well with most double-tuned overcoupled power amplifiers. For specific tuning information on a given transmitter, the manufacturer's instructions should be consulted.

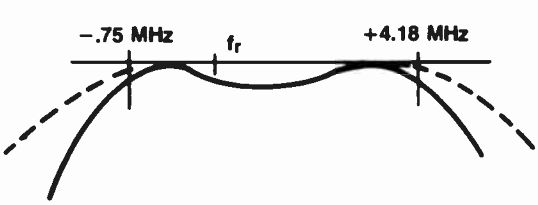
UHF TRANSMITTERS

There has been a significant amount of new technology introduced recently regarding UHF amplifiers used in TV transmitters. Most of this has been concentrated in enhancing the efficiency of the UHF amplifiers.

The introduction of the multiple depressed collector klystron and klystrode[5] to UHF TV transmitters has dramatically reduced transmitter power consumption. Those new technologies will be addressed in this section. However, a basic background is helpful to the understanding of the new devices.

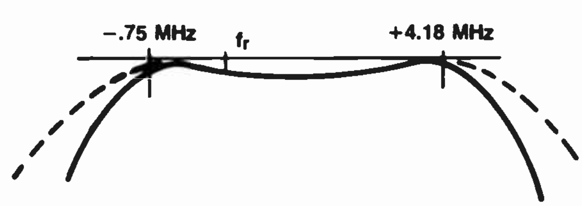
Basic Klystron Theory and Practice

The klystron uses velocity modulation to serve as an amplifying device. The electron beam emitted from the cathode is accelerated to high velocity by the



NOTE: The driver bandwidth is the dotted line, and the PA bandwidth is the solid line. The solid line also represents the overall transmitter bandpass.

Figure 62. Proper bandwidth.



NOTE: The PA bandwidth is the dotted line, and the driver bandwidth is the solid line. The solid line also represents the overall transmitter bandpass.

Figure 63. Improper PA bandwidth.

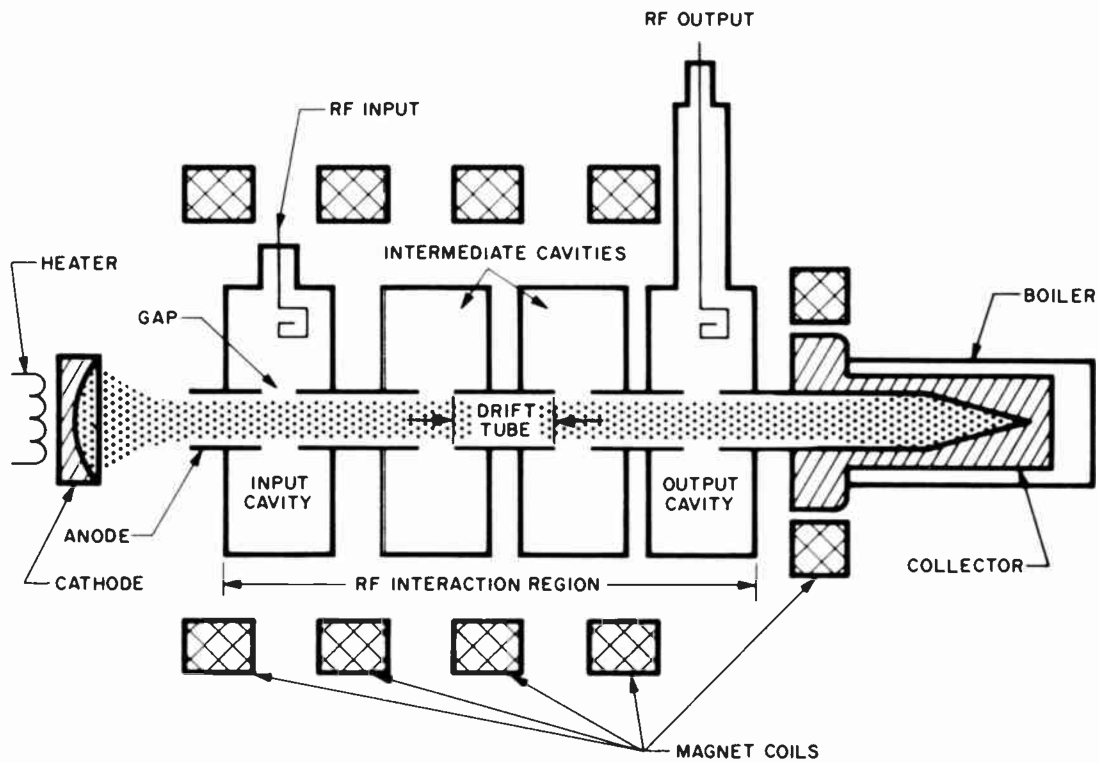


Figure 64. Principal elements of a klystron.

electric field between the cathode and anode and is directed into the RF interaction region, as shown in Fig. 64. An external magnetic field is employed to prevent the beam from spreading as it passes through the tube. At the other end of the tube, the electron beam impinges on the collector electrode, which dissipates the beam energy and returns the electron current to the beam power supply.

The RF interaction region, where the amplification occurs, contains resonant cavities and field-free drift spaces. The first resonant cavity encountered by an electron in the beam (the input cavity) is excited by the UHF signal to be amplified, and a RF voltage is developed across the gap. Since electrons approach the input-cavity gap with equal velocities and emerge with different velocities, the electron beam is said to be *velocity modulated*. As the electrons travel down the drift tube, bunching develops, and thus the density of electrons passing a given point varies cyclically with time.

The RF energy produced by this interaction with the beam is extracted from the beam and fed into a coaxial or waveguide transmission line by means of a coupling loop in the output cavity. The DC beam input power not converted to RF energy is dissipated in the collector.

The cavities can be mounted external to the klystron as shown in Fig. 65 or can be included in the vacuum envelope as shown in Fig. 66.

All cathodes have optimum ranges of operating temperature. The operating temperature of the cathode

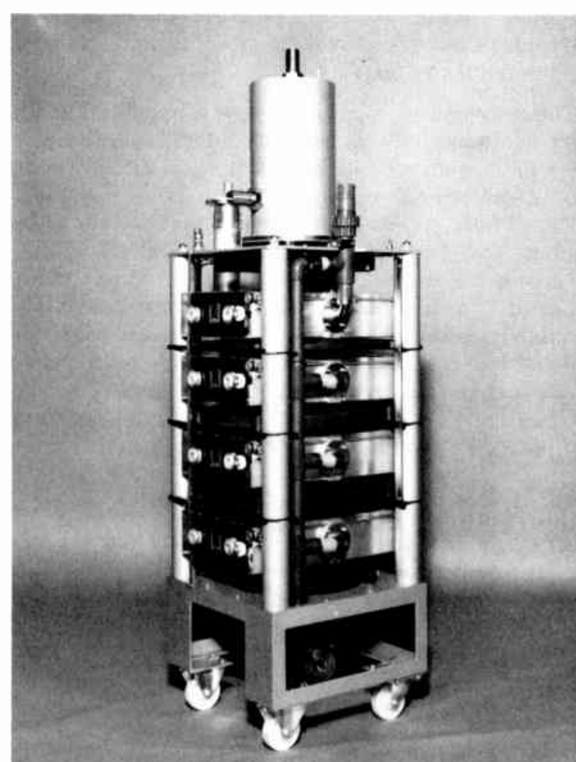


Figure 65. External cavity klystron.

Figure 66. Integral cavity klystron.

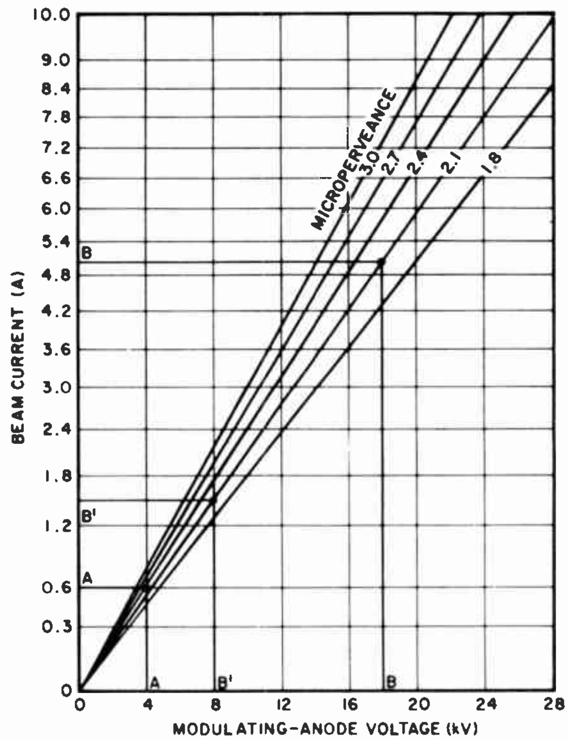


Figure 67. Beam current variation with modulating-anode voltage.

must be high enough to prevent variations in heater power from affecting the electron emission current (beam current) in the klystron. However, the temperature of the emitting surface must not be higher than necessary, since excessive temperature can reduce cathode life.

Perveance and the Modulating Anode

Perveance is a function of the geometry of the cathode-anode structure. In klystrons manufactured today, there are two electrodes which may control the beam current: The modulating anode, and a lower voltage (0 to 1,400 volts) electrode typically used for pulsing the beam current. (Examples are: Beam Control Device (BCD), Annular Control Electrode (ACE), and Annular Beam Control (ABC). If the low voltage electrode is connected to the cathode, the modulating anode voltage controls beam current which can be calculated using the following equation:

$$I_b = K \times E^{3/2}$$

where: K = Perveance constant of the klystron
I_b = Beam current in amperes
E = Beam voltage

Fig. 67 shows the relationship between beam current and voltage described in the above equation. Two examples for using the graph are given. In example A, if a modulating anode of 4,000 volts with respect to the cathode beam voltage produces a beam current of 0.6 amp, the intersection point lies on the 2.4 microperveance line. The perveance is expressed as 2.4×10^{-6} or 2.4 micropervs. Operating condition B illustrates a practical television transmitter situation in which a common beam supply of 18 kilovolts is used to power both the visual and aural klystron. At 18 kV, the visual tube operates at a beam current of approximately 5.0 amperes if the modulating anode is connected (through an isolating resistance) to the body of the tube, and the perveance is 2.1 micropervs. Since the aural output power required is much less, the DC input power is less than that required to operate the visual tube. Points B' indicate that if the modulating anode is supplied with only 8 kV (through a voltage divider) then the intersection with the 2.1 microperv line yields a beam current of only 1.5 amperes, thus accomplishing the necessary reduction of input power for aural service.

Magnetic Field

Electromagnetic coils are placed around the klystron to develop a magnetic field along the axis of the electron beam which controls the size of the electron beam and keep it aligned with the drift tubes. If the magnetic field is interrupted or insufficiently controlled the electron beam will land on surfaces other than the collector and may destroy the tube.

Cavity Tuning

The resonant frequency of each of the cavities of a klystron can be adjusted in two ways to the operating frequency of the transmitter. The inductance can be changed by changing the volume of the cavity in external cavity klystrons, or the capacitance of the drift-tube gaps can be changed in integral cavity klystrons.

Cavity/Transmission Line Coupling

Fig. 68 illustrates magnetic-loop coupling, where the RF energy is fed through a coaxial line with its center conductor inserted into the klystron cavity. The end of the center conductor is formed into a loop. This forms a simple one-turn transformer which couples RF energy into or out of the cavity through a coaxial transmission line. Intermediate cavities may have their loops coupled into RF loads to vary the "Q" of the cavities to change the overall bandpass characteristics of the klystron.

Effect of RF Drive Power on RF Output Power

Fig. 69 shows RF output power as a function of RF drive power applied to the tube. From this curve, we see that when the RF drive power level is low, the RF output power is low. As the level of RF drive power increases, RF output power increases until an optimum point is reached. Beyond this point, further increases in RF drive power result in less RF output power. Because of these effects, two zones and one point have been labeled on the curve. In the zone labeled "Underdriven", RF output power increases when the RF input power is increased. The point labeled "Optimum" represents the maximum RF output power obtainable. Klystrons are said to be saturated at this point, since any further increase in RF drive only decreases the RF output power. The zone formed at the right side of saturation is labeled "Overdriven". To obtain maximum RF output power from a klystron, sufficient RF drive power must be applied to the tube to reach the point of saturation on the curve. Operating at RF drive levels beyond the saturation point will only overdrive the klystron, decrease RF output power, and increase the amount of beam interception at the drift tubes (body current). Klystrons tuned for TV service are operated within the underdriven zone of Fig. 69.

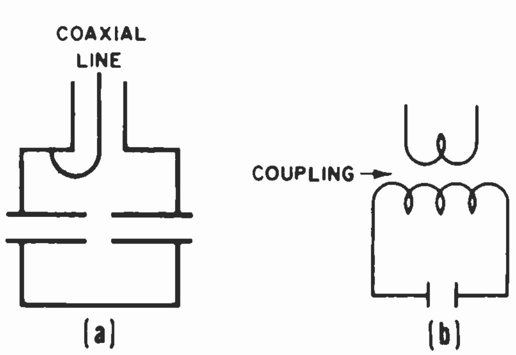


Figure 68. Loop coupling and equivalent circuit.

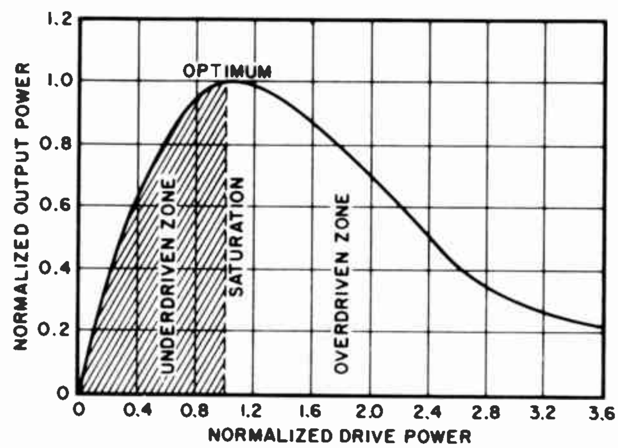


Figure 69. RF output power as a function of RF drive power.

Fig. 70 shows how RF output power changes with various levels of RF drive power applied to a klystron under different tuning conditions. Point A represents the drive saturation point for a synchronously-tuned tube. Point B shows a new point of saturation that is reached by tuning the penultimate (closest to the output cavity) cavity to a somewhat higher frequency. By tuning the penultimate (next to last) cavity still further, Point C is reached. There is a point, Point D, where increasing the penultimate cavity frequency no longer increases RF output power. Instead, it reduces the output power as shown at point E.

Klystron Efficiency

Klystron "efficiency" has often been measured by dividing Peak RF output power by the DC power input. Since klystrons have been compared using "peak RF output power" the term efficiency is not valid because it is possible to have greater than 100% efficiency with some amplifiers.

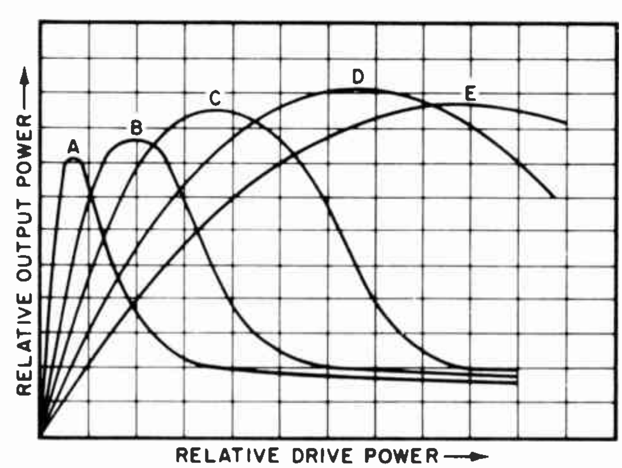


Figure 70. Output power variation with drive power under different tuning conditions.

The term "Figure of Merit" is a more valid terminology and is calculated as follows:

$$\text{Figure of Merit} = \frac{\text{RF power output}}{\text{DC input power}}$$

where DC input power is measured using a 50% APL signal and RF power output is the peak of sync value for visual.

This is a measure of the klystron stage only. Total transmitter or plant efficiency would include the power consumed by the magnets, heat exchanger blower, pumps, driver, and control circuits.

Tuning

There are a number of methods of tuning klystrons. They are:

- High gain tuning for monophonic aural service
- Broadband for MTS aural service
- Visual service with fixed beam current
- Visual service for pulsed beam current
- Tuning for integral cavity klystrons employing the variable visual coupler

It is best to consult the transmitter manufacturer for specific information regarding tuning and mode of operation desired. However, some basic information can be presented here:

AMPLIFIER DEVICE	FIGURE OF MERIT
TETRODE	.9 – 1.0
INTEGRAL CAVITY KLYSTRON	.65 – .75
EXTERNAL CAVITY KLYSTRON	.65 – .75
KLYSTRODE OR IOT	1.1 – 1.3
DEPRESSED COLECTOR KLYSTRON	1.2 – 1.3

General Klystron Tuning Considerations

The output cavity is generally tuned to the carrier frequency. It is essential to operate with the coupling loop adjusted so the output cavity slightly over coupled. Fig. 71 shows the relationship of output power to proper coupling loop adjustment. If the coupling loop is adjusted so that the cavity is under-coupled, arcing and ceramic fracture resulting in klystron failure may occur.

The input coupling is adjusted for the best tradeoff of minimum reflected power and best overall bandpass.

Intermediate cavities may be externally loaded to lower Q and increase bandwidth.

The penultimate or next to last cavity is tuned above the passband. The integral cavity klystron penultimate cavity tuning location is generally 10 MHz to 15 MHz above the passband, while the tuning location for the four cavity external cavity klystron is 6 MHz to 8 MHz above the passband.

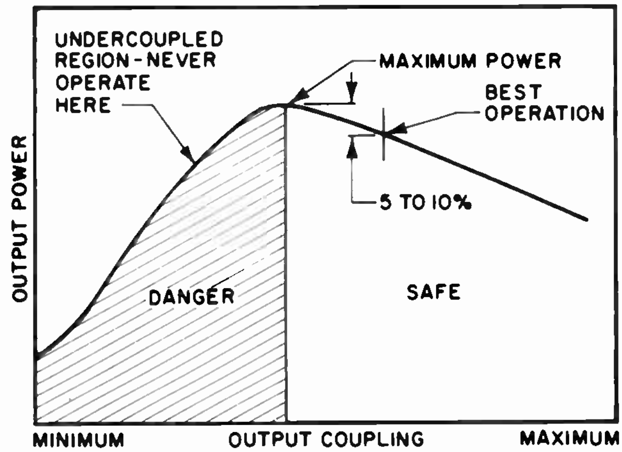


Figure 71. Adjustment of output coupling control.

Variable Coupler for Integral Cavity Tubes

The purpose of the variable load coupler is to improve the klystron Figure of Merit. This is accomplished by raising the impedance presented to the klystron. The variable load coupler as shown in Fig. 72 has a shorted transmission line stub which is tuned beyond a quarter wavelength to present a capacitive susceptance load to the output transmission line. This capacitance susceptance (or inductive reactance) in parallel with the load impedance is transformed back through an approximate ⅜-wavelength to present a substantially resistive load to the output coupling loop of the final klystron cavity.

Transmission line formulas aid in the analysis of how the variable coupler functions.

For example, assume that the transmitter and variable load coupler are terminated with a 50 ohm impedance. If the tuned short is exactly 0.25 wavelength long, the equivalent load impedance at the junction of the tuned short can be represented by:

This impedance, transfored back to the klystron ⅜-wavelength away is given by the equation:

$$Z_{in} = Z_0 \frac{[Z_L + jZ_0 \tan\beta l]}{[Z_0 + jZ_L \tan\beta l]}$$

In this case Z_L = 50 ohms and Z_0 = 50 ohms.

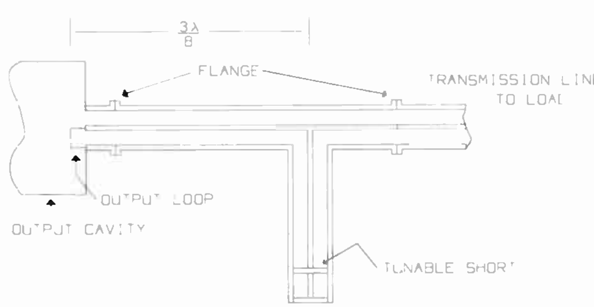


Figure 72. Variable visual coupler.

However, when the tuned short is longer than 0.25 wavelength, an inductive reactance is added in parallel with the 50 ohm load impedance.

Assume for example that the tuned short was lengthened such that the resultant load impedance was now 50 − j10 ohms.

Using the same equation, where $\beta = 2\pi/$(one electrical wavelength) and l = the electrical wavelength between the equivalent load impedance and the klystron.

In this case, l = ⅜-wavelength, $Z_0 = 50$, and $Z_L = 50 + j10$.

Solving this equation yields a new impedance at the klystron of about 61 ohms. Thus the impedance has been raised via the tuning stub. To raise the impedance further, the tuning short must be lengthened again.

The iterative procedure of small adjustments which lengthen the tuning stub allows a suitable compromise between efficient operation and klystron safe operation.

Beam Current Pulsing

For a number of years klystrons were operated at maximum beam current. The mod anode was tied to ground through a resistor. Early model klystrons would draw approximately 7.5 amps at 24 kV for 55 kW peak-of-sync visual operation. The development of more efficient klystrons allowed the reduction of beam currents to near 6.3 amps for 55 kW. With saturation set at 115% of needed power, the klystron was operated in the more linear part of the curve, but excess beam current was being consumed.

Operating the klystron at saturation will improve efficiency but requires more linearity and phase compensation. By using a pulser to switch to a higher beam current during sync and back to a lesser current during video the average beam current is significantly reduced.

Fig. 73 shows the horizontal line timing. Observe that sync is 8% of the duty cycle and video is the remaining 92% of the transmitter signal.

The practical limit of reduction of the video beam current is the point at which tip of burst and back porch signal distortion is not correctable.

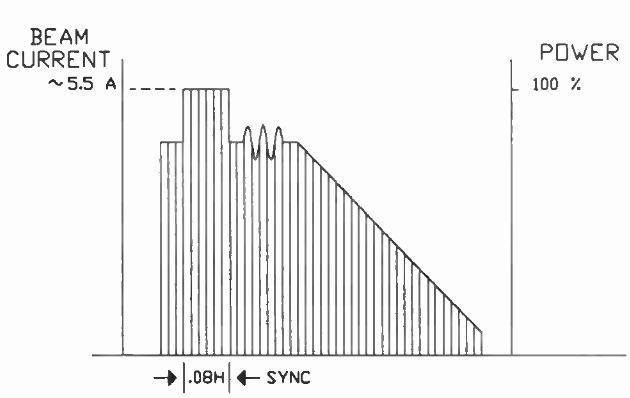


Figure 73. RF envelope vs. beam current.

With beam current pulsing, the effective "Q" of the electron beam and cavity combination is altered. This results as a passband tilt from visual carrier down to the upper edge of the passband. This requires readjustment of the cavities to obtain a flat response. Tilt of as much as 7 dB has been observed. Tuning for high efficiency with beam current pulsing can reduce the gain of a klystron.

Consider the following example of a typical 60 kW external cavity klystron.

In static (nonpulsed) operation, the tube is supplied with enough beam current to saturate at 100% power and a small amount of headroom added for changes in beam voltage. In pulsed operation only enough beam current is supplied to saturate at 100% power during sync. This would normally correspond to about 5.5 amps for a wideband external cavity klystron. The amount of beam current reduction during the video picture is dependent upon operating channel, klystron perveance, available drive power, and precorrection capability.

The optimum value of beam current must be experimentally determined. However, figures of merit of 77% have been achieved. It is important to note that the absolute value of the pulsed efficiency is directly proportional to, and therefore limited by, the efficiency obtained for nonpulsed operation. In order to prevent severe burst and back porch distortion, a small amount of amplifier headroom is needed, approximately 2% to 5%. Typical values of beam current during the video portion of the waveform are 3.4 amps to 3.9 amps.

Pulsed operation for klystrons is accomplished by connecting the beam control electrode to a voltage source of 0 to −1,400 volts with respect to the cathode voltage. During sync, the pulser operates at zero volts. During the video portion of the signal, values of −400 volts to −800 volts are used to achieve reliable high efficiency operation.

In some transmitters, sync is actually reduced or removed from the input video. As the beam current is pulsed, the klystron gain change increases the RF power to produce the proper sync level.

For pulsed operation, DC input power is calculated as follows:

$$\text{DC input power} = \text{Beam voltage} \times [(\text{Beam } I_{sync} \times \text{Sync duty cycle}) + (\text{Beam } I_{video} \times \text{Video duty cycle})]$$

In this example,

$$\text{DC input power} = 24\,[(5.5 \times .08) + (3.7 \times .92)] = 24 \times 3.84 = 92.2\ \text{kW}$$

$$\text{Figure of merit} = 60/92.2 = 0.65$$

Effect of Pulsing on Transmitter Precorrection

With increasing RF drive level, klystron amplifiers exhibit a RF phase change in addition to amplitude compression. This phase change is called incidental phase modulation (ICPM). Very little can be done to reduce ICPM by klystron tuning or selection of magnet current. Also, incidental phase distortion increases rapidly near saturation of the klystron.

The RF phase shift through the klystron will also change as beam current is varied. When the klystron is pulsed during sync the change in beam current causes the phase of the signal to change to a new value. When the klystron switches back to the video current level the previous value of phase returns. To combat this problem, ICPM correctors have been developed. These correctors generally operate at the exciter intermediate frequency (IF) and introduce a correction signal equal and opposite to the distortion produced by the klystron. A phase modulation stage in the exciter may be keyed by sync and adjusted to pre-correct for incidental phase distortions caused in the klystron during pulsing. This is most commonly done at the IF level because of the ease of implementation.

When amplifiers are operating very close to saturation during the color burst, more differential gain and differential phase correction may be needed. However, modern exciters can fully precorrect these conditions.

For maximum efficiency, operate the tube at saturation at sync tip. In a pulsed transmitter, the color burst and black picture content are near saturation as well.

The vector diagram of Fig. 74 illustrates first order klystron nonlinearities and the operation of the ICPM and linearity correctors. The desired TV output signal is represented by an amplified version of the desired instantaneous phasor. However, the transmitter output signal is phase shifted by a phase error and compressed in amplitude. To compensate for this, the signal in the exciter is precorrected by an amplitude expansion, and a correction in quadrature. When the resultant signal is amplified, the output signal will be a replica of the desired TV signal.

Sync pulse oscillations may appear as ringing on the sync pulse when the klystron is operated at saturation. This may exhibit itself as a tearing of the picture or sync. This ringing is believed to be caused by secondary electron feedback enhanced by the reverse gain of the klystron cavities. Rebiasing the tube slightly out of saturation will eliminate the ringing.

Multi-stage Depressed Collector (MSDC) Klystrons

Klystron amplifiers operate by converting energy from a beam of electrons to RF output power. At full output power, about half of the beam power is converted to RF output power. Correspondingly, half the DC input power remains on the beam as it exits the cavity region. In a standard klystron this "spent electron beam" energy is dissipated as heat. The MSDC klystron, however, operates on the spent electron beam, recovering its energy to reduce the dissipated heat.

After the electron beam has completed its job by providing the desired output power, significant energy still remains in the electron beam. The interaction process has produced a wide range of velocities for individual electrons, from nearly stopped to up to twice the initial energy.

In the early 1970s, researchers at NASA investigated collector designs. In 1984, NASA provided the depressed collector technology for development of UHF-TV klystrons. Power recovery in the collector region is accomplished by decelerating the electrons in the spent beam. This can be done by providing an electric field in the collector such that the electrons are slowed before they strike the collector wall. A collector composed of multiple elements is utilized, with each element operated at a negative potential with respect to ground potential.

Consequently, the collector potentials are referred to as being *depressed* below ground. The collector element geometry was carefully selected to provide an electric field shape which would sort the impinging electrons according to their velocity, reducing their energy as much as possible, yet ensuring that all electrons would strike one of the elements and not

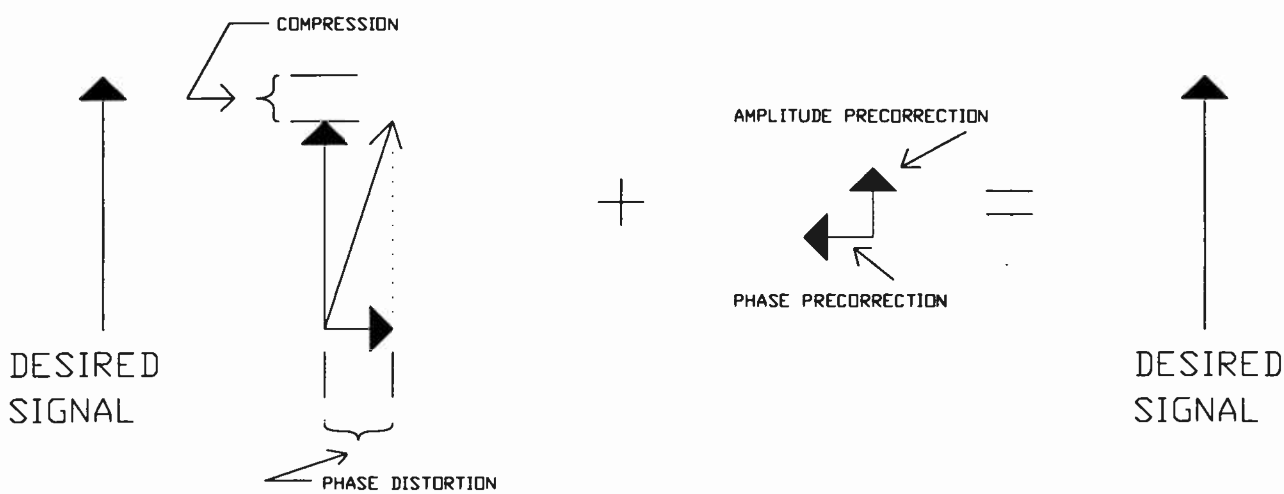


Figure 74. Vector representation of precorrection.

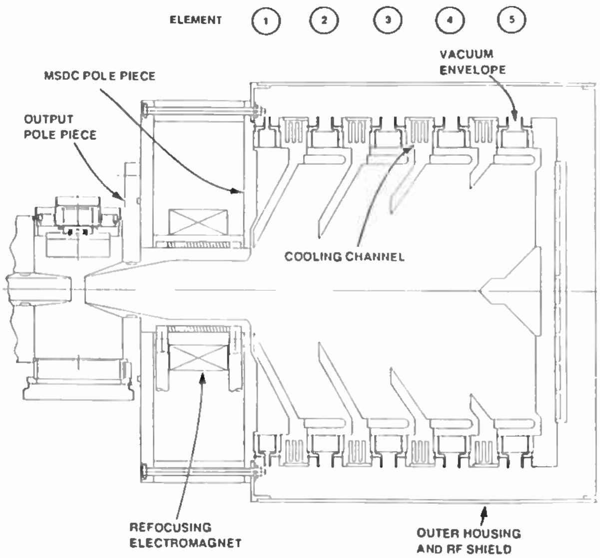


Figure 75. UHF-TV MSDC design.

be reflected back into the electron beam. Modern computer aided techniques facilitated this design effort.

Fig. 75 shows the resulting collector configuration. The collector is composed of five elements and is designed to operate with equal voltages between elements. For the 60 kW klystron, the voltage per collector stage is typically 6,250 volts. This means that element 1 is at ground potential, element 2 is at −6.25 kV, element 3 is at −12.5 kV, element 4 is at −18.75 kV, and element 5 is at full cathode potential of −25 kV.

To make sure that the electron beam is optimized before entering the collector region, a refocus coil is provided just ahead of the collector.

The significant benefit of the depressed collector klystron is the reduction in power consumption for a given power output. The individual collector beam currents will vary depending on the picture level being transmitted. Since the recovering of the spent beam takes place in the collector, it has essentially no impact on the tuning and normal precorrection of the klystron. Therefore, a multi-stage depressed collector klystron tunes the same as a standard klystron and requires the same precorrection as a standard klystron. The depressed collector technique can be applied to either external or integral cavity klystrons. The cooling of the cavities is not affected by the addition of the collector.

To date, these klystrons require the use of a vacuum ion pump. Since the power dissipated in the collector is reduced, only six gallons per minute of high purity water coolant is needed. Although the efficiency performance of the MSDC is dependent upon the transmitter configuration, figures of merit of 1.2 to 1.6 have been obtained in transmitter installations in the field.

Transmitter Design Using Multiple Depressed Collector Klystrons

The primary differences in this transmitter from a standard klystron transmitter are in the beam supply, the cooling system. Fig. 77 shows the power connections made to a MSDC klystron. Since the RF performance of the MSDC klystron is the same as the standard klystron, there are no differences in the RF driver chain.

The cooling system of a MSDC klystron transmitter uses a two-stage heat transfer system as shown in the cooling system block diagram of Fig. 78. The reason a two-stage system is chosen is to allow outside heat exchangers to be used. The cooling system consists of a high purity water loop and a glycol-water mixture

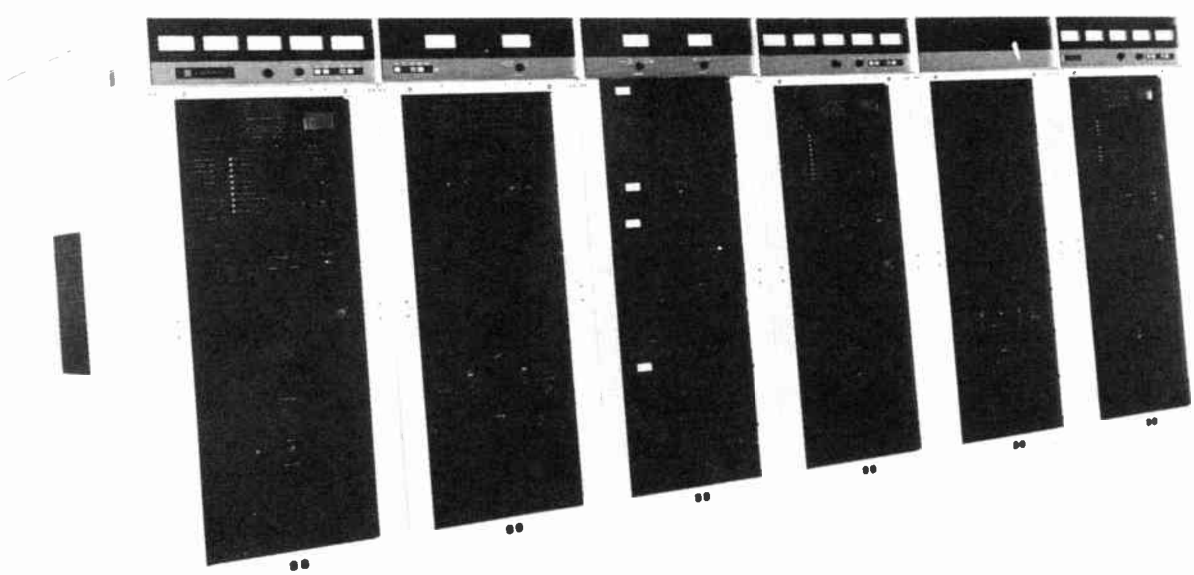


Figure 76. MSDC transmitter.

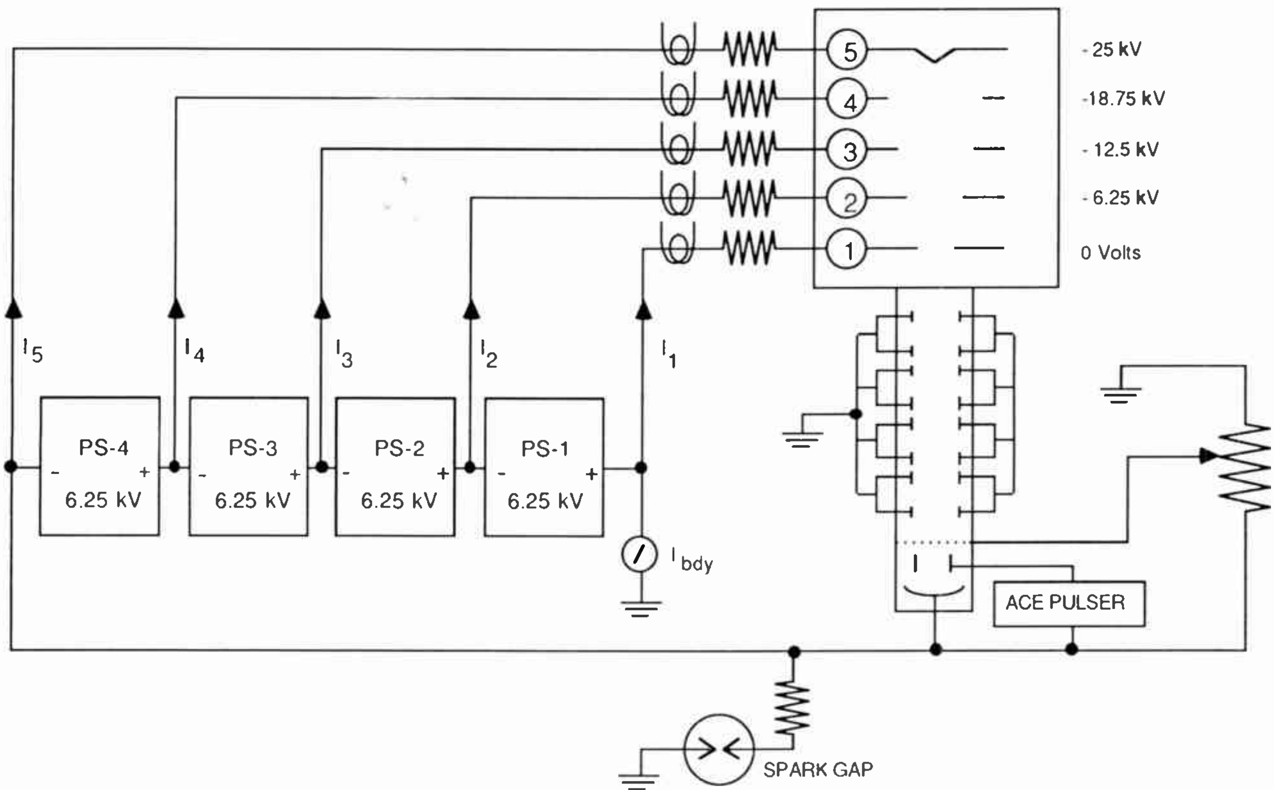


Figure 77. MSDC high voltage connections.

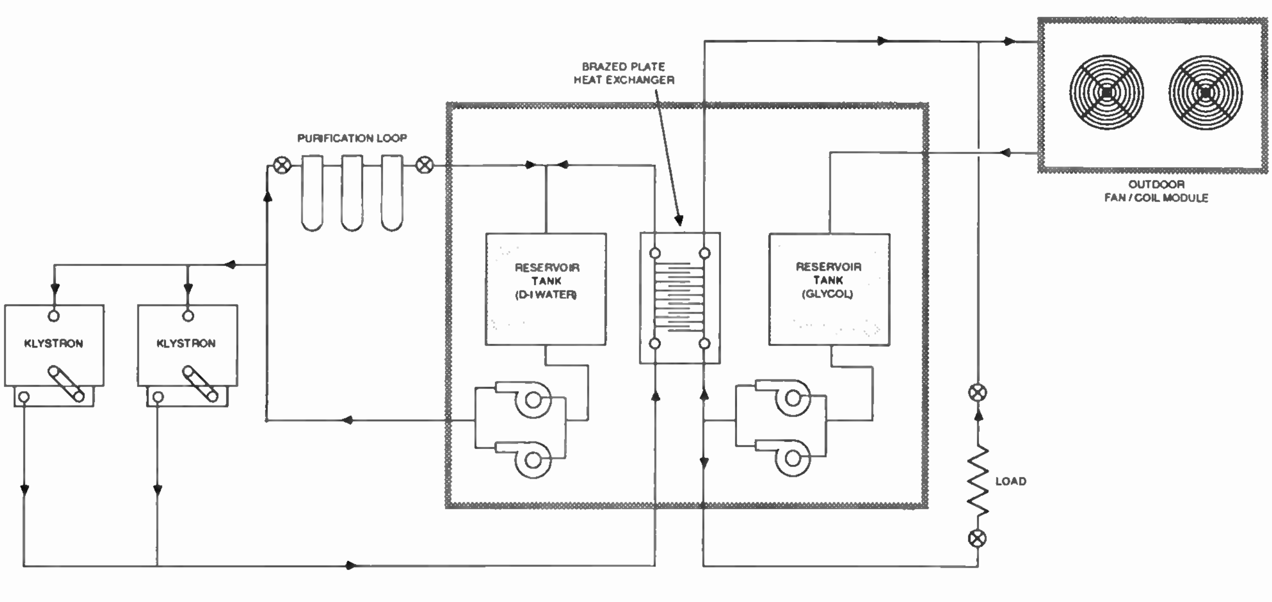


Figure 78. TV-60UM heat exchanger system.

loop. High purity water (resistivity of 200,000 ohm-cm or more) must be used with the cooling system to prevent current flow between collectors. To remove ions, free oxygen and other possible contaminants from the water, a three-stage filter system (purification loop) is used. The filter cartridges sample part of the water flow so the whole system is continuously cleaned. The filter cartridges can be replaced without taking the transmitter off the air.

Separate beam supplies for each klystron are frequently used in TV transmitters.

Since the currents from the beam supplies will change dependent on the power level, there will be video frequency currents present on the power supply leads. Therefore, sufficient bypassing and power supply high voltage wire shielding is required.

Monitoring each section of the beam supply current is required to obtain the currents for power dissipation calculations.

Protection circuitry needs to include magnet overcurrent and undercurrent trip points, beam supply overcurrent and overvoltage sensors (for each collector), water flow sensors, arc detectors within the 3rd and 4th cavities, and sufficient interlocks to prevent personnel from accidentally coming in contact with high voltage.

KLYSTRODE®

The Klystrode® basically combines features of a tetrode and a klystron. Hence, the name Klystrode® was created.

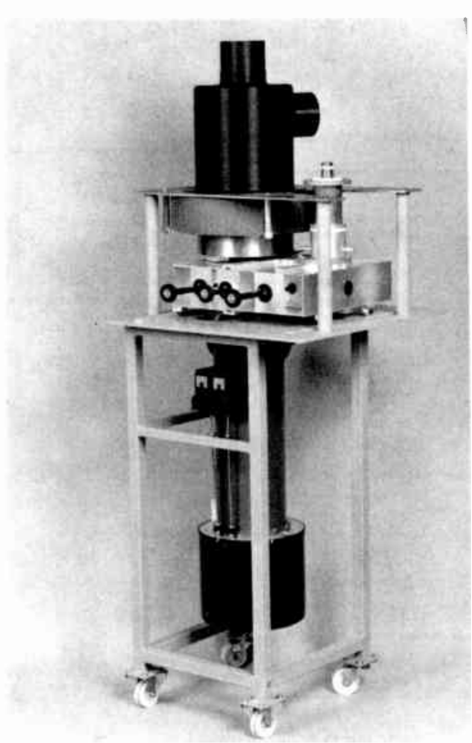


Figure 79A. Klystrode.

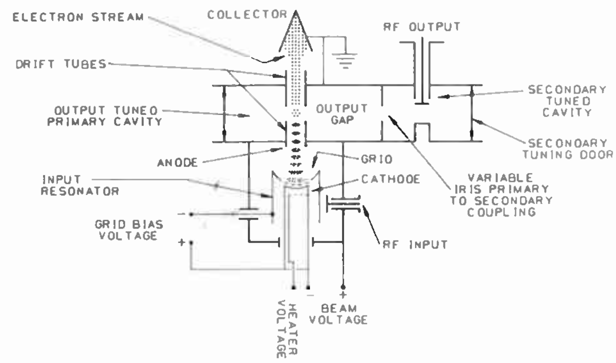


Figure 79B. Klystrode schematic.

A Klystrode® is shown in Fig. 79A. It is composed of an electron gun very similar to a klystron, a control grid, an input cavity, accelerating anode, drift tube, output cavity, and collector. It is physically smaller than a klystron and weighs approximately 90 pounds including its circuit assembly. The tube itself weighs about 45 pounds.

Refer to Fig. 79B. The electron beam is formed at the cathode, density modulated with the input RF signal applied to a grid, and then accelerated through the anode aperture. In its bunched form, the beam drifts through a field-free region and then interacts with the RF field at the drift tube gap in the output cavity. Power is extracted from the beam in the output cavity in the same manner as a standard klystron.

RF input power is applied to the control grid via a resonant cavity. The grid is usually biased negatively. The first part of the tube may be thought of as a triode with a perforated anode through which the electron beam is guided by electric and magnetic fields. The beam is bunched, or chopped, at the radio frequency and is accelerated by the high anode potential. It passes through the anode extension cylinder, which is an electrostatic shield, and then interacts with the RF field in the output gap. The spent beam is dissipated in the collector, separate from the output RF interaction circuit.

The tuned input and output circuits are external. The RF input circuit is a 4 stub tuner which matches the drive source to the high impedance grid. Also included is a circuit which provides a DC block for high voltage. The output circuit cavity is clamped to the body of the tube with techniques similar to those used for external cavity klystrons. The output circuit is double-tuned to achieve the bandwidth required for visual service. The double-tuned output circuit consists of a primary cavity clamped to the body and an iris-coupled secondary cavity. The output transmission line is probe coupled to the electric field in the secondary cavity. Variable controls are used to adjust primary and secondary cavity resonant frequency and the iris coupling.

The grid structure intercepts some electrons causing grid current to flow. On 60 kW models, some material from the cathode migrates to the grid and must be periodically boiled off. A design goal of future Klystrodes® is to eliminate this periodic procedure.

Because the tube only has two cavities it is much shorter than a klystron. The magnetic field requirements of the tube are about 7 volts at 30 amps.

The high voltage circuitry is contained within a shielded compartment on top of the input circuit. This circuitry consists of various filters to contain the UHF fields and prevent instabilities at video frequencies. High voltage, grid bias, and filament power enter this section via high voltage cables.

The fundamental benefit of the Klystrode® is that it may operated as a Class B amplifier. Thus the beam current is proportional to the RF drive signal. Although the efficiency performance of the Klystrode® is dependent upon the transmitter configuration, figures of merit of 1.2 to 1.4 have been obtained for 60 kW visual service from transmitter installations in the field.

In aural service, the Klystrode® is tuned the same as for visual service.

Power gain in either type of service is about 23 dB. Thus drive power is about 300 watts for the visual and 30 watts for the aural (assuming 10% nominal aural power).

The transfer characteristics of the Klystrode® is also a combination of a klystron and a tetrode. A typical transfer curve is shown in Fig. 80. The nonlinearity at white picture power levels resembles that of a tetrode and the nonlinearity approaching peak output power also resembles a tetrode. The amount of nonlinearity is similar to a tetrode and the shape of the precorrection curve is "S" shaped.

Klystrodes® at 15 kW, 40 kW, and 60 kW visual peak sync power levels have been constructed and are in service.

The Klystrode® may also be used as a multiplexed amplifier. This means the diplexer may be eliminated. For example, a tube may be used as a 30 kW visual with a 10% aural signal. Intermodulation products at ±920 kHz from visual carrier generated by the tube can be precorrected by low level IF circuitry.

Cooling of the Klystrode® at the 60 kW power level requires about 25 gallons per minute of water for the collector. The body of the tube is also water cooled.

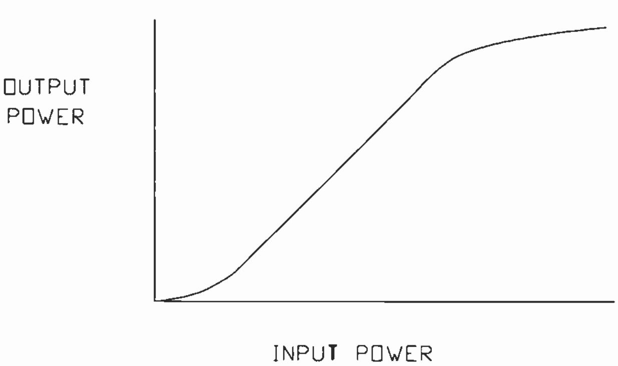


Figure 80. Klystrode transfer curve.

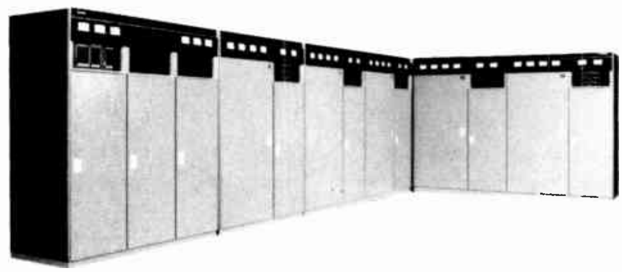


Figure 81. Klystrode transmitter.

A 50/50 water/glycol solution is typically used without any special water purification. The RF input and output cavities are air cooled.

The 15 kW and 40 kW Klystrodes® may be air-cooled. The volume of air and air pressure needed are comparable to a tetrode of similar power.

Two versions of Klystrodes® are needed to cover the entire UHF operating band.

To date, these devices require the use of a vacuum ion pump.

Transmitter Design Using Klystrode®.

Support circuitry for the Klystrode® consists of providing the necessary drive power, precorrection, different types of power supplies necessary, and the protection circuitry. A Klystrode® transmitter is shown in Fig. 81. A Klystrode® transmitter block diagram is shown in Fig. 82.

The Klystrode® uses a beam voltage of 32 kV. Since the beam current changes with modulation, there will be video frequency currents required from the beam supply. The beam supply must be designed to provide excellent regulation from no load to full load and to also provide a low source impedance for all video frequencies. In the event of a high voltage failure, the conventional beam supply should limit the energy dumped into an arc. The Klystrode® supply, being stiffer, requires a triggered "crowbar circuit" to limit the beam supply arc energy. A block diagram of a crowbar circuit is shown in Fig. 83.

The grid bias supply of −10 to −70 volts, with respect to the cathode, typically should "float" with the beam voltage. A simple method of developing the grid voltage is to use zener diodes connected between the power supply and the tube cathode connection and tap the grid to the appropriate zener to obtain the desired bias current. As with klystrons, the Klystrode® magnet power supply must have sufficient energy storage that the beam remains focused until the beam decays. Also, a power supply for the ion vacuum pump is needed with appropriate sensors for the protection circuitry.

Other protection circuitry provided in the Klystrode® transmitter is similar to that needed in the klystron transmitter.

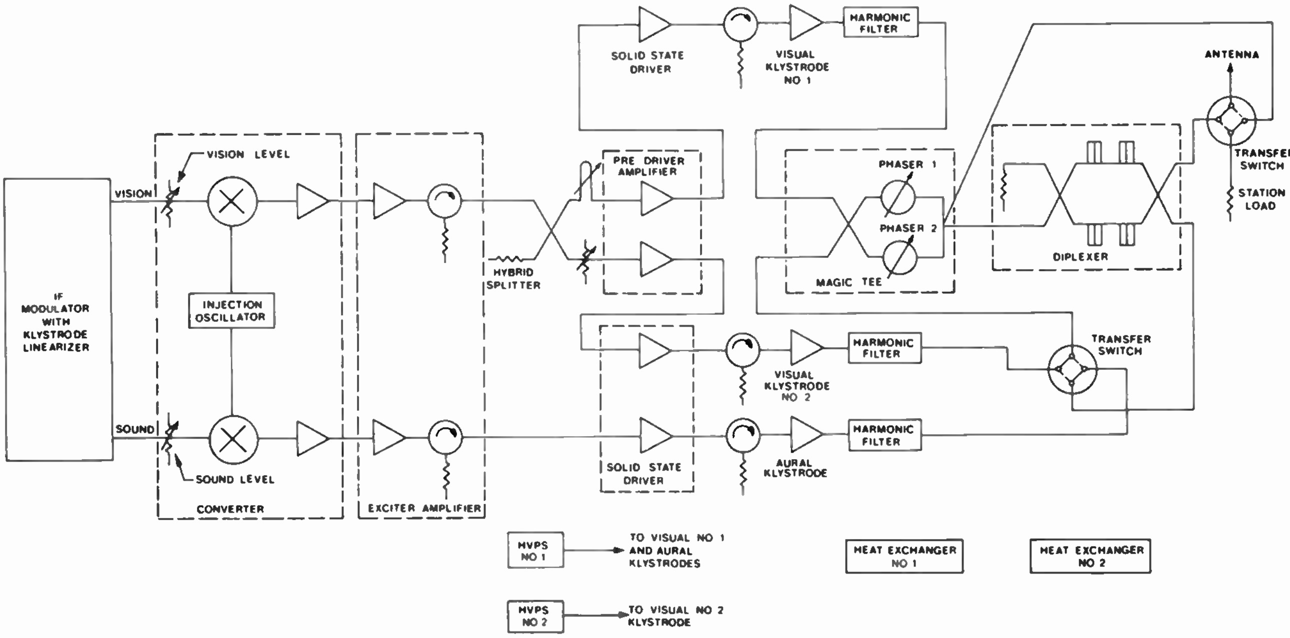


Figure 82. Klystrode transmitter block diagram.

TRANSMITTER COMBINING CIRCUITS AND RF SYSTEMS

A typical RF system for a transmitter will consist of hybrid combiners, (if more than one amplifier cabinet is used), harmonic filters, −3.58 MHz notch filter, and diplexer to combine visual and aural.

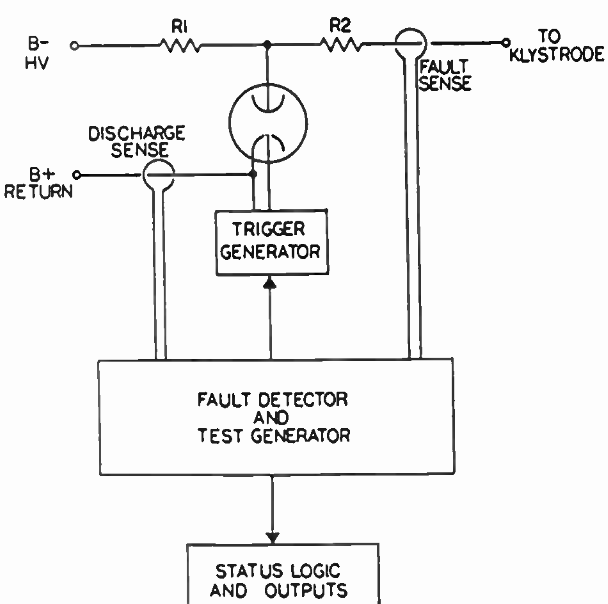


Figure 83. Triggered crowbar circuit.

At VHF frequencies the −3.58 MHz notch filter, harmonic filters, and diplexers are all quite large and are typically supplied as individual units. In some cases, the −3.58 MHz filter is incorporated with the notch diplexer. However, at UHF frequencies, components are smaller and waveguide technology is typically used so the −3.58 MHz notch filter, the notch diplexer aural cavities, and associated hybrids are often integrated into one assembly. Multiple aural notch cavities may be used at UHF frequencies to optimize multichannel TV sound performance and for power sharing when very large powers are used.

This section will review hybrid combiners, notch diplexers, switchless combiners, and Magic Tee RF systems.

Two different types of combining circuits must be considered:

1. Combining of sources with the same frequency (i.e., multiple power amplifiers).
2. Combining of sources with different frequencies (i.e., visual and aural).

For combining of sources of the same frequency, 90° 3 dB hybrid couplers have found almost universal acceptance. Fig. 84 shows the physical model of a 90° 3 dB hybrid coupler.

A 3 dB hybrid coupler consists of two identical parallel transmission lines mounted in a common outer conductor and coupled over a length equal or approximately equal to the quarter wavelength of the operating frequency.

The construction is symmetrical, i.e., both inner conductors have the same physical dimensions with

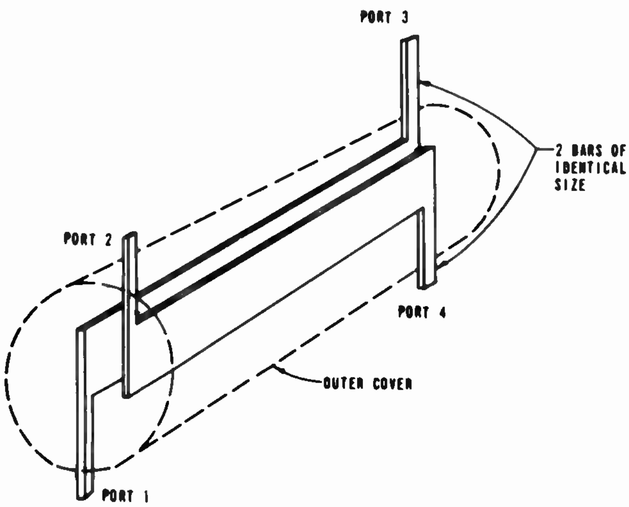


Figure 84. 90° 3 dB hybrid coupler.

respect to the outer conductor. The 3 dB hybrid coupler can be used as a power splitter or as a power combiner. Fig. 85 shows the 3 dB hybrid coupler as a power splitter, and Fig. 86 shows the 3 dB hybrid coupler as a power combiner.

One of the common uses of the 3 dB hybrid is to combine two frequencies that are relatively close (visual and aural) in a back-to-back hybrid configuration called a notch diplexer. The first step in building a notch diplexer can be shown by combining two 3 dB hybrids back to back as shown in Fig. 87A.

The input hybrid will act as a splitter while the hybrid at the output will act as a combiner. The visual signal is inserted into the left hybrid and will appear at the antenna port as shown in Fig. 87B.

If we now add the aural signal as in Fig. 88, we can see that it is passed through the 3 dB hybrids but does not appear at the antenna port as required.

If we now add two very high Q notch cavities tuned to the aural frequency, as depicted in Fig. 89, we can introduce a short circuit to the aural frequency. These shorts circuits cause the aural signal to be reflected back into the 3 dB hybrid where the same 3 dB hybrid now effectively acts as a power combiner and passes

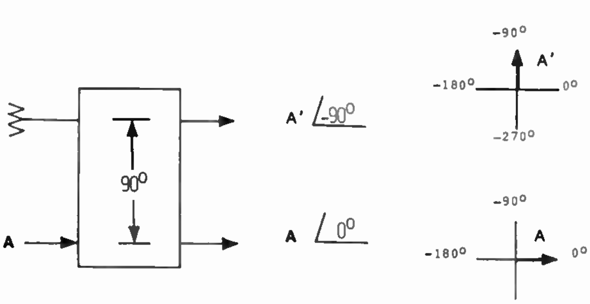


Figure 85. 3 dB hybrid as a power splitter.

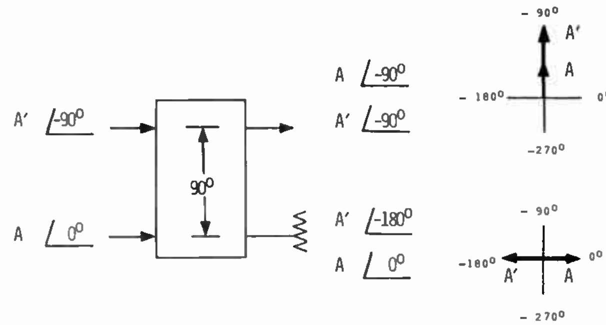


Figure 86. 3 dB hybrid as a power combiner.

the combined aural signals to the antenna port along with the visual signal.

The aural notch cavities are theoretically not seen by the visual signal since they are high Q cavities tuned to the aural frequency. In reality, the aural cavities do cause some amplitude roll-off and group delay of the high end of the visual passband.

Another use of the back-to-back hybrids can be seen in Fig. 90 which shows a single signal source driving a power splitter (Hy1) and dual amplifiers (A1 and A2). Real world amplifiers do not have exactly the same gain and phase shift through the amplifier especially if the amplifier has tuning controls. This hybrid arrangement can be used for visual or aural signals. Gain (AT1 and AT2) and phase adjustments are used to keep the amplifiers' output signals (A1 and A2) at equal output power and 90° phase difference so that they can be properly combined by the power combiner (HY2) with a minimum of reject load power.

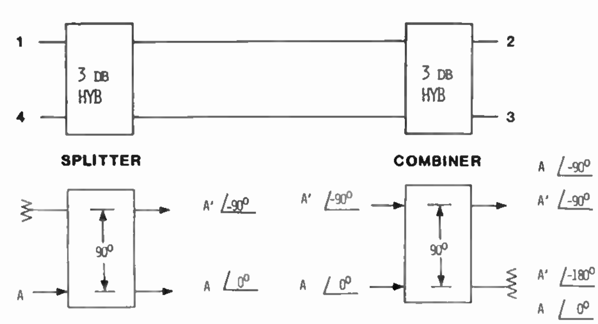


Figure 87A. Back-to-back hybrids.

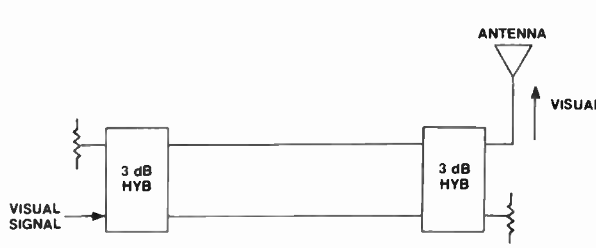


Figure 87B. Visual signal applied to back-to-back hybrids.

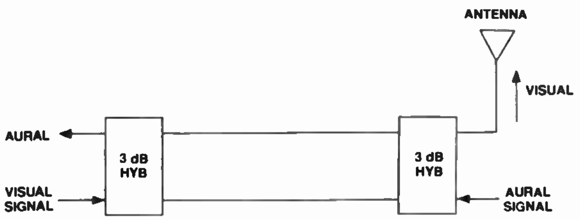


Figure 88. Aural signal applied to back-to-back hybrids.

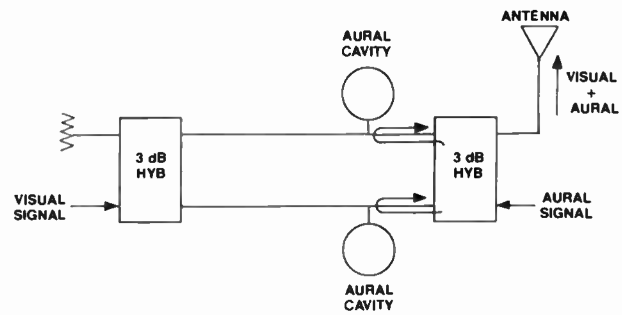


Figure 89. Aural and visual signals applied to back-to-back hybrids with aural notch cavities to reflect aural signals back to antenna port.

A typical use of this system would be the driving of two power amplifier chains in parallel with a single exciter. HY1 would be a low power splitter while HY2 would be a high power combiner. A1 and A2 would be the complete amplification chain consisting of multiple stages of solid state and/or tube amplifiers.

Differences in amplifier chain gain are compensated for by attenuators AT1 and AT2 which vary the amount of input signal applied to the amplifier chains. Differences in phase are compensated for by the length of cable between the attenuators and the amplifier chain. Optimum phase and gain adjustments are determined by minimum power being dissipated by the reject load and maximum power into output load as shown in Fig. 90.

Effective power transferred to antenna or load with respect to differences in phase or gain of amplifier chains can be seen in Figures 91 and 92 respectively. As can be seen from these figures, a 60° error in-phase will cause only a 25% reduction in output power. If one amplifier has only half the output power of its counterpart, a 3% reduction in power (in reference to the combined input powers) will result. For example, let transmitter A = 5 kW, transmitter B = 10 kW, K^2 = Pa/Pb = 0.5 which gives a K of 0.707 which equals approximately 3% in Fig. 92 or approximately 14.55 kW of useful power out of the hybrid. Worst case for amplifier gain differences occurs when one transmitter is not producing any output. In this case half of the remaining transmitter's output (25%) will be dissipated in the reject load and the other half (25%) will be applied to the antenna.

A typical dual transmitter with notch diplexer is shown in Fig. 93.

Switchless Combiners

As discussed in the previous section, dual transmitters may be combined using hybrids to obtain maximum power. However, if one of the transmitters were to be disabled, the combined power output would drop to 25% of the nominal value. In order to continue operating at a reasonable power level many stations have employed coax switches to bypass the hybrid as shown in Fig. 94 to boost the transmitter power back to 50% of nominal operating power. During the switching process, however, it is necessary to take the transmitter off the air.

An alternative which allows the transmitter to stay-on-the-air while changing power from 25% to 50% is to use a "switchless combiner." A switchless combiner uses phase shifting with back-to-back 3 dB hybrids to accomplish the correct direction of power output.

A diagram of a transmitter using a switchless combiner is shown in Fig. 95. As identified in Table 1, nominally both transmitters are producing equal power and

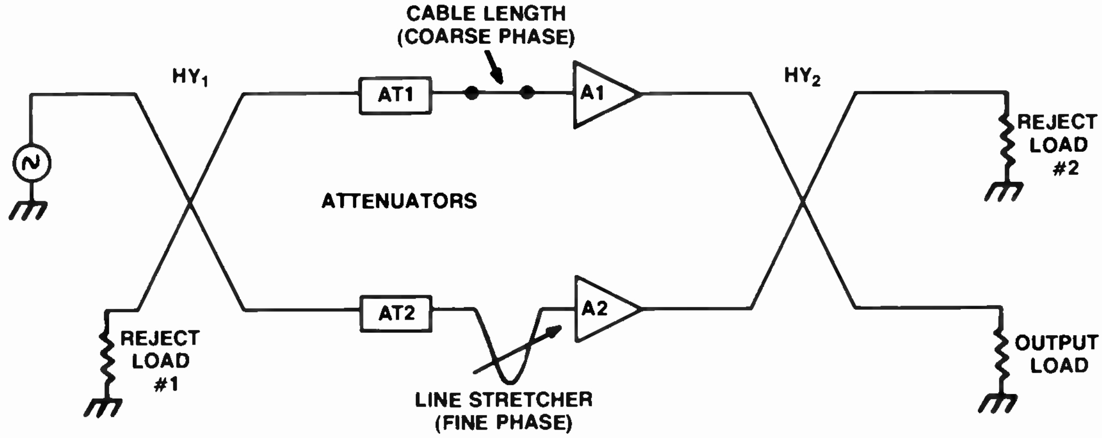


Figure 90. Typical parallel amplifier with single signal source.

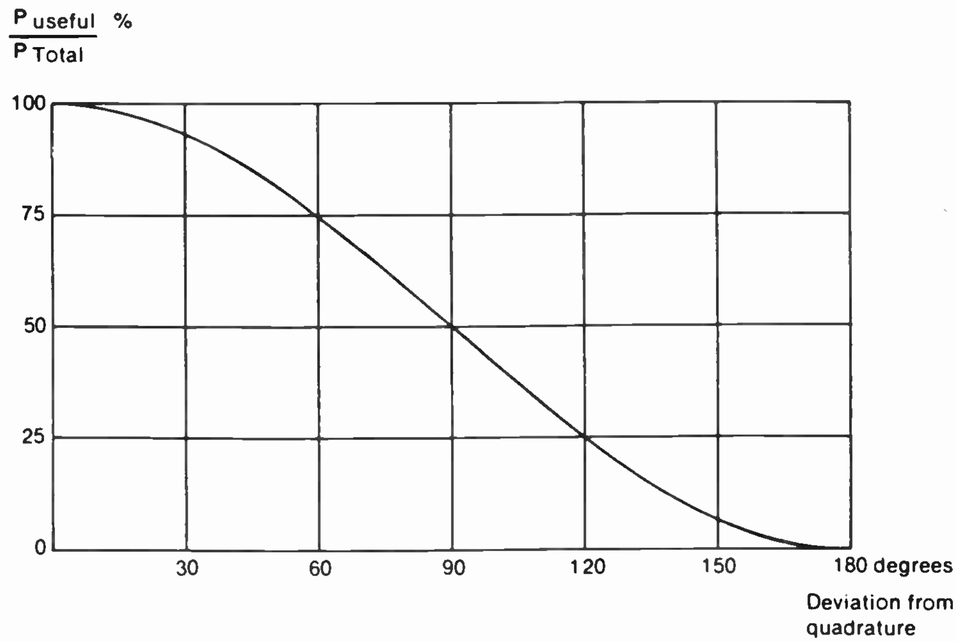


Figure 91. Hybrid coupler phase sensitivity.

are 90° out of phase. When transmitter A is disabled, Phase Shifter 1 (PS1) is energized changing the phase relationship as shown in Table 1. Transmitter A is routed to the switchless combiner load. When transmitter B is disabled (and assuming transmitter A is enabled) Phase Shifter 2 (PS2) is energized changing the phase relationship as indicated in Table 1. Now, transmitter B is routed to the switchless combiner load.

With one transmitter routed to the antenna and one transmitter routed to the combiner load, it is possible to perform adjustments on the transmitter connected to the load without impacting the on-air signal provided a separate exciter is used. (For more detail on this subject see "Subjective Tests on a Switchless Combiner TV Transmitter System" noted in the bibliography.)

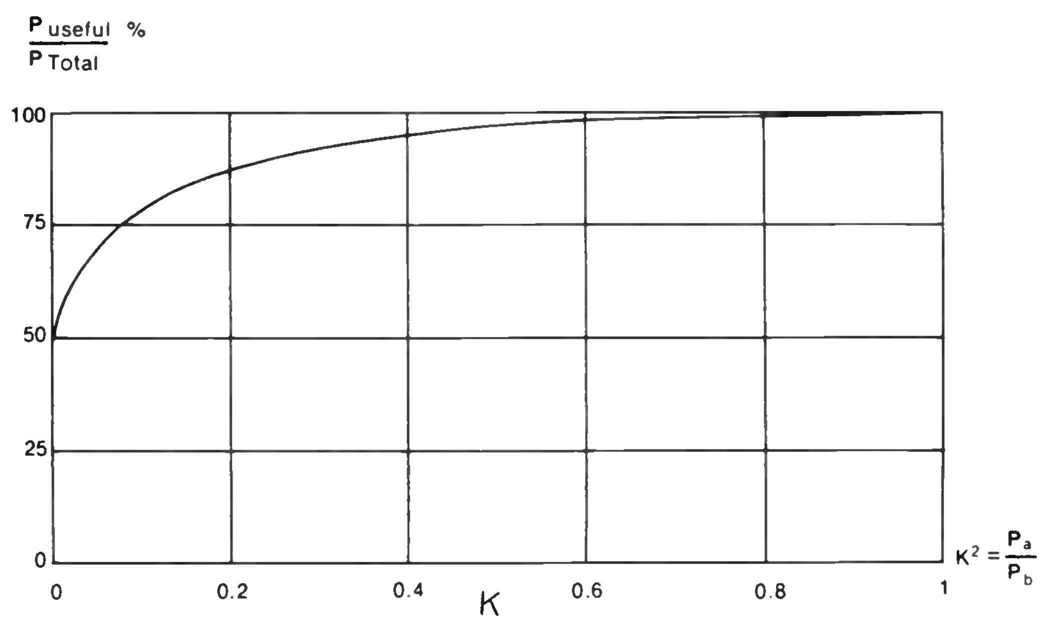


Figure 92. Power imbalance in hybrid couplers.

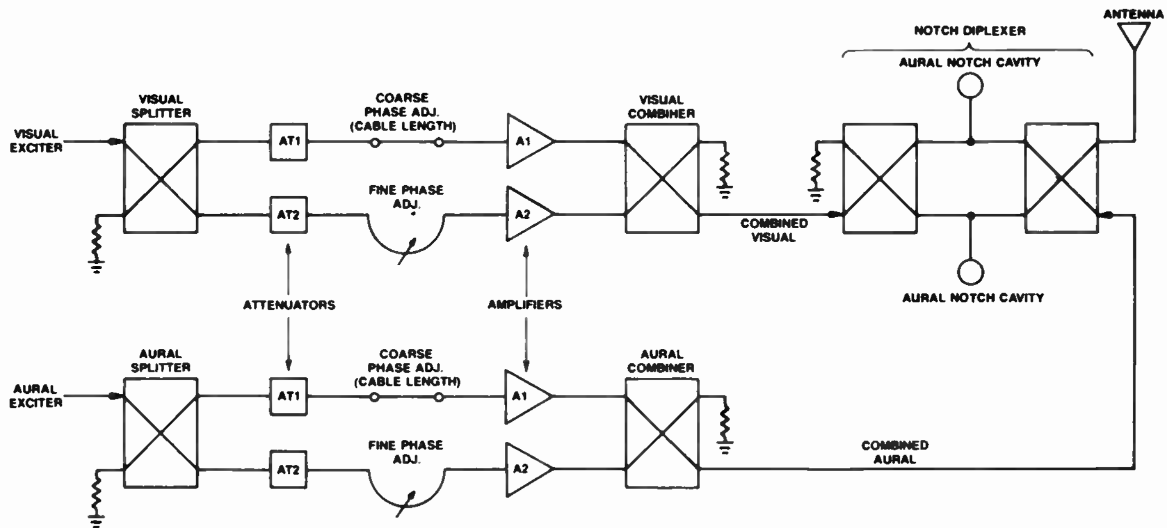


Figure 93. Typical dual transmitter with notch diplexer.

The switchless combiner concept has significantly increased the on-air availability of VHF transmitters.

UHF Magic Tee RF Systems

Since a similar problem exists for UHF transmitters, a similar concept can be applied again. However, since components are smaller many of them can be integrated into one assembly.

UHF transmitters typically are not dual systems but often do employ multiple visual amplifiers as shown in Fig. 96. A 180° hybrid implemented in waveguide accompanied by phase shifters located in the waveguide can accomplish the same function as the switchless combiner does with dual transmitters. See Table 2 for a diagram of what the phase relationships are and which amplifier is routed to which device.

One example of a unitized RF system, shown in Fig. 97, consists of a waveguide magic tee combining/switching system, a waveguide notch diplexer, two waveguide transfer switches, a 50 kW coaxial reject/test load, and a 100 kW waveguide test load interconnected and mounted within an open frame. The "magic tee" system provides visual amplifier power combining and power routing of either visual amplifier to the

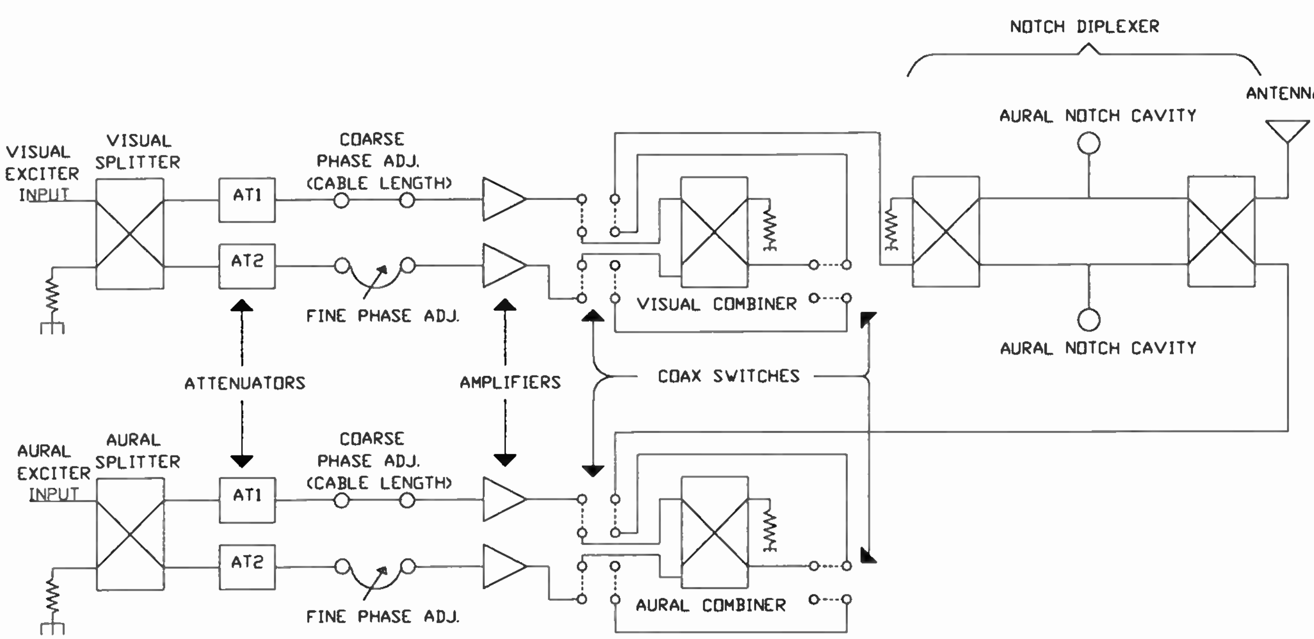


Figure 94. Typical dual transmitter with coaxial switchers.

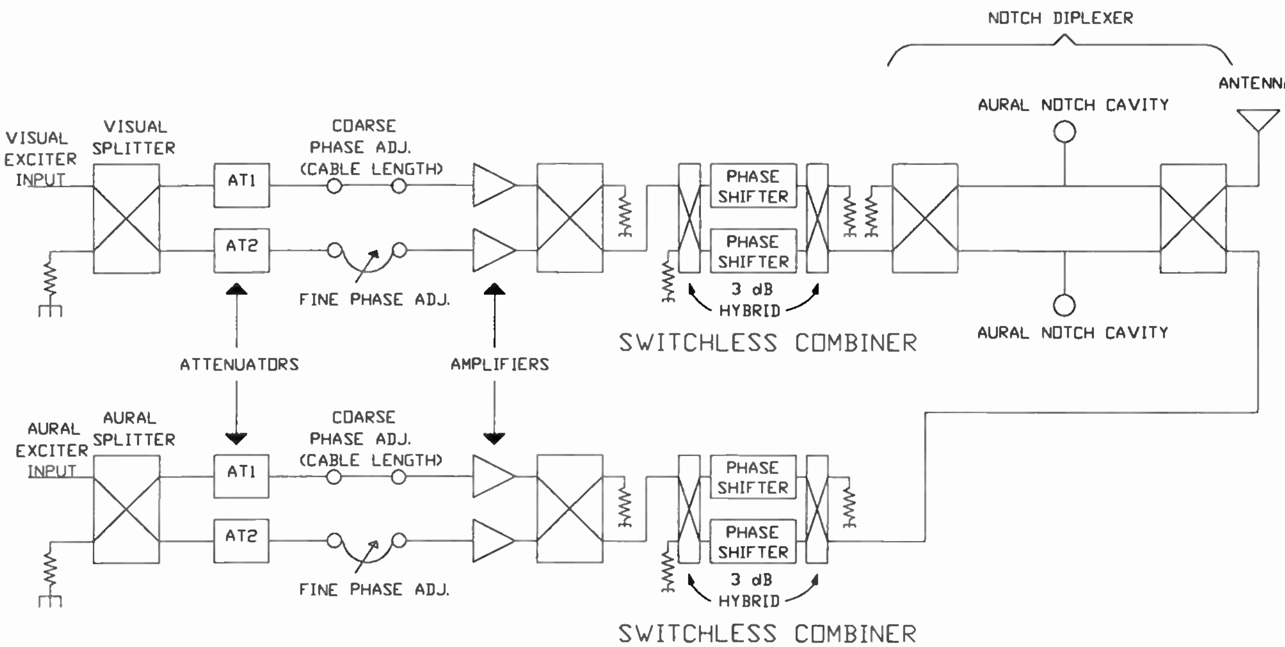


Figure 95. Typical dual transmitter with switchless combiner.

TABLE 1

A + B on AIR	180	270
A on AIR and B in TEST	180	180
B on AIR and A in TEST	90	270

TABLE 2

A + B	AIR	180	180
A AIR	B TEST	90	180
B AIR	A TEST	180	90

diplexer or 50 kW coaxial test load. The diplexer combines the visual and aural transmitter outputs. It is equipped with a lower color subcarrier notch filter and aural notch cavity de-tuning mechanisms to allow multiplexed visual/aural signals to be passed through the diplexer from visual input port to the antenna output port during emergency operation. One wave-guide transfer switch routes the diplexer output either

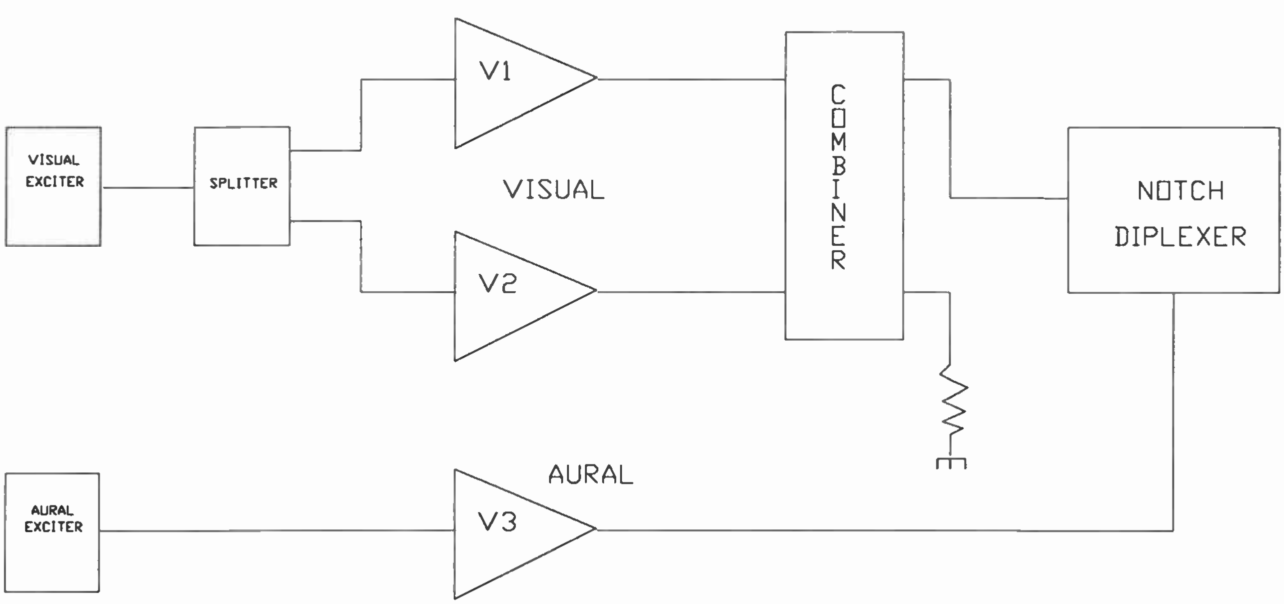


Figure 96. UHF transmitter with multiple visual amplifiers.

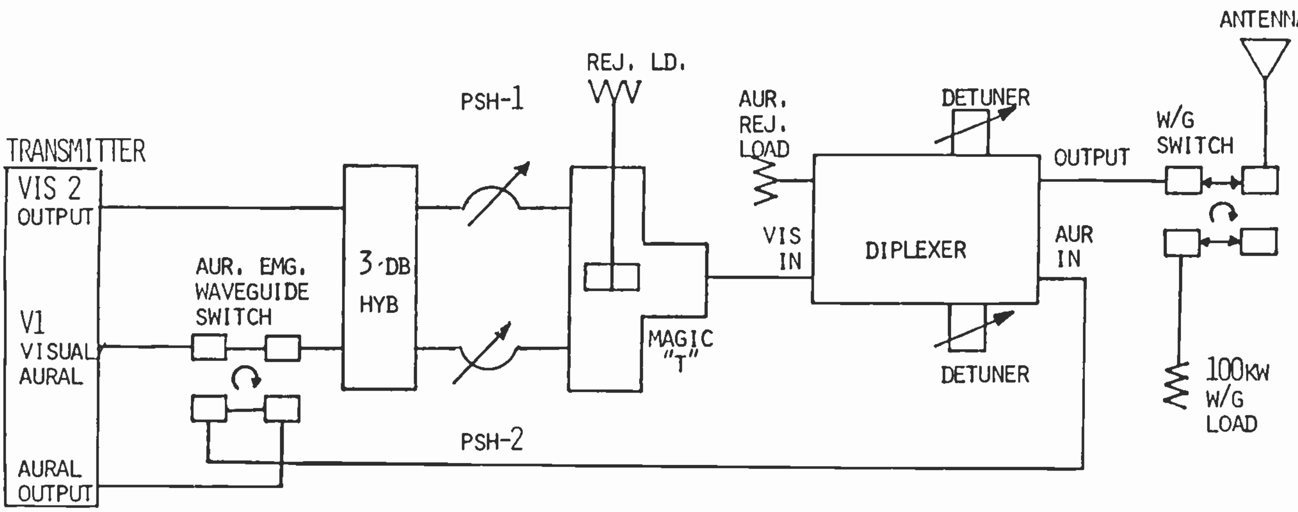


Figure 97. Unitized RF system.

to the antenna or to the 100 kW waveguide test load. The other waveguide transfer switch can route the output of a visual amplifier to the aural input of the diplexer. This feature allows a visual amplifier to become an aural amplifier in the event of failure of the normal aural amplifier. A control logic system is installed in the transmitter to coordinate transmitter and RF system operations. All functions of the RF system can be operated from the front of the transmitter or from a remote location.

PERFORMANCE MEASUREMENTS

The quality of the broadcast television transmitted signal is the resonsibility of the broadcaster. The introduction of IF modulation, solid state SAW vestigial filters and sophisticated precorrection circuits improved critical signal performance parameters two to one. Precision demodulators with synchronous detection and SAW filters are capable of near ideal detection. New test waveforms and digital signal synthesis now provide accurate test signal generation to complement the improved transmission facilities.

Home receivers employ high tech circuits including comb filters for luminance, stable SAW IF filters, high resolution display devices, and sophisticated LSI video processing circuits.

It is difficult to show the correlation between signal quality and audience ratings. The issue of viewer enjoyment is easier to demonstrate. Many scientific studies relating signal degradation to perceived quality have been made over the years.

The Electronic Industries Associates Standard RS-508 takes into account empirical quality factors and reflects a common denominator for new transmitter performance. This standard is a valuable reference document which describes performance parameters, standards and methods of measurements.

A thorough proof-of-performance at the time of installation is an invaluable record of normal operating performance. It also serves as verification of proper signal quality and emission standards. The number of detailed measurements required after the "proof" usually can be limited since several performance characteristics are a measure of the same impairment. Also some test waveforms are more useful for transmitter adjustment than others.

Transmitter measurements can be broken down into two primary categories as follows: frequency response and linearity.

Performance dependent on frequency response and group delay are sometimes referred to as linear distortions. That is, these distortions are not signal level dependent. In the time domain 2T, modulated sine-squared pulses, and multiburst waveforms are used in identifying and correcting distortions. In the frequency domain, swept amplitude and group delay measurements can be used but are normally reserved for out-of-service testing.

Linearity performance includes distortions which are signal level dependent and are a function of the instantaneous luminance level and on the average picture level over several lines. For example, level dependent chroma gain and phase are called differential gain and phase respectively. Luminance gain variation called low-frequency nonlinearity, and frequency response versus brightness which is the change in swept response with static luminance level are other level dependent distortions. Nonlinear responses can also change as a function of average picture level. To test for this, one line of the video waveform is alternated with four lines containing a static luminance level. Fig. 98 contains a modulated staircase with three average picture levels.

Modern transmitters are designed for unattended operation for extended periods of time. A transmitter operating with adequate cooling, regulated power lines and properly adjusted, may need only be checked in detail every three or four months or whenever a major component is replaced or repaired.

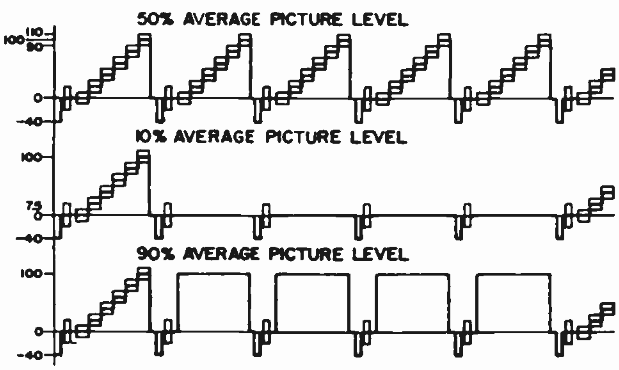


Figure 98. Modulated staircase waveform.

The following pretest checklist and sequence of tests are presented as a general guide for a properly adjusted transmitter.

Pretest Checks:

- Test equipment
- Input video
- Output transmission line, station load, and antenna
- AC mains input
- Record meter readings, adjustment settings, and key performance parameters

Transmitter Test Sequence

1. Exciter linear and nonlinear performance check.
2. Intermediate and driver linear amplifier performance check.
3. Swept frequency response.
4. Modulation depth (Include UHF pulser and sync ICPM adjustments).
5. Power output meter calibration.
6. Nonlinearity response checks. (Differential gain, ICPM, differential phase, etc.)
7. Linear response checks. (Group delay, pulse tests, composite waveform K factors, etc.)
8. Record meter readings, adjustment settings, and key performance parameters.

Many adjustments are interrelated and require returning to previous tests to verify proper performance. The sequence above is intended to minimize the number of adjustments.

At the time of a periodic measurement and adjustment sequence, it is a good practice to record meter readings and adjustment settings before and after the test. These recordings will indicate normal setting range and can aid in getting the transmitter back to near-normal in the event of mistuning errors during adjustments.

For off-air monitoring, a picture monitor is convenient for identifying gross video degradations quickly. Vertical interval test signals containing the composite waveform or other specialized test signals are useful for keeping track of depth of modulation, and several linear and nonlinear responses. These test signals are displayed on a waveform monitor and vectorscope. Records of transmitter meter readings and performance data are very useful in maintaining a transmitter at a high performance level. Maintenance logs should include date, time duration of outages, corrective action taken, and where possible, identify cause.

Monitoring TV Multichannel Sound

Maintaining a high quality aural signal requires high quality monitoring facilities. To be able to perform the proof-of-performance measurements is another important reason to have a monitor that demodulates the RF signal and separates the components in the composite MTS signal for analysis.

An ideal modulation monitor should consist of, but not be limited to, the following functions:

1. Demodulates the composite signal from the aural carrier.
2. Separates the components in the composite MTS signal for measurements.
3. Capable of off-air monitoring.
4. Covers all VHF and UHF channels.
5. Suitable for proof-of-performance use.
6. Contains a precision (BTSC) expander.

Equipment Needed

In order to make satisfactory measurements of video signals, a certain minimum of test equipment is required. Measuring the output of a television transmitter at baseband requires a precision television demodulator. For greatest accuracy, this demodulator should provide a zero carrier reference pulse for determining percentage of modulation, and ideally will have both a synchronous and envelope detection mode.

For most baseband measurements, a video waveform monitor is all that is needed, but to do a complete analysis of a system requires some additional equipment. The waveform monitor should include a provision for making measurements on vertical interval test signals (VITS) and provide filtering to allow separate examination of the luminance and chrominance components of color television signals.

In addition to the waveform monitor, a vectorscope must be used for making certain measurements on color signal components, particularly differential gain and phase measurements. The waveform monitor and vectorscope are all that are required to accomplish these measurements, but if a greater level of accuracy is desired in timing measurements, a conventional oscilloscope with an associated digital counter/timer would be a great asset.

PREVENTATIVE MAINTENANCE

A good preventative maintenance program should include periodic inspection and cleaning of the equipment. A vacuum cleaner is preferred to remove dust instead of compressed air which will simply blow the dirt into the air and let it fall back down on something else. A paint brush can be used to dislodge dust from delicate circuit boards. Avoid using a nylon bristled brush with a plastic handle as the static charge may damage CMOS or other static sensitive components. A natural bristle brush with wooden handle and metal binding is recommended.

High voltage wires and insulators must be cleaned with denatured alcohol, or other cleaner capable of removing the dirt without leaving any residue. Meter cases are cleaned with Glass Wax or other nonstatic cleaner.

Air filters should be replaced or cleaned as necessary to maintain adequate air flow to the equipment. A second set of washable filters can save time when using a single transmitter in critical service by quickly switching the clean filters and washing the dirty ones later. Blowers should be inspected to see if the curved fins are filled with any debris that would reduce air flow. Motor windings may collect a layer of dirt and interfere with the cooling of the motor itself. The fins of high-power tubes must be cleaned of any obstructions which may have passed through the air filters. Bearings should be lubricated and checked for excessive noise.

Color change in silver plated cavity parts can be a sign of over-heating and may require the disassembly of the cavity to check for obstructed air passages or loose connections. Set-screws in gear and chain drive tuning mechanisms should be checked for tightness. Black silver-oxide is a good conductor and need not be removed. Small parts should be dipped in Tarnix® for cleaning. Be sure to flush the parts after cleaning to remove any residue. Scotch-Brite® is a good nonmetallic cleaning pad for silver-plated parts. Remember that you do not want to remove the plating when cleaning these parts.

High current wires may move during turn-on surges and can suffer abrasions which may eventually cause an arc if not properly dressed away from sharp edges. Wiring on terminal boards may loosen through thermal cycling and vibration. All connections must be checked to be sure they remain tight. If wires need to be replaced, the correct gauge, voltage rating, and temperature rating must be considered when selecting the replacement.

Edge connectors on printed circuit boards should be cleaned with Cramolin® or other cleaner. A small amount of this cleaner is applied to the edge connector and then removed with a lint-free cloth. Do not use pencil erasers as this will remove gold or silver plating from the edge traces, and could degrade the connection or create an intermittent later as the sulfur in the eraser causes chemical reaction to the edge connector material.

Back-up systems or emergency modes of operation should be checked periodically. The worst time to discover trouble in the back-up equipment is when you need to put it on the air. Relays need some exercise to keep their contacts polished and in working order.

Transmitter site cleaning should also include a check of the building for such things as leaks in the roof which may cause damage to the transmitter, and presence of insects or small animals which may wander into unwanted areas.

Intake blowers with filters capable of creating positive pressure in the room can minimize the need for cleaning by keeping the dust out. A careful record must be maintained in order to establish a good history for future reference. Such a log will include a description of what was done, when it was done, and the name of the person performing the work. The question "How often must the transmitter be cleaned?", can be answered by "How often does it need cleaning?" Seasonal events such as harvesting in farm locations, severe weather, construction projects, in and around the transmitter building can require special action, but usually a pattern will emerge that allows the maintenance to be scheduled on a regular basis.

A complete set of meter readings taken when the equipment is working properly and updated weekly or monthly can greatly assist the engineer when trying to diagnose problems.

Create a maintenance program with weekly, monthly, quarterly, semi-annual, and annual tasks evenly spread throughout the year.

The following list of items can be used to develop a maintenance routine for any broadcast television transmitter facility:

Maintenance Items

Prefilter manometer readings
Inspect prefilters
Replace prefilters
Post-filter manometer readings
Cabinet input air manometer reading
Inspect transmitter air filters
Replace transmitter air filters
Vacuum cabinets
Clear tube fins or accumulated dust
Measure blower currents
Clean fan blades and motor windings
Connections checked for tightness
Inspection of MOVs

Recommended Data to be Recorded

Record all parameters on meters or user displays
Record DC input power and calculate dissipation (if applicable)
Record transmitter currents in black picture and at idle (no RF drive)

Visual Performance Checks

Ensure proper video level
Ensure proper sync level
Optimize differential gain
Optimize ICPM
Optimize differential phase
Optimize group delay using T pulses

Ensure proper power calibration
Ensure proper audio processor set-up
Optimize swept response

Aural Performance Checks
Ensure proper modulation levels
Ensure proper SCA input levels
Ensure proper power calibration
Ensure proper audio processor set-up
Optimize audio frequency response
Check and optimize distortion
Optimize stereo separation
Check and optimize crosstalk between MTS channels

Control System Checks
Verify proper operation of all interlock circuits
Verify proper operation of all overload circuits
Verify proper operation of all control processes (VSWR foldback, filament timing, coax switches, etc.)

RF amplitude response as a function of frequency is the variation in gain over the frequency range of the channel. The use of a television sideband analyzer or other frequency selective voltmeter provides a suitable means to measure a television transmitter amplitude response versus frequency.

At one time or another, just about every component part of a video signal must be measured or adjusted. It is good engineering practice that each part of a transmission plant including studio, distribution, transmitter, and monitor equipment provides minimal distortion to the video signal. Furthermore it is not recommended that one part of the system correct for another part except within a given system such as the precorrection circuits used in a transmitter to compensate for certain RF components. The signal going into the transmitter should be a standard NTSC video signal meeting its full specifications. The transmitter's correction circuits should not have to compensate for incoming video problems.

The overall amplitude of the signal and each of its component parts have strictly defined levels, and the relationship between the parts is also critical. Refer to RS-508 for a description of specific electrical performance parameters and standard test methods.

AIR SYSTEMS FOR TRANSMITTERS

Most electronic equipment which requires forced air cooling has the required blower or fan already designed into it. However, equipment such as transmitters, dissipate large amounts of power. Air which has already passed through the equipment and has picked up heat must be removed from the immediate vicinity of the equipment (exhaust air) in order to prevent the hot air from being recirculated. In addition to the air removal requirement, provisions must be made for sufficient make-up air (intake air) to replace that which has been circulated through the equipment and removed.

If the equipment user is to provide adequately for hot air exhaust and fresh air intake, the maximum and minimum environmental conditions which the equipment may operate in and the following information from the equipment manufacturer must be known:

Altitude: ______________________

Max. temp: ______________________

Min. temp: ______________________

Total air through the transmitter:

CFM ______________________

Pressure drop within the critical portion of the air circuit (across the transmitter tube for example. This is required for checking purposes after the system is built.)

Air pressure (Inches of water) ______________

Air temperature rise through the transmitter: ______

Air exhaust area: ______________________

Exhaust Air

Equipment layouts usually provide for heated air to exit from the top surface of the cabinet. The size and location of this exhaust area is usually shown on a manufacturer supplied outline drawing.

Most broadcast equipment internal air systems are designed to be operated into free space (back pressure of 0.0 inches of water) so any exhaust ducts must have minimum loss. A good practice usually is to design for no more than +0.1 inches of water pressure in the duct close to the exhaust area of the transmitter.

Any exhaust installation other than a large cross section duct (equal to the cross section of the transmitter exhaust port) with no bends, and with a long radius turn outside the transmitter building for weather protection, will need an exhaust blower or fan.

Keep in mind that the recommended system is sized only for cooling the transmitter and any additional cooling load in the building must be considered separately when selecting the air system components. The transmitter exhaust should not be the only exhaust in the room as heat from the peripheral equipment would be forced to go out through the transmitter.

The "sensible-heat" load is the sum of heat loads such as solar radiation, heat gains from equipment and lights and personnel in the area that is to be cooled.

The following hypothetical exhaust duct design illustrates the key cooling concepts:

Air flow (volume) through transmitter = 325 CFM (ft^3/min)

Air exhaust area = 3.4 square ft

$$\text{Air exhaust velocity} = \frac{325\ \text{CFM}}{3.4\ \text{ft}^2} = 94.5\ \text{ft/min}$$

The 94.5 feet per minute air velocity is relatively low which will allow a transition to a smaller diameter pipe if desired. Assume a transition down to a 10-inch diameter pipe, thus the following:

$$A = \frac{D^2}{4 \times 144} = \frac{100}{4 \times 144} = .545\ \text{square ft}$$

$$V = \text{CFM}/A$$

where: D = Duct diameter
A = Exhaust duct cross-sectional area
V = Exhaust duct air velocity

$$V = \frac{325\ \text{CFM}}{\text{dia. pipe } .545\ \text{ft}} = 596\ \text{ft/min air velocity in a } 10''$$

In a 10-inch diameter pipe the air friction chart in Fig. 99 gives 0.06 drop per 100 ft of pipe with a flow of 325 CFM. Assuming for this hypothetical design there is 20 ft of straight pipe to a roof, with two 90° elbows to turn the pipe down for weather protection, the total loss of the exhaust system may be estimated as follows:

Fig. 99 shows loss is 0.06″/100 ft. Since 20 ft is 0.2 of 100 ft,

0.2″ × 0.06″ = 0.012″ of water pressure drop in 20 ft of pipe

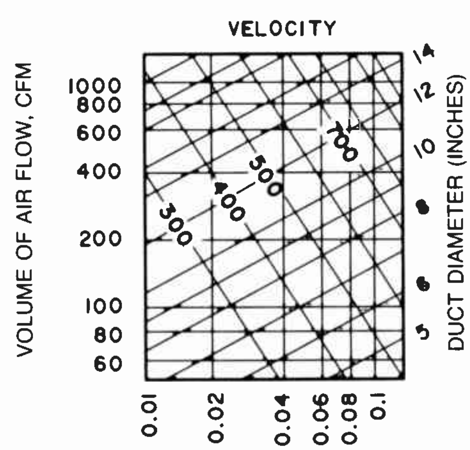


Figure 99. Friction loss in pipes.

Considering the static pressure drop of two 90° elbows, to give a 180° bend in the pipe, it is found from Fig. 100 that a 10 inch elbow at 900 ft/min gives less than 0.01 pressure drop, and there is just under 600 feet per minute in the system. Adding the two drops (one for each 90° elbow) and the 20 ft section of 10 inch pipe together, the result is:

0.012″ + 0.01″ + 0.01″ = 0.032″ of water total pressure drop

Therefore, no exhaust fan is necessary as the 0.032 inch of pressure is less than the 0.1 inch pressure level that requires a fan.

If the installer has a problem in exhausting the transmitter in this simple fashion (for example a roof exit is not available), and it is required to add two additional 90° elbows and a straight run of 10 inch pipe 100 ft long, the pressure drop in the hypothetical design exhaust system will be:

Friction loss in 20 ft of 10″ dia. pipe	=0.012
4 elbows × .01″ (water) each	=0.04″
Friction loss/100 ft of 10″ dia. pipe	=0.06″
	=0.112″ of water

The 0.112 inch of pressure now exceeds the level at which a fan is needed.

Because the transmitter manufacturer recommends no more than a 0.1 inch water pressure loss in an exhaust system, good engineering practice indicates that an exhaust fan is now required in this configuration. The performance curve on the fan shown in Fig. 101, indicates that it will deliver 325 to 330 CFM into 0.1 inches of water back pressure. This is sufficient to handle all of the air coming from the transmitter and overcome all of the estimated duct losses in this configuration.

The outline drawing illustrated in Fig. 102 shows a typical exhaust duct and blower system. The recommended minimum ceiling height to properly handle exhaust air as shown is 12 ft The outline drawing also indicates a typical air intake and prefilter system.

Intake Air

An intake vent and blower should be sized to provide a desireable slight positive room pressure. Installing a manometer to sense pressure drop across the filters can help determine replacement interval of prefilters.

If existing space on site will not permit the construction of the transmitter manufacturer's recommended air system, care must be taken to modify the design to fit the available space and still properly cool the transmitter.

Referring again to Fig. 101, examination of the curve shows that at 325 CFM delivery, this fan will develop a static pressure of 0.1 inch water which is sufficient to overcome the less than 0.01 inch pressure drop of a 10-inch diameter 90° elbow installed on the outside wall of the building for a weather hood, as well as a filter and screen to clean the incoming air.

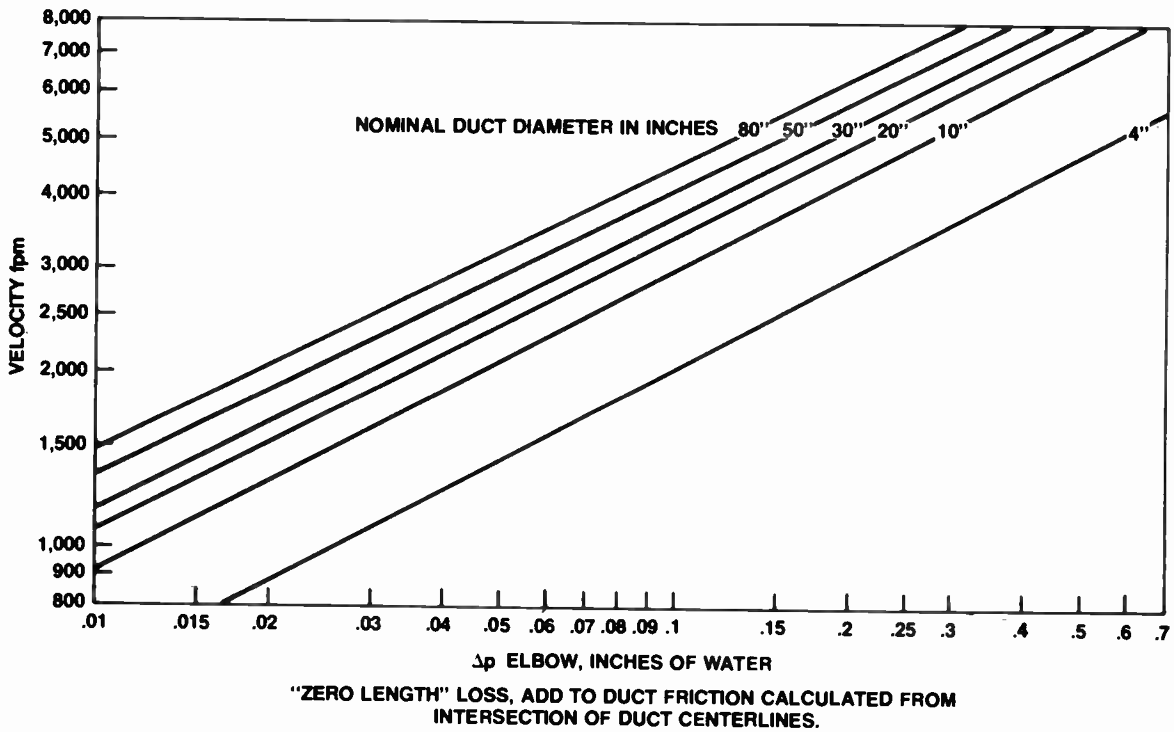


Figure 100. Friction loss in 90° elbow.

20 ft of pipe = .012"
2 elbows at .01" each = .02"
.032" total loss, maximum.

Additional flushing air is recommended for the removal of heat from any surrounding equipment that shares space with the transmitter. It is a recommended guideline to keep input air no greater than 5° C above ambient.

The above calculations are a very simple example of an air system. Most television transmitters, both UHF and VHF, have multiple enclosures with large volumes of air required. In these cases the equipment user is required to combine the heat exhaust air from several enclosures into a complex duct system. When these circumstances occur, the safest practice is for the user to contract for the services of a heating, ventilating, and air-conditioning (HVAC) consultant or company. The cost involved in having the best possible air exhaust and supply system will pay for itself in extended transmitter life and lower service costs.

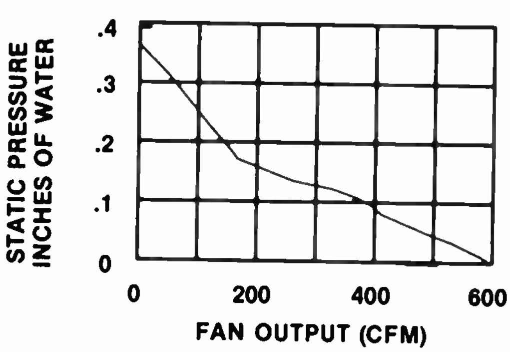


Figure 101. Vaneaxial Caravel fan performance curve.

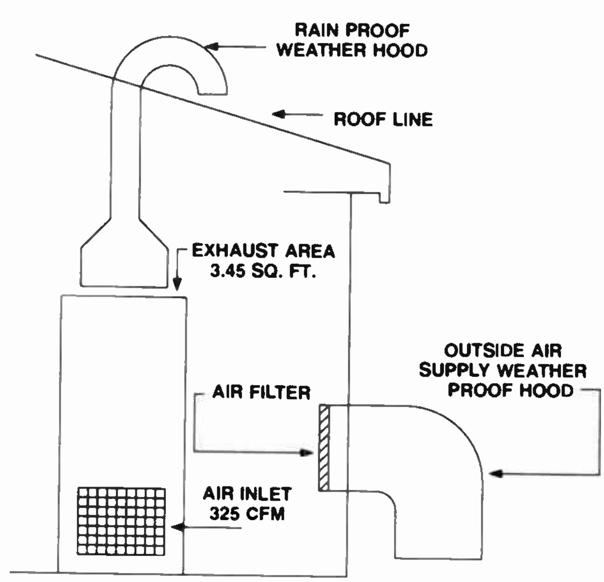


Figure 102. Suggested intake and exhaust duct installations.

Air Conditioning

It is a common practice to set the transmitter into a sealed wall to produce a plenum chamber and supply it with outside air, while providing separate air conditioning for the front side to cool personnel and source equipment.

In areas with severely polluted air, it may be necessary to run the transmitter on air-conditioned air to avoid bringing in corrosive salts or gaseous contaminants.

The amount of air conditioning will depend on several factors. It is strongly recommended that the air conditioning be shared by a number of units rather than one large central system so that operation can continue in the event of the failure of one unit. Air conditioning units are usually listed by "tons" of cooling capacity with "one ton" equal to 12,000 BTU per hour.

Again, consult experienced professionals in the area of HVAC design to achieve the desired result and to prevent problems from developing in the future.

ENDNOTE

1. Based on the stereo and auxiliary sound system recommended by the Broadcast Television Systems Committee (BTSC) of the Electronics Industries Association. See Chapter 3.6, "Multichannel Sound," of this Handbook.

BIBLIOGRAPHY

Weirather and Hershberger, "Amplitude Bandwidth, Incidental AM, and Saturation Characteristics of Power Tube Amplifiers for FM", IEEE Transactions on Broadcasting, March 1983, Vol BC-29, No. 1, p. 14.

W.L. Behrend and R.L. Rocamora, "Television Transmitter System Consideration for Multichannel Sound", EIA report on Multichannel Television Sound, Vol. 2A, October 31, 1983, Appendix D.

Electronics Industries Association. RS-508 "Electrical Performance Standards For Television Broadcast Transmitters".

Hans Schmid. "The Measurement Of Short-Time Waveform Distortion In NTSC TV Facilities." IEEE Trans. Broadcast, Vol 18, No. 3, Sept., 1971.

Charles Rhodes. "The 12.5T Modulated Sine-Squared Pulse For NTSC." IEEE Trans. Broadcast, Vol 18, Mar., 1971.

Pieter Fockens, Carl Eilers. "Inter-Carrier Buzz Phenomena Analysis And Cures." Transmitter Consumer Electric Vol 27, No. 3, August, 1981.

Bailey Neal. "Frequency And Amplitude Dependent Phase Effects In Television Broadcast Systems." IEEE Trans. On Consumer Electronics, Vol 23, No. 3, August, 1977.

T. M. Gluyas. "Influence Of RF Output Systems On TV Transmitter Picture Quality." NAB Convention Record, April, 1971.

L. E. Weaver. "The Multipulse Waveform." Application Note No. 30, Tektronix Corporation.

W. A. Resch. "Visual Performance Characteristics Which Affect Multichannel Sound." NAB Convention Record 1982.

W.L. Behrend. "Effects Of Incidental Phase Modulation Of TV Transmitters Or Other Circuits, On TV Signals." IEEE Trans On Broadcasting, Vol 19, No. 3, September, 1973.

Varian Associates, Inc. "Integral Cavity Klystrons for UHF TV Transmitters."

Varian Associates, Inc. "External Cavity Klystrons for UHF TV Transmitters."

Gregory L. Best. "Transmitter Conversion for TV Multichannel Sound." NAB Convention Conference Proceedings, 1984.

United Sheet Metal Bulletin 100-5-178 "Engineering Design Manual for Air Handling Systems".

Rotron Incorporated. "The Caravel Fan" Catalog Sheet E-3050C.

McGraw-Hill. Marks's Mechanical Engineers Handbook.

Robert Plonka. "Group Delay Corrector For Improved TV Stereo Performance." NAB Convention Conference Proceedings, 1989.

N. Ostroff, A. Whiteside, L.F. Howard. "An Integrated Exciter/Pulser System for Ultra High-Efficiency Klystron Operation." NAB Convention Conference Proceedings, 1985.

N. Ostroff, A. Whiteside. "Using Klystrode® Technology to Create a New Generation Of High Efficiency UHF-TV Transmitters." NAB Convention Conference Proceedings, 1986.

N. Ostroff, A. Whiteside, A. See, R.C. Kiesel. "A 120 KW Klystrode® Transmitter For Full Broadcast Service—The "SK" Series." NAB Convention Conference Proceedings, 1988.

N. Ostroff, A. Whiteside, A. See, R.C. Kiesel. "Klystrode® Equipped UHF TV Transmitters: Report On The Initial Full Service Station Installations." NAB Convention Conference Proceedings, 1989.

N. Ostroff. "Recent Advances In Klystrode® Equipped Transmitters." NAB Convention Conference Proceedings, 1990.

P. C. Turner. "Recent Developments In Solid-State TV Transmitters." NAB Convention Conference Proceedings, 1989.

G. Badger. "The Klystrode®, A New High-Efficiency UHF TV Power Amplifier." NAB Convention Conference Proceedings, 1985.

Donald Priest, Merrald Shrader. "The Klystrode®—An Unusual Transmitting Tube With Potential For UHF-TV." Proceedings of the IEEE, Vol. 70, No. 11, Nov. 1982.

George Badger. "Klystrode® Technical Performance For Modern High-Efficiency UHF-TV Transmitters." International Broadcasting Convention Proceedings, 1986.

Merrald Shrader, Donald Priest. "Klystrode® Update." IED Meeting, San Francisco, CA, December 1984.

Michael Chase. "The 15 kW Klystrode®: An Efficient Air Cooled Device for Television Broadcast Service." Engineering Conference, November 10-12 '87.

Robert Weirather. "A Distributed Architecture For A Reliable Solid-State VHF Television Transmitter Series." NAB Convention Conference Proceedings, 1989.

Martyn Horspool. "Designing An All Solid-State VHF TV Transmitter For High On-Air Reliability." International Broadcast Engineer, May 1989.

G. L. Best, F. A. Svet. "Architecture For A High Reliability VHF Transmitter." IEEE Broadcast Symposium, Sept. 1988.

Ulrich Gysel. "A New N-Way Power Divider/Combiner Suitable For High Power Applications." IEEE-MTTS-5, International Symposium Digest, 1975.

Gerald Collins. "Combiner For Platinum Series TV Transmitter." Harris Corporation Technical Paper, 1989.

R. S. Symons. " Depressed Collector Klystrons For High-Efficiency UHF-TV." Varian Technical Report, 1986.

Earl McCune. " MSDC Klystron Progress Report." SBE National Convention Proceedings, 1989.

James Pickard. "Field Performance of a Multiple Stage Depressed Collector Klystron." NAB Convention Conference Proceedings, 1990.

Greg Best, James Pickard, David Danielsons. "Subjective Tests on A Switchless Combiner TV Transmitter System." NAB Convention Conference Proceedings, 1987.

Electronic Industries Association. "BTSC System Recommended Practices." EIA Systems Bulletin No. 5, July 1985.

ANSI/IEEE C62.41 Standard (Sometimes Referred to as the IEEE-587 Standard).

George M. Badger, The Klystrode®, A New High-Efficiency UHF TV Power Amplifier. National Association of Broadcasters, Las Vegas, Nevada, 1985.

3.6 Multichannel Television Sound

Edmund A. Williams
Public Broadcasting Service, Alexandria, Virginia

INTRODUCTION

Multichannel Television Sound (MTS) is a generic term for the process of adding subcarriers to the aural carrier of a television station. The development of a stereo MTS system was accomplished through the Broadcast Television Systems Committee (BTSC) of the Electronics Industry Association (EIA) and is known as the BTSC system for television stereo sound. Some of the subcarriers in the BTSC system are designed to be received by the general public and may be used for a variety of purposes including: stereophonic sound, secondary audio channel for a second language or other program related aural service, commercial communications, and data. Other subcarriers may be used by the broadcaster for professional uses such as news crew cues, transmitter telemetry or other digital or analog, and aural or data services. Non-MTS subcarriers need not be related to the program on the video or aural portion of the television signal itself. The BTSC stereo system is completely compatible with monophonic receivers.

Modern broadcast television transmitters are designed to accommodate multichannel sound. Older transmitters are likely to require revised tuning procedures and more attention to overall performance will be necessary to prevent degradation to the stereo signal. Most importantly, new testing and measurement techniques are required to insure that the BTSC multichannel sound signals are transmitted to the receiver in good condition,

The introduction of BTSC multichannel sound, as an all-industry recommendation, permitted receiver manufacturers to develop new models that offer good stereo performance. In 1990, over 25% of all television receivers sold in the U.S. had BTSC stereo. Nearly 20% of all U.S. TV homes have stereo. Most new television programs are produced in stereo. More than 75% of all television stations in the U.S. have BTSC transmission capability. Virtually all video services available to the public (VCR, CATV, satellite, video disk) contain stereo sound tracks.

The addition of stereo to television broadcasting accelerated the use of stereo sound tracks for television program production. Channeling the stereo sound throughout the broadcast facility in turn created challenges to audio design and production, installation, distribution, operation, and maintenance. (See other chapters in this handbook for details.)

During the introduction of the new stereo service, many stations employed stereo synthesizers to provide viewers with an immediate sensation on newly acquired stereo television receivers. Because a large portion of original program material is now produced in stereo, the synthesizer is no longer needed in the audio chain. In fact, using the synthesizer anywhere in the audio chain other than in the studio can severely degrade mono and stereo sound (and enhancements such as surround sound) which in turn may cause distraction and discomfort to the viewer who may make incorrect adjustments to the receiver.

BACKGROUND

The development of the multichannel sound system was accomplished by broadcast industry organizations through the Broadcast Television Systems Committee (BTSC) of the Electronics Industry Association (EIA). The EIA established a laboratory and conducted transmission tests on several potential multichannel sound systems. The system selected by a vote of the BTSC members uses the transmission parameters developed by Zenith Electronics Corporation and the noise reduction system developed by dBx Incorporated. This combination, called the BTSC Multichannel Television Sound system, was recommended to the FCC for adoption as an industry-wide standard. The EIA recommendation was approved by the Federal Communications Commission (FCC) on March 29, 1984 in their Second Report and Order (Docket 21323). While not adopting the BTSC system as such, the FCC provided protection to the establishment of the BTSC system which in turn resulted in a well developed de facto standard. In summary:

1. The FCC protects BTSC equipped receivers from interference from other MTS subcarrier schemes (see FCC Rule Section 73.682(c)).
2. The FCC Office of Engineering and Technology published Technical Bulletin OET-60A, a document that contains the BTSC specifications that are described in FCC Rule Section 73.681.
3. The FCC permitted stations to implement not only the BTSC MTS system but any other MTS system of subcarriers a station desired so long as no interference resulted to receivers designed to BTSC specifications or to regular monophonic sound television receivers.

While the BTSC system is not a standard as such, the Commission provided the next best thing—protection. Therefore, transmitter and receiver manufacturers are able to produce equipment conforming to the BTSC system with full knowledge that a compatible, protected MTS system is recognized by both FCC and industry. The EIA also developed a set of recommended operating practices aimed mainly at transmitter and receiver manufacturers entitled "BTSC System: Television Multichannel Sound Recommended Practices."

Television stations may employ aural subcarriers for virtually any purpose as long as they do not interfere with the BTSC system or normal monophonic receivers' operation.

It should be noted that the multichannel sound system adopted for use in the United States is unique. The multichannel sound transmission schemes used in other countries have different configurations. In Japan, an FM subcarrier carries either a stereo sound channel or a second language channel but not both. West Germany adopted a two-carrier sound transmission system providing either stereo or second language. In Great Britain, a digitally modulated carrier for stereo sound is inserted above the normal monaural FM carrier. In comparison, the U.S. system with its combination of AM and FM subcarriers can, if desired, simultaneously transmit stereo and a separate program or second language, plus a subcarrier for professional or nonpublic uses. Further, the U.S. system permits stations to employ other subcarrier arrangements, when not using the BTSC system, to serve specific station requirements.

A noise reduction or companding system for the stereo and separate audio program channels was incorporated into the BTSC system from the beginning of MTS service. A high performance compressor is built-in to all BTSC encoders. A matching expander is built-in to all BTSC decoders. They are, therefore, part of the overall system design and not merely an add on feature.

THE BTSC SYSTEM

The design of the BTSC MTS system took into account several major performance objectives and considerations. Among the more important factors were:

1. Compatibility with existing monophonic sound television receivers.
2. Ease of modification of transmitters.
3. Full sound fidelity.
4. The ability of the MTS signals to pass through cable television systems, master antenna systems, and TV translators.
5. The ruggedness to withstand typical transmission impairments (noise, multipath, interference) in the path between the transmitter and receiver.

These factors influenced the design criteria with the result that the single aural carrier system with subcarriers was adopted with an audio companding (noise reduction) system. Also, a separate audio program channel for special programming purposes is included along with space for a nonpublic subcarrier for station use.

It is recommended that all stations contemplating stereo operation obtain a copy of Bulletin OET-60A entitled "Multichannel Television Sound Transmission and Audio Processing Requirements for the BTSC System" published by the FCC.

The BTSC multichannel sound system is designed to be compatible with existing monophonic television receivers and provide high quality stereo sound transmission. The system requirements and performance capability is given below.

BTSC System Requirements

1. The system is compatible with existing monophonic television receivers. Establishing the correct operating levels in the stereo encoder is essential for proper stereo and monophonic operation. Monophonic performance is also a function of the balance of the original audio program material and the distribution and processing within the broadcast plant.
2. The system provides both a stereo and separate audio program (SAP) channel simultaneously. The SAP channel is available on most BTSC equipped television receivers. The quality of the SAP channel, while somewhat limited compared to the stereo and mono channels, is a function of the original program material and the processing used in the broadcast facility.
3. There are provisions for other professional use (PRO) or station use subcarriers. The PRO channel may be used for a variety of purposes for station operations or commercial services. It is not capable of being received on consumer television receivers.
4. A noise reduction (companding) system is employed for both stereo and separate audio program channels. Nearly 30 dB reduction in noise is

achieved through the use of advanced circuit design, compression algorithm, precision components, proper setup, and routine maintenance. The decoders used in all consumer BTSC equipped television receivers are matched to the encoders used in the transmitter. The Electronic Industries Association published a detailed description (EIA Television Systems Bulletin #5) of the entire system entitled "Television Multichannel Sound BTSC System Recommended Practices." and is used by manufacturers of television transmitters, receivers, cable headend equipment, add on adapters, and other equipment which contains BTSC encoders and decoders.

5. The aural carrier deviation capability in the aural transmitter exciter and amplifiers must be increased to accommodate the new subcarriers. The deviation for monophonic transmission remains at 25 kHz. With stereo, SAP, and PRO the deviation may be increased to 73 kHz. Insufficient bandwidth in aural transmission components will result in distortion and crosstalk between channels in the receiver.

Excellent stereo audio quality and separation between stereo and SAP channels can easily be obtained from BTSC stereo exciters. However, the transmission system will degrade some performance parameters, depending upon how well the RF components are adjusted and maintained.

Fig. 1 shows the MTS performance objectives that can be achieved with properly operating BTSC and aural RF transmission equipment, routine maintenance, and good test equipment.

Channel	*Freq Response*	*Noise*	*Distortion*
Main (monophonic)	20–15kHz +/− 1dB	>58dB	<1.0%
Stereo (difference)	100–15kHz +/− 1dB	Dyn*	<1.0%
Separate Audio Program (SAP)	50–10kHz +/− 2dB	Dyn*	<2.0%
Professional Use (PRO)	100–3kHz +/− 3dB	>40dB	<5.0%
Stereo Separation	>40 dB (without NR encoding)		
Crosstalk into SAP or PRO	>50 dB		
Crosstalk into Main (any)	>60 dB (>40 dB for stereo into main)		

*Dyn indicates that the noise varies dynamically with program content and the characteristics of the noise reduction system. Subjectively, the noise will normally be 60 dB below peak operating levels.

Figure 1. MTS performance objectives.

Degradation to any of the above performance parameters will occur under the following conditions:

1. Crosstalk: nonlinear amplifiers, excessive incidental carrier phase modulation (ICPM).
2. Frequency response, noise, and distortion: audio circuitry in stereo encoder or exciter.
3. Noise reduction system artifacts such as pumping or changes in level: improper adjustments in encoder. (Do not adjust without knowledge of the system and proper instrumentation.)
4. Buzz in the received audio: excessive ICPM in the transmitter, video overmodulation, or video energy in the aural carrier passband.

The BTSC System Transmission Format

The transmission format of the BTSC baseband configuration is shown in Fig. 2. The stereophonic

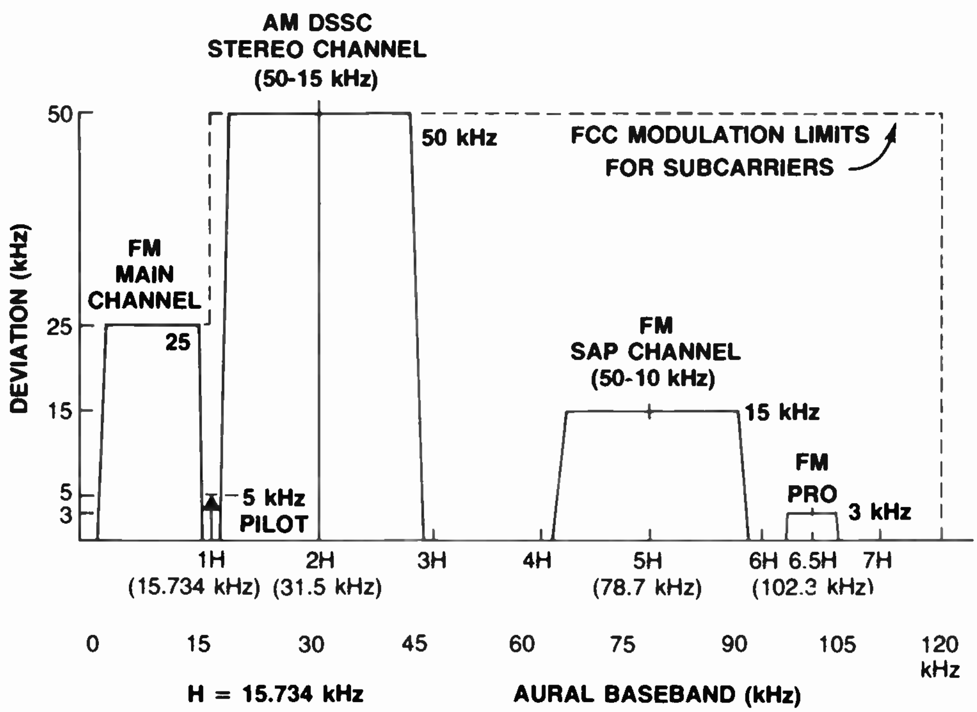


Figure 2. BTSC MTS baseband.

component is transmitted in the form of a difference, or left minus right (L – R), signal that is based on the current U.S. FM stereo broadcast model which is familiar to most broadcast engineers. Subcarriers for cues, data, telemetry, or other station related purposes may also be employed. Space for these professional use, nonpublic subcarriers is provided in the aural baseband. Fig. 3A shows how the BTSC signal is configured at the transmitter. Fig. 3B shows how BTSC is handled in a typical receiver.

Monophonic Channel

The monophonic, or sum channel (L + R), is maintained at its present place in the aural baseband and continues to have 25 kHz peak deviation and the standard 75 microsecond pre-emphasis. The addition of BTSC or MTS subcarriers will increase the overall deviation of the main carrier. Increasing the peak deviation of the aural carrier to 73 kHz (for BTSC) does not result in any degradation to the normal monophonic sound in existing receivers. The monophonic channel is maintained at 25 kHz deviation unlike FM radio broadcasting where the mono channel is reduced when transmitting stereo. This would result in lowering the signal-to-noise ratio of the aural channel and the loudness of the sum channel compared with stations not operating with the BTSC system. Increasing the overall deviation provides the room for the stereo, SAP, and PRO subcarriers to be added to the aural baseband successfully. The fidelity of the main channel monophonic audio is maintained as before. But, because of the closeness of the sidebands of the stereo subcarrier to the main channel, and to prevent crosstalk between channels, BTSC stereo encoder equipment incorporates sophisticated low-pass filters in the audio circuits to limit the frequency response of each audio channel to no more than 15 kHz. Nevertheless, the BTSC system is capable of providing audio quality essentially limited only by the quality of the main or monophonic channel.

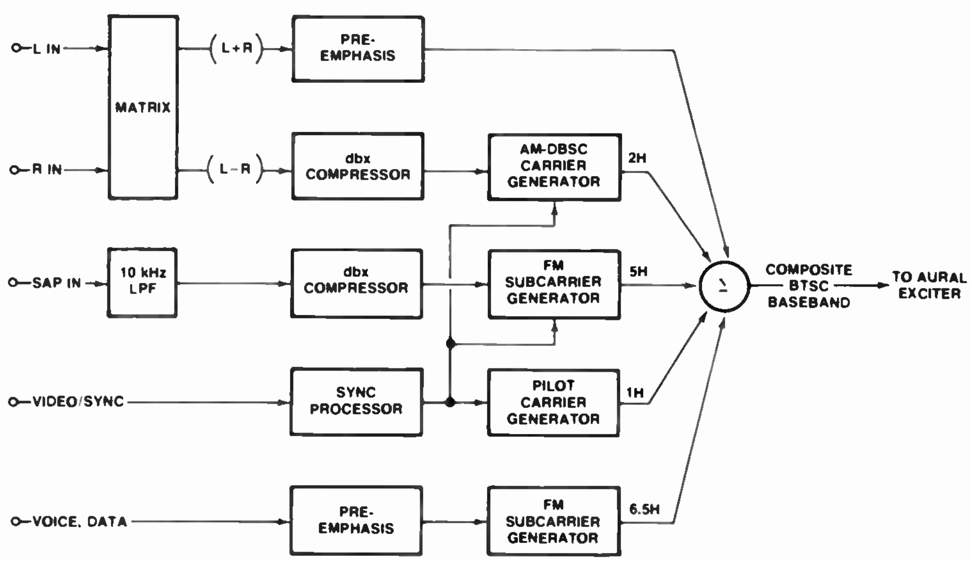


Figure 3A. Typical MTS transmitter encoder.

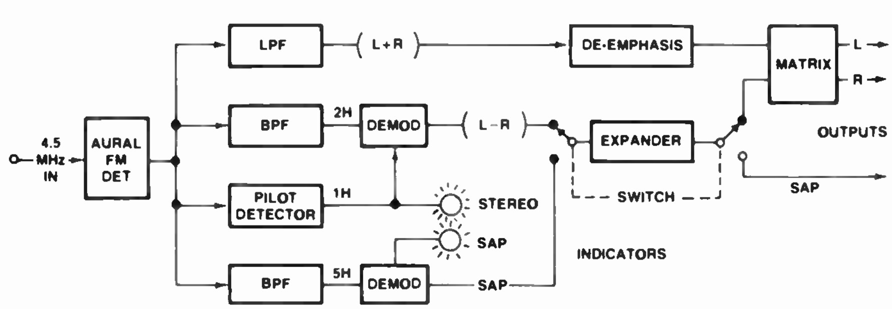


Figure 3B. Typical MTS receiver decoder.

Stereo Subcarrier

The stereophonic subchannel is an amplitude modulated, double sideband suppressed subcarrier. The subcarrier is suppressed to avoid wasting modulation capability. The energy in the sidebands, however, is permitted to deviate the main aural carrier by as much as 50 kHz. The subcarrier operates at twice the NTSC video horizontal scanning rate of 15734.3 Hz or 31468.6 Hz. Station video sync is fed to the encoder to maintain synchronism with the video. The subcarrier is modulated with the stereophonic difference signal, or left minus right (L − R). The maximum peak deviation of the main carrier (injection) by the stereo subcarrier is 50 kHz.

The injection level is related to the companding system algorithm and normally is quite high during stereo program material due to the compression. However, because of the nature of stereophonic sound (little correlation between the sum and difference channels) a form of interleaving takes place that in effect prevents the deviation of the main carrier from becoming the sum of the two components which would appear to produce a total of 75 kHz. Instead, the maximum rarely exceeds 50 kHz. This is because a left-only or right-only signal originating from the studio would produce only a 50% modulated signal in the mono channel even though it produces a high modulation level in the difference channel.

The difference channel of the BTSC system employs the dBx noise reduction circuit (the compression portion of the compander) which precisely controls the maximum modulation level of the difference signal. The preemphasis varies according to the level and frequency of the sound entering the compressor. While the combined deviation of main (L + R) and stereo (L − R) carriers rarely produces more than 50 kHz deviation of the main channel, most stereo encoder manufacturers incorporate some form of peak limiting device in these channels to prevent over deviation.

Note that the difference channel operates in a nonlinear manner by design. That is, the compander processes the audio signal differently at low levels than at high levels. When conducting tests on the difference and SAP channels, a 75 microsecond preemphasis network is substituted for the compressor. Consult the alignment instructions provided by the manufacturer for specific tests and setup requirements.

The selection of the center frequency of the stereophonic subcarrier was based on its relationship with the horizontal frequency (H) component of the video in both the transmitter and the receiver. Crosstalk caused by video signal circuits, power supply, and RF paths in the transmitter generally occur at multiples of H and, therefore, may be cancelled, or at least substantially reduced, in the receiver without significantly affecting overall stereo audio quality.

The first harmonic of H (15734 Hz), present to some extent in many transmitter and receiver audio circuits, is above the hearing range of most viewers and not normally audible. Receiver audio circuits may employ notch or low-pass filters to eliminate all but vestiges of H at the audio output. The second harmonic falls at 2H (31468.6 Hz) which is the frequency of the BTSC stereo subcarrier. By locking the stereo subcarrier frequency to the video sync, the harmonic and subcarrier remain in phase and no beat occurs.

Modulation of the 2H harmonic by 60 Hz (59.94 Hz) field rate sync components of the video signal can also occur in either the transmitter or receiver. The sharp rise time of the 60 Hz component causes the 60 Hz sound to have the characteristic buzz that is heard on some receivers. A related buzz-beat effect is reduced by the use of high-pass filters in the stereo subchannel at the receiver to remove unwanted audio components below about 100 Hz, which can beat with the 60 Hz buzz. While this has virtually no affect on the stereo audio quality, it substantially reduces the low frequency buzz and hum caused by the field rate modulation.

Placing the stereo subcarrier, with its 15 kHz sidebands, so close to the upper end of the main audio channel invites potential crosstalk problems. However, it is essential that the subcarrier frequency be an integer multiple of H as described above. Placing the stereo subcarrier at a higher or odd multiple of H is not desirable due to the much higher noise levels encountered in the upper portion of the aural baseband and the potential for problems with phase-locked loop audio detector circuits in the receiver. A higher subcarrier frequency would also have made the addition of the SAP channel more difficult to implement as well.

The build-up of noise in the aural baseband is shown in Fig. 4. This illustrates why it was important to place the stereo subcarrier as low as possible in the baseband as well as the need for the audio companding system. In FM modulation systems, noise increases at a rate of 6 dB for each doubling of the bandwidth. As a result, the noise the stereo subcarrier would add to the received dematrixed sound is about 23 dB. This is reduced by about 6 dB because the stereo subcarrier

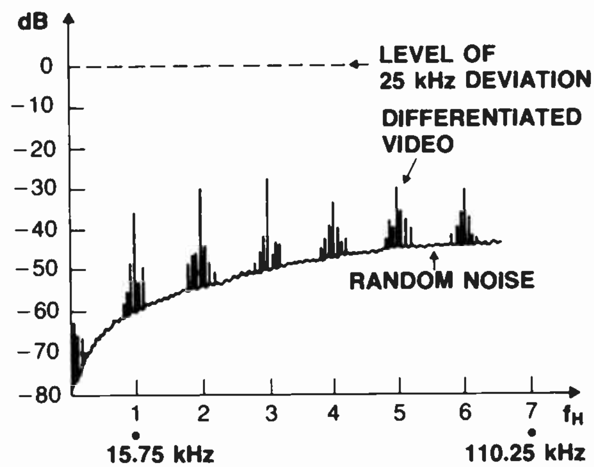


Figure 4. Spectrum of noise on the received aural baseband. (Courtesy of J.J. Gibson, RCA)

deviates the main carrier by 50 kHz compared to the main channel deviation of 25 kHz. The resulting degradation in signal-to-noise is only about 18 dB. Further, modulation components caused by the horizontal and vertical sync signals also add to the noise build-up. Both of these interference sources are easily overcome by the powerful 30 dB BTSC noise reduction system.

The use of an AM modulated, double-sideband, suppressed carrier subcarrier for the stereo channel also provides several other significant advantages over an FM subcarrier. AM subcarrier technology is well known and developed within the FM broadcast industry. Circuit performance refinements have been raised to near perfect levels. Moreover, AM offers (1) lower theoretical distortion levels than would be possible for FM, given the limited spectrum available (30 kHz) for transmission of the stereo subcarrier, and (2) provides an opportunity for further enhancements of the aural channel (by quadrature modulation or other compatible modulation schemes). The frequency response of the stereo subchannel is 15 kHz which properly compliments the main channel.

Pilot Carrier

The amplitude modulated stereo subcarrier is transmitted with the carrier suppressed. In order to provide a reference for the AM detector, an unmodulated pilot signal is transmitted at the TV horizontal line rate of 15.734 kHz which is one-half the subcarrier frequency. The pilot is used in the receiver to reinstate the carrier on the exact frequency and phase as the original. This is similar to the technique employed in broadcast FM stereo transmission systems.

For BTSC stereo, the pilot is locked to the video horizontal sync rate of 15.734 kHz. The pilot modulates the main aural carrier to a deviation of 5 kHz ±0.5 kHz. The stereo encoder requires a feed from the companion video signal in order to obtain its frequency reference. While it is possible that stereo sound receivers could use the horizontal sync information obtained from the receiver video circuits for the reference, there are receivers which do not process the video when decoding the stereo audio and therefore would not have a reference to synchronize the stereo detector.

The pilot also serves to signal that stereo is being broadcast by the station and is used to activate the stereo circuits and indicators in the receiver. Therefore, the pilot must be protected from extraneous signals near the pilot frequency in the aural baseband.

Pilot Carrier Protection

FCC Rule 73.682(c)(3) states that modulation components around the pilot frequency (±20 Hz) must be attenuated to 0.125 kHz deviation (46 dB below maximum modulation of 25 kHz). This is the case whether the station is operating with multichannel sound or monophonic sound.

A disadvantage with having the pilot at the horizontal rate is that there may be some residual horizontal sync modulation components and crosstalk in older transmitters or in program audio content that can cause false indications of stereo and malfunction in stereo decoders in BTSC receivers. Crosstalk of horizontal sync can occur in the exciter or amplifiers of older transmitters. Problems of this sort must be solved by broadcasters and manufacturers by adequately shielding or isolating the sync components of the video from the audio circuits in transmitting and receiving equipment. Modifications may be necessary to existing transmission equipment. The level of horizontal sync modulation and crosstalk components in the visual transmitter 4.5 MHz above the visual carrier can be controlled by installing low-pass video filters, adjusting transmitter visual RF circuits, or improving the 4.5 MHz rejection in the transmitter combiners or diplexers.

Low level sync modulation components with sufficient amplitude to cause false triggering of BTSC receivers can also occur by acoustic pick-up by microphones that are in close proximity of the mechanical vibrations caused by the horizontal scanning circuits in video monitors in studios, news sets, and in announce booths. Installation of identical high-quality low-pass or notch filters in the stereo audio path after the last switching point in the broadcast plant, will effectively eliminate the horizontal sync stray pick up in the studio from interfering with processors, STL facilities, and the BTSC encoder at the transmitter. The characteristics of such a low-pass filter are shown in Fig. 5. Filters should be installed ahead of devices with preemphasis so as not to compound the problem. For example, a standard 75 microsecond preemphasis network has the equivalent of about 15 dB gain at 15 kHz.

Operators of monophonic transmitters should also check their facilities for residual horizontal sync signals in either the aural transmitter or in the combining network at the output of the transmitters. The monophonic audio inputs to transmitters should incorporate a 15 kHz low pass filter.

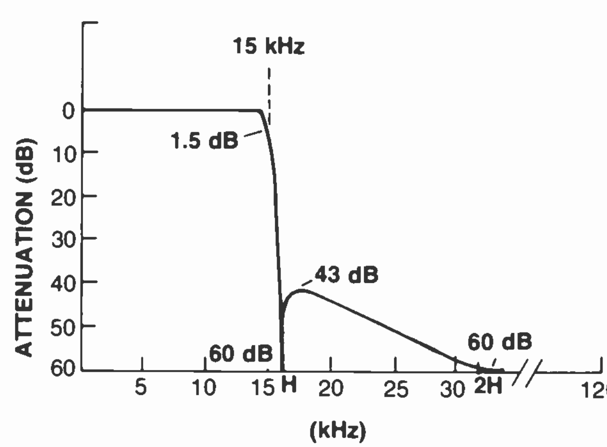


Figure 5. Audio input low-pass filter.

The Separate Audio Program Channel

Multichannel Television Sound transmission may include a separate audio program (SAP) channel in addition to stereo. This SAP channel may be program related or not, as the station chooses. The SAP channel is designed for reception by the public. If it is desired to use the SAP channel for nonpublic uses, the injection level of the subcarrier must be kept below the receiver threshold (about 5 kHz). If the audio information is not intended for the public, viewers with the SAP option switch may wonder about the material if it is not related to the picture in some manner.

The SAP channel subcarrier is located at exactly 5H (78.671 kHz) in the aural baseband. (See Fig. 2.) Although somewhat noisier at this location in the baseband than if it were placed at 4H, there is less chance for crosstalk with the main channel. The same BTSC noise reduction system on this channel provides an adequate signal-to-noise ratio. The audio frequency response is limited to a maximum of 12 kHz. The use of the companding system eliminates the need for separate preemphasis in the SAP channel.

The SAP subcarrier is frequency modulated by program material to a maximum deviation of 10 kHz. The subcarrier injection into the main carrier is limited to a peak deviation of 15 kHz. That is, the subcarrier modulates (deviates) the main carrier by 15 kHz. The subcarrier frequency is locked to 5H when not modulated.

The maximum injection level of 15 kHz of the main carrier by the SAP subcarrier is important. Receivers with SAP decoding equipment will detect the presence of the subcarrier and provide either automatic or manual switching and an indicator light. The receiver SAP decoder is designed to activate when the injection level reaches 8 kHz to 10 kHz. Below this level the receiver will ignore the presence of a subcarrier. To use this portion of the baseband for some purpose other than for program related audio, a subcarrier with an injection of 3 kHz to 5 kHz would probably be ignored by most SAP receivers. A nonpublic use of the SAP subcarrier cannot occur at the same time as regular SAP channel operation but could be used during stereo or monophonic operation.

The SAP channel uses the same noise reduction scheme as the stereo subcarrier. This permits a single expander decoder to be used in the receiver for either stereo or SAP. Most receivers will provide either stereo or SAP but not both at the same time. To make both available simultaneously a receiver would need two expanders. Alternatively, the SAP output could be provided without expansion, but the output would be highly compressed. The sound would be acceptable for most vocal material but would be objectionable for most music.

The Companding System

Companding (combines the functions of compression and expansion) is a term used frequently in the telecommunication industry to describe what broadcasters call noise reduction, a process well known to studio audio technicians. Its use is restricted to systems where there is control over both ends of the circuit. Examples include magnetic analog audio recorders, microwave and satellite circuits, and telephone lines used for program transmission. Simply stated, noise reduction is compressing the dynamic range at the sending end and expanding the signal an equal amount at the receiving end of the circuit. In the process of expansion the noise is reduced. Until the development of the BTSC MTS system, noise reduction techniques could not successfully be used for over-the-air transmission because a complimentary expander is required in the receiver. MTS required the use of expanders in the receiver from the beginning of the service. As a result, MTS is the first broadcast service to use companding as an integral part of the transmission system.

Some FM broadcast stations employ a limited range compression arrangement for an over-the-air noise reduction system. But because few receivers are equipped with the corresponding expander, the compression must be held to about only 10 dB in order to avoid producing objectional side effects and becoming objectional to a listener without a decoder.

Noise reduction was built-in to the BTSC MTS system from the beginning of the service and therefore a high performance system may be employed. Fig. 6 graphically depicts how a typical noise reduction system alters the dynamic range of the program material during transmission according to the content. Dynamic range is reduced from 70 dB to about 40 dB. A matching expander in the receiver restores the dynamic range.

The companding system employed by the BTSC MTS system was specifically designed to operate in the comparatively hostile environments presented by television transmitters, signal propagation, retransmission systems (CATV, MATV, translators) and the receiver itself. Thermal and impulse noise, multipath distortion, and buzz generated by transmitter and receiver all combine to present formidable obstacles for the companding system to overcome.

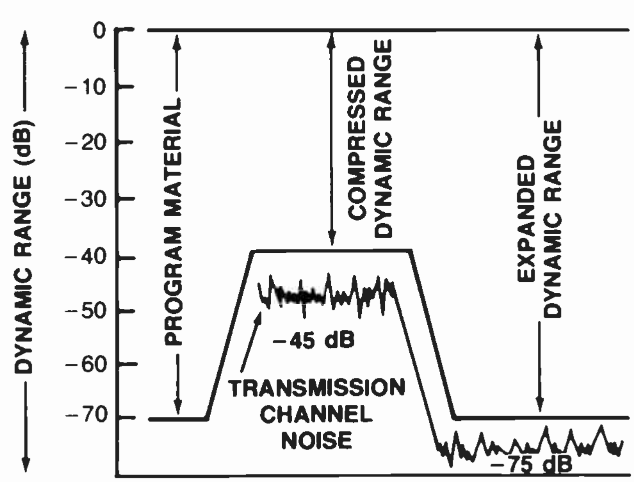


Figure 6. Typical noise reduction action.

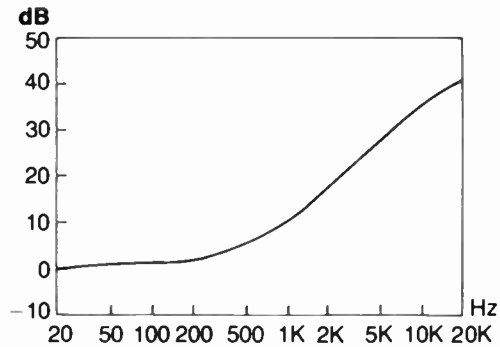


Figure 7A. Fixed preemphasis frequency response.

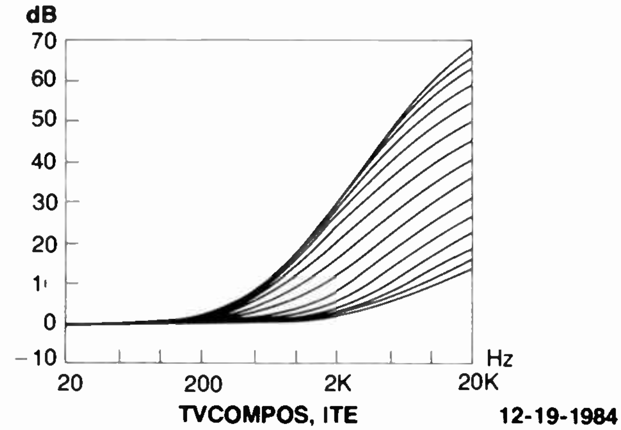


Figure 7C. Frequency-response range of spectral compressor and fixed preemphasis.

The companding system used in the BTSC system employs a combination of (1) fixed preemphasis, (2) spectral compression, and (3) amplitude compression. The fixed preemphasis combines the familiar 75 microsecond rising frequency response with a 390 microsecond network. The resulting curve is shown in Fig. 7A. By itself this extremely steep preemphasis curve would cause problems with high audio frequencies. To avoid this, a dynamic "spectral" compressor (variable preemphasis) circuit is employed to increase the gain of low level, high frequency material and reduce the gain of high level, high frequency audio. Only frequencies above about 200 Hz are affected. High frequency material is increased during transmission by as much as 30 dB at low levels but virtually not at all at high levels. Fig. 7B illustrates the dynamic characteristics of the spectral compressor. A complicated algorithm controls the fixed and variable preemphasis to produce the sharply rising frequency response shown in Fig. 7C.

The third feature of the companding system is the amplitude compressor which reduces the dynamic range, measured in dB, of the input signal by a factor of 2:1 for low frequencies, and 3:1 for high frequencies. In other words, a 40 dB dynamic range audio signal is reduced to 15 dB to 20 dB or less for transmission as shown in Fig. 8. In addition to dynamic range compression, the maximum input level applied to the stereo subcarrier is reduced to just below the maximum modulation of the subcarrier. This feature provides some headroom to accommodate instantaneous peaks without resorting to clipping. High level transients are clipped to avoid severe distortion during reconstruction of the compressed signal in the receiver expander.

The equivalent 75 microsecond mode in the companding system uses a precision 75 microsecond network in place of the dbx compressor when separation, crosstalk, and noise tests are conducted. In this mode the exciter contains only linear elements which allows routine adjustments on the exciter and transmitter and measurements of the main, stereo, and the SAP channels to be performed very much as conventional FM broadcast stereo exciters.

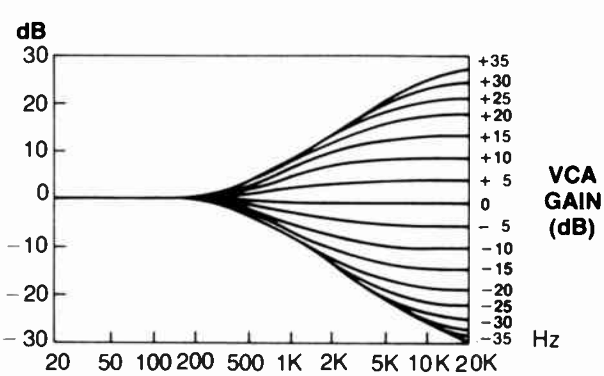


Figure 7B. Spectral compressor, frequency-response range.

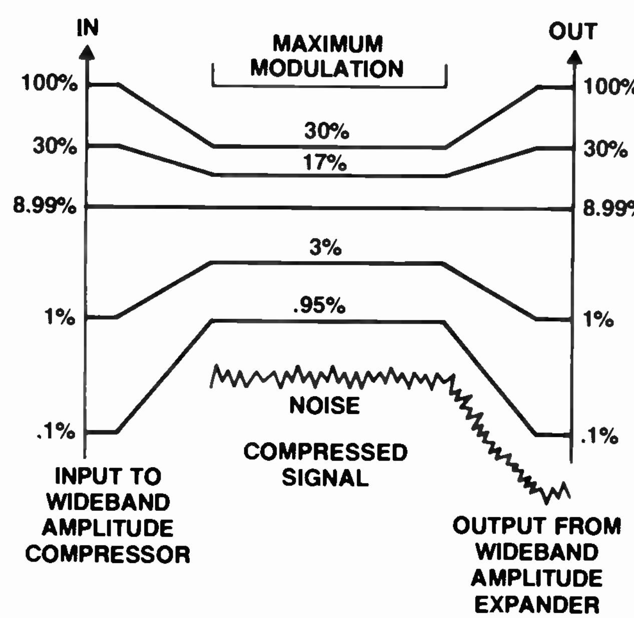


Figure 8. Companding of (L-R) and SAP. (Courtesy of dBx, Inc.)

The use of companders and noise reduction systems is gaining wide acceptance throughout the broadcast industry. It is not unusual for a given program to undergo noise reduction processing five to six times by the time it reaches the viewer. The mistracking of a single companding system in the chain may result in substantial errors in the resulting sound which may be more objectionable than the problems the compander was designed to solve. In order to prevent audible artifacts from occurring, only the highest quality processing and noise reduction systems should be employed, and these must be correctly operated and properly maintained.

Professional Use Channel

The PRO channel is not designed to be received by the public and may be used by the broadcaster for many purposes such as voice cues, signalling, and data transmission. The channel is located at 6.5H (102.28 kHz) ±500 Hz in the BTSC baseband as shown in Fig. 2. The maximum peak deviation of the subcarrier is 3 kHz. The peak deviation (injection) of the main carrier by the PRO subcarrier is 3 kHz. When used for voice or audio transmission, a modulation limiter, 3.4 kHz bandlimiting filter, and a 150 μsec preemphasis network are normally employed. For data transmission, the limiter, preemphasis, and filter are bypassed.

Other Subcarriers

In addition to providing for stereo and SAP subcarriers, the FCC allows television stations to use any other scheme of subcarriers on the aural carrier as long as they do not interfere with the normal operation of monophonic or BTSC television receivers. When not transmitting stereo or SAP, broadcasters may elect to use the aural baseband for subcarriers for nonprogram related purposes. Virtually any kind of subcarrier for program material or data may be transmitted using AM, FM, or other analog or digital modulation methods. Subcarriers may be placed in the aural baseband between 16 kHz and 120 kHz but must avoid the BTSC pilot and SAP frequencies. The maximum deviation of the main carrier by these subcarriers is limited to a total of 50 kHz. Any number of subcarriers may be used. The total modulation of the aural carrier including main channel sound and non-BTSC subcarriers may not exceed 75 kHz.

Stations should avoid placing nonpublic subcarriers within ±10 kHz of the SAP frequency unless the injection level of the subcarrier is kept below 5 kHz deviation. The use of a single subcarrier using the full 50 kHz deviation should be avoided because of the potential of crosstalk into the main channel in some receivers unless extensive compatibility testing is conducted.

Modulation Summary

The FCC permits the BTSC specified multichannel sound signal to modulate the main aural carrier a total peak deviation of 73 kHz. This is produced by the combined main and stereo subcarrier deviation of 50 kHz, the pilot subcarrier deviation of 5 kHz, the SAP subcarrier deviation of 15 kHz, and the PRO subcarrier of 3 kHz for a total of 73 kHz. (See OET-60A section D(b)(6)). Fig. 9 lists the MTS modulation components that should be monitored on a regular basis.

SIGNAL	MONITORED PARAMETER
MAIN CHANNEL (L+R)	25 kHz PEAK DEVIATION
STEREO CHANNEL (L+R)	50 kHz PEAK DEVIATION
MAIN AND STEREO	50 kHz PEAK DEVIATION
PILOT CARRIER INJECTION	5 kHz PEAK DEVIATION
SAP MODULATION	10 kHz PEAK DEVIATION
SAP INJECTION	15 kHz PEAK DEVIATION
OTHER SUBCARRIER	3 kHz PEAK DEVIATION
TOTAL COMPOSITE	73 kHz PEAK DEVIATION

Figure 9. Typical modulation monitor functions.

Other subcarrier arrangements may modulate the main carrier to a peak deviation of 50 kHz. With the main monaural channel at 25 kHz and non-BTSC subcarriers at 50 kHz, the total may not exceed 75 kHz deviation (FCC Rule 73.68(c)(9)).

TRANSMISSION OF BTSC

Transmission Requirements

Many of the technical rules in Part 73 of the FCC Rules and Regulations that described "quality" characteristics of the transmitted signal were deleted by the FCC in 1984. Those rules, requiring stations to meet certain distortion, frequency response, and signal-to-noise performance levels, were eliminated in favor of marketplace pressures on stations to maintain high quality levels. Therefore, there are no FCC Rules which require minimum audio performance levels in the AM, FM, or TV broadcast service. However, rules which act to control interference have been maintained.

The FCC Rules which permit stations to transmit MTS also do not specify performance objectives other than those which are directly related to interference to other services or to the monophonic sound channel. However, because the MTS is a relatively new service, several quality specifications are included in Technical Bulletin OET-60A, published by the FCC Office of Engineering and Technology. Television stations should have a copy of OET-60A. The Electronic Industries Association (EIA) in 1985 produced a set of recommended practices (EIA Television Systems Bulletin #5) for BTSC MTS system written especially for receiver and transmitter manufacturers. The EIA also published the Compandor Complexity Analyses that supplements Bulletin #5.

Transmitter and BTSC encoder manufacturers also publish booklets and manuals describing potential problems encountered in preparing transmission systems of BTSC audio along with methods for measurement and problem correction. Some of these are listed in the references at the end of this chapter.

Because the transmission of stereo sound for television is technically similar to that of FM broadcasting, a review of the chapter on FM broadcasting and stereophonic transmission elsewhere in this handbook provides additional background information. Portions of the BTSC MTS system description are also contained in the Television Transmitter chapter.

The MTS performance objectives achievable under normal operating conditions are shown in Fig. 1. Most of the performance objectives indicate the potential for high audio quality. However, the stereo separation figure may be of some concern to those familiar with the much higher performance of the FM stereo system. More than 30 dB separation is easily achieved in practice above which little subjective improvement is obtained. The nonlinear circuits in the companding system and the critical adjustments which must be made in the BTSC encoding system may produce somewhat less than the 40 dB or more separation easily achieved in broadcast FM stereo operation. The tracking gain between the mono and stereo channels is one of several important operating characteristics that determine the relative separation between channels as shown in Fig. 10A. While it may be possible to balance the gain to a tenth of a dB between the sum and difference channels in FM radio stereo transmission, typical balance is about 0.5 to 0.8 dB in the BTSC system. This is because the noise reduction system compresses the stereo channel by as much as 20 dB, which magnifies the difference when received, decoded, and expanded in the TV receiver. Circuit stability improvements may permit as much as 40 dB in stereo separation in the future.

Phase shift between audio channels will also reduce stereo separation. Only a few degrees difference reduces separation to 30 dB as shown in Fig. 10B. When combined with channel gain differences, separation will be further reduced.

Incorrect stereo channel carrier phase error and the pilot phase error also contributes to reduced stereo separation as shown in Fig. 10C.

A major feature of the BTSC MTS system is that there is no stereo noise penalty as there is in FM stereo. Without the companding system the stereo subchannel is subject to the increased noise in the channel baseband by as much as 23 dB. In FM radio the main or mono channel must be reduced by eight to ten percent to permit the addition of the pilot subcarrier. With TV MTS, the main aural carrier deviation is increased when adding the pilot and stereo subcarriers. There is no reduction in deviation of the mono signal during BTSC stereo transmission as there is when adding stereo in broadcast FM operation.

The crosstalk performance values are the single most important technical parameters with which the broadcaster must be concerned for the BTSC system. Most crosstalk is caused by systems external to the

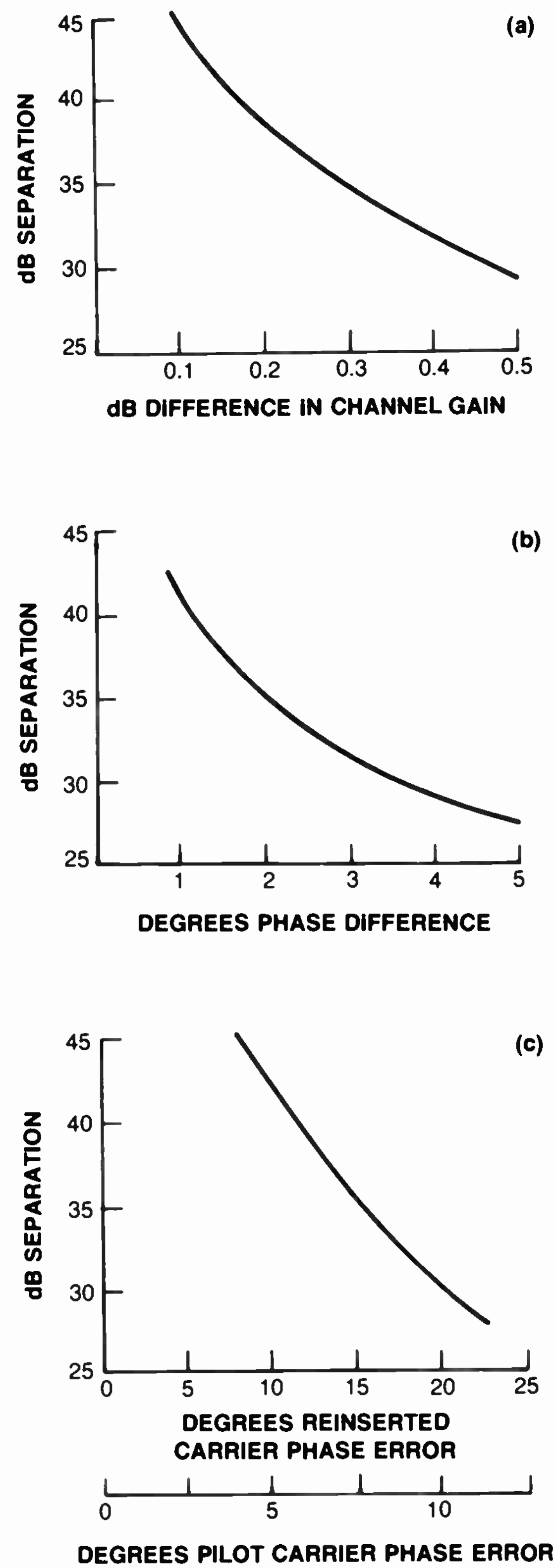


Figure 10. Stereo separation versus (a) amplitude difference between main and stereo channels, (b) phase difference between main and stereo channels, and (c) error in phase of pilot or reinserted carrier.

stereo transmission equipment itself. Nonlinearities in the TV aural exciter and power amplifier stages, combining networks, antenna switching components, and the antenna itself, constitute the main areas in the signal path where crosstalk can be generated. In general, crosstalk into the stereo signals from the SAP will be less than crosstalk from the main channel into the SAP signal. As long as crosstalk levels do not exceed the levels shown in Fig. 11, most crosstalk will be concealed by the companding system.

Other sources of crosstalk are (1) multipath distortion due to poor transmission line and antenna voltage standing wave ratio (VSWR), (2) multipath (ghosts) during the propagation of the transmitted wave, and (3) improper matching in the receiving antenna system. In general, the companding system can conceal noise and crosstalk in the stereo and SAP channels by as much as 30 dB.

Visual Over-Modulation

The most important way to reduce buzz on television receivers is for the broadcaster to insure that the video modulation level is never permitted to exceed 87.5%, the maximum permitted by the FCC. Varying video levels during program production and transients from character and graphics generators can easily cause the visual modulation level to approach or equal cut-off. At this point the aural detector in intercarrier type television receivers loses reference (the aural carrier), and the result is a strong buzz in the received audio. Careful adjustment to video processing equipment will eliminate this problem as a source of buzz. In some cases it may be necessary to add high quality low-pass filters to the video outputs of character and graphics generators to limit the high frequency content.

Incidental Carrier Phase Modulation

The control of incidental carrier phase modulation (ICPM) in the visual transmitter is the second most important performance characteristic of the transmitter that can affect the quality of multichannel sound.

ICPM in visual transmitters must be adjusted and maintained to less than 3° during the luminance portion of the video and to less than 5° during sync in order to keep buzz to acceptable levels as shown in Fig. 12.

Excessive ICPM in the transmitter will cause an audible and annoying buzz in intercarrier sound television receivers. The buzz will be modulated by the action of the expander which will increase the level of annoyance.

CROSSTALK	OBJECTIVES
MAIN INTO STEREO	>40 dB
STEREO INTO MAIN	>40 dB
MAIN INTO SAP	>50 dB
SAP INTO MAIN	>60 dB
OTHER SUB INTO MAIN	>60 dB
OTHER SUB INTO STEREO	>60 dB

Figure 11. Crosstalk objectives.

DEGREES ICPM →	1°	2°	3°	4°	5°
BASEBAND BUZZ LEVEL IN dB BELOW 25 kHz DEV. ("WORST CASE")	−50	−44	−40	−38	−36
STEREO SBR dB	56	50	46	44	42
SAP BUZZBEAT THD dB	−49	−43	−39	−37	−35

NYQUIST SLOPE EQV. ICPM = { 2.4° AT (L-R) SUBC. / 6° AT SAP SUBC.

Figure 12. Intercarrier buzz and buzzbeat levels. (Courtesy of J.J. Gibson, RCA)

Intercarrier type receivers, typical of most receivers, are used for multichannel sound reception because of their high immunity to common-mode phase modulation that can occur in nonlinear active devices or distribution and conversion systems that are found in the path of the television signal before it reaches the receiver. Common-mode phase modulation often occurs in TV translators, master antenna downconverters, and cable system headends and set-top converter equipment.

In intercarrier type receivers, the aural and visual carriers beat together in the video detector to produce the 4.5 MHz aural intermediate frequency (IF) signal. Because common-mode phase modulation affects both carriers (hence the term common-mode), the phase modulation is factored out of the 4.5 MHz carrier. Therefore, phase modulation that occurs on the visual carrier in the transmitter, which is independent of the aural carrier, will produce undesired phase modulation of the 4.5 MHz aural carrier in the receiver and the result will be the familiar buzz.

Phase modulation in the local oscillator of an up or down converter can be caused by the power supply or mechanical vibration. It can also occur if the phase of the composite signal is changed by a nonlinear amplifier. Phase modulation is common-mode when both aural and visual carriers are affected equally. This would be the case when the aural signal is carried as a subcarrier on the visual carrier rather than treated as separate signals as is the case in broadcast transmitters.

Intercarrier receivers carry the aural signal to the detector as a subcarrier. Translators and most CATV systems carry the aural through headend processors as a subcarrier. Therefore, if phase modulation is introduced in these systems, it is common-mode and will be rejected by the receiver. Intercarrier sound receivers cannot reject phase modulation that is introduced independently into the visual or aural carriers at the transmitter.

Phase modulation in the visual transmitter can be caused by hum in the carrier frequency or up-converter oscillators, and by nonlinearities in the visual modulator, IF, and power amplifier stages. Phase correction circuits in the modulator can compensate for most ICPM problems. UHF transmitters which use klystron anode pulsers to achieve higher operating efficiencies may have phase modulation introduced by the pulser modulating the visual amplifier power supply at the video field rate (59.94 Hz) and the line rate (15.734 kHz). The high frequency component is not heard by most viewers but the low frequency component is heard by all as a raspy buzz. An anode pulser on the visual amplifier can cause crosstalk through the common power supply into the aural amplifier. The resulting 15.734 kHz modulation, if high enough, can cause false indications of the stereo indicator on MTS television receivers.

Measuring and Correcting ICPM

ICPM is manifested as differential phase in the demodulated video. Measurement instrumentation requires a TV demodulator with both envelope and quadrature video detectors and a waveform monitor with the ICPM graticules. After measuring ICPM or listening for buzz on a stereo television receiver to determine if ICPM is present, phase adjustments in the modulator or upconverter and retuning the RF stages will result in substantial improvements. ICPM adjustments, if provided on the transmitter exciter, will also affect video differential phase. Once ICPM is corrected, the video correction circuits can be adjusted to compensate for differential phase.

ICPM will be more noticeable on television receivers equipped for multichannel sound. ICPM will affect subcarriers to a greater extent than the main channel. In addition, improved audio performance in the new multichannel receivers makes existing buzz more noticeable. Therefore, stations planning to implement BTSC multichannel sound should check their transmitters for ICPM with an ear toward improving the quality of the mono sound on mono receivers as well as stereo on new MTS receivers.

Synchronous Amplitude Modulation

Synchronous amplitude modulation (SAM), occurs when the frequency swing of the aural carrier is greater than the flattest portion of the tuned circuits through which the carrier passes. A reduction of the carrier occurs where the passband begins to roll-off as shown in Fig. 13. The result of effect is to impart amplitude changes on the RF carrier during FM modulation. The SAM is in sync with the FM, hence the name synchronous AM. SAM is made worse if the RF carrier is not centered within the passband of the tuned circuits in the aural transmitters, combiners, and diplexers.

SAM manifests itself as intermodulation distortion which in turn causes an increase in crosstalk between channels and decreased stereo separation. The SAP and PRO channels are more susceptible to SAM than the main channel especially during high levels of main channel modulation.

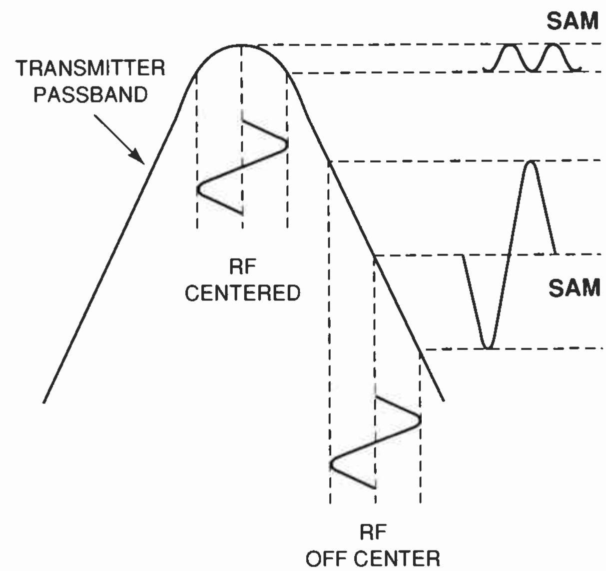


Figure 13. Synchronous Amplitude Modulation (SAM).

Compatibility With Monophonic Receivers

During the development of the BTSC multichannel sound system, considerable effort was spent evaluating the effect of a fully loaded BTSC system on existing monophonic television receivers. Extensive compatibility testing of a wide range of receiver makes and models, under adverse transmission conditions, revealed virtually no significant degradation to either the sound or picture of the desired or adjacent channel caused by the BTSC system, as noted by trained observers.

BTSC sound transmission will not introduce additional buzz or distortion in existing monophonic receivers. On the other hand, nonlinear circuits and the resulting AM-to-PM conversion will nearly always produce buzz which will be more objectional in stereo than in mono. Most of the buzz heard in mono or stereo receivers is usually created in the transmitter or the cable TV system in the form of ICPM or video over-modulation.

Modern monophonic television receivers should not experience significant degradation when receiving a TV MTS signal. In fact, monophonic receivers may actually exhibit a noticeable improvement in performance when tuned to BTSC stereo transmissions due to the improved performance of the stereo exciter, broadband aural RF systems, and lower ICPM in visual transmitter components. Program producers now pay increased attention to stereophonic sound production and processing. (See the chapter on television program production.)

These desirable side benefits easily offset the slight reduction in monophonic modulation level which occurs when transmitting a mostly left or right channel sound. Because the main channel is the sum of left plus right, a right- or left-only signal will produce up to 6 dB less modulation than the sum signal. Producers, however, seldom employ left- or right-only signals except for sound effects and music. As a result, any reduction in modulation will be something between 0 and 6 dB. It could be argued that these several dB reduce the apparent loudness of some program material on a stereo station when compared to a monophonic station. The problem will disappear as more television stations install stereo transmission facilities.

Compatibility With CATV Systems

Most broadcast television signals are carried by Community Antenna Television Systems (CATV) or cable TV. In 1990, more than 50 million subscribers, or over half the homes with TV, have at least one receiver connected to a CATV system. It is important, therefore, that broadcasters be concerned with the quality of the audio as well as the video signal as carried on cable television systems.

The BTSC MTS system was specifically designed to pass through properly adjusted MATV and CATV system headend and distribution equipment and subscriber terminals. The concern that the lower sideband of the upper adjacent visual carrier might affect the wider BTSC MTS modulated aural carrier of the lower channel has not been confirmed. In general, those CATV systems which process off-air signals in an RF mode (that is, where the aural carrier is not demodulated) will have little difficulty in successfully processing and passing an acceptable BTSC MTS signal through the system. The same is true of the RF conversion type set-top converters used by CATV systems.

On the other hand, cable systems with demodulating type headend or subscriber terminal equipment may not be able to pass MTS signals without substantial degradation to the stereo signal due to the use of low pass filters in the composite audio portions of demodulator facilities where stereo and SAP subcarriers are filtered out or severely distorted. Newer cable headend equipment, however, process BTSC stereo with very little degradation by carrying the BTSC composite signal through at IF (such as 4.5 MHz).

The baseband type set-top converter poses the greatest threat to BTSC stereo if the audio is demodulated within the device. This is often done to permit the audio level, as well as the channel, to be controlled by a remote control device. BTSC stereo will not reliably pass through such devices.

Most cable television systems maintain the visual-to-aural carrier ratio at 15 dB to 17 dB compared with the normal 7 dB to 10 dB for broadcast. This will add 5 dB to 10 dB more noise and buzz to the BTSC signal in the cable system. Therefore, it is important that the broadcaster limit degradation to the broadcast BTSC signal in order to insure that the carriage on cable systems is satisfactory.

The BTSC MTS signal should pass, without significant degradation, through the amplitude modulated link (AML) microwave systems used by many cable companies for wide area distribution of multiple channels.

Cable companies also routinely convert discrete stereo audio received from satellite program sources into BTSC stereo for distribution to subscribers.

Television broadcasters will find it to their advantage to determine how their stereo sound is processed in CATV systems on which their station is carried and conduct routine tests. Most television stations routinely check the quality of the visual signal on CATV systems. The cable system introduces several additional active elements between the broadcast transmitter and home receiver. Basic audio tests for noise, stereo separation, pilot level, SAP channel quality, and crosstalk between SAP and main channels should, at the very least, be conducted semi-annually at the output of a subscriber terminal on the largest cable systems carrying the station. This procedure will help insure that the cable company is properly handling (quality and loudness) the broadcast signal as well as the satellite channels. A BTSC receiver of known characteristics and standard audio test instrumentation is needed to conduct the tests on the BTSC MTS signal. Such tests can be conducted with ordinary signal generators and audio level meters.

Compatibility With Television Translators

Many television broadcast stations are carried to remote communities by the use of translators which may range in size from 1 Watt to 1 kW.

Television translators using heterodyne or direct RF conversion techniques should not cause significant degradation to BTSC MTS signals if the translator is properly maintained. Some translators, however, receive little attention unless they fail completely or viewers report poor performance. In general, unless the picture and sound are severely degraded, or visible or audible interference is present due to improper operation, then the BTSC MTS signal will pass through television translators with little degradation.

Remodulating type translators may require modifications or adjustments to accommodate the BTSC MTS signal. Improvements include: (1) a wider bandwidth for the composite audio signal, to accept the subcarriers in the BTSC signal, (2) good visual transmitter oscillator stability to reduce ICPM, and (3) correct RF tuning to reduce SAM. Newer remodulating translators process the audio as a 4.5 MHz subcarrier similar to the process used in cable TV systems and have improved RF bandwidth characteristics with the result that little degradation occurs to the BTSC signal.

Modulation Monitoring

Monitoring the modulation and conducting routine tests of the BTSC system can be accomplished by either a spectrum analyzer plus the standard audio instrumentation, or a BTSC modulation monitor which provides continuous modulation information on all critical operating parameters. Stations must have the ability to monitor main channel, stereo composite, pilot, SAP (and PRO channel, if used), and overall modulation levels during normal operation. Virtually all commercial modulation monitors provide the means to make accurate measurements of all operating parameters and technical tests. Fig. 9 lists the regularly monitored MTS transmission parameters.

OPERATING PRACTICES

A compilation of transmission impairments that can cause stereo audio impairments is shown in Fig. 14.

Stereo synthesizers, used by some stations to provide a sense of stereo during monophonic programming, no longer seem necessary as more stations convert to stereo and more programming is produced in stereo. Under some conditions synthesizers may actually cause loss of localization of stereo sound sources because some program material accurately generates center channel information. Further, the effect of surround sound and other special aural techniques may be lost if the synthesizer is employed.

Synthesizers should not be connected in-line with the transmitter input or any point downstream in the audio chain where a mixture of stereo and mono material will be encountered. Synthesizers may be employed in production facilities, however.

ICPM, visual carrier overmodulation, severe group delay, and phase distortion in the visual amplifier will cause buzz in the received audio and loss of stereo separation. These two characteristics are the most sensitive to problems with the stereo generator and RF systems.

Adjustment of the BTSC stereo encoder and dbx compressor circuits (used in both stereo and SAP channels) is not recommended without a thorough understanding of the overall system and appropriate test equipment. The service manuals for the stereo exciter contain the specific procedures for routine maintenance.

Ordinary harmonic distortion in the BTSC stereo baseband (50 Hz to 120 kHz) will result in crosstalk between channels. In order to keep crosstalk within limits the distortion must be maintained to less than 0.5%.

FCC Rules require all stations to provide protection to the pilot frequency by limiting modulation in the vicinity of the pilot 15,734 Hz (±20 Hz) to no more than 46 dB below 50 kHz deviation. The BTSC pilot frequency is immediately adjacent to the upper end of the frequency response for the main channel. It is essential that monophonic as well as stereo stations insure the pilot is protected to avoid causing improper operation of stereo receivers and false stereo indications. While all MTS stereo exciters incorporate low-pass filters with a sharp cut-off above 14 kHz, monophonic stations may need to install such a filter if adequate protection is not otherwise provided. (See Fig. 5.) Measuring the pilot protection can be performed with a BTSC modulation monitor or a spectrum analyzer.

Many stations prefer to locate the stereo encoding equipment at the studio and transmit a composite MTS signal to the transmitter over a studio transmitter link (STL). The STL in effect becomes another section of the RF path where degradation to the BTSC signal can occur. Stereo separation, a good indicator of overall performance, will suffer if the baseband BTSC signal is degraded. In order to maintain at least 30 dB separation at the output of the transmitter, the STL should have at least 40 dB separation to compensate for degradation in the transmitter and elsewhere in the RF path. This means that the phase and amplitude frequency response across the 53 kHz baseband of the STL must be maintained to less than 0.1 dB combined with no more than one degree phase error as shown in Fig. 15.

EFFECT / CAUSE	BUZZ	SAP BUZZ BEAT	NOISE & NOISE PUMPING	REDUCED STEREO SEP	DIST.	X-TALK
NON-LINEAR PIX	✓					
NON-LINEAR SOUND					✓	✓
MISMATCHES						✓
DIPLEX FILTER					✓	✓
ICPM	✓	✓				
MULTIPATH	✓	✓	✓	✓	✓	✓
WEAK SIGNALS			✓	✓		
CO-CHANNEL					✓	✓
MISTRACKING				✓		
REC. FILTERS	✓	✓		✓	✓	✓
FM DETECT	✓				✓	✓

Figure 14. Effects of transmission impairments. (Courtesy of J.J. Gibson, RCA)

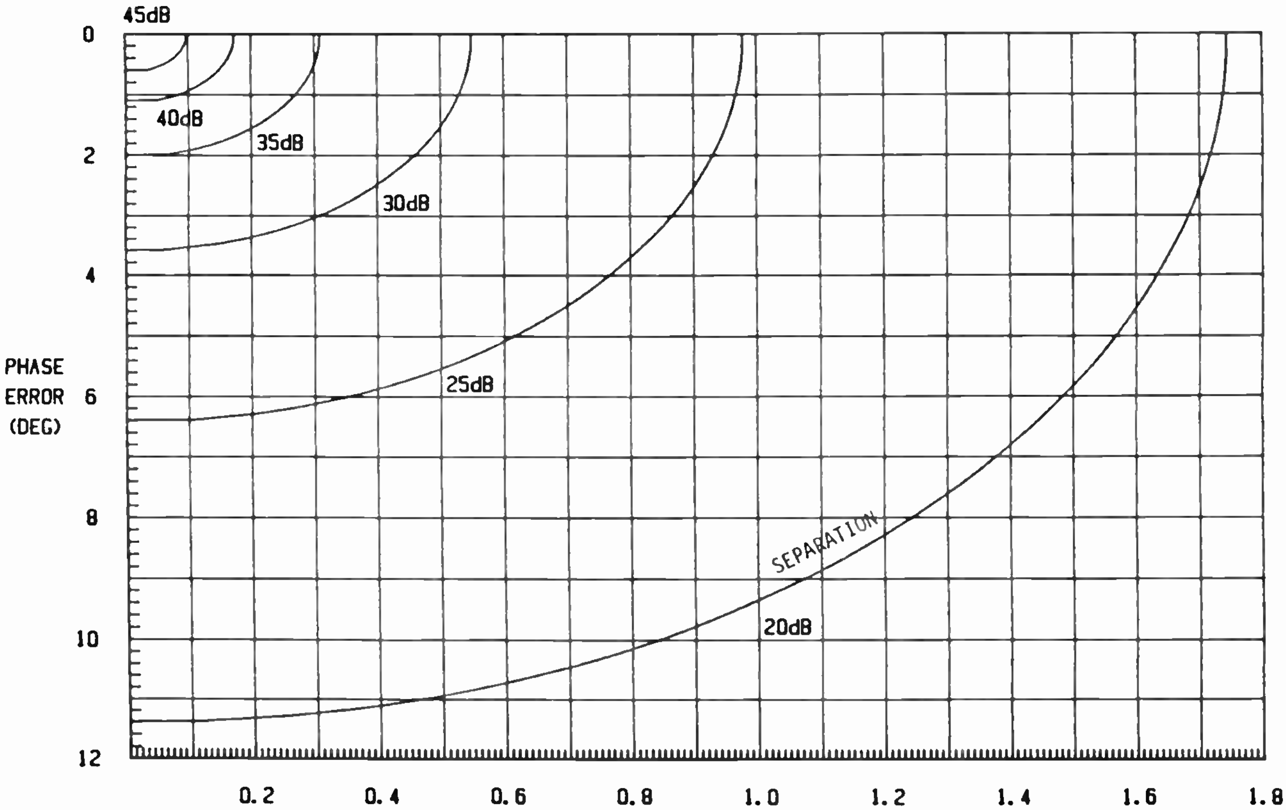


Figure 15. BTSC separation vs. combined amplitude and phase errors in the composite baseband. (Courtesy Broadcast Electronics, Inc.)

For good stereo operation it is necessary that the left and right channels remain well separated ahead of, within, and at the end of the transmission chain. The amplitude and phase of the left and right signals must be nearly identical and the phase of the pilot carrier must be maintained. Channel separation as a function of these three factors is given in the equation in Fig. 16.

Separation in dB =

$$20 \log_{10} \left[\frac{(\cos\theta + \frac{S}{M}\cos\phi)^2 + (\sin\theta)^2}{(\cos\theta - \frac{S}{M}\cos\phi)^2 + (\sin\theta)^2} \right]^{1/2}$$

where:

- M = the gain of the main (L + R) path and
- S = the gain of the stereo (L − R) path
- ϕ = the phase error of the reinserted 38 kHz subcarrier that is twice the phase error of the 19 kHz pilot carrier
- θ = the difference in phase between the (L + R) and (L − R) paths.

Figure 16. Formula for calculating phase and amplitude changes.

Total harmonic distortion (THD) and intermodulation distortion on the composite STL must not exceed 0.25% to keep crosstalk between channels to acceptable levels. If the SAP and PRO channels are also included in the composite baseband, the baseband performance values cited above must extend to 106 kHz.

Discrete stereo audio channels on the STL remain popular because most stereo encoders are located at the transmitter. Variations in the performance of discrete audio channels will produce less degradation to the integrity of the stereo signal than will the same performance characteristics of the composite channel.

The BTSC stereo system has not yet become attractive for use on satellite transmission circuits probably for the reasons cited above for composite STL systems.

Off-air monitoring is essential in the maintenance of good stereo and monophonic sound. Reversed channel polarity (sometimes referred to as phase reversal) on one channel may cause the stereo sound to be slightly degraded on stereo receivers, but severe degradation will result to the sound on monophonic receivers. By using a consumer receiver for audio monitoring, buzz caused by over-modulation or ICPM can be detected at the station. Many stations employ both stereo and monophonic monitoring of the transmitted audio in order to insure that stereo imaging and channel polarity are correct.

The implementation of multichannel sound in a television facility which has not dealt with stereo or multiple audio channels in the past will prove to be a challenging experience. Virtually all audio equipment in the station will need to be reviewed, reworked, rewired, or replaced when adding new audio channels. Videotape machines, routing switchers, master control switching and monitoring, studio audio facilities, and the STL will require close examination to determine how to implement extra sound channels. Video tape equipment with two sound channels may require new test procedures to determine if the phase relationship between channels is correct.

Special attention must be paid to audio processing. There will be a tendency towards maximizing the loudness of the signal rather than to moderate processing which will result in more acceptable stereo sound quality.

Multichannel sound equipped television receivers have substantially improved audio systems which will reveal audio faults now concealed by most television receivers. As a result, special attention must be devoted to the audio transmission facilities of all television stations, whether transmitting monophonic or multichannel sound.

BIBLIOGRAPHY

1. "Multichannel Television Sound Transmission and Audio Processing for the BTSC system, OET-60A," Federal Communications Commission, 1986.
2. "Multichannel Television Sound BTSC System Recommended Practices, Television Bulletin #5," Electronic Industries Association, 1985.
3. "Testing Television Transmission Systems For Multichannel Sound Compatibility, Geoffrey N. Mendenhall, Broadcast Electronics, Inc.
4. "Demystifying TV MTS Specifications," James R. Carpenter, Broadcast Electronics, Inc.
5. "BTSC Stereo: TV Aural Proof of Performance Guide," TFT, Inc.
6. "Understanding MTS Equivalent Mode, and Is Your Transmitter Stereo-ready," Eric Small, Modulation Sciences, Inc.
7. "Evaluating BTSC Stereo Generators and Audio Processors: Some Suggestions," Robert Orban, Orban Associates, Inc.
8. AM Stereo and TV Stereo—New Sound Dimensions, Stan Prentiss, TAB Books.

3.7 Transmitter Control Systems

Harold Hallikainen
Hallikainen & Friends, Inc., San Luis Obispo, California

INTRODUCTION

The Federal Communications Commission (FCC) requires broadcast transmitters to have various meters and controls. See FCC Rules Part 73 for relevant regulations: AM, 73.58; FM, 73.258; NCE-FM, 73.558; TV, 73.688. Other references to regulations will be shown in brackets [ ]. The required meters include those necessary to determine the transmitter final amplifier input power (final amplifier voltage and current) and the output power (transmission line meters for FM and TV, base or common point meters for AM). The required controls include those necessary to insure that the transmitter is operating within the terms of the station license. These controls have traditionally included a carrier on/off control, a day/night control for AM stations (power and/or pattern change) and a "power trim" control. In addition, directional AM stations and television stations must have additional monitors. Directional AM stations are required to have an FCC authorized antenna monitor [§73.69]. This monitor indicates the current ratios and phases for each tower, relative to a reference tower. Television stations are required to have a monitor capable of determining whether the transmitted visual signal meets FCC specifications [§73.691]. If the operator can view the transmitter and associated monitoring equipment from the normal duty position, then no further remote metering or control is legally required.[1]

If the operator cannot directly view the transmitter and associated monitoring equipment, the metering and perhaps the control must be extended to the operating position such that the required observations can be made. In some cases where the operator can directly observe the transmitter parameters, it is desirable to add equipment to gather the data from the various pieces of equipment, analyze that data, and present the operator with a concise display reporting the status of the transmission system. Such a system may also record the system parameters for later trend analysis. This "enhanced" monitoring and control simplifies the operator interface in complex systems. The *operator interface* equipment may be located at the transmitter site or at another location. Finally, stations are permitted to have a transmitter control system that automatically determines system parameters, checks limits, and makes adjustments. Should such a system detect an interference-causing condition such as excessive power, it may turn off the transmitter without further operator intervention. This *automatic transmission system* (ATS) does what an operator is supposed to do. However, even with an ATS, an operator is required to be present (at the ATS control/alarm point).

This chapter discusses the sampling and control of various transmitter parameters. These sampling and control concepts can be applied to extension metering, remote control, and ATS. The discussion then moves on to data conversion and transmission subjects important to remote control and remote ATS. This is followed by a discussion of the operator interface, data analysis, and FCC regulations. The chapter organization can be thought of as starting at a transmitter and working our way back to the control point.

REQUIRED LEVELS OF TRANSMITTER CONTROL

Extension Metering

If the transmitter operator cannot directly observe the transmitter and monitor indications from the normal duty position, extension meters may be installed. The normal duty position is required to be in the same building as the transmitter, within 30.5 meters, and within one floor above or below the transmitter [§73.1550]. These distance limitations allow the operator to quickly go to the transmitter to make any required adjustments, since extension metering does not typically include extension control.

Remote Control

If the normal duty position is further from the transmitter than authorized for extension metering, remote control (and metering) equipment may be installed. *Remote control* consists of a remote metering system (just as extension metering) and a method of making the required transmitter adjustments from the normal duty position.

Automatic Transmission Systems

Automatic transmission systems monitor various transmitter parameters, automatically adjust those parameters, and alarm any adjustment failures (failure of the adjustment to keep the parameter within licensed parameters). Further, should an interference causing condition exist for three minutes, the ATS is required to shut the station down [§73.1500]. While the existing regulations do not specify which parameters are to be monitored by the ATS system, previous regulations required the ATS to monitor power and modulation. Since Section 73.69 requires antenna monitor readings of a directional antenna system to be available to the operator (remoted), it is suggested that an ATS also monitor the DA parameters. The current regulations also do not specify which parameters the ATS must adjust. If an ATS is unable to adjust a specific parameter (such as DA parameters), it can continue to monitor that parameter, alarm it, and shut down the transmitter should it be found at an interference causing value.

ATS automatically does the routine work of a transmitter operator (parameter limit checking and adjustments). Should the ATS not be able to maintain proper operation, it "calls for help" (sounds an alarm). Should the problem remain uncorrected for three minutes, it shuts down the transmitter.

PARAMETER SAMPLING CIRCUITRY

Most transmitter parameters can be reduced to a "sample voltage" by using voltage dividers or current sense resistors. Extension (or remote) metering is accomplished by sending a signal representing this sample voltage to the remote metering point, then displaying a value representing the original parameter. The concept of sample voltages is important. Although we could measure the parameter directly from the remote metering point, wiring costs often make this impractical. For example, we would not want to send 5,000 volts to the remote metering point so we could measure the final amplifier plate voltage. We would also probably not want to send the several amperes of plate current or antenna current to the remote metering point. Instead, voltages (typically DC) representing the values of these parameters are sent.[2] We may send one volt for every kilovolt of plate voltage. If we calibrate the scale of the remote meter such that every volt of sample voltage represents another kilovolt of plate voltage, the remote meter can directly indicate the transmitter plate voltage without having to deal with high voltages.

In practice, the actual sampling ratio (sample voltage/parameter value) is not important as long as the sampling ratio is stable and the sample voltage is reasonable (low enough to be easily handled, high enough to not get lost in noise). Most transmitter manufacturers provide remote samples for the FCC required indicating instruments. Fig. 1 shows some typical voltage sampling circuits.

DC Voltage Sampling circuits

Fig. 1A shows how a voltage divider may be used to divide the plate voltage down to a safe sample voltage. This circuit is independent of the transmitter front panel plate voltage meter.

In extension meter applications, we may want to adjust R2 to provide a voltage that is an integer order of magnitude below the actual plate voltage. For example, if the transmitter plate voltage meter has a full scale value of 5 kV, we could adjust R2 to provide an output of 3.00 volts when the plate voltage is 3.00 kV. The extension meter could then read volts but be calibrated in kilovolts. The extension meter could be a 3.5 digit digital panel meter (a DPM that reads between −1.999 and +1.999), which typically reads in millivolts. R2 is adjusted such that the sample voltage is 300 mV when the plate voltage is 3.00 kV. The appropriate decimal point on the display can be enabled to directly indicate the plate voltage in kilovolts with a resolution of 0.01 kV.

Fig. 1B shows how the existing plate voltmeter "multiplier resistor" (R1) can be used to drive the front panel meter and provide a sample voltage. The additional 1K sampling resistor (R2) decreases the reading of the local meter by 0.01% (which is negligible).

Clamp Diode for Safety

In each of the sampling circuits shown in Fig. 1, the full plate voltage would be supplied to the remote indicator should the shunt resistor (R2) open (assuming minimal loading by the remote indicator). To protect wiring and personnel, it is common practice to add a zener diode across R2. Should R2 open, the sample voltage is clamped to the zener voltage. The zener will

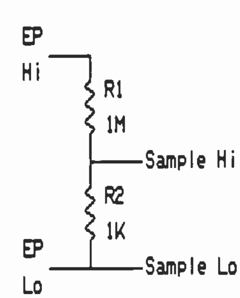


Figure 1A. Unbalanced plate voltage sampling circuit. Sample voltage = EP/1001

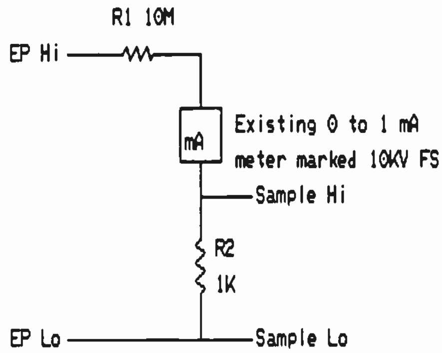


Figure 1B. Adding resistor to existing plate voltmeter circuit provides a remote sample.

typically fail to a short, protecting the external sampling equipment. This zener is commonly damaged in lightning storms, causing a loss of plate voltage sampling.

Common Mode Voltages

Input and output terminals in the circuits of Fig. 1 are labeled "high" and "low." The low terminal is the one whose voltage is closest to ground. In many cases, the low terminal connects directly to ground, allowing the sample voltage to be measured between sample high and ground. Many transmitters, however, do not measure the plate voltage between the plate supply and ground. These transmitters often have resistors between the final amplifier tube cathode and ground. These resistors are used to provide protective bias, cathode current sampling, and output power trimming. The final amplifier input voltage is then measured between the plate supply to the tube (generally +3 kV to +10 kV) and the cathode of the tube (0 volts to +250 volts). The circuits of Fig. 1 could possibly place the low side of the sample 250 volts above ground. This hazardous voltage would require special precautions in wiring. Further, many remote control systems assume the low side of the sample is grounded. Connecting such a remote control to such a transmitter would short out the circuitry in the cathode circuit, disabling the tube protective bias, disabling the current overload protection and causing the transmitter to run maximum power, regardless of the power trim setting.

Common Mode Attenuator

Fig. 2A shows circuitry similar to that presented so far. It is driven by +5 kV from the plate of the amplifier and +250 volts from the cathode. The "plate voltage" is the differential voltage (Va − Vb). The differential voltage is attenuated (multiplied by 0.001) to arrive at the differential sample voltage. Unfortunately, the circuitry shown does not attenuate the common mode voltage ((Va + Vb)/2) as much, multiplying it by 0.096. We end up with the dangerously high common mode voltage.

Fig. 2B shows a sampling circuit that attenuates both the common mode and differential mode voltages equally. As in Fig. 2A, the differential mode sample

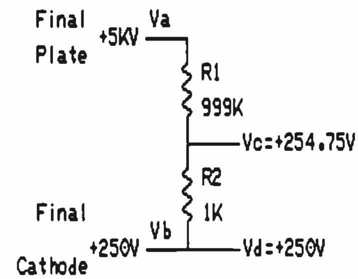


Figure 2A. Plate voltage sampling circuit without common mode attenuation gives dangerous sample voltage.

voltage is 0.001 times the differential mode input voltage (0.001 times 4.75 kV or 4.75 volts, in this case). However, the common mode sample voltage is also multiplied by .001, reducing it from 2.625 kV to 2.625 volts. This circuit can be used to sample high voltages that are not measured with respect to ground *provided* the remote metering equipment has differential inputs.

Common Mode Rejection

The differential output voltage of Fig. 2A (Vc − Vd) is independent of the common mode input voltage ((Va + Vb)/2) if all the resistors are precisely matched. This can be checked by applying a high common mode voltage with no differential voltage. When both Va and Vb are connected to the high voltage supply (+5 kV), one of the 1K resistors can be adjusted as needed to give no voltage between Vc and Vd.

Differential Input Circuitry

Several techniques are available for dealing with differential sampling voltages (where neither side of

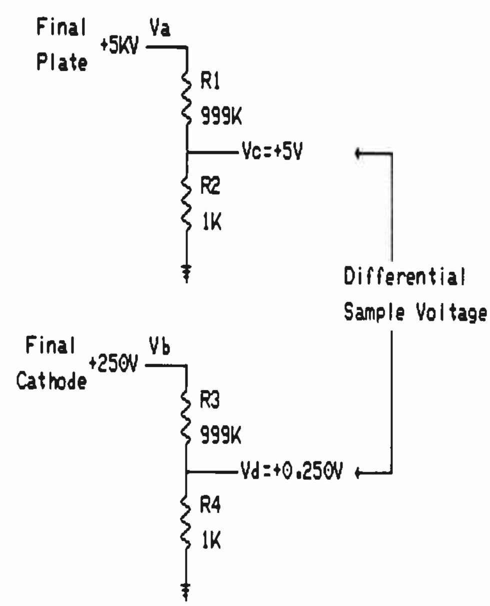


Figure 2B. This voltage divider reduces common mode voltage to a safe +2.65V.

the sample is grounded). For circuit simplicity, many remote control systems assume one side of the sample voltage is grounded. Other systems provide differential inputs. If differential samples must be measured, additional circuitry may be added outside the remote control to convert the sample to a ground referenced sample. Typical approaches to dealing with a differential input voltage include: use of an isolation amplifier, use of a differential amplifier, floating the remote control analog to digital converter, and calculating the differential voltage.

Isolation Amplifier

An isolation amplifier may be included inside a transmitter, inside a remote control, or between the two units. Isolation amplifiers provide isolation between the input and output by using magnetic, capacitive, optical, or thermal coupling between input and output. Many isolation amplifiers require a floating power supply to drive the input circuitry, which considerably adds to cost. Careful design is required to insure that the "encode" and "decode" process track each other (the output voltage is indeed proportional to the input voltage). Some isolation amplifiers use feedback techniques to reduce the nonlinearities induced by the coupling method.

Flying Capacitor Isolator

Fig. 3D shows a "flying capacitor" isolator. C1 "flies" between the floating sample and the ground referenced output. When the switches are to the left, C1 charges to the sample voltage. When the switches are to the right, C1 discharges into C2. After several cycles (depending upon the relative capacities of C1 and C2), C2 will charge to the sample voltage. The switch can be either an electromechanical relay or a solid-state switch (such as the Maxim MAX343). The switching frequency is typically 1 Hz to 100 kHz, depending on the switch type. Since similar switches are needed for the analog multiplexor of a remote control, it is possible to combine the flying capacitor isolator with the analog multiplexor, giving a substantial parts savings.

Differential Amplifier

In most cases, electrical isolation between the sampling circuitry and the remote metering equipment is not required. Instead, a differential voltage must be changed to a "single ended" voltage (one side grounded). This can be accomplished using one of the differential amplifiers shown in Fig. 3. Fig. 3A is a basic differential amplifier. If the resistors are precisely matched, the output voltage will be a constant (the differential gain) times the differential input voltage. If the resistors are not precisely matched, the amplifier will have a "common mode gain" other than 0. The output voltage is $(Va - Vb)Ad + ((Va + Vb)/2)Acm$, where Ad is the differential gain (amplification factor), and Acm is the common mode gain. If the resistors are precisely matched, $Acm = 0$.

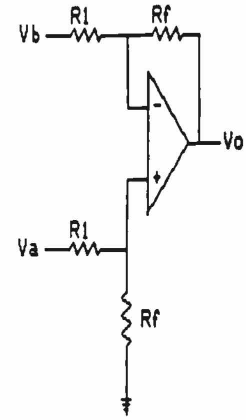


Figure 3A. Basic differential amplifier. Vo = (Rf/R1) (Va – Vb)

The circuit of Fig. 3A suffers from a limited input voltage range. If the amplifier is set up for a gain of 1 (R1 = Rf), the circuit will only accept about twice the input range of the operational amplifier. The circuit of Fig. 3B allows up to 200 volts of common mode input voltage.

The circuit of Fig. 3C (the instrumentation amplifier) loads the sample circuitry less than the other circuits, has better common mode rejection (lower common mode gain), and the gain can be altered by changing only one resistor (R3). The input voltage range is, however, limited to the voltage range of the operational amplifiers.

Calculated Differential Voltage

Each of the schemes so far has been directly measuring the differential voltage or calculating it using operational amplifier (analog computer) techniques. If we go back to Fig. 2B and measure the sample voltages at points C and D with respect to ground (Vc and Vd), the differential voltage can be calculated. A typical remote control could put Vc on one "channel" and Vd

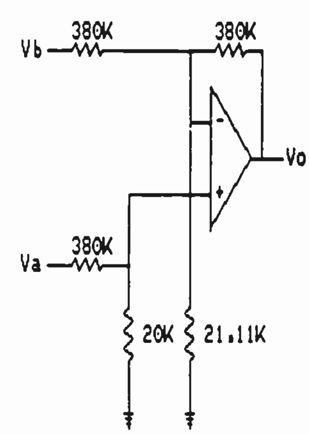


Figure 3B. Burr-Brown INA117 differential amplifier with 200 volt input range. Vo = Va – Vb

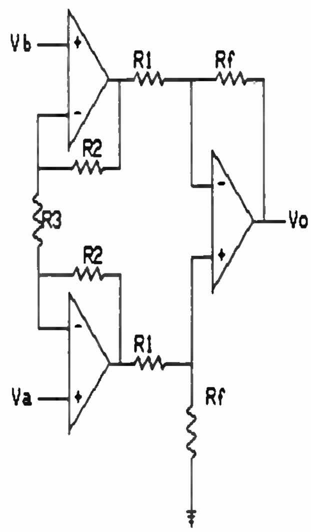


Figure 3C. Standard instrumentation amplifier. Gain can be varied by changing R3.

on another channel. The remote indication would show Vc at 5 volts, which indicates Va is 5 kV. The remote indication would show Vd at 0.250 volts, which indicates that Vb is 250 volts (0.250 kV). Subtracting, the transmitter operator could determine the plate voltage (5.000 kV − 0.250 kV = 4.750 kV). If the remote control has a software subtraction capability, the differential plate voltage could be displayed directly.

This approach does require the voltages to be relatively stable, as most systems do not sample all the channels simultaneously.

DC Current Samples

Fig. 4A shows the use of a current sense resistor to give a voltage sample that is proportional to current. Unfortunately, the common mode voltage of this sample is very high (about 5 kV). An isolation amplifier would probably be required to safely utilize this sample.

Of course, the plate current also flows through the

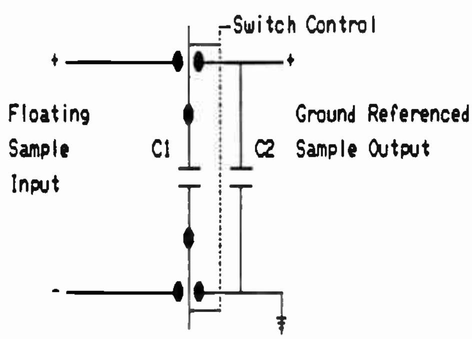


Figure 3D. Flying capacitor isolator.

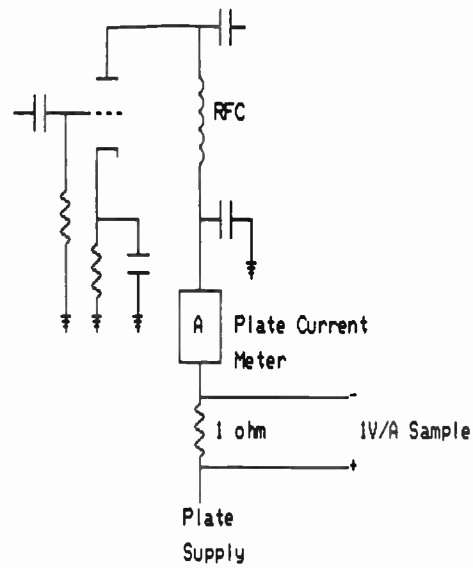


Figure 4A. Plate current sample with dangerously high common mode voltage.

cathode, so a ground referenced sample is available there. However, the cathode current also includes the control grid current and screen grid current (if the tube has a screen grid). If the grid and screen circuits are returned "above" the sense resistor, an isolated sample of plate current is available (as shown in Fig. 4B). If these circuits are not returned above the sense resistor, samples for each of these currents can also be derived. These currents can then be subtracted from the measured cathode current using analog or software techniques.

If a high voltage supply is powering only the plate of the RF amplifier (such as the grounded grid amplifier

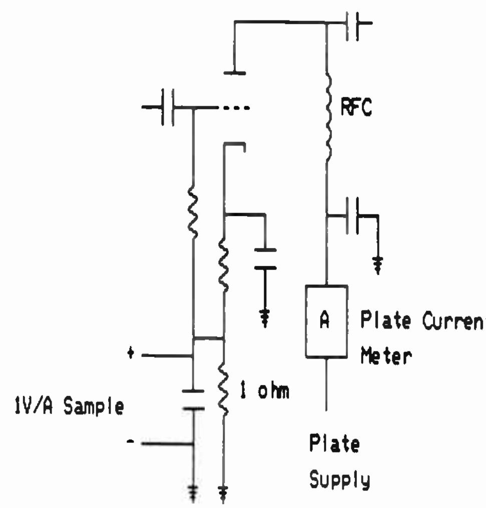


Figure 4B. Plate current sampling in cathode circuit. Note that grid is returned above sample resistor.

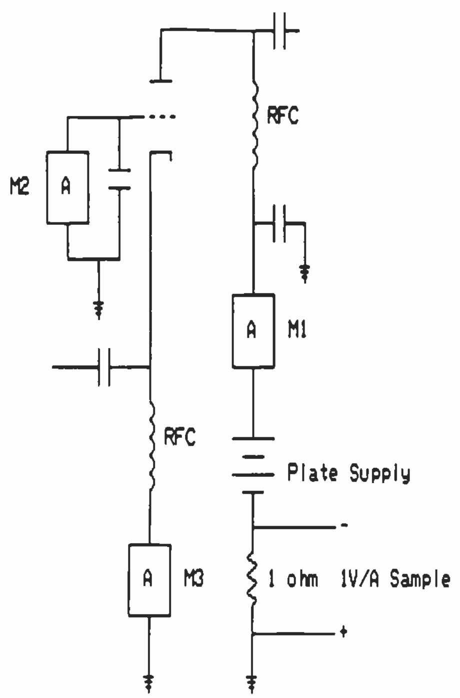


Figure 4C. Plate current sampling in negative side of plate supply.

shown in Fig. 4C), it is possible to measure the plate current by measuring the current into the negative side of the supply. Instead of grounding the negative side of the supply (as typically done), a current sense resistor is placed between the supply and ground. The voltage across this resistor is proportional to the plate current. Note that the current through this resistor is the same as the plate current indicated by M1, while M3 (the cathode current) is the plate current (M1) plus the grid current (M2). Note also that the sample voltage is negative. Most remote control equipment can handle bipolar sample voltages.

Some pulse duration modulated AM transmitters operate with the plate of the final amplifier at DC ground potential. The cathode of the final idles at −7.5 kV and peaks at −15 kV on +100% modulation peaks. The same sampling techniques can be used.

Resistor Stability

The initial tolerance of resistors used as current sense resistors or in voltage dividers is not critical. An inexact resistance will cause the sample voltage to be other than that expected, but, whatever sample voltage appears, the remote meter can be calibrated to agree with the "local" meter. The resistance of resistors changes with time and temperature. The variation with temperature is the major error contributor.

A carbon composition resistor may have a temperature coefficient of about ±600 ppm/°C for resistors up to 1K. This increases to ±1875 ppm/°C for resistors up to 1M. If these resistors were used in the plate voltage sampler of Fig. 1A, the sample voltage *could* change 22% over a 0°C to 50°C temperature range. Carbon film resistors are available with temperature coefficients that range from 0 to −300 ppm/°C for low resistance values to 0 to −1000 ppm/°C for high resistance values. Metal film resistors are available with temperature coefficients down to 15 ppm/°C. Bulk metal resistors are available with temperature coefficients of 5 ppm. If the resistors in a voltage divider are at the same temperature, the matching between the temperature coefficient of resistance (TCRs) is more important than the actual TCR. Since the TCR matching is generally better than the TCR of an individual resistor, sampling systems should be designed to keep the resistors in the network at the same temperature.

In differential amplifiers, the matching of resistor values is very critical. To insure the resistance values track with temperature, it is common to use a resistor network made using hybrid film techniques or integrated circuit techniques. Since the resistors are made at the same time and of the same materials, the TCRs will match closely. Since they are on the same substrate, they will be at about the same temperature.

Other considerations in resistor selection include the power and voltage rating of the resistors. For example, a typical 2 watt carbon film resistor is rated at 750 volts. A resistor with a higher voltage rating, or several in series, would need to be used to sample a high voltage.

AC Voltage Sampling Circuits

Fig. 5A shows a simple transformer isolated AC line voltage sensor. The DC output voltage will be approximately proportional to the differential line voltage (as measured between hot and neutral). The linearity of the sample is limited by the diode "knee" voltage as a proportion of the transformer secondary voltage. This line voltage sampler may be as simple as a "calculator supply" directly driving the analog input of the remote control.

Fig. 5B shows another line voltage sampling circuit. This circuit provides a differential output. Ideally, the neutral line voltage is about zero, so that portion of the circuit may be deleted. If the circuit is used with a remote control that does not have differential inputs, an external differential amplifier may be used, or the hot and neutral samples can drive two input channels of the remote control. The operator or the remote control software can then calculate the actual differential voltage.

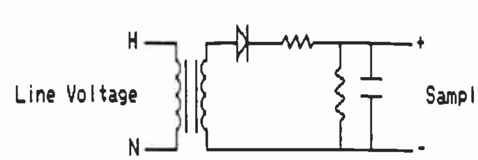


Figure 5A. Transformer isolated line voltage sampler.

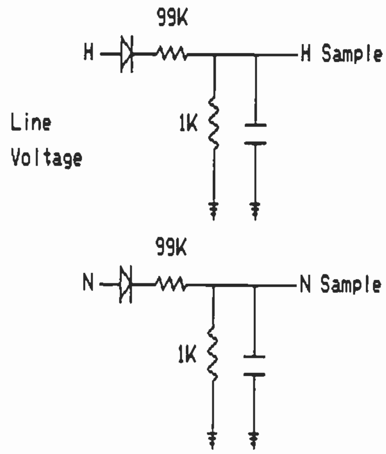


Figure 5B. Non-isolated line voltage sampler with differential sample output.

Three-Phase Line Voltage Sampling

The "hot" portion of the circuit in Fig. 5B can be duplicated for each phase of a three-phase system. The remote control is then presented with four samples, one for each phase voltage referenced to ground, and one for the neutral line voltage. Note, however, that this circuitry will not necessarily indicate a "loss of phase." Should a phase open, a load between two phases may "pull them together," causing the correct voltage to be on each phase, but the proper phase relationship will not be present. Such an open phase can be detected with the "phase imbalance detector" of Fig. 6.

Fig. 6A shows the circuit driven by a three-phase wye source. The three 10K, 2 watt resistors form an "averaging circuit," adding the three voltages and dividing the result by three. As shown in Fig. 6A, these three voltages add up to 0 volts at V1, indicating the circuit is balanced.

Fig. 6B shows the same circuit, but here a phase is "lost." Phase B is following phase A. This places 69.28 Vrms at −30° at V1. The remaining circuitry provides a DC voltage proportional to the half wave average of the voltage at V1. Use of this simple circuit

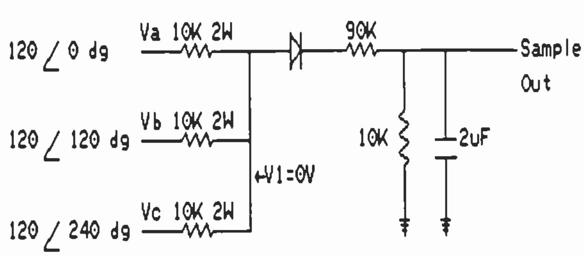


Figure 6A. AC line phase imbalance detector. Note that the sample output voltage is 0 volts when AC line is balanced.

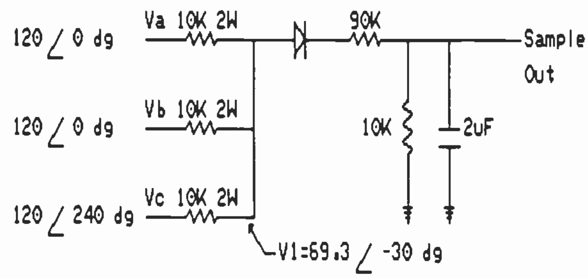


Figure 6B. AC line phase imbalance detector with phase B shorted to phase A, giving sample voltage representing imbalance.

detects loss of a phase while merely measuring phase voltages does not.

Note that the standard wiring of a three-phase delta source results in voltages that are not balanced with respect to ground. The typical voltages are 120V at 0°, 120V at 180°, and 208V at 90°. If these are fed to the circuit of Fig. 6, we get V1 = 69V at 90° and a sample output of 13.13 VDC. This sample voltage could be monitored. Significant deviation would indicate a power line problem. Another approach would be to add three transformers between the line and the phase imbalance detector to change the delta source to a wye.

Note that in all these circuits, the diode is placed before the voltage divider (the 90K and 10K) to minimize the nonlinearity due to the diode knee voltage.

Finally, none of these circuits will detect a "phase reversal." The actual phase sequence is generally unimportant to transmitter power supplies. However, many transmitters use three-phase blowers. A phase reversal will cause the blower to rotate the wrong way, severely deteriorating the transmitter cooling. Phase sequence detecting relays are available that will release a set of contacts on a missing phase or a phase reversal. These contacts could drive an analog or a status input of a remote control.

Filament Voltage Sampling

Filament voltage sampling can be done using the same techniques described above for line voltage sampling. If an isolation transformer is used, the voltage ratio would need to be adjusted to yield the proper sample voltage. If an isolation transformer is not used (instead, just a diode and voltage divider are used), care must be taken to insure that you are only sampling the filament voltage. Grounded grid amplifiers have a high RF voltage on the filaments, so the filament voltage must be measured before the RF isolation inductors. AM transmitters often have a DC common mode voltage on the filament (typically due to circuitry for protective bias, cathode current sensing, and power trim). This can be removed by using an isolation transformer or capacitively coupling the AC voltage to the sampling circuitry.

Tower Light Sampling

The monitoring of tower lighting systems is required by Section 17.47 of the FCC Rules. This section requires a daily observation of tower lights or a properly operating tower light indicator. Section 17.47(a)(1) requires the remote indicator (or alarm) to indicate the failure of any lamp on the tower. Section 17.48(a) requires the failure of a steady burning lamp at the top of a tower or the failure of any flashing obstruction lamp to be reported to the FAA. A daily observation of tower lights requires that the lights be viewed directly from a location which would allow any lighting failure to be observed. A properly operating tower light indicator is one that will report the proper operation of the lighting system as well as the failure of a required light. Such systems usually operate by sensing the current driving the tower lights. Stations operating by remote control often read this sensed current at the remote control point and interpret proper operation by observing the indicated current.

Fig. 7A shows a possible tower light sensor. T1 is a current sensing transformer (Toroid Corporation of Maryland TR-3025) with a primary to secondary current ratio of 300. Ideally, the secondary current is the primary current divided by 300. As circuitry is added to the secondary (secondary not shorted), the secondary current decreases and the linearity of the secondary versus primary current curve decreases. This is due to the output voltage limitations of the transformer (it is not an ideal current source). Fig. 7B shows the peak voltage output of the circuit when measuring the current driving two flashing 620 watt lamps and two steady 116 watt lamps (all rated at and driven with 130 VAC). Note the high output when the beacon is first lit. This high "inrush" current (due to the filament cold resistance) increases the initial beacon current by about 70%. The time constant for the exponential decay down to the steady state current is about 91 mS. If the remote control system uses an integrating (dual slope) analog to digital converter, it will measure the average DC voltage across the resistor during the sample time. A typical dual slope A/D does between 3 and 30 conversions per second. It is not sampling the input during the total conversion time (time is spent in the autozero and reference integrate phases). If we assume 3 conversions per second, 2.73 conversions can be done during the beacon off time (using the typical flash rate and duty cycle). 5.46 conversions can be completed during the beacon on time. By watching the minimum and maximum conversions, it is possible to determine if all lamps are working properly. Many people will be tempted to add a capacitor across the output of the sampler to provide a DC sample to the remote control. This, however, will cause samples taken during the beacon off time to include a component due to the beacon on current and vice versa. This interference between the measured on and off currents is eliminated by allowing the A/D integrator to remove the AC line frequency (actually twice line frequency, since it is full wave rectified) component. A/D sampling rates are typically chosen to minimize the effects of 60 Hz and its harmonics.

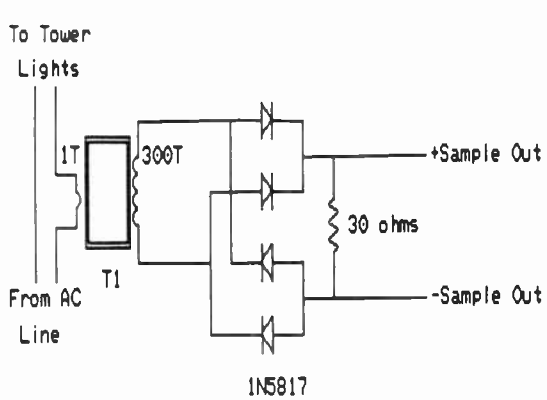


Figure 7A. Tower light current sampler provides 84 mV DC average/RMS AC amp.

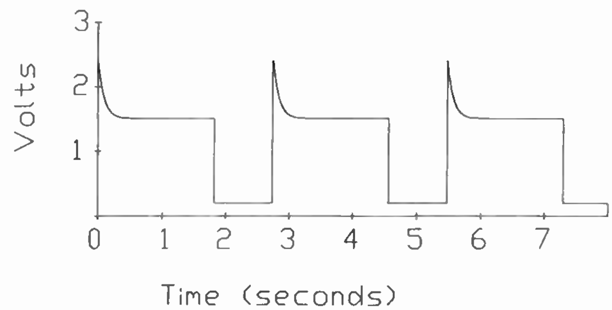


Figure 7B. Tower light sampler output waveform.

Table 1 shows typical minimum, maximum, and average for a series of samples taken over a one minute period with 3 conversions per second. Note that these voltages are less than the peak voltages shown in Fig. 7B (measured with an oscilloscope), since the voltages in Table 1 are "average DC" voltages, measured with an integrating A/D converter.

Note that the circuit of Fig. 7A measures tower light current, which will vary with tower light voltage. FCC Rule Section 17.54 requires the voltage at the base of tower lamps to be no more than 3% less than the rated voltage. The tower lamps tend to be constant current devices. As the line voltage drops, the filament cools, decreasing the filament resistance, increasing the cur-

TABLE 1

Average DC voltages (as opposed to RMS) from circuit of Fig. 7A under various tower lamp failures, 3 samples per second for 60 seconds.

Light Condition	Minimum Volts	Maximum Volts	Average Volts
All ok	0.132 (100%)	1.402 (100%)	0.766 (100%)
Side light out	0.060 (45.5%)	1.217 (86.8%)	0.687 (89.7%)
Beacon out	0.132 (100%)	0.766 (54.6%)	0.465 (60.7%)

rent, though not back to the original value. With a typical tower lighting load (two 620 watt lamps and two 116 watt lamps), a 3% decrease in line voltage causes a 1.57% decrease in lamp current. A 6% decrease in line voltage causes a 3.30% decrease in line current. Since these variations due to line voltage fluctuations (staying within FCC required line voltage limits) are much less than the variations due to a lamp failure, it is simple to differentiate between lamp failures and permissible line voltage variations.

If a remote control returns voltage indications corresponding to tower light current (using a circuit similar to that suggested), minimums and maximums can be checked by an operator or system software to determine if the tower lights are operating properly. If the output of the circuit is run through a low-pass filter with a cutoff frequency well below the beacon flash rate, a steady DC voltage corresponding to the average voltage would be available. This voltage is proportional to the current drawn by the steady burning lamps plus the product of the current drawn by the flashing beacons and the duty cycle of the flashing beacons. A variation of more than 5% in this voltage would indicate a lamp failure.

All the discussion has been based on a "standard" tower lighting load. This consists of two flashing 620 watt lamps and two steady 116 watt lamps. Tall towers require more lamps. It is suggested that the lighting circuits for such towers be sampled individually, breaking the system down to several "standard" lighting loads. Increasing the load substantially will make it difficult to detect the loss of one small lamp.

No matter what approach is used to tower light sampling, it is suggested that a single 116 watt lamp be added to each lighting circuit. This light should remain on at all times or switch on when the other tower lights are turned on. For AM stations, this lamp might be in the antenna tuning unit at the base of each tower. FM and TV stations might use it as a transmitter site "night light." The tower light sensing circuitry and software can be tested by turning off this lamp, simulating the failure of a steady burning lamp. An operator, system hardware, or system software should be able to detect this tower lamp "failure." This is an easier method of testing the system than climbing the tower to unscrew a light bulb!

FCC Rules Sections 17.39 through 17.42 permit the use of high intensity (strobe) lighting as an alternative to tower painting and "standard" lighting. Due to the complexity of the required control circuitry, strobe control systems typically provide outputs that can be used to determine the state of the system (high/medium/low intensity and failure of any lamp). These outputs can be returned to the transmitter operator using standard status monitoring circuitry.

FM AND TV OUTPUT POWER SAMPLING

FM and TV transmitters are required to have a metering circuit that measures the output power. Most transmitters have a remote output of this sensor that can be used to drive remote or extension meters. These meters are generally directional couplers that sample the forward and reflected power at the transmitter output. For those transmitters without a remote output power sample (or those stations using combined transmitters), a transmission line directional coupler with an associated wattmeter (with remote sample) can be added. Television transmitter output power meters include a peak detector circuit, since the actual output power varies substantially with picture content. The peak amplitude (tip of sync) should be a constant power.

The DC output of the reflectometer has a relatively high source impedance and is easily loaded. Many transmitters include isolation circuitry so the remote or extension meter does not load the directional coupler, causing inaccurate indications on the transmitter front panel meter.

The sample voltage out of a directional coupler is proportional to the square root of the power (directly proportional to the voltage or current), except at very low powers, where the diode knee voltage again causes nonlinearities. The output power meter on most transmitters is a mechanical meter with a nonlinear scale, allowing it to read output power directly. An analog extension or remote meter can also use the same scale, allowing direct reading of the transmitter output power. Digital meters, however, generally use a linear analog to digital converter. Getting an accurate remote indication involves squaring the sample voltage either before or after the analog to digital converter. The sample voltage can be squared prior to the A/D using an analog multiplier (or balanced modulator) and tying the two inputs together (multiplying the sample voltage by itself). Devices that do this analog squaring are often called "power to linear converters." The sample voltage can also be squared after the A/D converter. This may be done by giving the transmitter operator a chart (indicated power versus actual power). Moving to the computer age, you can give the operator a calculator! Finally, most microprocessor based remote controls make this calculation automatically.

AM POWER SAMPLING

AM stations generally determine power by measuring the RF current into the antenna (system). For nondirectional stations, this is the base current of a series fed antenna or the feedwire current of a shunt fed antenna. A directional station measures the current into the "common point" of the array. In each case, the measured current is squared and multiplied by the resistance (base, feed wire, or common point) to arrive at the power into the antenna. The Rules also permit the use of direct reading power meters that measure the voltage, current, and phase relationship between them to determine the power, but these meters are rare.

RF currents have traditionally been measured with thermocouple meters, which measure the true RMS current. Remote thermocouple meters are available, but suffer from nonlinearity between the RF current and the sample voltage. Further, when an AM transmitter is modulated, the RMS current into the antenna increases (corresponding with the addition of power to the AM sidebands). The FCC places limits on the unmodulated antenna current. Having the antenna current indication vary substantially with modulation would require us to check the antenna current without modulation. For these reasons, "diode meters" are now generally used for remote RF current indications.

Diode meters utilize an RF current sensing transformer to develop a DC voltage that is proportional to the RF current. The DC sample is obtained by running the rectified RF through a low-pass filter, removing the AC components (RF and audio). Ideally, the DC component of the rectified RF is constant, since the transmitter AC couples the audio into the final amplifier (using typical modulation techniques). The sample voltage will be the same with or without modulation, allowing measurement of the unmodulated antenna current without interrupting modulation. Again, there are potential nonlinearities due to diode knee voltages. However, these may be overcome by placing the current sense transformer terminating resistor after the rectifier, as was done in the tower light sensor described above. Other techniques are available to eliminate this nonlinearity. These include developing a bias voltage to "get over the knee" and the use of zero crossing driven FET switches (similar to a "synchronous rectifier").

A well designed diode meter (such as the Delta TCA series) also serves well as the "local" antenna current meter. These meters provide a local indication and a sample voltage suitable for driving extension or remote metering.

In practice, the indication of a diode-type RF ammeter will vary with modulation. This is generally due to less than perfect carrier amplitude regulation (carrier shift) in the transmitter. Often the transmitter high-voltage power supply is loaded by the modulators as the modulation level is increased, decreasing the high voltage available for the final amplifier. This can cause a decrease in the indicated antenna current. Variations in the antenna impedance with frequency may also cause carrier shift, as the various sideband frequencies "see" a varying load impedance. If these variations in indicated antenna current are excessive (remote antenna current meters are required to agree with the local meter within 2%), the readings would have to be taken without modulation, or perhaps some electrical compensation could be added (see "Carrier Shift Compensation," below).

Some stations take a sample of the RF voltage at the antenna input (base, feed wire, or common point) and calibrate the remote indication to agree with the local ammeter. If we assume the impedance of the antenna is constant, then the RF voltage will indeed be proportional to the current.

Finally, directional stations are equipped with a very high-quality RF detector circuit in their antenna monitor. Stations that operate nondirectional some portion of the time may use the indication of the antenna monitor current sample for the nondirectional tower as a remote base current sample. The actual RF sample is typically from an RF current transformer at the base of the tower or a current sampling loop part way up the tower. Stations can also add a current sense transformer just prior to the common point meter. This transformer can drive the antenna monitor to give a remote sample of the common point current.

Carrier Shift Compensation

The circuit of Fig. 8 may be added between a Delta TCA RF ammeter and a remote control (input resistance of 1 M or greater) to compensate for variations in indicated antenna current due to carrier amplitude shift with modulation. Fig. 9 shows the remote sample voltage out of a Delta TCA20EXR with modulation. Normally, the AC component (due to modulation) is removed by a low-pass filter in the remote control or the "mechanical low-pass filter" created by the movement of mass in a meter movement. If R2 is set to the center of its range, the circuit of Fig. 8 forms a low-pass filter (Fh = .034 Hz). The output voltage will be the DC component of the input, regardless of any AC component (due to modulation). If, during modulation, the DC component drops (due to carrier shift), the DC output voltage will drop. If, however, the wiper of R2 is moved towards D1, C1 will tend to charge to the peak input voltage instead of the average voltage. If the output voltage without modulation is measured (remote control calibrated without modulation), then R2 adjusted for the same indication with modulation, most of the effects of carrier shift can be removed. Testing of the circuit on a 5 kW plate modulated transmitter operating into a three-tower directional array reduced variations in indicated common point current from −1.2% to ±0.1%.

Note that the circuit must not be loaded by the remote metering circuitry. If a low resistance (less than 1 M) is to be driven, a standard operational amplifier voltage follower circuit should be added between the carrier shift compensator and the remote metering system.

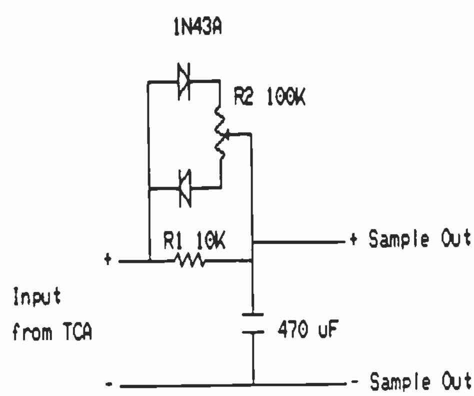


Figure 8. Remote ammeter carrier shift compensator.

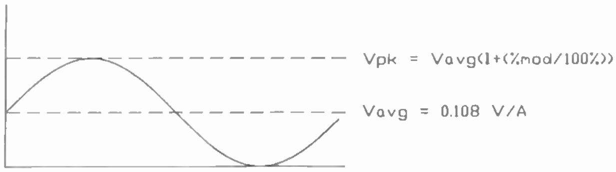


Figure 9. Remote sample output waveform of Delta TCA20EXR. Note that Vavg will typically decrease with modulation due to carrier shift. Vpk increases with modulation.

DIRECTIONAL ANTENNA MONITOR

AM radio stations with directional antenna systems must have an antenna monitor to verify the proper operation of the antenna system when operated in a directional mode. The indications of a directional antenna monitor may also drive remote or extension meters when designed for such use and when authorized by the FCC [§73.53(a)(9)]. Most antenna monitors require a tower select input and usually a pattern select input in order to measure the amplitude (or ratio) and phase of the selected tower. As a result, the outputs of such monitors are not continuously available to remote control equipment. Additional difficulties are caused by the necessity to properly indicate phase relationships that may be positive, negative, or very close to 0°. In extension metering systems, the tower select controls can be extended to the metering point, allowing the operator to select which tower the monitor will measure.

Most (though not all) antenna monitors use one or two amplitude selectors and one phase detector. This technique requires that the detector circuitry be shared among several towers through the use of an RF multiplexor which selects the desired tower input samples to the detectors. Monitors with two amplitude detectors leave one connected to the reference tower at all times. The other (or only) amplitude detector is switched between the RF samples from the various towers in the antenna system. One input of the phase detector is also constantly connected to the reference tower. The other input of the phase detector follows the amplitude detector, connecting to one tower sample at a time. When a tower is selected on such an antenna monitor, the response time of the RF multiplexor and the settling time of the detection circuitry prevent an accurate indication from being immediately available. At least one antenna monitor avoids this problem by providing complete detector circuitry for each tower input sample. This approach eliminates the need for tower input selection and can provide continuous indications of all ratios and phases to remote control equipment.

In remote control applications, when tower select inputs are required, the control relays of the remote control system are typically used to drive the tower select inputs of the antenna monitor. A fairly common practice has been to assign a tower to a channel, then use the RAISE control to read the phase and the LOWER control to read the loop current or ratio. This technique, however, requires interface circuitry between the remote control and the antenna monitor (the control contacts are selecting which indication, phase or ratio, as well as the tower to be sampled) and may result in calibration difficulties, since two dissimilar readings will need to be calibrated on a single analog telemetry input. A simpler approach is to have one or more channels assigned to phase and one or more channels assigned to loop current. The analog inputs of these channels can connect directly to the antenna monitor phase and loop current sample outputs, and each channel will be calibrated to only one type of indication. The RAISE and LOWER control lines then drive the tower select inputs of the antenna monitor. Table 2 lists some sample channel assignments for such a system.

Some remote controls provide a channel selected output that notifies external sampling equipment that the channel is selected. With such a system, the channel selected output can directly drive the tower select input of the antenna monitor, allowing the operator to get an antenna monitor reading in the same manner as any other reading (i.e., by simply selecting a channel). Such systems can also scan the monitor channels automatically. In this case the response time of the RF multiplexor and settling time of the detector circuits must be considered. Most antenna monitors will present a valid sample about one second after a tower is selected.

To simplify the remote interface of antenna monitors, "sample and hold" devices are available for use with antenna monitors. These devices scan through the tower selects on the antenna monitor and store the resulting sample voltages. The device then provides a steady output sample voltage for the phase and loop current (or ratio) for each tower in the system. This makes the antenna monitor appear like other samples at a transmitter site (continuously available). Operators and system software must allow for the possibility that such a sample and hold system may have acquired

TABLE 2

Suggested channel assignments for remote monitoring of a 4 tower DA. This arrangement removes the requirement for interface relays between the antenna monitor and the remote control. The "current" indications may be sample currents (typically from a sampling transformer at the tower base or a sampling loop on the tower) or may be current ratios.

Parameter	Control (Channel)
Tower 1 phase	Raise(10)
Tower 2 phase	Lower(10)
Tower 3 phase	Raise(11)
Tower 4 phase	Lower(11)
Tower 1 current	Raise(12)
Tower 2 current	Lower(12)
Tower 3 current	Raise(13)
Tower 4 current	Lower(13)

some data with the station operating in one pattern, and the remainder of the data in another pattern. In addition it is possible for sample and hold systems to retain valid antenna monitor readings, even though the transmitter may be off the air. Operators should not rely on such an indicator to determine anything other than antenna parameters.

Getting a remote indication of the sign of the phase angle indication on an antenna monitor is difficult. Most antenna monitors output a positive voltage (typically 0 to 1.800 volts for 0° to 180°), whether the sign of the phase is positive or negative. Some antenna monitors introduce a delay in the reference tower circuitry when a *phase sign* switch is pressed. This delay will cause an increase in the phase indication if the phase is positive (selected tower leads reference tower), or a decrease in the phase indication if the phase is negative. Although this switch could potentially be remote, it rarely is. Some antenna monitors use digital techniques (such as a D flip-flop) to detect the sign of the phase. While the phase sign is brought out for remote display (typically as a TTL level), the analog output for driving the remote control still outputs a positive voltage regardless of the phase sign. This phase sign logic level can be returned to the operator through a remote control status channel. The operator (or system software) can negate the phase indication, when appropriate.

Since it is rare to have the sign of a DA tower phase change (unless it is very near 0° or 180°), many stations wire the antenna monitor to the remote control such that the proper sign is indicated (reverse input leads to get a negative reading), if the remote control has a floating analog input. Others may merely note the sign of the phase for each tower in the operator instructions at the remote control point. Microprocessor based remote controls may use a negative calibration or scaling factor to cause a positive input voltage to display a negative value (corresponding to a negative phase).

Finally, note that most antenna monitors will require a "pattern select" control input. This input may change reference levels and/or reference towers. This input is typically driven by the antenna phasing equipment control system. This insures that the antenna monitor and the antenna system agree on the pattern.

PROBLEMS IN ANALOG SAMPLING INTERFACING

Due to a lack of industry standards, one cannot always connect the remote outputs of a transmitter directly to a remote control. A few typical problems and suggested solutions are listed here.

Sample Voltage Too High

To eliminate errors contributed by voltage dividers inside a remote control, many manufacturers run the sample voltage directly through an analog multiplexor into the analog-to-digital (A/D) converter. Voltage limitations on the multiplexor and the A/D limit the maximum sample voltage the system will accept. Exceeding the multiplexor maximum voltage will often introduce "leakage" into other channels, making all readings appear to be out of calibration. Exceeding the A/D maximum input voltage will cause an erroneous reading on that channel (or an overrange indication) and may cause erroneous readings on other channels (due to A/D overload recovery time).

Some transmitters provide adjustable sample voltages. In this case, the sample voltage should be adjusted down to match the optimum input of the remote control.

Sometimes a voltage divider can be modified by adding an external resistor across the remote output of the transmitter, reducing the sample voltage. This will not be the case if the transmitter includes an isolation amplifier for the metering circuitry. The added shunt resistor should be a high-quality, low-TCR resistor.

The sample voltage output of the transmitter can be run through another voltage divider, reducing the sample down to the optimum for the remote control. Again, high-quality, low-TCR resistors should be used.

Sample Voltage Too Low

Too low a sample voltage will cause a lack of resolution in remote metering circuitry. Many systems will not allow you to calibrate a system if the sample voltage is less than 100 times the A/D step size, since this would give a resolution of less than 1%. About the only solution to low sample voltages is the addition of an amplifier between the sample and the remote control. This may be required on FM or TV power meters (reflectometers or directional couplers).

Some parameters are ideally zero (such as reference tower phase, three-phase line imbalance, or reflected power). The sample voltage for these parameters are also ideally zero, making it difficult to calibrate remote metering.

If the antenna monitor has a "180° calibrate" button, this switch can be held while the remote control is being calibrated. The system is calibrated to the indicated 180°. When the calibrate button is released, the remote control should agree with the local meter.

A three-phase balance detector can be calibrated by purposely unbalancing the three-phase input. This is typically accomplished by disconnecting one input of the detector and grounding it (setting it at 0 volts). The unbalance can be measured with a voltmeter. The remote metering can then be calibrated to agree with this indication.

The remote indication of a reflected power sample can be calibrated by calibrating it to the forward power. If the transmitter uses the same circuitry (same size power sampler, etc.) for forward and reflected power, the input of the remote control can be temporarily connected to the forward power sample. The remote metering is then calibrated to agree with the forward power indication. The remote control input is then reconnected to the proper point.

Some transmitters utilize rotatable "slugs" to sample the forward and reflected power. The reflected power slug can be rotated to measure forward power. If the reflected power slug is more sensitive than the forward power slug (full-scale reflected is less than full-scale forward), then the transmitter power should first be trimmed down to the maximum the reflected sampler can handle. If the transmitter includes VSWR shutdown circuitry, it will have to be disabled during this adjustment. Otherwise, the transmitter will be immediately shut down.

Similarly, some transmitters have coaxial cables going out to external forward and reflected power samplers. These cables can be temporarily reversed, giving a high reflected power indication, allowing calibration. Again, VSWR shutdown circuitry would have to be disabled during this time.

Sample Voltage Differential, Remote Single Ended

Often a transmitter provides differential sample voltages while the remote control expects one side of a sample voltage to be grounded. Various techniques for handling this have been suggested above. One solution (the historically typical approach) is to add a differential amplifier between the transmitter and the remote control. Another solution suggested involves measuring each side of the differential voltage with respect to ground, then determining the differential voltage through calculation (either an operator calculation or a computer calculation).

Sample Voltage Not Linear

Generally, the sample voltage is the measured parameter multiplied by a constant. Sometimes it is the square root of the measured parameter (typically for power indicators). Other times it may be linearly related to the parameter, but offset (zero sample voltage does not correspond to zero parameter, often the case for temperature indicators).

In each case, the "calculation" required to convert the sample voltage back to the parameter can be done before the remote control, in the remote control, or after the remote control. *Power-to-linear converters* can square the sample voltage so it linearly tracks the power. Differential amplifiers often include an offset control that can be used to subtract out whatever may be necessary to make the sample voltage track the parameter properly. The remote control itself may contain calculation software to perform these calculations after the A/D conversion. Finally, the remote control may just indicate the sample voltage. The operator may use a calculator or a look-up table to determine the actual parameter.

STATUS SAMPLING

It is often desirable to remotely indicate the status on/off indicators. Many transmitters utilize back lit push button switches for local control and status indication. The switch that corresponds to the current status is lit. For example, if there are a pair of switches marked "plate off" and "plate on," pressing the "plate on" switch will cause the indicator in the "plate off" switch to extinguish, while the indicator in the "plate on" switch will be illuminated. This sort of indication is very "user friendly." An operator could determine that the plates had been turned on by looking at the plate voltage or current meter, but seeing that the button lit is a more intuitive indication that the desired action took place. Other status indications that are often available at a transmitter site include filaments on/off, power high/low, stereo/mono, overload tally, RF switch settings (including DA pattern select relays, transmitter select relays, and combiner relays), STL carrier presence, STL selected, audio processor selected, fire alarm, and burglar alarm. Each of these is an on/off indication. Many are isolated switch contact closures. Some are switch (or open collector transistor) closures to ground. Some are switch (or open collector) closures to some supply voltage. Most remote controls accept closures to ground. If the status input is open, the remote control pulls the status line to +5 volts. This input design is generally used due to low cost. The inputs drive CMOS (complementary metal oxide semiconductor) or TTL (transistor transistor logic) inputs through input protection circuitry.

If a transmitter status output is not compatible with the remote control status inputs, additional interface circuitry may be added. This circuitry may consist of relays or optical couplers, which provide complete isolation between the transmitter and the remote control. Resistor/transistor circuits or integrated circuits may also be used for "level translation," converting the transmitter's status levels to levels appropriate for the remote control. Finally, the transmitter status outputs can drive analog inputs to the remote control. These analog samples can be compared with a threshold (by the operator or system limit checking software) to determine whether the status is true or false.

CONTROL OUTPUTS

Transmitter remote controls have traditionally provided momentary isolated relay contact closures for control. Two control outputs were provided for each analog input. These outputs were designated *Raise* and *Lower*. If the operator had analog channel 1 selected (perhaps plate voltage), pressing the *raise* button closed a pair of contacts that were wired to the plates on control input of the transmitter. The transmitter plates would go on, and the operator would see the plate voltage go up. Pressing *lower* while plate voltage was selected turns the plates off. If channel 2 were selected (perhaps plate current), *raise* and *lower* might change the transmitter between high and low power. Use of momentary controls allows the remote control to be connected in parallel with the local control (transmitter front panel) while allowing control from both locations simultaneously. Latched control lines (nonmomentary) would disallow control from more than one location.

Use of isolated relay contacts allows the remote control to drive a variety of control circuits. Older transmitters used 120 to 240 volts AC to drive the control relays. Isolated relay contacts were suitable for driving these transmitters, though such transmitters did place hazardous voltages on the remote control. Some current transmitters are expecting the remote control to provide a momentary contact closure to some positive voltage (generally +12 or +28 volts). Again, remote controls with isolated relay contacts can drive these directly.

Many remote control systems provide open collector control outputs. This is generally a cost saving technique. A single integrated circuit may provide four, eight, or even 32 open collector outputs. This compares quite favorably to one control output per relay. Further, open collector outputs that all switch to a common ground also allow a higher wiring density.

Fig. 10 shows the interconnection of a remote control open collector output to a transmitter control input. The diode across the relay coil clamps the control voltage when the transistor turns off, preventing the transistor from breaking down due to excessive voltage. A remote control may include clamp circuitry to protect the control output transistors.

If a transmitter uses "high side" control switching (expecting a closure to a positive voltage), it may be possible to reverse the power supply leads to the control interface portion of the transmitter, converting it to ground switching. This eliminates the need for "repeat" relays between a remote control with open collector outputs and the transmitter. If the control relays include clamp diodes across the coils, they will have to be reversed. Further, if the transmitter uses magnetic latching relays in the control circuit, their actions will probably reverse due to the change in current direction.

Finally, at least one transmitter expects contact closures to ground, but interrupts the ground with the

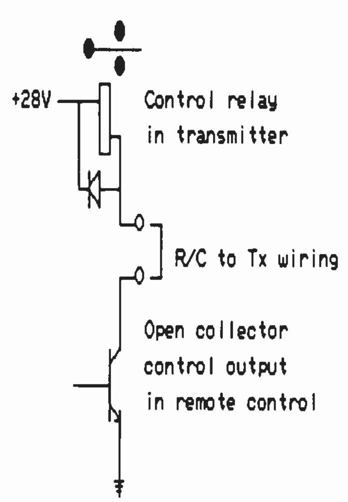


Figure 10. Open collector remote control output driving a transmitter control relay.

local/remote switch. If the remote control uses relay outputs to switch the control inputs to the "switched ground" of the transmitter, all works fine. However, if the remote control switches the control lines directly to ground (as an open collector output would), the remote/local switch is bypassed. This may present a safety problem to someone unfamiliar with the transmitter site. The remote control would still be able to turn on the transmitter, even though the local/remote switch was in local. It may be possible to move the wiring of the remote/local switch to the top of the control relay coils. The switch would then interrupt the positive supply to the relays when the switch was in local.

Control Interface Panel

In all but the simplest transmitter sites, it is common to have a relay interface panel between the remote control and the various pieces of transmitter site equipment. Besides any repeat relays that may be required to convert from open collector drive to whatever the transmitters may require, such a relay interface panel will generally use *relay ladder* logic to interlock various transmitter site functions. Such an interface panel may handle the sequencing of RF switches in transmitter or DA switching, insure that a transmitter drives a load before it can be enabled, and insure that dummy load cooling is on before power is sent to a dummy load.

In the design of a transmitter site, there is an interesting "division of responsibility" problem. Various portions of the site control and monitoring can be assigned to the remote control, control interface, or the transmitter. It has been common practice to have a transmitter accept control pulses and report its current status back to external equipment. Many transmitters now include an automatic power control, which moves some of the control into the transmitter. Some other transmitters have moved much or all of what has typically been considered a remote control's job into the transmitter itself. In most stations, however, there is a fair amount of "decision making" required outside the transmitter. A fairly common decision problem is the RF sequencing of an AM station pattern change. Similar sequences are required for transmitter changes in other AM stations, FM stations, and TV stations.

AM Pattern Change Sequencing

Many stations have used time delay sequencing circuitry to do a pattern change while insuring the RF relays did a "cold switch" (switching while no RF present). A typical sequence might be:

1. Drop transmitter carrier.
2. Wait one second.
3. Switch RF relays.
4. Wait one second.
5. Bring up transmitter carrier.

There is another approach that results in faster pattern changes and insures the station will not come back up if all the RF relays do not operate properly. This "interlocking" of DA RF relays was an FCC requirement (under the operator requirement rules) when the FCC first reduced operator requirements for DA stations to less than a First Class Radiotelephone Operator. The interlocking rule has been deleted, but is still good practice (insuring against operation with other than the desired pattern, and protecting equipment from excessive currents or voltages should a relay stick). An interlocking and sequencing system can be easily set up using relay ladder logic in a remote control interface panel. In relay ladder logic, logical AND gates are formed by wiring relay contacts in series, while logical OR gates are formed by wiring relay contacts in parallel. It is also fairly common to form OR gates using diodes in combination with the relays, reducing the number of contacts a relay may require. The logic for a typical DA pattern change is listed below.

IF NightSelected AND OnDayPattern AND PlatesOn THEN CommandPlatesOff
IF NightSelected AND OnDayPattern AND PlatesOff THEN RFSwitchToNight
IF NightSelected AND OnNightPattern AND PlatesOff THEN CommandPlatesOn

This represents three series strings of relay contacts. NightSelected is the pattern change command from the operator. OnDayPattern and OnNightPattern represent the current condition of all the RF relays. Neither will be true during the switch (as the relays change the system configuration). The PlatesOn and PlatesOff status lines correspond to the transmitter's state. Actually, we are interested in whether there is any RF in the system. A interface panel may use an RF sample from the final amplifier of the transmitter (often a modulation monitor tap) to drive a relay that indicates whether RF is present or not.

Some AM transmitters have a very fast carrier interrupt input. This on/off input drives the modulator to −100% (envelope voltage zero) in pulse duration, digital, and series modulated transmitters. On receiving a carrier drop command, the transmitter drops the carrier in 100 microseconds or less. Since it takes substantially longer than this for an RF relay to open a contact, the carrier drop and relay drive signals can be supplied simultaneously.

The carrier interrupt input on some pulse duration modulated transmitters expects a floating short (in a switched +28 volt DC line) to keep the carrier up, and an open to drop the carrier. This can be driven with floating relay contacts. In many cases, a resistor (10 K typical) may be added across the carrier interrupt terminals. The carrier drop circuit input (the side of the resistor that is driven by the +28 volts, not the side that provides the voltage) may be grounded with a relay contact or an open collector driver to interrupt the carrier.

Finally, in the design of the control system for AM directional stations, it is suggested that the relay interface panel send low voltage DC (typically 24 volts) to repeat relays in the tuning units at the base of each tower. Standard 25 pair telephone cable can be "looped past" each tower. This cable carries these low voltage control signals. It also carries the status of each RF contactor in each antenna coupling unit. The cable can also be used to send sample voltages for RF base currents, tower light sample currents, and to provide a telephone at the base of each tower. The pattern change repeat relays switch the coil voltage to the RF contactors. The contactors can be powered by the same wiring used to drive any tower lights, reducing the requirements for heavy wiring. It is suggested that the contactors be run on the highest voltage available (typically 208 or 240 volts), reducing the current requirements and the resulting voltage drop in the power wiring.

REMOTE CONTROL HARDWARE DESIGN

The basic purpose of a transmitter remote control system is to extend the transmitter metering and control from the transmitter site to the operator. This may be accomplished by running a pair of wires for each meter and another for each control. This is the technique used in "extension metering." It can be used in "remote control" also, but it is generally more economical to combine several functions onto a single pair of wires. The basic transmitter site remote control unit multiplexes the metering, status, and control signals onto a single communications circuit. This circuit may be a DC circuit, voice grade circuit, or digital circuit. The characteristics of various circuit types are discussed in more detail later.

A DC circuit is a pair of wires that has DC continuity from one end to the other (a "metallic pair"). DC and low frequency AC (typically to 3 kHz) may be sent through these circuits. Early broadcast remote controls used two DC circuits: one for control and the other for metering. The metering sample selected by the operator is put directly on the metering pair. A current or voltage meter at the control point displayed the transmitter parameter. The various control functions could be sent over a single control pair by feeding various levels and polarities of voltages between the control conductors, or between each conductor and ground.

Later "single pair" remote control systems utilized audible tones for control and continued to use DC for metering. This allowed both signals to be on the same pair of wires using a form of frequency division multiplexing.

Voice grade circuits typically pass a 300 Hz to 3 kHz band of frequencies. Voice grade circuits are available over a very wide range of media: including telephone lines, radio links, subcarriers on microwave, broadcast, and satellite systems. Most broadcast remote control equipment is designed to work over voice grade circuits, since they are so widely available. Many signaling techniques are available using voice grade circuits.

Digital circuits are circuits that present to the customer provided equipment, at each end of the circuit, in digital form. Between the customer terminations, the circuit may have taken any of several forms. It may have been run through modems, converting the data for transmission over a voice grade circuit. It may have been combined with other digital data into a high speed data stream for transmission over microwave, satellite, or fiber optic cable. At the other end of the high speed medium, the data are split off to each customer.

A remote control must combine the various control and telemetry signals, then encode them in a manner suitable for transmission over the medium available.

Analog Multiplexing

Since we generally have several sample voltages to send over one available circuit, a switch is used to select the sample voltages, one at a time, for transmission to the control point. In many remote control systems, this switch is under the direct control of the operator. When the operator selects a metering channel, the switch at the transmitter site selects the appropriate sample and sends the result to the control point for display. These systems may be thought of as a "voltmeter with very long leads." Just as a technician moves voltmeter probes from point to point in a circuit to get an overview of circuit operation, a transmitter operator remotely selects which sample to monitor. Early remote control systems used the electromechanical *stepper switch* from the telephone industry to select among the metering samples. The step pulses for the stepper were generated with a mechanical telephone dial.

More recent remote controls use individual mechanical relays (reed relays or other relays designed for dry switching) for analog multiplexing.

The most recent remote control designs use integrated circuit multiplexors for analog multiplexing. These circuits act like a mechanical rotary switch. A typical multiplexor may accept three address inputs to determine which one of eight inputs will be connected to the output. The IC multiplexors cost less than relays, are faster, do not have a *cycle limit*, and take less space. They do, however, have a higher leakage current than relays and do not tolerate voltages higher than the supply voltage. If the voltage on an input is higher than the supply voltage, stray currents will generally develop in the circuit, causing erroneous readings on all inputs. With appropriate precautions, however, integrated circuit multiplexing is well suited to the job.

Telemetry Encoding

Once the analog multiplexor has selected a sample for transmission to the control point, it is necessary to encode the data into a form suitable for the communications medium involved. If the circuit has DC continuity, the output of the multiplexor can be applied directly to the circuit. If the circuit does not have DC continuity, further encoding is necessary.

Analog FM Encoding of Telemetry

A low cost method of converting a DC sample voltage to a form suitable for transmission over a voice grade or other AC circuit (such as subcarrier) is to use the sample voltage to drive a voltage-to-frequency converter (V/F) or *frequency modulated oscillator*. A block diagram for a system using this technique is shown in Fig. 11. A zero sample voltage results in one

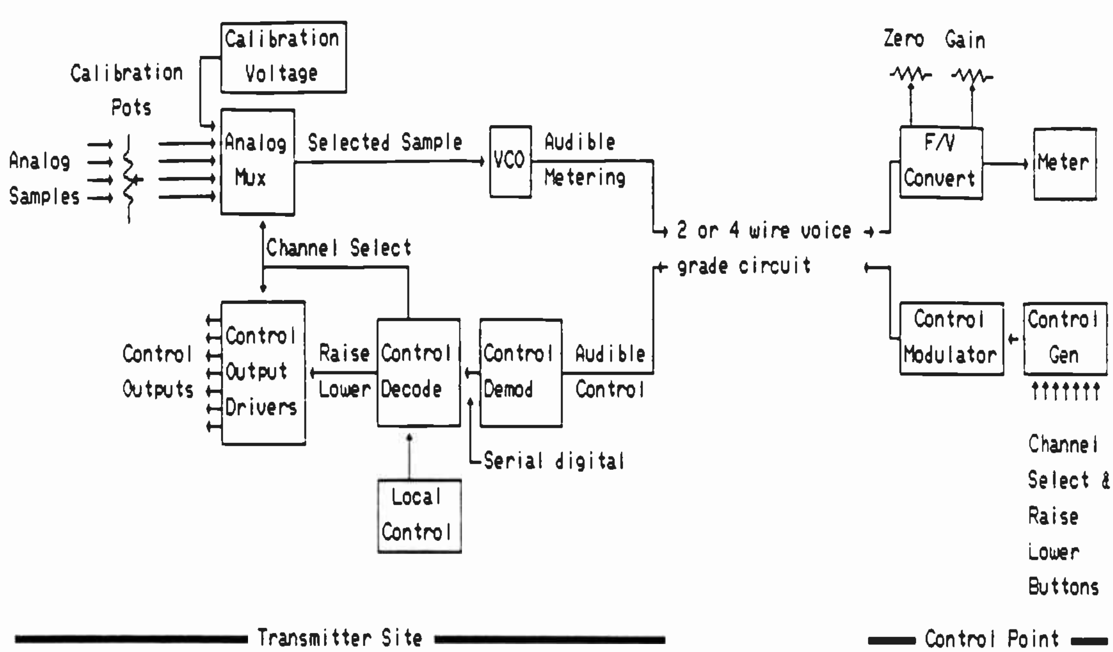


Figure 11. Typically analog remote control block diagram.

frequency, full scale results in another. A half scale sample results in a frequency half-way between the two. One such system utilizes 800 Hz for zero sample voltage and 1200 Hz for full scale. This is suitable for transmission over any voice grade circuit. A similar scheme uses 20 Hz to 30 Hz to transmit analog metering below program audio on AM stations or subcarriers. This system can also be used to generate higher frequencies for use as subcarriers. It is more common, however, to use the audible or subaudible metering as input to a subcarrier generator. This may result in FM (the remote control V/F converter) on FM (the subcarrier generator) on FM (an FM broadcast or TV aural transmitter). This approach is used because of the common *voice grade* interface.

At the control point, an F/V converter (or FM demodulator) develops a DC voltage that corresponds to the sample voltage. This voltage drives the meter or display at the control point.

Accuracy of these systems is limited by the tracking between the V/F and F/V converters. Typical errors include zero drift, gain drift, and nonlinearity. The zero and gain errors can be minimized by frequently adjusting the F/V converter offset and gain while driving the V/F with zero and full scale samples.

Calibration of these systems typically involves a transmitter technician using the telephone to call the operator at the remote control point. The transmitter technician then adjusts calibration controls at the transmitter site until the operator at the control point says the reading matches the reading at the transmitter site.

A/D With FSK Transmission

Another method of transmitting the analog sample voltage to the control point consists of splitting a digital voltmeter into two parts (see Fig. 12). The analog-to-digital converter is put at the transmitter site, and the display is put at the control point. This typically involves adding some data formatting circuitry to the output of the A/D converter and driving the transmit side of a UART (universal asynchronous receiver transmitter). The serial data out of the UART drives the transmit side of a modem. This modem typically uses frequency shift keying at 300 to 1,200 bits per second. The digital output of the A/D also drives a local display, allowing calibration at the transmitter site.

At the control point, the incoming audible FSK data is demodulated back to the original serial digital form. The serial digital data then drives the receive portion of a UART. The parallel output of the UART then drives data decoding and display circuitry. Since digital coding was used in transmitting the output of the A/D to the control point, the control point display will indicate the same as the transmitter site display. If there is an error in the received digital data (which is generally detected by a parity check on each byte received), the display is blanked and an error lamp is lit.

Control Encoding

As mentioned previously, if DC circuits are available, pulses of loop current or voltage may be used to

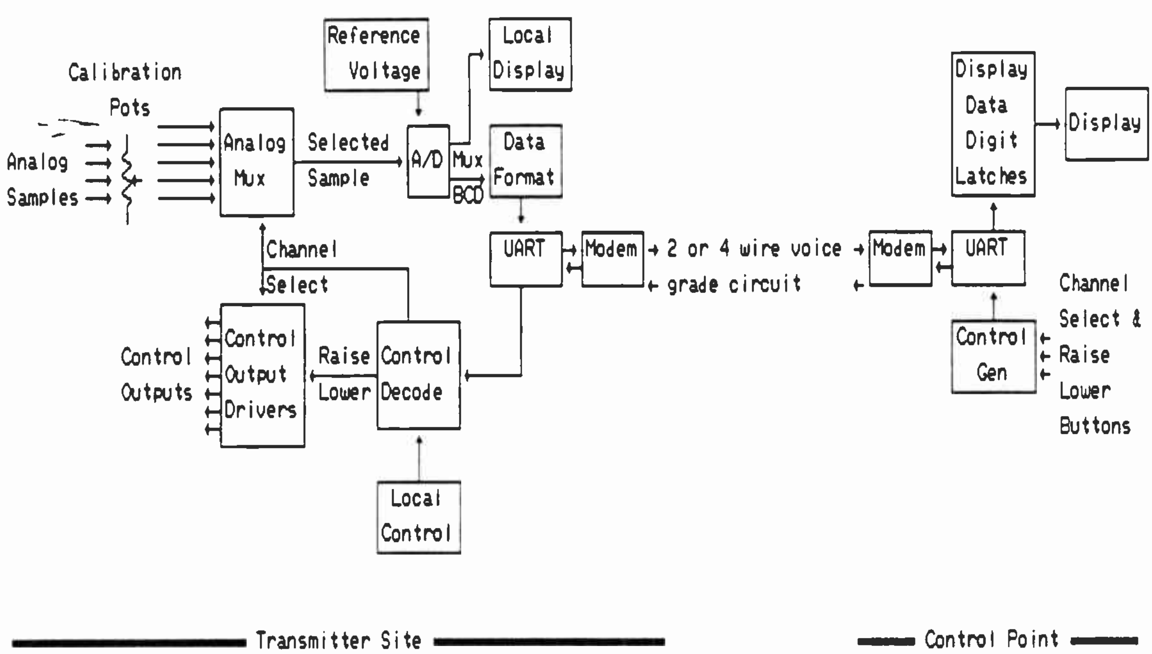


Figure 12. Typical digital remote control without processor.

send control to the transmitter site. Many recent systems use some form of tone encoding.

Frequency Shift Keyed Control

The ready availability of voice grade circuits has encouraged the design of systems utilizing audio tones for the transmission of control signals. The most common technique is the use of frequency shift keying, the same technique described in the A/D with FSK section above. Early systems utilized a telephone dial or integrated circuits emulating a telephone dial to generate a string of pulses. The number of pulses in the string determined what channel was selected. More recent systems utilize binary coding. Since a UART normally sends eight data bits, six could be allocated to channel select, and two to control (one raise, one lower). This allows a simple 64 channel system to be built (since six bits can code decimal numbers 0 to 63). Since an error in a control byte transmitted can have serious consequences (turning transmitters on and off), many systems send the data at least twice and only accept it as error-free if the two transmissions agree.

Typical Nonprocessor-Based Designs

Fig. 11 shows the design of a typical analog remote control system. Digital circuitry is used to encode operator instructions through the channel select and raise/lower buttons. The serialized control data is sent to a frequency shift keyed oscillator (the control modulator). This generates an audio tone at 400 Hz for one digital input level, and 300 Hz for the other digital input level.

This audio control signal may be applied directly to a two-wire or four-wire telephone line (or other audio circuit). It may also be further modulated onto a subcarrier for carriage above programming on an STL transmitter.

At the transmitter site, the FSK audio is demodulated back to the original serial control data stream. This drives the control decode circuitry, which determines what channel the operator selected and whether a raise or lower button was pushed. The channel number is sent to the analog multiplexor (which may be relays or integrated circuit analog multiplexors) and the control output drivers (which may be transistors, integrated circuit output drivers, or relays). When a channel is selected *and* a raise or lower button is pressed, the selected channel control line is enabled.

Sample voltages from transmitters and other site equipment are scaled (multiplied by a constant) by *calibration pots*. During system calibration, these are adjusted so the control point meter reads the same as the meters at the transmitter site.

The outputs of the calibration pots drive the analog multiplexor. This sends the sample voltage selected by the operator to the VCO. Note that a calibration voltage is also shown. Such systems will typically allocate one or two channels of the analog multiplexor to a zero and full scale (occasionally half scale) voltage. These reference voltages allow the operator at the control point to adjust the F/V converter offset (zero) and gain controls to compensate for drift in the VCO and F/V.

The output of the analog multiplexor drives a voltage controlled oscillator (VCO). As described above, the VCO will typically generate 800 Hz for a zero reading and 1200 Hz for full scale. If subaudible metering is used, 20 Hz represents zero and 30 Hz full scale. The 800 Hz to 1200 Hz metering frequency range is well separated from the 300 Hz to 400 Hz control frequency range, allowing both signals to be put on a single two-wire telephone line, yet be easily separated at each end using filters. This combining of two signals onto a single circuit by using different frequencies is a form of frequency division multiplexing.

Finally, the 800 Hz to 1200 Hz metering signal gets back to the F/V converter at the control point. The DC output voltage (proportional to the deviation from the nominal frequency) drives a front panel meter, which the operator reads. By selecting each channel and noting the meter indication, the operator can monitor several transmitter site parameters. Further, by selecting a channel and using the raise and lower controls, various transmitter site controls may be activated.

Nonprocessor Digital Remote Control

Fig. 12 shows the design of a typical digital remote control that does not use a microprocessor. It is quite similar to the analog system described above. The differences are described below.

The calibration pots, analog multiplexor and control output drivers are the same as the analog system. The digital system does not, however, require calibration voltages to be sent to the analog multiplexor. The analog-to-digital converter (A/D) typically accepts a reference voltage (its output is the ratio of the unknown voltage to the reference voltage). The A/D automatically calibrates and zeroes itself during each conversion.

The output of the A/D in most such systems is multiplexed *binary coded decimal* (BCD). The decimal code for each digit is represented as a 4-bit number on four wires. The code for each digit is multiplexed onto these same four wires. Digit Select lines out of the A/D tell the outside circuitry which digit's data is present at any particular time. This data may be latched and decoded to drive LCD or LED displays. It may also be merely decoded (not latched) to drive multiplexed LED displays. The output of the A/D drives a display at the transmitter site, allowing the calibration pots to be adjusted at the transmitter site, since the transmitter site display will exactly track the control point display.

The output of the A/D also drives data formatting circuitry which converts the multiplexed BCD to something suitable for the UART. The transmit side of the UART accepts up to 8 bits of parallel data and a strobe pulse. It then serially sends the data bits (along with some synchronizing bits). Typical data formatting circuitry selectively gates the digit select outputs of the

A/D to the transmit strobe input of the UART, causing each digit of BCD data to be transmitted. The multiplexed BCD is connected directly to four input lines of the UART. The digit select gating allows the A/D data to be sent to the UART without latching the multiplexed BCD data out to parallel BCD (which would take lots of parts). Typically, a couple more bits of the UART are used to transmit a "digit identification code," corresponding to the digit being transmitted. At the receive end, the digit identification code is used to route the incoming BCD data to the proper digit of the display.

The output of the UART is serial digital data. This data is sent to a modem for transmission over a voice grade circuit. Often, Bell 103 modem tones are used. These modem tones allow the transmission of 300 bits per second in each direction over a single two wire voice grade circuit (again, using frequency division multiplexing). Bell 103 uses frequency-shift keying with 1270 Hz representing a 1 and 1070 Hz representing a 0 in one direction. Data going the other direction uses 2225 Hz for 1, 2025 Hz for 0.

The audible metering data is received by the modem at the control point. It is converted back to serial digital data by the receive side of the modem. The UART changes the serial digital data back to the original parallel form (up to 8 data bits plus a strobe). If the previously described encoding scheme was used (4 bits of BCD, 2 bits of digit select), the digit select code bits may be used to drive a *demultiplexor* integrated circuit. This "demux" routes the receive data strobe from the UART to the clock input of a latched display driver for each digit. The latched display drivers for each digit are being fed the received BCD data from the UART, but only capture the data for the appropriate digit when the strobe for that digit is presented.

In a similar manner, front panel controls may be sent to the transmit side of the UART, where the data is serialized for transmission to the transmitter site. The transmit side of the control point modem converts the data to a form suitable for transmission over a voice grade circuit. At the transmitter site, the receive side of the modem converts the control data back to the original serial digital form. The receive side of the UART converts the control data back to the original parallel form.

Microprocessor-Based Remote Control

Each of the techniques described so far have used standard nonprogramable integrated circuits to build a system dedicated to transmitter control. Each system was a digital voltmeter with a remote input select switch and display. Making it do much beyond this is quite difficult. The microprocessor changed it all.

The basic components of the remote control remain the same. There is still an analog multiplexor, an A/D converter, and a modem. Instead of being connected to each other in a manner fixed in hardware, they are all connected as input/output (I/O) devices on a computer bus. The computer processor then addresses each of these devices, telling them what to do and picking up data. The complexity of the system is moved from hardware to software. The use of computer techniques allows features to be added to the system at very low production cost (as compared to providing the same features with dedicated hardware).

Fig. 13 shows a block diagram of a rather large microprocessor based remote control transmitter site unit. Most systems will not contain all these elements, but processor based systems can be designed to allow for substantial expansion. We'll look at each of the blocks of this system.

Processor

The processor executes a computer program. It reads instructions and data from the system memory. It can read data from input ports and write data to output ports (commonly combined into I/O ports). It can make calculations and decisions based on those calculations. It can also write the results of calculations back out to memory or I/O.

The processor has three buses to communicate with the rest of the system. These include the control bus, the address bus, and the data bus. In simple systems (such as a typical remote control), the address bus is driven only by the processor. The processor uses this bus to tell memory and I/O devices what location is to be read or written. In transmitter remote control equipment, the address bus is normally between 11 and 20 bits wide.

The data bus is a bidirectional bus. It is driven by the processor when the processor is writing (sending) data to a memory location or an I/O device. It is driven by the memory or I/O device when the processor is reading data. The data bus in transmitter remote control equipment is generally 8 or 16 bits wide. The processor uses the address bus and the control bus to tell other devices on the data bus whether data is for that particular device, and where in that device the data is supposed to go or be read from.

The control bus varies between different processor types. In Motorola processors (6800, 6802, 6805, 6809, 68000), a +read/−write line has the same timing as the address lines. It is high during a processor read and low during a processor write. These processors then have a common (between read and write) strobe line (called E or phase 2). This line goes high during a read or write. Data is enabled on the leading edge of the strobe and latched and disabled on the trailing edge. Intel processors (and similar ones, such as Zilog) use two separate strobes, one for read and one for write. These have about the same timing as the common strobe in Motorola processors.

Another variation in the control buses deals with how I/O devices are handled. Motorola processors have no input/output instructions. Instead, memory instructions are used for I/O. If the address that data is sent to has an I/O device instead of a memory device, the data is output. The processor doesn't know (or care) where the data is sent. It just puts an address on

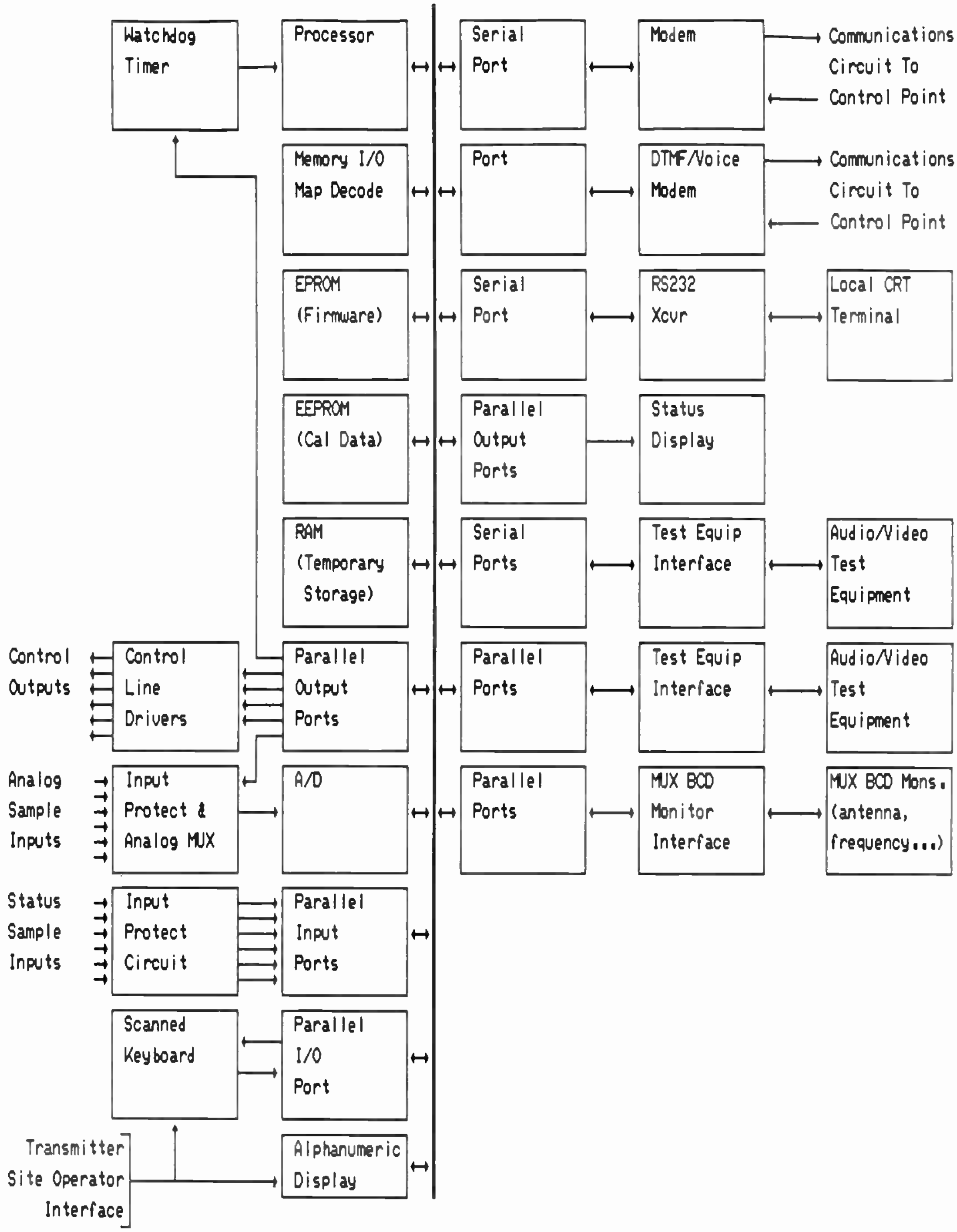


Figure 13. Transmitter site unit of microprocessor based remote control.

the address bus, puts the data on the data bus, then strobes the Enable line. What device (if any) picks up the data is not important to the processor. Similarly, the processor can read data from either memory or I/O. Motorola processors use the same instructions for both. This is called "memory mapped I/O."

Intel processors (and similar) have a "separate I/O map." Separate I/O instructions set up the address bus and put data on the data bus, just as memory write instructions. However, an I/O write strobes the I/O write strobe instead of the memory write strobe, allowing I/O data to go to the I/O device instead on a memory device, which may be at the same address. Systems based on Intel processors may also use memory mapped I/O.

The other main control bus lines are the *interrupt lines*. When an I/O device enables an interrupt line, the processor stores away its current state, then executes an *interrupt service routine* (ISR). The number of interrupt lines varies from processor to processor. When a particular interrupt line is enabled, the processor vectors to a particular address to find the ISR for that device. Several I/O devices may share the same interrupt line. When this is the case, the ISR polls each device using that interrupt line to determine which one generated the interrupt. On completion of the ISR, the processor executes a "return from interrupt" instruction, which restores the processor state to what it was prior to the interrupt. The main program continues on as though nothing happened.

Interrupts are commonly used to allow a processor to appear to do more than one task at a time (multitasking).

Processors commonly used in broadcast transmitter remote control systems include the 6802, 6805, 6809, and 8088. Table 3 gives a brief description of these processors.

Watchdog Timer

Occasionally, something will go wrong in a computer system. When this occurs, the processor must be reset. The reset loads the program counter (or instruction pointer) with the beginning address of the system program, allowing it to start over. System programs typically include an instruction to set a bit of an output port high at one point in the program, then set it low at another point. If this fails to occur within some time period, the watchdog timer "times out" resetting the system. This insures that a system crash will not disable the system.

Memory I/O Map Decode

This portion of the remote control may be done in one chip, or may be spread over several chips. The map decoder recognizes when particular addresses are present on the address bus, then generates a "chip select" for the device that is to be selected.

EPROM

The Erasable Programmable Read-Only Memory (EPROM) holds the program the system executes when the power is first applied. Depending upon the system, this may be the only program the system ever executes. Some systems may hold additional software in EEPROM or RAM. The program that tells the system how to be a remote broadcast transmitter controller (instead of, say, a microwave oven controller) is held in this EPROM. The program is just a series of binary numbers. Under direction of the processor, the EPROM puts the contents of its various memory locations (more than 131,072 8-bit locations are available in today's EPROMS) on the processor bus. The processor reads these instructions, then executes them.

The program in the EPROM is often called "firmware." "Software" is a program that is loaded into memory (typically RAM) to be executed, but does not permanently reside in memory. "Hardware" consists of the physical parts of the system (the chips).

TABLE 3

Characteristics of typical processors used in remote controls.

Characteristic	6802	6805	6809	8088
Accumulator(s)	2 @ 16 bit ea.	1 @ 8 bit	2 @ 8 bit ea.	4 @ 16 bit ea.
Index register(s)	1 @ 16 bit	1 @ 8 bit	2 @ 16 bit ea.	See note.
Stack pointer(s)	16 bit	5 bit	2 @ 16 bit ea. See note.	See note.
Program counter	16 bit	12 bit	16 bit	16 bit
Condition codes (flag) register	8 bit	5 bit	8 bit	16 bit
External data bus	8 bit	?	?	?
External address bus	16 bit	?	?	?
Memory, I/O map	64 kb	4 kb	64 kb	1 Mb

NOTES

6802 Motorola 6800 with internal clock generator and 128 bytes of RAM. Can support larger memory, I/O map with bank switching. Motorola control bus.

6805 Motorola single chip microcontroller. Various chips are available in this family. They typically are a full computer on a chip. The single chip includes the processor core, ROM (or EPROM or EEPROM), RAM and I/O. Most broadcast control applications require more I/O than the standard chips have, so they are operated in an expanded mode, in which the memory and I/O operations are typically moved off the microcontroller chip, often merely using the microcontroller chip as a microprocessor (like the 6802 or 6809). Can support larger memory, I/O map with bank switching

6809 Enhanced version of Motorola 6800. Two 8 bit accumulators that can be combined for 16 bit arithmetic. One 16 bit "user" stack pointer is often used for procedure and function parameter passing, a second 16 bit "hardware" stack pointer is used to hold interrupt and subroutine return address. An 8 bit "direct page" register allows shorter "zero page" instructions to be used anywhere in the memory map, somewhat similar to a segment register. Can support larger memory, I/O map with bank switching

8088 Intel processor used in the IBM PC and XT. The four 16 bit data registers are also accessible as 8 bit half-registers. These data registers serve many general purposes. Some instructions use these registers for special purposes. The address registers are: SI and DI (source and destination index) used in data block moves, SP (the stack pointer) and BP (the base pointer, used to locate parameters and local variables on the stack). To allow access to more than 64 K bytes of memory, the processor includes four 16 bit segment registers. These are offset four bits to the left and added to the address registers, resulting in a 20 bit address, allowing access to 1 M byte of memory. Use of segmented memory is somewhat similar to memory bank selecting (or paging). It can make execution of programs larger than 64 K bytes and access to data structures larger than 64 K difficult, but recent compilers can deal with this by breaking the program into units, each smallerthan 64 K bytes, but whose total may exceed 64 K bytes.

The firmware may be written in one or more of several computer languages. Common languages for transmitter control systems include assembly, Basic, C and Pascal. The higher level languages (all but assembly) are compiled down to assembly language. This is then converted to machine language by an assembler. The resulting machine language (the "object code" or "executable code") is put on the EPROM.

EEPROM

Most systems now include EEPROM (Electrically Erasable Programmable Read-Only Memory) or battery-backed RAM to hold system setup information. EEPROM is similar to EPROM in that it is nonvolatile (it retains the stored data when power is lost). It is similar to RAM since it can be reprogrammed in circuit (EPROM is typically removed from the circuit, erased with ultraviolet light, reprogrammed, then put back in the circuit). Most recent EEPROMs use the same bus interface as static RAMs, easing system design considerably. EEPROMs can be written to, the same as RAM, except that a time delay of 1 mS to 10 mS is required before another read or write to the chip. The system firmware must take the EEPROM write delay into account (it cannot stuff calibration data into the chip at a typical processor speed of a byte per microsecond).

The "ease of programmability" of system EEPROM has its drawbacks. If there is a processor crash, it is easy for the processor to write garbage into the EEPROM, causing loss of calibration data. Most systems, therefore, include extra write protect hardware for the EEPROM. This hardware typically consists of an output port that drives gating circuitry for the write to the EEPROM. Prior to writing data to the EEPROM, the processor must write a specific code to the write protect port. After the data has been sent to the EEPROM, the processor writes another code to the write protect port, disabling further write strobes from reaching the EEPROM. It is unlikely that a crashed processor would write the proper code to the write protect port, thus protecting the EEPROM data until a watchdog timeout resets the system.

The only other limitation on EEPROM is the limited number of write cycles. Each byte of EEPROM may typically be written up to 10,000 times. Since calibration and setup of transmitter control systems is typically done no more than once a week, this limitation is generally not serious.

Some systems use battery-backed CMOS RAM to hold system calibration data. This has several advantages over EEPROM. The RAM can operate at full processor speed (writes at more than one byte per microsecond) and there is no limit on the number of write cycles. The RAM is normally battery-backed with a lithium battery, which has an expected shelf life of ten years. The addition of the battery and other power switching circuitry makes the system design more complex, though some companies sell battery backed RAM subsystems (the memory, battery and all support circuitry contained in a standard 24 pin package, the same package used by RAM or EEPROM). One battery backed RAM subsystem also includes a battery backed real time clock and all associated circuitry, again in the 24 pin package. Use of this chip further simplifies system design, when a real time clock is required. The battery backed RAM suffers the same "too easy to write" problem as EEPROM. Many systems, therefore, include write-protect circuitry for the battery-backed RAM.

The use of nonvolatile memory for system calibration and setup has vastly increased the capabilities of transmitter control systems. A calibration pot may be replaced with 4 bytes of EEPROM. A single byte of EEPROM can replace 8 DIP switches. EEPROMS (or battery backed RAM) in transmitter control systems typically include 2 K to 32 K of nonvolatile memory (where 1 K is 1,024 bytes).

The data included in the nonvolatile memory may include the following:

For each analog channel:

1. A scaling factor, replacing the calibration pot and decimal point positioning logic.
2. An alphanumeric label identifying the parameter (such as ICP for common point current).
3. An alphanumeric units identifier (such as kV for kilovolts).
4. A curve or formula identifier (such as linear or square law).
5. A "sample delay" to allow for the settling time of external monitoring equipment (such as AM antenna monitors).
6. Alarm and action limits on the parameter.
7. Alarm and action codes for each limit. These may adjust transmitter power on one set of limits and alarm on another set of limits.

For each status input:

1. Active high or low (light indicator when input high or low).
2. Latch or momentary. It may be desirable to have some status indicators remain on even after the status input has gone back to the false state. The status indicators are then reset by operator command.
3. Status change alarm or action code. A status change may cause some control command to be issued or enable an audible alarm.

Other site specific data:

1. Communications speed for intersite communications (typically between the transmitter and the transmitter control point).
2. Site identification number. Many systems use the same firmware in control point and transmitter site equipment. The site identification number tells the processor which portions of the firmware apply to this site. This site number is also used to recognize messages intended for this site in multisite systems.
3. In multisite systems, site intercommunication data

such as site access slot timing, highest site number, and site message routing tables.
4. Communications speed for any other serial devices (such as CRT terminals, test equipment, video monitors, etc.).
5. Coding for interface to specific monitors at the site (such as those connected to the remote control through RS232 or other serial links, IEEE488 or other parallel links and multiplexed BCD).
6. Time specific data, such as generating control pulses at specific times (perhaps AM pattern change), or changing limit tables based on time (different limits used day or night).

The variety of data held in the system nonvolatile memory is limited only by the system firmware. The use of specific locations can be changed through system firmware updates (EPROM swaps), allowing features to be added later with no hardware changes. Finally, the packages and pinouts for all "bytewide" memory devices have been standardized. This allows more memory to be added to the system (through addition of chips or the substitution of larger chips in the same sockets) or the memory balance (between EPROM, RAM, and nonvolatile memory) to be adjusted as the system design advances. This is a vast improvement over calibration pots and DIP switches!

RAM

Random access memory (RAM) is used for temporary storage in the processor based system. The two major RAM types are "static" and "dynamic." Static memory remembers data as long as the power is applied. Dynamic memory remembers data for about two milliseconds. After two milliseconds, dynamic memory must be "refreshed." Due to the simplicity of design of systems utilizing static memory, it is commonly used where 64 K bytes of RAM or less is required. In larger systems, the added expense of refresh circuitry is offset by the lower cost of dynamic RAM.

The processor uses RAM to hold "return addresses" so it can find its way back after a subroutine call or interrupt. RAM is also used to form buffers for input/output devices. With appropriate firmware, the system can send large blocks of data (such as a full screen of characters) to a "device driver." This firmware then holds the data in RAM and sends it to the output device (such as a serial port) a byte at a time, as the port is ready to receive more data.

RAM buffers may be used in receiving data from an input port. This allows more flexibility in the system firmware, since it need not continuously poll a specific input device. Without an input buffer or continuous polling, some input data could be lost (the firmware may be dealing with the A/D converter when data arrives from the modem, causing modem data to be lost). The RAM input buffer holds received data until the main firmware is ready to deal with it. Another RAM input buffer is used with the communications port that receives data from the control point. To insure data is received error-free, it is typically sent in "packets." Each packet of data includes error detecting data at the end of the packet. The packet cannot be dealt with until the entire packet is received. The RAM buffer holds received packets until they are complete. The system firmware then determines if the packet is error free. If so, the packet may be acknowledged (depending on system design) and passed on to other sites or firmware at this site. If a packet has an error, a negative acknowledgement is typically passed back to the originating site.

The amazing flexibility of processor based systems is again evident in the uses of system RAM. All these uses (and more) may reside in the same chip. The allocation of that memory may change under program control and with system firmware upgrades (EPROM swap). In a pure hardware (nonprogrammed) system, each buffer would probably be implemented using another chip, considerably adding to hardware complexity.

Parallel Output Ports

A parallel output port may consist of an 8-bit latch. When the processor does a write to the appropriate address, the memory I/O map decoder clocks the latch, causing it to capture the data then on the data bus. This data is held until the processor does another write to the same address. As can be seen from Fig. 13, parallel output ports serve several purposes. Whenever we need to generate a logical level under program control, a parallel output port is used.

One bit of a parallel output port is used to drive the watchdog timer (as described above). Additional lines of a parallel output port (not shown) drive the EEPROM write protect circuit.

Additional port outputs drive the control line drivers. The output port itself is typically an LS (Low-power Schottky transistor logic) or HC (High speed complementary metal oxide semiconductor) compatible device, which has limited driving capabilities. The port outputs drive control line drivers, which may be discrete transistors or peripheral driver chips. These open collector control outputs may be sent directly to the rear panel (through RF filters) to drive transmitter site equipment, or may drive interface relays within or outside the remote control unit. Under firmware control, these control lines may be set high or low or pulsed high or low. How these lines are pulsed or set is typically a combination of user action (command keystrokes), system parameters (automatic control) and user data in the EEPROM. These lines replace the raise and lower control lines of nonprocessor based systems. For convenience, many processor based systems still refer to the control lines by a raise/lower designation.

Additional parallel output ports may be used to drive status indicators (typically LEDs). These may be driven directly by the output port or may be driven by a peripheral driver chip. The various LEDs can be lit or extinguished under processor control. These may reflect the status inputs at this site or another site, or other on/off conditions.

Additional parallel output port lines drive the address lines of the analog multiplexor. These determine which of the many analog samples are routed to the A/D converter.

A/D Converter and Analog Multiplexor

The A/D converter here serves the same purpose as that in the nonprocessor based digital remote control. It converts the analog sample voltage from the analog multiplexor to a digital code. Of the several A/D techniques available, transmitter control systems typically use integrating (dual slope or residue multiplication) or successive approximation. For the same cost, an integrating A/D offers higher resolution (12 bits versus 8) while the successive approximation A/D is faster (50,000 conversions per second versus 20). Although the same A/D used in the nonprocessor based system could be used with a processor, A/Ds are available designed to interface directly with processor buses. The processor may read the A/D to see if a conversion is complete and to pick up the converted data. It may write to the A/D to start a conversion, select an input range, select AC or DC or select an input (some include analog multiplexors).

The A/D is driven by the analog multiplexor. Any of the previously described analog multiplexing techniques may be used. Further, the sample selected by the multiplexor may be directly selected by the operator (who may be interested in only a particular channel at the moment), or may be scanned through all the inputs that are used. When the multiplexor is scanned, the resulting conversions are usually held in RAM, making them immediately available to the firmware and the operator.

Solid-state analog multiplexors typically will not accept voltages higher than the supply voltages (either positive or negative). Many multiplexors are protected from damage due to higher than supply input, however, they will typically have excessive leakage into other channels, causing erroneous readings. The input protect circuitry typically protects the multiplexor from damaging voltages and those voltages that may cause improper operation.

Scanned Keyboard Port

A transmitter site operator needs to interact with the system for setup and calibration. The local display and keyboard serve this purpose. With the addition of four pullup resistors, a single parallel I/O port (4 bits input, 4 output) can be used to scan a 16-key keyboard (typically arranged as four rows and four columns).

Alphanumeric Display

This display is used by the processor to "talk" to the operator. In simpler systems, the display may only display numeric data. In more complex systems, the display typically displays 32 alphanumeric characters (see Fig. 14). The use of a full alphanumeric display allows the processor to explain the operation of the system. Each key may have several uses (reducing the keyboard hardware required). The display allows the processor to inform the operator the meaning of each key at this particular point in the operation of the system.

Many systems use alphanumeric display modules that interface directly to the processor bus. A write by the processor to the display may put a character on the display, position the cursor, clear the display, or various other control functions. The processor can also read data from the display module. It can typically read a status bit indicating whether the display controller is ready for another character (the display controller is generally slower than the system processor), read the cursor position and various other parameters of the display controller.

Alternate Display Techniques

When large amounts of information are to be displayed, it is difficult to beat the economics of a

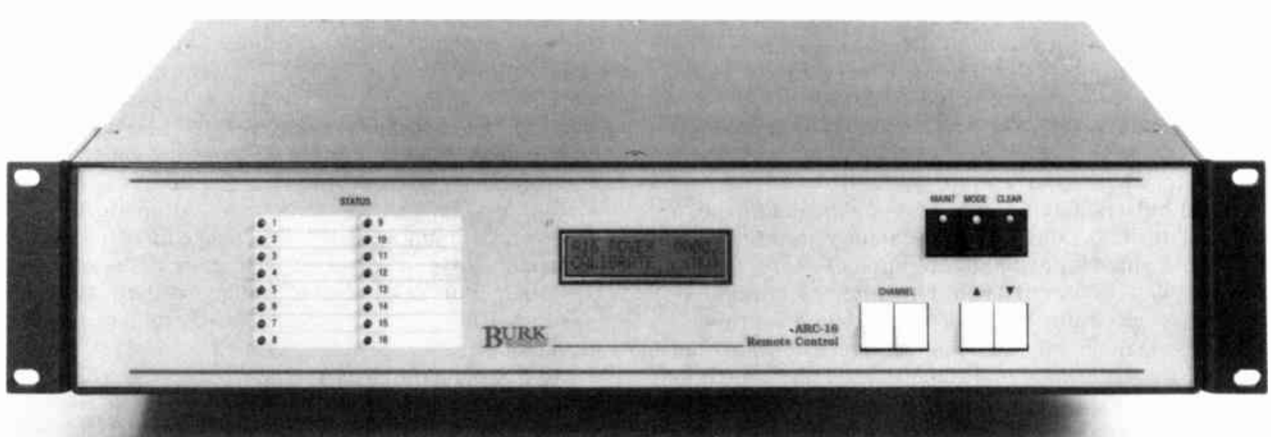


Figure 14. Burke ARC-16.

CRT display. Transmitter and control point operator interaction with the system can be accomplished with a full alphanumeric keyboard and CRT at the site.

Alternate Interaction Techniques

At least one remote control system does all operator interaction (including transmitter site setup and calibration) without a display. After describing systems with lots of lights, buttons and displays, it is interesting to find a system with two LEDs and one switch on the front panel (see Fig. 15). All operator interaction with the system is through DTMF (Touch Tone) and voice. A standard telephone replaces the keyboard and display previously described. The processor sends data to a voice synthesizer. The output of the voice synthesizer is sent to the local (or a distant) telephone. The operator responds to inquiries from the system using the DTMF keyboard on the telephone. This technique has proved quite popular and has been awarded a U.S. patent. Again, however, when large amounts of data are to be displayed, it is hard to top a CRT display. Most systems with DTMF/Voice also offer serial ASCII interfaces (to drive a CRT) as an option.

Digital Monitor Interface

Not all transmitter site parameters can be easily changed to DC voltages. Some monitors digitally derive a parameter, then convert it to an analog sample for interface to the remote control. This is a rather round about method of getting data from one instrument to another (digital data converted to an analog voltage, then digitized again). A direct digital connection would be simpler and more error-free (no D/A, A/D tracking problems). Several methods exist for handling the digital interface. These can be divided into serial and parallel interfaces.

Serial Monitor Interface. Various serial interface standards exist. The most common is EIA-232-D (RS-232) with asynchronous data. Typically, not all the EIA-232-D lines are used. The two signal lines most commonly used are transmit serial data and receive serial data. The remote control processor can send requests to the external monitor by writing to the appropriate serial port in the remote control. The monitor sends data back to the remote control through the same serial port. The processor can read this data from the same serial port.

EIA-232-D allows one port to talk to one device. It also uses "single ended" data transmission (one signal line fed against ground). EIA-422-A (an electrical specification that is part of EIA-449), however, uses balanced lines (two wires driven differentially). Some EIA-422-A applications operate in a "multidrop" mode, allowing several devices to share the same port. When operating in a multidrop mode, each device connected to the port is addressed by the processor (polled) when the processor needs data from that device. This technique allows several devices to share a single serial port, saving on hardware.

Parallel Monitor Interface. Various parallel interfaces exist. For example, a frequency monitor may present its data in parallel BCD, where each digit of its indication is transmitted over four wires. Obviously, this requires a lot of wires and a lot of interface hardware.

An antenna monitor may provide multiplexed BCD data. Four wires carry all the BCD data and additional "digit select" lines identify the digits. An antenna

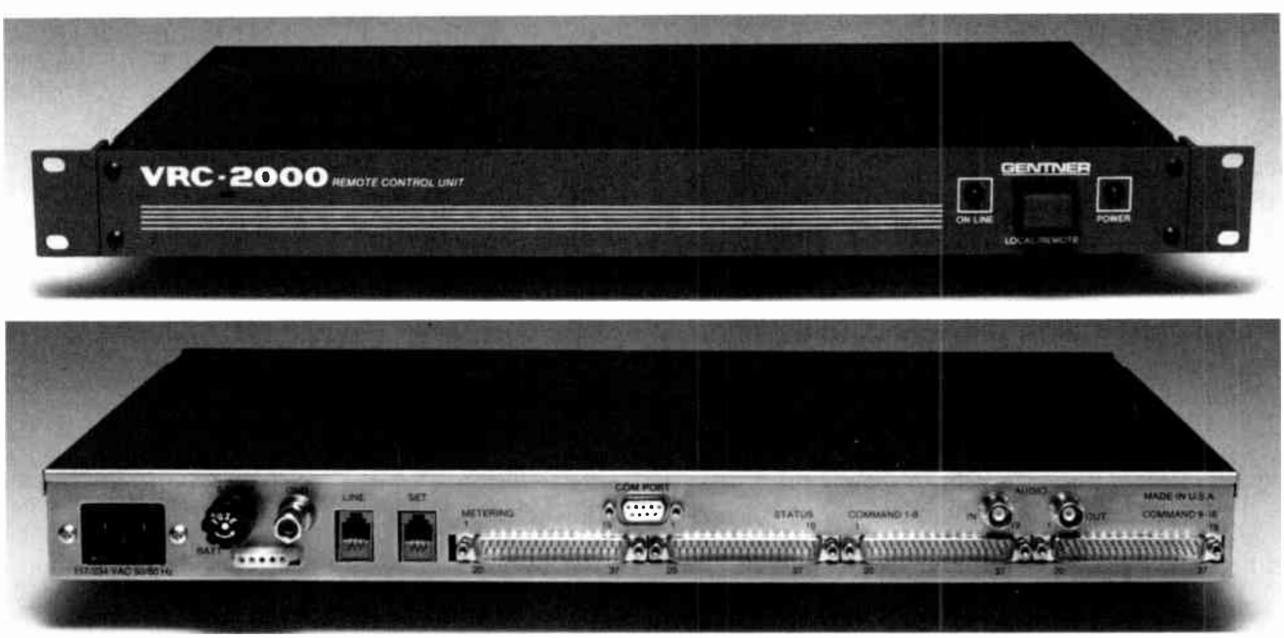


Figure 15. Gentner VRC2000 front and rear panels.

monitor will also require various control signals, such as tower selects. Again, a lot of wires and interface hardware is required.

Further complicating the BCD interface is the wide variety of software that would be required to deal with each instrument.

Hewlett Packard recognized the problem with existing test equipment interfaces and invented the HPIB (Hewlett Packard Interface Bus). This eventually became the IEEE488 bus (also known as the GPIB or General Purpose Interface Bus). The IEEE488 bus standard includes well-defined hardware specifications and addressing specifications (it is a multidrop parallel bus). Some standards also exist for command and response messages. At this writing, the IEEE488 bus is not widely used in broadcast remote control applications. This will probably change.

Intersite Communications

Most transmitter remote control systems utilize low speed (300 to 1,200 bits per second) asynchronous data to communicate between the control point and the transmitter. In Fig. 13, this is shown as a serial port driving a modem. Some systems offer a variety of modems that plug into the system. One modem may be optimized for two wire full voice grade circuits. Another may be optimized for four wire voice grade circuits. Another may directly generate signals suitable for use as subcarriers on STL and broadcast transmitters. It is also possible to bring out a serial digital interface to communicate with the control point. The station can then transmit the data digitally to the control point. This digital transmission may involve the use of external modems, a fiber optic link, or multiplexing the data into a higher speed digital link.

The data that is sent to the control point must, of course, be of a form that the control point equipment can deal with. This data may be ASCII data ready for immediate screen display (directly driving a CRT terminal), or may be a stream of site data where the control point equipment does error checking and extracts the required data for display. In each case, the data and its encoding (modem type) is optimized for the control point equipment and the communications channel. One widely available piece of control point equipment is the standard Touch Tone (DTMF) telephone. Since this does not have a display, data is sent to it using voice synthesis. Commands from the phone are sent to the transmitter site using the DTMF tones. The voice synthesizer/DTMF detector can be thought of as another modem that takes digital data and formats it for transmission over a voice grade circuit.

Summarizing Transmitter Site Hardware

Fig. 13 shows many of the elements that *may* be present in the transmitter site unit of a transmitter control system. Some systems may have more I/O devices "hung on the bus," others may have less. There are microcontrollers available (such as the 6801, 6804, 6805, 8048, and 8051) that combine many of the hardware requirements for such a system on one chip. Some microcontrollers contain a processor, EPROM, EEPROM, A/D converter, analog multiplexor, serial port, and parallel I/O on one chip. If one of these chips can meet system requirements, a very low chip count system can be built. However, once the design does not fit on a single chip, the system design may become more complex using microcontrollers in their "expanded" mode as opposed to using a standard processor.

Control Point Hardware Design

The control point equipment is to display data gathered by the transmitter site equipment (or announce it, in a voice/DTMF system), and transmit operator commands back to the transmitter site. In most cases (again, excepting voice/DTMF systems), the control point hardware is another microprocessor based system using many of the same components as the transmitter site equipment. It has been fairly standard practice to use the same hardware for control point and transmitter site equipment, generally eliminating the analog inputs and control outputs and some other special I/O at the control point. The control point equipment would include the processor, memory, a display, a keyboard, status display, and modem (for communicating with the transmitter site). If the system were to drive a control point CRT terminal, typically the intelligence for driving the terminal would be located in the control point end of the remote control system. Since the CRT terminal is also a microprocessor-based system that displays received data and transmits operator keystrokes (which just happens to be the job of the control point end of a remote control system!), it appears possible to replace the control point end of the system with a standard CRT terminal and a modem. This is the basis of some more recent remote control systems. The use of a standard CRT terminal as the control point equipment reduces the system cost. However, most standard CRT terminals do not include a watchdog timer, allowing the terminal processor to crash in a power dip. An external watchdog timer may be added, if desired.

System Firmware Considerations

The firmware design of a minimum system (one emulating the nonprocessor based system of Fig. 12) can be implemented relatively easily. However, in an effort to reduce parts count (and manufacturing costs), there is a trend towards moving as much of the design as possible from hardware to software. For example, the calibration pots are replaced with scaling factors and multiplication routines. Analog multipliers (power to linear converters) are replaced with software. Once the system includes a processor, powerful features can be added with minimal manufacturing cost (though the design cost may be the same or higher than designing a similar feature in hardware). These added features may include automatic limit checking, transmitter adjustment, transmitter logging, full screen display, etc. A typical system has a firmware assembly language listing that is 384 pages long. The resulting object code occupies 22 Kbytes of EPROM. We'll look at a few areas of system firmware, including intersite communications, limit checking, automatic control and operator interface.

Intersite Communications. In a simple system with a single control point and a single transmitter site, data sent by the transmitter site to the control point is displayed and operator keystrokes are sent to the transmitter site for appropriate action. This communications is normally carried out using simple data packet techniques so that any erroneous data is ignored. If an error-free data link can be established, the display (perhaps a CRT terminal) can be connected directly to the link. Such an error-free link can be established by moving the error correction responsibility from the control point box to the modem. In such a system, it is typical to use a standard QAM (quadrature amplitude modulation) modem with MNP (Microcom Network Protocol) error correction and data compression built into the modem.

It gets more interesting when there are multiple transmitter sites to deal with. If dial-up lines are used, the telephone company handles all the site selection and data contention problems. DTMF/voice systems may use a single telephone with an automatic dialer. If a site has a problem to report, it calls it in. If the operator wants to poll a site, a single key on the dialer is pressed, and the operator is shortly in communications with that site.

A similar approach may be used with CRT terminal based dial-up systems. In such systems, programmable function keys on the terminal may be used to generate modem dialing codes for each of the transmitter sites. If a transmitter site calls in with an alarm report, the modem answers and puts the site information on the screen. The dial-up phone system is handling all the data routing problems. Further, only the data from one site at a time is displayed. If the CRT terminal is replaced with a computer and a modem, it can gather the last reported data from each site and put together a display showing a summary of the site information.

If dedicated circuits are used to interconnect the sites, it is the remote control's responsibility to route the data to the appropriate site. Two major methods exist. The first utilizes half duplex modems and is similar in design to local area networks (LANs) and packet radio data transmission. The second approach uses full time digital links (which may include modems to allow use of voice grade or other analog circuits) and is more similar to the switched telephone network or wire line (or fiber optic) packet data networks.

Multisite Data Packet Communications. In multisite communications, a single link is used to carry messages that may originate from any of several sites and be sent to any of several sites. To keep the messages straight, each one is put in a packet that includes the number of the originating site and the destination site. This information is included in the message header. The header consists of a flag identifying the beginning of a packet, the address information, and a message length count. A flag consists of a sequence of bits or bytes that is unique and guaranteed to not occur in the data stream (which would cause a false begin of header detection).

The structure of the header varies system to system, but will often include the destination address (site number), the return address (site number), and a byte count for the message being sent in the packet. The destination address is used by each site in determining whether to accept a message, route the message to another site or ignore the message. The return address is used by the destination site to send back any requested data, a message acknowledgement, or a negative acknowledgement. The message byte count is used to determine where the end of the packet is. An alternative method is to use another flag sequence to identify the end of the packet. The header may also include a sequence number. As each packet is transmitted, the sequence number is incremented (and rolls over, back to zero). The receive site can then acknowledge a whole sequence of packets instead of each one individually. This generally speeds up the data transmission (a bunch of packets can be sent in one burst instead of sending them one at a time and waiting for acknowledgements). The actual contents of the data field of the packet may have varying meaning. Often one byte of the packet will identify the type of data contained in the packet. This packet-type byte is then used to identify the contents of various fields of the data packet. The data packet ends with a frame check sequence (FCS). This is generally an 8-bit or 16-bit error detecting sequence, often a checksum, cyclic redundancy check, or some similar error detection system.

Half Duplex Multisite Communications. These multisite systems set up a single analog communications link among all sites in the system, such that all sites can "hear" all other sites. A message transmitted by any site is heard by all other sites in the system. This is similar to many LAN and packet radio systems. Each site receives all the packets that are exchanged in the system, but ignores all packets except those where the destination address matches the site number of that site. This is a fast communications system, since the message immediately goes to the desired site (and all others). Some method of keeping the sites from interfering with each other is required.

Each site operates "half duplex," keeping its modem carrier off line unless it has some data to transmit. Each site is continuously listening to the line. As such, it is somewhat similar to a tristate bus in a computer system (each device listens to the bus, but only the addressed device puts data on the bus). A variety of techniques are available to coordinate the data transmission on the channel. These include "master/slave," "token passing," "Aloha," "Carrier Sense Multiple Access with Collision Detection" (CSMA/CD), and "minislotted access."

Full Duplex Multisite Communications. Another approach to routing data packets among the several sites in a system is based on a continuous two-way data link linking each site to at least one other site in the system (a "mesh" topology). When a packet is received from any of the incoming links, the packet is first checked for errors. If the packet is error-free, the destination address is examined. If the destination address matches this site, the message is passed to the remainder of the firmware at this site (perhaps requesting a parameter from the A/D handler, etc.). If the destination address does not match this site, a routing table (typically in EEPROM) is consulted. Based on the destination address, the routing table information and circuit loading, the received packet is placed in the output buffer for a particular output port, which drives a communications link to another site in the system. The routing table holds one or more port numbers for each site in the system. Where more than one port number is assigned, the system makes a decision as to which port is to be used based on system traffic. If the first choice port (the most direct path) has a lot of data in its output buffer, the second choice port may be used, spreading the traffic for optimum transmission speed (is it faster to use the indirect route or get in line for the busy port?). On receiving a packet, a site does a similar evaluation. Note that a particular communications link will be carrying data destined for many different sites. A properly received packet that needs to be relayed to another site is put into a FIFO buffer for transmission to that site. Since each message must be fully received (for error detection) before it is passed on, the delay in getting a message from one end of the system will be more than in the previously described half duplex systems. This delay can be reduced by using smaller packets, though the header overhead then reduces overall system efficiency. This routed system does result in higher overall throughput since the system may carry more than one message at a time, while half duplex systems carry only one message at a time. Further, there are no delays due to carrier bring up, shut down, and carrier sense delays. Each site acts as a data switch and multiplexor along with its normal data acquisition and control activities. Each message is fully demodulated, error checked, retimed and remodulated when sent on to another site. This "digital repeating" prevents the accumulation of noise, distortion, and errors in the data.

Parameter Limit Checking. A simple transmitter control system (such as the nonprocessor based systems) typically extends all the transmitter site meter indications to the control point, but makes no evaluations of these indications. Once a processor is added to the system, parameter limit checking can be accomplished with no additional hardware. After each parameter (or the whole group of parameters) is acquired by the analog to digital converter, each reading is compared with various limits. If a reading is above or below a specified limit, a specified action is taken. The software to accomplish this limit checking may reside at any point in the system (such as the transmitter site or the control point). Further, the limit checking may reside in the remote control system or may be in a computer that gathers data from the remote control system (division of software responsibilities is discussed later).

Remote control equipment is not legally required to have limit alarms. If, however, the station utilizes automatic logging (often done by the remote control system), then Section 73.1820(b)(4) requires the system to include an aural alarm. Early systems used analog comparator integrated circuits to detect a sample going over an alarm limit. A potentiometer adjusted the threshold voltage for the comparator. Current systems handle the problem in software (although many automatic power control systems in transmitters themselves utilize a pair of analog comparators).

A typical remote control system will assign four limits to each analog channel. These limits divide the total possible range of the parameter into five bands. We may assign limit codes to these bands as shown below:

Limit Code	Description
+2	Above upper upper limit
+1	Below upper upper limit, but above upper limit
0	Inside all limits
−1	Below lower limit, but above lower lower limit
−2	Below lower lower limit

A system may allow for different limit tables depending upon the state of the transmitter site. For example, different limits will apply depending on whether an AM antenna system is in day or night pattern. Different limits will apply if a different transmitter is on the air. Different limits will apply if no transmitter is on the air (after station sign off, we don't care if the antenna current is below the licensed minimum!). A system could evaluate the state of the transmitter site and determine which limit table to use based on the state discovered. This may, however, cause some problems. If a transmitter goes off the air and the "Determine State" software determines that the current state is "off air," it may decide to use the "off air" limit table (ignoring all alarms except tower lights, burglar alarm, and fire alarm). For this reason, a system will often change state based only on operator commands or automatic commands (scheduled pattern changes, etc.).

Associated with each limit code and state is a limit action. The limit action tells the system software what to do when this limit code occurs. For example, a limit code of +1 may cause the system to trim down the transmitter power while a limit code of +2 may shut down the transmitter (if the transmitter got to this limit code, the power trim must have not been successful). A system will also typically store the previous limit code for each parameter so that transitions can be detected. A system should log the transition of a limit boundary rather than continuously printing that the power is too low. One log entry when the power went out of limits and another when the power went back in is generally sufficient. Of course, one of the limit actions will be to alert the operator of the limit transition (alarm on limit transition). A system may highlight the out-of-tolerance parameter on a display, display a user specified alarm message, or some other action. In some systems the out of tolerance parameter or alarm message will flash and an audible alarm will sound until the operator acknowledges the alarm by pressing a key. Once the alarm is acknowledged, the audible alarm and flash stop, but the alarm message remains visible as long as the parameter is out of tolerance.

Time of Day Actions. If the time of day (along with day of week and date) are stored in the system as a number (generally integers), then the same limit checking routines used for analog parameters can be used for time of day automatic control. If a lower limit is set to the sunrise time, and the upper limit is set to sunset time, actions (pattern/power changes) can be generated by the limit checking software. Many systems allow for time of day programming using this or similar techniques.

Operator Interface. A wide variety of operator interface options are available. These range from the numeric display of a single metering channel along with channel select and raise/lower buttons through DTMF/voice to CRT screen displays. The single channel display is traditional (processor based systems emulating nonprocessor based systems) and quite adequate where a small number of parameters are to be monitored.

The DTMF/voice interface has proven quite popular due to the ready availability of Touch Tone phones. A station technician already has a control point terminal at home and perhaps even in the car. With a large number of parameters to be monitored, the interface can become awkward, but is certainly better than no interface at all.

Analog Parameter Display. When a large number of parameters are to be displayed, it is difficult to beat a CRT screen. Fig. 16 shows a CRT display of analog parameters gathered from an FM transmitter site. Out of tolerance parameters are shown in inverse video. Full label and unit designators are included for each channel. An operator can view this one screen and have a good idea of the overall condition of the transmitter site.

Alphabetic Control Menu. Fig. 17 shows a possible control menu. Instead of a channel select and raise/lower controls, the operator is given a display of all possible choices. The display combines the functions of a control menu and a status display. Note that on line A, "Gates Exciter A" is in bright video while exciter B is in dim video. The bright video indicates the current state of a status line which is driven by an RF relay (the exciter select relay). Pressing "A" or "a" on the terminal keyboard will send the appropriate control signals to change exciters. When the status line changes, reflecting the exciter change, the screen is updated to show the new state of the system. Pressing the "A" key again will toggle the selected exciter back, again updating the display. The use of a single menu letter for two functions (exciter A select and exciter B select) allows more data to be placed on the screen (higher screen density) when compared with providing a separate letter for each. If a separate letter were used for each function, half the time one or the other would be an invalid choice. Selecting exciter A when it is already selected is not valid. The "toggle" command provides the operator only with valid choices.

The use of alphabetic keys (as opposed to numeric keys or function keys) has several advantages. Numeric keys could be used, but then the system would be limited to 10 menu choices (0 to 9) for each menu. Since a typical screen may have 24 or 25 lines, this is often not adequate. Function keys generate a variety of codes that vary terminal to terminal. This lack of standardization makes it difficult to design a system that will work with a wide variety of terminals. Further, most terminals have fewer function keys than alphabetic keys.

It is often desirable to get a command confirmation from the operator before a command is executed. If

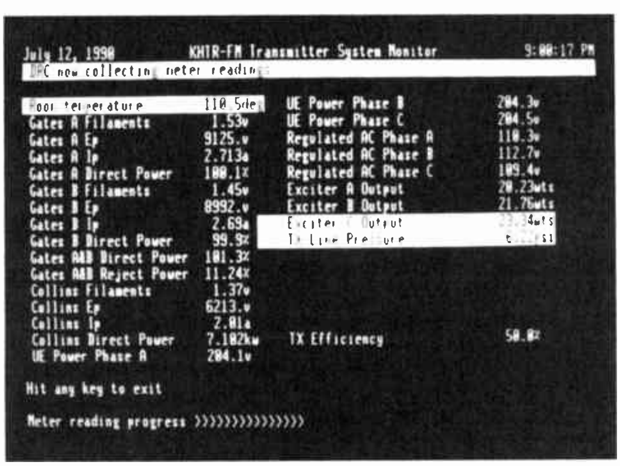


Figure 16. Analog reading CRT display.

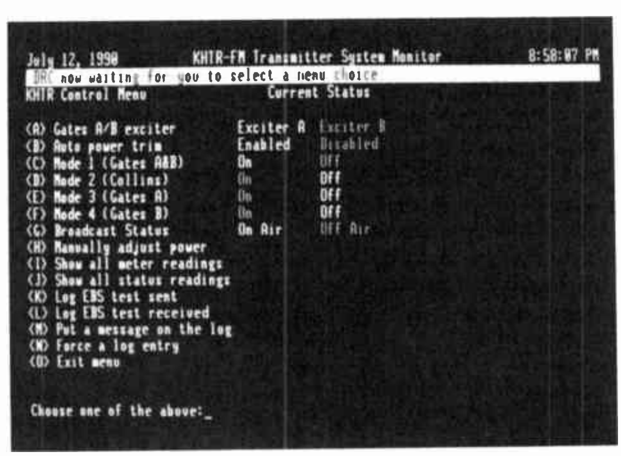


Figure 17. CRT control menu.

the operator were to select item G (Broadcast Status On Air/Off Air), the terminal beeps and puts up a message, "Are you sure you want the station off (Y/N)?" The operator is given an opportunity to respond. A lack of response causes the system to time out and go back to normal operation without taking any action.

Touch Screen Control. A terminal alphanumeric keyboard looks too much like a computer for some transmitter operators. A Touch Screen display combines the advantages of a CRT display with a more "user friendly" interface.

Fig. 18 shows a typical Touch Screen transmitter control screen. This is similar to the alphabetic menu system described above, except that the operator touches "targets" on the screen instead of keys on a keyboard. The screen emulates a panel of backlit pushbutton switches. Note that the "ON" targets are bright, indicating the filaments, plates, and auto power control are on. If the plate voltage off target is pressed, the "Command Confirm" target starts flashing and the "now line" (the inverse video text on line 3) prompts the operator to confirm the plates off command. If the command is confirmed, the command is executed. Otherwise, the command confirm prompt times out and the system takes no action.

Use of Screen Graphics. Fig. 19 is an example of how CRT screen graphics can be used to display the status of a transmitter site. The remote control is evaluating about 20 status inputs to draw the block diagram. It is checking on/off status for each piece of equipment to determine whether to write the text for that block in bright or dim video. It is evaluating the position of each coax switch (two status lines per switch) to draw the lines interconnecting the blocks. Such a block diagram is easier to interpret than the 20 individual status indicators.

Some transmitter sites have a panel of LEDs showing the settings of all the RF relays. Often this panel is engraved with a block diagram of the site and the LEDs indicate the setting of each relay in the block diagram. This screen is emulating such a display panel, again moving a hardware problem to software.

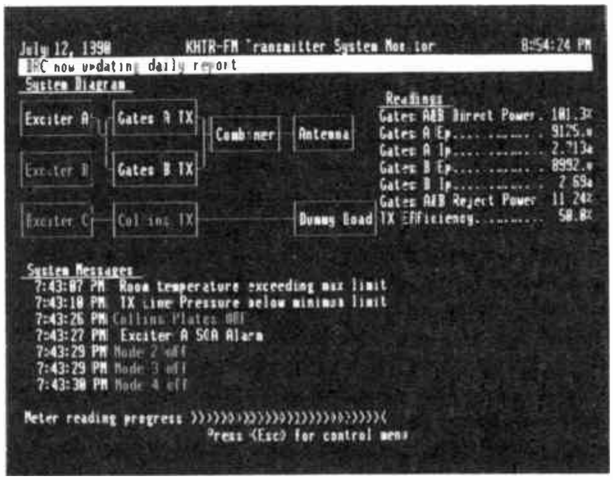


Figure 19. CRT block diagram display.

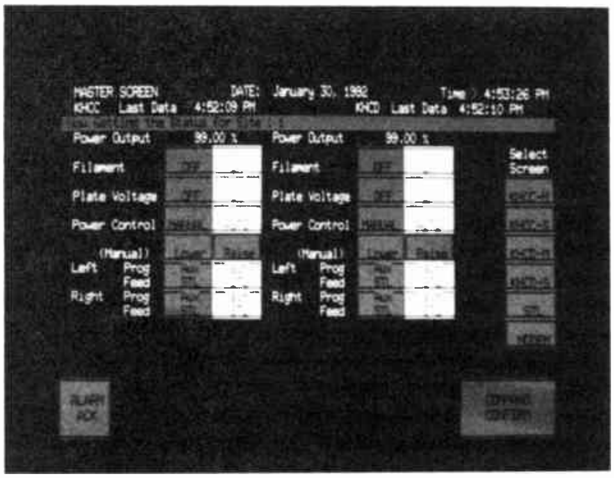


Figure 18. Touch screen control menu.

Automatic Logging. Fig. 20 shows a log generated by the automatic logging software of a remote control system. Such systems typically log all measured and calculated parameters on a scheduled basis and on a significant change in any parameter. In addition, any automatic or operator initiated commands are typically logged, giving an "audit trail" of system operation. The automatic logging software may also accumulate statistics on the system, reporting minimum, maximum, and average for each parameter as part of a daily report.

Programming The System. Prior to processor based remote controls, system setup and calibration consisted of adjusting the calibration pots. As features are added to processor based systems, more and more user input is required to describe how the system is to act. Keeping the system flexible while not overwhelming the user is a challenge for system designers. As with any other computer program, various methods of user data entry exist. Most systems utilize a series of menus, data tables, or an adaptation of a standard computer language.

TRANSMITTER CONTROL DIVISION OF RESPONSIBILITY

Automatic transmitter control systems are often broken into several hardware boxes, each responsible for a portion of the job. Some of the control responsibility reside in the transmitter. Some reside in a remote control transmitter site unit. Some reside in a remote control control point unit. Some may reside in a control point computer. We'll consider some of the factors used in determining how this responsibility should be divided.

KIIS AM/FM Transmitter Log For Wednesday, August 6, 1986, All Time PDT Page 1

Time	AM Readings									FM Readings				Comments
	EP	IP	IA	ICP	Power	DA Parameter	T1	T2	T3	EP	IP	In. TPO	Reflc.	
1:00:18 AM	5153.	1.34		10.1	102.0	Phase (deg)	51.3	0.0	140.7	6302.	1.95	96.1	1.1	
						Loop Current	47.4	99.7	46.8					
3:00:43 AM	5153.	1.34		10.1	102.2	Phase (deg)	51.2	0.0	140.7	6324.	1.96	97.0	1.1	
						Loop Current	46.8	98.7	46.7					
5:00:23 AM	5170.	1.34		10.1	102.8	Phase (deg)	51.2	0.0	140.6	6317.	1.96	97.0	1.1	
						Loop Current	47.2	99.5	46.4					
														AM Power % out of limits: 5:04:52 AM
5:05:09 AM	5142.	1.35		10.2	103.0	Phase (deg)	51.3	0.0	140.7	6280.	1.94	95.6	1.1	
						Loop Current	47.8	100.3	47.0					
														AM Power % back in limits 5:05:46 AM
5:06:06 AM	5184.	1.35		9.9	99.6	Phase (deg)	51.3	0.0	140.7	6280.	1.95	95.7	1.1	
						Loop Current	47.5	100.1	46.7					
														Trimming AM power down: 6:14:39 AM Changed to day pattern: 6:17:36 AM
6:17:56 AM	4992.	1.31	11.6		102.5					6273.	1.94	95.2	1.1	
														Trimming AM power down: 6:25:07 AM AM Power % out of limits: 6:25:54 AM Trimming AM power down: 6:25:59 AM
6:26:08 AM	4973.	1.32	11.7		102.5					6295.	1.95	96.1	1.1	
														AM Power % back in limits: 6:26:45 AM
6:26:56 AM	4916.	1.30	11.6		102.9					6288.	1.95	96.1	1.1	
8:00:33 AM	4931.	1.29	11.4		98.9					6236.	1.92	93.6	1.1	
10:00:43 AM	4940.	1.30	11.4		98.9					6244.	1.92	93.7	1.1	
12:00:04 PM	4939.	1.30	11.4		99.3					6295.	1.93	95.0	1.1	
2:53:58 PM	4934.	1.28	11.3		97.9					6266.	1.91	93.8	1.1	
4:00:08 PM	4939.	1.29	11.3		98.5					6229.	1.90	92.9	1.1	
6:00:43 PM	4942.	1.30	11.4		99.3					6222.	1.90	92.8	1.1	
														Changed to night pattern: 7:47:23 PM
7:47:52 PM	4997.	1.31		9.7	95.9	Phase (deg)	51.1	0.0	140.4	6200.	1.90	92.3	1.1	
						Loop Current	46.3	97.7	46.0					
														Trimming AM power up: 8:02:40 PM
9:00:22 PM	5083.	1.34		10.1	102.7	Phase (deg)	51.1	0.0	140.7	6244.	1.92	93.8	1.1	
						Loop Current	46.5	99.0	46.5					
														Tower lights OK at 10:00:16 PM
11:00:23 PM	5087.	1.33		10.0	100.3	Phase (deg)	51.2	0.0	140.7	6236.	1.92	93.7	1.1	
						Loop Current	47.0	98.8	46.6					

Figure 20. Sample Transmitter Operating Log generated by remote control.

Extended Voltmeter

The simplest remote control systems utilize an "extended voltmeter" for telemetry. In response to operator requests, the transmitter site analog multiplexor connects the voltmeter input to the various parameter sample. The operator logs the resulting readings, does limit checking and makes required adjustments. The transmitter site equipment has minimal responsibility.

Transmitter Site Polling

Automatic equipment may be placed at the control point to poll the various transmitter parameter samples (just as the operator did). This equipment may then log the readings, do limit checking and make adjustments. The "intelligence" of the system is at the transmitter control point and may be in the remote control unit itself, or in an external computer (through a computer interface on the remote control).

Continuous Telemetry Transmission

Another design approach calls for the transmitter site equipment to continuously scan all the parameter samples and send the resulting data to the control point. This data is then held in memory by the control point. As new data for a specific parameter is received, the old data is overwritten. The operator may then select (through a standard "channel select switch") which parameter is to be displayed. The data is immediately pulled from local RAM. If a CRT display is used at the control point, the screen is continuously updated with the received data, giving a full screen display of all transmitter site data.

Exception Reporting

In the previously described systems, telemetry data is continuously transmitted, even if it duplicates data that was sent a few seconds earlier. Another approach is to send only the data that has changed. This requires the addition of identification information to each piece of data, but will generally result in a faster update time. This also allows the introduction of noncontinuous data communications between the control point and the transmitter. The typical noncontinuous data communications link utilized is the dial-up telephone. The telephone company charges for the line during the time it is used. A dedicated circuit is assumed to be used at all times. The "break even" point between dial-up and dedicated voice grade circuits is typically at two to three hours per day.

By placing some of the "intelligence" at the transmitter site, the amount of information transmitted to the control point can be reduced considerably. This allows more sites to share a dedicated circuit (due to lesser loading of the circuit), or allows the use of noncontinuous circuits, such as dial-up circuits.

A transmitter site may transmit the status of the site due to any of several conditions. These include the existence of an alarm condition, a change in a parameter, or a scheduled report.

Systems that have dedicated circuits to the control point may immediately transmit the current state of the site based on any of these conditions. Systems that utilize dial-up circuits will typically originate a call only on an alarm condition. These systems may buffer all the parameter changes. This buffer is "dumped" to the control point on it filling, on an alarm condition (requiring immediate notification of the operator), and on a scheduled basis. One multisite dial-up system utilized by a television network has each transmitter site do a buffer dump each morning between 6:00 and 7:00. If all goes well, this may be the only report from the site during that day. It occurs when the telephone charges are the lowest. It also confirms that each site is operating properly as the broadcast day begins.

Data Analysis

Each week, the designated chief operator of a broadcast station is required to review the station logs and certify that the station operation was in compliance with FCC requirements [FCC Rule §73.1870(c)(3)]. In addition, the chief operator will typically look for drift in parameters in an effort to anticipate problems at the transmitter site. Statistics gathering may be done by hand (looking at manually or automatically printed logs), by the remote control (especially programmable systems), or by a separate computer.

Some remote control installations use a standard computer(s) for the control point terminal. This computer emulates a standard CRT terminal with additional features. As data is received from a transmitter site, some of the data is sent to the screen for immediate display, some is sent to a printer for immediate review, and some is sent to disk for later analysis. Through the use of multisession communications software, the computer may communicate with several transmitter sites simultaneously (over dial-up lines). The data sent to disk is available for immediate or later analysis by standard data base management programs running on the same computer or another computer (on a local area network).

Once the control point terminal is a computer (instead of a terminal), different data techniques may be appropriate. The ability to "talk to a terminal" would be retained to allow communications with portable terminals (lap top computers), however, communications with the fixed control point(s) is accomplished in a machine readable form. The control point computer runs several programs simultaneously. For each incoming port (modem port), a program handles incoming modem data. That data is in a form directly suitable for use with a data base program. As each *record* (the entire state of the transmitter site at a specific time) is received, it is appended to the existing data base file. The *main program* running on the computer is the data base analysis program. This program presents various reports to the operator based on the data stored in the data base. This report may be a summary of the conditions of all sites based on their last reported data. With the data base analysis program, the operator can also choose any site for closer inspection. The screen report generation of the data base analysis program may be able to put up screens that use a combination of color graphics, text, and the data pulled from the data base file records. The operator can step forwards and backwards through time (stepping through the data base records from the chosen site), watching a particular piece of equipment fail. Each site may have several *screen descriptor files* that present the data from the site in different formats. One screen may present an overall block diagram of the site, while others show more detailed looks at each piece of equipment. While a particular site is displayed, the operator can press a "hot key" (a key that activates a "pop up" or Terminate and Stay Resident program) that establishes communications with the site. This immediately dumps the buffer of the transmitter site equipment to the computer, updating the data base file and, if selected, updating the screen. It also enables direct control of the site. Control menus on the screen become active. If the operator selects a command, it is sent to the transmitter site and executed. The resulting changes in the transmitter site status are immediately reported back, updating the data base file and the screen.

Data Reporting

Various devices may be used at the control point to make the transmitter site data available to the operator. These include a visual display (whether a small single channel display or a CRT display), a printer, a DTMF/voice interface (through a telephone), and a fax machine.

Visual displays show the last data received. Printers show all the data received, forming a printed log.

DTMF/voice systems report the current data and alarm conditions. Systems with visual displays often include printers to generate a printed log. DTMF/voice systems typically do not generate a printed log (the Touch Tone phone doesn't have a printer port). However, many stations now have facsimile machines connected to the dial up telephone network. These may be utilized by DTMF/voice systems to generate a printed log. The transmitter site equipment buffers all parameter changes, alarm conditions, operator alarm acknowledgements, automatic adjustments and operator adjustments. After each day is completed (some time after midnight), the transmitter site unit calls the fax machine and prints the log. If several transmitter sites are reporting, these calls may be scheduled to avoid conflicts.

The use of the fax machine allows a station to get an automatically printed log using existing equipment.

REMOTE DIAGNOSTICS

The systems discussed so far have merely presented the transmitter site data to an operator. They have done minimal analysis of the data. One broadcast software company has developed transmitter analysis software. Through extensive analysis of the failure modes of transmitters, the system is able to determine what component most likely failed to cause the existing transmitter parameters. This approach is an automated "logic tree," similar to the flow charts equipment manufacturers supply to aid in the troubleshooting of equipment.

COMMUNICATIONS CIRCUITS

Various circuits are available to connect a control point with transmitter sites. They include the following:

- Telco "metallic" pair
- Telco voice grade pair
- Telco dial-up voice grade
- Telco digital circuits
- STL subcarrier (voice grade or digital)
- FM subcarrier
- TV aural subcarrier
- TV vertical blanking interval
- AM subaudible
- Dedicated radio link (P channel)
- FM squared satellite
- SCPC satellite
- Satellite VSAT

Telco Metallic Pair

The telephone system is designed for voice communications. Individual pairs of wires connect each telephone instrument to the local central office. Once the pair is terminated in the central office, the voice is generally digitized (8 bits per sample, 8,000 samples per second) for switching and transmission to other central offices. The links between central offices are now high speed digital lines. As such, it is very difficult to get long distance "metallic" pairs (DC continuity end-to-end). Within the coverage area of a single central office, however, these lines are sometimes available. They may be ordered as a "3002 Local Area Data channel" line with DC continuity. These lines are suitable for DC metering and control. They can also be used to transmit digital telemetry and control by using "short haul modems" or "line drivers." These condition the data for transmission over the long distance balanced lines where there is likely to be substantial common mode voltage. A low-cost, short haul modem can typically transmit 9,600 bits per second data 2.2 miles. At 1,200 bps, this increases to 6 miles. These short haul modems normally require a four-wire circuit (one pair in each direction), though, with appropriate drivers and software, it is possible to set up a single half duplex "current loop" circuit that allows communications one direction at a time over the single pair.

Telco Voice Grade Pair

If a 3002 Local Area Data channel goes through more than one central office, it will probably not have DC continuity. The link between central offices was multiplexed onto wire pairs, coaxial cable, microwave, or optical fiber. The telephone company can multiplex 24 voice grade circuits onto a single pair, so it is not economical for them to allocate a pair to your one voice grade circuit.

The voice grade circuit can, of course, carry voice (and dual-tone multifrequency [DTMF]), making it suitable for DTMF/voice control. It can also carry analog remote control telemetry that is based on audible tones. Finally, with the addition of modems (whether inside the remote control or outside), it can carry digital data, allowing its use on microprocessor and nonprocessor based digital control systems.

Note that the telephone company central office normally digitizes the incoming voice grade circuit at 64,000 bits per second. Based on this, we *should* be able to send 64 Kbps through our voice grade circuit. However, due to noise, distortion and delay distortion on the local loop between the customer and the central office, current modems are limited to about 9,600 bits per second. Eventually, it is expected that the 64 Kbps digital encoding utilized at the central office will be extended the "last mile" to the customer premises. This entirely digital system is ISDN (integrated services digital network).

Telco Dial-up Voice Grade Circuit

In December 1984, the FCC dropped the "failsafe" requirement from the Rules, making the use of temporary circuits legal for the control of broadcast transmitters. These "plain old telephone" circuits are inexpensive, reliable and very widespread. They are voice grade, allowing various data encoding techniques. The most common data encoding techniques for dial-up lines are standard dial-up modems (for digital data) and DTMF/voice. The wide availability of personal computers has reduced the cost on dial-up modems, making them suitable for driving control point terminals and computers. DTMF/voice systems utilize *dual-tone multifrequency* (Touch Tone) signals for control and synthesized voice for telemetry and status reporting. Use of the dial-up telephone network allows large transmitter control networks to be built, since the telephone company equipment handles all the data routing between the transmitter sites and the control points. The FCC does, however, require the broadcast transmitter operator to have the capability to turn off the transmitter at any time. They have stated that a single dial-up line is not sufficient. The alternatives are outlined in their clarification of the remote control rules (see 1988 Federal Register, page 37762).

Telco Digital Circuits

The digital circuits available from telephone companies are based on the 64 kilobit per second rate used to encode voice. Allocating a full "voice grade" digital circuit to broadcast transmitter telemetry severely underutilizes the circuit. However, some stations are now utilizing T1 circuits (1.544 Mbps, designed to carry 24 voice circuits) as a digital STL. With appropriate encoding, it is possible to run composite or discrete program quality stereo through a single T1 circuit. A small portion of the T1 circuit is typically allocated to be a low quality voice circuit or low speed digital circuit for transmitter control and telemetry. In each case, the 1.544 Mbps circuit is "broken" into several lower speed data channels suitable for carrying the digitized audio and/or other data. Optical fiber is being introduced throughout the telephone industry. It is typically used between central offices to handle very high speed digital data (generally representing many voice circuits). At this time, most telephone subscribers do not require sufficient data bandwidth to extend the fiber the "last mile" to customer premises. The one possible exception is where television stations utilize digital fiber for the transmission of digitized video. As with T1, some data capacity can be "stolen" from the digitized video channel to provide a digital data circuit between the studio and the transmitter, suitable for transmitter control and telemetry.

STL Subcarrier

Many stations use an aural or television STL to get programming from the studio to the transmitter. There is usually additional bandwidth available for the addition of a voice grade or digital data channel. In addition, television stations may insert control data in the vertical blanking interval of the video going to the transmitter site. This data may be replaced with telemetry data prior to being fed to the broadcast transmitter.

The most common technique of sending control data over an aural STL is to use an FM subcarrier to establish a voice grade circuit. Standard modems (inside or outside the remote control) then drive this new "voice grade circuit". Other modulation techniques are available, including frequency shift keying of the subcarrier, and quadrature amplitude modulating the subcarrier with digital data. The following frequencies are typically used:

- 26 kHz: Monaural or discrete stereo aural STLs
- 110 kHz: Composite stereo aural STLs
- 135 kHz: Composite stereo aural STLs
- 152 kHz: Composite stereo aural STLs
- 185 kHz: Composite stereo aural STLs
- 7.5 MHz: Video STLs
- 8.3 MHz: Video STLs

FM Subcarrier

Transmitter site telemetry may be returned to the control point (and anywhere else there is a subcarrier receiver) over a subcarrier on the FM broadcast channel. This is generally a voice grade FM subcarrier driven by a modem in the remote control. Other modulation techniques are available, including *frequency shift keying* and *quadrature amplitude modulation*. In addition, the subcarrier is sometimes further multiplexed by using the 20 to 30 Hz range to return the metering (either analog or digital) and the range above 50 Hz for other programming. Typical FM broadcast subcarrier frequencies are (assuming stereo operation) 67 kHz and 92 kHz.

TV Aural Subcarrier

In 1971, the FCC authorized TV stations to use an aural subcarrier with an instantaneous frequency between 20 and 50 kHz. Much of this spectrum is now taken by BTSC stereo. BTSC does, however, set aside a subcarrier frequency for station use (the PRO channel). This is at 6.5 times the horizontal scan rate. It is frequency modulated with a voice grade signal, or may be frequency shift keyed with a digital signal.

Vertical Blanking Interval

FCC Rule Section 73.682(a)(21) allows lines 17 through 20 (except line 19, which is reserved for VIR) of both fields for test, cue, control and identification signals. On an STL, these may be used to send transmitter control information. The data could be replaced with telemetry data at the transmitter site.

AM Subaudible

In 1969, the FCC authorized AM stations to use frequencies below 30 Hz on the AM carrier for telemetry. The modulation level of the telemetry signal was limited to 6%. Analog metering systems typically use 20 Hz to 30 Hz (20 Hz representing zero, 30 Hz representing full scale). Subaudible signalling may also be used to transmit digital data (although at a low rate, in the tens of bits per second). The subaudible digital signalling may use any of various modulation techniques. Frequency-shift keying and phase-shift keying have been used. The low speed capability of the data channel may be a limiting factor, but over a full day, a lot of data can be sent. If the telemetry system sends only the changes in parameters while the control point terminal stores the most recently received data (perhaps using a standard CRT terminal), a reasonable telemetry system is possible. The current FCC Rule Section 73.127 governs this multiplex use of AM carriers. It allows the use of either amplitude or phase modulation (the original rule required the use of AM). The rules no longer set a maximum injection level or frequency, though the multiplex signal must not be audible or cause radiation outside the channel.

AM multiplex signals are also finding other applications. One such application is electric utility load management. The subaudible "band" is also used by AM stereo system for a pilot signal to trigger the operation of stereo demodulators. These other applications may limit the use of this range of frequencies for transmitter telemetry.

Broadcast Subcarrier Limitations

Broadcast subcarrier (whether AM subaudible, FM or TV aural subcarrier, or TV VBI) telemetry transmission is low cost, but has one serious limitation. When there are problems with the broadcast transmitter, no telemetry is available, making remote diagnostics and troubleshooting difficult. For this reason, many prefer to use a telemetry link that will continue to operate on failure of the broadcast transmitter.

Dedicated Radio Link (P Channel)

The FCC has set aside eight frequencies in the 450 to 455 MHz area for operational communications, including telemetry and control. See FCC Rules Sections 74.402(c)(9), 74.462(b), and 74.482(d). It is most common to treat these radio links as a voice grade circuit, then use standard voice grade modem techniques to send data. It is also possible to directly digitally modulate the UHF carrier using any of the standard techniques (FSK, etc.). Many stations leave the transmitter up continuously, making the frequency unavailable to other stations in the area. Simple circuitry can be added between a standard remote control and the telemetry return link (TRL) transmitter to only key the transmitter when readings are being taken. Such systems could also use the dial-up technique of holding all the data changes in a buffer until the buffer fills, an alarm condition occurs, or a scheduled report time. Another technique available for frequency sharing is "packet radio." Here, a terminal node controller (TNC), packet modem, or PAD (packet assembler/disassembler) coordinates the use of the RF frequency, adds addressing to the packets (setting up "virtual circuits" between sites in the system) and adds error correction. The RF packet communications system provides a transparent serial digital circuit between two points over a shared RF channel. Some remote control equipment includes packet hardware and software in its design. Other systems may be adapted to packet radio by adding an external packet modem.

Finally, note that Section 74.482(d) outlines the specifications for automatically identifying the P channel radio link. Some remote control systems include the ID system. Some P channel transmitters also include an ID system.

FM Squared Satellite

Some satellite networks providing transmitter operating services rely on dial-up telephone lines for routine control and telemetry of transmitters. To provide an independent method of shutting down transmitters on the failure of the dial-up system, the networks may use subaudible tones on the program channel or interrupt the program audio. This is typically used on "FM squared" satellite channels where the program audio is carried as an FM subcarrier above video (which is also frequency modulated on the satellite, although some systems delete the video, just running an FM carrier with several FM subcarriers). Since subaudible data channels can only handle low data rates, some FM squared systems run a separate data channel (on another video subcarrier). This data channel normally handles addressed packet data. It includes program cues, individual station transmitter commands and text messages to subscriber stations (electronic mail). Loss of the data channel may be used as a "failsafe" way of shutting down the transmitters.

SCPC Satellite

Single Channel Per Carrier (SCPC) satellite systems run several FM audio carriers in a single video transponder, as opposed to carrying the audio as subcarriers on a single video carrier. The bandwidth of the SCPC carriers normally allows the addition of a subcarrier immediately above the program audio for each channel. Through the use of data packets and addressing, this data channel may be used for program cues, individual station transmitter commands, and text messages to subscribers. Again, routine transmitter control and telemetry is accomplished over dial-up telephone lines, since each station has a receive-only earth station.

VSAT Satellite

Very Small Aperture Terminal (VSAT) satellite systems are widely used for point-to-point and point-to-multipoint data distribution. They are now being used for broadcast transmitter control. These differ from the previously discussed satellite systems in that a transmit/receive terminal is put at each station. Through data packet switching, the VSAT system sets up a large number of "virtual circuits" through a single satellite transponder, on a single frequency. The techniques are similar to those used in packet radio. The single RF channel is divided among several users through time division multiplexing. However, instead of allocating a fixed time slot to each user, a time slot large enough to handle the data a site needs to transmit at that particular time is allocated (as in statistical multiplexing). The VSAT system appears to the user as a "very long RS-232 cable." Data put in one end comes out the other. The complexity in the middle is handled by someone else.

This "very long RS-232 cable" allows the control terminal of the transmitter control system to be located almost anywhere. To decrease the amount of time utilized on the satellite circuit, the remote control (or an interface computer) uses exception reporting, reporting only changes in system parameters (and often updating the data every half hour). Sending the same data over and over again (as would result from a stable transmitter) wastes time on the satellite that would otherwise be available to other users.

The two-way satellite link allows the complete control and telemetry of the broadcast transmitter to be handled by the satellite circuit. The previously described receive-only satellite systems relied on dial-up telephone circuits to get telemetry back to the control point. The VSAT system can also handle digitized audio to allow the control point to hear the EBS receiver and send EBS audio back to the station (although dial-up telephone may also be used).

FCC REGULATION

The FCC first authorized the remote control of non-commercial educational FM (NCE-FM) transmitters 10 watts and under in 1950. Since that time, the rules have gone through many changes. This section reviews current regulatory issues in remote control.

As with all regulations, these are subject to change. Please refer to the most recent version of the Rules, Part 73, as referenced at the beginning of this chapter or contact the FCC, a consulting engineer, or communications lawyer to determine the status of specific regulations.

Since the FCC simplified the remote control regulations in 1984, there has been considerable confusion as to what is required in the remote control of a broadcast transmitter. By reviewing various rules, FCC correspondence and FCC enforcement records, we can get a fair idea as to what is required. Generally, a station that complied with the remote control requirements prior to the 1984 reregulation probably still complies.

Remote Control Authorization

Nondirectional AM, FM, and TV stations are authorized to commence remote control operation without prior FCC approval. If a remote control point is established at other than the transmitter or main studio location, the FCC must be notified within three days [§73.1400]. New directional AM stations or those making modifications (and thereby required to install approved sampling systems as specified in §73.68) may request remote control authorization when applying for a construction permit (FCC form 301 or 340).

Existing directional AM stations with approved sampling systems may apply for remote control authorization on FCC form 301-A. Existing directional stations with nonapproved sampling systems may apply for remote control authorization on FCC form 301-A and including a showing that the array is in proper adjustment along with a one year stability analysis.

Required Metering

Reviewing the history of FCC regulation of remote control, we find that the remote control rules have never specified what parameters must be remotely metered. Instead, the rules have generally required sufficient metering to allow the operator at the remote control point to perform the functions required by the Commission's Rules. Since the rules at that time required specified entries in the transmitter operating log (at specified intervals), it was clear that these particular parameters must be telemetered. The logging rules no longer specify particular parameters that must be logged. The Commission has left to the station licensee the decision as to what parameters are to be logged and at what intervals. Correspondence with the Commission, reviews of their enforcement actions and reviewing other sections of the FCC Rules reveals the minimum parameter telemetering capability the FCC is expecting.

All stations with required tower lighting are required to make a daily inspection of the tower lighting or a remote indicating device (or alarm) that properly indicates the failure of any lamp. This is often accomplished by allocating an analog metering channel to monitor the tower light current.

From FCC correspondence and review of enforcement actions, it appears that AM stations are expected to have the capacity to determine the transmitter final amplifier voltage and current, and the antenna or common point current at the remote control point. These parameters make it possible for the operator to determine power by the direct and indirect methods. Differences in power determined between these methods indicate a change in transmitter efficiency or antenna (or common point) resistance.

FCC Rule Section 73.69(a) implies that the operator must be able to observe the antenna monitor indications. It is generally accepted that these indications will be telemetered.

FM stations determining power by the indirect method generally telemeter the final amplifier voltage and current. Those determining power by the direct method also telemeter the transmission line power meter.

Television stations normally telemeter the final amplifier voltage and current indications for both visual and aural transmitters (although the FCC deleted this requirement on visual transmitters in 1974, since logging of these parameters was not required). In addition, stations normally telemeter the visual and aural transmitter output power meters. Stations are required to have such additional monitors as required to insure the transmitted signal meets FCC requirements. These normally are a waveform monitor, vectorscope, and aural modulation monitor at the remote control point. These instruments monitor the broadcast signal received at the remote control point. There are automatic video analyzers available that can be located at the transmitter site. The data out of these analyzers (generally serial digital) can then be sent to the control point over a dedicated circuit or multiplexed into the remote control data.

Required Metering Accuracy

In establishing the rules permitting remote control operation, the FCC continually looked for a remote metering accuracy of 2% (remote meter to agree with local meter within 2% of reading). In 1975, rules were established specifying this accuracy. In 1984, the specified accuracy for remote metering was deleted. Currently, Section 73.1410(c) requires remote meters to be calibrated as often as necessary to ensure proper operation.

Section 73.57(d)(2) requires remote reading RF ammeters utilized by AM stations to agree with the local meter within 2%. Section 73.53(b)(9) sets requirements on the accuracy of external meters used with antenna monitors. No other specifications exist for required accuracy of remote control metering (although Section 73.1550(c)(2) requires 2% accuracy on extension meters).

FCC enforcement policy appears to require remote meters agree with local meters within 2% of the reading (as was previously required). This may not be sufficient to guarantee operation within licensed parameters, however. For example, if a directional AM station has a licensed phase of 150° (± 3°), and the remote meter disagrees by 2% (3°), the station is operating out of tolerance if the remote metering indicates anything other than 0° of error or 150°. Correspondence with the FCC has suggested a "tightened window" concept for remote metering calibration. Local meters are assumed to be correct. If (as in our previous example) the remote phase indication agrees with the antenna monitor plus or minus 1°, the station should keep the remote indication within two degrees to insure the indication at the transmitter site is within 3°. Reviewing FCC enforcement actions, it appears a station will be cited if the remote meters disagree with the local meters by more than 2% *or* if either the remote or local meters indicate an out of tolerance condition.

Required Control Capability

Section 73.1410(a) requires stations operated by remote control to provide sufficient control to assure compliance with FCC regulations. In the original discussion leading to the authorization of remote control of VHF television transmitters, it was suggested that stations extend control circuits that required daily adjustment. Traditionally, stations have extended the carrier on/off control, power trim (minor adjustment in output power to adjust for line voltage fluctuations) and any mode switch controls (day/night for AM stations). At this point, it appears that a station could "get by" with just on/off control available at the remote control point. However, should the operator find the station operating in an interference causing condition (over power, over modulation, out of tolerance DA, etc.), the operator's only way of correcting the problem is shutting down the transmitter. Extending other controls to the remote control point allows the operator to adjust the transmitter to correct some interference causing conditions, allowing the station to continue operation.

Fail Safe

Prior to 1984, the FCC required interruptions to the control circuit to shut down the transmitter. In addition, television stations were required to shut down within one hour of the loss of the telemetry circuit or the loss of any required telemetry sample. In 1984, the rules were revised to only require that control circuit interruptions (or other faults) must not activate the transmitter or change modes. The allowed interruption in the control circuit allowed the use of dial-up circuits for transmitter control (where the control and telemetry circuit is purposely interrupted). There was considerable confusion as to what additional control was required to insure the proper operation of a dial-up system. In 1988, the FCC issued a public notice explaining that it is permissible to use dial-up lines as the primary means of broadcast transmitter control and telemetry, subject to certain restrictions. The major concern is that the dial-up line is always available, or some other method of shutting down the transmitter is always available. Suggested methods of insuring these requirements are met include dedicating the line for the exclusive use of the duty operator, addition of a second dial-up line allowing the duty operator to override other system access, and the use of other continuous circuits between the control point and the transmitter to terminate transmitter operation. Suggested continuous links include STL carrier, program audio, and "P channel" radio links.

Alternate Control Points

Section 73.1860 requires transmitter operators to be licensed. Section 73.1800 requires the station log to be maintained by "station employees competent to do so, having actual knowledge of the facts required." In the early 1970s, the FCC considered the "employee" requirement to prohibit the use of contract duty operators or operating services. Since that time, however, the FCC has permitted a variety of transmitter operating services. These range from operation of the transmitter from other transmitter control points, the manager's home, an alarm service company, program networks, and dedicated transmitter operating services.

When some co-owned AM/FM/TV stations were split (again, in the early 1970s), the stations had been operating with the TV operator responsible for all the transmitters. When the ownership of the stations split, some stations continued to have the TV operator be responsible for the operation of the AM or FM transmitter. They may have paid the operator some token amount to make him/her legally an employee.

More recently, some FM stations sharing a transmitter site extended the remote metering and control lines from one station's remote control to the other. In addition, audio switching (actually, composite stereo switching) allowed a remote controlled simulcasting, allowing the EBS emergency transmission requirements to be quickly met. These additions allowed the operator of one station to operate the other, removing the requirement for an employee to "babysit" a transmitter and program automation system overnight.

With the introduction of dial-up transmitter control, it is easy to extend most of the transmitter control, and metering functions are available anywhere there is a telephone. Typical "off premises" control points include the manager's home, an alarm service company, an answering service, a program network and transmitter operating services. These are permissible to the FCC as long as there is always a qualified operator at a fixed control point. These operators must be licensed, properly trained in the operation of the broadcast transmitter and EBS equipment, and not be assigned to other duties that will result in inattention to the operation of the broadcast station. All ramifications involved in off-premises control should be carefully considered before starting such operation. The FCC must be notified of this control point within three days of its first use (see §73.1400(c)). Some alternate provision must be made for the operator to shut down the transmitter (or one of the other options outlined in the FCC clarification of the remote control rules, Federal Register, 28 September 1978, page 37762). Many stations have arranged transmitter site equipment so that dropping an STL carrier or interrupting program audio will drop the broadcast carrier. This appears to meet the FCC requirements as long as the operator has control over the STL carrier or program audio. If the operator is not at the studio, the operator may not have such control. In addition, the operator needs to comply with the EBS requirements. These include monitoring the EBS receiver and being able to properly respond to an emergency action notification (running announcements, running EBS tones, and running emergency programming). One approach to these requirements has been to install a second "remote control unit" at the studio. This unit controls the STL transmitter, the EBS encoder, program audio switching, and monitors the EBS receiver. It also satisfies the FCC's suggestion that a second dial-up line is suitable as an alternate means of shutting down the transmitter.

Some program networks offer transmitter operating services to the stations they own and to program subscribers. These generally operate the same as other dial-up systems. The satellite delivered program channel can be interrupted by the network to provide the FCC required alternate transmitter shutdown method. Network data channels may be used to provide a backup means of shutting down individual transmitters. Interruption of the data channel may also shut down all transmitters, providing a system "failsafe". In a national emergency, the program network may originate all EBS programming, including required announcements and EBS tones (although special precautions may be necessary to insure the EBS tones modulate the transmitter 40% each, as specified by Section 73.906(c)). These networks typically get a waiver of FCC Rules to allow the network to notify a local "on call person" of a local or state emergency (since EBS participation in these emergencies at management discretion, as specified in Sections 73.936(d) and 73.937(d)). This "on call person" decides whether to travel to the studio to participate in the state or local EBS. Other approaches available to these networks include listening to the EBS receiver over the dial-up line and remotely enabling the EBS tones, broadcasting the required announcements and the emergency programming (typically the CPCS station).

Finally, one transmitter operating service is using VSAT (very small aperture terminal) satellite transceivers to provide a dedicated two-way "virtual circuit" between many transmitter sites and the control point (see Fig. 21). This allows continuous supervision of multiple transmitters anywhere in the satellite coverage area. This differs from the program network system described above in that the data communications path is two way and (appears) dedicated. Digital control techniques are used to gather transmitter readings, receive alarm messages and generate transmitter control signals. The digital link can also be used to send digitized EBS audio in each direction (or dial-up telephone lines can be used, as is the case for the program network system). The digital circuits through the VSAT network also are used to distribute text messages (electronic mail) and digitized program audio between subscribers, or between the subscribers and the network control point.

MANY OPTIONS

Section 318 of the Communications Act of 1934 requires licensed operators to be in charge of all licensed transmitters (those that require a station license). The degree of supervision is not specified in the Act. The FCC Rules provide a bit more detail as to the degree of supervision, but leave the majority of the decisions to the station licensee. Advances in communications and computer technology have vastly expanded the available options to help the operator evaluate transmitter operation and make required adjustments. These advances are expected to continue. This chapter should be used as the starting point in the design of a control system for your station(s). Further research is suggested.

ENDNOTES

1. The requirement that transmitter monitors and meters be "within the operator's view" is based on Sections 73.69(a)(1), 73.1860(b) and 73.1550(a)(3). The FCC is considering authorizing stations to place required meters, monitors and controls within 30.5 meters and one floor above or below the normal operating position (see 55 FR 17093, item 4013). Such a rule change would allow the station licensee considerable flexibility in the placement of meters, monitors, controls and remote control equipment (co- owned stations could place monitors and remote controls in a common equipment room that

Figure 21. National Supervisory Network Control Point.

the operator visits as often as necessary to insure proper operation of the transmitters). Such a rule change would also eliminate the requirement (by the FCC) for extension meters. Stations could use extension meters provided observations of these meters indeed insured that the station operated properly (the extension meters were accurate).

2. Note that the process control industry normally transmits a current representing the measured sample, with 4 mA representing a value of zero, and 20 mA representing full scale. The use of current transmission eliminates the effects of variations in transmission loop resistance. The use of 4 mA to represent a parameter of zero causes different indications for a true parameter value of zero and a loop failure. Due to the short loop length and the complexity of developing 4 to 20 mA samples, the broadcast industry normally uses voltage samples.

3.8 Frequency and Aural Modulation Monitoring

Joe Wu
TFT, Inc., Santa Clara, California

GENERAL

Various types of modulation monitors are discussed in this chapter, together with the FCC Rules with which modulation monitors and transmitter emissions must comply.

Modulation monitors are specialized, highly accurate, and very low distortion instruments. They replace several more complicated, general purpose test instruments of similar accuracy as well as provide a convenient means for complying with the FCC Rules.

Broadcast station monitors fall into two categories: *frequency monitors* and *modulation monitors*. These two functions are sometimes combined into one unit and are sometimes packaged as separate units.

Broadcast aural modulation monitors are also classified according to the type of modulation to be monitored:

- AM monaural
- AM stereo
- FM monaural (and SCA as needed)
- FM stereo (and SCA as needed)
- TV monaural
- TV stereo (and multichannel sound)

Frequently, monaural and stereophonic functions are combined in one instrument for the convenience of the user.

WHY MONITOR THE TRANSMITTER

The three major reasons for using frequency and aural modulation monitors at a broadcast station are described in the following sections.

Coverage

Because coverage of the broadcast service area is enhanced by a high modulation level, it is desirable to maintain the modulation level at the maximum legal limit in order to maintain good coverage and improve signal-to-noise ratio for the audience.

Proof-of-Performance

By using the monitor as a precision test instrument, the station engineer can perform proof-of-performance measurements to insure that the transmitter is within the modulation and frequency technical specifications of the manufacturer and the operating requirements of the FCC. The FCC no longer requires a regular "proof-of-performance." However, stations are still required to meet certain technical requirements.

Most aural modulation monitors have built-in facilities to measure baseband audio frequency response and signal-to-noise ratio. Outputs are provided for total harmonic distortion measurements by an external distortion analyzer. Modulation monitors with stereo and SCA functions can measure stereo separation, subcarrier injection levels, crosstalk between service channels, and provide demodulated outputs for stereo and SCA channels. Aural broadcast modulation monitors are often comprised of several integrated pieces of test equipment to test many functions of aural transmitters.

Compliance with FCC Requirements

The monitor enables the station engineers to operate the transmitter in accordance with FCC Rules regarding aural modulation levels and carrier frequency tolerances. In addition, Section 73.1590(a) requires proof-of-performance measurements of all main transmitters, except class D noncommercial educational FM stations operating under 10 watts, to be taken as follows:

- Upon initial installations
- Upon modifications of transmission facilities
- Installation of AM stereo

- Installation of FM SCA or stereo
- Installation of TV stereo or subcarrier
- Annually on AM stations
- When required by other, special provisions of the station license

FCC RULES AFFECTING MONITORING

The FCC no longer specifies the type of aural modulation monitor or measuring equipment a broadcast station must use. Therefore, it is the responsibility of the station licensee to decide what monitoring equipment is needed to ensure that the station transmitter emissions comply with FCC frequency and modulation requirements.

Frequency Monitoring

Section 73.1540(a) of the FCC Rules requires that center frequencies of AM, FM, and TV stations must be measured or determined as often as necessary to ensure that they are maintained within the tolerances stated in FCC Rule Section 73.1545 as follows:

AM stations: The carrier frequency for monophonic transmissions or the center frequency for stereophonic transmissions must be within ±20 Hz of the assigned frequency.

FM stations: The center frequency must be within ±2000 Hz of the assigned frequency (±3000 Hz for transmitters having a power output of 10 watts or less).

TV stations: The visual carrier frequency must be within ±1000 Hz of the assigned frequency. The aural carrier frequency must be 4.5 MHz ±1000 Hz above the actual visual carrier frequency.

Aural Modulation Limits

Section 73.1570 of the FCC Rules states that modulation percentage is to be maintained at the highest level consistent with good transmission quality and broadcast service, not to exceed the following limits:

AM stations: Modulation of the carrier must not exceed 100% on negative peaks of frequent recurrence, or 125% on positive peaks at any time. There are additional regulations for AM stereo and telemetry transmissions.

FM stations: Total modulation must not exceed 100% on peaks of frequent recurrence referenced to 75 kHz deviation. However, stereo stations simultaneously providing subsidiary communication services (SCA) on subcarriers may increase the total peak modulation 0.5% for each 1.0% of subcarrier injection modulation; but the total carrier modulation must not exceed 110%. If two or more SCA subcarriers are used in conjunction with the stereo channel, the maximum allowable peak deviation is ±82.5 kHz or 110%.

TV stations: In general, the total modulation of the aural carrier must not exceed ±25 kHz deviation (monaural) on peaks of frequent recurrence. Stations transmitting multiplexed subcarrier signals on the aural carrier must limit the modulation of the aural carrier by the arithmetic sum of the subcarrier(s) allowable deviation and the total modulation must not exceed ±75 kHz deviation.

Modulation requirements for stations transmitting aural subcarriers as part of encoded subscription programs are stated in the application for FCC approval and contained in FCC Rule Section 73.682(b).

Modulation requirements for BTSC stereo sound are subject to the criteria set forth in FCC Rule Section 73.682(c) and FCC OET Bulletin 60A. Also see Chapter 3.6 on "Multichannel Television Sound."

ESSENTIAL FEATURES OF FREQUENCY MONITORS OR METERS

The primary standard of frequency measurements is the standard frequency maintained by the National Institute of Standards and Technology (NIST), which was formerly the National Bureau of Standards (NBS), or their standard broadcast signals of stations WWV, WWVB. Frequency monitors or meters (such as shown in Fig. 1) must be capable of accurately measuring and displaying the carrier frequency requirements. Recommended resolution for AM monitoring is 1 Hz, while FM and TV monitors should provide a resolution of 10 Hz or better. These tolerances call for an extremely accurate and stable internal frequency standard with an aging rate of 1 ppm per year or better which is traceable to NIST.

If the transmitter is to be monitored at some distance from the transmitter site, a built-in or external preselector is generally required to raise the input level for receiving the signal off-the-air.

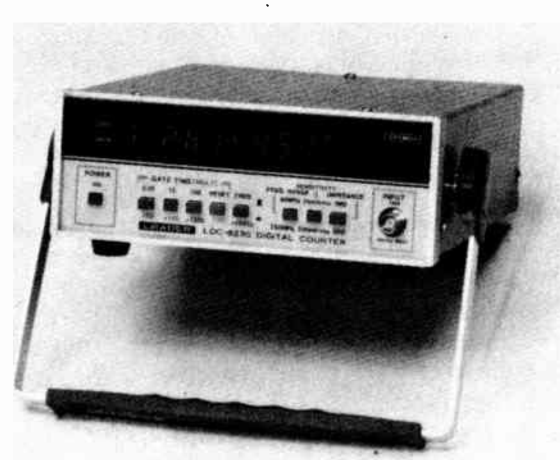


Figure 1. A general purpose frequency counter may be used to measure carrier frequency as well as other frequencies within its range. Specified accuracy and resolution for the model shown here make it suitable for measuring carrier frequencies of AM, FM and VHF-TV. (Courtesy Leader Instruments Corp.)

The following features are highly desirable in a frequency monitor:

- Digitally tuned RF preselector for multiple off-air applications
- Digitally tuned RF preselector for multiple off-air applications combined with the aural modulation monitor
- An output to operate an alarm when preset frequency limits are exceeded
- Provision for calibrating the internal frequency standard against an NIST station or other highly accurate standard
- An output for automatic logging

Peak Modulation Duration

Although there is no precise definition given by the FCC regarding the maximum allowable peak modulation duration in the rules today, some manufacturers of modulation monitors are using the pre-1983 de facto rules as their design guidelines. The de facto rules allowed a 1 millisecond response time for the peak modulation indicators and the permissible over-modulation limit was ten counts per minute.

ESSENTIAL FEATURES OF AN AURAL MODULATION MONITOR

General

As a minimum, all aural modulation monitors should have a quasi-peak reading modulation meter to give a direct indication of modulation percentage. Meter accuracy should be ±4% or better. Other desirable features common to all modulation monitor are:

- Park indicators, accurate to ±2% or better, to indicate when maximum positive and negative modulation peaks are occurring
- Adjustable peak indicator trigger points to indicate when modulation peaks exceed preset levels
- Adjustable peak modulation duration detecting circuit in the event FCC and the broadcast community agree upon the maximum allowable peak duration
- An internal modulation level calibrator to check the accuracy of the modulation meter and peak flashers and a means for recalibrating the meter and peak flasher circuits
- An output to operate an over-modulation alarm or a built-in alarm
- An output to operate an external alarm or a built-in alarm for when the modulation drops below a certain level (i.e., 10%) for a specified period of time
- Outputs for a remote meter and peak flasher

AM Monitors

AM monaural modulation monitors should possess the features described in the preceding paragraph. If an RF preselector is used for off-air monitoring with the monitor, its sensitivity should be approximately 100 μV for a 35 dB SNR and 1 mV for a 50 dB SNR. Selectivity should be at least −40 dB at ±40 kHz, and image rejection should be at least 50 dB. The RF preselector must be very linear to avoid causing erroneous readings on the modulation monitor.

It is also desirable that the AM modulation monitor be equipped with new National Radio Systems Committee (NRSC) de-emphasis circuit so that the audio output of the monitor matches that of the AM transmitter modified with the NRSC audio response characteristics.

AM stereo monitors should have the following additional features:

- L+R and L−R channel decoding and outputs
- L and R channel separation measurement capability
- L and R channel signal-to-noise ratio measurement capability
- L and R channel frequency response measurement capability
- Pilot carrier injection level measurement capability
- Channel crosstalk measurement capability
- Signal output for distortion measurements

FM Monitors

FM monitors should have at lease a 70 dB signal-to-noise ratio. The discriminator must have a distortion figure of 0.1% or better and a baseband frequency response of at least 25 Hz to 100 kHz, so that it can pass and accurately measure an SCA channel up to 92 kHz. FM monitors often consist of both frequency and modulation monitors in one package (see Fig. 2).

If an RF preselector is used for off-air monitoring, its IF amplifier should have a linear phase response curve, yet be narrow enough to reject adjacent channels. A built-in multipath detector is highly desirable to help minimize multipath interference. The FM monitor should be equipped to measure synchronous and non-synchronous AM noise of the FM carrier.

FM stereo monitors should be able to measure:

- The L+R channel level (30 Hz to 15 kHz)
- The L−R channel level (23 to 53 kHz)
- The 19 kHz pilot injection level
- The 38 kHz subcarrier level
- Crosstalk between main channel and subcarriers
- Separation of left and right channels (up to 60 dB)

If one or more SCA subcarriers are transmitted, an SCA monitor should be used in addition to the FM stereo modulation monitor. The SCA monitor is usually an add-on to the main unit, which takes a composite feed from the demodulated output of the baseband monitor. Some manufacturers offer an optional RF and baseband demodulator, so that the SCA monitor can be used independently from the main monitor. The capability for user selection of SCA frequencies is important for future expansion of SCA service.

The SCA monitor should be able to measure:

- Modulation percentage
- SCA injection level on the composite signal
- Signal-to-noise ratio
- Crosstalk

Figure 2. A combined frequency and modulation monitor. (Courtesy TFT, Inc.)

The SCA modulation measurement should be selectable for ±4 kHz or ±6 kHz as the level for a meter indication of 100%. FM monitors often include both modulation and frequency monitoring in one package (see Fig. 2).

TV Monitors

The features for TV monaural and stereo monitors are similar to those for FM monitors. When transmitting Broadcast Television Systems Committee (BTSC) multichannel sound, the operator should be able to monitor the main channel and stereo channel as well as SAP and PRO channels (if they are utilized). Monitoring the modulation level of the BTSC signal is extremely important for achieving good stereo separation (see Fig. 3).

BTSC Stereo Separation and Modulation Accuracy

Stereo separation in the BTSC format is very sensitive to gain and phase errors in the transmission path. This is because the L + R and L − R signals are treated differently. In particular, L − R is companded while L + R is simply preemphasized and deemphasized. The L − R and L + R signals must arrive at the receiver's decoder matrix, which yields L and R, with very small errors in gain and phase across the aural baseband from 50 Hz to 50 kHz. Fig. 4 shows how stereophonic separation is affected by gain and phase errors in the L − R signal relative to the L + R signal at the input of the final matrix.

Subjective tests have shown that an average listener begins to "perceive" a loss in the spatial character of stereophonic music material when the separation drops below 18 dB. A separation of 15 dB ±3 dB is considered "adequate" by the average listener.

Although the subjective effects of separation depend on the spectral distribution and other aspects of the audio material, it appears that a good engineering objective for the entire system is for the separation to exceed 20 dB in the mid-range, decreasing somewhat at frequencies above 8 kHz. Fig. 4 shows that a separation of 20 dB requires a gain error smaller than 1 dB, and a phase error of less than 10 degrees. The BTSC standards require that the separation of the radiated signal exceed 30 dB in the mid-band from 100 Hz to 8 kHz, but that it may decrease at low frequencies to 26 dB at 50 Hz, and to 20 dB at 14 kHz. This

Figure 3. For remote monitoring of BTSC signal off-air or on-site, this monitor can also take an input from an IF modulated TV transmitter or a BTSC stereo generator. A plug-in distortion and baseband signal analyzer is optional. (Courtesy TFT, Inc.)

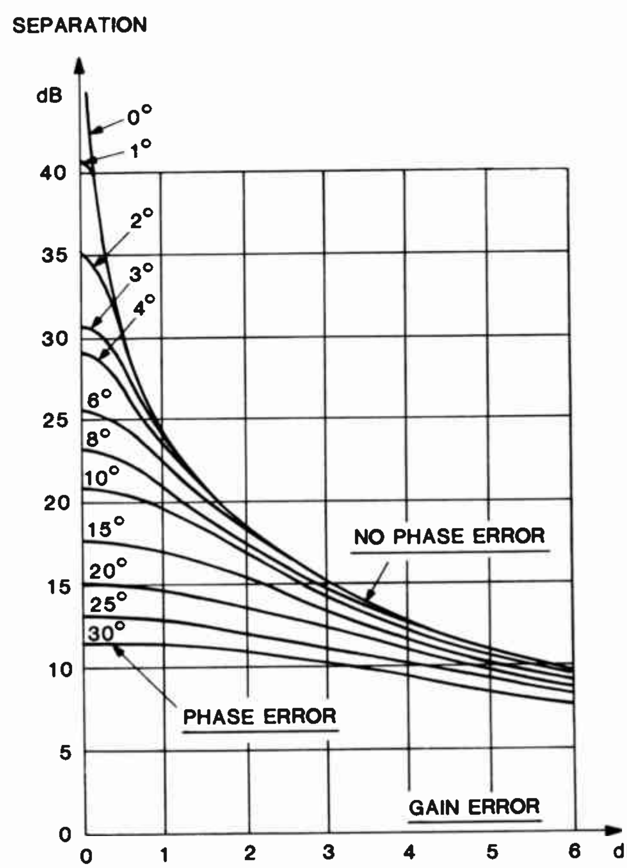


Figure 4. Stereo separation as a function of gain and phase errors in the L − R versus the L + R paths. (Reprinted with permission from TFT, *TV Aural Proof of Performance Guide*.)

requires that the gain and phase errors in the mid-band be less than 0.3 dB and 3.0 degrees, respectively.

The total modulation level accuracy in BTSC stereo is more critical than in FM stereo radio broadcasting in order to produce acceptable stereo separation. Because the L+R and L−R signal paths in BTSC are processed differently, a small change of modulation level in the BTSC system will affect the stereo separation. This is because the amplitude and phase relationship between the L+R and L−R channel is altered. If the total modulation level of the BTSC system is not maintained accurately, the dbx decoder in the receiver will see an incorrect RMS level and reproduce an L−R signal with altered amplitude and phase. That is, if an incorrect L−R signal is fed to the decoding matrix, the consequence will be poor stereo separation. Because the RMS level to the input of the decoder is directly proportional to the total modulation level, the total modulation level in the BTSC transmitter must, therefore, be accurately monitored in order to maintain good stereo separation and high quality audio performance.

ON-SITE MONITORING TECHNIQUES

In a studio transmitter collocated operation, the monitor is normally connected directly, or through an attenuator, to a RF sampling point of the transmission line feeding the antenna via a directional coupler. It is important to know the RF voltage level of this sampling point so that it meets the input requirement of the monitor.

AM Monitoring

Most AM frequency monitors display carrier frequency error rather than actual carrier frequency. Typically, a front-panel digital display indicates the carrier deviation in Hz from its assigned frequency, and a lighted "+" or "−" indicates whether the carrier is above or below the carrier frequency.

The modulation percentage of a monaural AM carrier is normally displayed directly on a front-panel meter.

Peak flashers, used on some monitors, are intended to catch fast transients and peaks that the meter cannot respond to. There may be one flasher to indicate maximum allowable negative peaks (100%) and another to indicate maximum allowable positive peaks (125%). The monitor may also have an adjustable peak flasher which can be preset by means of digital switches so that it flashes when the modulation percentage exceeds the preset switch value.

For monitoring modulation of an AM stereo transmitter, an AM stereo monitor or a stereo monitor plus a compatible AM modulation monitor is required. The equipment should permit the operator to simultaneously read the modulation percentage on both left and right channels and to measure separation between channels and crosstalk between the main channel and subchannels.

FM Monitoring

A typical FM frequency monitor digitally displays carrier frequency error in Hz (rather than display actual carrier frequency), and indicates whether the carrier is above or below its assigned frequency. The RF input level must be adjusted so that sufficient signal is available for measurement but not so high that overloading occurs. In addition to monitoring the FM stereo transmitter carrier frequency, it should be possible to monitor the stereo pilot carrier frequency as well.

On a typical monaural FM monitor, the modulation percentage is normally displayed on a front panel meter. Generally, either positive, negative, or combined modulation peaks can be selected for monitoring. Some monitors also provide peak flashers to indicate modulation peaks that exceed a preset percentage. The RF input must be adjusted to the correct level, as described in the instruction manual, before accurate readings can be taken.

For monitoring a stereo FM transmitter, a modulation monitor with a compatible stereo monitor must be used. Left channel, right channel, and total modulation are generally read on front panel meters. With the typical monitor, these meters can also be used to measure separation between the left and right channels as well as crosstalk between the main channel and subchannels.

SCA monitors are usually accessories to the baseband monitor. They are either fed from the FM composite signal or from the RF carrier through a separate FM demodulator.

TV Monitoring

The typical TV frequency monitor will provide separate displays of the visual carrier, aural carrier, and intercarrier frequency errors (rather than actual carrier frequency). The RF input to the monitor should be adjusted to the correct level as described in the instruction manual before measurements are made.

Aural modulation percentage can be read from a front panel meter. If the TV transmitter is also transmitting BTSC stereo sound, a monitor with a stereo decoder for left and right channels must be used. If a SAP channel is employed, a SAP monitor should also be used to monitor modulation and frequency.

Due to the critical relationship in BTSC stereo separation between the recovery of the companded L−R channel and the modulation level of the transmitter described previously, it is essential to maintain accurate modulation levels when broadcasting multichannel television sound. One convenient method of achieving this objective is to monitor the pilot carrier injection level, which has a constant level. After calibration of modulation levels for maximum stereo separation, the pilot carrier injection level is the best reference for maintaining maximum stereo performance.

OFF-SITE MONITORING TECHNIQUES

For remote, off-air monitoring, the monitor should incorporate a built-in preselector (RF amplifier) or an external RF preselector connected to an outdoor antenna. Some preselectors are capable of measuring and indicating carrier frequency error. Such preselectors, when combined with a modulation monitor, provides a compact complement of monitoring equipment.

Off-site monitoring techniques are generally the same as on-site techniques. RF input level to the preselector or monitor must be carefully adjusted, as described in the manufacturers' instruction manuals. It is even more important to know the RF level so that it does not overload the RF preselector and create intermodulation products. Some monitors are equipped with a multipath detector which enables the user to rotate the receiving antenna for minimum multipath interference.

CALIBRATION AND MAINTENANCE

Frequency Monitor Calibration

Frequency monitors should be calibrated periodically to ensure accurate measurement of transmitter carrier frequency error. The time between calibrations depends on the frequency of the monitor's internal standard, which in turn depends on the operating frequency of the transmitter. For a monitor having an internal crystal standard that has been correctly calibrated once and that has an aging rate of 1 part per million (ppm) per year, the following calibration schedule is recommended:

Transmitter Frequency	Calibration Interval
AM band	Every 12 months
FM band	Every 12 months
Low VHF TV band (Ch 2–6)	Every 12 months
High VHF TV band (Ch 7–13)	Every 6 months
UHF TV band	Every 3 months

After the first few calibrations, the interval may be lengthened if the drift observed in the first few checks warrants it.

There are a number of ways the internal frequency standard in a monitor can be calibrated:

1. Use a high quality frequency counter having a resolution of 1 Hz display and a time base that is calibrated against a secondary standard having an accuracy of at least 1×10^{-8}.
2. If the frequency display can be operated in a general-purpose counter mode, a frequency standard of higher accuracy than the monitor's internal standard can be used with the counter to calibrate the internal standard. The external standard is connected to the counter input, and the internal standard (which furnishes the time base for the counter) is adjusted so that the counter displays the exact frequency of the external standard.
3. If a receiver capable of receiving one of the NIST stations (WWV at exactly 5 MHz, 10 MHz, and 15 MHz) is available, and if the monitor has an output from its internal standard that is an exact subharmonic of the frequency of the NIST station being received, the monitor internal standard can be calibrated by adjusting its frequency to zero-beat with the NIST station frequency.

Aural Modulation Monitor Calibration

The aural modulation meter and peak flashers should be calibrated regularly. Most monitors have built-in calibrators, so that meter and flasher accuracy can be checked by simply pressing front-panel switches and observing the peak flashers and the meter reading. If the reading is in error, a simple adjustment usually corrects the error.

If the built-in modulation calibrator is of the frequency marker type for establishing the ±100% peak modulation level, and the frequency markers are generated from single crystal sources, the accuracy can usually be maintained within ±1%. If this type of monitor is being used, it is not necessary to calibrate the FM modulator using a Bessel null method.

If an FM modulation monitor has no internal calibrator, the monitor must be calibrated against a laboratory standard or by means of a Bessel null measurement using a spectrum analyzer and a precision audio frequency generator.

If an AM modulation monitor has no internal calibrator, the monitor can be calibrated using an RF generator. The generator must be capable of very low distortion amplitude modulation, and the level of this modulation should be accurately observed by using a high quality oscilloscope.

An AM monitor is usually calibrated by comparing the peak amplitude of the waveform against an amplitude reference established by an oscilloscope with high linearity. A digitally-generated reference is frequently used as a built-in calibration standard as in an FM modulation monitor.

Maintenance

A broadcast station monitor should be maintained in the same way as other precision laboratory instruments. If should be calibrated regularly as described in the preceding paragraphs. Manufacturers also offer this type of service to their customers.

APPENDIX*

How to Monitor Frequency and Modulation without a "Monitor"

Carrier frequencies can be measured and monitored with a frequency counter that has the required accuracy

*This material was provided through the courtesy of Arno Meyer, President, Belar Electronics Laboratory, Inc., Devon, Pennsylvania.

and frequency range. Characteristics to be aware of include:

For AM—The carrier frequency must be measured with no modulation or at low modulation, since frequency counters will usually miscount or lose cycles on negative modulation peaks.

For FM—A time base of longer than one second is necessary to average out the modulation swings when a frequency counter is used while modulation is present.

For TV—When a frequency counter is used on TV carriers, the visual carrier and the aural carrier or intercarrier frequencies must be measured separately. The visual carrier must be measured with no modulation in order not to lose count in the modulation troughs. The aural or intercarrier should be averaged the same way as for FM.

Aural Modulation

AM modulation can be measured without a demodulator but both FM and TV aural require a high quality, calibrated demodulator to measure modulation.

For AM—the common method is to use an oscilloscope to look at the AM envelope. 100% positive modulation is defined as twice the unmodulated carrier and 100% negative modulation is equal to zero carrier or carrier pinch-off.

Positive peaks of program material are difficult to see, but a digital storage scope can capture the peaks. 100% negative peaks can easily be observed as carrier shut-offs.

For FM—Monitoring FM modulation without a specialized monitor requires a high quality demodulator. The deflection on the y-axis of an oscilloscope is calibrated for 100% using the Bessel null technique. For increased accuracy, the y-axis is expanded and the zero is offset.

When the baseband signal consists of stereo and SCA subcarriers, the peaks are difficult to observe and a digital storage scope with a high sample rate readily captures the peaks when used in a long persistence mode. This is the method used by the FCC enforcement units when measuring total peak modulation for compliance with FCC Rule Section 73.1570.

For TV—Monaural TV and the total modulation of a TV stereo transmission may be monitored in a similar fashion to the FM method described above. To monitor the 25 kHz main channel requires a means of separating the main channel from the subcarriers.

Convenience

The above methods of monitoring work but may not be as convenient as using specialized modulation and frequency monitors. The cost of a specialized monitor may actually be less than the cost of general purpose test instruments needed for measuring frequency and modulation manually.

4.1 Television Signal Transmission Standards

Edited by Edmund A. Williams
Public Broadcasting Service, Alexandria, Virginia

INTRODUCTION

This chapter is intended to provide the television broadcast engineer with standardized measurement methods and performance objectives for various transmission systems used to distribute television signals. It consists of (1) an introduction and definitions section, (2) a section listing of video and audio performance objectives for terrestrial and satellite transmission systems from EIA/TIA-250-C,[1] (3) a section with the text of NTC Report Number 7, which provides the method for making most of the video performance measurements, and (4) a section with additional material on video and audio signal measurements from EIA/TIA-250-C.

Electronic Industries Association Standard EIA/TIA-250-C "Electrical Performance for Television Transmission Systems" (formerly RS-250-B), produced in 1990, reflects industry wide consensus on measures of performance which may be used to evaluate short-, medium-, and long-haul microwave circuits, satellite circuits, and various end-to-end combinations of microwave and satellite circuits. Many of the measurement techniques presented in EIA/TIA-250-C are based on the procedures that were established in NTC-7.

NTC-7 (Report Number 7, "Video Facility Testing") was prepared by the Network Transmission Committee (NTC) of the Video Transmission Engineering Advisory Committee, a joint committee of the television networks and AT&T. Because AT&T was the primary carrier of television network signals prior to the development of satellite systems, the objective of the document was to provide standard methods for making performance measurements and to establish guaranteed minimum performance criteria for coast-to-coast, long-haul microwave and coaxial cable transmission systems operated by AT&T and other common carriers and local telephone companies.

NTC-7, produced in 1976, defines transmission parameters, test signals, measurement methods, and performance objectives. The performance objectives are applicable to an overall (end-to-end) *terrestrial transmission system*, including both local and intercity facilities. NTC-7 does not include audio channel performance objectives. Therefore, excerpts from the audio section of EIA/TIA-250-C has been included in this chapter.

NTC Number 9 (NTC-9), produced in 1988, defines the measurement methods (most of which are based on NTC-7) and means for determining performance criteria for *satellite transmission circuits* that are provided by AT&T and other common carriers. Performance objectives are not established in NTC-9 but may be found in the SAT/NTC-9 column in the table.

Organizations utilizing terrestrial microwave and satellite circuits on a regular basis should acquire copies of the several standards described above. In addition to video and audio performance measurements methods, the documents provides a wide range of definitions and measurement techniques. NTC-9 especially describes the RF aspects of satellite earth terminals. Up-to-date versions of the television performance standards documents cited in this chapter are essential for specifying, operating and maintaining video transmission circuits.

Further information on obtaining complete and up-to-date versions of all three standards is contained in the reference section at the end of the chapter.

Definitions Applicable to This Chapter

Short-Haul Transmission System: A transmission system (generally one section) for the transmission of television signals between two points which can range from a few feet to about 20 route miles.

Medium-Haul Transmission System: A transmission system of generally more than one section between two points over a distance of about 20 to 150 route miles.

Long-Haul Transmission System: A transmission system of generally more than one section between two points over a distance of 150 to about 3,000 route miles.

Satellite Transmission System: A transmission system between a transmitting earth station and a receiving earth station via a satellite repeater.

End-To-End Network: The interconnection of more than one of the above transmission systems. An end-to-end configuration might, for example, consist of a short-haul microwave circuit, a satellite circuit, and a medium-haul microwave circuit.

IRE Units: An IRE unit is 1/100 of the luminance value range, of a television signal, that is between blanking (0 IRE) and reference white (100 IRE). 100 IRE represents the video portion (0.714 volts) of the composite signal (1.0 volt). The synchronization portion of the composite signal is 40 IRE units (0.286 volts). The term IRE units was developed by the Institute of Radio Engineers prior to that organization becoming part of the Institute of Electrical and Electronic Engineers (IEEE).

Average Picture Level (APL): The average level during active scanning time excluding blanking and synchronizing signals that is integrated over a time period of one frame (1/29.94 second or 33.4 ms).

VIDEO AND AUDIO PERFORMANCE OBJECTIVES

The chart on the next several pages (Fig. 1) presents a summary of most of the tests and performance objectives of EIA/TIA-250-C, NTC-7 and NTC-9. They are listed in the order consistent with EIA/TIA-250-C, but there are references to the NTC-7 section with measurement procedures which are provided on the following pages. By referring to the chart and the complete test procedure, the tests and performance objectives for a particular application may be selected.

The original order and numbering of NTC-7 tests have been preserved. To supplement NTC-7, excerpts from the audio section of EIA/TIA-250-C has been included. Audio test information appears at the end of Section 2. The whole of NTC-7 is reproduced in Section 3.

Additional excerpts from EIA/TIA-250-C are provided at the end of the Chapter.

Audio Performance Standards

In addition to the signal-to-noise ratio, signal availability, and time differential parameters shown in the chart, the following values, from EIA/TIA-250-C, are applicable to all television transmission circuits (terrestrial and satellite):

Test Measurement	Value	EIA/TIA-250-C Sec.
Amplitude/Frequency Response	See Fig. 3	7.1
Insertion Gain	0 dB	7.4
Total Harmonic Distortion	0.5%	7.2
Gain Difference Between Stereo Channels		7.5.1
50—100 Hz	1.0 dB	
100—7500 Hz	0.5 dB	
7500—15000 Hz	1.0 dB	
Phase Difference Between Stereo Channels		7.5.2
50—100 Hz	10°	
100—7500 Hz	3°	
7500—15000 Hz	10°	
Crosstalk Between Stereo Channels	56 dB	7.5.3

Amplitude vs. Frequency Response

Definition: The audio amplitude versus frequency characteristic of a television relay system is an expression of amplitude variation as a function of audio frequency when applied to the system audio input and measured at the system audio output. The amplitude variation is expressed in decibels (dB).

Standard: The audio amplitude versus frequency characteristic shall be within the applicable limits cited in Fig. 3. This standard applies equally to short-haul, medium-haul satellite, long-haul, and end-to-end performance on the assumption that audio will be demodulated from its subcarrier only once in any system of interconnection.

Method of Measurement: The measuring equipment shall terminate the output of the audio channel under test in a standard load impedance. The grounded or balanced-to-ground connection normally used shall be maintained. A standard sinusoidal 400 Hz test tone shall be applied to the input of the system and rated modulation established. The output level control shall be set to deliver standard output.

If pre-emphasis and de-emphasis are not used, the response shall be measured while maintaining constant input to the system and measuring output variations.

If pre-emphasis and de-emphasis are used, after adjusting the circuit as described above, the input to the system shall be decreased 20 dB and maintained at this level during the complete frequency response measurement. The output level control shall not be reset. This requirement is necessary to prevent overload of the audio channel at the higher audio frequencies due to the boost provided by the pre-emphasis network.

Harmonic Distortion

Definition: Audio frequency harmonic distortion is the production of harmonic frequencies at the output of the system caused by nonlinearities when a nondistorted sinusoidal signal is applied to the input of the system.

Figure 1 part 1.

TEST PARAMETER (Test signal)	NTC-7	EIA/TIA-250-C					SECTION (Add Rate)
		Short Haul	Medium Haul	Sat/NTC-9	Long Haul	End-to-End	
Gain/Frequency Response	+3–7 IRE See Sec. 3.8	±2.5 IRE Max See Fig. 2	±7 IRE Max See Fig. 2	±7 IRE Max See Fig. 2	±12 IRE Max See Fig. 2	±12 IRE Max See Fig. 2	6.1.1 (RSS)
(Multiburst, Multipulse or Sweep)	Note: The effect of luminance signal high-frequency response roll-off in the video passband is poor resolution or softness observed in the displayed picture. Non-linear phase characteristics can appear as a consequence of poor high frequency response of the video passband. If the video passband is peaked at the upper limit, the display will show an unusual crispness or enhancement in picture. The use of high frequency peaking circuits in video equipment and transmission channels will often result in increased appearance of noise and other high frequency artifacts.						
Chrominance to Luminance Gain Inequality (CLGI)	±3 IRE See Sec. 3.6	±2 IRE	±2 IRE	±4 IRE	±7 IRE	±7 IRE	6.1.2.1 (RSS)
(Modulated Stairstep)	Note: This problem will affect the saturation levels in the received picture which will appear as over or under saturated colors. This distortion is sometimes referred to as "saturation errors."						
Chrominance to Luminance Delay Inequality (CLDI)	±75 ns See Sec. 3.7	±20 ns	±33 ns	±26 ns	±54 ns	±60 ns	6.1.2.2 (RSS)
(Modulated Stairstep)	Note: This distortion, a difference in transmission time between the luminance and chrominance signals, is most noticible with red lettering smearing into a neutral background. In more severe examples of this distortion, "color ghosting," or misregistration of colors, may be displayed on the received picture.						
Field Time Waveform Distortion	< 4 IRE P.P See Sec. 3.3	±4 IRE P.P	±4 IRE P.P	±4 IRE P.P	±4 IRE P.P	±4 IRE P.P	6.1.4 (Linear)
(Field Bar)	Note: Obvious field time errors are displayed as objectional vertical shading from top to bottom in the received picture.						
Line Time Waveform Distortion	< 4 IRE P.P See Sec. 3.4	0.5 IRE P.P	1.0 IRE P.P	1.0 IRE P.P	1.5 IRE P.P	2.0 IRE P.P	6.1.5 (RSS)
(Line bar)	Note: This impairment may appear as smearing or streaking from left to right in received picture. An uneven shading effect from left to right or poor contrast may be observed as well.						
Short-Time Waveform Distortion	±6 IRE See Sec. 3.5	2% SD ±4 IRE	2% SD ±4 IRE	2% SD ±4 IRE	3% SD ±6 IRE	3% SD ±6 IRE	6.1.6 (3/2 Power)
(0.125μ sec Sin^2 step)	Note: This type of distortion affects horizontal resolution, sometimes referred to as definition. Also, "contouring" or "ringing" can be observed when excessive high frequency response of the video passband is present.						
Long Time Waveform Distortion (Bounce)	5 IRE 1 second See Sec. 3.12	8 IRE 3 seconds	8 IRE 3 seconds	8 IRE 3 seconds	8 IRE 3 seconds	8 IRE 3 seconds	6.1.7.1
(10% to 90% Bounce)	Note: Intermittent and slow variation in picture brightness may be seen over long periods of time. Flicker is also an example of a long-time distortion.						
Insertion Gain Variation	±3 IRE/5 sec ±1 IRE/hr See Sec. 3.2	±0.15 dB ±1.7 IRE	±0.3 dB ±3.5 IRE	±0.2 dB ±2.3 IRE	±0.45 dB ±5.3 IRE	±0.5 dB ±6 IRE	6.1.8 (RSS)
(18 μs bar-window)	Note: May be observed as over or under contrasted pictures in monochrome transmissions and over or under contrasted and saturated pictures in color transmissions.						
Luminance Nonlinearity	10% See Sec. 3.9	2%	4%	6%	8%	10%	6.2.1 (3/2 Power)
(Modulated Stairstep)	Note: Non-linear transfer characteristics in a video channel will result in contrast errors relative to the transmitted picture for the luminance. More serious problems occur to the chrominance signal, such as hue and saturation errors, that result from the non-uniformities in the luminance component relative to the chrominance information.						

Figure 1. Television signal performance standards.

Figure 1 part 2.

TEST PARAMETER (Test signal)	NTC-7	EIA/TIA-250-C					SECTION (Add Rate)
		Short Haul	Medium Haul	Sat/NTC-9	Long Haul	End-to-End	
Differential Gain	< 15 IRE < 15% (Sec. 3.13)	2 IRE or %	5 IRE or %	4 IRE or %	8 IRE or %	10 IRE or %	6.2.2.1 (3/2 Power)
(Modulated Stairstep)	Note: This impairment may be observed as saturation errors in the received picture as a function of amplitude variations. For example, low or high contrast colors may be over or under saturated.						
Differential Phase	< 5 deg (Sec. 3.14)	0.7 deg	1.3 deg	1.5 deg	2.5 deg	3 deg	6.2.2.2 (3/2 Power)
(Modulated Stairstep)	Note: The effect of this impairment when displayed on the received picture are changes or errors in color tint, or hue, as a function of luminance amplitude variations.						
Chrominance-to-Luminance Intermodulation	< 3 IRE (Sec. 3.15)	1 IRE	2 IRE	2 IRE	4 IRE	4 IRE	6.2.3 (3/2 Power)
(3-Level Chroma Signal)	Note: The result of this impairment is the incorrect reproduction of highly saturated color areas of the received picture.						
Chrominance Non-linear Gain	10 IRE or % (sec. 3.10)	1 IRE	2 IRE	2 IRE	4 IRE	5 IRE	6.2.4.1 (3/2 Power)
(3-Level Chroma Signal)	Note: The visual effect for this video impairment may also be the incorrect reproduction of highly saturated colored areas. However, the effect may not be apparent by observing the received picture without comparing it to the transmitted information.						
Chrominance Non-linear Phase	< 5 deg (Sec. 3.11)	1 deg	2 deg	2 deg	4 deg	5 deg	6.2.4.2 (3/2 Power)
(3-Level Chroma Signal)	Note: The effect of this impairment on the received picture would be a shift in color tint or hue in highly saturated areas.						
Dynamic Gain of Video Signal	±3 IRE (Sec. 3.12)	2 IRE	3 IRE	4 IRE	5 IRE	6 IRE	6.2.5.1 (3/2 Power)
(Stairstep with Variable APL)	Note: If the system gain is reduced as a function of APL, the dynamic range of the video signal will typically be compressed. This will contribute to all non-linear distortions and be displayed as reduced contrast ratios on the received picture.						
Dynamic Gain of the Sync Signal	±2 IRE (Sec. 3.12)	1.2 IRE	1.6 IRE	2 IRE	2.4 IRE	2.8 IRE	6.2.5.2 (3/2 Power)
(Stairstep with Variable APL)	Note: This effect may be due to a non-linear transfer characteristic of the video channel. This problem contributes to all non-linear distortions as a function of APL.						
Transient Sync Signal Non-linearity	—	1 IRE	2 IRE	3 IRE	4 IRE	5 IRE	6.2.6 (3/2 Power)
(Flat field with Variable APL)	Note: This effect is observed as changes in sync level as a result of a sudden and large change in APL.						
Signal-to-Noise Ratio	≥ 53 dB (Sec. 3.16)	67 dB	60 dB	56 dB	54 dB	54 dB	6.3.1 (RSS)
(10 kHz to 4.2 MHz Weighted)	Note: Excessive noise results in a "snowy" or "grainy" picture displayed on monochrome monitors. On color monitors, the effect is noted as colored snow, also referred to as confetti.						
Signal-to-Low Frequency Noise Ratio	—	53 dB	48 dB	50 dB	44 dB	43 dB	6.3.2 (RSS)
(0–10 kHz)	Note: Typical sources of low frequency noise are power line or power supply related. The observed visual effect for power line or 60 Hz problems would be one slowly moving, wide horizontal bar across the screen. The appearance of two narrow bars indicates 120 Hz ripple from full-wave power supplies.						

Figure 1 part 3.

<table>
<tr><th rowspan="2">TEST PARAMETER (Test signal)</th><th rowspan="2">NTC-7</th><th colspan="5">EIA/TIA-250-C</th><th rowspan="2">SECTION (Add Rate)</th></tr>
<tr><th>Short Haul</th><th>Medium Haul</th><th>Sat/NTC-9</th><th>Long Haul</th><th>End-to-End</th></tr>
<tr><td rowspan="2">Signal-to-Periodic Noise Ratio

(300 Hz-4.2 MHz)</td><td>>52 dB (Sec. 3.18)</td><td>67 dB</td><td>62 dB</td><td>63 dB</td><td>58 dB</td><td>57 dB</td><td>6.3.3 (RSS)</td></tr>
<tr><td colspan="7">Note: The observed visual effect of an extraneous tone in the designated passband (300Hz to 4.2 MHz) is displayed on the received picture by multiple horizontal lines of varing intensity across the screen. This would occur below the horizontal line frequency (15.734 kHz) or at multiples of line frequency. The displayed video could include from one wide vertical column to several lines observed to move across the received picture. However, extraneous signals are seldom at line frequency. Therefore, diagonal lines racing across the display would be seen.</td></tr>
<tr><td rowspan="2">Signal-to-Impulse Noise Ratio

(100% Flat Field)</td><td>> 23 dB
1 per min
(Sec. 3.17)</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td colspan="7">Note: The visual effect noted for impulse noise displayed on a monochrome display would be black and white dots racing across the screen in a haphazard arrangement. In color systems, colored, as well as black and white, speckles would be seen.</td></tr>
<tr><td rowspan="2">Signal to Crosstalk Noise Ratio

(100% Flat Field)</td><td>> 60 dB
(Sec. 3.19)</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td colspan="7">Note: This distortion could typcially produce many different types of erratic patterns displayed on the received picture. These observations as noted in the displayed picture would be indicative of other video or audio signals or extraneous tones crosstalking into the desired signal.</td></tr>
<tr><td rowspan="2">Continuity (Availability) of Video Service</td><td>—</td><td>99.99%
for 37 dB S/N</td><td>99.99%
for 37 dB S/N</td><td>99.99%
for 37 dB S/N</td><td>99.99%
for 37 dB S/N</td><td>99.99%
for 37 dB S/N</td><td>6.4</td></tr>
<tr><td colspan="7">Note: Availability is based on a year's time or 53 minutes of outage time. An outage is considered to occur when is less than 37 dB S/N for terrestrial circuits, 46 dB S/N for satellite circuits or the performance, due to other impairments, is deemed to be "unusable."</td></tr>
<tr><td rowspan="2">Color Burst Amplitude

(40 IRE Nominal)</td><td>40 IRE
±4 IRE</td><td>±0.5 IRE
(RS-250-B)</td><td>±2 IRE
(RS-250-B)</td><td>±2 IRE
(RS-250-B)</td><td>±3 IRE
(RS-250-B)</td><td>±3 IRE
(RS-250-B)</td><td>—</td></tr>
<tr><td colspan="7">Note: Low amplitude of the color burst will result in unstable, or loss of, color synchronization. Loss of color sync will first be noticed with a red, blue and green rainbow effect across the entire video display. This is sometimes referred to as the barberpole effect. Further reduction of the color burst will turn a color transmission into a monochrome display.</td></tr>
<tr><td rowspan="2">Color Burst Relative Gain Error

(Relative to 100 IRE Luminance Signal)</td><td>±1 IRE</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td colspan="7">Note: High or low color saturation observed in the received picture may be the consequence of a misadjusted processing amplifier inserted in the signal channel but not be apparent by observation alone.</td></tr>
<tr><td rowspan="2">Color Burst Relative Phase Error

(Relative to chroma Channel Signal)</td><td>±1 IRE</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td colspan="7">Note: A change of variation in color tint of hue observed in the received picture. As a consequence of a misadjusted video processing amplifier inserted in the signal channel. The existence or non-existance of a processing amplifier in the signal path may not be apparent by visual observation alone.</td></tr>
<tr><td></td><td colspan="7">The above notes are based on those provided by Michael J. Shumila of the David Sarnoff Research Center in Princeton, N.J.</td></tr>
<tr><td></td><td colspan="7">SELECTED AUDIO PERFORMANCE STANDARDS FROM EIA/TIA-250-C</td></tr>
<tr><td>Audio Signal-to-Noise Ratio</td><td>—</td><td>66 dB</td><td>65 dB</td><td>58 dB</td><td>57 dB</td><td>56 dB</td><td>7.3
(RSS)</td></tr>
<tr><td>Audio Signal Availability</td><td>—</td><td>99.99%
for 25 dB S/N</td><td>99.99%
for 25 dB S/N</td><td>99.99%
for 34 dB S/N</td><td>99.99%
for 25 dB S/N</td><td>99.99%
for 25 dB S/N</td><td>7.6</td></tr>
<tr><td>Audio-to-Video Time Differential</td><td>—</td><td>+25 to
−40 ms</td><td>+25 to
−40 ms</td><td>+25 to
−40 ms</td><td>+25 to
−40 ms</td><td>+25 to
−40 ms</td><td>7.7
(linear)</td></tr>
</table>

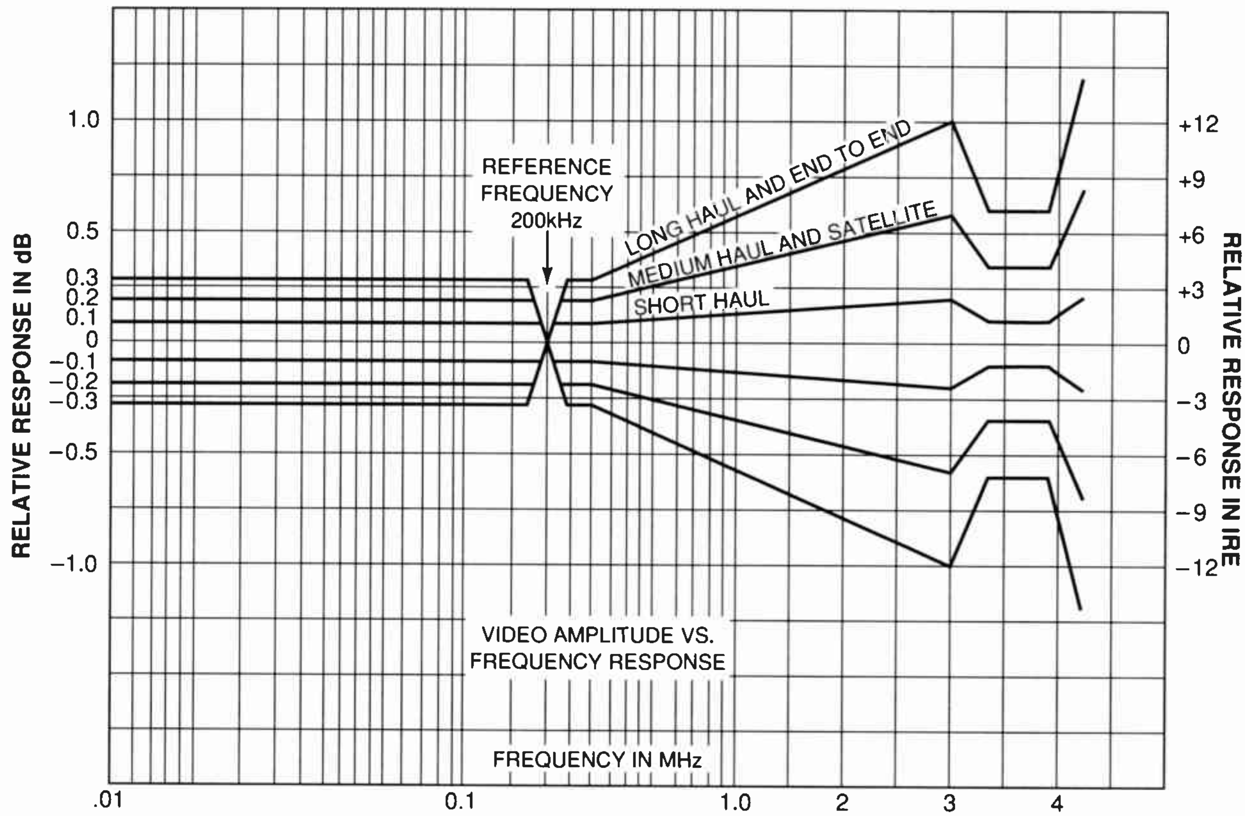


Figure 2. Video amplitude vs. frequency response characteristic.

Standard: The harmonic distortion of the audio frequency signal, including all harmonics up to 30 kHz, shall not exceed 0.5% over a frequency range of 50 Hz to 15 kHz.

This standard applies equally to short-haul, medium-haul satellite, long-haul, and end-to-end performance on the assumption that audio will be demodulated from its subcarrier only once in any system of interconnection.

Method of Measurement: The audio channel shall be operated at standard input and output test tone level and the measuring equipment shall terminate the circuit in a standard load impedance. A test tone of 400 Hz with less than 0.1% harmonic distortion shall be applied to the audio input of the system and rated maximum modulation established.

If pre-emphasis and de-emphasis is not used, distortion shall be measured at this level with readings taken on as many frequencies as necessary to ensue compliance.

If pre-emphasis and de-emphasis is used, the input level to the system shall be adjusted at each measurement to operate the system at rated maximum modulation. This will require reducing the input voltage level with increasing frequency as many dB as the pre-emphasis curve rises. The input and output level controls shall not be readjusted during this procedure.

Signal-to-Noise Ratio

Definition: The audio signal-to-noise ratio of the system is the ratio of root mean square (RMS) standard test tone voltage to the RMS noise voltage at the system output terminals. Noise is any extraneous output voltage in the frequency band from 50 Hz to 15 kHz.

Standard: The signal-to-noise ratio for the various transmission system configurations is given in Fig. 1 above.

Method of Measurement: The system shall be operated at standard input and output test tone levels, and the measuring equipment shall terminate the system in a standard load impedance. A test tone of 400 Hz shall be applied to the audio input of the system and rated maximum modulation established. The system output shall be measured with an RMS indicator. The test tone shall then be removed, the system input terminated with a standard load impedance, and the noise of the system measured with the RMS indicator.

NOTE 1: If the audio system is multiplexed with a video signal, the audio signal-to-noise measurement shall be made in the presence of a standard composite picture under standard input and output conditions. The modulated stairstep signal with a 50% APL shall be used for this test.

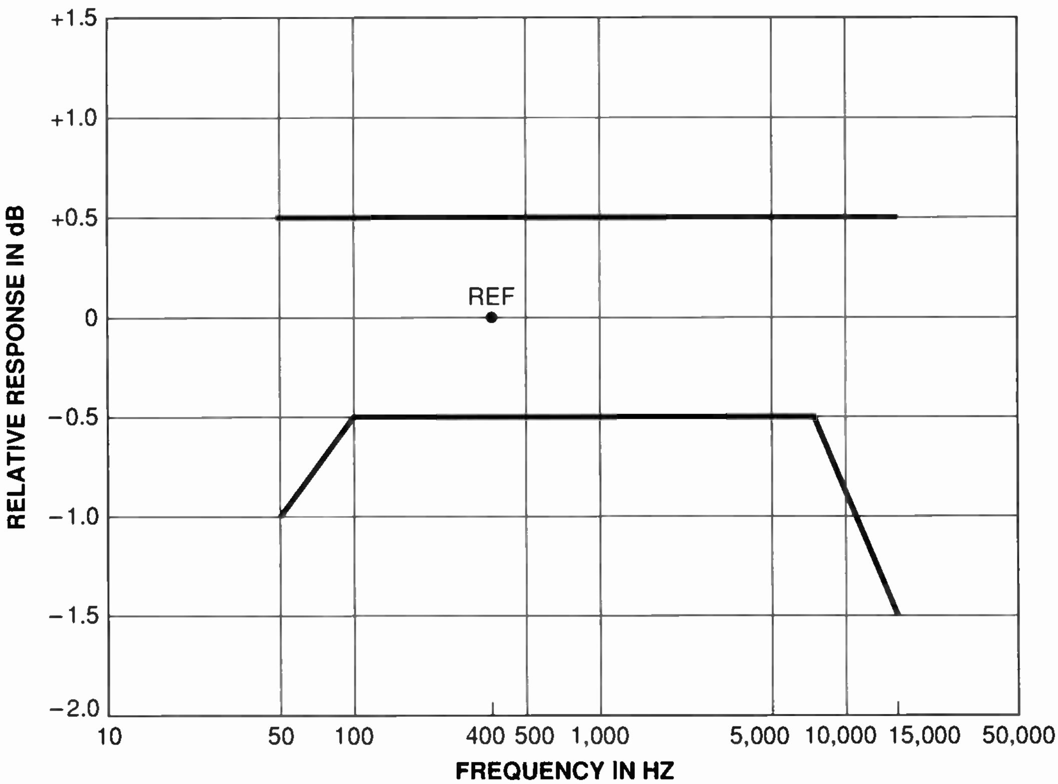


Figure 3. Audio amplitude vs. frequency response characteristic.

NOTE 2: Most direct dBm reading instruments are calibrated to read correctly across 600-ohm circuits. A correction factor of 10 log(Z_L/600) is to be applied to the reading when measuring across a termination of some other value (where Z_L is the load impedance).

Insertion Gain

Definition: The insertion gain of a transmission system is the ratio in dB of the output signal level with respect to the input signal level.

Standard: The preferred standard of insertion gain is 0 dB. In order to accommodate a standard input and output level, audio equipment may have input and output level adjustments to allow insertion gains of up to ±18 dB.

Method of Measurement: An audio signal 10 dB below nominal test signal level shall be applied to the input of the system. The output of the system is measured, and the difference between input and output expressed, in dB. Generally accepted audio measuring equipment shall be used.

Continuity of Audio Service

Definition: Continuity (or availability) of service means the time the signal is available for use and not unavailable or unusable. Availability is based on the time of one year. 99.99% represents 53 minutes per year during which the audio signal may have a signal-to-noise ratio of less than that specified in Fig. 1 or be degraded such that it is considered unusable.

Standard: The signal is considered unavailable for use if (1) the signal-to-noise falls below 25 dB for terrestrial circuits and 34 dB for satellite circuits, or (2) the signal is degraded to the extent that it is considered unusable.

Method of Measurement: Generally accepted audio measuring equipment shall be used to determine the availability of the audio channel.

Audio-to-Video Time Differential

Definition: Audio-to-video time differential is the departure from equality of the transmission time of associated audio and video signals.

NOTE: A time differential may be caused by the use of digital processing devices and frame stores (which can cause significant delays) in either the video or audio channel. A time differential can also occur if the audio and video signal take different paths.

Standard: The transmission time differential between audio signal and the associated video signal shall be with the range of 25 ms lead to 40 ms lag for all transmission systems.

Method of Measurement: The measurement shall be made using a storage oscilloscope with a dual-trace amplifier. Provisions shall be made at the input of the system to simultaneously initiate or interrupt the audio and video signals. The transmission time differential is read on the oscilloscope display at the output of the system.

Determining Rates of Accumulated Degradation of Performance

Each link in a transmission will add to the degradation of performance in various ways. Accumulated degradations, when two or more of the systems, identified in EIA/TIA-250-C, are connected in tandem, may be calculated from the following, assuming a baseband interconnection:

Accumulation Rate	*Accumulated Degradation D*
Linear	$D = d_1 + d_2 + d_3 + \ldots$
3/2 Power	$D = (d_1^{3/2} + d_2^{3/2} + d_3^{3/2} + \ldots)^{2/3}$
RSS	$D = (d_1^2 + d_2^2 + d_3^2 + \ldots)^{1/2}$

where: d_n = degradation due to the n^{th} subsystem

The expected rate of degradation accumulation for each parameter is given in the EIA/TIA Section column in the chart (Fig. 1).

The amount of degradation for each portion of the system (d_n) is determined as a percentage of the whole and summed according to one of the rates above. For example, in a three segment system (short-haul, satellite, short-haul systems in tandem), if the accumulation rate was linear for a particular parameter, the degradation for each segment would be additive. That is, 25% (1 IRE) + 50% (2 IRE) + 25% (1 IRE) = 100% (4 IRE) if 4 IRE units is the maximum degradation permitted for the parameter of interest.

If the rate is 3/2 power, then the same amount of degradation for a given parameter using the same values as above would be only 71.4% (or 2.8 IRE) of the maximum (4 IRE) permitted in the example. That is, the total is 2/3 power of the sum of the 3/2 power of each d value.

If the rate is RSS (root sum square), then the same amount of degradation for a given parameter using the same values as above would be only 72.1% (or 2.8 IRE) of the maximum (4 IRE) permitted in the example. That is, the total is the square root of the sum of the squares of each d value.

However, if the parameter is expressed in dB, the ratio must first be converted to a linear value before the particular accumulation rate is applied. The power addition curve in Fig. 4 can be used to combine dB ratios.

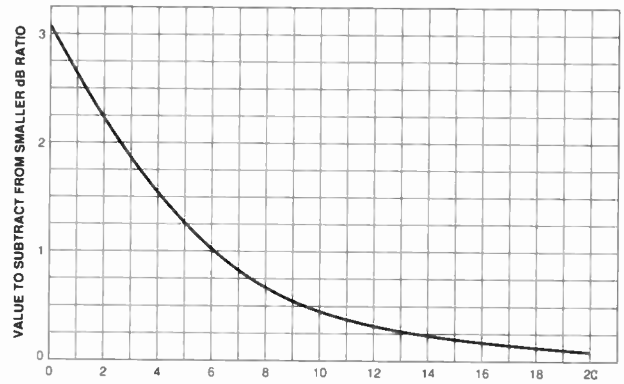


Figure 4. Power addition curve.

Additional Transmission Considerations

FM Terrestrial and Satellite Systems

1. Video Emphasis—FM transmission systems usually employ emphasis to reduce noise upon reception. Pre-emphasis, spectrum shaping, is performed in a pre-emphasis network prior to or in the FM modulator. It is a specified and standardized boost of high frequencies. De-emphasis is the complementary function of pre-emphasis and is performed in a network within or immediately following the demodulator. When combined, the resulting signal should have a "flat" amplitude/frequency characteristic.

   Emphasis introduces 3.1 dB improvement to the video unweighted signal-to-noise ratio (SNR) measurement. The emphasis characteristic standard is from CCIR Recommendation 405.
2. Noise Weighting—A noise weighting network is used for the video SNR to account for the psychovisual effect of the human eye with respect to the interfering noise spectrum found in FM modulation systems.

   The noise weighting network is used in conjunction with a 10 kHz high-pass filter and 5.0 MHz bandlimiting filter for the SNR and provides 8 to 12 dB improvement based on the spectrum shape of the noise.

   The noise weighting network is defined in CMTT Report 637. The filters specified above are readily available for use with noise measuring instruments.
3. Deviation—The deviation of frequency by the video signal is typically 8 MHz peak-to-peak with the white portion of the signal moving the carrier up in frequency and the black (or sync) portion moving the carrier down in frequency. The amount of deviation is somewhat dependent upon the particular FM modulation section employed and the particular user.

   When pre-emphasis is employed, the deviation is set to its nominal rated value (i.e., 8 MHz p-p) with a 1.0 volt sinewave at 761 kHz. This frequency is the reference point on the pre-emphasis curve.
4. Intermediate Frequency—Most video transmission systems employ a 70 MHz IF for interconnection to other equipment. The IF is generally available on most systems at an impedance of 75 ohms and a level of +1 to +5 dBm with a return loss of about 30 dB. Do not interconnect systems with different deviations.

5. FM Subcarriers—Audio and data associated with a particular video signal is usually carried on subcarriers above the 4.2 MHz video cutoff frequency. Filters are used in the video circuit to remove the subcarrier energy from the video signal. The filters used on the subcarriers to prevent contamination from video components may introduce degradation to the audio and data signals. Therefore, only phase equalized filters should be employed and tandem connection of baseband filters should be avoided.

   When choosing a subcarrier frequency, avoid the second harmonic of the color subcarrier ($3.58 \times 2 = 7.16$ MHz) that could introduce intermodulation distortions.

   The level of the subcarrier is typically 20 dB below maximum deviation produced by the video signal. For example, if the video signal deviates the carrier by 8 MHz then the audio subcarrier deviates the carrier by 20 dB less or 0.8 MHz ($20 \log d_v/d_a = \text{dB}$).
6. Audio Pre-emphasis—Audio pre-emphasis is employed on most audio subcarriers using the standard 75 μs curve. A 400 Hz tone is used to establish rated deviation at the maximum audio input level (typically +18 dBm). A different deviation may be used on different FM modulation sections. The deviation for most audio subcarriers is ±75 kHz p-p.

   When large numbers of subcarriers are employed, the deviation of each may have to be reduced to prevent contributing too much to the total deviation that also includes the video signal. Lowering the deviation will lower the SNR. Companding systems employing 2:1 or 3:1 compression are often used to improve SNR under normal operating conditions. However, the companding system may produce undesirable effects when the FM system is operated at or near threshold.

Audio Practices

While many different level and impedance standards are use in the wide variety of audio facilities employed by broadcasters, terrestrial and satellite FM audio channels enjoy only a few standards.

1. Reference or Normal Program Level—This level is normally used for lining up facilities. The level is typically set to be +8 dBm or 0 VU.
2. Test or Peak Program Level—This level is the maximum or peak level the system is capable of while still maintaining the performance objectives referred to in this chapter. It is the level at which SNR and harmonic distortion measurements are made. The level is typically +18 dBm or about 10 dB above the normal program level.
3. Program Clip Level—In order to provide enough headroom for most audio program material, some FM audio modulators will accommodate up to 6 dB more input than the peak level.
4. Input and Output Impedance—The input impedance to most audio FM modulators is 600 ohms balanced. The output circuit is designed to operate into 600 ohms balanced.
5. Audio Limiting—Because FM subcarriers are easily over-deviated by peaks in program material it is common practice (but not standardized) to employ limiters at the +18 dBm level. At this level the limiter should not permit more than a one or two dB increase in deviation for an additional 10 dB increase in audio input level. The result of audio channel over-deviation is heavy distortion at the receiving end and interference to the video channel or other subcarrier channels.

PARAMETERS, MEASURING TECHNIQUES AND PERFORMANCE OBJECTIVES[2]

General and Waveform Technology

This section describes the transmission parameters, measuring techniques and performance objectives that are applicable to all video transmission facilities leased from the Bell System.

It should be noted that except where full-field test signals are essential to the measurement of a particular transmission parameter, e.g., field-time waveform distortion, the measuring technique and its associated performance objective specified for each transmission parameter apply equally to both vertical interval test signals (VITS) and full-field test signals. Furthermore the performance objectives apply irrespective of the average picture level (APL) within the APL range 10% to 90%. This is an important point to remember when making VITS measurements, particularly during program transmission periods where control cannot be exercised over the APL value of picture signal. Many of the transmission parameters can be markedly affected by APL variations. Accordingly, the operator should allow sufficient time when making VITS measurements to ensure a good portion of the APL range is explored by the picture signal before recording the test signal measurement. In every case the highest distortion measurement observed during this period should be recorded and then compared with the performance objective to determine whether or not the facility is within the stated objective. (See Appendix A).

Waveform measurement techniques described in this report are based on the IRE scale units of measurement (see Fig. 1) except where specifically noted otherwise. The waveform technology used throughout the report is in accordance with the definitions shown in Fig. 2 wherein the standard composite color video signal is defined.

The two principal test signals that are required to conduct the various measurements described in this report are:

A. The composite test signal shown in Fig. 3, which consists of a line bar, a 2T pulse, a chrominance pulse, and a five-riser staircase signal.

B. The combination test signal shown in Fig. 4, which consists of a white flag, a multiburst and a three-level chrominance signal.

When conducting in-service VITS measurements the composite test signal shall be inserted on line 17, field 1, and the combination test signal shall be inserted on line 17, field 2.

NOTE: The performance objectives specified in this report pertaining to television transmission facilities, apply only when the recommended test signals are originated directly at the designated broadcaster/TELCO interface. No broadcaster maintained and operated equipment shall be used between the test signal generator output and the interface point.

Insertion Gain

Definition: Insertion gain is defined as the difference, in IRE units, between the peak-to-peak amplitude of a specified test signal at the receiving end of a television facility and the nominal amplitude of the test signal at the sending end.

Measurement: The line bar portion of the composite test signal shown in Fig. 5 is used when measurement insertion gain. The test signal's amplitude must be accurately adjusted at the sending end prior to then commencement of the test. Similarly, the waveform monitor at the receiving end should be properly calibrated.

Following the above, the peak-to-peak amplitude of the test signal should be measured, in IRE units, at the receiving end using the points approximating to b_1 and b_2 as shown in Fig. 5. The difference between the measured amplitude of the test signal and its nominal amplitude of 100 IRE units is the insertion gain of the television facility. If the measured amplitude is less than 100 IRE units then the insertion gain should be recorded with a negative sign. Such an example is shown in Fig. 6.

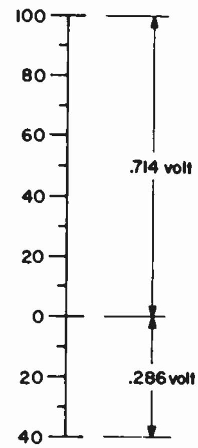


Figure 5. The IRE scale units. (For a 1V P-P composite signal.)

Performance Objectives:

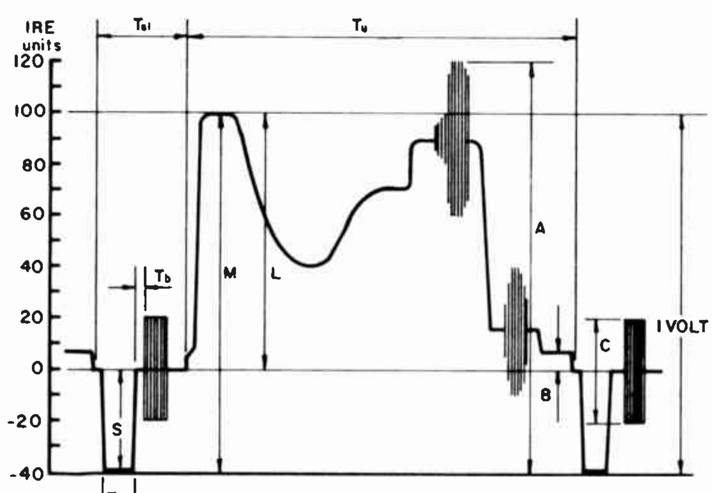
A. The insertion gain shall not exceed zero ±3 IRE units.

B. Variations in insertion gain shall not exceed:
   1. Short period (e.g., 5 seconds): ±1 IRE unit.
   2. Medium period (e.g., 1 hour): ±2 IRE units.

These variations in gain are permissible only within the range specified in objective A.

NOTE 1: The performance objectives shown above apply equally to full-field and in-service VITS measurements.

NOTE 2: Insertion gain may also vary as a function of APL. This is termed dynamic gain, and is covered in section 3.12.



WAVEFORM TERMINOLOGY

A: The peak-to-peak amplitude of the composite color video signal
B: The difference between black level and blanking level (set-up)
C: The peak-to-peak amplitude of the color burst
L: Luminance signal—nominal value
M: Monochrome video signal peak-to-peak amplitude (M = L + S)
S: Synchronizing signal—amplitude
T_b: Duration of breezeway
T_{sl}: Duration of line blanking period
T_{sy}: Duration of line synchronizing pulse
T_u: Duration of active line period

Figure 6. The standard composite color video signal.

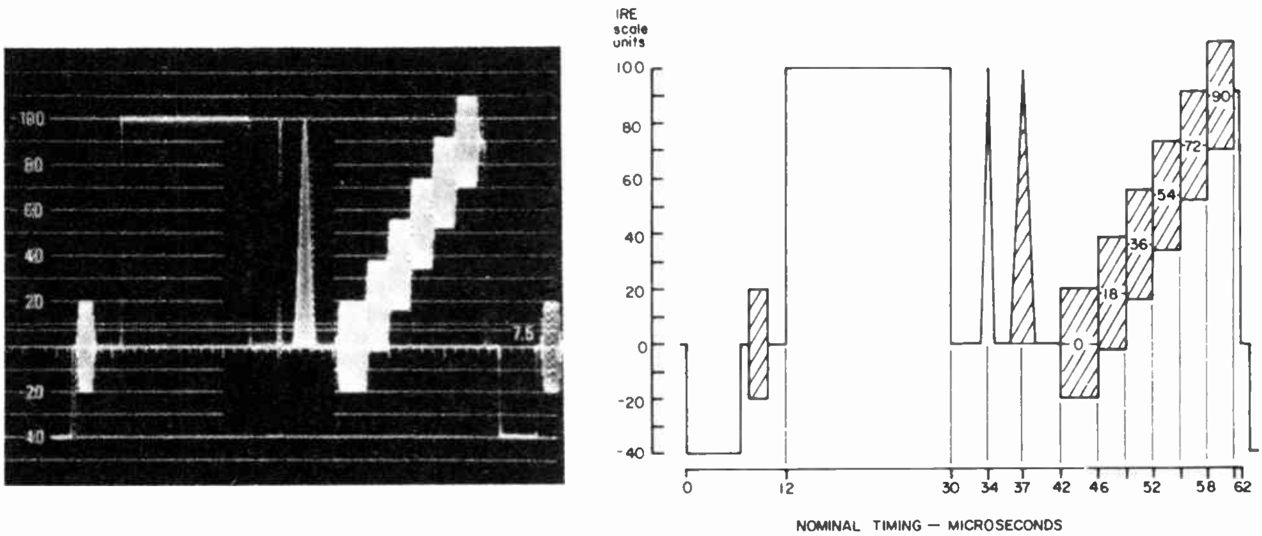


Figure 7. The composite test signal.

Field-Time Waveform Distortion

Definition: When a television test signal having a period of one television field, and of reference white amplitude is applied to the sending end of a television facility, the field-time waveform distortion is defined as the change in shape of the top of the test signal at the receiving end. The beginning and end of the test signal, equivalent to a few scanning lines, are excluded from the measurement.

Measurement: The field bar test signal shown in Fig. 7 is used when measuring field-time waveform distortion. The test signal's amplitude must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the magnitude of the distortion is obtained by measuring, in IRE units, the peak-to-peak change in amplitude of the bar top with the amplitude of the bar center adjusted to 100 IRE units. In order to avoid leading and trailing overshoots, the first and last 250 microseconds (approximately four television lines) are ignored in this measurement. An example of field-time waveform distortion is shown in Fig. 8.

Performance Objective: The peak-to-peak excursions of the bar top shall not exceed 4 IRE units.

Line-Time Waveform Distortion

Definition: When a television test signal having a period of one television line and of reference white

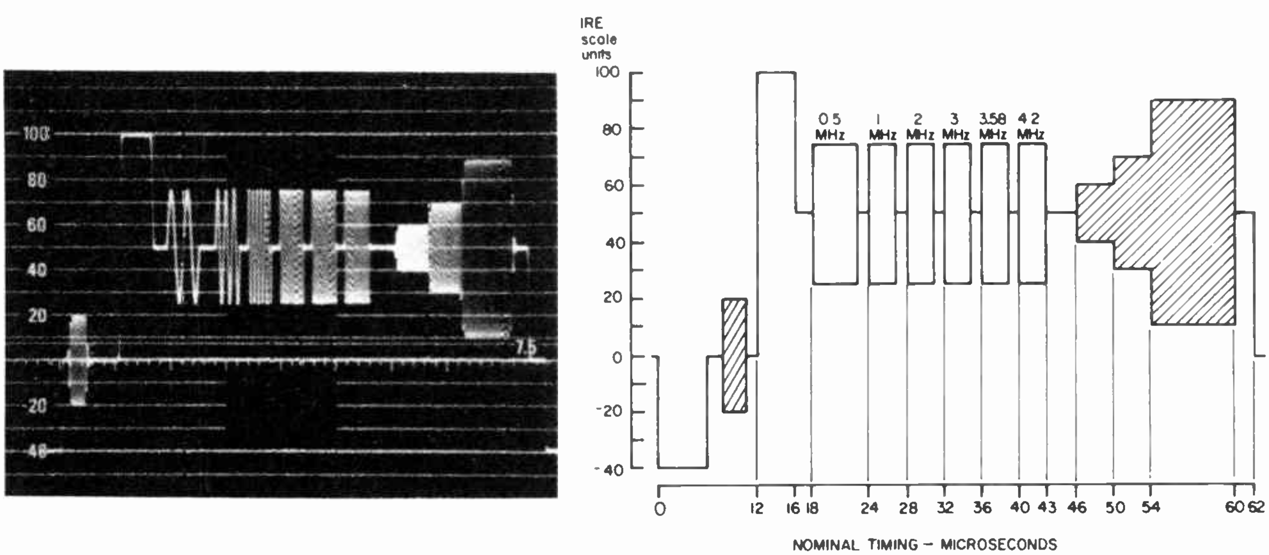


Figure 8. The combination test signal.

amplitude is applied to the sending end of a television facility, the line-time waveform distortion is defined as the change in shape of the top of the test signal at the receiving end. The beginning and end of the test signal, equivalent to a few pictures elements, are excluded from the measurement.

Measurement: The line bar test portion of the composite test signal shown in Fig. 9 is used when measuring line-time waveform distortion. The test signal's amplitude must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the magnitude of the distortion is obtained by measuring, in IRE units, the peak-to-peak change in amplitude of the bar top with the amplitude of the center adjusted to 100 IRE units. The first and last one microsecond are ignored in this measurement. An example of line-time waveform distortion is shown in Fig. 10.

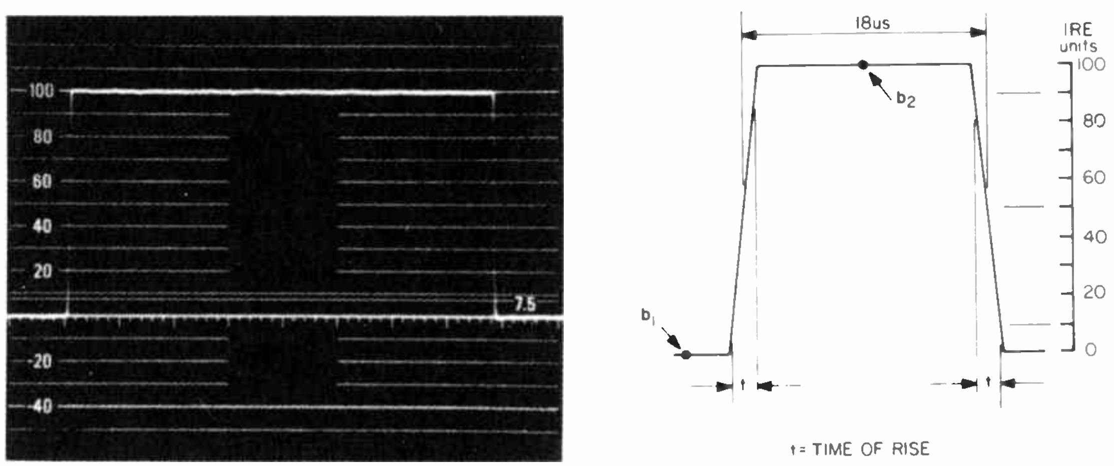
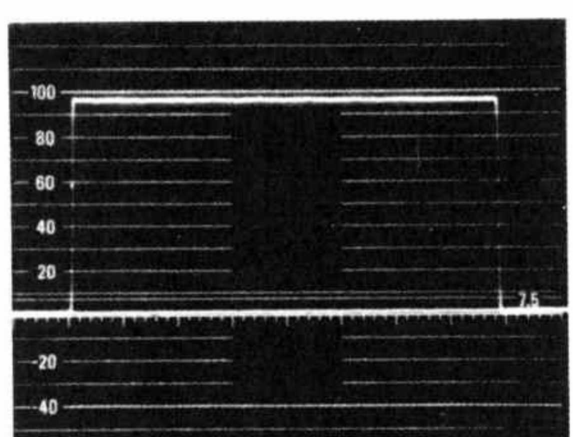


Figure 9. The line bar test signal.

Generator Output Specifications

Peak amplitude	:	100 ± 0.5 IRE units *(reference white amplitude)*
Line-time waveform distortion	:	less than 0.3 IRE units
Time of rise and time of fall of bar edges (10%-90%)	:	t = 125 ± 5 nanoseconds with integrated sine-squared shape



Insertion gain is 3 IRE units

Figure 10. An example of negative insertion gain.

Performance Objective: The peak-to-peak excursions of the bar top shall not exceed 4 IRE units.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

Short-Time Waveform Distortion

Definition: If a short pulse or rapid step function of reference white amplitude and defined shape is applied to the sending end of a television facility, the short-time waveform distortion is defined as the departure of the output pulse or step from its original shape. The choice of the half amplitude duration of the pulse or the rise-time of the step is determined by the nominal cutoff frequency of the television facility.

Measurement: The line bar portion of the composite test signal shown in Fig. 9 and the 2T pulse test signal shown in Fig. 11 are used when measuring short-time waveform distortion. The test signal's amplitude must be accurately adjusted at the sending end prior to commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the amplitude of the 2T pulse test signal is measured, in IRE units, having previously adjusted the amplitude of the line bar test signal to exactly 100 IRE units.

The peak-to-peak variations within the 1 microsecond intervals on either side (preceding and following) the T-step transitions (rise and fall) are then measured with the amplitude of the line bar test signal adjusted to 100 IRE units as measured between blanking and a point approximately 2 microseconds from the bar edge. (A graticule method is currently under study by IEEE.) Examples of short-time waveform distortion are shown in Figs. 12 and 13.

Performance Objective:

A. The 2T pulse amplitude shall be 100 ±6 IRE units.
B. The peak-to-peak amplitude variations preceding or following the T-step transitions to the line bar test signal shall not exceed 10 IRE units.

NOTE: The performance objectives shown above apply equally to both full-field and in-service VITS measurements.

Chrominance-Luminance Gain Inequality

Definition: When a test signal having defined luminance and chrominance components is applied to the sending end of a television facility, the chrominance-luminance gain inequality is defined as the change in amplitude at the receiving end of the color component of the test signal relative to the luminance component.

Measurement: The chrominance pulse portion of the composite test signal shown in Fig. 14 is used when measuring chrominance-luminance gain inequality.[3] The test signal's amplitude must be accurately adjusted at the receiving end should be properly calibrated.

Following the above, the amplitude of the Chrominance Pulse Test Signal is measured in IRE units, having previously adjusted the amplitude of the line bar test signal to exactly 100 IRE units. This method is accurate to within 2% with up to 300 nanoseconds of chrominance-to-luminance delay present. The convention of Fig. 17B shows how the chrominance pulse will look with different types of gain and delay distortion. If harmonic distortion is present on the chroma pulse, as evidenced by multiple irregularities of the baseline, this method is invalid and an accurate measurement cannot be made. Methods to make measurements in the presence of harmonic distortion are presently under study.

An example of low chrominance amplitude is shown in Fig. 15.

Performance Objective: The amplitude of the chrominance pulse shall be 100 ±3 IRE units.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

Chrominance-Luminance Delay Inequality

Definition: When a test signal having defined luminance and chrominance components is applied to the sending end of a television facility, the chrominance-luminance delay inequality is defined as the change in relative timing, at the receiving end, of the chrominance component of the test signal relative to the luminance component.

Measurement: The chrominance pulse portion of the composite signal shown in Fig. 14 is used when measuring chrominance-luminance delay inequality.[4] The test signal's amplitude must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the amplitude of the chrominance pulse test signal should be adjusted to exactly 100 IRE units. The nomogram show in Fig. 17A should be used to compute the magnitude of the chrominance-luminance delay inequality. If the chrominance component of the test signal starts with a positive going lobe then the chrominance-luminance delay inequality should be recorded as delayed chroma. The delay inequality can also be computed by the formula: CLDI (RCT) in ns $= 20\sqrt{Y_1 \times Y_2}$. If harmonic distortion is present, as evidenced by multiple irregularities of the baseline, this method is invalid and an accurate measurement cannot be made. Methods to determine chrominance-luminance delay inequality in the presence of harmonic distortion are currently under study.

An example of chrominance-luminance delay inequality is shown in Fig. 16 below.

Performance Objective: The chrominance-luminance delay inequality shall be no greater than 75 nanoseconds, advanced or delayed chroma.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

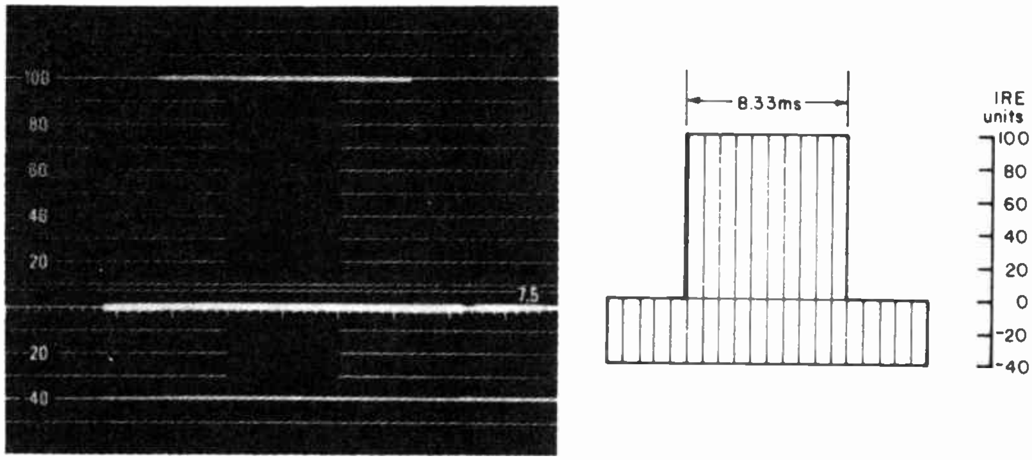
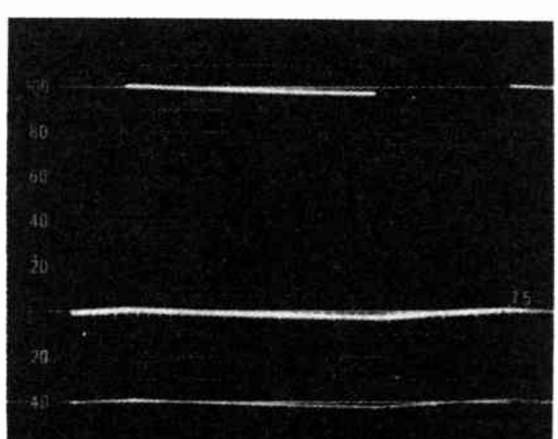


Figure 11. The field bar test signal.

Generator Output Specifications

Peak amplitude of luminance signal	:	100 ± 0.5 IRE units
Peak amplitude of synchronizing signal	:	40 ± 0.5 IRE units
Field-time waveform distortion	:	less than 0.3 IRE units
Horizontal component	:	100 IRE unit flat field of 52.45 microsecond nominal duration

NOTE: This signal should be generated with field and line synchronizing pulses.



Field-Time Distortion is 3 IRE units. (+1, −2)

Figure 12. An example of field-time distortion.

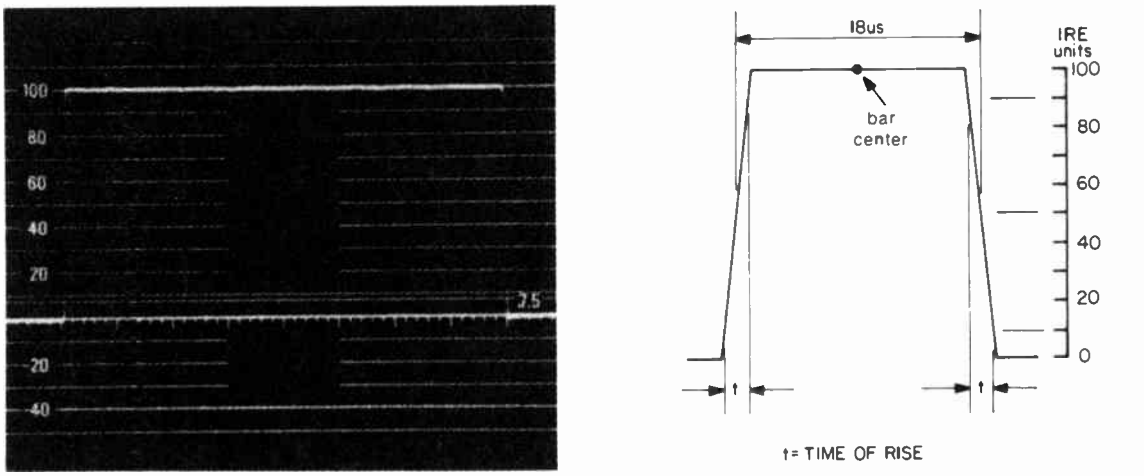
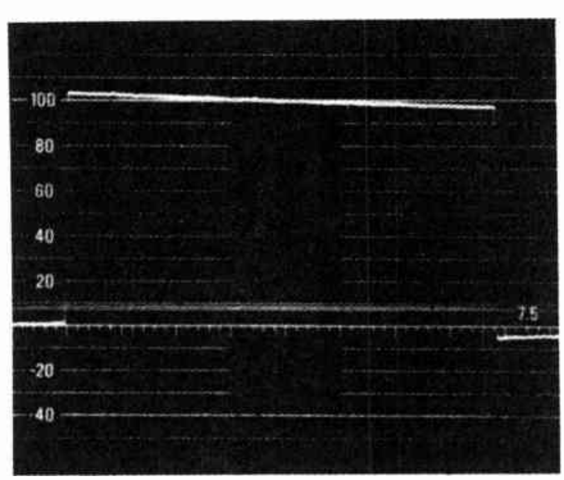


Figure 13. Line bar test signal.

Generator Output Specifications

Peak amplitude	:	100 ± 0.5 IRE units
Line-time waveform distortion	:	less than 0.3 IRE units
Time of rise and time of fall of bar edges (10%-90%)	:	t = 125 ± 5 nanoseconds with integrated sine-squared shape



Line-Time Distortion is 4 IRE units. (+2, −2)

Figure 14. An example of line-time distortion.

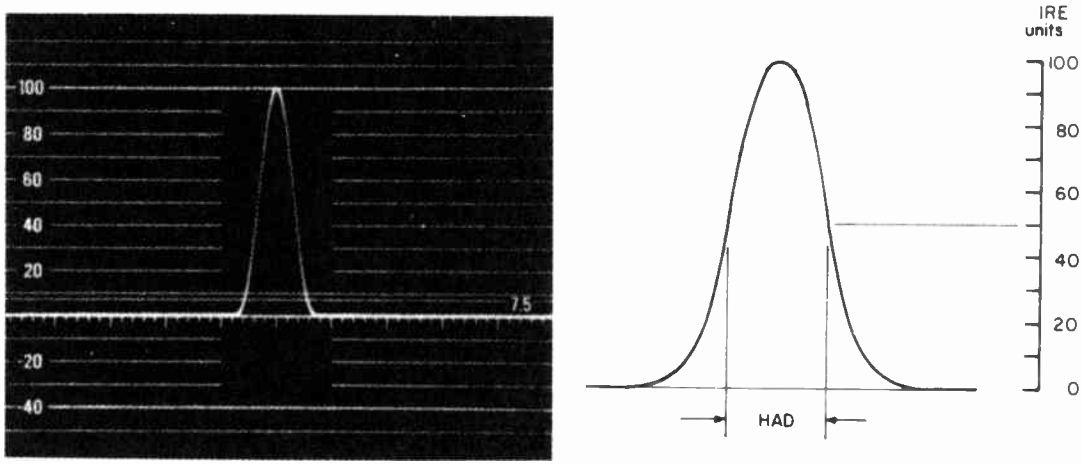


Figure 15. The 2T pulse test signal.

Generator Output Specifications

Peak amplitude	:	100 ± 0.5 IRE units
Half amplitude duration (HAD)	:	250 ± 10 nanoseconds

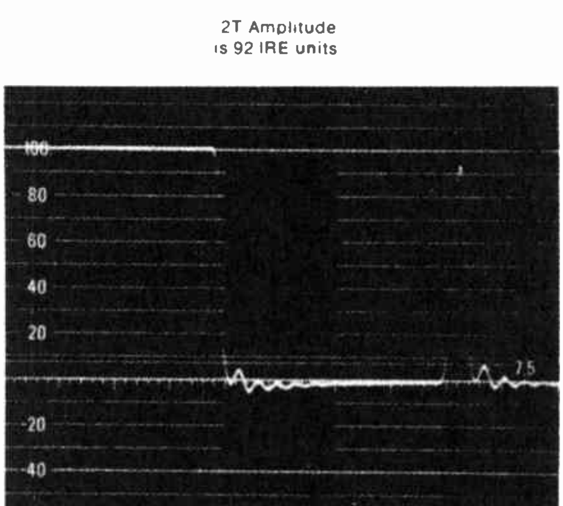


Figure 16. An example of short-time distortion as it affects the 2T pulse amplitude.

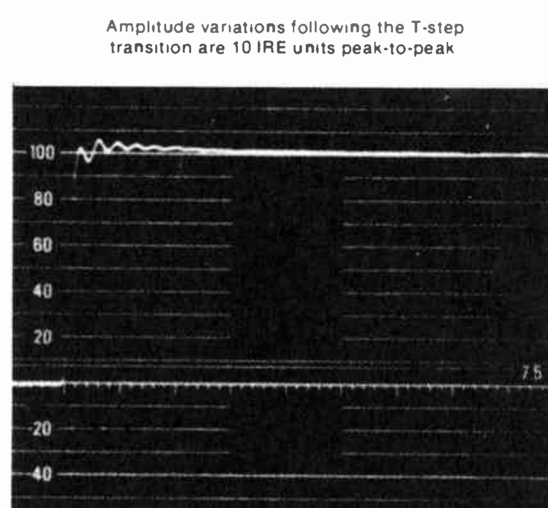


Figure 17. An example of short-time distortion as it affects the edge response of the line bar.

Gain/Frequency Distortion

Definition: The gain/frequency distortion of a television facility is defined as the variation in gain over the frequency band extending from the television field repetition frequency to the nominal cutoff frequency of the facility, relative to the gain at a suitable reference frequency.

Measurement: The multiburst portion of the combination test signal shown in Fig. 18 is used when measuring gain/frequency distortion in the range from 500 kHz to 4.2 MHz. The test signal's amplitude must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the amplitude of the white flag should be adjusted to exactly 100 IRE units and then the peak-to-peak amplitude of each burst frequency should be measured and recorded. An example of gain/frequency distortion is shown in Fig. 19.

Performance Objective: With the white flag amplitude adjusted to 100 IRE units:

A. All frequency burst amplitude shall be 50 +3, −5 IRE units.
B. Color burst amplitude shall be 40 ±4 IRE units.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

Luminance Nonlinear Distortion

Definition: For a particular value of average picture level (APL), the nonlinear distortion of the luminance signal is defined as the departure from proportionality between the amplitude of a small unit step function at the sending end of a television facility and the corresponding amplitude at the receiving end, as the level of the step is shifted from blanking level to white level.

Measurement: The modulated five-riser staircase portion of the composite test signal shown in Fig. 20 is used when measuring luminance nonlinear distortion. The test signals amplitude at each step level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal is passed through a differentiating and shaping network of the type shown in Fig. 21 with the output of the network connected to the waveform monitor being used for the measurement. The function of the network is that of transforming the test signal into a train of five pulses of equal amplitude under zero distortion conditions. The gain of the waveform monitor should be increased to the point where the largest pulse amplitude is 100 IRE units and then amplitude of the smallest pulse can be measured and recorded. This is the luminance nonlinear distortion at 50% APL. The above measurement procedure should be repeated using the same test signal transmitted on every fifth television line with intermediate lines set at blanking level for a 10% APL value and then at peak white for a 90% APL value.[5]

An example of luminance nonlinear distortion is shown in Fig. 22.

Performance Objective: With the largest pulse amplitude adjusted to 100 IRE units, the difference between it and the smallest pulse amplitude shall not be greater than 10 IRE units at 10%, 50%, or 90% APL.

Chrominance Nonlinear Gain Distortion

Definition: For fixed values of luminance level and average picture level, the nonlinear gain distortion of the chrominance signal is defined as the departure from proportionality between the amplitude of the chrominance subcarrier at the sending end of a television facility and the corresponding amplitude at the receiving end as the amplitude of the subcarrier is varied from a specified minimum value to a specified maximum value.

Measurement: The three-level chrominance portion of the combination test signal shown in Fig. 23 is used when measuring chrominance nonlinear gain distortion. The test signals' amplitude at each chrominance level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal is passed through a high-pass filter network[6] and the output of the network is connected to the waveform monitor being used for the measurement. The gain of the waveform monitor is adjusted to the point where the middle *subcarrier* amplitude is exactly 40 IRE units and then the amplitude of the largest and smallest subcarrier levels are measured and recorded. An example of chrominance nonlinear gain distortion is shown in Fig. 24.

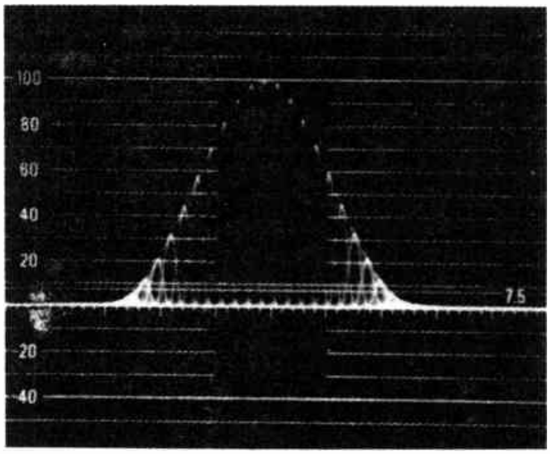
Performance Objective: With the middle subcarrier amplitude adjusted to 40 IRE units the amplitude of the other two subcarrier levels shall be:

A. The smallest subcarrier amplitude shall be 20 ±2 IRE units.
B. The largest subcarrier amplitude shall be 80 ±8 IRE units.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

Chrominance Nonlinear Phase Distortion

Definition: For fixed values of luminance signal level and average picture level, the nonlinear phase distortion of the chrominance signal is defined as the variation in phase of the chrominance subcarrier at the receiving end of a television facility as the amplitude of the subcarrier is varied from a specified minimum value to a specified maximum value.



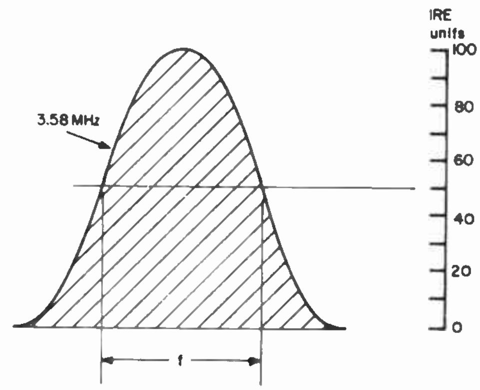
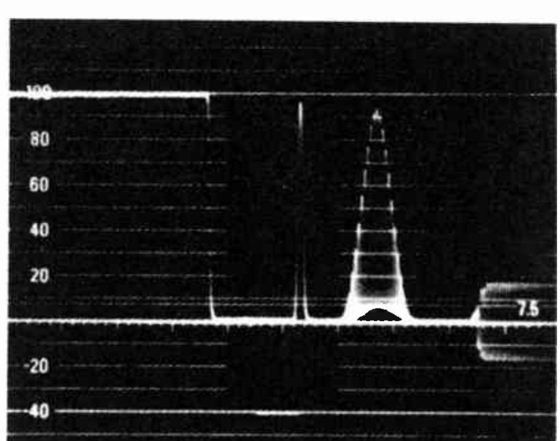


Figure 18. The chrominance pulse test signal.

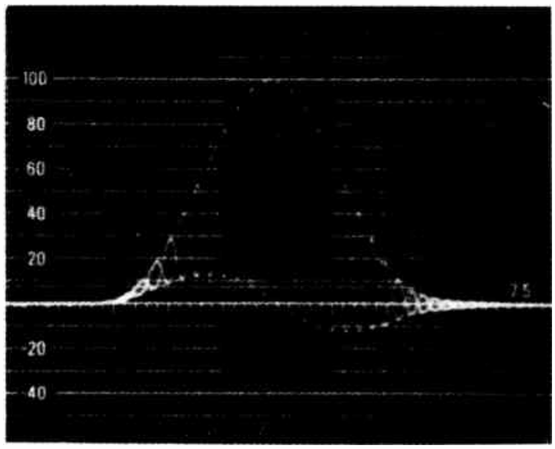
Generator Output Specifications

Peak amplitude	:	100 ± 0.5 IRE units
Half amplitude duration	:	t = 1562.5 ± 50 nanoseconds
Inherent chrominance luminance		
a) gain inequality (RCL)	:	less than ± 0.5 IRE (± 1%)
b) delay inequality (RCT)	:	less than 5 nanoseconds, delayed or advanced
Subcarrier harmonic distortion	:	less than 1%
Irregularities in the pulse base line	:	less than ± 0.5 IRE units



Chrominance Pulse Amplitude is 94 IRE units (RCL = − 12%) no delay inequality is present

Figure 19. An example of chrominance-luminance gain inequality.



CLDI(RCT) is 240 Nanoseconds Delayed Chroma with no gain inequality present

Figure 20. An example of chrominance-luminance delay inequality.

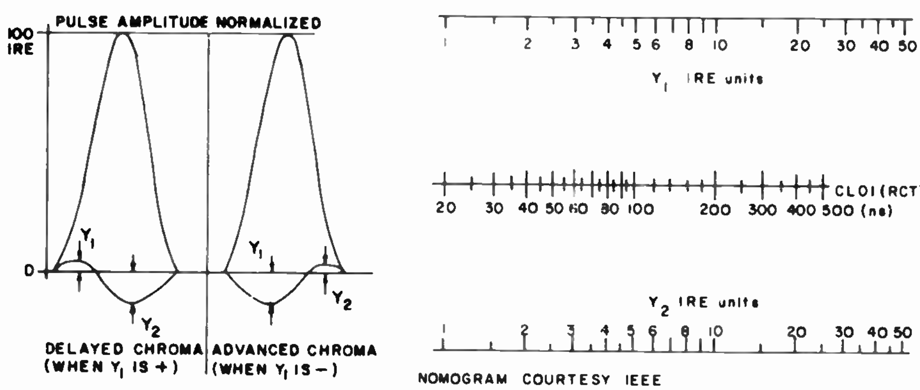


Figure 21A. Chrominance-luminance delay monogram with measurement convention.

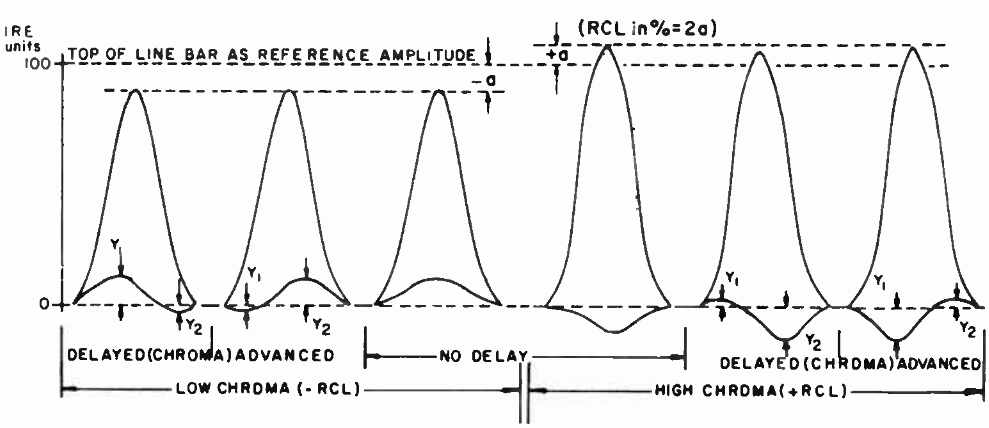


Figure 21B. Chrominance-luminance gain measurement convention.

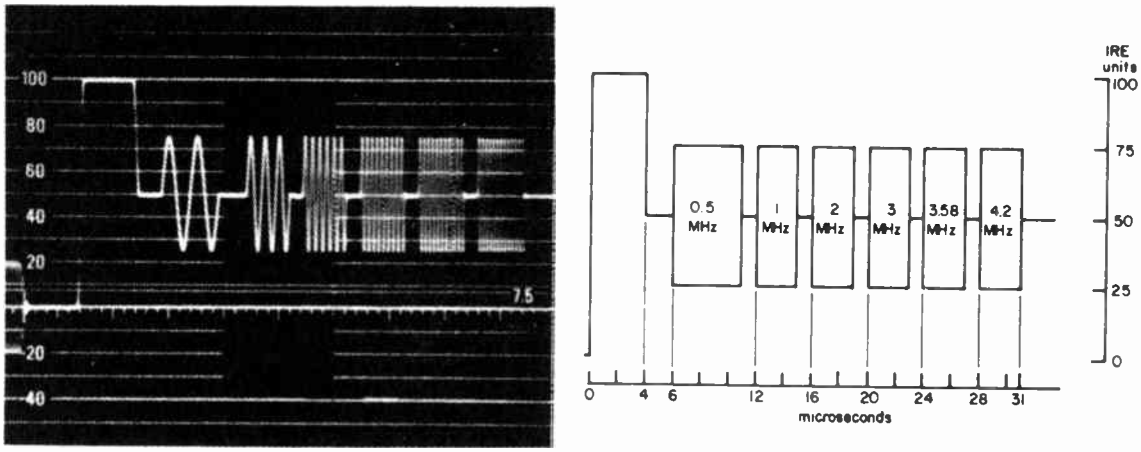


Figure 22. The multiburst test signal.

Generator Output Specifications

A) **White Flag**

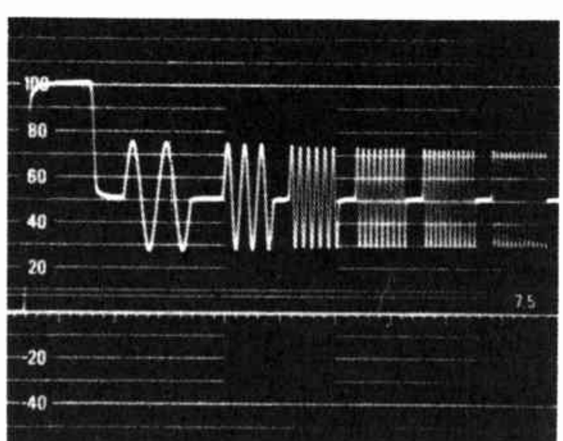
Peak amplitude	:	100 ± 1 IRE unit.
Time of rise and time of fall of flag edges	:	derived from the shaping network of the 2T sine-squared pulse.
Overshoot	:	less than ± 1 IRE unit.
Tilt	:	less than ± 1 IRE unit.

B) **Multiburst Frequencies**

All half-amplitude points of all burst frequencies	:	50 ± 1 IRE unit.

(The starting point of each burst frequency shall be at zero phase of each sinewave.)

Peak-to-peak amplitude of all bursts		50 ± 0.5 IRE units.



In this example, high frequency roll-off is shown. The burst amplitudes vary from 50 IRE units in the 500 kHz burst to 42 IRE units in the 4.2 MHz burst.

Figure 23. An example of gain/frequency distortion.

Measurement: The three-level chrominance portion of the combination test signal shown in Fig. 23 is used when measuring chrominance nonlinear phase distortion. The test signal's amplitude and phase at each chrominance level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor and phase comparator (e.g., vectorscope) at the receiving end should be properly calibrated.

Following the above, the test signal should be fed through a high-pass filter network to the phase comparator (or directly to the vectorscope). Under zero distortion conditions the phase at each level of the three-level chrominance test signal should be $-90°$ relative to the phase of the color burst (see Fig. 25).

The phase of the three levels of the test signal should be measured relative to the phase of the color burst and recorded.

The peak-to-peak variation of the phase of the three-level chrominance test signal is the difference between the largest and smallest readings obtained.

It should be noted that processing amplifiers can change the phase of the test signal chrominance information relative to the phase of the color burst. Care should be taken to ensure a processing amplifier is not in the circuit during tests of this kind on television facilities.

An example of chrominance non-linear phase distortion is shown in Fig. 26.

Performance Objective:. The peak-to-peak phase variation of the three-level chrominance test signal shall not exceed 5°.

NOTE: The performance objective shown above applies equally to both full-field and in-service VITS measurements.

Dynamic Gain Distortions

Definitions: If, at the sending end of a television facility, the average picture level (APL) of a video signal is stepped from a low value to a high value, or vice versa, the operating point within the transfer characteristic of the system may be affected and introduce various distortions on the receiving end. This section covers two such distortions known, respectively, as long-time waveform distortion and dynamic gain.

Long-Time Waveform Distortion (Bounce)[7]

Measurement: The flat-field 'bounce' test signal, as shown in Fig. 27, is used when measuring long-time waveform distortion. The test signal is switched between 10 and 90 IRE units, at an appropriate rate, while the following measurements are made on a properly calibrated de-coupled oscilloscope or waveform monitor, with any internal clamp in the instrument disabled.

With the waveform monitor on the field rate, observe any instantaneous peak excursion of the blanking level when the switch from 10 to 90 IRE, or vice versa, is made and record in IRE units. Also, observe the time necessary for blanking to settle to within 1 IRE unit of its final position. This may be as long as several seconds peak-to-peak changes before stabilizing.

If an oscilloscope is used, a sweep time of 1 sec./cm will generally enable the entire transient to be observed. The Tektronix 1480 MOD. 6 or 7 is presently the only waveform monitor with a slow sweep feature.

A photograph of this slow sweep is the best method of observing these distortions.

An example of long-time waveform distortion is shown in Fig. 28.

Performance Objectives: Peak overshoot at blanking shall be no greater than 5 IRE units and the time for blanking to settle to within 1 IRE unit of its final level shall be less than 1 sec.

NOTE: This distortion only occurs when APL change is accompanied be a change in dc component.

Dynamic Gain[8]

Measurement: Switch the flat-field test signal out of the 'bounce' mode and select 50% APL (Fig. 29A). Next, switch the waveform monitor to display the composite test signal on line 17, field one (Fig. 3). If necessary, normalize the gain of the waveform monitor for 100 IRE units at bar center. This is the reference amplitude for the following measurements.

Next, switch the test signal to 10% APL and observe and record, in IRE units, any change in amplitude of the line bar and/or synchronizing pulses. Finally, select 90% APL and record again any change in amplitude of the line bar and/or synchronizing pulses from the reference amplitude at 50% APL.

Examples of dynamic gain distortion are shown in Fig. 29B and C.

Performance Objectives: At 100% or 90% APL, dynamic gain-picture shall not exceed ±3 IRE units and dynamic gainsync shall not exceed ±2 IRE units referenced to the amplitude at 50% APL.

NOTE: This distortion may occur with or without a change in dc component accompanying the APL change.

Differential Gain

Definition: If a small constant amplitude of chrominance subcarrier, superimposed on a luminance signal, is applied to the sending end of a television facility, the differential gain is defined as the change in amplitude of the subcarrier at the receiving end as the luminance varies from blanking level to white level, the average picture level being maintained at a particular value.

Measurement: The modulated five-riser staircase portion of the composite test signal shown in Fig. 30 is used when measuring differential gain. The test signal's amplitude at each step level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal should be fed through a high-pass filter network[9] and the output of the network connected to the waveform monitor being

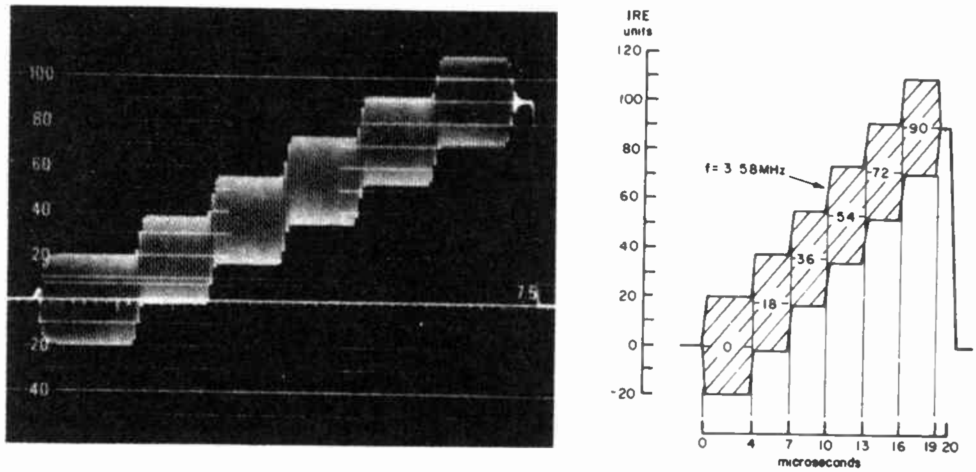


Figure 24. The modulated five-riser staircase test signal.

Generator Output Specifications

Peak amplitude of each riser	:	shall be such that the luminance non-linear distortion shall be greater than 99.5 IRE units.
Rise time of each riser	:	250 ± 10 nanoseconds
Step duration	:	3.0 ± 0.1 usecs; final step 4.0 ± 0.1 microseconds (4.0 ± 0.1 microseconds at blanking level).
Tilt on any step	:	less than 0.3 IRE units
Overshoot on any step	:	less than 0.3 IRE units

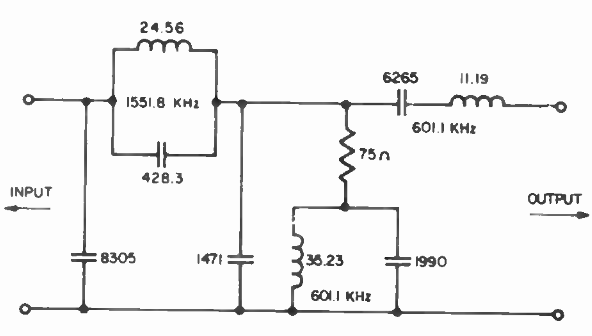


Figure 25. Differentiating and shaping network for luminance nonlinearity measurement.

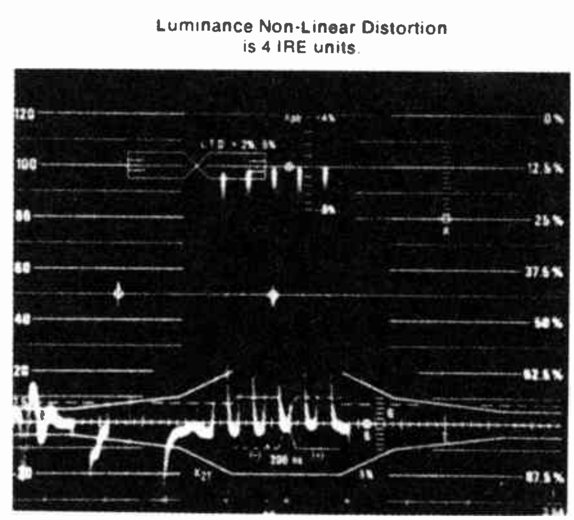


Figure 26. An example of luminance nonlinear distortion.

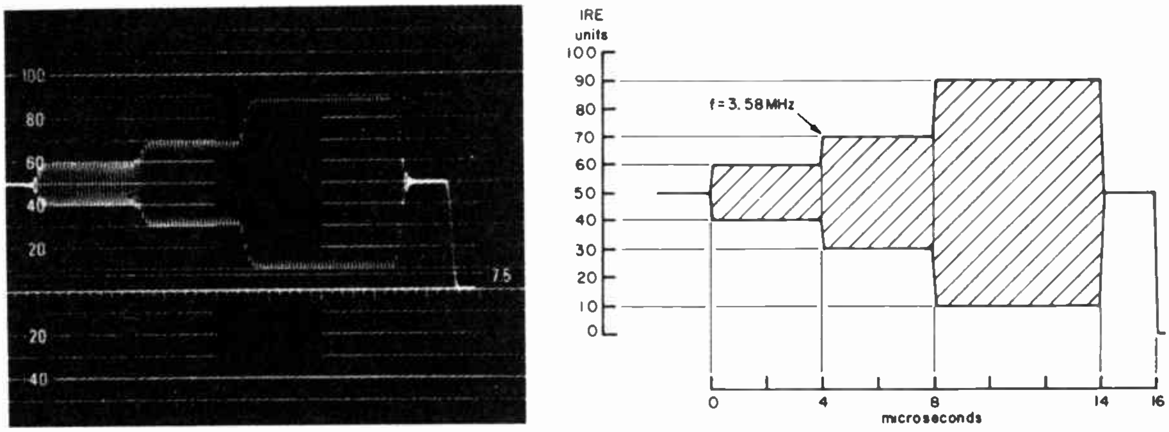
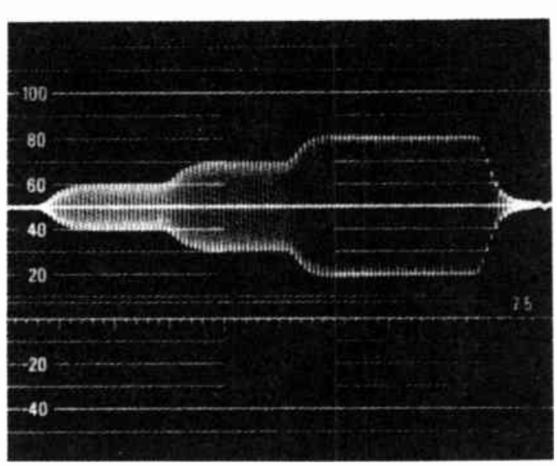


Figure 27. The three-level chrominance test signal.

Generator Output Specifications

Peak-to-peak amplitudes of the 3-levels	:	20,40 and 80 IRE units ± 0.5 IRE units
Sub-carrier frequency phase	:	90° ± 1° relative to reference burst; 0° ± 0.2° any one relative to the other two.
Rise and fall of chrominance envelopes	:	400 ± 25 nanoseconds
Duration of entire signal	:	14 ± 0.5 microseconds



The chrominance non-linear gain distortion is evidenced by the largest burst being equal to 64 IRE units.

Figure 28. An example of chrominance nonlinear gain distortion.

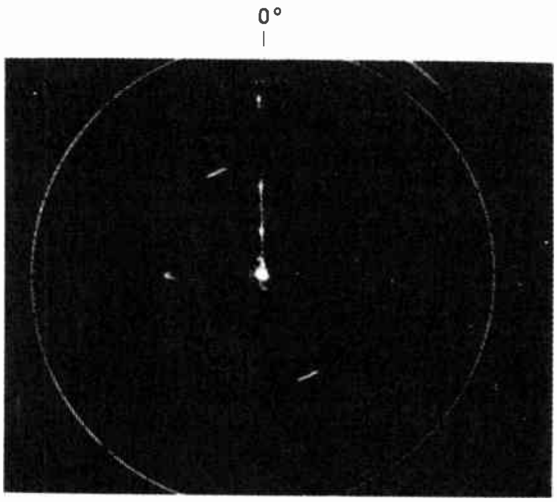


Figure 29. Chrominance nonlinear phase distortion display showing zero distortion.

used for the measurement. The gain of the waveform monitor is then adjusted until the highest subcarrier peak-to-peak amplitude of the lowest subcarrier is then measured. The difference between the highest subcarrier amplitude and the lowest subcarrier amplitude is the differential gain distortion at 50% APL. The above measurement procedure should be repeated using the same test signal transmitted on every fifth television line with intermediate lines set at blanking level for a 10% APL value. The maximum differential gain distortion measured should be recorded.[10]

An example of differential gain distortion is shown in Fig. 31.

Performance Objective: At 10%, 50%, and 90% APL, the differential gain shall not exceed 15% (15 IRE units).

Differential Phase

Definition: If a constant amplitude of chrominance subcarrier without phase modulation, superimposed on a luminance signal, is applied to the sending end of a television facility, the differential phase is defined as the change in the phase of the subcarrier at the receiving end as the luminance varies from blanking level to white level, the average picture level being maintained at a particular value.

Measurement: The modulated five-riser staircase portion of the composite test signal shown in Fig. 30 is used when measuring differential phase. The test signal's amplitude and its subcarrier phase at each step level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor and phase comparator (e.g., vectorscope) at the receiving end should be properly calibrated. A vectorscope display of differential phase with zero distortion is shown in Fig. 32.

Following the above, the test signal should be fed through a high-pass filter network to the phase comparator (or directly to the vectorscope). The differential phase distortion is the measured peak-to-peak change in subcarrier phase at 50% APL. The above measurement procedure should be repeated using the same test signal transmitted on every fifth television line with intermediate lines set at blanking level for a 10% APL value and then at peak white level for a 90% APL value. The maximum differential phase distortion should be recorded.[11]

An example of differential phase distortion is shown in Fig. 33.

Performance Objective: At 10%, 50% and 90% APL the differential phase shall not exceed 5°.

Chrominance-to-Luminance Intermodulation

Definition: If a luminance signal of constant amplitude is applied to the sending end of a television facility, the intermodulation is defined as the variation of the amplitude of the luminance signal at the receiving end resulting from the superimposition on the luminance signal of a chrominance signal of specified amplitude.

Measurement: The three-level chrominance portion of the combination test signal shown in Fig. 34 is used when measuring chrominance-to-luminance distortion. The test signal's amplitude at each chrominance level must be accurately adjusted at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated. A waveform monitor display of chrominance-to-luminance intermodulation with zero distortion is shown in Fig. 35.

Following the above, the test signal is passed through a low-pass filter network[12] the output of which is connected to the waveform monitor being used for the

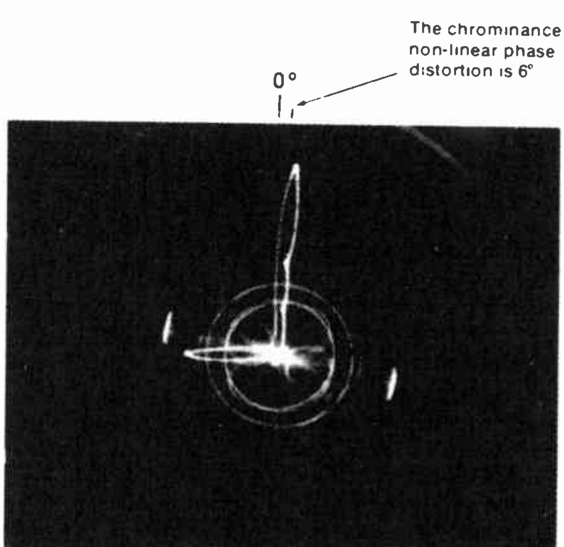


Figure 30. An example of chrominance nonlinear phase distortion.

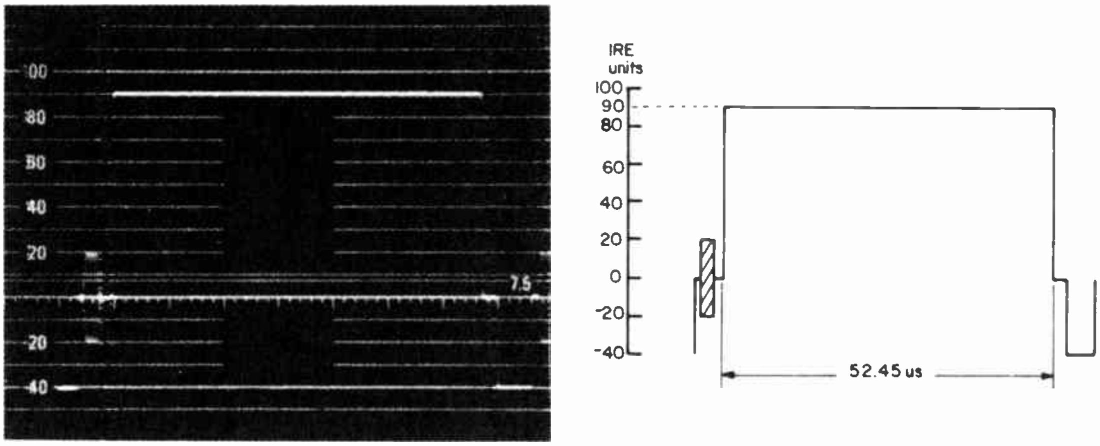


Figure 31. The flat-field bounce test signal (90% APL value shown).

Generator Output Specifications

Time of rise and fall of bar edge	:	derived from the shaping network of the sine-squared pulse.
Bar Duration	:	nominally 52.45 microseconds
Bar Amplitude	:	(High) 90 ± 1 IRE unit (Low) 10 ± .5 IRE units
Tilt	:	Less than 1 IRE unit (10-90 IRE units)
Time of Transition (10-90 or 90-10)	:	Less than 10 microseconds
Rate of Transition	:	≤ 5 sec.

Peak overshoot is 50mV (7 IRE), settling time .75 sec.
sweep = 1 sec./CM gain = .2V/CM (DC—Coupled)

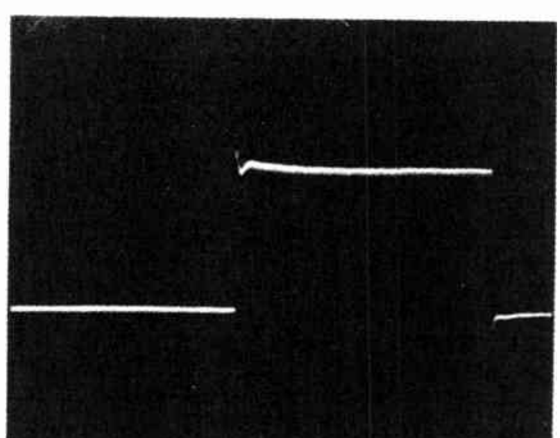
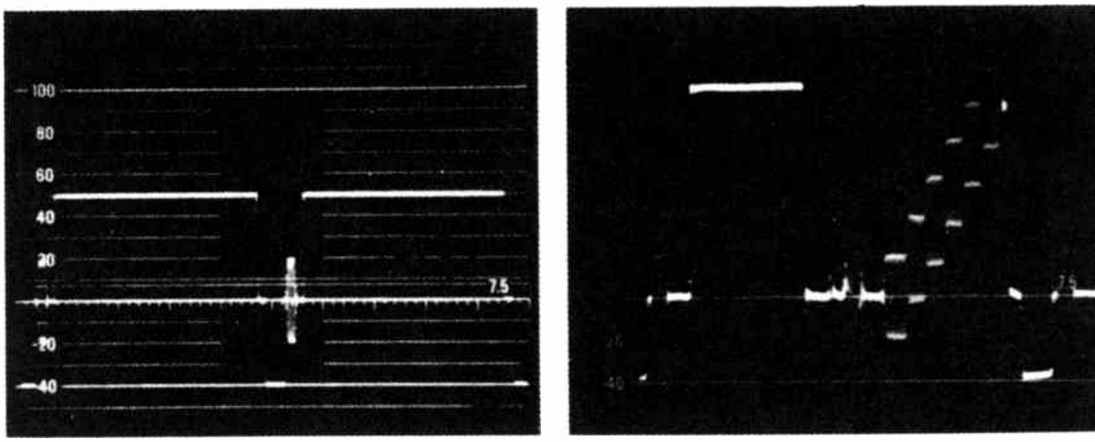


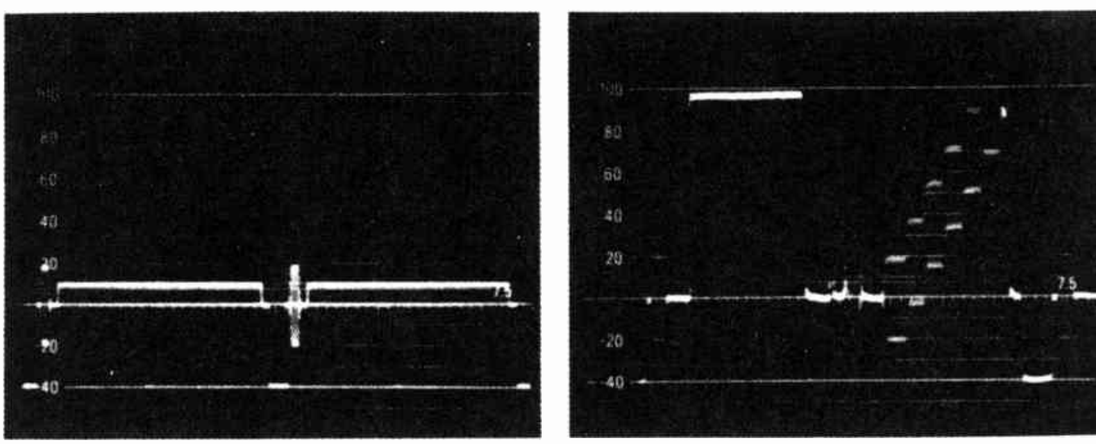
Figure 32. An example of long-time waveform distortion.



Full field test signal at 50% APL

Figure 33A.

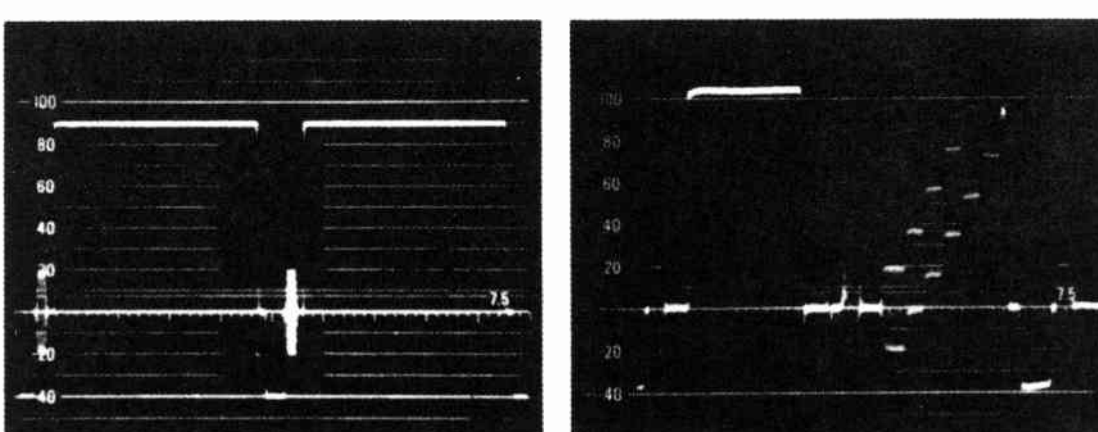
Line 17 field one, reference amplitude
Picture 100 IRE units
Sync 38 IRE units



Full field test signal at 10% APL

Figure 33B.

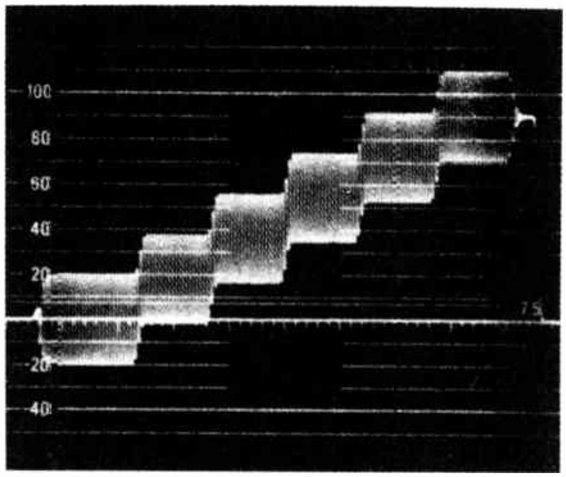
Dynamic gain distortion at 10% APL is
Picture 3 IRE units
Sync +2 IRE units



Full field test signal at 90% APL

Figure 33C.

Dynamic gain distortion at 90% APL is
Picture, +3 IRE units
Sync 0 IRE units



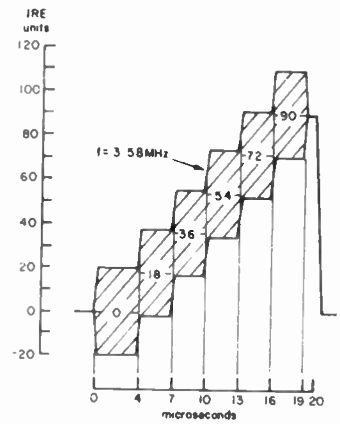
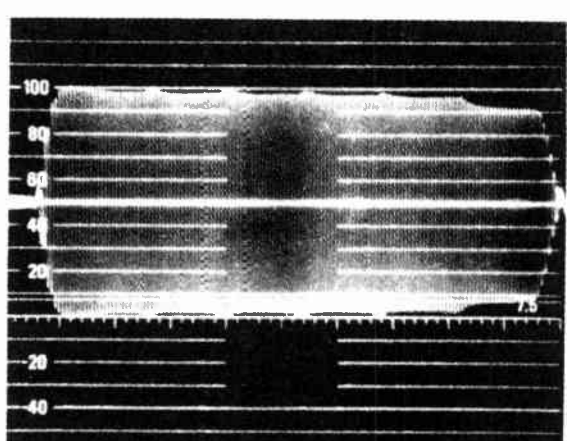


Figure 34. The modulated five-riser staircase test signal.

Generator Output Specifications

(In addition to the luminance specifications shown with Figure 20 in Section 3.9)

Chrominance amplitude	:	40 ± 0.5 IRE units
Inherent differential gain	:	less than 0.5 percent
and differential phase	:	less than 0.2°
Rise and fall times of the modulation envelope	:	400 ± 25 nanoseconds
Phase of chrominance signal relative to reference burst phase	:	0° ± 1.0° over the range 10% to 90% APL



Differential gain is 16 percent.

Figure 35. An example of differential gain distortion.

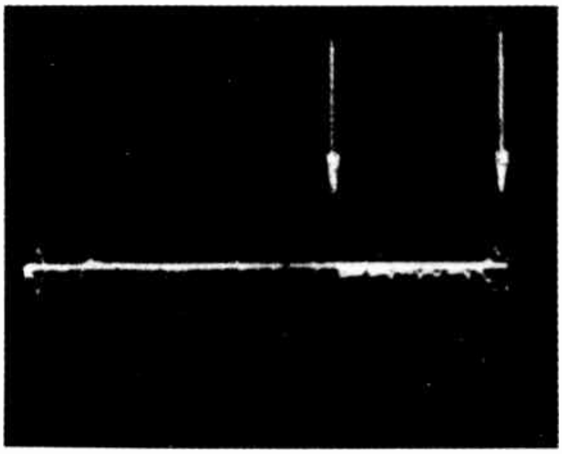


Figure 36. Vectorscope display of differential phase showing zero distortion (between arrows).

measurement. The chrominance-to-luminance intermodulation is the maximum amplitude departure in IRE units of the filtered luminance pedestal which did not contain the superimposed subcarrier, with the pedestal adjusted to exactly 50 IRE units.

An example of chrominance-to-luminance intermodulation is shown in Fig. 36.

Performance Objective: Chrominance-to-luminance intermodulation shall displace the luminance component no more than 3 IRE units from the 50 IRE unit reference level.

Random Noise-Weighted

Definition: The signal-to-weighted noise ratio of a television facility is defined as the ratio, expressed in decibels, of the nominal amplitude of the luminance signal (100 IRE units) to the RMS amplitude of the noise measured at the receiving end after band-limiting and weighting with a specified network. The measurement should be made with an instrument having, in terms of power, a time constant or integrating time of 0.4 seconds.

Measurements: In general, the random noise-weighted measurement should be made with an RMS reading instrument with the video signal removed from the television facility and the input to the facility terminated in its characteristic impedance. However, the measurement can also be made either with a flat-field test signal as shown in Fig. 37 or using a specified line in the vertical blanking interval which is kept free of picture information, provided the measuring instrument accurately integrates the measured noise over the full-field period. In all of the above measurements the band-limiting and weighting networks shown in Fig. 38 shall be inserted at the receiving end of the television facility prior to the noise measuring instrument.

The signal-to-weighted noise ratio in decibels can be computed using the following formula:

Signal-to-weighted noise (dB) =

$$20 \log_{10} \frac{\text{p-p signal amplitude}}{\text{RMS weighted noise}}$$

An example of random noise is shown in Fig. 39.

Performance Objective: The signal-to-random noise ration shall be greater than or equal to 53 dB.

NOTE: Appendix B describes a technique for approximating weighted video random noise.

Impulsive Noise

Definition: The signal-to-impulsive noise ratio of a television facility is defined as the ratio, expressed in decibels, of the nominal amplitude of the luminance signal (100 IRE units) to the peak-to-peak amplitude of the noise.

Measurement: The flat-field test signal shown in Fig. 37 is used when measuring impulsive noise. The test signal's amplitude must be accurately adjusted to 100 IRE units at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal's amplitude should be adjusted to be exactly 100 IRE units at the receiving end and the waveform display then examined closely to determine if impulsive noise interference (occasional random spikes or transients) is present. The peak-to-peak amplitude of the interference is then measured, in IRE units, and recorded. To compute the signal-to-impulsive noise ratio in decibels, the following formula can be used:

Signal-to-impulsive noise (dB) =

$$20 \log_{10} \frac{100 \text{ IRE units}}{\text{peak-to-peak amplitude of impulsive noise}}$$

Alternatively, the actual peak-to-peak amplitude of the impulsive noise, in IRE units, can be used directly

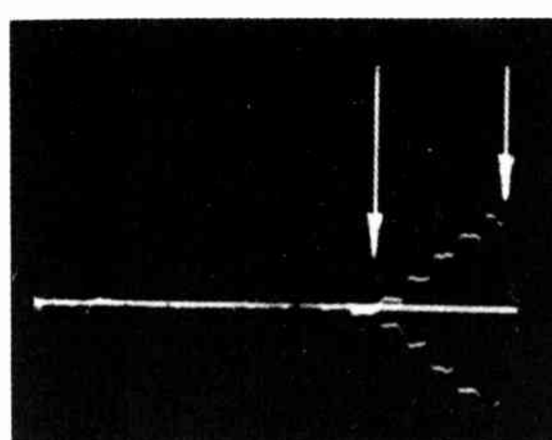


Figure 37. Vectorscope display showing differential phase distortion (between arrows).

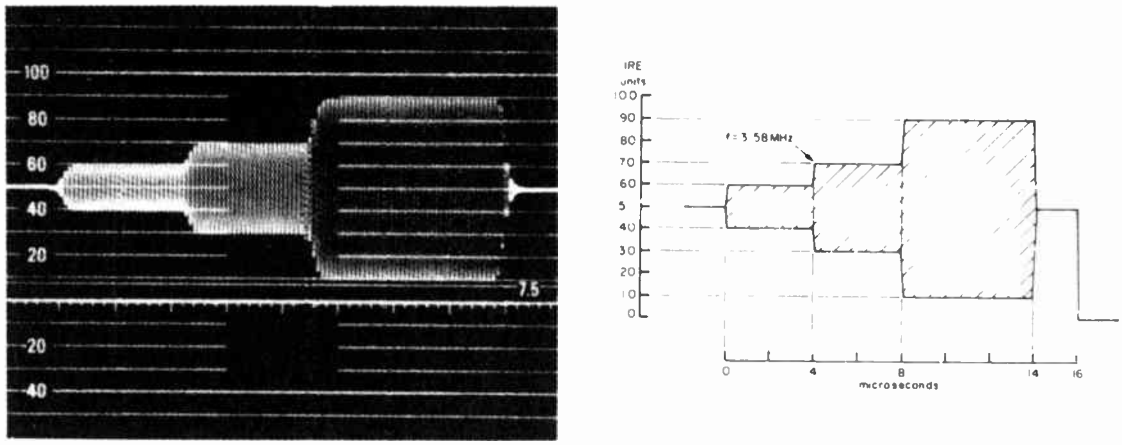


Figure 38. The three-level chrominance signal.

Generator Output Specifications

Peak-to-peak amplitudes of the 3-levels :	20,40 and 80 IRE units ± 0.5 IRE units
Sub-carrier frequency phase :	90 ° ± 1° relative to reference burst; 0° ± 0.2° any one relative to the other two.
Rise and fall of chrominance envelopes :	400 ± 25 nanoseconds
Duration of entire signal :	14 ± 0.5 microseconds

The chrominance-to-luminance intermodulation is + 3 IRE units.

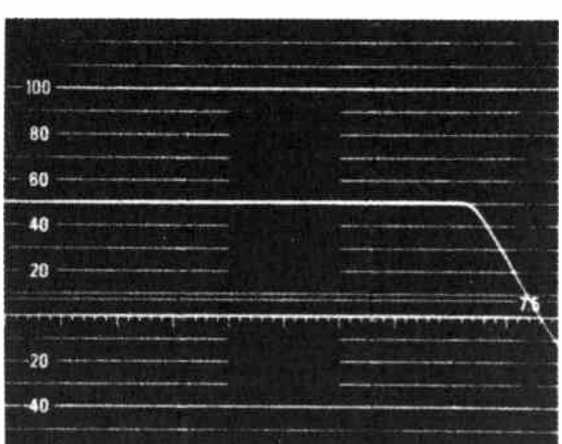


Figure 39. Waveform monitor display showing zero chrominance-luminance intermodulation.

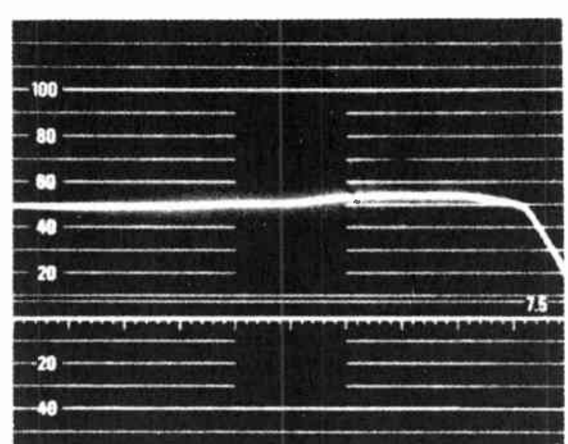


Figure 40. Waveform monitor display showing chrominance-luminance intermodulation distortion.

to determine if the television facility meets the performance objective stated below.

An example of impulsive noise interference is shown in Fig. 40.

Performance Objective: The impulsive noise shall be no greater than 7 IRE units peak-to-peak which equals a signal-to-impulsive noise ratio of no greater than 23 dB. The frequency of occurrence of noise impulses shall not exceed one per minute.

Periodic Noise

Definition: The signal-to-periodic noise ratio is defined as the ratio in decibels of the nominal amplitude of the luminance signal (100 IRE units) to the peak-to-peak amplitude of the noise. Different performance objectives are sometimes specified for periodic noise (single frequency) between 1 kHz and the upper limit of the video frequency band and for power supply hum, including low-order harmonics.

Measurement: The flat-field test signal shown in Fig. 37 is used when measuring periodic noise. The test signal's amplitude must be accurately adjusted to 100 IRE units at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal's amplitude should be adjusted to be exactly 100 IRE units at the receiving end and the waveform display then examined to determine if periodic noise interference is present. The peak-to-peak amplitude, in IRE units, of periodic noise which is low-frequency in nature (power-supply hum, etc.) should be measured and recorded separately from the periodic noise in the nominal frequency range 1 kHz to 4.2 MHz.[13] Examples of periodic noise interference are shown in Figs. 41 and 42. To compute the signal-to-signal noise ratio in decibels the following formula can be used:

Signal-to-periodic noise (dB) =

$$20 \log_{10} \frac{100\ \text{IRE units}}{\text{peak-to-peak amplitude of periodic noise}}$$

Performance Objective: The signal-to-periodic noise ratio,

A. below 1 kHz (including power-supply hum and lower order harmonics) shall be greater than or equal to 50 dB.

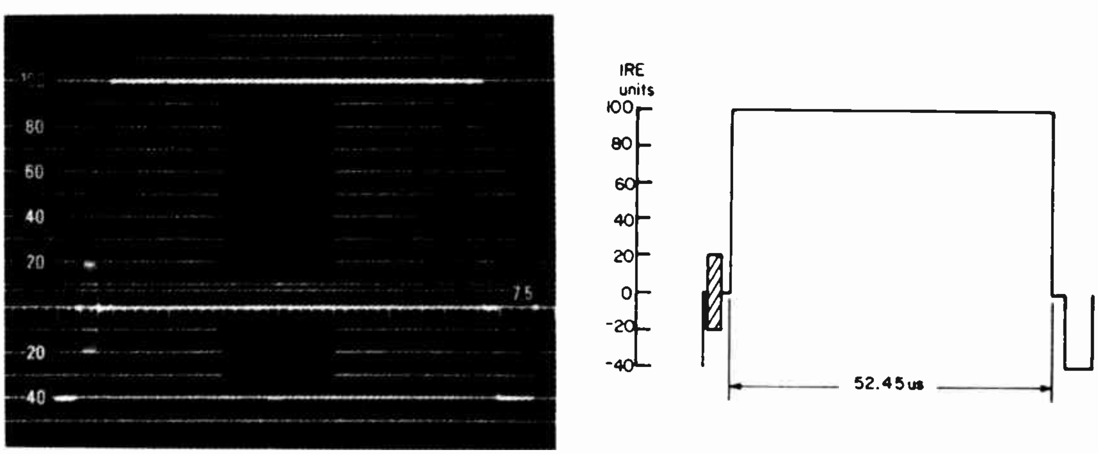


Figure 41. The 100% flat-field test signal.

Generator Output Specifications

Time of rise and fall	:	derived from the shaping network of the sine-squared pulse.
Amplitude	:	100 ± 1 IRE unit.
Tilt	:	less than 1 IRE unit.
Time of bar duration	:	nominally 52.45 microseconds.

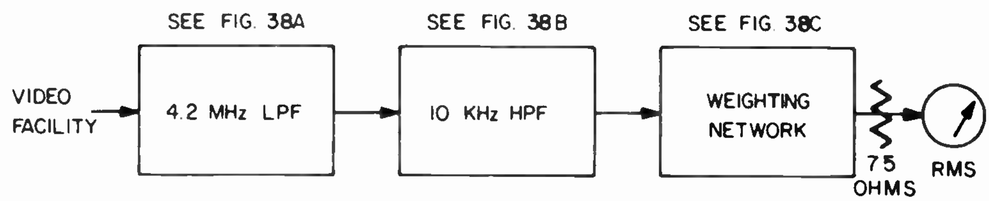
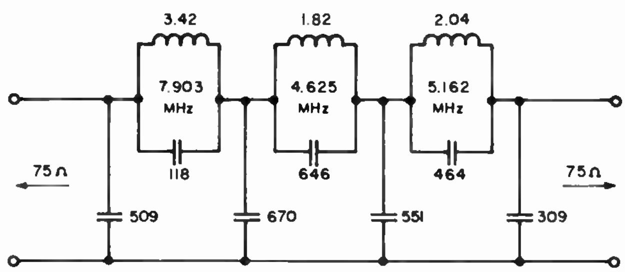
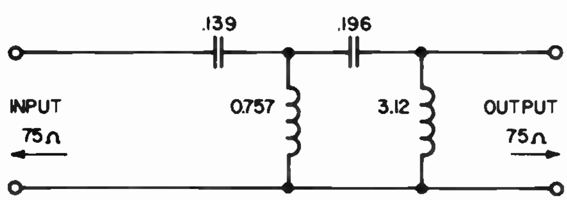


Figure 42. Band limiting and weighting network for measurement of random noise—weighted.



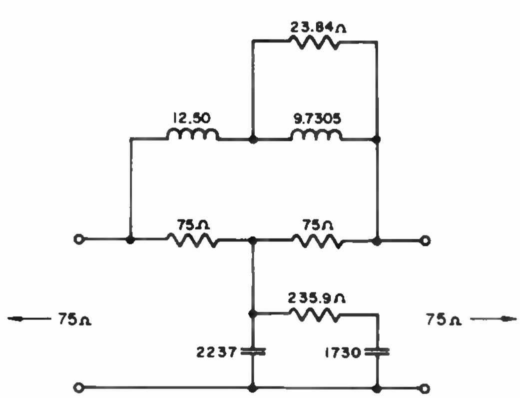
Inductances are given in uH, and capacitances in pF. Q measured at 5 MHz is between 80 and 125 for all inductors.

Figure 42A. Low-pass filter to use in noise measrements (f_c = 4.2 MHz).



Inductances are given in MH, capacitances in pF. Q measured at 10 kHz should be 100 or more.

Figure 42B. High-pass filter (f_c = 10 kHz).



Capacitances are given in pF, inductances in uH.
Capacitor and resistor tolerance ± 1%.

$$\text{Insertion loss (dB)} = 10 \log_{10} \frac{[1+(f/f_1)^2][1+(f/f_2)^2]}{[1+(f/f_3)^2]},$$

where f_1 = 0.270 MHz, f_2 = 1.37 MHz, and f_3 = 0.390 MHz.

Figure 42C. Random noise weighting network.

B. between 1 kHz and 4.2 MHz[14] shall be greater than or equal to 50 dB.

NOTE: In specific instances where an exact measurement must be made, a spectrum analyzer or a frequency-selective voltmeter should be used. As these devices are normally RMS reading instruments, 9 dB should be subtracted from the performance objective (to convert from peak-to-peak noise to RMS noise).

Crosstalk Noise

Definition: The signal-to-crosstalk ratio is defined as the ratio, in decibels, of the nominal amplitude of the luminance signal (100 IRE units) to the peak-to-peak amplitude of the interfering waveform.

Measurement: The flat-field test signal shown in Fig. 37 is used when measuring crosstalk. The test signal's amplitude must be accurately adjusted to 100 IRE units at the sending end prior to the commencement of the test. Similarly the waveform monitor at the receiving end should be properly calibrated.

Following the above, the test signal's amplitude should be adjusted to be exactly 100 IRE units at the receiving end. The waveform display and a suitable broadcast-quality picture monitor can then be examined to determine if crosstalk interference is present. The peak-to-peak amplitude of the interfering waveform is then measured, in IRE units, and recorded.[15]

To compute the signal-to-crosstalk ratio in decibels the following formula can be used:

Signal-to-crosstalk (dB) =

$$20 \log_{10} \frac{\text{100 IRE units}}{\text{peak-to-peak amplitude of periodic noise}}$$

Performance Objective: The signal-to-crosstalk ratio shall be greater than or equal to 60 dB.[16]

BIBLIOGRAPHY

EIA/TIA-250-C "Electrical Performance for Television Transmission System" may be obtained from: The Electronics Industries Association, Engineering Standards Department, 2001 Pennsylvania Avenue, Washington, D.C., 20006.

NTC-7 "Video Facility Testing" may be obtained from: Public Broadcasting Service, Engineering Department, 1320 Braddock Place, Alexandria, VA, 22314.

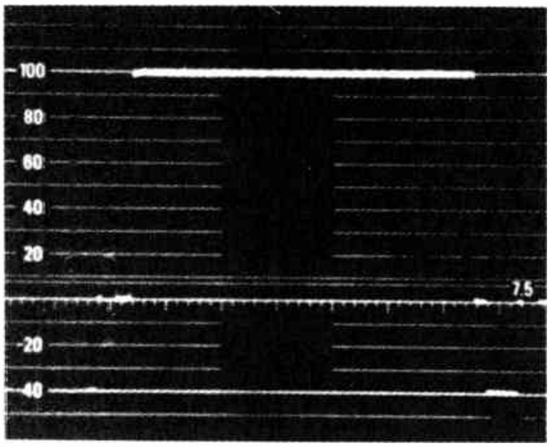


Figure 43. An example of random noise.

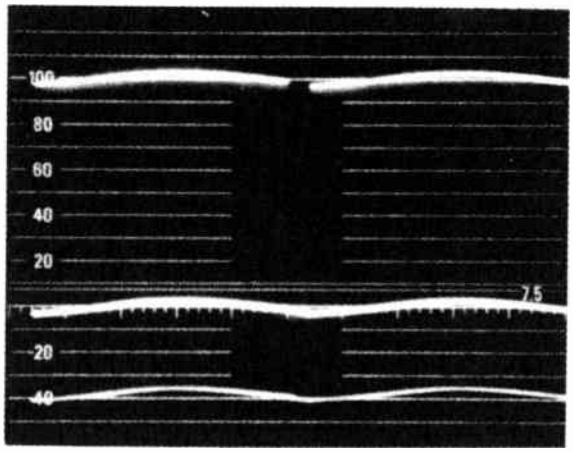


Figure 45. Example of below 1 kHz (i.e., 60 Hz) periodic noise—field rate display.

Impulsive noise is 12 IRE units peak-to-peak.

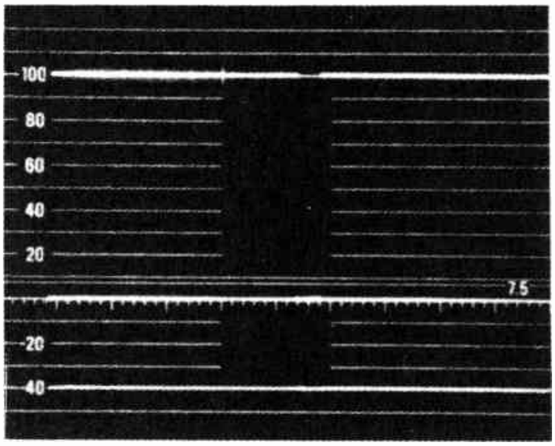


Figure 44. An example of impulsive noise interference.

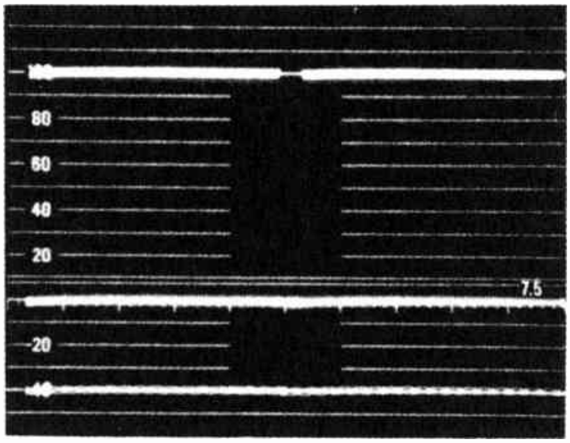


Figure 46. an example of above 1 kHz (i.e., 5 kHz) periodic noise—field rate display.

NTC-9 "Satellite Operation Technical Reference" may be obtained from: CapCities ABC Engineering, 47 W 66th Street, New York, NY, 10023.

ENDNOTES

1. Excerpts reprinted with permission of the Electronics Industries Association, Washington, DC.
2. Section and test numbers are original NTC-7 numbers.
3. This parameter is also called Relative Chroma Level (RCL), and is expressed as a percentage of P-P chrominance referenced to the line bar amplitude, as shown in Fig. 15. Hence, the performance objective expressed as RCL would be ±6%.
4. This parameter is also called the Relative Chroma Time.
5. For in-service VITS measurements, the procedures outlined in the section "Parameters, Measuring Techniques, and Performance Objectives" should be followed.
6. The chroma filter network incorporated into most television waveform monitors is suitable for this test.
7. For in-service measurement of long-time waveform distortion, the procedure outlined in the section "Parameters, Measuring Techniques, and Performance Objectives" should be followed.
8. For in-service measurements of dynamic gain, the procedure outlined in the section "Parameters, Measuring Techniques, and Performance Objectives" should be followed.
9. The chroma filter network incorporated into most television waveform monitors is suitable for this test.

10. For in-service VITS measurement of differential gain, the procedure outlined in the section "Parameters, Measuring Techniques, and Performance Objectives" should be followed.
11. For in-service VITS measurement of differential phase, the procedure outlined in the section "Parameters, Measuring Techniques, and Performance Objectives" should be followed.
12. The low-pass filter network incorporated into most television waveform monitors is suitable for this test.
13. It may be necessary to use the high-gain setting on the waveform monitor when making this measurement. If so, care should be taken to ensure that it is properly calibrated.
14. The low-pass filter network shown in Fig. 38A should be used when measuring periodic noise interference in the range 1 kHz to 4.2 MHz.
15. It may be necessary to use the high-gain setting on the waveform monitor when making this measurement. if so, care should be taken to ensure that it is properly calibrated.
16. The low-pass filter network shown in Fig. 38A should be used when measuring crosstalk noise interference.

APPENDIX A

Example of a Weekly VITS Log

WEEKLY VITS LOG

WEEK BEGINNING SUNDAY ________

STATION ________________ LOCATION ____________

PARAMETER	SIGNAL	LIMIT	SUN	MON	TUES	WED	THUR	FRI	SAT
INSERTION GAIN (See Sec. 3.2)	BAR	±3 IRE							
LINE TIME DISTORTION (See Sec. 3.4)	BAR	4 IRE PK-PK							
SHORT TIME DISTORTION (See Sec. 3.5)	2 T	±6 IRE P/B							
	BAR EDGE	10 IRE PK-PK							
CHROMA-LUMINANCE GAIN INEQUALITY (See Sec. 3.6)	CHROMA PULSE	100 ±3 IRE							
CHROMA-LUMINANCE DELAY INEQUALITY (See Sec. 3.7)	CHROMA PULSE	(CHROMA) 75nSec ADV or DLY							
GAIN-FREQUENCY (See Sec. 3.8)	MULTI-BURST	45-53 IRE 0.5							
		1.0							
		2.0							
		3.0							
		3.58							
		4.2							
DIFFERENTIAL GAIN (See Sec. 3.13)	STAIR-CASE	15 %							
DIFFERENTIAL PHASE (See Sec. 3.14)	STAIR-CASE	5°							
CHROMA NON-LINEAR GAIN DISTORTION (See Sec. 3.10)	3 LEVEL CHROMA	40 REF: 20±2 IRE 80±8 IRE							
CHROMA NON-LINEAR PHASE DISTORTION (See Sec. 3.11)	3 LEVEL CHROMA	5°							
CHROMA-LUMINANCE INTERMODULATION (See Sec. 3.15)	3 LEVEL CHROMA	3 IRE							
RANDOM NOISE (See Appendix II)	"GRASS" ESTI-MATE	53 dB							

NOTES: OBSERVATIONS TO BE MADE AHEAD OF ANY STATION EQUIPMENT.

THESE NETWORK TRANSMISSION OBSERVATIONS SHOULD BE SUPPLEMENTED, AS APPROPRIATE, BY OBSERVATIONS OR MEASUREMENTS OF THE OTHER PARAMETERS COVERED IN NTC REPORT NO. 7.

SECTION NO. UNDER PARAMETER REFERS TO APPROPRIATE SECTION OF NTC REPORT NO. 7.

APPENDIX B

TECHNIQUE FOR APPROXIMATING WEIGHTED VIDEO RANDOM NOISE

When an RMS reading noise meter is not available, an alternate method of determining video S/N should be used. The one suggested here is accurate to within 2 dB if care is taken in setting scope brightness and operator estimation of the peak-to-peak noise is done in a consistent manner.

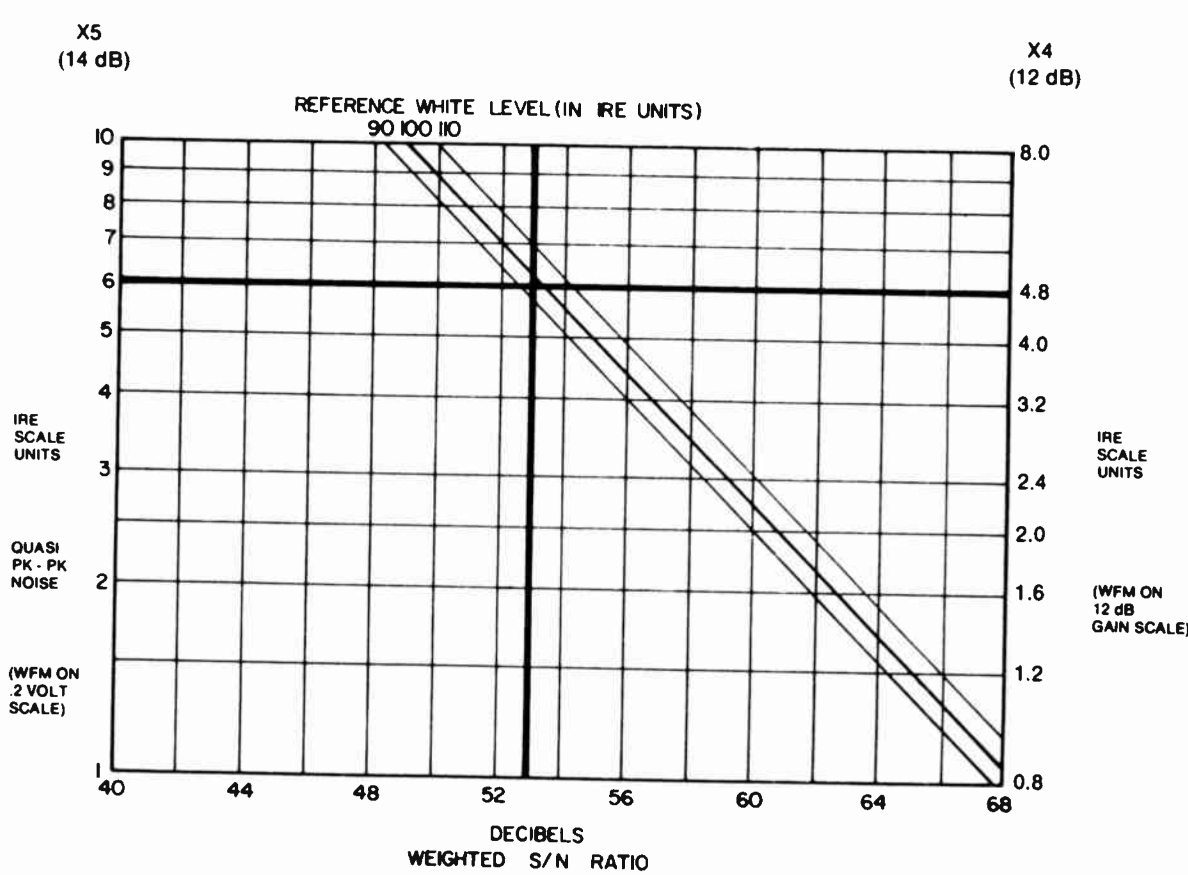
This procedure uses the low-pass filter and video weighting networks shown in Fig. 38A and 38C, respectively, of Sect. 3.16. Place these two networks in tandem between the video facility and the waveform monitor. After noting the amplitude of the line bar (for VITS) or the 100% flat field (for a full-field test) place the waveform monitor vertical gain to the 0.2 volt range* and estimate the quasi peak-to-peak** amplitude of the grass at blanking level. Transfer this reading to the S/N nomograph and determine the weighted video S/N ratio.

The objective is 53 dB or greater as shown by the heavy line.

*14 dB gain.

**Quasi peak-to-peak is the average level ignoring large occasional spikes of noise.

NOTE: The network of Fig. 38B, 10 kHz high-pass filter, is not needed as the line selector acts as a digital filter and cuts off frequencies below 15 kHz.



4.2 Microwave Engineering for the Broadcaster

Ernest M. Hickin
Myton Associates Inc., Winchester, Massachusetts

INTRODUCTION

Microwave transmission has become increasingly important in the broadcast engineer's operation. The Studio-to-Transmitter Link (STL) has been the backbone of the industry for decades. Equipment design today allows for essentially transparent, full color video transmission, as well as sub-carriers capable of full FM stereo audio bandwidth and multiple voice and remote control functions.

The more recent development of Electronic News Gathering (ENG) and Field Production—as well as Satellite Earth Stations (TVRO) which require local back-haul, Satellite Studios and Regional News Networks—has demanded a continuous awareness of microwave practices and principles on the part of the broadcast engineer.

In most cases, microwave equipment suppliers are prepared to support broadcast engineers in their design needs. Path profiling, surveying, and predicting path performance, are services available from consultants and suppliers.

The following discussion is presented to aid broadcast engineers in their own understanding of this technology.

PATH PERFORMANCE

Free Space Loss

Free space loss for any electromagnetic wave arises from the spreading of the wavefront radiating from the source, like ripples on a pond after a stone has been thrown into the water. There is a four-fold loss in power (6 dB) every time the range doubles once outside the near field of the antenna. This relationship is known as the "inverse square law."

Free space loss (L), the loss which is independent of ground or atmospheric effects, is given by:

$$L(\text{dB}) = 36.6 + 20 \log F(\text{MHz}) + 20 \log D(\text{miles})$$
$$\text{or} = 96.6 + 20 \log F(\text{GHz}) + 20 \log D(\text{miles})$$

This is the loss between isotropic antennas (theoretical antennas which radiate or receive equally in all directions). The gain (G) of a microwave antenna is then expressed in dBi (gain relative to an isotropic antenna). For a parabolic reflector antenna of diameter "d" feet, with an efficiency of 55% (which is typical of all but the smallest antennas) gain is given by:

$$G(\text{dBi}) = 20 \log d(\text{ft}) + 20 \log F(\text{MHz}) - 52.6$$

Strictly speaking, the "gain" of a directional antenna is the amount by which the radiation in the desired direction has been increased by re-directing energy which would have been radiated in unwanted directions by an isotropic antenna. It should be noted that below 1 GHz it is usual to express gain relative to a dipole where 0 dBd = 2.2 dBi; thus the gain of a VHF or UHF TV transmitting antenna is most likely to be quoted in dBd.

System Path Loss

The ratio of received power to transmitted power between two correctly aligned antennas of gain G when D miles apart will be given by:

$$G1 - 36.6 - 20 \log F - 20 \log D + G2$$

This number will be negative, and the numerical value is referred to as "path loss."

As examples, if the antennas are 6-foot in diameter and the path is 30 miles long then:

At 2 GHz: $28.6 - 132.2 + 28.6 = -75.0$ dB; path loss = 75.0 dB

At 7 GHz: $39.8 - 143.0 + 39.8 = -63.4$ dB; path loss = 63.4 dB

At 13 GHz: $45.1 - 148.4 + 45.1 = -58.2$ dB; path loss = 58.2 dB

Similarly, if the antennas are 2-ft diameter and the path is 10 miles long then:

At 18 GHz: $38.5 - 141.7 + 38.5 = -64.7$ dB; path loss = 64.7 dB

At 23 GHz: $40.7 - 143.8 + 40.7 = -62.5$ dB; path loss = 62.5 dB

Note that in these equations, doubling the frequency increases the antenna gain by 12 dB (2×6), while the path loss increases by 6 dB. This means that lower frequency systems tend to require larger antennas, more power, or both, for a given received carrier level.

Other Sources of Path Loss

The free space loss is easily calculated, and antenna manufacturers issue slide rules from which that loss, and the antenna gains, can be read.

Atmospheric and other conditions will cause the loss to vary from time to time and these effects must be considered when planning a system.

Multipath signals arriving at the receiving antenna by reflections from water, hills, buildings or atmospheric discontinuities as well as by the direct path, can add to or cancel the signal causing an increase of up to 6 dB or a decrease of more than 50 dB.

Atmospheric bending of the wave front due to abnormal changes of temperature and/or humidity with height can cause loss of signal due to diversion from the desired direction. This effect is discussed below under "K Factor."

Rainfall, and to a lesser extent snowfall, can attenuate the signal; this effect increases with increasing frequency and is the over-riding factor at 18 and 23 GHz.

To offset these effects, a system is designed to give a more than adequate signal level at the receiver under free space loss condition (which is the norm for at least 90% of the time). The excess of signal over the minimum required for satisfactory service is called the *fade margin*. Typical fade margins are in the range of 26 to 46 dB and will be larger for higher frequencies, longer paths, and over-water or similarly difficult situations.

The choice of fade margin depends upon the specific situation, for example, when existing towers or masts limit the size of the antenna. It should also be influenced by the environment, more margin being desirable in humid, flat country and less in dry mountainous regions.

Multipath Outages

Outage time (time out of service due to propagation) or availability, which is (1 − outage) x 100% when outage is expressed as a fraction of time, can be calculated from an accepted formula. This formula assumes that the path has adequate clearance for the area, and is based on frequency, path length, fade margin, and two empirical factors related to terrain roughness (the rougher the better as this breaks up the atmosphere) and humidity (the lower the better as moisture accentuates atmospheric effects). The formula, derived by Barnett & Vigants from the average of many Bell System microwave paths, is:

$$t = a \times b \times 2.5 \times 10^{-6} \times f \times D^3 \times 10^{(-10/F)}$$

where:
- a = 4 for very smooth terrain, including over water, 1 for average terrain with some roughness, 0.25 for mountainous, very rough or very dry areas
- b = 0.5 for Gulf Coast or similar low, humid areas, 0.25 for normal interior, temperate or northern areas, 0.125 for mountainous or very dry areas
- f = frequency in GHz
- D = path length in miles
- F = fade margin in dB
- t = time out of service as a fraction
  $(1 - t) \times 100$ = availability (%)

This formula can be useful in predicting performance for links where D is between 10 and 40 miles, and for fade margins between 25 and 55 dB.

Rainfall Outages

Attenuation due to rainfall, however, is not covered by the formula which was derived from 4 GHz data where rain effects are negligible. Figs. 1 and 2 show estimated outage times in hours/year for 30 mile paths at 2 GHz and 20 mile paths at 13 GHz in both Wyoming and Alabama (extremes of rainfall). Clearly rainfall becomes of increasing importance at higher frequencies.

To calculate outages due to rainfall a five-step procedure is used.

1. The fade margin is divided by the path length in kilometers (km). (km = miles × 1.609)
2. By reference to Fig. 3, from the International Radio Consultative Committee (CCIR) Report 721-2, the average rainfall rate (in mm/hour) along the whole path which would just cause a loss equal to the fade margin is derived; note that this number is different for vertical and horizontal polarizations.
3. By reference to Fig. 4, from Barsis and Samson, find the point corresponding to that average and to the path length in question; read the ratio number at the left and use it to divide the average rate and get the point rainfall.
4. By reference to Fig. 5, note the rainfall zone in which the link is to be established.
5. Finally, read the outage time by reference to the table, Fig. 6.

For longer paths at 13 or 15 GHz, and for all paths at 18 and 23 GHz, the rainfall outage predominates and multipath effects as derived from the B & V formula can usually be ignored.

K Factor

The pressure and hence the density of the atmosphere surrounding the earth varies with height, de-

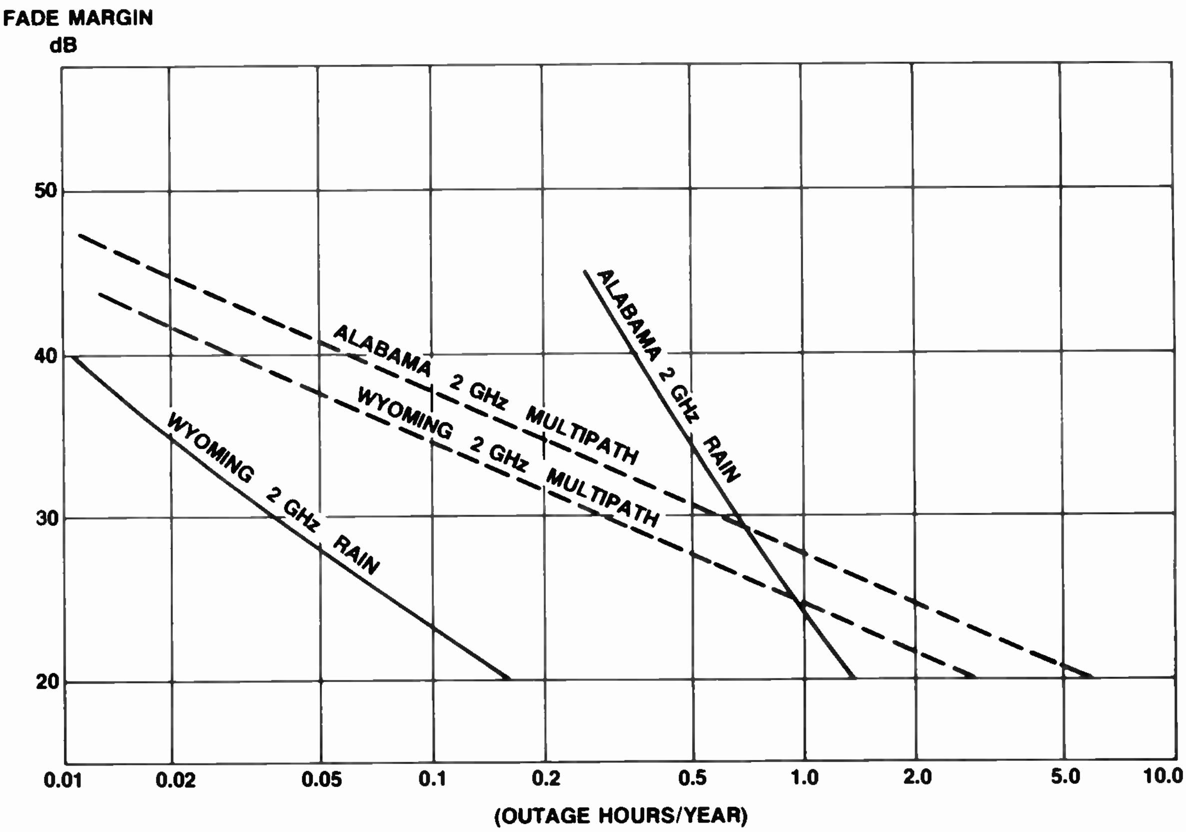


Figure 1. Outage time in hours for differing fade margins on 2 GHz 30-mile paths in Wyoming and Alabama.

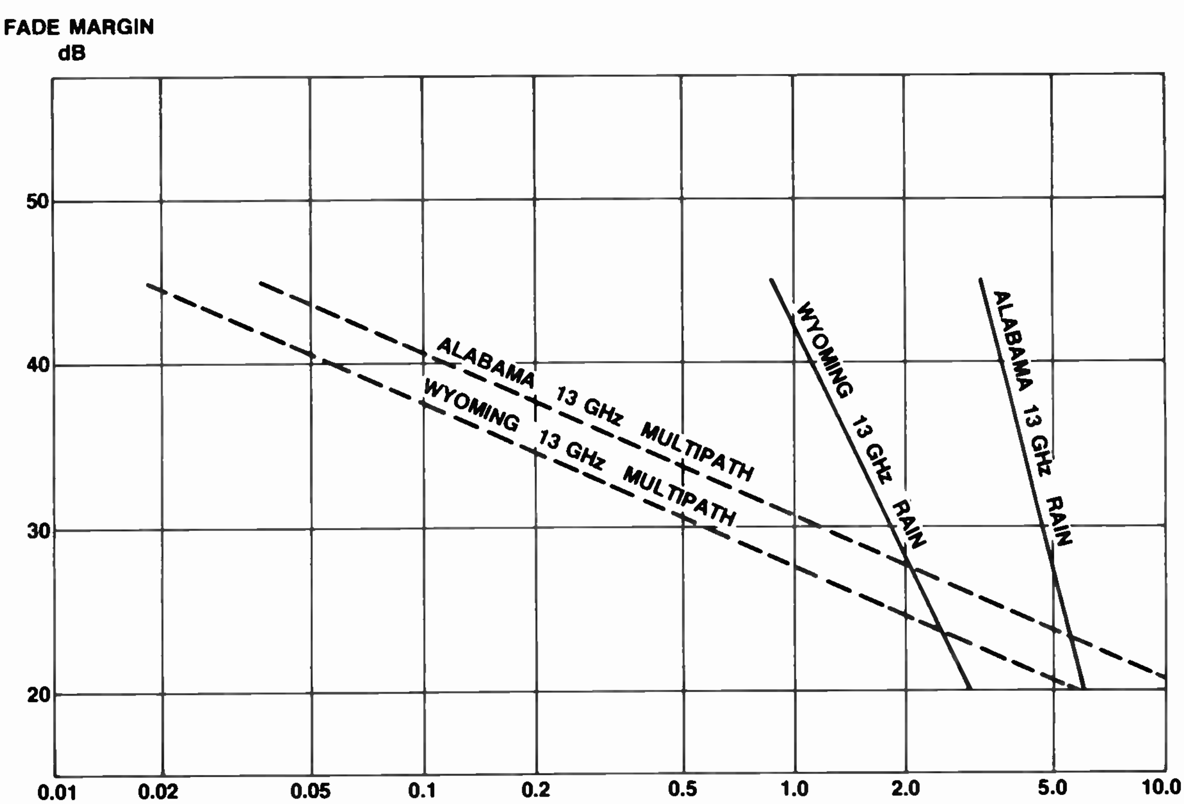


Figure 2. Outage time in hours for differing fade margins on 13 GHz 20-mile paths in Wyoming and Alabama.

creasing as the height increases and as the weight of the air above decreases. As a result, the dielectric constant also decreases with height and this has a prismatic effect causing microwaves (and light waves) to bend towards the earth. Under normal conditions the bending is less than the curvature of the earth, but nonetheless microwaves will go further than simple geometry would suggest. A convenient way in which to allow for this is to increase the radius of the earth until the microwaves appear to be travelling in straight lines.

The ratio of this effective earth's radius to the true earth's radius is called K and its value is approximately 4/3 or 1.33 for over 90% of the time in most parts of the world. There are times, however, when K can be anything from infinity to as low as 0.45. K = infinity (flat earth) is a condition where mirages are seen and radar echoes are received from hundreds of miles away. The potential for interference between systems is increased; fortunately, it is a relatively rare condition. K values between 1 and 0.45 can occur for a few percent of the year and it is necessary to allow for this if a reliable link is to be achieved. Fig. 7 is a map of the continental USA showing contours of equal minimum K factor; this is based on refractive index measurements made by the Central Radio Propagation Laboratory.

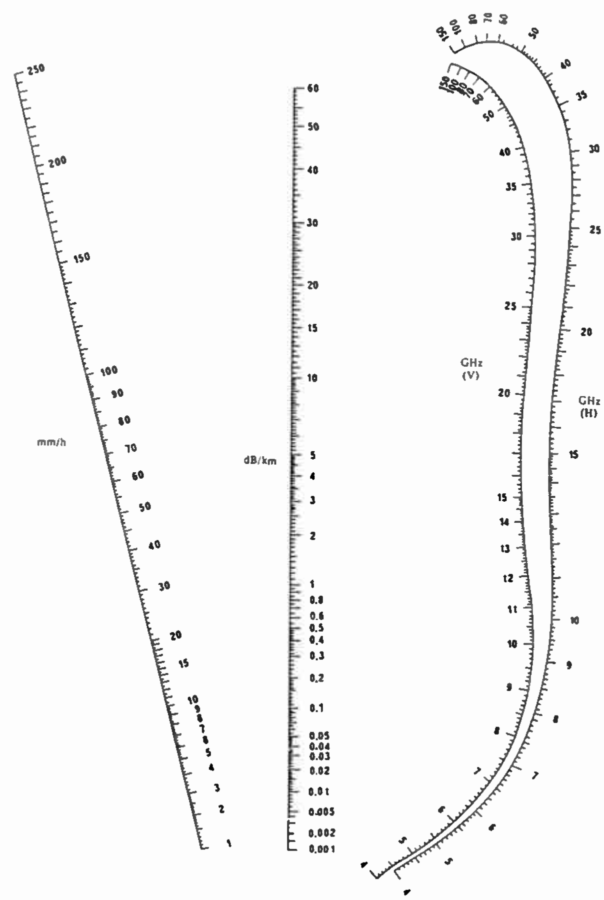


Figure 3. Specific attenuation due to rain (H): horizontal polarization (V): vertical polarization (from CCIR Report 721-2).

Clearance Requirements

Fig. 8 shows the way in which the microwave signal is attenuated when the path is close to an obstruction. Clearance is stated as the first Fresnel zone clearance which is a function of frequency.

The first Fresnel zone clearance (FZC) (feet) is given by:

$$FZC = 2280\sqrt{\frac{d(D-d)}{fD}}$$

where: f = frequency (MHz)
D = path length (miles)
d = distance to the obstruction from either end of the path

From Fig. 8 it will be seen that the attenuation with a clearance of 0.6 FZC is equal to the free space attenuation. However, as noted in Fig. 6, K factor variations will mean that more than this minimum clearance will have to be built in to allow for K values less than 4/3. A typical design technique is to plan for 0.3 FZC for the lowest K factor expected on the path, as shown by Fig. 6. While such a clearance will introduce 2 to 8 dB of loss, this is well within the fade margin of a well-designed link.

Path Profiles

Sites for microwave link terminals are likely to be studios, earth stations, or transmitters, all of which are predetermined. Choice of sites for repeaters, if they are necessary, may be constrained by availability and access. Since the access road is often the most costly part of a site, it can prove more economic to avoid the highest point and use a taller tower close to a road.

To determine if any combination of sites has a clear path between them, and to determine the antenna heights at each end of the path, it is necessary to plot a profile.

The profile, a plot of ground height against distance, can be drawn on simple squared paper or special K factor paper (usually K = 4/3). Squared paper has the advantage that any convenient scales can be chosen for the x and y axes; 4/3 paper has the advantage that potential obstructions are more obvious and the effect of path length on necessary clearance stands out. Whichever paper is used, the next step after plotting the profile is to add an allowance to high points for trees or buildings (as the map shows ground elevations). *Never accept a profile that has not been checked on the ground by travelling the path to check tree heights and other potential obstructions.* The effect of low values of K must then be added to these obstructions using one of the following formulas:

a) If using squared paper then add to each obstruction:

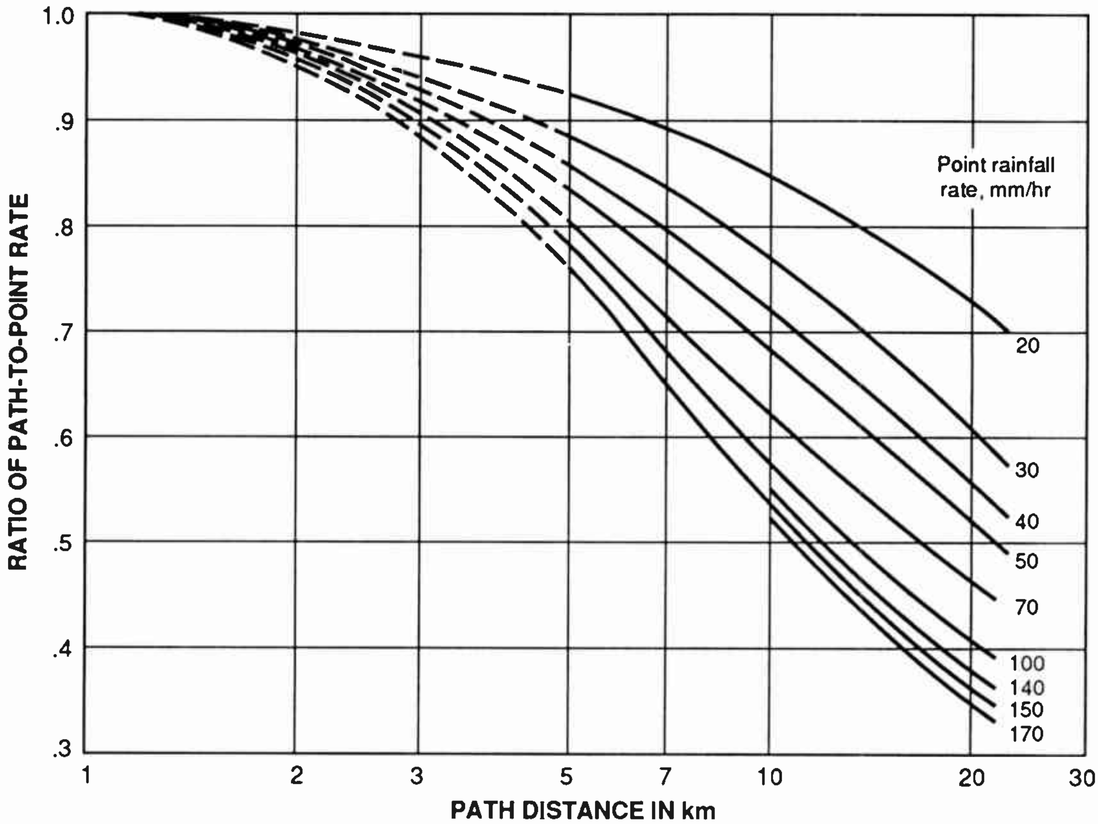


Figure 4. Path average to surface point rainfall as a function of path length.

(1) 0.6 FZC + d(D-d)/2 + T

(2) 0.3 FZC + d(D-d)3K/2 + T

b) If using 4/3 paper then add to each obstruction:

(1) 0.6 FZC + T

(2) 0.3 FZC + d(D-d)3K/4 + T

where T is the height of trees or other obstructions. Then a straight line drawn through the upper points (usually two except for short paths and lower frequencies) will pass through the terminals at the minimum antenna heights for adequate clearances. If most obstructions are near one end of the path it may be more economic to raise the antenna at that end and save a greater amount at the other end.

Figs. 9 and 10 show examples of the same profile plotted on squared and 4/3 papers and the final selection of antenna heights. The worst K factor assumed was 2/3, shown as (2), on each profile at the main obstruction point at 8 miles.

ANTENNA AND WAVEGUIDE CONSIDERATIONS

Antenna Systems

Several different types of antenna are used in a TV microwave system. Omnidirectional vertical and directional horizontal stacked arrays are used for portable and vehicular applications. Parabolic reflector antennas are used for STL, intercity, educational television (ETV) and Community Antenna Relay Service (CARS) systems. Horn or shrouded parabolic antennas may be used for high front-to-back ratios. Simple dipoles may be used for back-pack portable systems. The reader is referred to Electronic Industries Association/Telecommunications Industries Association (EIA/TIA) Standard 195A for mechanical and electrical specifications.

Isotropic Antenna

As noted above, the isotropic antenna is a hypothetical antenna used as a reference against which the gain

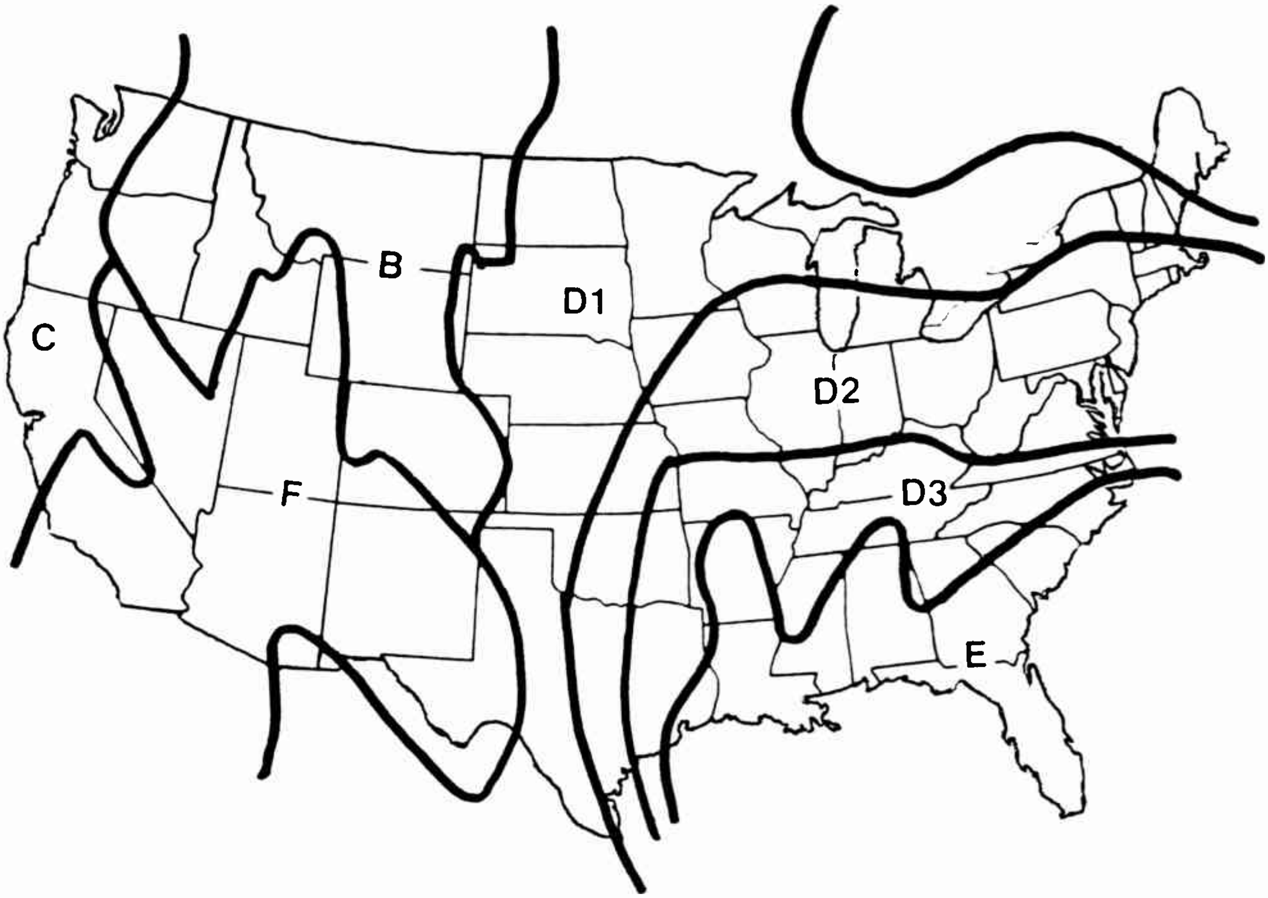


Figure 5. Geographic regions with similar rainfall statistics. (From Crane & CCIR.)

PERCENT (YR)	RAIN CLIMATE REGION										MINUTES (YR)	HOURS (YR)
	A	B	C	D_1	D_2	D_3	E	F	G	H		
0.001	28.0	54.0	80.0	90.0	102.0	127.0	164.0	66.0	129.0	251.0	5.3	0.09
0.002	24.0	40.0	62.0	72.0	86.0	107.0	144.0	51.0	109.0	220.0	10.5	0.18
0.005	19.0	26.0	41.0	50.0	64.0	81.0	117.0	34.0	85.0	178.0	26.0	0.44
0.01	15.0	19.0	28.0	37.0	49.0	63.0	98.0	23.0	67.0	147.0	53.0	0.88
0.02	12.0	14.0	18.0	27.0	35.0	48.0	77.0	14.0	51.0	115.0	105.0	1.75
0.05	8.0	9.5	11.0	16.0	22.0	31.0	52.0	8.0	33.0	77.0	263.0	4.38
0.1	6.5	6.8	72.0	11.0	15.0	22.0	35.0	5.5	22.0	51.0	526.0	8.77
0.2	4.0	4.8	4.8	7.5	9.5	14.0	21.0	3.8	14.0	31.0	1052.0	17.50
0.5	2.5	2.7	2.8	4.0	5.2	7.0	8.5	2.4	7.0	13.0	2630.0	43.80
1.0	1.7	1.8	1.9	2.2	3.0	4.0	4.0	1.7	3.7	6.4	5260.0	87.56
2.0	1.1	1.2	1.2	1.3	1.8	2.5	2.0	1.1	1.6	2.8	10520.0	175.30

Figure 6. Point rain rate distribution values (mm/hr) as a function of percentage of a year that a given rain rate is exceeded.

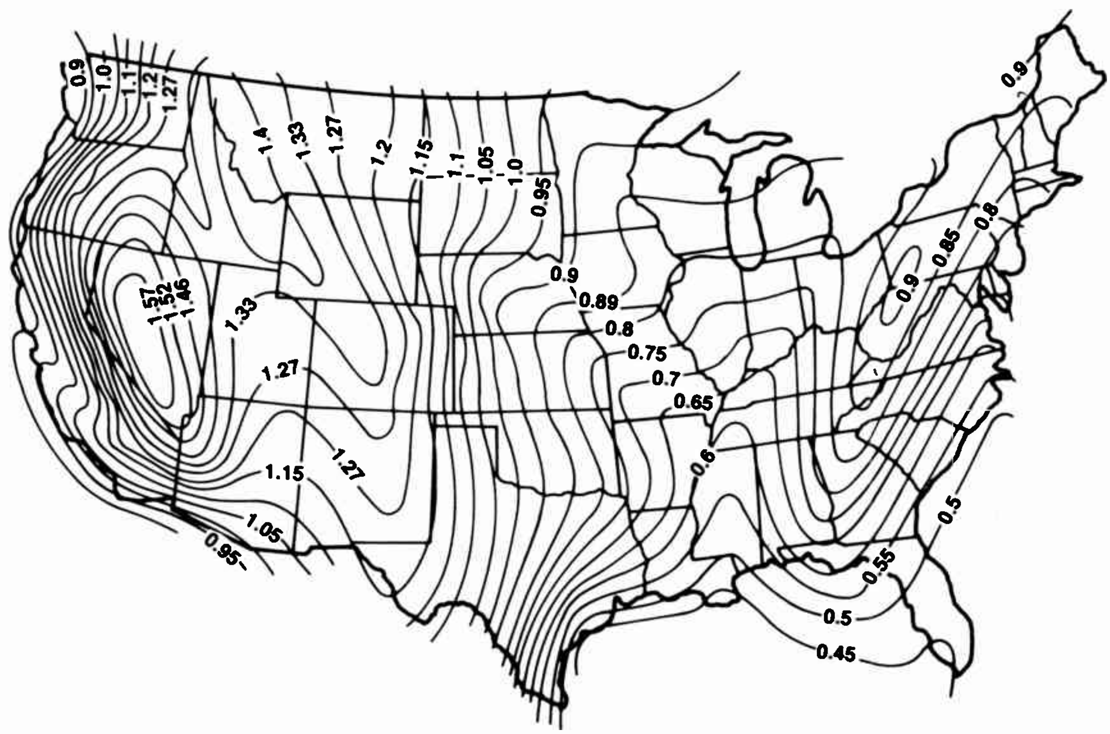


Figure 7. Estimated minimum K-factor for the continental United States.

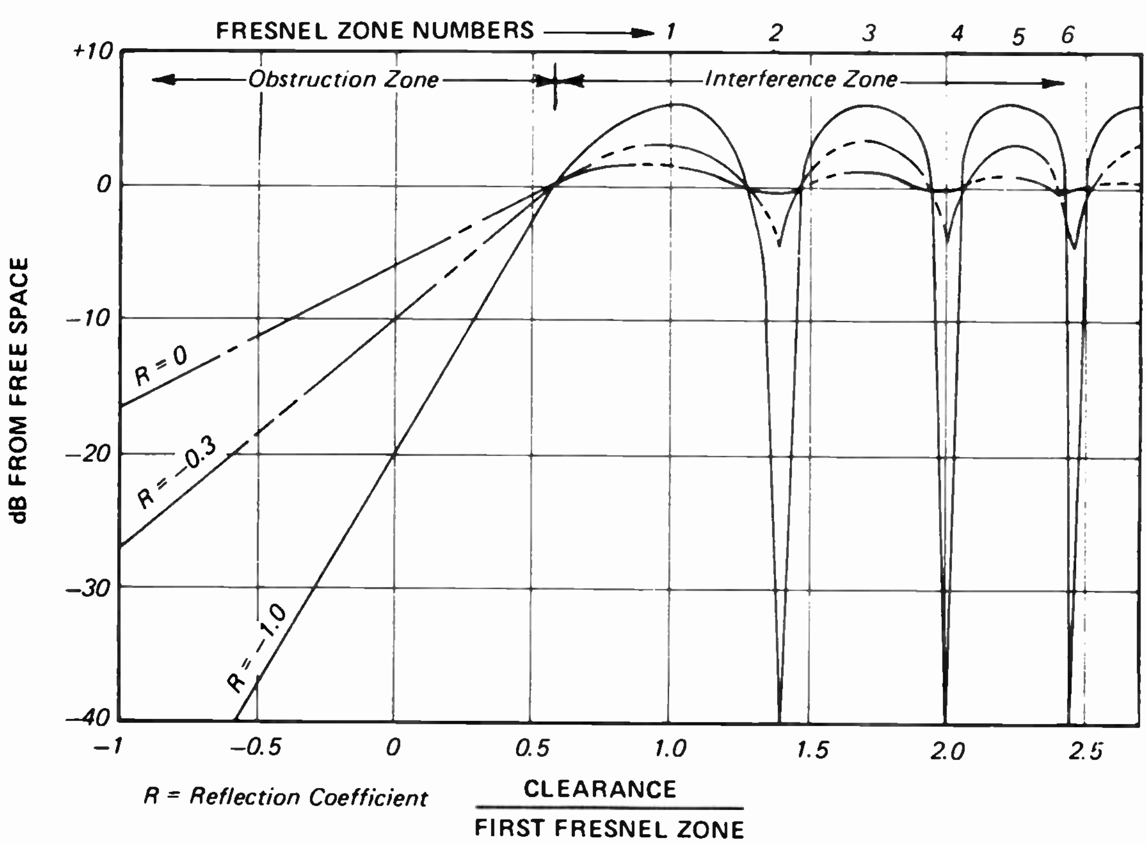


Figure 8. Attenuation versus path clearance for various types of obstruction.

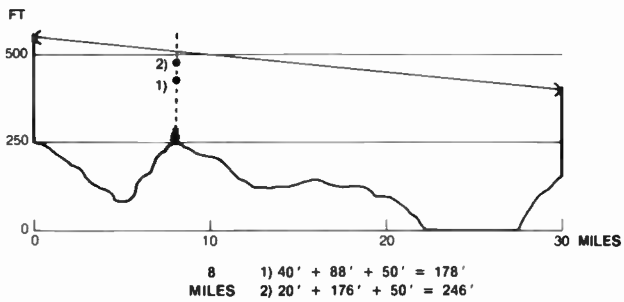


Figure 9. Example of profile plotted on squared paper, with required clearances for (1) 0.6 first Fresnel zone when K = 4/3 and (2) 0.3 first Fresnel zone when K = 2/3 (50-ft tree allowance).

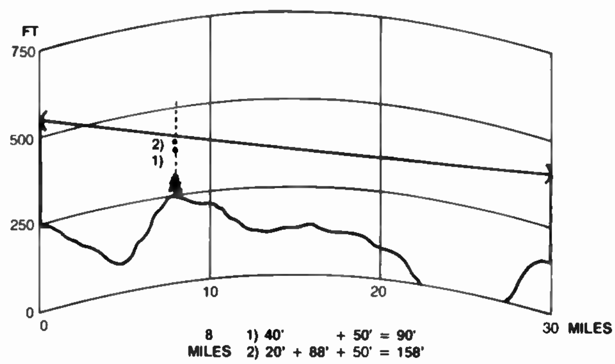


Figure 10. The profile of Figure 9 plotted on K = 4/3 paper: (1) and (2) marked as on that figure.

of practical microwave antennas can be measured. Gain is given in dBi, gain relative to an isotropic antenna. The isotropic antenna has a radiation pattern which is a perfect sphere. Other antennas are designed to enhance their radiation in a given direction at the expense of all other directions.

The gains (dB) of various types of antenna can be calculated as follows:

Type	Gain (dBi)
Parabolic:	$10 \log 5.5(Df)^2$
Horn (optimum):	$10 \log 10.3(f)^2$
Omni (stacked array):	$10 \log 2L(f)^2$

where: D = diameter of the parabola (feet)
A = area of horn mouth (square feet)
L = length of stack (feet)
f = frequency (GHz)

Parabolas

At frequencies up to about 3 GHz the reflector of a parabolic antenna may be built up from parallel rods; above that the rods have to be too close to give any advantage in reduced wind resistance. These grid antennas can only be used on the polarization which is parallel to the rods. Typical feeds are coaxial-cable fed dipoles which can be foam-filled to avoid the need for dry-air pressurization.

Above 3 GHz the reflectors are solid, usually spun aluminum, and the feeds are waveguide. Dual cross-polarized feeds, with 25 dB isolation, can be used to increase the number of signals sharing the antenna. Dual band antennas (eg. 7 GHz and 13 GHz) can be built, each band having one polarization. Gains in excess of 50 dB are seldom used.

Radomes, heated or unheated, can be used to reduce icing effects and to reduce wind loading. Cylindrical shrouds will give major reductions in side and back lobe radiation. One-foot to four-foot antennas, now very common at 18 and 23 GHz, usually mount on 3-inch vertical pipes; larger antennas use 4.5-inch pipes. All waveguide or air-dielectric coaxial feeds must be airtight, as any moisture will attenuate the signals; they are usually fed with dry air from the waveguide system (see below).

Horns

Several variations of the horn antenna are available and they provide wider bandwidth and better radiation patterns than parabolas. Their weight and wind loading, however, make severe demands on the tower or mast structure, and their use is limited to severely congested areas.

Omnidirectional

The omnidirectional antenna is characterized by a doughnut-shaped radiation pattern. Gain is achieved by reducing radiation in the vertical direction. These antennas are often used to eliminate the need for tracking in mobile systems, with typical gains of 6 dBi on the vehicle and 10 dBi at fixed points. The higher gain antenna is not used on helicopters or motorcycles as the signal loss during vehicle banking would be excessive.

Passive Reflectors

Tower or mast mounted passive reflectors have often been used at 7 and 13 GHz to avoid the cost and loss associated with long runs of waveguide. The disadvantage is the stray radiation from the lattice structure of the tower which reduces the back-to-front ratio and prevents re-use of frequencies at a given site. The installation comprises a parabolic antenna near the foot of the tower (but preferably far enough away to avoid problems with falling ice) pointing nearly vertically to a flat reflector mounted at approximately 45° at the required height on the tower. Typical reflector sizes are 6 × 8 feet to 10 × 15 feet. The associated parabolics range from 4 to 8 feet diameter, the larger sizes being used for larger antenna to reflector spacings. It is important to match the antenna and reflector sizes; over or under concentration of energy at the reflector can reduce the efficiency of the system.

The "gain" of a passive system is defined as the difference between the effective radiated power of the antenna-reflector combination and that of the feeding

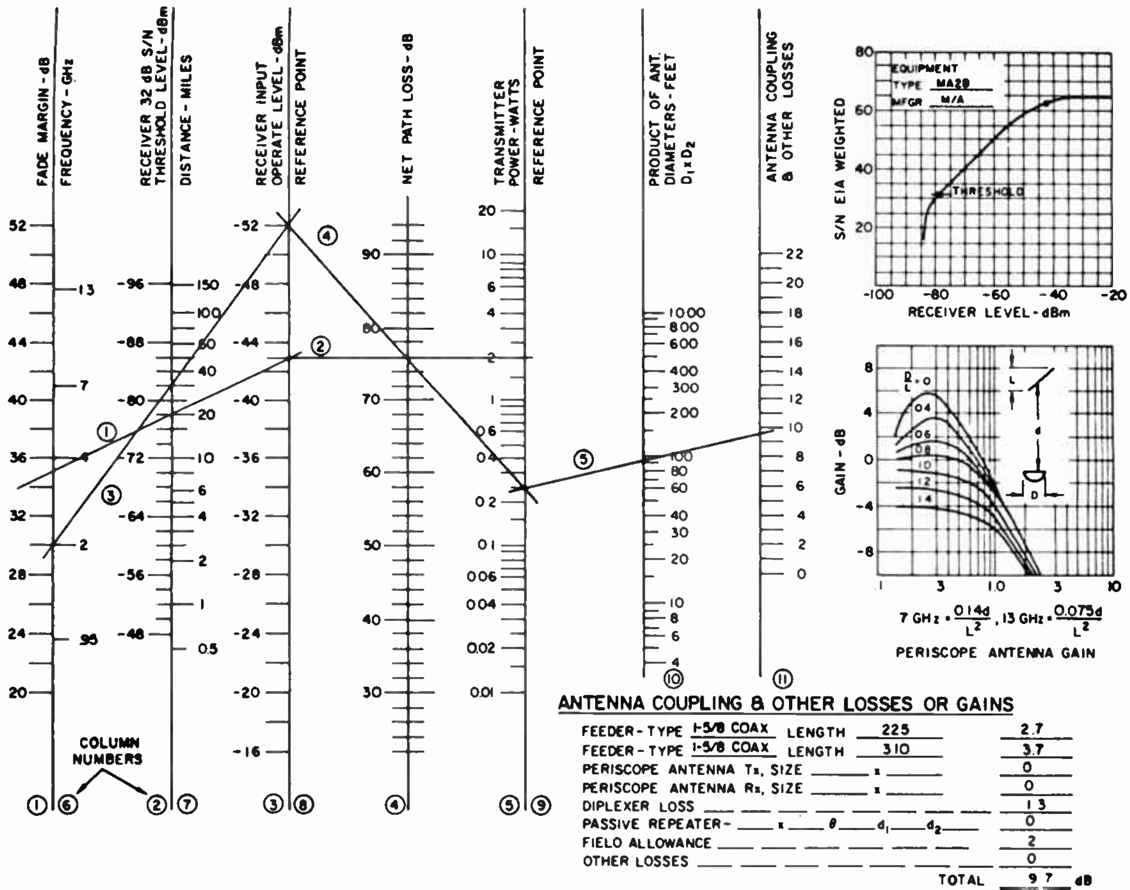


Figure 11. Transmission calculation worksheet.

antenna; this can be a positive number when the reflector has significantly more area than the feeding antenna. Gain can be estimated from the nomograph in Fig. 11.

Transmission Lines

Microwave equipment is connected to the antenna with either coaxial cable or waveguide. Cable is used below 3 GHz and has foam or air dielectric. Waveguide is used above 3 GHz and may be either elliptical flexible, rigid rectangular, or circular. Short sections of flexible rectangular waveguide are often used at the equipment interface and behind the antenna to simplify installation and alignment.

In addition to the transmission line, various hangers, clamps, adapters, grounding kits, bending tools, hoisting grips, and pressurizing equipment are required to complete an installation. Flexible guides and cables should be installed in continuous lengths to avoid problems with splices or joints. Chapter 2.3 of this *Handbook* provides detailed information on transmission lines and installation.

Care must be taken during installation to avoid pressure leaks, dents, and discontinuities, any of which will affect the system performance. Professional, experienced riggers are normally used with supervision by the equipment manufacturer. Foam-filled coaxial cable is the easiest to install but it has higher loss than the air-dielectric type; it must only be used with a foam-filled feed in the antenna.

Waveguide or air-dielectric cable must be maintained at a positive pressure (0.5 to 5.0 psi) above atmospheric with dry air or nitrogen. Compressor/dehydrator units to supply this dry air can be fully automatic (the desiccant is dried out periodically by the unit) or semi-automatic (where the desiccant must be dried out or replaced by the operator). To avoid rapid cycling of the compressor, the volume of air in the system should not be less than 1.5 cubic feet; this can be achieved in small systems by adding a storage tank.

Passive Repeaters

When two microwave terminals are less than 10 or 15 miles apart, but the direct path is blocked by an obstruction, it may be possible to employ a passive repeater to bypass the obstruction. The passive design is dependent upon the included angle between the two sections of the path at the repeater. If the angle is less than about 120 degrees a single flat panel can be used as the reflector; typical sizes at 7 or 13 GHz are from

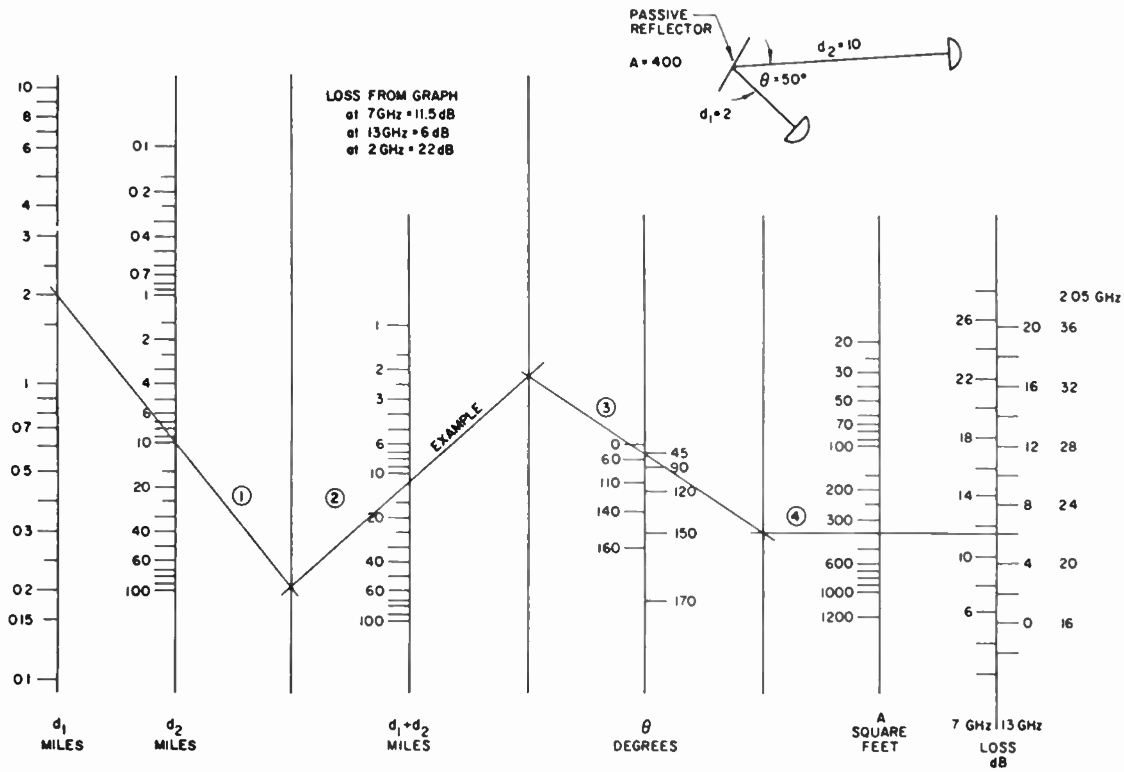


Figure 12. Passive repeater loss relative to direct transmission distance over $D_1 + D_2$.

8 × 10 feet to 20 × 32 feet. Even bigger ones can be used at 7 GHz while special ones with carefully controlled surfaces are available for 18 and 23 GHz in smaller sizes.

For angles greater than about 120°, an arrangement of two closely-spaced flat reflectors, or two parabolic antennas back-to-back with a short section of waveguide between them, can be used. In the latter case, a low-noise amplifier can be fitted in the waveguide to improve performance; care must be taken to prevent overloading the amplifier if the repeater is near one terminal of the system and to avoid feedback between the antennas.

Fig. 12 is a nomogram with which the transmission loss through a passive repeater relative to the direct free-space loss can be read. It is clear that the loss is significantly less at higher frequencies, hence passive repeaters are seldom used below 4 GHz.

Because the passive repeater requires no power, and little maintenance, its use should not be overlooked when designing microwave systems. For more information on the design, performance and installation of passive repeaters, see ref. 7.

RADIO CONSIDERATIONS

Transmission of video and audio signals over the last forty years has been almost exclusively carried out using frequency modulation of the microwave carrier by an analog signal.

Since the broadcast television transmitter uses that analog signal, it is probable that analog video transmission will predominate for another decade; by that time the requirements of high definition television (HDTV) may lead to new broadcast transmitter standards and a whole new look at link design. Audio information is generally frequency-modulated onto a subcarrier above the video, though digital modulation in the sync pulse is widely used in Europe, and the use of digitally-modulated subcarriers is increasing.

The microwave transmitter terminal thus comprises a baseband processor and frequency modulator (directly to RF or via an intermediate frequency). The microwave receiver terminal comprises a down-convertor (invariably to an intermediate frequency), demodulator, and baseband processing.

Repeaters

The choice of transmitter type is determined by the design of the repeater. Two approaches are possible. The received signal may be demodulated to baseband and applied directly to an RF oscillator in the following transmitter; this is the *remodulating repeater*. Alternatively, the intermediate frequency from the receiver can be mixed with a local oscillator in the following transmitter and the desired sideband selected and amplified to produce the transmitted signal; this is the *heterodyne repeater*. The advantages of the remodulating approach are that the modulated oscillator can be a relatively high power device so that little or no RF amplification is needed. Because the carrier is regenerated at each repeater there is no cumulative error in the carrier frequency. Both these factors lead to lower costs. The disadvantage is that multiple demodulations lead to level instabilities and bounce; after about three repeaters it is desirable to fit clamping circuits. For this reason, remodulating systems are usually less than six or seven hops.

The advantages of the heterodyne repeater are that once the modulation process has been accomplished at the initial transmitter, the deviation (and hence the level) will not change; other sources of noise and distortion associated with the modulator occur only once in a system of any length. The disadvantages are that to ensure that a carrier frequency error does not build up over a multi-hop system, either a very stable oscillator, or an automatic frequency control (AFC) system, is required in the receiver. Also, as the transmitter involves a mixing process, the power level is low and considerable RF amplification is needed. These factors lead to higher costs. The heterodyne repeater is used in longer microwave systems.

Transmitters (Remodulating)

The basic remodulating transmitter (Fig. 13) comprises a video amplifier (with preemphasis), an oscillator which can be frequency modulated, a crystal-referenced AFC loop, and a power amplifier and/or multiplier to deliver the required RF power. Various optional items are available to extend the usefulness of the transmitter; these include audio subcarrier modulators, pilot generators, hot-standby protection switches, and off-air monitors. Portable equipment may have multichannel switching. The output will have a channel filter and provision for combining the signal with other transmitters and/or receivers.

The baseband amplifier will be wideband (10 MHz or so) to ensure linearity and permit multiple subcarriers above the video signal. Preemphasis is used to improve the signal-to-noise ratio and reduce the effects of the DC component of the video signal. Combining circuits are used to inject the subcarriers and gain controls allow the deviation of the primary carrier and the subcarriers to be set correctly.

The transistorized voltage-controlled oscillator (VCO) typically runs at 2 GHz where high power is readily available. The output may be sufficient to use directly; more often a 10 dB amplifier is used. To generate frequencies in the 6–8 GHz band the 2 GHz signal is amplified and then multiplied by 3 or 4 times (the deviation of the primary oscillator being reduced to allow for this); for 13 GHz a further doubler is used. The power amplifier consists of one or more broadband stages to provide a power of 8 to 10 watts at 2 GHz, giving 1 to 2 watts at 7 GHz and 0.5 to 1 watt at 13 GHz. The multipliers use varactor diodes, matching circuits, and filters for efficient operation.

Two approaches for AFC circuits are common. In one, a crystal controlled source feeds one side of a

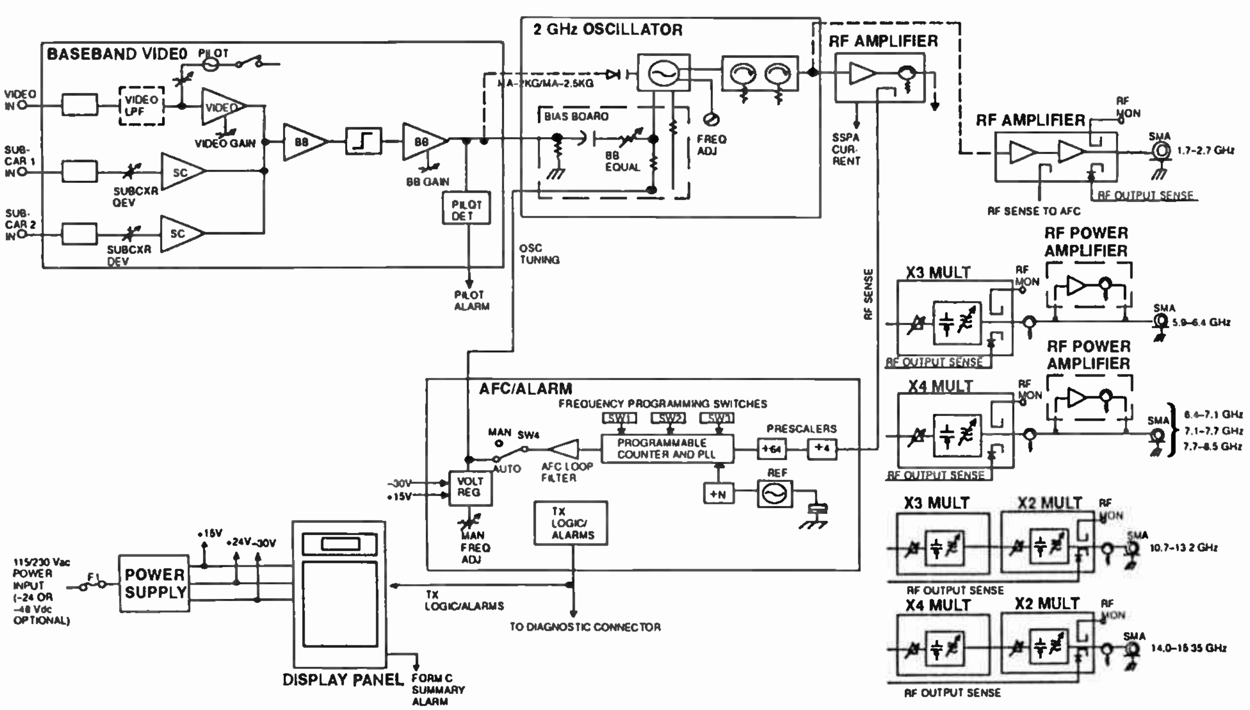


Figure 13. Remodulating transmitter—Functional block diagram.

mixer; the other side is fed with a sample of the oscillator frequency. The intermediate frequency from the mixer feeds a stable discriminator whose DC output is fed via an amplifier to control the oscillator frequency. Fig. 13 shows a more state-of-the-art method in which a microwave IC divider samples the output of the 2 GHz oscillator and compares the phase of the divided signal with a crystal oscillator giving an exact phase lock rather than an approximation. The phase-lock approach involves less hardware and is thus more reliable and less expensive.

Portable equipment which may have to operate on any channel in the band will have either multiple crystals or, more frequently, a frequency synthesizer as the reference in the AFC circuit. Since other circuits are broadband, no tuning is necessary when the reference is switched to another channel.

Fixed transmitting equipment is usually supplied with a channel filter; this reduces spurious emissions and permits the transmitter to share the antenna and feeder with other transmitters or receivers. Portable equipment may have a filter. For portable applications a weatherproof housing is required which can be easily fitted, with an antenna, to a tripod. If remote control is used, special wiring and a connector are installed for mating with the remote control cable.

Equipment for mobile applications must be packaged in a robust manner to withstand the high level of vibration and shock experienced in most vehicles. Also, since size and weight are restricted, care is needed to prevent overheating.

To prevent interaction with broadcast transmitting equipment, special precautions must be taken to reduce radio-frequency interference (RFI) by filtering and shielding power, baseband, and IF circuits and connectors. Both STL and portable microwave are often required to operate close to high-power transmitters.

Transmitters (Heterodyne)

A typical heterodyne transmitter block diagram is shown in Fig. 14. It comprises an intermediate frequency (IF) amplifier and a microwave source both of which feed an up-converting mixer. One sideband, usually the upper one, is selected by a filter or by the use of an image-rejecting mixer. The mixer output, at a level of around 0 dBm, is then amplified by 30 to 40 dB to provide the transmitter output.

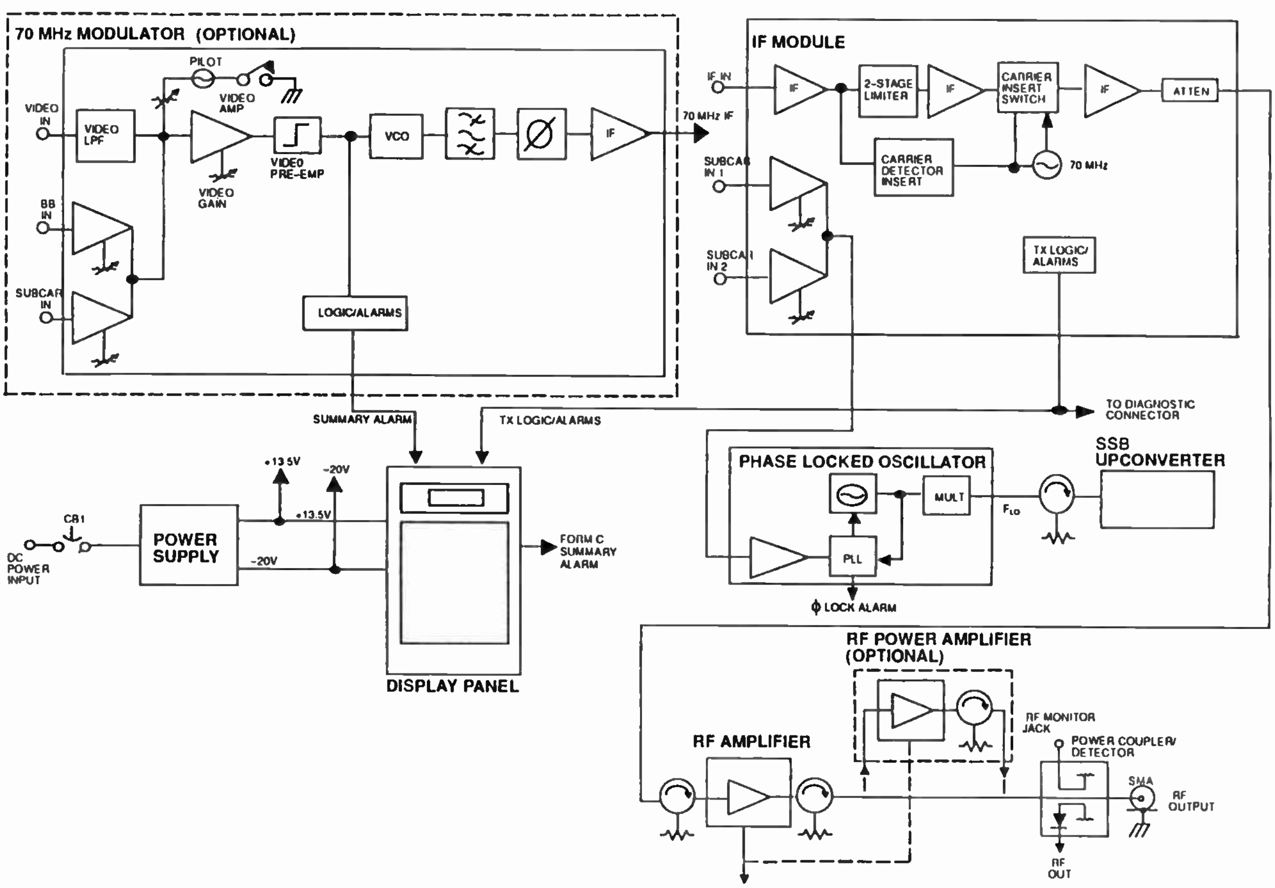


Figure 14. Remodulating transmitter—Functional block diagram.

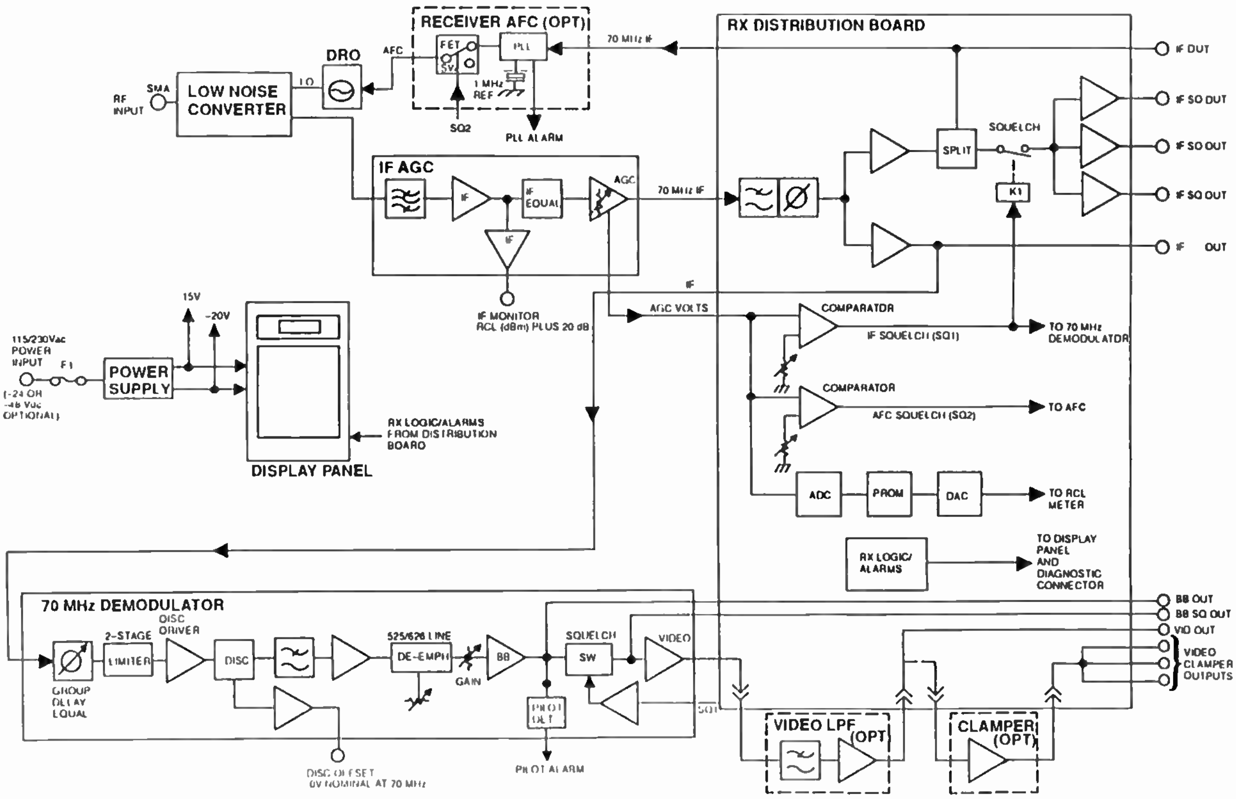


Figure 15. Receiver—Functional block diagram.

At the transmit terminal, the modulation is applied to a 70 MHz frequency modulator which produces a constant voltage at a frequency proportional to the applied voltage; the deviation sensitivity is such that a 1 volt peak-to-peak signal causes a deviation of 8 MHz peak-to-peak. The modulator has a baseband amplifier with provision for the injection of subcarriers identical to that feeding the oscillator of the remodulating transmitter. At a repeater, the 70 MHz output from the receiver is used instead of that from the modulator. To maintain output in the event of failure of the IF input, the IF amplifier is preceded by a monitor which can insert a locally generated 70 MHz carrier (*carrier reinsertion*). This is important as the RF output is not being regenerated at each transmitter but is an amplified version of the original input to the system.

Note that in Fig. 14 the local oscillator has provision for modulation; this is an important feature which permits the injection of a single-sideband suppressed-carrier signal with order-wire or supervisory signals even though the through signal has not been demodulated. Since this injection follows the carrier-reinsertion feature, alarms generated at a repeater site can still be sent out, even though the preceding receiver output has failed.

Receivers

The same basic receiver design is used with both remodulating and heterodyne transmitters. Fig. 15 is a typical block diagram. In fixed equipment the RF input is invariably via a channel filter; this facilitates multiplexing with other receivers or transmitters, suppresses local-oscillator radiation, and eliminates any response at the image frequency. In Fig. 15 the input is amplified at RF in a low-noise amplifier (LNA) before the image-rejecting mixer. Following the mixer is an amplifier, a filter-equalizer which provides a flat passband and high adjacent-channel rejection, and a further high-gain amplifier with automatic gain control (AGC). This latter feature ensures a constant level of the 70 MHz signal at the receiver output despite level changes of 50 to 60 dB in the received carrier level (RCL).

In the heterodyne repeater, one IF output is used to drive the following transmitter via the carrier reinsertion unit; another output may be used to drive a demodulator for the single-sideband suppressed carrier (SSB-SC) orderwire and supervisory channel, or for monitoring.

The local oscillator may be a Gunn or dielectric-resonance (DRO) type oscillator. In the remodulating case a stability of 0.005% is sufficient as any errors are not passed on to the following transmitter. In the heterodyne case frequency errors can add up in the system and either a stability of 0.001% or better, or an AFC circuit which will ensure a 70 MHz output, are necessary; Fig. 15 shows the latter arrangement.

Optimum receiver noise performance is achieved when the minimum necessary bandwidth is employed.

This is typically 20 to 25 MHz between the 3 dB points for TV with up to four audio subcarriers and a service channel. For portable equipment with one or at most two audio subcarriers, the spacing of channels may require an IF bandwidth of 12.5 MHz. Recently, some major educational networks have been established with 4.2 MHz video bandwidths and a single subcarrier at 4.5 MHz using a 10 MHz IF bandwidth. By reducing the peak deviation, the quality can be maintained close to broadcast standards. In all cases the characteristics of the IF filter-equalizer must be carefully controlled for good phase and amplitude linearity.

The main IF amplifier follows the filter and provides 60 to 65 dB of gain. Since the input microwave level can vary between −20 and −85 dBm, while the output must remain substantially fixed at about +5 dBm, an automatic gain control (AGC) must be incorporated in the main IF amplifier. When very large variations are expected, as in portable applications where the path lengths can be very short, the AGC action may be extended to the LNA: this can prevent overloading of the mixer. The receiver should have multiple IF outputs available to allow for repeaters where the signal is being sent in more than one direction as well as feeding a discriminator: for preference these outputs should be available squelched or unsquelched.

The demodulator comprises a limiter, discriminator, deemphasis, and a baseband amplifier. The limiter is an essential part of an FM receiver in that it removes AM before the discriminator: threshold performance is significantly affected by limiter design. The discriminator demodulates the FM signal back to video. In a TV system the discriminator must be linear over a wide bandwidth about the 70 MHz IF, at least ±5 MHz for video, and more where multiple audio subcarriers are being used. Any nonlinearities will produce distortion and noise by generating intermodulation products.

Finally, the deemphasis network and baseband amplifier restore the baseband response and the required 1 volt peak-to-peak level of the video signal. The video output will need a low-pass filter to remove the audio and other subcarriers from the picture signal. Typically an attenuation of 35 to 40 dB is used.

Power Supplies

Microwave transmitters and receivers require voltages in the 5 to 30 volt range. These are obtained from either a 60 Hz AC supply or from a DC source (usually at 24 or 48 volts).

Fig. 16 is a block diagram of a typical power supply system. AC units comprise a transformer and rectifiers to obtain the unregulated operating voltages. Various types of transistor regulators are used to keep the operating voltages within a narrow range. DC units, when used, are chopped to provide an AC input at a higher frequency than 60 Hz to reduce transformer size and improve efficiency. Good designs tend to employ protection circuits on both inputs and outputs to guard against overloads such as lightning surges and accidental short-circuits.

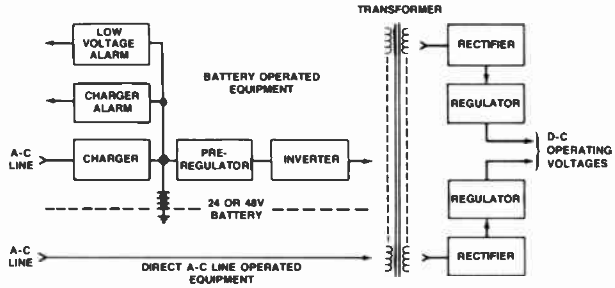


Figure 16. Power supply—Functional block diagram.

Subcarriers

In addition to carrying the video signal, a microwave link can carry additional information above the video portion of the baseband. As the NTSC video bandwidth does not extend above 4.5 MHz, the region above this, after allowing a suitable guard band, is available for other traffic. Four different uses are made of this region: one or more audio program subcarriers; engineering order-wire; supervisory signals (alarms and controls); and a continuity pilot. All these subcarriers generally deviate the carrier at 20 to 26 dB below the 4 MHz video peak.

FM audio subcarriers are generally deviated at 75 kHz peak at 1 kHz. To improve the signal to noise ratio, the audio signal is preemphasized. Maximum audio bandwidth is 15 kHz, giving an information bandwidth, using Carson's rule, of 180 kHz. To allow for filtering a minimum subcarrier spacing of 500 kHz is used, with increasing separation at higher frequencies to avoid beats between subcarriers falling into other channels. Typically a 15.7 kHz trap filter is fitted to reduce the effects of intermodulation with the sync pulses.

Up to four high quality program audio FM subcarriers are used, for TV stereo left and right, and for other signals such as second language or co-sited FM transmitters. As noted above, the number of subcarriers may be restricted if the RF channel spacing requires an IF bandwidth of less than 20 or 25 MHz. An encoded stereo signal (left plus right, left minus right, second audio channel and pro channel) can be carried on a single subcarrier, permitting the stereo encoder to be in the studio; this requires a special modulator and demodulator. Digital audio, using a modulator which transmits stereo audio as a 1.544 MHz T1 bit stream, is also used; the increased bandwidth required is compensated by an improved signal/noise ratio.

Program audio equipment may be supplied as a separate unit or packaged in with the radio terminals. In the FM case, the transmit end comprises a baseband amplifier with level setting controls to permit the recommended deviation to be achieved with inputs from 0 to +18 dBm, a preemphasis network, and a frequency modulator. The receive end comprises a filter followed by either a subcarrier amplifier or a mixer and IF amplifier, limiter/discriminator, deemphasis network, and baseband amplifier with level-setting capability similar to that at the transmit end. Unlike the video interface there is no universally accepted audio interface level; the impedance is invariably 600 ohms balanced at the transmit end and the same, or a lower impedance, at the receive end.

Often a SSB-SC AM signal is carried above the audio, at 8.59 or 8.85 MHz for example. By using a suppressed carrier, information can be inserted at repeater stations without carrier beats. In a typical SSB-SC system the bandwidth can be up to 108 kHz, with 0 to 4 kHz being used for an engineer's order-wire omnibus telephone system, 4 to 12 kHz for supervisory and control signals, and 12 to 108 kHz for up to 24 International Telephone and Telegraph Consultative Committee (CCITT) telephone channels.

In addition, a continuity pilot, at the CCIR recommended frequencies of 8.5 or 9.023 MHz, is used; this is essential if a protected system (hot-standby, twin-path or 1 + N type) is to be provided.

POWER PLANTS

With improvements in equipment and systems engineering the reliability of the primary power source becomes a major factor in overall system availability. Where terminals are in studios or TV transmitter buildings the AC supplies may well be protected with standby generators; in this case AC sourcing is the obvious choice. At repeaters the AC may be supplied to a remote site by overhead lines and these may be subject to interruption under adverse weather conditions. Since solid-state microwave equipment has essentially low power requirements, batteries are the most popular form of standby power. Where long outages are expected (say more than eight hours), an AC generator will be needed to recharge the batteries; this can be fueled by diesel or liquid gas. In addition, the use of solar power is increasing, sometimes with a small gasoline engine back-up to keep the battery capacity to an economic size.

Industrial lead-acid batteries are the most common type. Some can be sealed to prevent loss of liquid, reduce maintenance, and the need to ventilate the explosive gas given off when a battery is charged at too high a rate. Lead-acid cells are floated across the charger. Since the different types of lead-acid cell have differing float voltages (2.15, 2.2 and 2.23 V for lead-antimony, lead-calcium, and lead-selenium respectively), it is important that the charger and battery are matched; for this reason it is recommended that both are ordered from the same supplier.

Keep in mind that a battery of 12 cells with a nominal voltage of 24 volts will float at 26.8 volts, dropping to 24 if the charger fails, and finally falling to 21 volts at the end of the standby time. At that time a low-voltage disconnect should operate to prevent damage to the cells. The radio and other equipment running from the battery must be capable of accepting both the elevated running voltage and the end of standby voltage.

To determine the ampere-hour (AH) capacity of a cell required to give a particular standby time (usually 8 hours) the steady current drain must be determined (A amperes). Depending on the type of regulation used in the equipment power supplies this may be higher at 26.8 or at 21 volts; take the higher value. To arrive at the required cell capacity multiply A by the standby time H (usually 8 to 12 hours). This gives the capacity in ampere-hours. The minimum size of charger (C amperes) to meet the station load and at the same time to recharge a discharged battery in R hours is given by:

$$C = (A + AH) \times 1.1 / R$$

Since charger failure would lead to station failure after H hours, it is common practice to use a duplicated charger for full protection at key sites; in this case each charger need only have the capacity A rather than C as calculated above. The chargers must be designed to share the load.

The battery acts as a large capacitor, reducing any ripple voltage generated by the charger and protecting against power-line surges. The station load is supplied by the charger output under normal conditions, with the battery floating across the load. A good charger will have a low-voltage disconnect to isolate the battery when discharged, and overvoltage protection and charge rate limiting to protect against surges and gassing by the cells.

When diesel generators are used as the primary source of power they should be run at 75% or more of their rating (after allowing for any de-rating for altitude) to prevent oiling up.

Solar power using photo-voltaic cells is attractive as there are no moving parts requiring maintenance. However, the batteries used to maintain the supply during hours of darkness or heavy overcast (which could exist for several days in some parts of the country) can be the most expensive part of the installation. It may pay to have a small gasoline generator which could recharge the batteries if the solar input fails for more than three days, rather than batteries to maintain the supply for six days; such a generator could also power lights and test gear during routine site visits.

FREQUENCY PLANNING

Frequency coordination for the broadcast microwave industry has been primarily a local case-by-case selection process through cooperation of the respective chief engineers. Most areas, through the local chapter of the Society of Broadcast Engineers (SBE), maintain an up-to-date log of existing licenses.

Ideally, frequency planning for future growth should follow certain simple ideas. Foremost among these is the concept of dividing the available band into "high" and "low" parts, and trying to keep transmitters at a given site in one or the other part of the band. This ensures the maximum utilization of frequencies at every site. Adjacent channels should be cross polarized; second channels on a given route should be co-

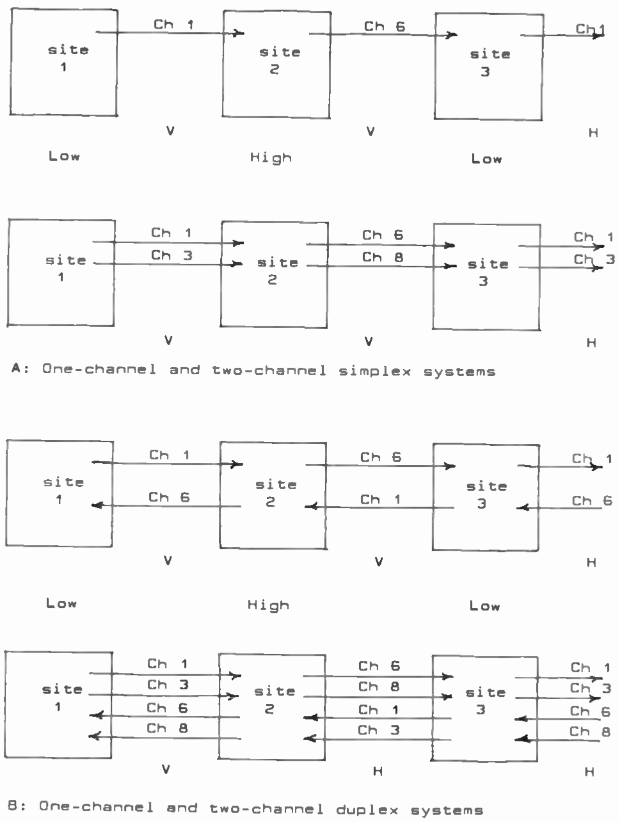


Figure 17. Recommended initial and expanded frequency plans; channels 2, 3, 7, 9 can be added on the opposite polarization, and channels 5, 10 can be used for spur routes.

polarized on the next-but-one channel; if a "go" channel is on an odd channel the corresponding "return" channel should be on an even channel to avoid intermodulation products (2A-B).

The diagrams of Fig. 17 illustrate good practices.

Frequency Re-Use

Microwave antenna discrimination provides a degree of isolation that very often allows the re-use of the same channel at the same or a nearby site. On parallel routes where the terminals at each end are several hundred yards apart it is often possible to use the same channel on both routes but in different directions. Every case has to be studied carefully to take account of the free-space losses and the angular discrimination of the antennas.

EIA/TIA Telecommunications Systems Bulletin 10E, "Interference Criteria for Microwave Systems," specifies a 60 dB co-channel carrier/interference (C/I) ratio. This is ideal but in many cases may not prove possible. Lowering the C/I ratio will raise the effective noise level and hence reduce the fade margin. Given a threshold of −85 dBm, say, the noise level will be approximately −95 dBm. Interference at −95 dBm then, will double the noise and raise the threshold to −82 dBm. This, or an even higher level, may be acceptable in a difficult situation. The interferer must at all times be below the threshold of the receiver or capture will occur if the wanted signal fails.

FCC LICENSING

All broadcast microwave systems, whether for fixed or mobile applications, must be licensed by the FCC. The application is filed under Part 74 of the Rules on Form 313, with pertinent characteristics and Type Acceptance Identification Number provided by the manufacturer of the chosen equipment.

Frequency coordination for clearance of the applied-for-frequencies is the responsibility of the local area engineers. In contrast, licensing by Common Carriers, Industrial Service, and CARS Band operators is supported by formal frequency coordination studies. Broadcast engineers should be aware and careful when applying for channels in the 12.7–13.25 GHz band. This band is shared with CARS and Cable TV operators and is in wide use in most areas.

The details in Form 313 are straightforward, and the explanatory booklet which is supplied by the FCC is very thorough. Note that the output power given in the Type Acceptance is a maximum; when quoting antenna input power and effective radiated power (ERP) the true transmitter power less any feeder losses (for input power), plus the antenna gain (dBi) for ERP, should be quoted.

The Emission Designator nomenclature has been changed since the seventh edition of this *Handbook*; 2 GHz systems are now 17MOF8W, while 7 GHz and 13 GHz systems are 25MOF8W.

Portable frequency-agile transmitters are typically licensed for all the channels that they can provide. For example, at 2 GHz, all seven channels are submitted on the single license application.

Broadcasters and others using this Chapter should refer to the FCC Rules & Regulations and to their attorneys for complete up-to-date information.

REFERENCES

1. Standard EIA/TIA-250-C: Electrical Performance Standards for Television Relay Facilities; Electronic Industries Association, Washington, DC, January 1990.
2. C.C.I.R. Documents of the Sixteenth Plenary Assembly, Dubrovnik, Yugoslavia: ITU, Geneva, Switzerland, 1986.
3. Reference Manual for Telecommunications Engineering: Roger L. Freeman; Wiley-Interscience, 1984.
4. Planning and Engineering of Radio Relay Links, 8th Edition: Brodhage & Hormuth; Siemens Aktiengesellschaft, Munich, FRG, 1977.
5. Statistics on Attenuation of Microwaves by Intense Rain: D.C. Hogg; Bell System Technical Journal, November 1969.

6. Andrew Catalog 34: Andrew Corporation, Orland Park, IL 60462, 1988.
7. Passive Repeater Engineering: Microflect Company, Inc., Salem, OR, 97309-0985, 1989.
8. NAB Engineering Handbook, Seventh Edition, Chapter 4.2: Carroll, Hickin, and McCarthy; National Association of Broadcasters, Washington, DC, 1985.

BIBLIOGRAPHY

Andrew Corporation. *Andrew Catalog 34.* Orland Park, Illinois: Andrew Corporation, 1988.

Brodhage & Hormuth. *Planning and Engineering of Radio Relay Links, 8th Edition.* Munich, FRG: Siemens Aktiengesellschaft, 1977.

Carroll, J.C., E.M. Hickin, and D.J. McCarthy. "Microwave Engineering for the Broadcaster." In *NAB Engineering Handbook, Seventh Edition*, edited by E.B. Crutchfield. Washington, DC: National Association of Broadcasters, 1985.

Electronic Industries Association. Standard EIA/TIA-250-C: *Electrical Performance Standards for Television Relay Facilities.* Washington, DC: Electronic Industries Association, January 1990.

Freeman, Roger L. *Reference Manual for Telecommunications Engineering.* Wiley-Interscience. 1984.

Hogg, D.C. "Statistics on Attenuation of Microwaves by Intense Rain." *Bell System Technical Journal*, November 1969.

International Radio Consultative Committee. *C.C.I.R. Documents of the Sixteenth Plenary Assembly, Dubrovnik, Yugoslavia.* Geneva, Switzerland: ITU, 1986.

Microflect Company, Inc. *Passive Repeater Engineering.* Salem, Oregon: Microflect Company, Inc., 1989.

4.3 Aural Broadcast Auxiliary Links

M.E. McClanahan
Marti Electronics, Inc., Cleburne, Texas

INTRODUCTION

The evolution of aural broadcast auxiliary communications has resulted in the development of higher audio quality, improved reliability, and cost effectiveness. This chapter covers the types of auxiliary communications and equipment most needed at radio stations, the information or data required for system design, the design of simple one-hop line-of-sight paths, ways to deal with problem paths, and, finally, how to provide back-up communications to survive auxiliary equipment failure.

SYSTEM REQUIREMENTS AND EQUIPMENT

Studio-Transmitter Link (STL)

950 MHz Bands

Several thousand 950 MHz STLs are in operation in the U.S. These line-of-sight (LOS one-way simplex) links are used to carry broadcast program channels as well as remote control and other communications on subcarriers from the studio to the remote transmitter location. 950 MHz studio-to-transmitter links (STLs) have proven to be cost effective, reliable links over typical paths of three to 15 miles. Reliable paths of 15 to 40 miles are possible over favorable terrain using tall towers and 6, 8, or 10 feet. diameter parabolic antennas. Details of radio path analysis, frequency coordination, and licensing for 950 MHz STLs will be covered later in this chapter. Most 950 MHz aural STLs use FM modulation and are licensed under Part 74, Subpart E of the FCC Rules. As of July 1, 1990 all 950 MHz aural STLs must have an FCC equipment authorization signified by a label bearing a unique "FCC ID" number, except that some equipment manufactured prior to that date has been "grandfathered" by the FCC until July 1, 1993.

A monophonic (mono) 950 MHz aural STL is simply a transmitter and receiver having a 600 ohm balanced program audio input and output with a frequency response of 20 Hz to 15 kHz and optional subcarrier capability from 26 kHz to 92 kHz (or higher), for remote control or other narrowband uses (Fig. 1). Stereophonic (stereo) STLs may take the form of the *composite* (Fig. 2) or the *dual* which is a single channel per carrier (SCPC) system (Fig. 3). Both systems must fit within the bandwidth prescribed by the FCC. The composite STL is modulated by the same composite waveform used in FM stereo broadcasting, and therefore, can be fed directly into the exciter of the FM broadcast transmitter. For AM stereo, this composite waveform must be decoded into left and right channels before being fed into the AM stereo generator at the transmitter. A dual channel AM STL can be fed directly into the AM stereo generator or through audio processing equipment.

The following general comparison can be made between the composite and the dual aural STLs in the 944-951 MHz band:

1. The audio processor/stereo generator is located at the studio in the composite STL system, while the

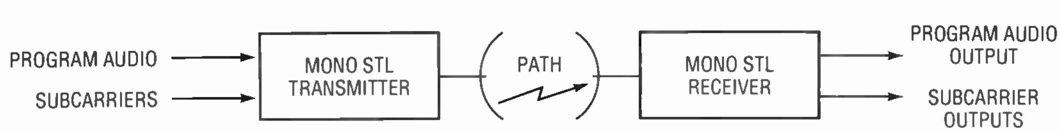


Figure 1. Monophonic STL with subcarriers.

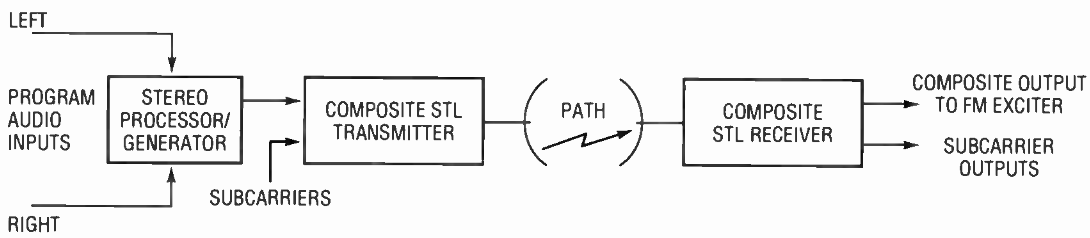


Figure 2. Composite stereo STL with subcarriers.

stereo generator must be located at the transmitter with the "dual" STL system.

2. Failure of a composite transmitter or receiver (without backup) can interrupt station operation. A similar failure in a dual channel STL would allow the station to continue operating in mono.
3. Composite receivers require about ten times as much RF signal level (in microvolts) as a dual channel receiver to produce the same noise quieting.

Duplex (STL/TSL) operation is not allowed under Part 74, Subpart E 950 MHz frequencies, however, transmitter-studio links (TSLs) are allowed under Part 94 of the FCC Rules in the 952-960 MHz as well as other bands. This will be covered in a later section of this chapter.

Digital 950 MHz STLs. Excessive bandwidth requirements have made transmission of digitized program material in the 300 kHz to 500 kHz bandwidth channel allocations of the 950 MHz aural STL spectrum impractical. Recent advances in audio source coding algorithms, along with high speed channel coding techniques, have led to the introduction of digital STL systems. These systems can, in effect, be thought of as high speed modems used to convey digital audio.

Advantages of a digital STL are numerous. The STL can now convey CD quality audio to the broadcast facility. Digital transmission requires less system gain than analog STLs. Higher system gain substantially reduces the investment a broadcaster must make in high-gain antennas low loss transmission line to achieve adequate signal-to-noise ratio (SNR) over longer STL paths. In analog systems, SNR depends on received carrier power. Digital systems deliver full SNR all the way down to the digital threshold. Fades have no affect on SNR. Because audio is multiplexed digitally, problems with crosstalk, phase, and gain are eliminated. In multiple hop configurations, analog STLs add noise and distortion. With digital STLs' SNR, frequency response and crosstalk remain consistent throughout the entire signal, regardless of the number of hops.

18 and 23 GHz STLs

Frequencies in the 18 and 23 GHz bands are available to radio broadcast stations for simplex (STL) as well as duplex (STL/TSL) use. These frequencies are regulated under Part 94 (Private Operational-Fixed Microwave Service) of the FCC Rules. 18 GHz channels are 5 MHz wide and 23 GHz channels are 50 MHz wide, which will allow use of digital modulation. 18 GHz quipment may use a maximum transmitter power output of 10 watts with a carrier frequency tolerance of 0.003%. 18 GHz equipment is therefore expensive and has found limited application in aural STLs. 23 GHz equipment is limited to a maximum transmitter output of 100 milliwatts and carrier frequency tolerance of 0.05%, which can be met by free-running Gunn diode oscillators. This power level and frequency tolerance has allowed manufacturers to produce equipment that is relatively low cost, resulting in many installations as aural STLs.

Signals in the 18 GHz band, and to a greater extent, the 23 GHz band, are severely attenuated by heavy

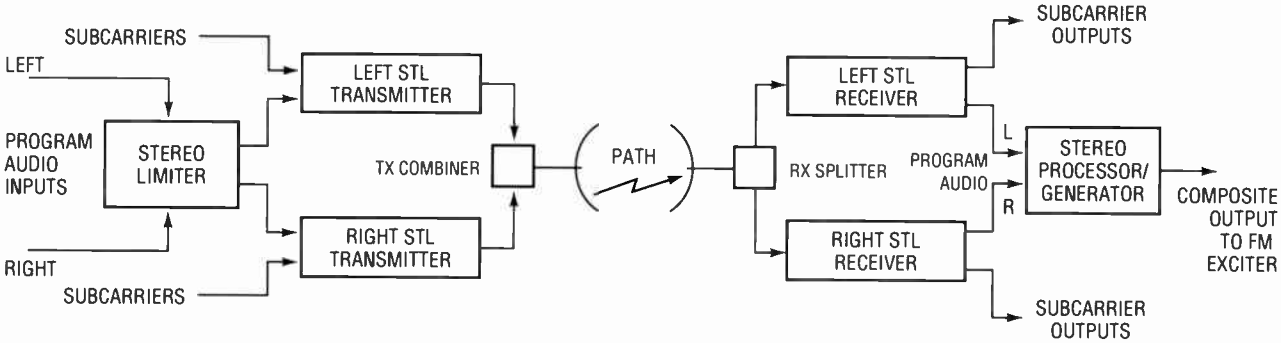


Figure 3. Dual stereo STL with subcarriers.

rainfall in the microwave path. Depending upon local annual rainfall statistics, signal interruptions of minutes or even hours per year can be anticipated on 23 GHz links of 7 miles or more. (Refer to Chapter 4.2, Outages: Rainfall.) While any program chain interruption is not tolerable in broadcasting, engineers have found ways to switch to backup STLs (950 MHz or wire) during 23 GHz outages. Accessibility of 23 GHz transmitters and receivers should also be considered. Because of cost and efficiency, the RF sections of 23 GHz systems are typically located at the antenna. This requires maintenance to be performed at installation height or equipment to be removed from the tower or supporting structure.

Low cost 23 GHz equipment, normally configured for video use, and having a baseband lowpass-filtered for video (0 to 4.5 MHz) and highpass-filtered for subcarriers (5 to 8 MHz), can be used for aural STLs. Audio can be carried on these systems with the addition of analog and/or digital processors. Some of the processing options are:

Video format PCM. Digital 16-bit pulse code modulation (PCM) combination encode/decode consumer grade equipment has been used successfully for an aural STL on 23 GHz video STL systems. This equipment usually has unbalanced audio input/output connections which can be a problem around high RF environments. Designed primarily for recording stereo audio on VCRs, the format contains additional bits not necessary for STL use. The 16-bit PCM performance can provide a 90 dB dynamic range, 80 dB stereo separation, and distortion less than 0.005%. With the introduction of digital audio tape (DAT) recorders, this relatively low cost consumer type equipment is vanishing from the market.

Satellite audio subcarriers. There are several sources of analog and digital subcarrier equipment designed primarily for use with satellite transponders that find application in 23 GHz aural STLs. The cost of these encode/decode systems can equal or exceed the cost of the 23 GHz RF system.

DS-1 based multiplexers. Digital multiplex equipment, based upon the telephone network DS-1 (T-1) standard 1.544 Mb/s data transmission rate, is finding application in 23 GHz STLs. Encode/decode systems for link use are available for transmission of two high quality channels, or, as shown in Fig. 4, transmission of digitized FM composite stereo signal plus optional audio channels.

STLs Available from the Local Telephone Company

The oldest form of STL is the equalized line furnished by the local telephone company. While still in use by some radio stations, these analog lines generally do not meet current requirements for FM broadcasting, and have been replaced by 950 MHz STLs. Renewed interest in telephone company facilities has resulted from the introduction of digital coder-decoder equipment mentioned in "DS-1 based multiplexers" above. These systems are compatible with the public DS-1 standards and can deliver 14-bit audio quality, which exceeds FM broadcast requirements. Cost and availablity of DS-1 service varies widely, so technical and economic considerations must be worked out with the local telephone company. The application of this technology is especially appealing in large cities where radio STL frequencies are crowded. In full duplex configuration, the STL and TSL needs of a radio station can be met.

Transmitter-Studio Links (TSL)

Remote transmitter sites not only require program audio and control data from the studio, but the studio must receive telemetry data from the transmitter site as well. If the STL is a duplex system, as is possible with some 23 GHz DS-1 systems, all requirements are met in a single system. The FCC Rules Part 74, Subpart E do not permit TSL operation in the 950 MHz STL band. Stations using 950 MHz STLs have the following options:

Narrow-Band (3 kHz) Telemetry On "P" Channels

Eight channels, 450.01, 450.02, 450.98, 450.99, 455.01, 455.02, 455.98, and 455.99 are allocated for this purpose (Part 74D). In congested areas, these few channels will not meet the needs of dozens of stations. "P" channel systems typically use Yagi directional antennas, transmitters of 30 watts or less, with automatic station identification in international Morse code.

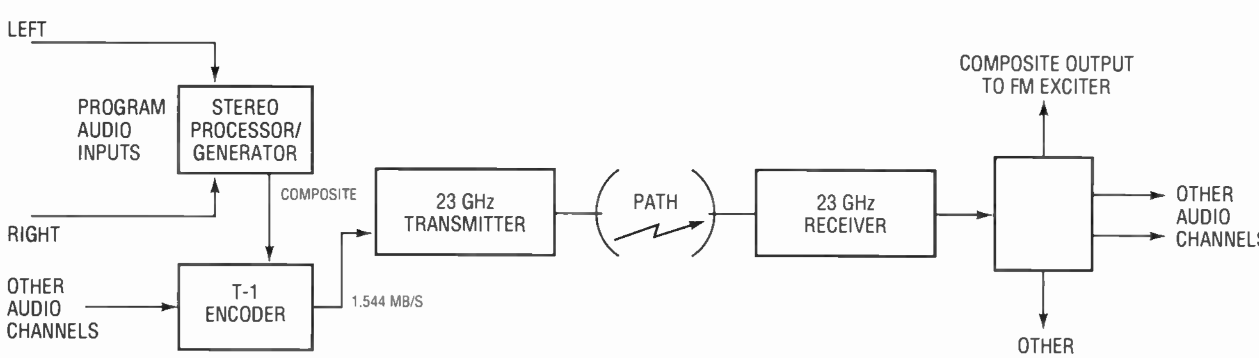


Figure 4. 23 GHz STL with T-1 encode/decode.

Telephone Lines

Unconditioned 3 kHz telephone lines are used by some stations for telemetry data return. Most remote control system modems use less than the maximum capacity of these circuits.

Telemetry on 952-960 MHz Part 94

When telco lines or "P" channels are not available or adequate FCC Rules, Part 94 Private Operational-Fixed Microwave Service (OFS) Link channels of 50, 100, or 200 kHz bandwidth can often do the job. 200 kHz channels can be multiplexed to carry one or more telemetry channels as well as remote pickup (RPU) backhaul audio and satellite network audio from the remote site. These links use channels coordinated by local frequency coordinators or frequency search firms and require directional antennas like those used for 950 MHz STLs. In some instances antennas for STL and TSL can be duplexed for economic or other reasons such as wind loading and tower space. Path calculation and frequency coordination are covered in following sections.

Intercity Relay (ICR) and Other Links

Programming opportunity or necessity has greatly increased the demand for broadcast-quality links between studios, from satellite receiving sites to studios, and between transmitter sites fed with duplicate programming. Many satellite network news and music channels are available, which has led to demand for links between remote satellite receive stations and studios to carry up to 15 program channels, requiring several multiplexed 950 MHz links. These links may use the 944-952 MHz band frequencies regulated under Part 94 of the FCC Rules or telephone company lines. Selection of radio links requires consideration of equipment cost, availability of tower sites, power, terrain, and frequencies. These considerations will be discussed in greater detail in the following section.

PRELIMINARY SYSTEM PLANNING CONSIDERATIONS

It is essential for all stations to have the latest FCC Rules and Regulations. They may be ordered by calling the Government Printing Office order desk at (202) 783-3238. Order 47 CFR Parts 70 to 79 and 47 CFR Parts 80 to end. Credit cards are accepted.

Frequency Availability, Coordination, and Licensing

The critical factor in broadcast auxiliary link planning is frequency availability. By FCC Rules, if a new link causes interference to existing links, the newcomer must solve the problem. Therefore complete and accurate frequency coordination information is required. In cases where a frequency is not available, a T-1 link from the telephone company, previously discussed, could be the answer.

Frequency coordination requires that the exact coordinates in degrees, minutes, and seconds at each end of the proposed link be known. From this, the azimuth of the main lobe of the transmitting antenna is determined. Often the same frequency can be shared by two or more links in the same area if their paths are on differing azimuths and opposite antenna polarization is used.

STL and ICRs in the 942-944 MHz Part 74 band are coordinated at the local level. Refer to Chapter 1.3, "Frequency Coordination" of this *Handbook*. If there is no local frequency coordinator, the frequencies in use by all local stations must be determined and a clear frequency selected from this data. License application is made on FCC Form 313.

A radio station may be authorized a maximum bandwidth of 500 kHz (300 kHz for FM STL and 200 kHz for AM STL) assembled from 25 kHz center frequency segments between 944.0125 and 951.9875 MHz. In noncongested areas, the FCC will continue to license 500 kHz bandwidth channels centered on the following frequencies: 944.5, 945.0, 945.5, 946.0, 946.5, 947.0, 947.5, 948.0, 948.5, 949.0, 949.5, 950.0, 950.5, 951.0, 951.5 MHz.

Part 94 frequencies in the 952-960 MHz, 18.76-18.82 GHz, 19.1-19.16 GHz, and 21.8-23.55 GHz bands require formal frequency coordination. License application is made on FCC Form 402. Certified supplemental showings of frequency coordination and interference analysis pursuant to the applicable Rules Section of Part 94 must accompany the application. Because of the time and extensive database required for Part 94 frequency coordination, it is advisable to have a consulting engineer or national frequency search firm do this work and file the application.

Terrain Considerations

The design of a broadcast auxiliary link must include a study of the terrain between the points of communication. All systems covered in this chapter require line-of-sight (LOS) paths, which means that no hill, building, tree, or other object can lie within the direct path, plus an additional clearance called the "Fresnel zone." Refer to "Clearance Requirements" in Chapter 4.2 for the procedure to determine Fresnel zone clearance. Topographic maps of your area can be obtained from local map dealers. First, request the "Catalog of Topographic and Other Published Maps" for your state and find the file number, name, and reference code of maps of the area covered by the radio path. While maps are accurate for terrain elevation, they do not include trees and structures in the path. These can be found only by traveling the path and noting estimated height above ground on the path profile.

Cost

In addition to frequency availability and path clearance, the economics of the auxiliary link must be considered. The initial cost of the equipment, installation, and maintenance must be determined. Equipment manufacturers or representatives will supply detailed proposals of equipment and costs. For comparison purposes, each proposal should include the complete system: the radio equipment, antennas, coaxial cable with connectors, jumper cables, weatherproofing and grounding kits, and other items. In the process of budgeting or justification for capital expenditures, management will consider the capital cost (and tax implications), the time value of money, operating and maintenance costs versus the return on investment including saving over other forms of communication, and intangible factors such as improved audio quality and reliability.

PATH ANALYSIS

By accounting for all gains and losses affecting the radio frequency energy emitted by the transmitter, the signal level available at the receiver can be accurately determined. Transmitter output power, antenna gains, and receiver sensitivity are considered system gains; while free space attenuation, coaxial cable, connectors, combiners, and terrain, are considered system losses.

The data and procedures for a 950 MHz path analysis are presented below. Refer to Chapter 4.2 which contains useful information for path analysis of 18 and 23 GHz links (including rain attenuation).

Path Profile

Accurately locate each end of the path on a topographic map and draw a straight line between these points. Using the map scale, mark one mile (or less) intervals on this line. Transfer all significant peak elevations along this line onto 4/3 earth curvature paper (available from equipment suppliers). Units of measure should be consistent (kilometers/meters or miles/feet). Add to the peak elevation points the heights of trees or other objects. See Fig. 10, Chapter 4.2.

Using the formula

$$H = 1368\frac{\sqrt{d_1 d_2}}{f(d_1 d_2)}$$

where H is the 0.6 first Fresnel zone clearance in feet, d_1 is the distance in miles from one end of the path to a point along the path, and d_2 is the distance from that point to the other end. Calculate the 0.6 Fresnel zone clearance at equal intervals along the path.

Add the proposed antenna height at each end of the path and draw a straight line between these elevations. Locate the calculated 0.6 Fresnel zone clearance points below this line and connect these points to form an elliptic curve of required clearances.

If ground elevation plus tree height intersects the 0.6 Fresnel clearance line at any point, it will be necessary to raise antenna height (preferably at the end nearest the obstruction) to provide clearance.

Having determined the required antenna heights, the coaxial cable lengths can be determined. The loss per 100 feet of typical cable is shown in Table 1.

TABLE 1

Coaxial cable and miscellaneous losses.

Cable Type	Attenuation dB/100 ft. at 950 MHz
Low Loss 50 ohm Foam Dielectric Coaxial Cable	
½ inch diameter	2.4
⅞ inch diameter	1.4
1⅝ inch diameter	0.8
Air Dielectric Coaxial Cable	
½ inch diameter	2.6
⅞ inch diameter	1.3
1⅝ inch diameter	0.7
Flexible Coaxial Cable	
RG-8/U, RG-214/U (use less than 2-foot lengths for jumpers.)	8.5

Miscellaneous Losses (at 950 MHz)

2-foot jumper cable with "N" connectors	0.5 dB
"isocoupler" tower isolater-950 MHz	0.5 dB
2-transmitter power combiner	3.9 dB
4-transmitter power combiner	6.9 dB
2-receiver power divider	3.5 dB

Path Attenuation Between Antennas

Free space path attenuation is calculated by:

$$\text{Path Attenuation (dB)} = 36.6 + 20 \log f + 20 \log d$$

where f is the frequency in MHz and d is the path length in miles.

Table 2 gives attenuation at 950 MHz for path lengths of 1-30 miles.

TABLE 2

Path attenuation at 950 MHz.

Distance in Miles	Loss in dB	Distance in Miles	Loss in dB
1	−96.2	16	−120.2
2	−102.2	17	−120.7
3	−105.7	18	−121.3
4	−108.2	19	−121.7
5	−110.1	20	−122.2
6	−111.7	21	−122.6
7	−113.1	22	−123.0
8	−114.2	23	−123.4
9	−115.3	24	−123.8
10	−116.2	25	−124.1
11	−116.9	26	−124.5
12	−117.7	27	−124.8
13	−118.4	28	−125.1
14	−119.1	29	−125.4
15	−119.7	30	−125.7

TABLE 3
Recommended fade margin versus path length.

Path Length	Fade Margin
5 miles	5 dB
10 miles	7 dB
15 miles	15 dB
20 miles	22 dB
25 miles	27 dB
30 miles	30 dB

Fading And Fade Margin

Signal strength variation on 950 MHz links designed for adequate clearance is due primarily to multipath wave propagation. Two or more signals arriving at the receiving antenna produce time-variable voltage vectors observed as increased or greatly decreased received signal. Multipath may be caused by atmospheric conditions or by reflection from the surface or other objects. Long paths over water usually experience strong multipath reflection and may require space diversity receiving systems. Those designing over water paths are referred to space diversity in the literature referenced at the end of this chapter. Table 3 gives recommended fade margins for various path lengths.

Other Losses

Coaxial cable length and its loss has been previously discussed. Other losses associated with antenna systems may include isocoupler, power combiner, power splitter, bandpass or notch filters, and coaxial connectors. Typical losses in dB for these items are shown in Table 1.

Transmitter Power (Gain)

It is convenient in path analysis to convert transmitter power in watts to dBm, which is the power level in dB above one milliwatt. Therefore:

$$\text{dBm} = 30 \log P(\text{watts}) \times 1{,}000$$

Table 4 gives dBm for typical transmitter power levels.

Receiver Sensitivity

Receiver sensitivity (gain) may be specified by the manufacturer in dBm or microvolts at the 50 ohm input connector for a certain signal-to-noise ratio. For path analysis purposes, select the receiver input level in dBm which corresponds to the minimum signal-to-noise ratio that can be tolerated for the link. If this level is in microvolts, use the microvolts to dBm conversion chart in Table 5 to obtain dBm.

Antenna Gain

Antenna gain is the variable most used in radio link design to satisfy path loss and fade margin requirements. Table 6 lists the gain in dBi for various full parabolic reflector diameters available for 950 MHz. When antennas are to be installed on existing towers,

TABLE 4
Transmitter power to dBm.

Transmitter Power (Watts)	dBm
1	30
2	33
3	34.7
4	36
5	37
6	37.8
7	38.5
8	39
9	39.5
10	40
11	40.4
12	40.8
13	41.1
14	41.5
15	41.8

For other power levels, use the formula:

$$\text{dBm} = 10 \log (\text{Watts}) \times 1000$$

TABLE 5
Conversion of microvolts to dBm (50 ohms).

Microvolts	dBm	Microvolts	dBm
2	−101	140	−64
3	−97	160	−63
4	−95	180	−62
5	−93	200	−61
6	−91	250	−59
7	−90	300	−57
8	−89	350	−56
9	−88	400	−55
10	−87	450	−54
11	−86	500	−53
12	−85	600	−51
14	−84	700	−50
16	−83	800	−49
18	−82	900	−48
20	−81	1000	−47
25	−79	1200	−45
30	−77	1400	−44
35	−76	1600	−43
40	−75	1800	−42
45	−74	2000	−41
50	−73	2500	−39
60	−71	3000	−37
70	−70	3500	−36
80	−69	4000	−35
90	−68	4500	−34
100	−67	5000	−33
120	−65		

TABLE 6
Typical isotropic gain for full parabolic antennas at 950 MHz.

Antenna Diameter	Isotropic Gain at 950 MHz.
4 feet	18.9 dBi
5 feet	21.0 dBi
6 feet	22.0 dBi
8 feet	25.0 dBi
10 feet	27.0 dBi

check with the tower manufacturer to verify that the tower will support the additional wind loading of the antennas. For regions subject to ice storms, use wind loading data for ice.

Path Data Calculation Sheet

Using the data obtained from the above topics, the overall path performance can be calculated by summing the system gains and losses to determine the received signal strength. First assume antenna diameters (gains) for transmitting and receiving in the Table 7 calculation sheet. If the resultant signal level does not provide the necessary signal level plus the required fade margin, then increase antenna diameter or transmitter power or make other system adjustments until the requirements are met. Conversely, if more than adequate signal is indicated, reduce antenna size or transmitter power to reduce system cost.

TABLE 7

Path data calculation sheet.

System Losses	
1. Path Attenuation _____ miles	_____ dB
2. Coaxial cable loss:	
(A) Transmitting ___ dB/100′ × ___ ft.	_____ dB
(B) Receiving ___ dB/100′ × ___ ft.	_____ dB
3. Total connector, jumper cable	_____ dB
4. Transmitter combiner loss (dual system)	_____ dB
5. Receiver coupler loss (dual system)	_____ dB
6. Isocoupler, pre-selector or other loss	_____ dB
Total loss	_____ dB
System Gains	
7. Transmitter output power	_____ dBm
8. Antenna gain:	
(A) Transmitting	_____ dBi
(B) Receiving	_____ dBi
Total gain	_____ dB
Subtract system gain totals from loss totals to obtain net system loss:	
Total System Loss	_____ dB
(subtracting) Total System Gain	_____ dB
Signal Level at Receiver	_____ dB
Subtract the actual receiver signal level (above) from the required signal level specified for desired signal/noise ratio to obtain system fade margin:	
Actual Receiver Signal Level	_____ dB
(subtracting) Required Receiver Signal Level	_____ dB
Fade Margin	_____ dB
If the fade margin obtained above does not meet system reliability requirements, re-calculate the system using larger antennas.	

Problem Paths

When an obstruction lies in the direct path, one of several alternatives may be used to "bend" the signal around or over the obstruction. Such solutions involve placing a relay station or repeater at a location that is line of sight to the desired points of communication. Each of the two paths thus substituted for the obstructed direct path are calculated as previously outlined. At the relay point, enough gain must be provided (antenna gain plus additional gain) to deliver the necessary signal level at the receivers and fade margin.

For example, consider a ten-mile path with a hill at the midpoint. A repeater on the hilltop will provide line of sight plus the necessary Fresnel zone clearance for the two five-mile paths. The original ten-mile path had a loss of 116 dB and each five mile path has a loss of 110 dB, for a total loss of 220 dB.

Passive Repeater

The simplest solution to the above problem would be to point high gain antennas at the STL transmit antenna and receive antenna and connect them with a short piece of coaxial cable. No electrical power would be required in such a passive repeater. The limitation of passive repeaters can be seen by working through the calculations previously outlined. In this example it will be found that even ten foot diameter antennas with 27 dB gain each at the repeater would give a marginal signal level at the receiver. Such repeaters are useful, however, on short paths or when located close to one end of the system. Passive repeaters at microwave frequencies are also covered in Chapter 4.2.

Active Repeaters

Active repeaters require a primary power source (commercial, solar/battery, etc.) and by providing large amounts of gain at the repeat point, can overcome the limitation of the passive repeater. The merits of several types of active repeaters used at 950 MHz follows.

Remodulating Repeater

A remodulating repeater is simply a receiver tuned to the first path frequency connected to a transmitter tuned to a different frequency for the second path. The interconnection is at the demodulated audio or composite signal band, which can add some distortion to the system.

Heterodyne Repeater

A heterodyne repeater down-converts the signal from the first path to an intermediate frequency (IF), amplifies it, then up-converts to a different frequency for transmission on the second path. Since remodulation is avoided, this repeater contributes little distortion to the system. A frequency stabilizing or "locking" feature may be required to prevent cumulative frequency error. (Repeaters are also covered in Chapter 4.2)

Microwave Booster

Much of the limitation of the passive repeater can be eliminated by simply adding an amplifier between the antennas. The booster amplifier has the advantage of spectrum conservation and lower cost than the remodulating and heterodyne repeaters. There is a limit to the amount of gain or amplification that can be added by the booster. To prevent oscillation or instability of this type of repeater, signal path attenuation between input and output antennas must be greater than the amplifier gain. By separating and cross-polarizing the antennas,it is possible, in a well designed 950 MHz booster repeater, to operate the booster amplifier at a power gain of 60 dB. To calculate the performance of a booster, determine the signal level at the input antenna, and add input and output antenna gains to the booster amplifier power gain to arrive at the signal power available for the second path. Analysis of the second path gives the level at the receiver. Good system criteria for a booster amplifier suggests an input level of −40 dBm to −33 dBm and output levels of +20 dBm to +27 dBm. Boosters must be operated within their linear range.

Active repeaters and boosters are licensed by the FCC under Part 74, Subpart E. Equipment that is FCC-approved under the Notification Process (74.550) is required and application is made on Form 313. The cost of a repeater is determined by the size of the antennas, power source, and equipment selected. Most manufacturers of broadcast auxiliary equipment will provide equipment proposals as well as technical assistance.

Backup Aural Broadcast Links

To prevent loss of valuable air time, backup for STL, TSL, repeater systems, and their power sources must be planned, implemented, and tested well in advance of need. The best, but most costly, backup system includes automatic switched hot standby transmitters and receivers for each radio link. Most modern auxiliary link equipment is designed for "hot standby" operation with the addition of the standby units and automatic switchers.

Aural broadcast remote pickup (RPU) equipment can be used as backup for STL (mono) and TSLs in an emergency, if the necessary antennas, and associated facilities are installed and tested prior to need. Station personnel should be given training and written instructions in the set-up and operation of backup equipment.

Emergency power for auxiliary links, especially remotely located equipment, can be critically important. Refer to the section titled "Power Plants" in Chapter 4.2.

STL/TSL Antenna Installation

Good system design can be neutralized by errors made during antenna assembly and installation. A technical representative of the radio station should supervise all work, checking each detail. Some of the details to check are:

1. Check for correct assembly of antennas. Grid reflector antennas must have the feed dipole parallel with the reflector grid.
2. Antennas must have the same polarity (vertical or horizontal) at each end of the path.
3. Pay special attention to field installation of coaxial connectors. Carefully follow manufacturers instructions. Check for short circuits with ohmmeter.
4. Do not strap STL/TSL coaxial cable to the broadcast antenna cable. Locate these cables on opposite sides of the tower to prevent coupling high RF levels into the STL/TSL system.
5. Do not allow coaxial cable to be damaged by sharp bending, crushing, or dents.
6. Bond outer conductor of coaxial cable to tower, using approved grounding kits, at top and bottom of cable run.
7. Tighten and moisture-proof all connectors of the system exposed to moisture.
8. Don't guess at antenna headings. On long paths, initially aim antennas by compass, then orient antennas in azimuth and elevation for maximum receiver signal. CAUTION: *Antennas have one "major" lobe and several "minor" lobes in their directivity patterns. Make sure the "major" lobe is selected.*
9. Feed power to each antenna system, and using the transmitter's directional wattmeter or an external wattmeter, determine the antenna VSWR. The VSWR should be less than 1.5:1 at the operating frequency.

BIBLIOGRAPHY

Capabilities of a 100 Milliwatt Nontranslating Repeater in the Aural Broadcast STL Service, G.W. Marti and M.E. McClanahan, Marti Electronics, Inc., 1973.

Communications System Engineering Handbook, Hamsher; McGraw-Hill.

Engineering Considerations for Microwave Communications Systems, GTE Network Systems, North Lake, IL.

Reference Data for Engineers: Radio, Electronics, Computer, and Communications. Seventh Edition. H.W. Sams Co.

Reference Manual for Telecommunications Engineering, Freeman: Wiley 1984.

4.4
Satellite Earth Stations

James H. Cook, Jr.
Scientific-Atlanta, Inc., Atlanta, Georgia

Gary Springer
Jorge B. Vespoli
Scientific-Atlanta, Inc., Norcross, Georgia

INTRODUCTION

The use of satellites for communications of all types is a technology that is highly developed and in widespread use throughout the world. It has progressed in two decades from a technology of limited acceptance to one of routine provision of services, including television and radio broadcast services. This acceptance is based on the unique characteristics of satellites placed on the geostationary satellite orbit. The capability of a single quasi-stationary repeater in the sky, visible to all of the contiguous continental states, offers unique distribution capabilities for broadcast services. Satellite communications is also particularly useful for long-distance communication services, for services across oceans or difficult terrain, and for point-to-multipoint services.

Satellites in the geosynchronous orbit rotate from west to east. They appear fixed in space to earth stations on the ground because they orbit in synchronism with the earth's rotation. A satellite that is closer to the earth orbits faster; one that is beyond synchronous orbit rotates slower than the earth. Compare the 90-minute orbit of the Space Shuttle, that operates roughly 150 miles above the earth, with the 28-day orbit of the moon. Satellites located in the 22,300-mile high geosynchronous orbit have direct lines of sight to almost half the earth, as shown in Fig. 1, therefore geostationary satellites are in effect unmanned relay stations. Except for small regions near

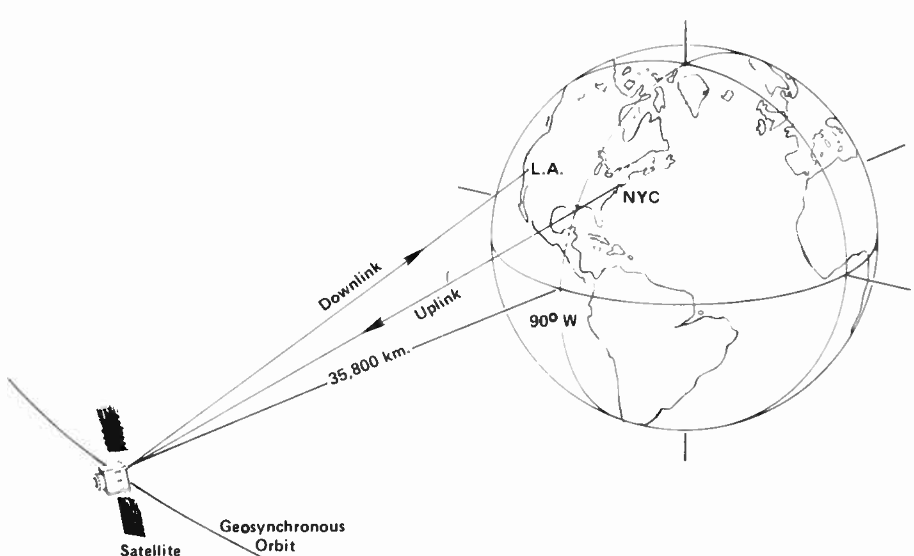


Figure 1. Satellite in geosynchronous orbit.

the North and South Poles, widely separated earth stations can be seen from a single satellite.

Communication by satellite was made possible by parallel advances in space technology and electronics. Arthur C. Clarke, the noted British scientist and science fiction writer, proposed relay stations in geostationary orbit for satellite communications in 1945. It took until 1963 for advances in technology, in solid state electronics, and in the thrust capability of rockets, to allow the placing of a satellite into a stationary orbit.

Communication by satellite is completely different from that by long-distance radio. Long-distance communication at radio frequencies is possible because the ionosphere, produced by radiation bombardment of the upper atmosphere by the sun, usually acts as a mirror to reflect certain radio waves back to earth. As the frequency increases, a critical point is reached where the ionosphere ceases to act as a reflector, letting the waves pass through into space. Of course, television signals in the frequency range above 54 MHz do not usually lend themselves to long-distance transmission. Therefore, long-distance transmission of television signals was accomplished by either coaxial cable or terrestrial microwave links prior to the advent of satellite communications. Transmission of radio frequency (RF) signals through the medium of fiber optics offers another choice for television distribution today.

Frequency

Communication satellites operate at microwave frequencies as shown in Fig. 2. At microwave frequencies the ionosphere is always virtually transparent regardless of sunspot activity or time of day, permitting continuous, almost loss-free transmission to and from satellites in orbit. In the United States the domestic commercial communications satellite networks operate in the Fixed Satellite Services (FSS) frequency bands as defined by the Federal Communications Commission. Most of the domestic systems operate in either the C-band (6 GHz and 4 GHz) or Ku-band (14 GHz and 12 GHz) frequency ranges with C-band generally preferred because of superior propagation characteristics. The Ka-band frequencies (30 GHz and 20 GHz) are also set aside for FSS operation but the use of these frequencies is still in the future except for experimental satellites (i.e., NASA Advanced Communications Technology Satellite, (ACTS)).

International systems provide services on a global basis (global beams) to all countries visible from a single orbit location, and on a regional basis with spot beams. Intelsat and Intersputnik are examples of this type of system, using both C-band and Ku-band. The international satellite communication frequency bands are similar to the U.S. frequencies at C-band, but somewhat different at Ku-band. The frequency bands are determined by joint negotiations by the countries of the world through the auspices of the International Radio Consultative Committee (CCIR) of the International Telecommunications Union (ITU). International systems have begun to use an extended C-band frequency range of 3.4 GHz to 4.2 GHz rather than the 3.7 GHz to 4.2 GHz U.S. band as well as 11.2 GHz to 11.7 GHz for downlink transmissions.

Satellite Stationkeeping

It was stated that the synchronous satellite appears stationary in space. Actually, a synchronous satellite is never perfectly stationary, because a number of forces including the pull of the sun and the moon perturb its orbit. If left alone, the satellite would eventually drift out of orbit. To overcome this, the position of the satellite is continuously monitored by an earth station, called a TT&C (telemetry, tracking, and command) station, and small jets of propellant such as hydrazine are used to keep it in position within a stationkeeping box. The stationkeeping box is typically a square ±0.1° or less on each side and oriented with the sides parallel and perpendicular to the orbital plane. Sufficient rocket propellant must be carried on board to last for the satellites predicted life, usually from seven to 10 years. In the international arena, Comsat has introduced a technique to extend the life of a satellite during its latter years. This

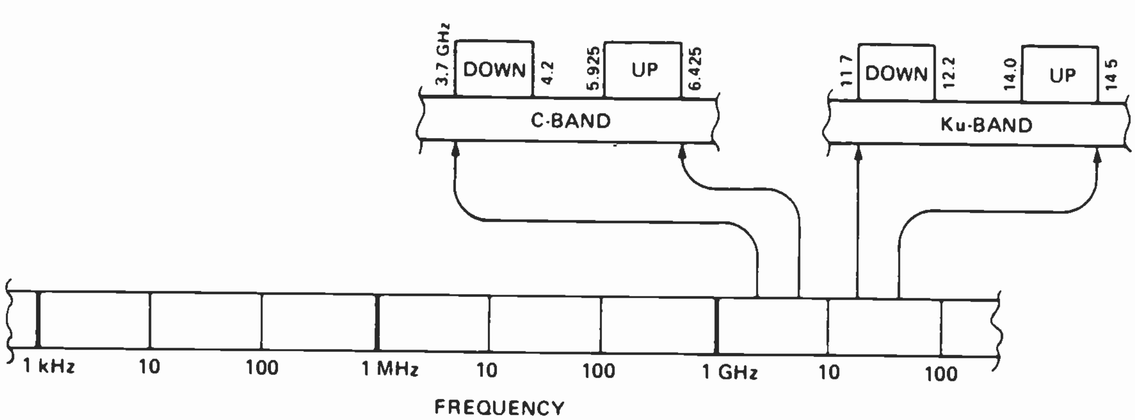


Figure 2. Frequencies in the microwave range of communications satellites.

technique, the "Comsat maneuver," allows the satellite to drift north and south an increasing amount as the satellite approaches its end of life with excursions of several degrees by the time the rocket propellant is entirely spent. The Comsat maneuver increases the service life of the satellite by a number of years, but places the burden of tracking the satellite position on the earth station.

Satellite Footprint

The transmitting and receiving antennas on the satellite are designed to cover only desired regions of the earth's surface. This has several purposes. It concentrates the power radiated from the satellite into desired directions, increases the sensitivity of its receiving antennas, and helps prevent interference with signals from other satellites. The part of the earth's surface covered by a satellite is called the satellite's *footprint*. The footprint may cover one or more relatively localized regions of the earth or almost a complete hemisphere. A typical footprint is shown in Fig. 3. The footprint is, of course, not sharply defined. Signal strengths tend to peak near the center of the footprint and roll off steeply past the 3 dB contour. Global or regional beams are usually shaped such that a particular defined section of the earth's surface is illuminated by the satellite's radiated signal. For example, a United States domestic satellite would probably limit the footprint such that its neighboring countries were illuminated with minimum signals to minimize interference possibilities. This is necessary since several countries share the same portions of the orbital arc.

Polarization

Electromagnetic waves and antennas are always polarized in some manner. The polarization may be linear, circular, or elliptical. For the purposes of discussion, we will dismiss elliptical polarizations as being nonideal cases that are intended to be either linear or circular. Linear polarizations and circular polarizations are aligned in space as in Fig. 4. A linearly polarized antenna receives maximum power from an incident linearly polarized wave if the tilt angles of the wave and the antenna polarizations are aligned in space as in Fig. 4A. The wave is then said to be *co-polarized* or *polarization matched*. As the tilt angle of the wave or antenna rotates from co-polarization, the received power decreases. When the tilt angles are 90° apart as in Fig. 4B, the antenna is cross polarized to the wave and receives no power from it. The antenna and the wave then have orthogonal polarizations. A given wave can have two orthogonal polarizations which exist simultaneously and carry different information without interference. This principle, *frequency reuse*, is used to increase the information capacity of satellites and of the geosynchronous orbit. The term frequency reuse will be explained in more detail below.

Circular polarizations have either right-hand (RHC) or left-hand (LHC) senses. RHC and LHC polarizations are orthogonal. A circularly polarized satellite and a circularly polarized earth station are co-polarized if they have the same senses and are cross-polarized if they have the opposite senses. The relative tilt angles of circular polarized antennas and waves are of no consequence and are not defined. This represents an advantage of circular polarization over linear polarization since the tilt angle of the earth station does have to be adjusted for a particular satellite. On the other hand, circularly polarized antennas tend to cost more than linearly polarized antennas due to the increased complexity of the feed components. Most domestic satellites are linearly polarized while INTELSAT satellites are for the most part circularly polarized.

Satellite System Characteristics

The design of a satellite communication system is an intricate process, involving tradeoffs between many

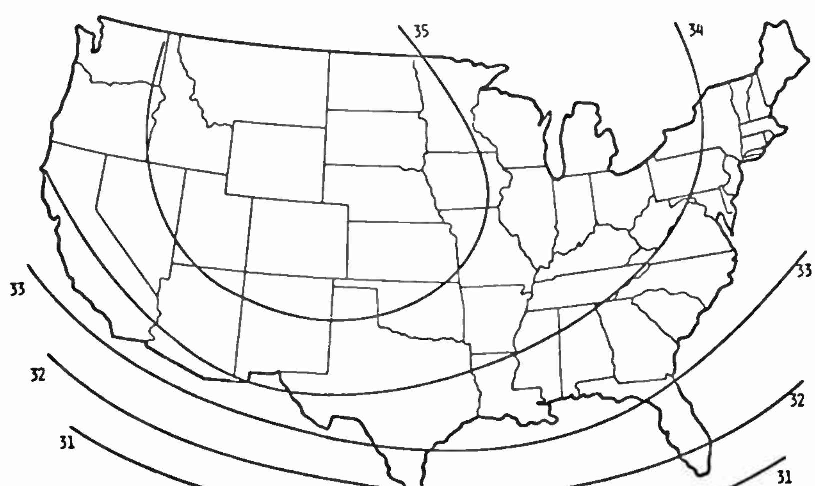


Figure 3. Satellite footprint.

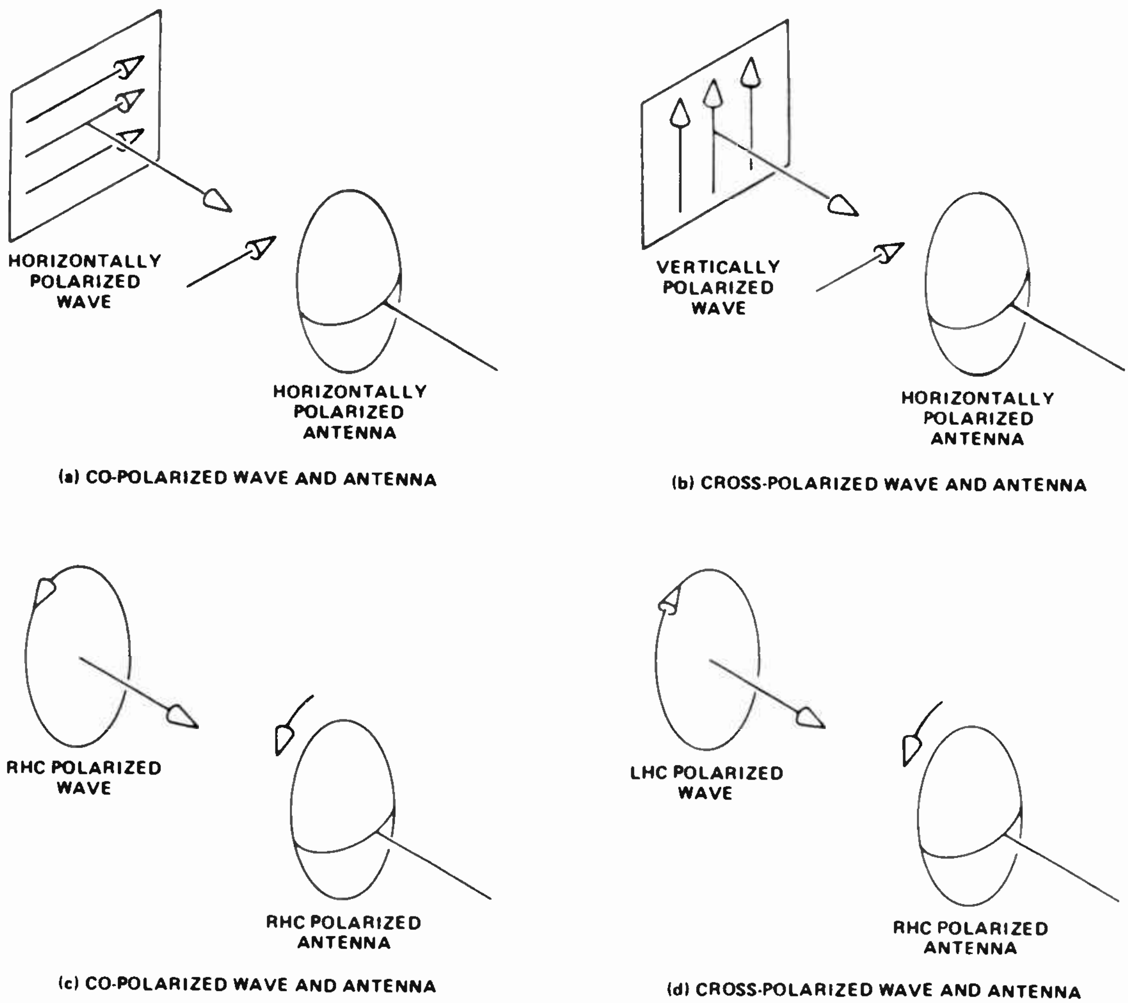


Figure 4. Linear and circular polarizations.

variables to obtain maximum performance at a reasonable cost. The major cost and complexity tradeoff occurs between satellite and earth stations or, more generically, between space segment and ground segment. The dominating design factors in both segments for systems using geostationary satellites are:

Space segment:

1. Weight and size of satellite.
2. DC power generated on board.
3. Dimensions and complexity of satellite antennas.

Ground segment:

1. Allocated frequency bands.
2. Earth stations antenna size and RF capabilities.
3. Earth stations multiple access techniques.

The weight of the satellite is limited by the high cost of launching a spacecraft into geostationary orbit, typically $40,000 to $50,000 per kilogram. For a satellite of limited weight and size, a limited number of solar cells can be deployed which defines an upper limit on the DC power available for the communication transponders. The size and power limitations translate into the fact that the spacecraft has a limited RF output power, which then must be transmitted onto particular areas of the earth, i.e., the continental United States.

Furthermore, power densities over the earth's surface are limited, depending on operating frequency bands, to allow interference-free coexistence with terrestrial systems operating in the same frequencies. The result is the signals arriving from communication satellites are inherently weak, typically -120 to -160 dBW/m^2 and therefore relatively large receiving ground antennas must be utilized.

Multiple access and multiple destinations are distinctive virtues of satellite communications. The methods by which a large number of earth stations share one satellite or one transponder providing the required connectivity (multiple access techniques), have also a significant impact on system design. The multiple access can be achieved by sharing the transponder bandwidth in separate frequency slots (frequency division multiple access [FDMA]), or the transponder availability in discrete time slots (time division multiple access [TDMA]). A third technique, code division multiple access (CDMA) or spread spectrum, shares the transponders by allowing coded signals to overlap in time and frequency.

A satellite communication system must be designed to meet certain minimum performance standards, within limitations of transmitted power, RF bandwidth, and antenna sizes. The most important performance criterion for analog systems is the signal-to-noise ratio (SNR or S/N) in the information channel or baseband. In digital systems the performance measurement criterion is bit-error rate (BER).

SNR and BER depend on a number of factors, such as the predetection carrier-to-noise-density ratio (C/N_0) and the carrier-to-noise ratio (C/N) in the receiver, the type of modulation, and the RF and baseband bandwidths. In the following section the design and analysis of satellite communications links in terms of C/N_0 will be conducted. Therefore, the carrier power received in an earth station receiver and the noise power density in the receiver need to be calculated to establish the operating link C/N_0.

Since satellites are inherently power limited, there will be invariably a modulation technique whereby a trade-off of bandwidth for power will make the baseband S/N larger than the RF C/N in analog systems and will optimize BER in digital systems.

Satellite Transmission Modes For Television

Analog and digital formats are used for the transmission of television signals, but the use of analog frequency modulation is by far the most common. The advantages of FM for satellite transmission are:

1. It minimizes the effects of nonlinearities in the transmission channel.
2. It is immune to AM noise.
3. Power-limited systems can take advantage of the wider bandwidth to increase the C/N.
4. Various processing techniques can be employed to optimize video transmission; i.e., multiplexing, preemphasis, and threshold extension.

The choice of the optimum modulation index m, and the ratio of the FM deviation to the highest modulating frequency f_m, is critical. Bandwidth and modulation index are related by: Bandwidth $= 2(m + 1)f_m$. The spectral distribution of a FM signal as a function of m is shown in Fig. 5. The selection of the optimum frequency deviation, Δf, must be based on the number of channels to be transmitted, the type of baseband signal (i.e., component or composite), the signal quality requirement, the power received, and the available bandwidth. A typical value of peak for a C-band

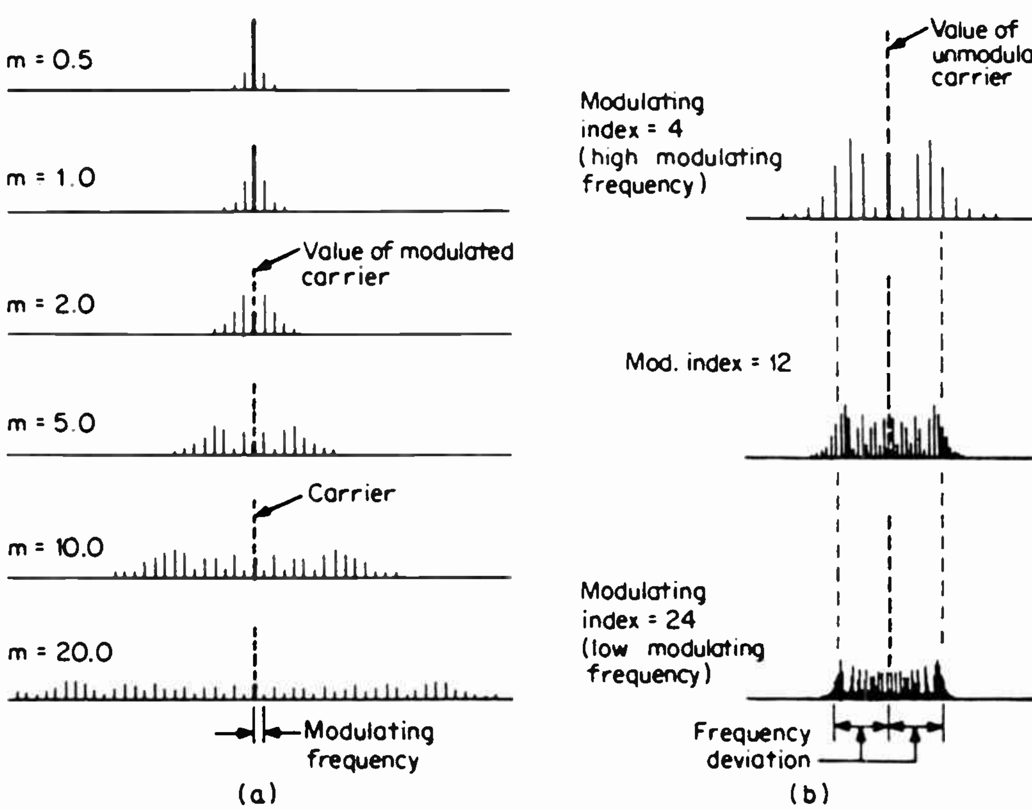


Figure 5. Spectra of frequency modulated signals *(a)* frequency spectra with increasing frequency deviation and constant modulating frequency; *(b)* frequency spectra with constant frequency deviation.

satellite transponder with a nominal bandwidth of 36 MHz is 10.75 MHz.

For satellite transmission, in addition to the deviation of the carrier by the signal, it is usually necessary to subject the main carrier to a low-frequency deviation. This spreads the high concentration of carrier and sideband energy over a larger range of the spectrum and permits higher satellite effective radiated power (ERP) without exceeding the FCC's limit on watts/meter^2/kHz downlink power density.

Preemphasis/deemphasis is employed in FM systems for the transmission of video to compensate for the increase in thermal noise with increasing frequency. CCIR Recommendation 567 specifies a standard 75 microsecond preemphasis, for example.

A characteristic of FM is that the detected signal-to-noise ratio (S/N) for the video signal is higher than its C/N ratio. This difference is the FM improvement factor; satellite transmission takes advantage of this improvement, provided the received C/N is greater than the receiver operating threshold.

Threshold extension demodulation (TED) is a common technique used in FM receivers to reduce video impulse noise when the carrier-to-noise ratio (C/N) drops below the receiver's operating threshold. Above threshold, the receiver acts like a standard discriminator; when C/N drops below threshold, TED circuitry automatically switches to a narrow bandwidth.

Digital Transmission

Satellite systems are used for the transmission of digital information. A typical digital transmission uplink consists of a modem, upconverter and power amplifier. The modem converts digital information to and from a modulated-carrier. The center-frequency of the modulated-carrier positions the signal within a satellite transponder. The upconverter converts the modulated-carrier to a satellite frequency and thus selects the transponder of the satellite.

The earth station component used to convert digital information to a format suitable for transmission by satellite is referred to as a modem. The modem accepts a digital data input signal and outputs an intermediate frequency, typically a range centered on either 70 MHz or 140 MHz, containing the modulated digital information.

A modem for use in satellite systems is similar to modems used for telephone circuits. However, satellite modems generally operate at much higher bit-rates and, in addition, contain special features specifically for satellite link use.

Transmitted digital data is first applied to the encoder section of the modem for forward-error-correction (FEC) encoding. This process appends additional bits to the original information to provide error-detection and correction. The signal is scrambled using a standard algorithm to ensure random data.

The aggregate data (i.e., original data plus error-correction bits) is applied to the modulator for frequency modulation onto an IF-carrier. The IF-carrier is selectable, typically in the range of 50 MHz to 90 MHz for 70 MHz operation, or 100 MHz to 180 MHz for 140 MHz operation. The center frequency of the modem modulator is tuned to position the signal within the satellite transponder.

A satellite digital transmission system is characterized by data interface, data rate, code rate, and modulation scheme. Data rate refers to the number of bits per second transmitted by the modem. The data rate is typically front-panel selectable. Typical ranges of rates are from 32 kbps to 3 mbps selectable in increments as small as 1 bps. Modems typically support a number of data interfaces. The data interface refers to the connector and signal levels. Typical data interfaces are DS1, CEPT, EIA-422, V.35 and MIL-188/114. Code rate refers to the FEC-encoding scheme. In some modems, the code rate is selectable. The code rate configuration is referred to as *m/n*. *m* refers to the number of information bits per block of transmitted bits. *n* refers to the number of information bits plus error-correction bits per block of transmitted bits. Thus, a code rate of ¾ means that for every three information bits, four bits are transmitted. Thus a 1024 kbps modem operating with a code rate of ¾ would transmit 1365 kbps over the satellite channel.

The modulation scheme refers to the method of indicating data bits. Two common modulation schemes employed in satellite transmission systems are bi-phase shift key (BPSK) and quadrature-phase shift key (QPSK). These modulation schemes generate a periodic set of phase-shifts referred to as symbols. The symbol rate (i.e., the number of symbols per second) and modulation scheme determine the amount of bandwidth required in the channel.

In BPSK, two phase-shifts are used representing two states. For this case the symbol rate is equal to the transmission rate. The QPSK scheme uses four phase-shifts thus transmitting two bits per symbol. Thus, QPSK uses a symbol rate that is half the transmission rate. QPSK requires less bandwidth than BPSK, but requires increased performance from the channel and more complicated modulators and demodulators.

C-Band Satellites

C-band was initially favored for communications satellites because of the favorable propagation characteristics for these frequencies. The specific bands in most common use are the 5925 MHz to 6425 MHz (uplink) and the 3700 MHz to 4200 MHZ (downlink) band pair. U.S. domestic FSS requires the use of 36 MHz bandwidth channels placed on 40 MHz centers. A satellite using a single polarization can provide 12 such transponders, although all new satellites are mandated to be capable of frequency reuse and provide 24 such transponders. Frequency reuse is implemented by the use of orthogonal polarizations and by staggering the microwave carriers of alternate transponders. As an example of a typical satellite, the transmit and receive frequency plans of a GE/RCA Satcom satellite are shown in Fig. 6. The numbered brackets represent

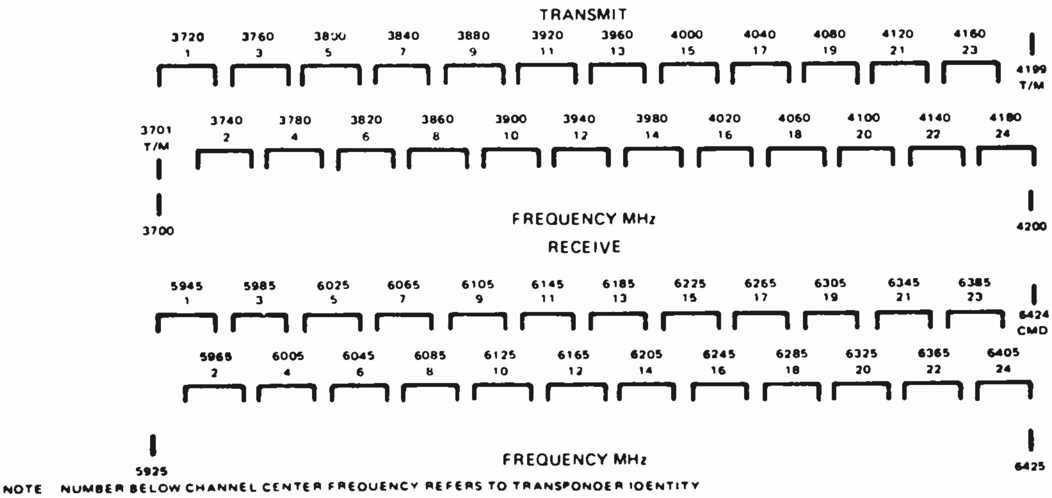


Figure 6. GE/RCA Satcom satellite frequency plans.

each channel. The bandwidth of the channel is represented by the width of the bracket. The carrier frequency, shown above the channel number, is centered on each channel. The signals of alternate transponders in the frequency plan of Fig. 6 are nominally orthogonal. If they were exactly orthogonal and the associated earth stations were ideal (with respect to polarization), there would be no interference caused by the overlapping sideband energy of adjacent transponders. In practice, the polarizations of the antennas of the satellite and earth stations are not ideal. Some small amount of interference occurs, but the combination of nearly orthogonal polarizations and the use of the staggered frequency plan provides for high quality transmission under almost all weather conditions.

Ku-Band Satellites

The first systems using the 14.0 GHz to 14.5 GHz (uplink) band and the 11.7 GHz to 12.2 GHz (downlink) band were launched in 1976 by SBS. The higher propagation losses characteristic of these frequencies require higher spacecraft equivalent isotropic radiated power (EIRP) to achieve the same transmission performance as C-band frequencies, and this is obtained from the use of greater spacecraft antenna gains. These are readily achievable at the higher frequencies. Since the Ku-band frequencies are not shared with terrestrial systems, the power flux density (PFD) limitation is less stringent and there is no requirement for coordination with terrestrial microwave systems. The high powers permit the use of very small earth station antennas at or near the user's premises. This results in important economic advantage for many services and makes the use of this frequency band very attractive. Even so, a good part of the higher satellite power achievable is necessary to offset the additional attenuation that is experienced at these frequencies during heavy rain conditions.

There is no mandated frequency plan for transponders in this frequency band, although typical transponder bandwidths are 36 MHz to 72 MHz. Since the bandwidth is the same as C-band it is possible to have a similar 24-transponder, 36 MHz frequency plan with 40 MHz channel spacings when frequency reuse is utilized.

Two of the more important differences between C-band and Ku-band are the following:

1. C-band FSS service shares frequencies with terrestrial microwave systems. This places constraints on the location of C-band earth stations, and it limits the permissible downlink power density.
2. Ku-band signals are subject to significant attenuation in heavy rainfall.

The advantages and disadvantages of C-band and Ku-band which result from these and other differences are summarized in Table 1. Merits of C- and Ku-band for satellite communications.

Regulatory Issues

Satellite communication systems are governed by the Federal Communications Commission in the United States and by the International Telecommunications Union (ITU) on the international level. The governing agencies assign frequency bands of operation, satellite performance characteristics and orbit location, and provide technical specifications of radiated power density and radiation gain patterns for the earth stations. The FCC is the licensing body for all transmit earth stations in the United States and licenses C-band receive-only earth stations at the owner's request. The FCC Rules and Regulations, Part 25, form the basis of the applicable documents which must be followed for the planning and implementation of any FSS band satellite communication system.

The FCC amends and interprets the rules as the

TABLE 1

Merits of C- and Ku-band for satellite communications.

C-band Advantages

1. Less susceptible to rain outages.
2. Established manufacturing infrastructure.
3. Antenna surface tolerance can be achieved by various techniques that lend themselves to low cost manufacturing.

C-band Disadvantages

1. Frequency band is congested because it is shared with terrestrial microwave, making frequency coordination a requirement.
2. Requires relatively large antennas because of low satellite EIRP levels and the necessity of narrow half-power beamwidth to allow two degree spaced satellites.
3. Avoiding terrestrial interference can make site selection a difficult process.
4. The use of artificial shielding to block interference can increase total system cost.
5. Faraday rotation of polarization can affect system performance.
6. Satellite dispersal signal is required to prevent harmful interference to terrestrial stations, resulting in more stringent video receiver clamping specifications.

Ku-band Advantages

1. Frequency band is only used for satellite communication.
2. Smaller antennas may be used because of higher gain and higher satellite EIRP.
3. Easier site selection because of smaller size of antenna and lack of terrestrial interference.
4. Narrower antenna beamwidth is desirable in reduced orbital spacing.
5. Suitable for direct-to-home application.
6. Lower reception equipment cost.
7. Flexibility in channelization plan.
8. Not affected by Faraday rotation.
9. No satellite dispersal signal disadvantages.

Ku-band Disadvantages

1. Affected by rain attenuation and depolarization.
2. Narrow beamwidths of antennas may require more rigid mounts.
3. Reflector surface tolerance requirements restrict manufacturing techniques and increase cost.
4. Waveguide and coaxial transmission line losses are quite high.
5. High noise temperature of low-noise amplifiers may cause the use of large antennas to achieve desired G/T antenna gain-to-noise temperature).

technology and the requirements of satellites changes through amendments, decisions, and declaratory orders; therefore, it is recommended that the FCC be contacted at the time of system planning to obtain the latest Rules and Regulations. Some of the more important parts of the regulations for earth station antennas are discussed below. The FCC established precedents for the minimum diameter apertures and sidelobe gain envelopes for earth station antennas operating in the fixed-satellite service bands at the beginning of these services in the early 1970s to minimize interference between terrestrial systems and satellite systems and between satellite systems. These precedents have been modified through the years as the use of satellite services has increased. The more significant recent rulings pertaining to earth station antenna performance have resulted in improved antenna radiation patterns in the close-in sidelobe region and have established maximum radiated power densities for antennas less than 9 meters in diameter for C-band operation and 5 meters in diameter for Ku-band operation.

The FCC Rules and Regulations Section 25.209 pertaining to antenna gain envelopes is mandatory for all transmit antennas. The standard is as follows:

A. The gain of any antenna to be employed in transmission from an earth station in the fixed-satellite service shall lie below the envelope defined below:

1. In the plane of the geostationary satellite orbit as it appears at the particular earth station location:

$[29 - 25 \times \log(\Theta)]$dBi	$1° < \Theta < 7°$
$+ 8$ dBi	$7° < \Theta < 9.2°$
$[32 - 25 \times \log(\Theta)]$ dBi	$9.2° < \Theta < 48°$
$- 10$ dBi	$48° < \Theta < 180°$

where Θ is the angle in degrees from the axis of the main lobe, and dBi refers to the dB relative to an isotropic radiator. For the purposes of this section, the peak gain of an individual sidelobe may not exceed the envelope defined above for Θ between 1° and 7°. For Θ greater than 7°, the envelope may be exceeded by 10% of the sidelobes, but no individual sidelobe may exceed the envelope by more than 3 dB.

2. In all other directions:

Outside the main beam, the gain of the antenna shall lie below the envelope defined by:

$[32 - 25 \times \log(\Theta)]$ dBi	$1° < \Theta < 48°$
$- 10$ dBi	$48° < \Theta < 180°$

where Θ is the angle in degrees from the axis of the main beam, and dBi refers to dB relative to an isotropic radiator. For the purpose of this section, the peak gain of an individual sidelobe may be reduced by averaging its peak level with the peaks of the nearest sidelobes on either side, or with the peaks of the two nearest sidelobes on either side, provide that the level of no individual sidelobe exceeds the gain envelope given above by more than 6 dB.

B. The off-axis cross-polarization isolation of any antenna to be employed in transmission at frequencies between 5925 MHz and 6425 MHz from an earth station to a space station in the domestic fixed-satellite service shall be defined by:

$[19 - 25 \times \log(\Theta)]$ dBi	$1.8° < \Theta < 7°$
-2 dBi	$7° < \Theta < 9.2°$

C. Any antenna licensed for reception of radio transmission from a space station in the fixed-satellite service shall be protected from radio interference caused by other space stations; only to the degree to which harmful interference would not be expected to be caused to an earth station employing an antenna conforming to the standards defined in paragraphs A and B of this section.

D. The standards specified in paragraphs A and B of this section shall apply to all new antennas after July 1, 1984 and to all antennas after January 1, 1987.

E. The operations of any earth station with an antenna not conforming to the standards of paragraph A and B of this section shall impose no limitations upon the operation, location, and design of any terrestrial station, any other earth station, or any space station.

The FCC further acknowledged within the text of the Federal Register Statement 47 CFR Part 25, Vol. 48, No. 173, September 6, 1983, that the envelope defined above is only a reference envelope in the receive band. Receiving antennas do not have to conform to this envelope to be eligible for licensing. Facilities with performance worse than the reference envelope must, of course, accept correspondingly potential, higher interference levels. The interference levels should be calculated based on typical measured radiation patterns, site location, and for a desired satellite or satellites. This analysis may result in acceptable receive-only carrier-to-interference performance for antennas meeting $[32 - 25\log(\theta)]$ envelope even with orbital spacings as small as 2°, since discrimination, that is, peak on-axis gain to sidelobe gain, is the important determining factor, not an arbitrary sidelobe gain performance envelope relative to isotropic.

FCC License

The FCC *requires* licensing of transmitting earth stations and *permits* licensing of receive-only (RO) earth stations. It is desirable for a broadcaster to license a C-band RO earth station, since licensing protects the station from future interference from domestic microwave systems.

The application for a C-band RO terminal is relatively simple. It is a narrative (there is no standard form) with the following information:

1. The nature of the proposed service.
2. A statement of public interest.
3. The name of the person to receive correspondence (usually includes a technical contact and counsel).
4. The applicant's technical qualifications.
5. A paragraph identifying that the application is a minor action for environmental considerations. Generally, the action is minor if the antenna is less than 30 feet (9.144 meters) in diameter and is not located in certain specific areas (e.g., historic districts, scenic areas, floodplain, and some others listed in 47 CFR1.1305(a)(6)).
6. Description of the site (plot plan) and its availability. This should include the elevation of the antenna base above sea level and the maximum antenna height above the base.
7. A functional block diagram.
8. Description of the antenna mounting and range.
9. A statement of compliance with the requirements for antenna sidelobe performance.
10. A statement of FAA compliance, if required (the same rules apply as for any antenna).
11. The report of frequency coordination and interference analysis.
12. A technical showing that the antenna meets the performance criteria for a 4.5 meter antenna, if the antenna is less than 4.5 meters.
13. A completed FCC Form 403.
14. A technical summary of the proposed station including a description of the equipment to be used, major structures, the type of communication being received, frequencies and polarizations received, satellites, and transponders being used.
15. A statement that the applicant will only receive program material for which it has obtained rights.
16. The applicant's certification.
17. The technical certification by a technically qualified person that the engineering data are complete and accurate.

These same items are required for transmitting stations, although naturally some of the items will be more involved (such as frequency coordination and technical description of the system). There are a few additional requirements for transmitting stations:

1. A completed FCC Form 430 (legal and financial qualification). This form is required if common carrier status is sought.
2. A statement of compliance with regulations for the protection of employees and the general public against excessive radiation.
3. A statement of compliance with FCC regulations regarding quiet zones.

SYSTEM PERFORMANCE ANALYSIS TECHNIQUES

Considering an RF link as illustrated in Fig. 7 with transmit power P_t and transmit gain G_t, the effective

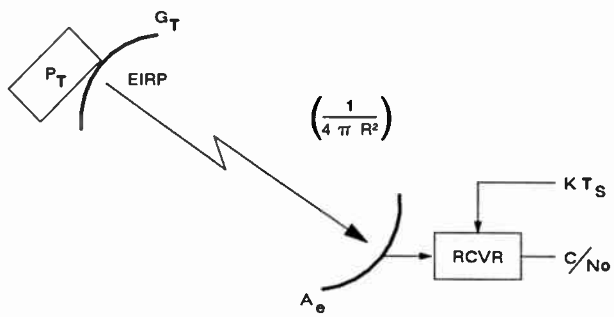


Figure 7. RF link diagram.

isotropic radiated power (EIRP) for the station along the main beam of the antenna is the product $G_t \times P_t$. At a distance R meters from the transmitter, the radiated flux density, S, becomes:

$$S = (G_t P_t) \frac{1}{4\pi R^2} k_a \quad \text{watts/m}^2 \qquad (1)$$

where: k_a = Atmospheric attenuation factor < 1

If an antenna with an effective area in square meters, A_e, is receiving this flux density, the received carrier level, C, at the antenna output is:

$$C = SA_e = (P_t G_t A_e) \frac{1}{4\pi R^2} k_a \quad \text{watts} \qquad (2)$$

At the same antenna output point, the effective noise power density, N_o is given by:

$$N_o = kT_s \quad \text{watts/hertz} \qquad (3)$$

where: k = Boltzmann's constant = 1.38×10^{-23} joules/K or −228.6 dB.
T_s = System noise temperature
K = Temperature in degrees Kelvin

We can assume the atmospheric attenuation factor to be unity with very small error, therefore, the carrier-to-noise density ratio C/N_o can be expressed by:

$$\frac{C}{N_o} = (P_t G_t A_e) \left(\frac{1}{4\pi R^2}\right) \left(\frac{1}{k\,T_s}\right) \qquad (4)$$

A fundamental relationship in antenna theory is that the gain, G_r, and the effective area of an antenna, A_e, are related by:

Substituting this relation into the expression for C/N_o,

$$A_e = G_r \left(\frac{\lambda^2}{4\pi}\right) \qquad (5)$$

$$\frac{C}{N_o} = (P_t G_t G_r) \left(\frac{\lambda}{4\pi R}\right)^2 \frac{1}{kT_s} \qquad (6)$$

OR

$$\frac{C}{N_o} = \text{EIRP} \frac{G_r}{T_s} \left(\frac{\lambda}{4\pi R}\right)^2 \frac{1}{k} \qquad (7)$$

The factor $(\lambda/4\pi R)^2$ is often inverted and defined as the spreading loss or space loss factor. This spreading loss can also be expressed as:

$$L_s = (4\pi Rf/c)^2 \qquad (8)$$

where: c = Speed of light = $3 \times 10^{+8}$ meter/second
f = Frequency in hertz

Link calculations are usually carried out in dB rather than directly from the above relations because of ease of working in common logarithms. C/N_o in dB can be calculated by:

$$(C/N_o)\,\text{dB} = 10 \log (C/N_o) \qquad (9)$$
$$(C/N_o)\,\text{dB} = \text{EIRP} - L_s + (G/T) + 228.6$$

where: EIRP = $10 \log (G_t P_t)$ dBW

$$L_s = 20 \log (4\pi Rf/c) \quad \text{dB} \qquad (10)$$
$$= 92.45 + 20 \log R\,(\text{km}) + 20 \log f\,(\text{GHz})$$

$$(G/T) = 10 \log (G_r/T_s) \quad \text{dB/K} \qquad (11)$$

Alternately, C/N_o can be expressed in terms of flux density, S, as:

$$(C/N_o)\text{dB} = S + (G/T) - A_i + 228.6 \quad \text{dBHz} \qquad (12)$$
$$S = \text{EIRP} - L_s + A_i \quad \text{dBW/m}^2 \qquad (13)$$

where A_i is the effective aperture of an isotropic radiator in dB:

$$A_i = 10 \log (4\pi/\lambda^2) \qquad (14)$$

Eq. 9 is a fundamental tool for characterizing space link performance. It will be utilized later when calculating overall satellite link performance.

Earth Station Receive Figure-Of-Merit G/T

G/T is the figure-of-merit of a receive system. It is primarily a function of the gain of the antenna along with the antenna noise temperature, first amplifier noise temperature and losses between the antenna and the first amplifier. The importance of the term G/T in Eqs. 9 and 12 cannot be overstated. Examination of the C/N_o expression shows that for a given available transmitting power and information format (and thus bandwidth), the only available method of controlling the received signal quality that can be used by the downlink operator is through the system G/T. Note that the G/T provides a direct dB relationship with C/N_o.

Fig. 8 shows a block diagram of a typical receive system. Each device in the RF path has an associated gain or loss and a noise temperature. These contributions are combined to reflect the noise power weighted by the gain distribution through the chain. The earth station G/T is given by:

$$(G/T) = G_a - 10 \log (T_s) \quad \text{dB/K} \qquad (15)$$

where: G_a = Antenna gain referenced to LNA input (dBi)
T_s = System noise temperature referenced to LNA input (K)

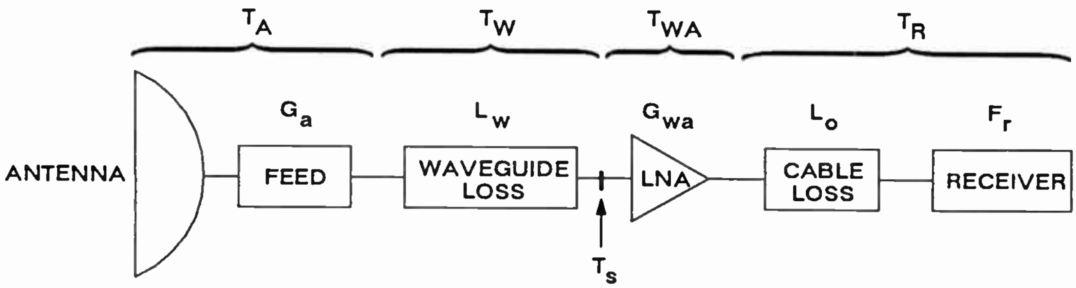


Figure 8. G/T system diagram: receive-only earth station.

The system noise temperature, T_s, referenced to the LNA input can be calculated by adding as noise powers the equivalent noise temperatures of all noise contributors, weighted by the net gain between the point in which that noise is being added and the LNA input, that is:

$$T_s = (T_a/L_w) + T_0(L_w - 1)/L_w + T_{vswr} + T_{lna} + [(L_t - 1) + L_t(F_r - 1)]\, T_l/G \quad (16)$$

where: T_a = Antenna noise temperature (K)
L_w = Waveguide loss between antenna and LNA (linear power ratio)
L_t = Transmission loss between LNA and receiver
T_o = Ambient temperature (K)
T_l = 290° K
T_{lna} = LNA noise temperature (K)
T_{vswr} = LNA-Antenna impedance mismatch noise temperature (K)
F_r = Receiver noise fig. (linear power ratio)
G = Net gain between LNA input and receiver input (includes interconnect cable loss)

The antenna temperature is usually minimum at zenith, typically 15° to 25° for a low-loss, C-band antenna with low wide-angle sidelobes. As the elevation angle decreases, the antenna temperature increases because more of the higher-level sidelobes look at the earth, which has a temperature of about 290K. A typical curve of the variation of noise temperature with elevation angle is illustrated in Fig. 9. Similarly, Figs. 10A and 10B show typical G/T system performance for different antenna diameters as a function of elevation angle for C-band and Ku-band, respectively.

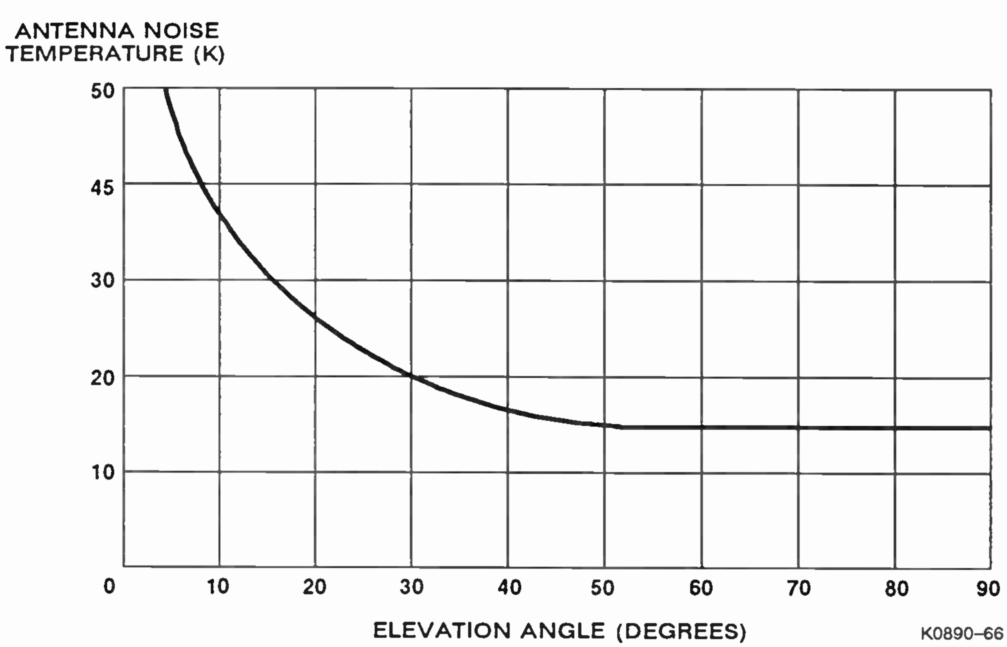


Figure 9. Typical antenna noise temperature variations with elevation angle.

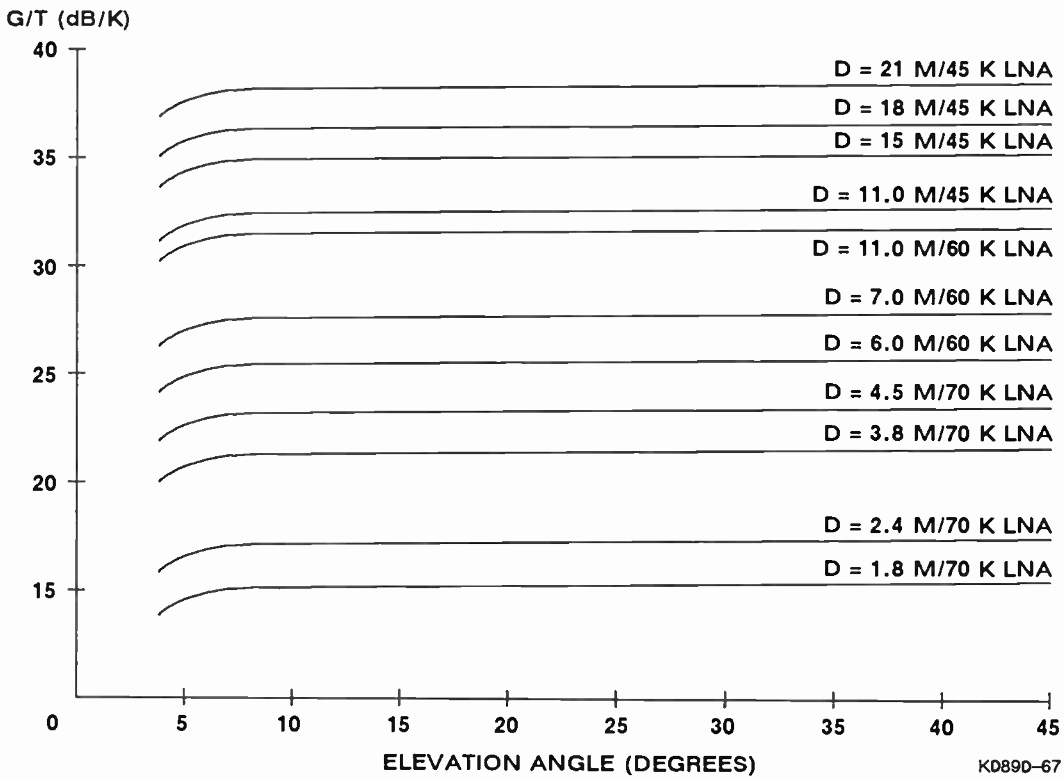


Figure 10A. Typical C-band G/T system performance versus elevation angle for different commonly used antenna diameters.

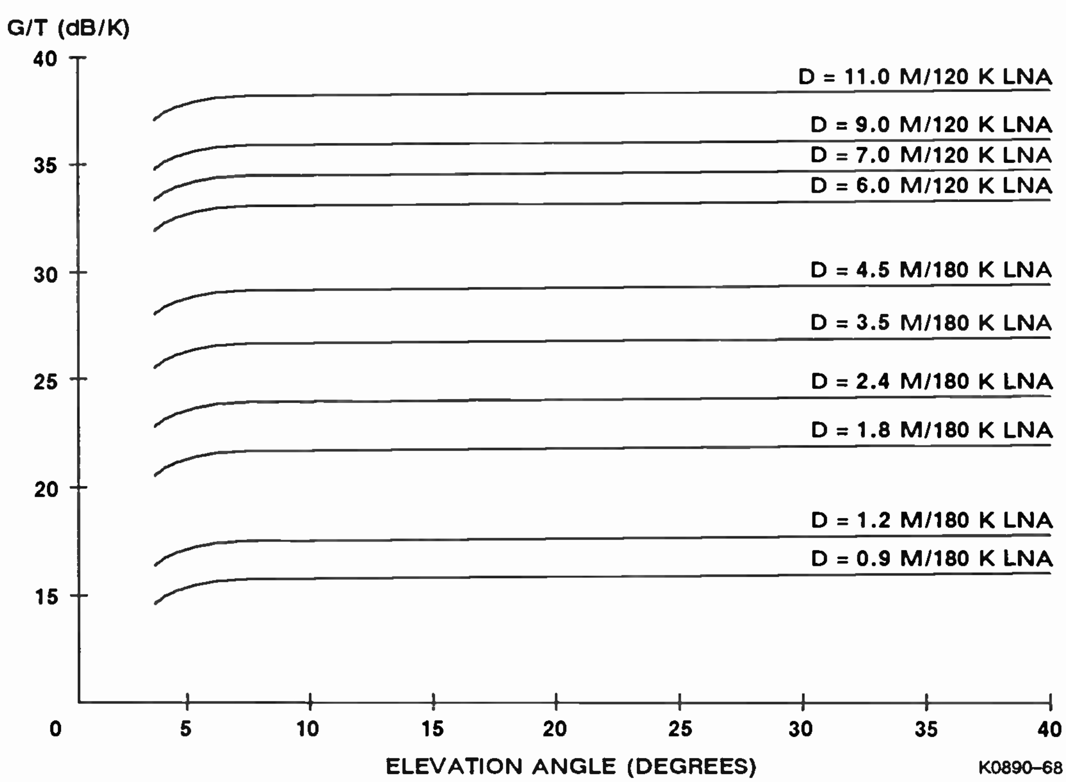


Figure 10B. Typical Ku-band G/T system performance versus elevation angle for different commonly used antenna diameters.

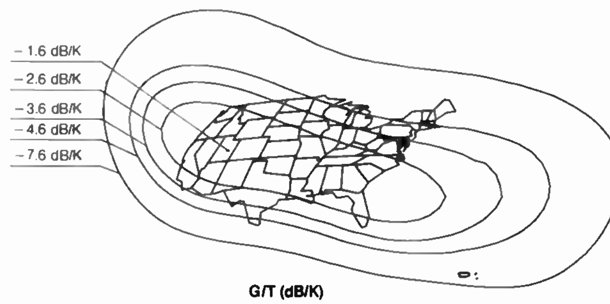
Satellite Transponder

The orbiting spacecraft provides a one-hop carrier relay over a wide geographic area. In C-band systems the uplink signal is transmitted near 6 GHz, received by the satellite, amplified, translated in frequency, filtered and retransmitted near 4 GHz. Likewise, in Ku-band systems the uplink occurs in the 14 GHz range and the downlink in the 12 GHz.

Since the satellite serves as a transmit/receive station, it must be characterized by a G/T for the uplink side and by saturated EIRP for the downlink side. To couple the uplink and downlink signal strengths, and as a definition of the transponder sensitivity, the uplink RF flux density required at the satellite to saturate the transponder is also specified (SFD). These three satellite parameters also vary with geographic location. Contour maps called footprints are usually available for assessing these variations. Typical footprints for C-band and Ku-band satellites are shown in Figs. 11 and 12, respectively.

Another important consideration to characterize the transponder performance is the input/output power transfer and intermodulation response. Both performance parameters are normally specified in terms of input (BO_i) and output (BO_o) back-off, that is, as a function of the power reduction expressed in dB with respect to saturation. Figs. 13 and 14 show typical transponder response for a satellite equipped with traveling wave tube (TWT) power amplifier.

SPACENET II (69° WL)
Expected 36 MHz C-Band G/T Performance



	G/T (dB/K)
ATLANTA	−1.9
BOSTON	−4.8
CHICAGO	−3.5
DALLAS	−2.2
HOUSTON	−3.1
LOS ANGELES	−3.5
NEW YORK	−4.4
ORLANDO	−2.2
SAN FRANCISCO	−2.5
SAN JUAN	−6.8
SEATTLE	−2.4

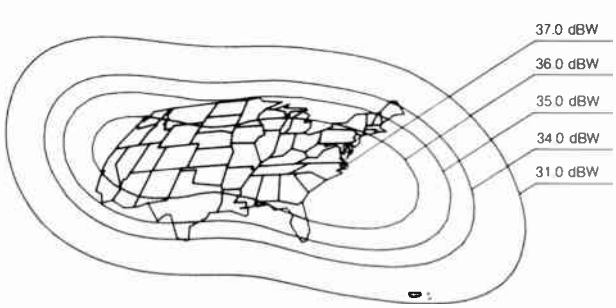
Note SPACENET 36 MHz C Band transponder saturation flux density (SFD) is ground commandable to a nominal value of 86 or 80 dBW/m² corresponding to a G/T level of 5 dB/K

Figure 11B. C-band satellite G/T footprint.

Satellite Link Analysis

With the preliminary procedures and formulations described above, link calculations can be conducted. First, the distance or slant range from the satellite to the earth station needs to be determined so the space

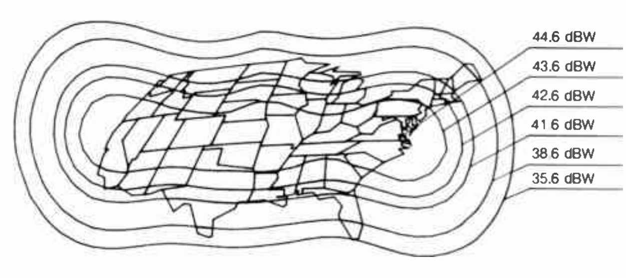
SPACENET II (69° WL)
Expected 36 MHz C-Band EIRP Performance



	EIRP (dBW)
ATLANTA	36.7
BOSTON	35.4
CHICAGO	36.4
DALLAS	36.0
HOUSTON	35.3
LOS ANGELES	35.3
NEW YORK	36.1
ORLANDO	35.9
SAN FRANCISCO	35.9
SAN JUAN	31.6
SEATTLE	36.0

Figure 11A. C-band satellite EIRP footprint.

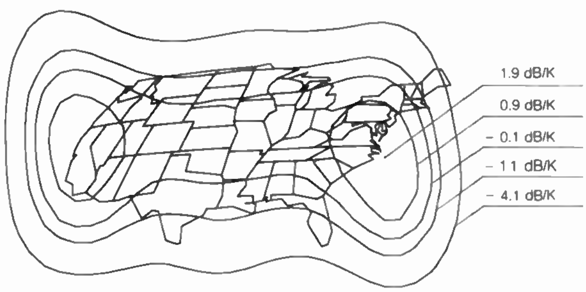
SPACENET II (69° WL)
Expected 72 MHz Ku-Band EIRP Performance



	EIRP (dBW)
ATLANTA	43.2
BOSTON	41.7
CHICAGO	42.9
DALLAS	43.2
HOUSTON	41.3
LOS ANGELES	43.0
NEW YORK	43.4
ORLANDO	39.8
SAN FRANCISCO	44.1
SEATTLE	42.0

Figure 12A. Ku-band satellite EIRP footprint.

SPACENET II (69° WL)
Expected 72 MHz Ku-Band G/T Performance



	G/T (dB/K)
ATLANTA	– .1
BOSTON	– 1.1
CHICAGO	.2
DALLAS	– .2
HOUSTON	– 1.0
LOS ANGELES	1.2
NEW YORK	.7
ORLANDO	– 1.4
SAN FRANCISCO	1.5
SEATTLE	– .3

Note: SPACENET Ku Band transponder saturation flux density (SFD) is ground commandable to a nominal value of –86, –80 or –74 dBW/m² corresponding to a [illegible] level of [illegible] 2 dB/K.

Figure 12B. Ku-band satellite G/T footprint.

loss is calculated. From orbit geometry and Eq. 10 above, the space loss expressed in dB is found to be:

$$L_s = 185.05 + 10\log[1 - 0.295\cos(H)\cos(AL)] + 20\log f \qquad (17)$$

where: H = Latitude of earth station
AL = Difference in longitude for earth station and satellite
f = Frequency in GHz

The overall satellite link can now be calculated.

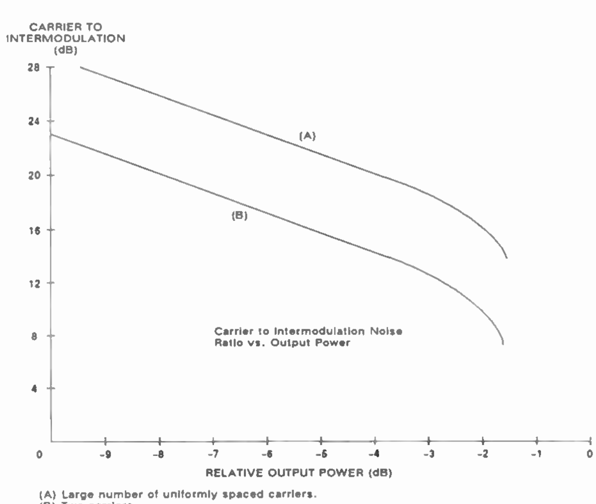


Figure 13. Output power normalized to single carrier saturation point.

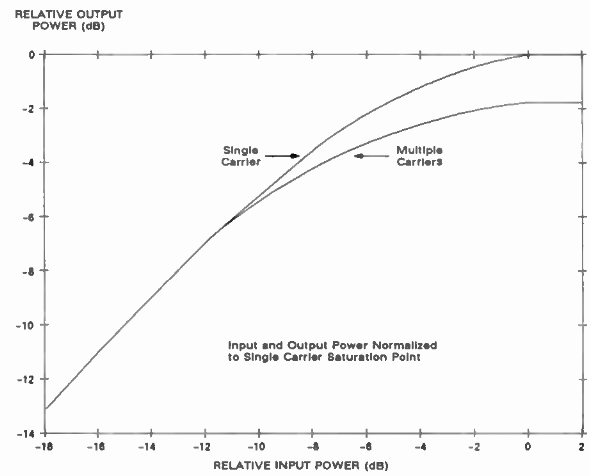


Figure 14. TWT power transfer characteristics.

Uplink C/N

From Eq. 9, the uplink $(C/N_0)_u$ becomes:

$$(C/N_0)_u = EIRP_u - L_u + (G/T)_s + 228.6 \qquad (18)$$

or

$$(C/N_0)_u = S - A_i + (G/T)_s + 228.6 \qquad (19)$$

and

$$S = SFD - BO_i \qquad (20)$$

where: S = Flux density (dBW/m^2)
L_u = Uplink space loss (dB)
$EIRP_u$ = Uplink EIRP (dBW)
$(G/T)_s$ = Satellite G/T (dB/K)
A_i = 21.5 + 20log f (GHz) (dB/m^2)
SFD = Saturation flux density (dBW/m^2)
BO_i = Transponder input Back-Off (dB)

Downlink C/N

Likewise the downlink $(C/N_0)_d$ can be calculated by:

$$(C/N_0)_d = EIRP_d - L_d + (G/T)_{e.s} + 228.6 \qquad (21)$$

and

$$EIRP_d = EIRP_s - BO_o \qquad (22)$$

where: $EIRP_d$ = Downlink EIRP (dBW)
$EIRP_s$ = Saturated EIRP (dBW)
L_d = Downlink space loss (dB)
$(G/T)_{e.s}$ = Earth station G/T (dB/K)
BO_o = Transponder output Back-Off (dB)

It is important to note that Eqs. 20 and 22 are related by the nonlinear power transfer function of the transponder, therefore, for transponder operation below saturation the input and output relationship needs to be resolved graphically with the aid of Fig. 14 or its equivalent.

Once uplink and downlink noise contributions are determined, the composite link performance, in terms of total carrier-to-noise-density ratio $(C/N_0)_t$, can be readily obtained by simple noise power addition since the uplink and downlink contributions are incoherent. This yields:

$$(C/N_0)_t = \{(C/N_0)_u^{-1} + (C/N_0)_d^{-1}\}^{-1} \qquad (23)$$

This equation represents a simplified situation in that only thermal noise is added to the carriers. In actuality there are other sources of perturbations and interference. Transponder nonlinearity is the cause of some of the more important ones. As shown in Fig. 14, operating the transponder near maximum power for better efficiency implies that compression, due to the instantaneous nonlinear transfer characteristic of the amplifier relative to the signal level, becomes more significant. Under this condition, when more than one frequency is amplified, interaction between the signals occur and consequently a spectrum of spurious frequencies or intermodulation products are generated.

Particularly the so called third-order intermodulation product of the form $(2f_1\text{-}f_2)$, a consequence of the third-order nonlinearity of the transponder, constitutes a significant interfering signal, since it is the largest product and it falls in the same operating bandwidth as the information signal. Fig. 13 shows how carrier-to-intermodulation ratio varies as a very sensitive function of the transponder operating output back-off.

System C/N

Fig. 15 depicts a complete satellite link. Other sources of interference have been added, such as uplink interference due to off-beam radiation from other earth stations and uplink cross-polarization isolation, combined, represented by the quantity $(C/I_0)_u$. Similarly in the downlink the quantity $(C/I_0)_d$ represents the combined effects of downlink cross-polarization isolation and adjacent satellite interference. When all these terms are considered the total link $(C/N_0)_t$ can be calculated by:

$$(C/N_0)_t = \{(C/N_0)_u^{-1} + (C/I_0)_u^{-1} + (C/N_0)_d^{-1} + (C/I_0)_d^{-1}\}^{-1} \qquad (24)$$

Fig. 16 shows the typical interaction of the different terms in Eq. 24 as a function of transponder input back-off and in the presence of thermal and intermodulation

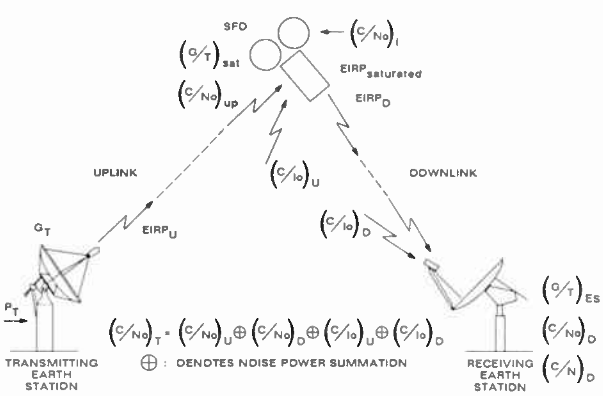


Figure 15. Satellite link model.

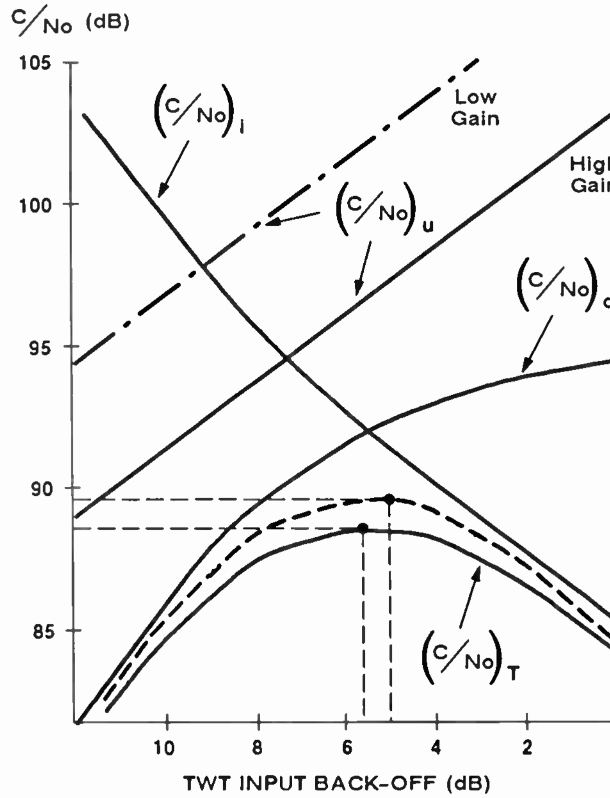


Figure 16. Optimum TWT operation.

noise. The total $(C/N_0)_t$ can be maximized by reducing the transponder input drive and adjusting the transponder gain. Backing off the TWT reduces $(C/N_0)_u$ and also $(C/N_0)_d$ (through the input/output relationship of the transponder), but as $(C/N_0)_i$ increases rapidly when the input drive is reduced, an optimum value of $(C/N_0)_t$ is obtained at a specific back-off level. Interference noise can be kept down by proper antenna design, transponder sensitivity, and frequency coordination of satellite services.

Rain Affects On System Performance

Rain is the dominant factor in satellite propagation for frequencies above 10 GHz . Rain propagation has been studied intensively[10] since the late 1960s, and only a brief discussion will be presented here. Due to the basic interaction of electromagnetic waves with water in liquid form, raindrops cause absorption, scattering, and depolarization phenomena. Absorption and scattering result in signal attenuation and an increase in sky noise temperature, with the consequent degradation of the received C/N_0. Depolarization has an effect on dual polarization systems of creating interference between cross-polarized signals.

Signal Attenuation

The amount of attenuation depends fundamentally on the rain intensity or rain rate and the path length in rain. Rainfall data are available for most parts of the world; different types of climates have been defined and boundaries of their regions identified. Figs. 17 through 19 show the NASA rain rate climate regions. The long term behavior of rain is described by the cumulative probability distribution or exceedence curve. This gives the percentage of time that the rain rate exceeds a given value. Table 2 gives the rain-rate distribution values versus percent-of-the-year for the various rain climate regions of Fig. 17 through Fig. 20. Figs. 20A and B plot the rain rate cumulative probability distributions for the regions presented on the previous maps.

The calculation of the rain attenuation involves two basic steps. The first step is to determine the rain rate in mm/hr as a function of the cumulative probability of occurrence. This probability will be defined by the grade of service or availability of the link to be provided. The second step consists in the calculation of the actual rain attenuation associated with the rain rate that was exceeded with such probability.

The attenuation per unit of length (specific attenuation), λ(dB/Km), is tied to the rain rate R (mm/hr), by the empirically derived relationship:

$$\lambda r = a(f) \times R^{b(f)} \quad (\mathrm{dB/Km}) \qquad (25)$$

where a(f) and b(f) are frequency dependent coefficients. For the frequency range between 8.5 and 25 GHz, Eq. 25 becomes:

$$\lambda r = 4.21 \times 10^{-5} f^{2.42} \times 1.41 \times f^{-0.0779} R^{(1.41f^{-0.0779})} \qquad (26)$$

The attenuation per unit length is heavily frequency dependent. Fig. 21 shows frequency dependence of λr for various rain rates.

Introducing the concept of equivalent path length, $L_e(R)$, the total rain attenuation in decibels is simply:

$$Ar = \lambda r L_e(R) \quad (\mathrm{dB}) \qquad (27)$$

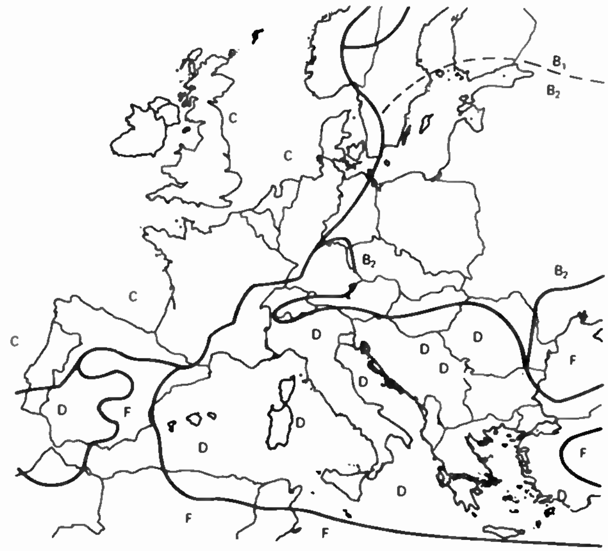


Figure 18. Rain rate climate regions for Europe. (From *NASA Propagation Effects Handbook for Satellite System Design*, ORI TR 1679.)

Equivalent path length is primarily determined by the height of the freezing level or zero-degree isotherm, which depends on latitude, season and rain rate, the cosecant of the elevation angle and site altitude. For latitudes within ±30°, the freezing level is at 4.8 km. Curves of equivalent path lengths versus elevation angle and for different rain rates are shown in Fig. 22.

The rain attenuation is required to be added to the satellite link as a margin to allow the specified availability under fading conditions. Figs. 23A and B show typical rain attenuations versus rainfall rate in the transmit and receive Ku-bands for different elevations angles.

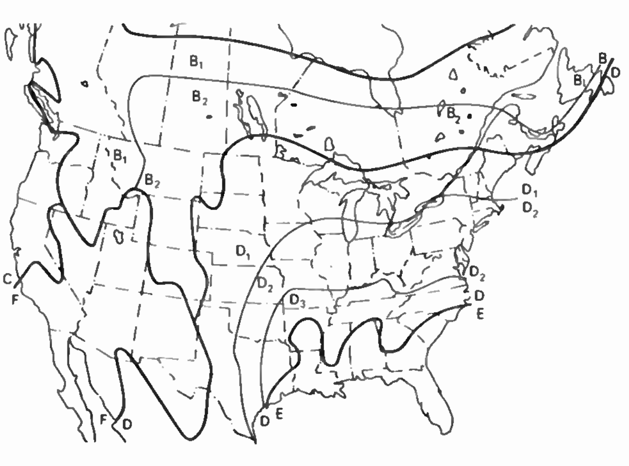


Figure 17. Rain rate climate regions for the continental United States showing the subdivision of Region D. (From *NASA Propagation Effects Handbook for Satellite System Design*, ORI TR 1679.)

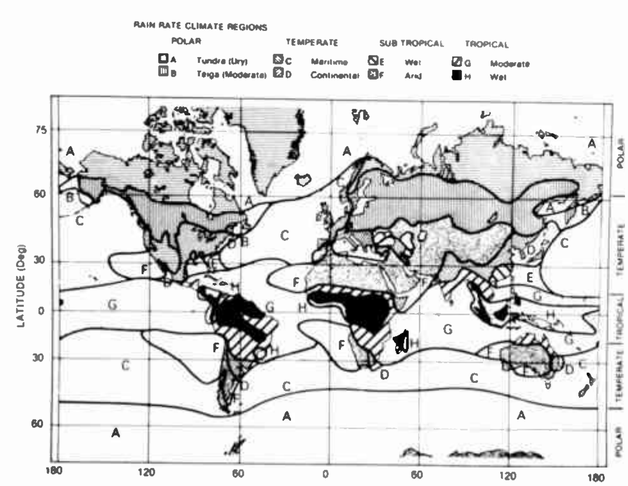


Figure 19. Global rain rate climate regions, including the ocean areas. (From *NASA Propagation Effects Handbook for Satellite System Design*, ORI TR 1679.)

TABLE 2

Point-rain-rate distribution values (millimeters per hour) versus percent-of-the-year rain rate is exceeded.

Percent of Year	A	B	C	D_1	D_2	D_3	E	F	G	H	Minutes per Year	Hours per Year
	Rain Climate Region											
0.001	28.0	54.0	80.0	90.0	102.0	127.0	164.0	66.0	129.0	251.0	5.3	0.09
0.002	24.0	40.0	62.0	72.0	86.0	107.0	144.0	51.0	109.0	220.0	10.5	0.18
0.005	19.0	26.0	41.0	50.0	64.0	81.0	117.0	34.0	85.0	178.0	26.0	.44
0.01	15.0	19.0	28.0	37.0	49.0	63.0	98.0	23.0	67.0	147.0	53.0	.88
0.02	12.0	14.0	18.0	27.0	35.0	48.0	77.0	14.0	51.0	115.0	105.0	1.75
0.05	8.0	9.5	11.0	16.0	22.0	31.0	52.0	8.0	33.0	77.0	263.0	4.38
0.1	6.5	6.8	7.2	11.0	15.0	22.0	35.0	5.5	22.0	51.0	526.0	8.77
0.2	4.0	4.8	4.8	7.5	9.5	14.0	21.0	3.8	14.0	31.0	1052.0	17.50
0.5	2.5	2.7	2.8	4.0	5.2	7.0	8.5	2.4	7.0	13.0	2630.0	43.80
1.0	1.7	1.8	1.9	2.2	3.0	4.0	4.0	1.7	3.7	6.4	5260.0	87.66
2.0	1.1	0.2	1.2	1.3	1.8	2.5	2.0	1.1	0.6	2.8	10520.0	175.30

Source: NASA Propagation Effects Handbook for Satellite Systems Design, ORITR 1679

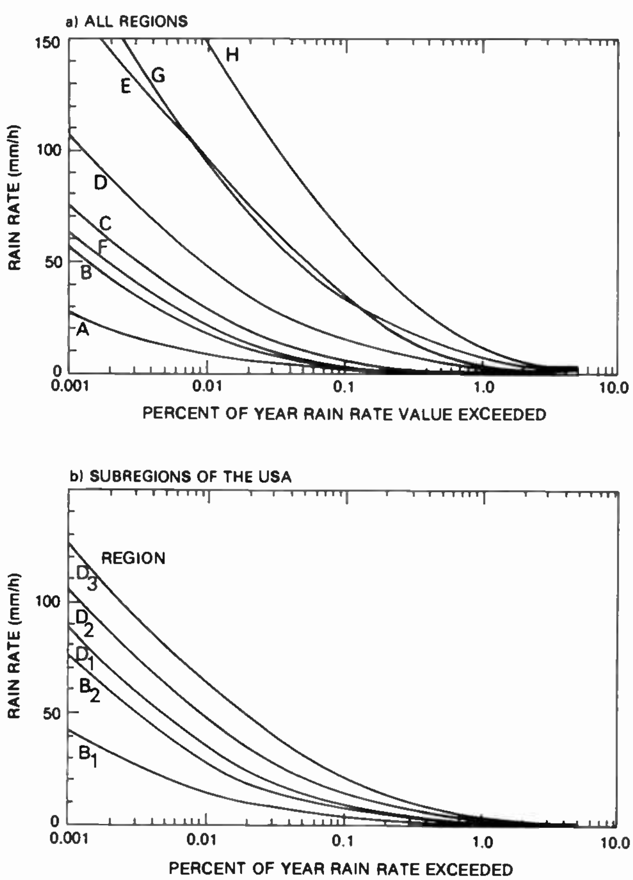
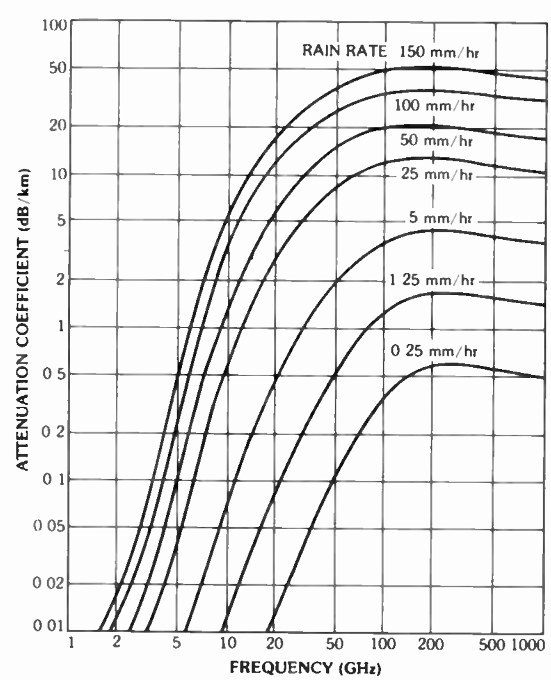


Figure 20. Rain rate cumulative probability distributions for the regions presented on the previous maps. (Reprinted from Louis J. Ippolito, R. D. Kaul, and R. G. Wallace, *Propagation Effects Handbook for Satellite Systems Design* [NASA Reference Publication 1082(03)], National Aeronautics and Space Adminstration, Washington, DC, June 1983. Courtesy of NASA.)

Noise Contribution

In addition to the attenuation, rain also degrades the performance of a satellite link by increasing the earth station antenna noise temperature. In clear weather the antenna sees the cold background of space, but in rain it receives thermal radiation from the raindrops. The increase in antenna noise temperature due to rain, T_r, may be estimated by:

$$T_r = 280(1 - 10^{-A/10}) \ \text{(K)} \tag{28}$$



Raindrop size distribution: Laws and Parsons, 1943
Terminal velocity of raindrops: Gunn and Kinzer, 1949
Dielectric constant of water at 20°C: Ray, 1972

Figure 21. Attenuation per unit length versus frequency and rain rate. (From K. Miya, ed., *Satellite Communications Technology*. Tokyo: KDD Engineering and Consulting, Inc., 1982.)

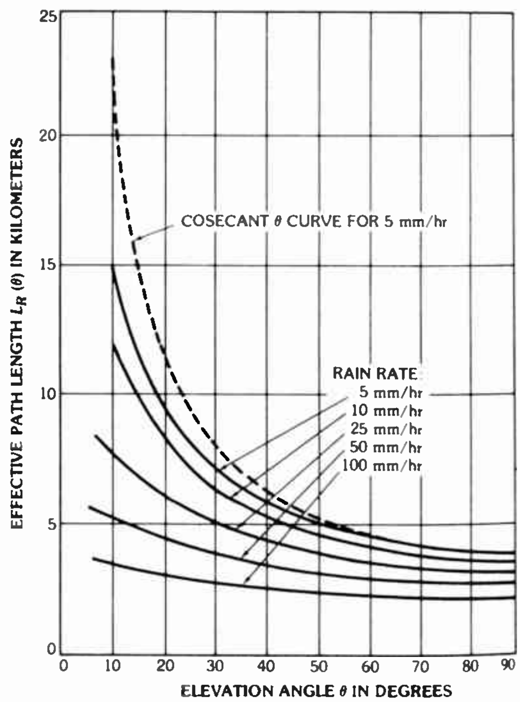


Figure 22. Equivalent path length versus rain rate and elevation angle. (From K. Miya, ed., *Satellite Communications Technology*. Tokyo: KDD Engineering and Consulting, Inc., 1982.)

where A is the rain attenuation in decibels. Fig. 24 shows the impact of the rain contribution of noise temperature on the normal clear sky G/T for different clear sky system temperatures. The G/T degradation corresponding to the rain attenuation for the stipulated link availability also must be added to the satellite downlink. This is to provide sufficient margin to compensate for the combined rain effect of signal attenuation and noise increase.

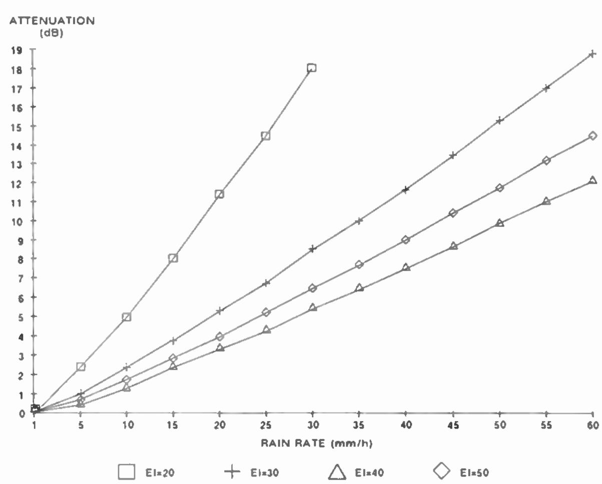


Figure 23A. Rain attenuation (11.95 GHz) 4.8 km zero-degree isotherm.

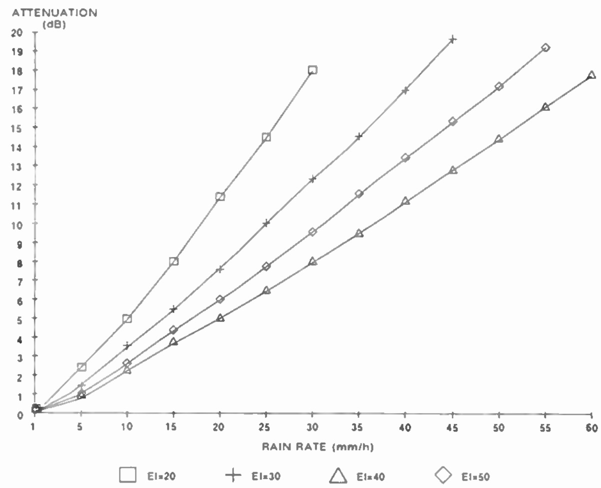


Figure 23B. Rain attenuation (14.25 GHz) 4.8 km zero-degree isotherm.

The allocations of rain fade margins in the uplink and downlink can be done independently, corresponding to specific availability requirements of the uplink or downlink and consistently with the availability requirement of the total link. The assumption is, due to the localized nature of the rain fades, the uplink fade and downlink fade can be considered as two statistically independent processes. Therefore, total link availability can be obtained as the reciprocal of the summation of the uplink and downlink outages calculated as if they occurred independently and one at a time.

Example of System Link Calculation

Table 3 shows a typical satellite link budget for a video uplink and downlink for Ku-band operation where a 5 dB uplink power control has been applied to mitigate the effects of rain fade in the uplink. A C-band link budget would contain similar terms but the uplink power control to mitigate the effects of rain would not be necessary.

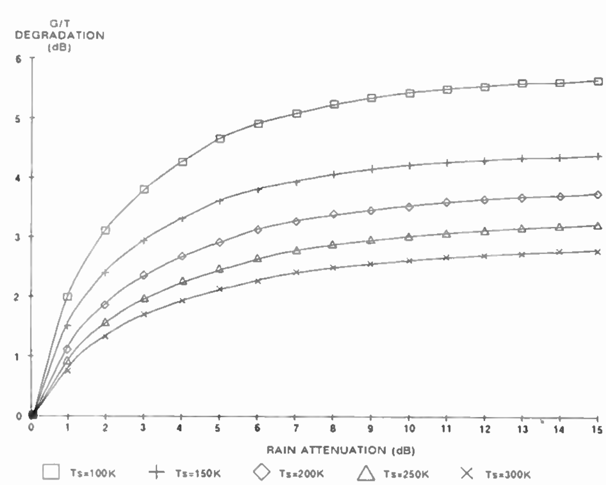


Figure 24. G/T rain degradation.

TABLE 3
Link budget for typical Ku-band video satellite link.

Satellite	Spacenet-II	meter
Beam Type	Conus	MHz
Type of Service	FM/Video	MHz
Transmit/Receive Connectivity	7.0/4.5	MHz
Occupied Bandwidth per Carrier	30.0	
Available Bandwidth per Carrier	36.0	
Transponder Bandwidth	72.0	

Parameter	Values			
I. UPLINK NOISE	Clear Sky	Uplink Fade	Downlink Fade	Units
Earth Station EIRP per Carrier	75.0	80.0	75.0	dBW
Pointing Losses	0.5	0.5	0.5	dB
Path Loss	207.0	207.0	207.0	dB
Isotropic Antenna Area	44.5	44.5	44.5	dB/m²
Saturation Flux Density	−81.0	−81.0	−81.0	dBW/m²
Rain Attenuation	0.0	6.0	0.0	dB
G/T Including Footprint Advantage	−1.1	−1.1	−1.1	dB/K
Input Backoff per Carrier	7.0	8.0	7.0	dB
Uplink Thermal C/N	20.2	19.2	20.2	dB
Co-channel Interference	−27.0	−24.0	−27.0	dBc
Off Beam Emissions Interference	−26.0	−25.0	−26.0	dBc
Total Uplink C/(N+I)	18.5	17.2	18.5	dB
UPLINK AVAILABILITY		99.99		%
II. INTERMODULATION NOISE	−20.0	−19.0	−20.0	dBc
III. DOWNLINK NOISE				
Satellite Saturation EIRP	43.0	43.0	43.0	dBW
Transponder Output Backoff/Carrier	4.5	6.5	4.5	dB
EIRP per Carrier	38.5	36.5	38.5	dBW
Path Loss	206.0	206.0	206.0	dB
Rain Attenuation	0.0	0.0	3.0	dB
Pointing Losses	0.5	0.5	0.5	dB
Earth Station G/T	29.5	29.5	29.5	dB/K
G/T Degradation	0.0	0.0	2.2	dB
Downlink Thermal C/N	15.3	13.3	0.1	dB
Co-channel Interference	−27.0	−25.0	−19.8	dBc
Total Downlink C/(N+I)	15.0	13.0	9.7	dB
DOWNLINK AVAILABILITY		99.85		%
IV. TOTAL C/(N+I) NOISE				
Total C/(N+I)	12.6	10.9	8.8	dB
Occupied Bandwidth per Carrier	74.8	74.8	74.8	dB-Hz
$C/(N_o+I_o)$ Total	87.4	85.7	83.6	dB-Hz
Required C/N_o	83.5	83.5	83.5	dB-Hz
Margin	3.9	2.2	0.1	dB
TOTAL LINK AVAILABILITY		99.75		%

Baseband Performance Link Analysis

As stated above, the overall quality of the delivered baseband signal can be expressed in analog systems by the S/N or in digital systems by the BER. Other performance parameters such as phase and frequency linearity and intersymbol interference, can be utilized to fully characterize quality as is common in any transmission system. However, since satellites are inherently wide band repeaters, S/N and BER are the most sensitive performance factors that depend on link operational parameters.

The C/N_o versus S/N performance of different modulation schemes can be characterized by rather simple mathematical equations. The most common analog modulation technique in satellite transmission is frequency modulation (FM). Because of its simplicity and the low cost of the receivers and demodulators, FM is widely utilized in the transmission of television signals. Eq. 29 allows the computation of the S/N as a function of C/N_o and modulation parameters.

$$(S/N)_w = C/N_o + 10\log(12AF^2/B_v^3) + W \qquad (29)$$

or in terms of carrier-to-total-noise power ratio in a bandwidth B,:

$$(S/N)_w = C/N + 10\log(12 \times AF^2/B_v^3) + 10\log(B/B_v) + W(30) \qquad (30)$$

where: $(S/N)_w$ = Weighted signal-to-noise ratio (dB)
C/N_0 = Carrier-to-noise density ratio (dB)
AF = Peak composite video deviation (MHz)
B = IF predetection noise bandwidth (MHz)
B_v = Video filter bandwidth (MHz)
W = Deemphasis and weighting improvement (dB)

For NTSC format with 30 MHz bandwidth these parameters are typically: AF = 10.75 MHz, B = 28 MHz, B_v = 4.2 MHz, and W = 13.8 dB. Fig. 25 shows the result of Eq. 29 with the above transmission parameters at high levels of C/N. The departure from a linear relationship at low values of C/N is not predicted by Eq. 29 but represents the actual performance of a typical FM demodulator. This phenomenon is known as the threshold effect.

In digital satellite systems the most common modulation technique is *phase shift keying* (PSK). This technique is also known as binary PSK or BPSK in the case that logic symbols zeros and ones are mapped into RF signals 180° apart in phase, and as quadrature PSK or QPSK when the phases are 90° apart. The BER performance in these systems is evaluated as a function of the energy per bit of information transmitted (E_b) versus noise density, E_b/N_0. Sophisticated digital coding and decoding techniques exist that, by adding error correction bits to the information data stream, allow substantial improvements in BER that can translate into transmit power reductions of up to 5 dB. The ratio between the uncoded data rate and the coded one

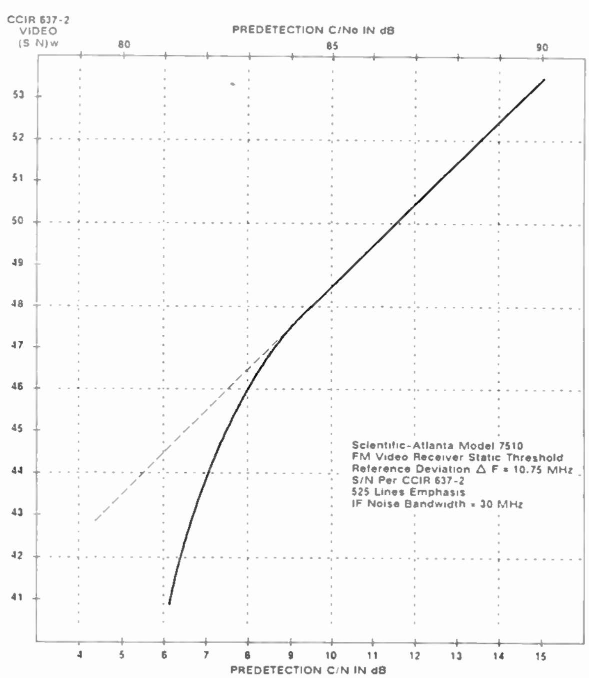


Figure 25. Signal-to-noise ratio performance of FM video demodulator vs. C/N and C/N_0.

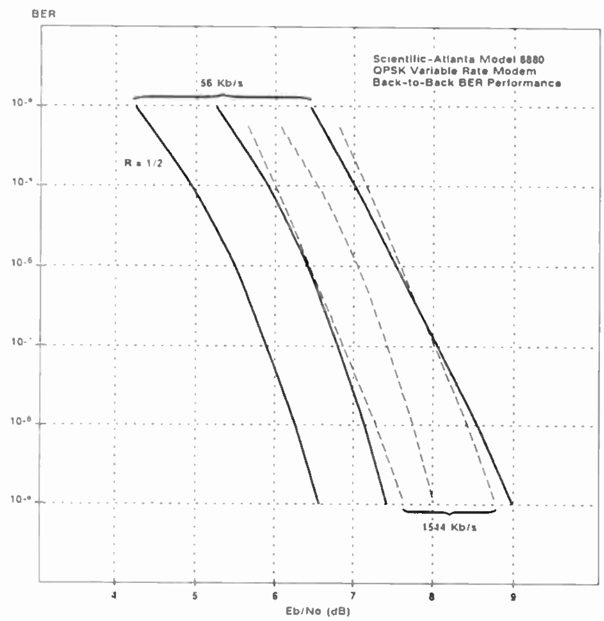


Figure 26. BER modem performance vs. bit energy over noise density ratio.

is called coding rate (R). Fig. 26 shows the performance of a typical QPSK modulator and demodulator for different data and coding rates.

Earth Station Block Diagram

The block diagram of Fig. 27 depicts an earth station capable of providing uplink services for both encrypted video and data in the vertical and horizontal polarizations simultaneously. All subsystems are redundant for maximum reliability. The back-up video exciter and HPA are also capable of occasional video uplinks through transmit couplers in either polarization. A computer based monitor and control system, by means of a serial control bus, offers centralized operation of the complete earth station with the ability of monitoring all status and controlling all variable parameters of every subsystem from the local or remote terminals.

Fig. 28 shows the block diagram of the corresponding dual polarization receive-only terminal. This low cost earth station, with an L-band Inter-Facility-Link, IFL, (950 MHZ to 1450 MHz) can provide simultaneous reception of encrypted video and data.

Interference Analysis

The consideration of interference in a satellite communication system is important, not only for being interfered with, but generating interference into existing systems. It is mandatory for a proposed transmit system in the United States to submit to the FCC a coordination filing that includes an interference analysis. This analysis must show the impact of the proposed system on existing operational systems and must satisfy the allowable interference requirements of the

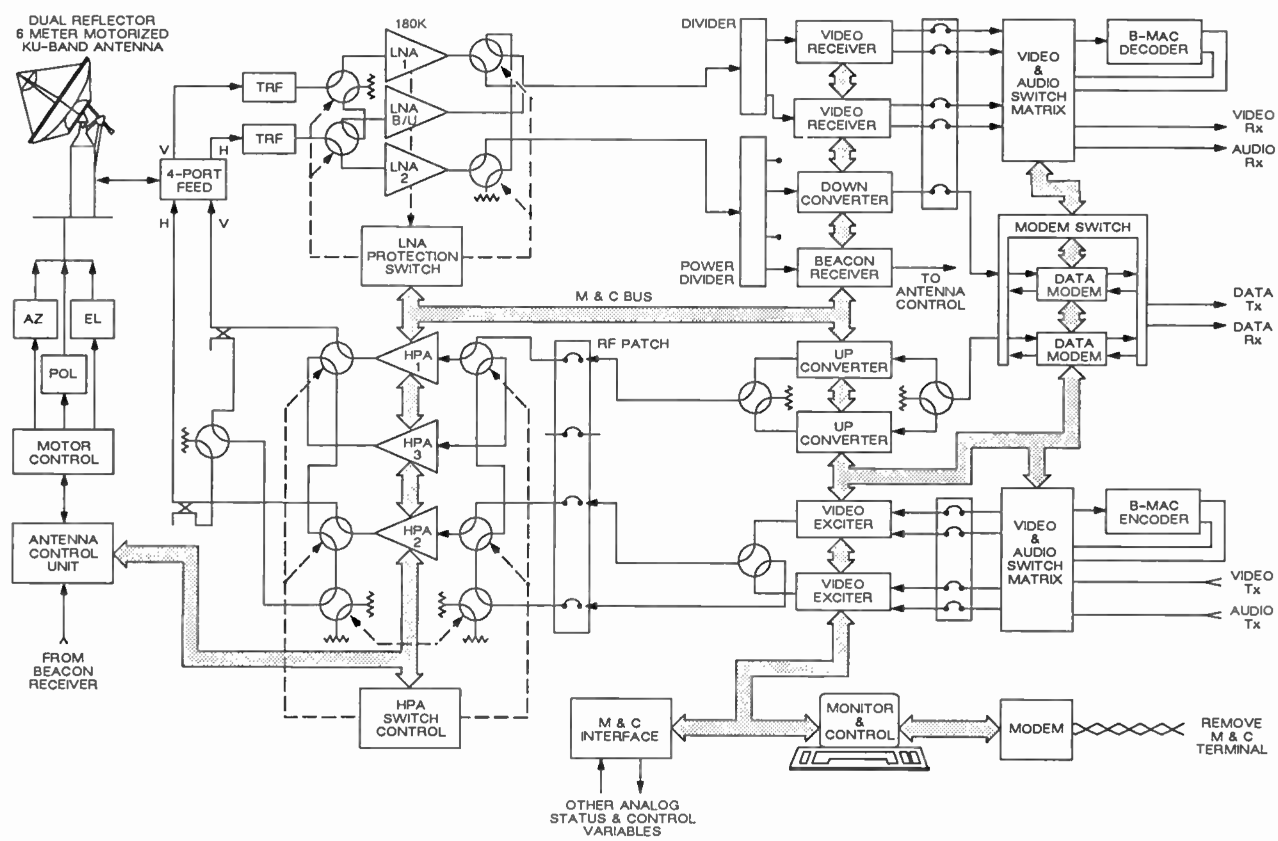


Figure 27. Encrypted video and data broadcast earth station block diagram.

FCC. Coordination for receive-only systems is not mandatory and is only necessary when a C-band system desires interference protection from future terrestrial systems.

The FCC, in August of 1983, finalized a new satellite orbital assignment plan based on a frequency and polarization plan to allow satellite spacing to be reduced from the previous four degrees to two degrees with an interim average spacing of 2.5° at C-band. The implementation of this plan depends on several important technical achievements including:

1. The adoption of a [29–25log(θ)] dBi peak sidelobe envelope for angles off boresight between one and seven degrees.
2. All frequency reuse satellites.

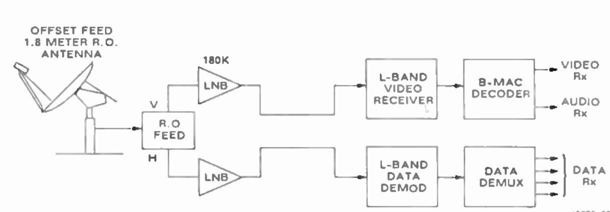


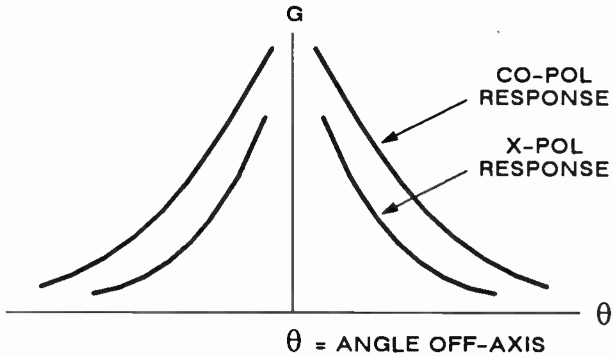
Figure 28. Receive-only earth station for encrypted video and data broadcast.

3. Adjacent satellite, same frequency transponders orthogonally polarized.
4. Homogeneity of satellite EIRP and saturation flux density characteristics for minimum spacing.

Model For Interference Analysis

The interference analysis presented in the following sections are based on models for the antenna characteristics, the spectral characteristics for the desired and interfering carriers, and assumes the final implementation of a uniform two degree spaced satellite environment about the geostationary orbital arc.

Antenna characteristics. The primary characteristics of the antenna which affect the interference analysis are the angular discrimination, the gain differential between the on-axis gain, and the gain of an off-axis angle for the interfering source. For this analysis, the copolarized radiation patterns of the assumed earth station antennas are characterized by the [29–25log(θ)] dBi envelope. The cross-polarized radiation patterns are characterized by [19–25log(θ)] dBi envelope (see Fig. 29). In actuality, the antenna radiation sidelobe may fall below or above these reference envelopes by some predetermined acceptable level. Any sidelobes that are below the reference envelope and at the appropriate pointing angles of adjacent satellites would



GREATER SEPARATION BETWEEN SPACECRAFT PROVIDES INCREASED DISCRIMINATION AND LOWER INTERFERENCE

Figure 29. Earth station characteristics.

reduce the interference and, conversely, any sidelobes above the envelope pointing at adjacent satellites would increase the interference. The cross-polarization discrimination of 10 dB is assumed to apply for clear sky conditions. It should be pointed out that during periods of heavy rain, the polarization of the incoming/outgoing signals may be affected such that the full 10 dB is not realized.

Satellite characteristics. The analysis that follows is based on a satellite deployment model with co-frequency transponders on adjacent satellites being cross-polarized with each other. This model may not necessarily exist in the early 1990s due to the mix of end-of-life satellites and new satellites. Nevertheless, the calculations are performed with this model to demonstrate the expected results in a uniform satellite environment. Three cases of this model are examined:

1. A homogeneous model in which interfering and desired satellites have the same saturation flux density and radiated EIRP. (The radiation patterns yield the same signal strength at any given location on the ground.)
2. A model in which the interfering satellite EIRP exceeds the desired satellite EIRP by 2 dB.
3. A model in which the interfering satellite EIRP exceeds the desired satellite EIRP by 4 dB.
4. The spacecraft antennas are assumed to have a minimum cross-polarization discrimination of 35 dB.

EIRP is a very important consideration. Antenna and transponder characteristics are such that their initial EIRP contours on the earth's surface are not identical. Differences in the initial EIRP contours and differences in transponder aging must be considered in a practical system. An orbital spacing plan that is predicated on differential EIRPs of less than 2 dB represents an impractical burden, both on the satellite manufacturers and on the FCC in assuring compliance with a more stringent specification. Therefore, it is suggested that the calculations for the second case (2 dB variations in EIRPs) be taken as representing a practical case.

The calculations will also be based on geosynchronous rather than topocentric angles and do not include station keeping inaccuracies. An average topocentric angle for the continental United States (CONUS) can be estimated by multiplying the geocentric angle by 1.08.

Sources Of Interference

Interference into a geostationary satellite communication system can originate from several sources, including the following:

- Adjacent satellite signals
- Internal cross-polarization signals (half-transponder frequency offsets)
- Terrestrial microwave signals (does not apply for Ku-band)

These three are analyzed separately in the following paragraphs and then combined to determine the total interference into the system.

Adjacent satellite interference. Interference from adjacent satellites occurs in two ways: uplink interference from earth stations transmitting to adjacent satellites and downlink interference from adjacent satellite transmission into the desired earth station. The interference in both the uplink and downlink consists of many signals (23 of 24 channels for a fully loaded frequency reuse satellite), but it is primarily caused by the co-frequency channels/or transponders and the two half transponder bandwidth offset-channels in a frequency reuse system, Fig. 30. The particular interferers for the C-band example system are the following:

1. The co-frequency, cross-polarized channel on the first adjacent satellite on each side.
2. The two 20 MHz offset-frequency, co-polarized channels on the first adjacent satellite on either side.
3. The co-frequency, co-polarized channel on the second adjacent satellite on either side.
4. The two 20 MHz offset-frequency, cross-polarized channels on the second adjacent satellite on either side.

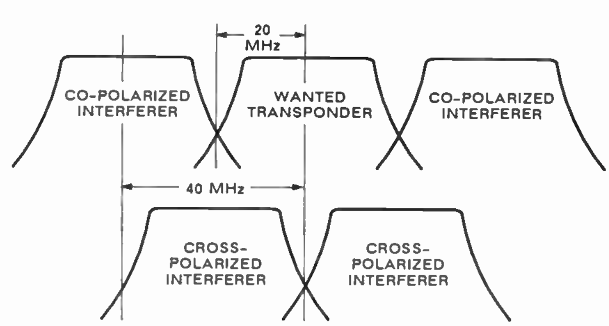


Figure 30. Frequency reuse transponder plan.

The contribution to interference from satellites at orbital positions greater than four degrees from the desired satellite tends to be noise-like in that it is the result of a number of small, relatively noncoherent signals.

The equations for calculation of the adjacent satellite interference are given below:

$$(C/I)_u = (EIRP)_{es} - \sum_{i=1}^{N} \oplus \{(EIRP)_i - [G_i - G(\theta_i)] + F_i + P_i\} \tag{31}$$

where: $\sum_{i=1}^{N} \oplus$ = Series power summation

$(EIRP)_{es}$ = Earth station radiated power in dBW

$(EIRP)_i$ = Effective radiated power of interfering earth station in dBW

G_i = Peak gain of the interfering earth station in dBi

$G(\theta)_i$ = Gain of the interfering earth station in direction (θ) in dBi

F_i = Frequency discrimination factor for *i*th earth station

P_i = Polarization discrimination factor for *i*th earth station

$$(C/I)_d = (EIRP)_{sat} + G_{es} - \sum_{i=1}^{N} \oplus \{(EIRP)_i + G_{es}(\theta_i) + F_i + P_i\} \tag{32}$$

where: $(EIRP)_{sat}$ = Effective radiated power of satellite in the direction of the receive earth station in dBW

G_{es} = Gain of the receive earth station in dBi

$G_{es}(\theta_i)$ = Gain of the receive earth station in the direction θ_i

F_i = Frequency discrimination factor

P_i = Polarization discrimination factor

N = Number of transponders considered (N >= 3)

M = Number of adjacent satellites considered (M >= 4, (typically 8; 4 on each side)

The total adjacent satellite interference is then calculated by combining the uplink and downlink contributions in a power summation manner.

$$(C/I)_{adj.sat} = (C/I)_u \oplus (C/I)_d \tag{33}$$

The polarization discrimination factor in the above equations is the system discrimination rather than that of the receive or transmit antenna alone. A well-designed dual linearly polarized antenna can achieve excellent cross-polarization discrimination on or near the main beam axis (greater than 35 dB or 40 dB relative to the co-polarized energy) and reasonable rejection of the cross-polarized signals in the close-in sidelobe regions. The adjacent satellite signals are received through the sidelobes of the earth station; the $[19-25\log(\theta)]$ envelope is assumed in the analysis. This assumption, rather than being conservative, may be optimistic when one considers the interactions of the ionosphere and atmosphere on the transmitted signals (from the earth station and/or the satellites), and the polarization angle alignment between satellites. A more conservative analysis may assume a slightly reduced discrimination of perhaps $[21-25\log(\theta)]$ for the cross-polarized sidelobe energy in the off-axis regions.

The frequency discrimination factor is related to the spectra of the desired and undesired signals. This factor can range from 0 dB to 12 dB depending upon the interfering power from the different services. For example, the F_i term where an FM/TV signal is interfering a FM/TV signal occupying the same bandwidth would be 0 dB. For a 20 MHz offset-frequency FM/TV signal, with a 36 MHz bandwidth, F_i could range from 3 dB to 15 dB depending on the characteristics of the video signals. It is suggested that for typical FM/TV a value 6.5 dB to 8.0 dB is realistic. The 8.0 db value is suggested by the FCC, but a conservative value of 6.5 dB will be used in the example analysis.

Internal interference. The internal interference in a satellite system is primarily due to the two adjacent 20 MHz offset-frequency channels. The interfering power from different services has been calculated by convolving the power spectra of the individual services. These data, taken together with the appropriate polarization discrimination term, determine the amount of interference and are given in Table 4. The polarization discrimination term is dependent on the climatic conditions which are dealt with in Table 4 by a percent of time condition.

Terrestrial interference. Terrestrial microwave carriers are centered on frequencies offset by 10 MHz

TABLE 4

Summary of internal interference for TV/FM service.

I.		Uplink Pol Disc		Downlink Pol Disc	
	Satellite	35.0 dB		35.0 dB	
	Ground Station	35.0		35.0	
	Faraday	35.0		29.0	
	Resultant	28.0 dB		26.0 dB	
II.	Atmospheric Effect—25 Degree Elevation Angle				
	% Time	Rain Rate	Uplink	Downlink	
	99.0	0.5 in/hr	30.5 dB	33.0 dB	
	99.9	1.5 in/hr	21.0	25.0	
	99.99	3.0 in/hr	16.5	20.0	
III.	Resultant Polarization Discrimination				
	% Time		Uplink	Downlink	
	99.0		23.1 dB	22.8 dB	
	99.9		17.8	19.5	
	99.99		14.5	16.5	
IV. Carrier-to-Interference TV/FM					
% Time	$(XPD)_u$	$(XPD)_d$	$(C/I)_u$	$(C/I)_d$	$(C/I)_s$
99.0	23.1	22.8	27.1	26.56	23.81
99.9	17.8	19.5	21.2	21.26	18.22
99.99	14.5	16.5	18.6	17.96	15.26

from the satellite carriers. To analyze the effects of terrestrial carriers on the FM/TV system, it is necessary to determine the power level of the interfering signal and the spillover of the terrestrial carrier spectra into the passband of the receiver. The first factor involves site details, such as angular discrimination and distance to the interfering transmitter. The second factor can be computed from the spectral distribution projected for the terrestrial carrier and the filter characteristic of the receiver. For the purpose of this analysis, it is assumed the C/I due to terrestrial microwave is 25 dB.

Interference Analysis For FM/TV Service

The interference for the FM/TV service is based on the following parameters:

Parameter	Specification
Transponder EIRP	34 dBW
Antenna Size	10 meter
Transmit	3 meter, 4.5 meter
Receive	7 meter, or 10 meter
Uplink EIRP	80 dBW
Transmit Power	4.5 kW

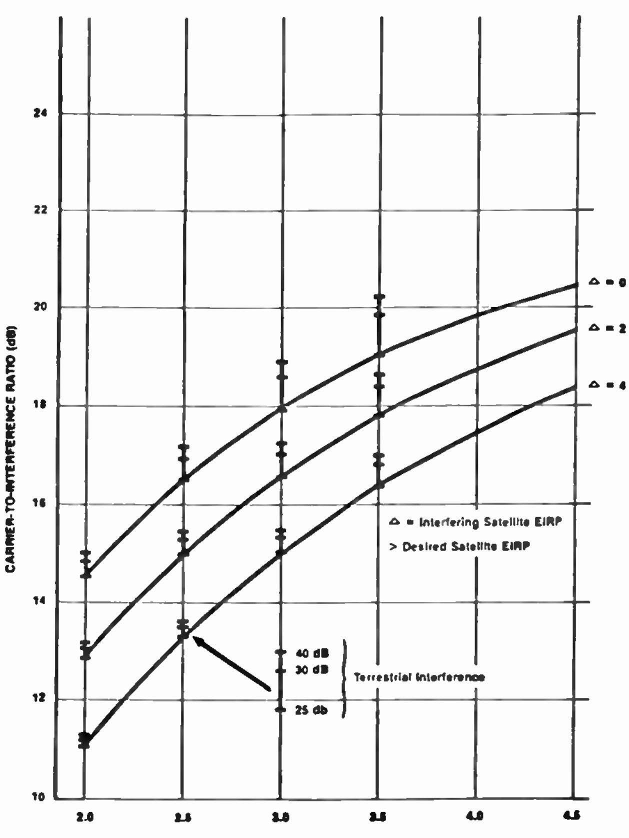


Figure 31. Carrier-to-interference ratio vs. satellite spacing for a 10-meter uplink and a 3-meter downlink. FX = 6.5 dB, PX = 10 dB. Co-Pol 29-25 Log ⊖ envelope and adjacent satellite polarization interleaving.

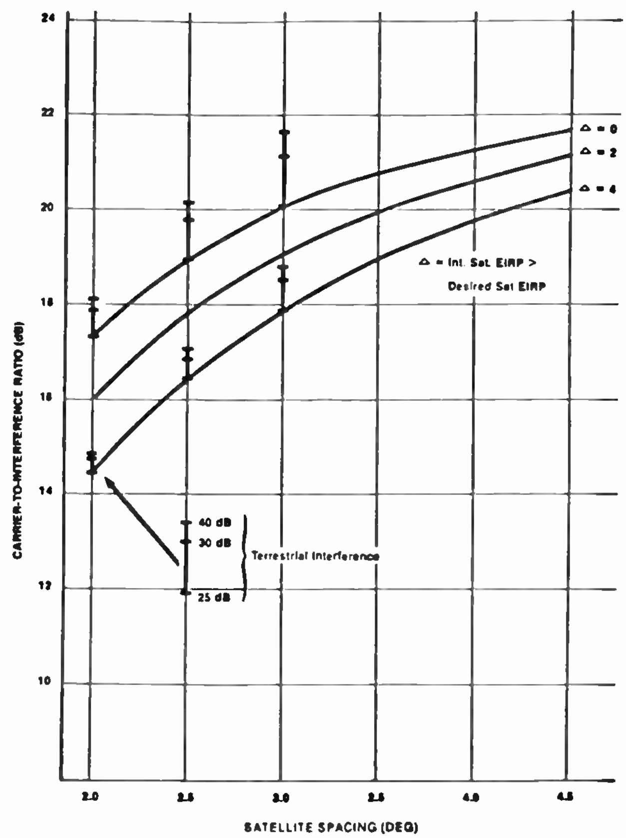


Figure 32. Carrier-to-interference ratio vs. satellite spacing for a 10-meter uplink and a 4.5-meter downlink. FX = 6.5 dB, PX = 10 dB. Co-Pol 29-25 Log ⊖ envelope and adjacent satellite polarization interleaving.

Each of the antennas listed above is presently used in FM/TV systems. Many are licensed and regulated and therefore protected from interference in certain respects. Many receive-only earth stations are unlicensed and not protected. The result of the analyses are presented in Figs. 31 through 34. Each figure includes three cases of desired signal EIRP relative to the interfering signal EIRP, and the effect of variable terrestrial interference is shown.

Sun Transits and Eclipses

Communication satellite systems experience predictable service interruptions involving the sun. A sun transit outage occurs when the pointing angles from a receiving earth station to a satellite and to the sun so nearly coincide that the additional noise power presented by the sun renders transmission unusable. A solar eclipse occurs when the earth shadows the sun from the satellite. The eclipse event is not as serious as the sun transit event since the satellite has battery back-up systems to augment the solar primary power.

Daily sun transits of all geostationary satellites serving an earth station occur during one week in the spring and again in the fall. The exact dates depend

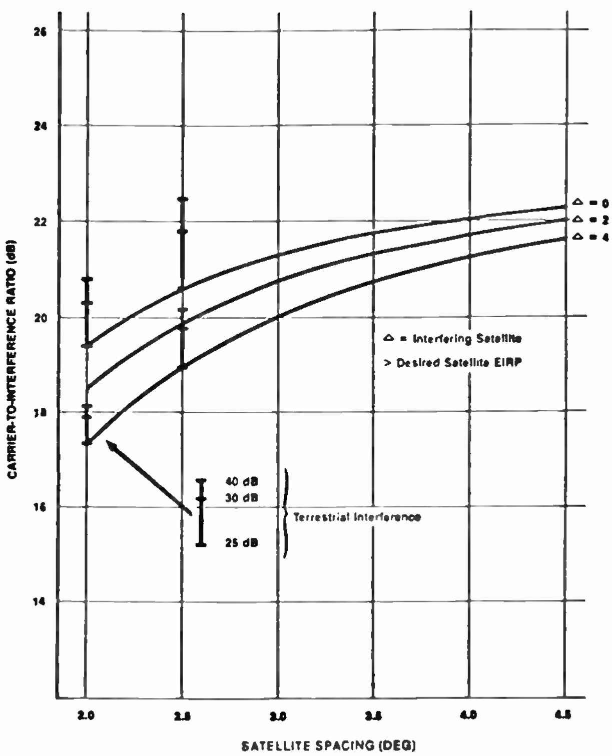


Figure 33. Carrier-to-interference ratio vs. satellite spacing for a 10-meter uplink and a 7-meter downlink. FX = 6.5 dB, PX = 10 dB. Co-Pol 29-25 Log ⊖ envelope and adjacent satellite polarization interleaving.

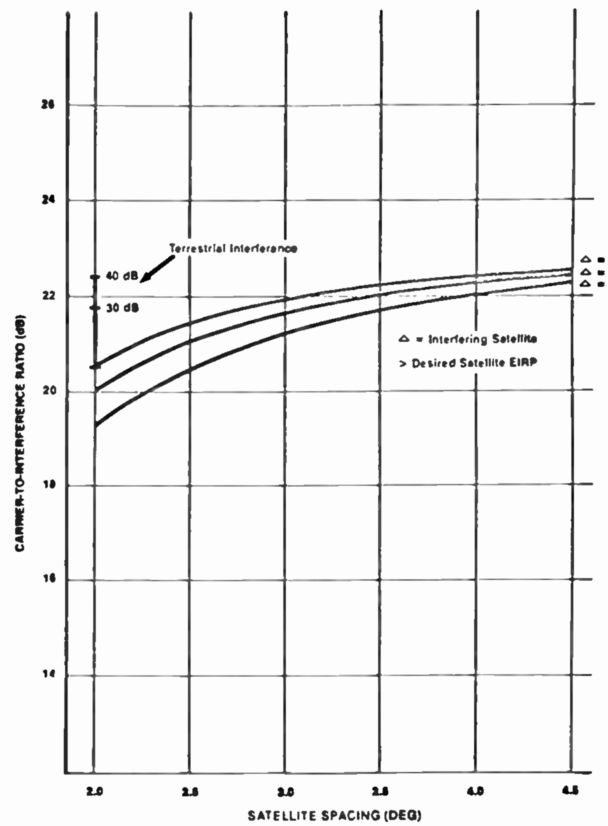


Figure 34. Carrier-to-interference ratio vs. satellite spacing for a 10-meter uplink and a 10-meter downlink. FX = 6.5 dB, PX = 10 dB. Co-Pol 29-25 Log ⊖ envelope and adjacent satellite polarization interleaving.

primarily upon the latitude of the receiving earth station. The geometry and duration associated with a sun transit are controlled by (1) the off-axis gain of the earth station antenna, (2) the receiving system noise temperature, (3) the solar noise power profile, and (4) the minimum acceptable signal-to-noise ratio. In late February or early March, short daily outages affect earth station systems situated near the United States-Canadian border. Two or three days later these systems experience maximum outages lasting five minutes or more, depending upon transmission parameters and permissible S/N. Outages at these earth station locations end after an additional two to three days and the sun transit outage paths progress southward at a rate of about three degrees latitude per day. All outages affecting United States earth station antenna systems above north latitude 26° cease prior to mid-March. Conversely, in the fall the daily outages progress from south to north, affecting southern United States earth stations about October 1 and ending in the north about mid-October.

Eclipses of geostationary satellites can be expected for a total of about 90 evenings per year in the spring and fall. Eclipses occur near apparent midnight of the time zone at each satellite's longitude, beginning in late February or early March and ending mid-April. Fall events begin about September 1 and end mid-October. Eclipses of about 70 minutes duration occur on the dates of the spring and fall equinoxes. Communication satellites are provided with batteries to prevent circuit outages and to maintain pointing, attitude control, station-keeping, telemetry, and command capabilities during eclipses.

EQUIPMENT CHARACTERISTICS

An earth station system is made up of four major subsystems: the antenna subsystem, the transmitting subsystem, the receiving subsystem, and the monitor and control subsystem. These four major subsystems consist of many component parts, each of which will be dealt with in the following paragraphs.

Antenna

The antenna is one of the more important component parts since it provides the means of transmitting signals to the satellite and/or collecting the signal transmitted by the satellite. The antenna not only must provide the gain necessary to allow proper transmission and reception but also must have radiation characteristics that discriminate against unwanted signals and minimize interference to other satellite or terrestrial systems. A further function of the antenna is to provide the means of polarization discrimination of unwanted signals. The individual communication system operational parameters dictate to the antenna designer the necessary electromagnetic, structural, and environmental specifications necessary for the antenna.

Antenna requirements can be grouped into several major categories: electrical or RF, mechanical, and miscellaneous system considerations. Table 5 summarizes many of the more important parameters of an earth station antenna.

Electrical Performance

The primary electrical specifications of an earth station antenna are gain, noise temperature, VSWR, power rating, receive/transmit group delay, radiation pattern, polarization, axial ratio, isolation, and G/T. All of the parameters except the radiation pattern are determined by the system requirements. The radiation pattern should meet the minimum requirements set by the FCC and/or the CCIR of the ITU. Earth stations that operate in a regulated environment in the United States domestic system must meet the requirements set forth in the FCC regulations for earth station antennas pertaining to antenna aperture diameter, sidelobes, (See Section 25.209 of the FCC Rules) and/or radiated power density (See FCC Declaratory Order DA 87–391, Routine Licensing of Earth Stations in the 6 GHz and 14 GHz Bands Using Antennas Less than 9 Meters and 5 Meters in Diameter, Respectively, for Both Full Transponder and Narrowband Transmissions).

The desired radiation properties to satisfy the communication system design dictate the choice of the type of antenna to be employed as an earth station. The three most important radiation properties are gain, sidelobe performance, and noise temperature. Most earth station antennas are designed to maximize gain and minimize noise, thereby maximizing G/T. These two criteria have led to the predominance of reflector type antennas for earth station applications although other types of antennas such as arrays have been used.

TABLE 5

General considerations for earth station antenna design.

Electrical Performance	Frequency (Bandwidth)
	Gain
	Noise Temperature
	Radiation Pattern
	Polarization
	Axial Ratio
	VSWR
	Power Handling Capability
	Port-to-port Isolation
	Out-of-band Emissions
Mechanical Performance	Angular Travel
	Drive Speed and Acceleration
	Pointing and Tracking Accuracies
	Compatibility with Environmental
Conditions	Reflector Surface Accuracy
	Physical Dimensions
	Weight
System Considerations	Operational Functions
	Local and/or Remote Operation
	Availability and Maintainability
	Design Lifetime
	Interface Conditions with Other Subsystems

Types Of Earth Station Antennas

Several types of earth station antennas are now in use within the United States and abroad. These antennas can be grouped into two broad categories, single-beam antennas and multiple-beam antennas. A single beam earth station antenna is defined as an antenna which generates a single beam which is pointed toward a satellite by means of a positioning system. A multiple-beam earth station antenna is defined as an antenna that generates multiple beams by employing a common reflector aperture with multiple feeds illuminating that aperture. The axes of the beams are determined by the location of the feeds. The individual beam identified with a feed is pointed toward a satellite by positioning the feed without moving the reflector. The dual-frequency antennas may be considered another class of antennas as they produce two coincident simultaneous beams, and as such are categorized as single-beam antennas.

Single-beam antennas. The majority of the earth station antennas in use today are single-beam antennas. Single-beam antenna types used as earth stations are paraboloidal reflectors with focal point feeds (prime focus antenna), dual reflector antennas such as the Cassegrain and Gregorian configurations, horn reflector antennas, offset-fed paraboloidal antennas, and offset-fed, multiple reflector antennas. Each of these antenna types has its own unique characteristics, and the advantages and disadvantages have to be considered when choosing an antenna for a particular application.

Axisymmetric dual-reflector antennas. The predominant choice of most system operators has been the dual-reflector Cassegrain antenna. Cassegrain antennas can be divided into three primary types:

1. The classical Cassegrain geometry employing a paraboloidal contour for the main reflector and a hyperboloidal contour for the subreflector (Fig.

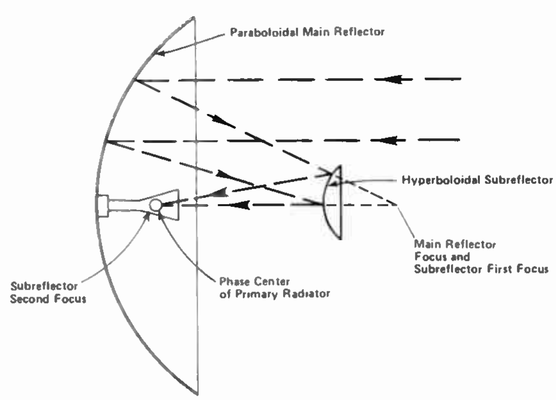


Figure 35A. Cassegrain antenna geometry.

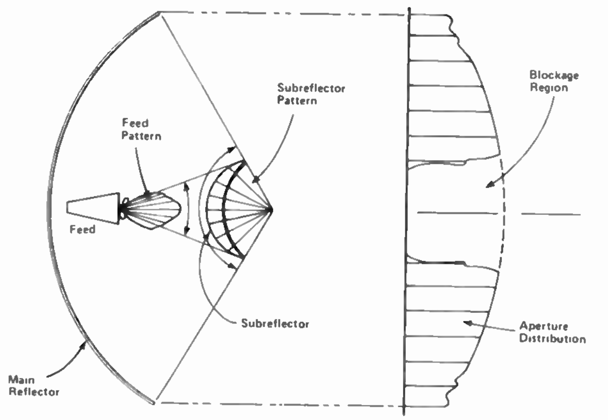


Figure 35B. Aperture distribution of a cassegrain antenna.

35). The paraboloidal reflector is a point focus device with a diameter D_p and a focal length f_p. The hyperboloidal subreflector has two foci. For proper operation, one of the two foci is the *real focal point* of the system and is located coincident with the phase center of the feed; the other focus, the *virtual focal point*, is located coincident with the focal point of the main reflector.

2. A geometry consisting of a paraboloidal main reflector and special-shaped, quasi-hyperboloidal subreflector. The geometry of Fig. 36 is appropriate for describing this antenna. The main difference between the classical Cassegrain mentioned above and this antenna is the subreflector has been designed such that the overall efficiency of the antenna has been enhanced, thereby yielding improved gain performance. This technique is especially useful with antenna diameters of approximately 30 to 100 wavelengths; for example, a 5-meter antenna in the 6/4 GHz frequency band.

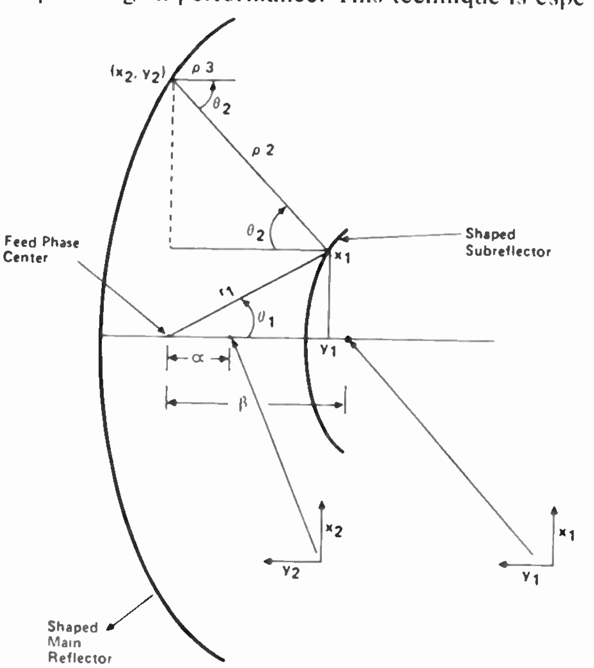


Figure 36A. Dual-shaped reflector geometry.

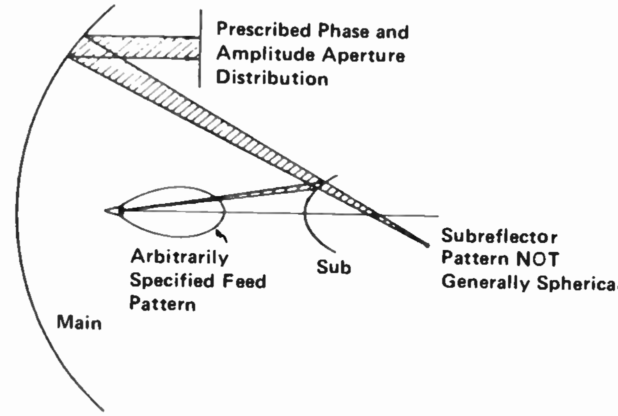


Figure 36B. Circularly symmetric dual-shaped reflectors.

3. A generalization of the Cassegrain geometry consisting of a special-shaped, quasi-paraboloidal main reflector and a shaped, quasi-hyperboloidal subreflector. The subreflector is shaped to redistribute its incident energy such that the illumination of the main reflector is optimized for high gain and desired radiation pattern. The main reflector is then shaped to correct the phase of the aperture field such that it is in-phase. The feed must have a high beam efficiency and its radiation pattern should be circular symmetric. This technique allows the antenna designer to synthesize the surfaces to achieve an arbitrary aperture distribution.

The dual reflector antenna offers excellent gain performance and for aperture sizes larger than approximately 75 wavelengths the sidelobe performance can meet the FCC pattern requirements. Dual reflector designs are employed for earth station antennas for apertures as small as 50 wavelengths to as large as 500 wavelengths.

Prime-focus-fed paraboloidal antenna. The prime-focus-fed paraboloidal (PFFP) antenna is another of the most often employed antennas for earth stations. This type of antenna can have excellent sidelobe performance in all angular regions except the spillover region around the edge of the reflector, but even in this region the pattern requirements of the FCC can be met. This antenna configuration has a lower cost than dual reflector antennas and offers a good compromise choice between gain and sidelobes. Its basic limitations are in its location of the feed for transmit applications and for aperture sizes less than approximately 30 wavelengths, the blockage of the feed and the feed support structure raises the sidelobes with respect to the main beam such that it becomes exceedingly difficult to meet the FCC sidelobe requirements. The PFFP antenna is used for many receive-only earth station antennas as well as for transmit/receive applications when only one transmit polarization is required.

Offset-fed reflector antenna. The offset-fed reflector antenna, Fig. 37, has been used primarily in small aperture antennas for VSAT applications. The offset-fed reflector antenna can employ a single main reflector or multiple reflectors, with two reflectors the more prevalent of the multiple reflector designs. The offset, front-fed reflector, consisting of a section of a paraboloidal surface, eliminates the direct aperture blockage from the feed and feed supports, and minimizes diffraction scattering by removing the feed and feed support structure from direct illumination of the aperture current distribution. The sidelobe patterns of small apertures can meet the FCC and CCIR requirements. The limitations of the offset-fed single-reflector antenna are in its polarization performance, reduced cross-polarization performance off-axis for linear polarizations, and beam squints in opposite directions for two orthogonal circular polarizations. The offset geometry typically means higher manufacturing cost except in the case of small, single piece reflectors such as employed in the VSAT applications.

The offset dual reflector antenna, Fig. 38, can be designed to have all the desirable characteristics of an axisymmetric antenna with increased gain and lower sidelobes. The polarization problems associated with the single offset reflector design can also be compensated for with a two reflector antenna design. The only disadvantages of the offset dual reflector antenna are its high cost of manufacturing for large apertures consisting of multiple sections, and the complexity of its mount geometry and its associated high cost.

Beam waveguide antenna. The beam waveguide antenna, Fig. 39, utilizes a beam waveguide transmission system to minimize the loss between a multiple reflector antenna and its feed, which is located below the rotating axis of the positioning system. This configuration allows the HPA and LNA subsystems to be housed conveniently for maintenance and repair. The beam waveguide design is applicable for very large

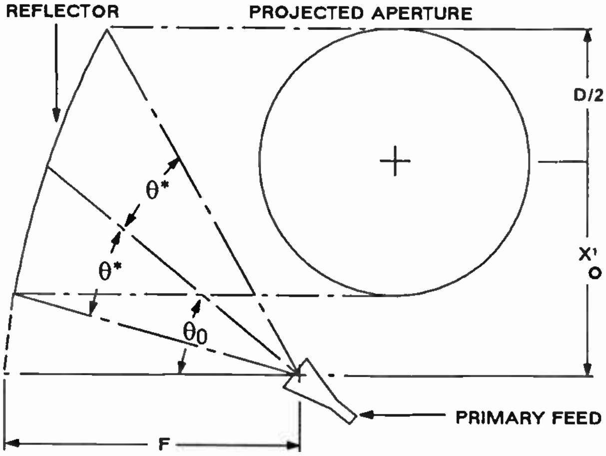


Figure 37. Single offset reflector antenna.

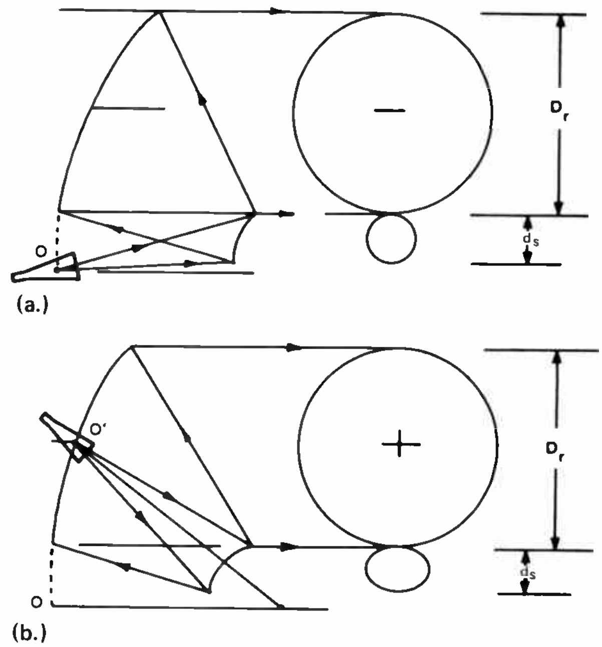


Figure 38. Offset dual-reflector geometries (a) double-offset geometry, (feed phase center and paraboloidal vertex at 0); (b) open cassegrainian geometry (feed phase center located at 0'; paraboloidal vertex, at 0).

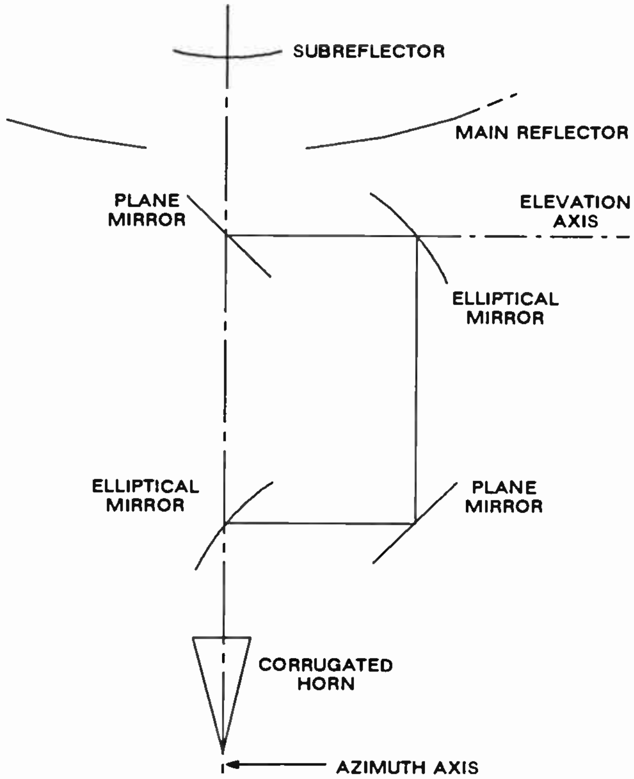
apertures and is very costly to manufacture. Beam waveguide systems have been used primarily for 30 meter C-band Standard A earth stations and for some TT&C earth stations.

Multiple Beam Antenna

Several multiple-beam antenna (MBA) configurations, Fig. 40, are used for earth station applications. These include the spherical reflector, the torus antenna, and a class of offset-fed, Cassegrain antennas. All of these configurations employ multiple feeds to generate the multiple beams. The multiple feeds must be physically small such that the individual beams may be pointed at desired satellites. When the desired satellites are spaced as close as two degrees apart, the MBA antenna may not be practical. The obvious advantage of the MBA antenna is that a single antenna installation can transmit to or receive signals from several satellites simultaneously. The disadvantages are the complexity of the feed arrangements for maintaining pointing for several satellites at the same time when the primary antenna aperture, the main reflector, remains fixed with respect to the earth's coordinates and the stringent requirements for the initial installation of the antenna system. As a result of these disadvantages the MBA antenna is not often employed as an earth station antenna.

Reflector Feed Configurations

There are many different feed configurations used in earth station antennas. The feed configurations are



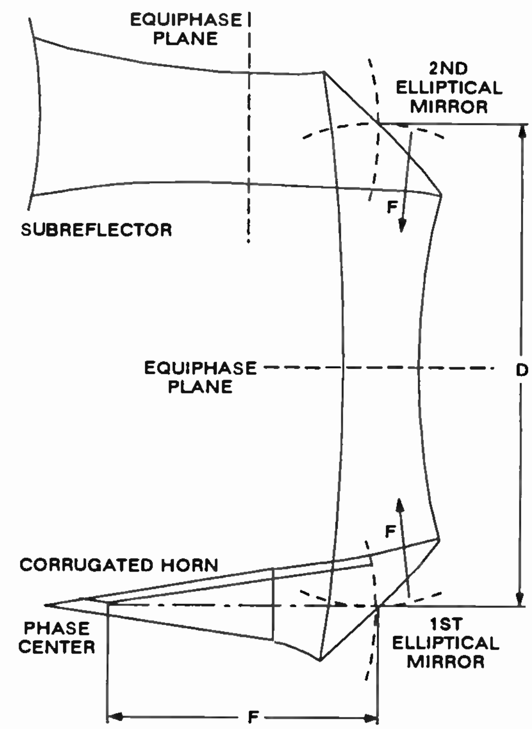


Figure 39. Beam waveguide antenna.

typically classified by the number of transmit and receive ports available. The frequency bands of operation are those specified above for FSS operation in the United States or the appropriate FSS bands for international services. Note that C-band or Ku-band refers to the frequency segments for both transmit and receive bands.

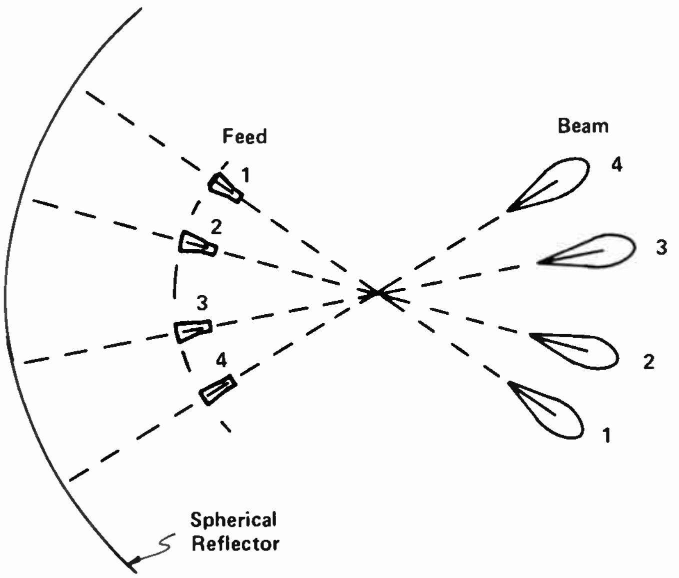
The more popular feed systems are classified as follows:

1. The two-port feed. This configuration may have two orthogonally polarized receive ports or a single transmit port and a single receive port. The transmit and receive ports may be either co-polarized or cross-polarized with respect to each other. The two-port feeds are available in either C-band or Ku-band.
2. The three-port feed. The three-port feed has two receive ports and a single transmit port. The receive ports provide for two orthogonal polarizations. The three-port feeds are available in either C-band or Ku-band.
3. The four-port feed. The four-port feed provides for dual polarization capability for both the transmit and receive applications. This feed configuration is also referred to as a frequency reuse feed. The four port feed configurations are available for either C-band or Ku-band.
4. The dual band feed. The dual band feed provides for the simultaneous reception of C-band signals and Ku-band signals from a hybrid satellite. The dual band feeds are typically a single aperture, that is the C-band and Ku-band radiating apertures occupy the same space and usually sacrifice gain and sidelobe performance to provide the dual frequency operation. This is true for the single reflector and dual reflector designs.

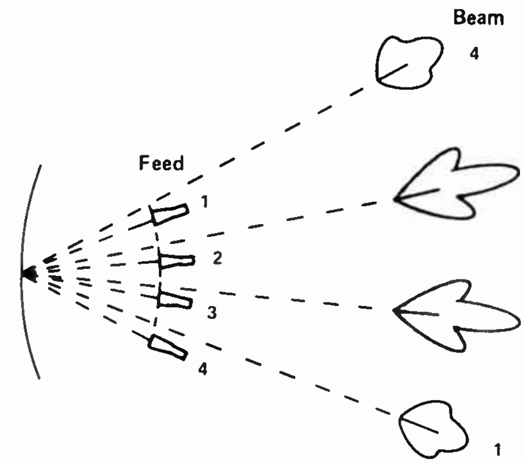
An alternate configuration that does not sacrifice radiation performance utilizes a dual reflector geometry where the subreflector is a frequency selective surface (FSS). This configuration typically uses a prime focus C-band feed, a fss subreflector (which is transparent to C-band and reflective for Ku-band), and a Ku-band dual-reflector feed. The Ku-band feed may be as simple as a single-port feed to a full frequency reuse, four-port feed, whereas C-band is a prime focus, receive-only feed.

Mechanical Performance

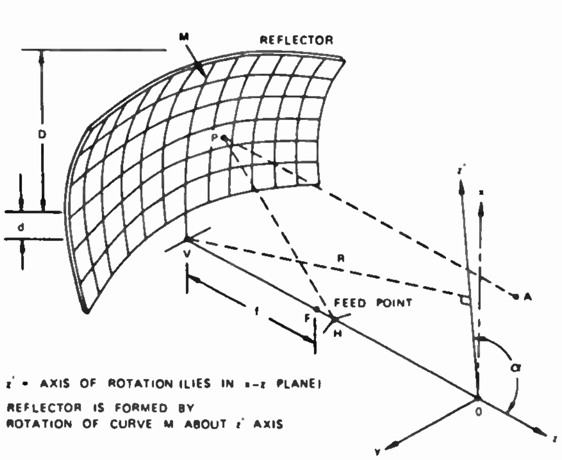
The mechanical design of an earth station antenna must provide the structural integrity to accurately point the antenna beam towards the desired satellite and to maintain the pointing accuracy within the environmental conditions for the locale. Further, the mechanical design of the antenna must ensure the



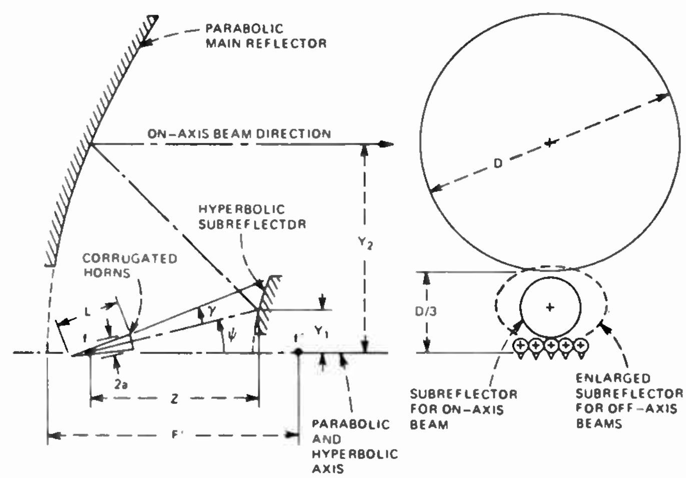
Conventional spherical multibeam antenna
using extended reflector and multiple feeds



Alternative spherical multibeam antenna using
minimum reflector aperture with scanned beam feeds



Torus-antenna geometry
(Copyright 1974, *COMSAT Technical Review*.
Reprinted by permission.)



Geometry of the offset-fed multibeam Cassegrain antenna.
(Copyright, 1974, *American Telephone and Telegraph Company*.
Reprinted by permission.)

Figure 40. Multiple beam configuration.

required tolerance of the radiating surface such that the radiation performance of the antenna is not compromised. The antenna pedestal must also provide the means to steer the antenna beam to the satellites of interest.

The location and size of an earth station antenna system (antenna, pedestal or mount, electronics and control housing) usually make it subject to local building codes. The code which is almost universally accepted is the American Standard Building Code Requirements for Minimum Design Loads in Buildings and Other Structures, ANSI A58.1. Paragraph 1.3 of that code states:

> Buildings or other structures, and all parts thereof, shall be designed and constructed to support safely all loads, including dead loads, without exceeding the allowable stresses (or ultimate strengths when appropriate load factors are applied) for the materials of construction in the structural members and

connections. When both wind and earthquake loads are present, only that one which produces the greater stresses need be considered, and both need not be assumed to act simultaneously.

The loads that must be safely supported by an earth station antenna system are the weight of the antenna and the attached equipment, the expected ice and snow load, earthquake load, and wind load. Of these, the wind load is usually the largest single contributor to the stress and deflection of the structure.

Earth station antennas have a specification that is variously called maximum wind, survival wind, or withstand wind. These terms should be considered synonymous. At the manufacturer's specified survival wind, the system must be safely supported without exceeding the allowable stresses for the materials. Survival wind as defined herein, when combined with ice and dead weight, results in the Design Load as defined in Standard EIA-222C. In addition to survival wind, two other sets of wind conditions are usually specified, the operational wind velocity and the drive-to-stow wind velocity. The operational wind velocity is the maximum value at which the antenna system fully meets the performance specifications. The drive-to-stow wind velocity is the maximum value the antenna may be driven through the azimuth and elevation actuators to the prescribed stow position (usually zenith).

Positioning Systems

There are two broad classes of positioning systems used for satellite earth station antennas. One class consists of orthogonal two-axis configurations; the other is the one-axis or single-axis configuration. The two-axis systems are characterized by the orientation of the lower-most axis with respect to the earth. A two-axis system having its lower axis perpendicular to the ground (Fig. 41A) is called an elevation-over-azimuth. One that has its lower axis parallel to the ground (Fig. 41B) is called X-Y. One that has its lower axis parallel to the earth's axis of rotation (Fig. 41C) is called an hour angle-declination or polar. Each of the three positioning systems has the beam-axis or

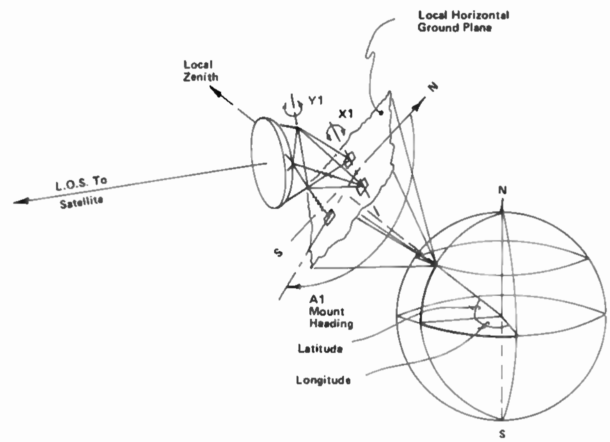


Figure 41B. X-Y mount geometry.

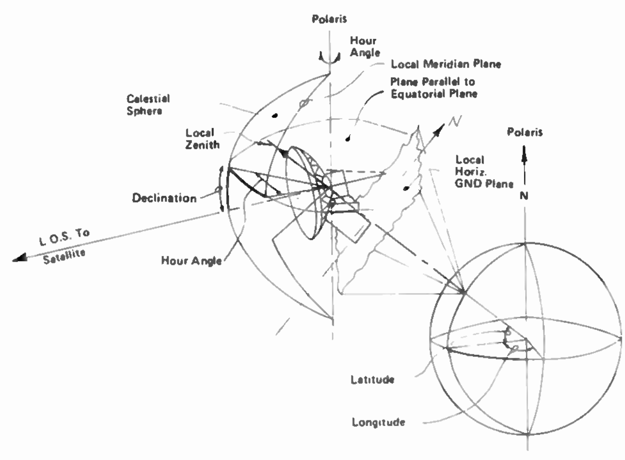


Figure 41C. Polar mount geometry.

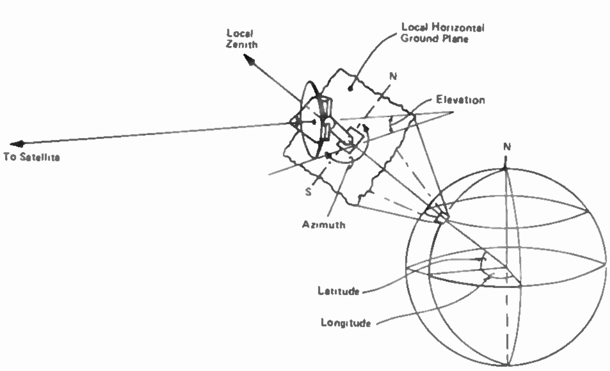


Figure 41A. Elevation-over-azimuth geometry.

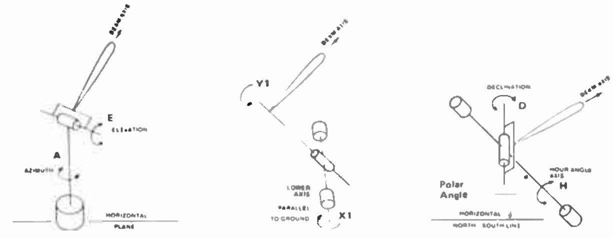


Figure 41D. Schematic illustration of two-axis earth station positioning systems.

pointing direction perpendicular to the upper axis. Providing there are no physical limitations, all three types can theoretically point in any direction.

Development of single-axis antenna mounts was brought about by efforts to reduce the costs and to simplify the positioning of the antenna beam with respect to the geostationary orbit. This geometry has restricted applications due to its lack of capability to follow satellites in inclined orbits and its inherent error as the pointing transverses the geostationary arc.

The elevation-over-azimuth positioner has become the choice for most systems. Fig. 42 is a graph of azimuth and elevation angles versus a particular site latitude and longitudinal difference between the satellite and the site. The horizontal and vertical rectangular coordinates are site latitude and difference longitude, respectively. The curved lines running toward and labeled at the top of the graph are the required azimuth angles (add 180° if satellite is west of site, subtract from 180° if satellite is east of site). The curved lines running toward and labeled at the left margin (down to 15°) are the required elevation angles. (The elevation lines of 10° and below are labeled at the top of the graph.) To determine the required azimuth and elevation pointing angles, find the satellite site longitudinal difference and move vertically on this line until it intersects the horizontal latitude line. At this intersection interpolate between bounding azimuth and elevation curves for the requiring angles.

The azimuth and elevation angles to a particular geostationary satellite can be calculated using the satellite longitude, Z; the site longitude, Y; the site latitude X; and the equations listed below:

$$C = Z - Y \tag{34}$$

$$A(\text{azimuth}) = 180 + \tan^{-1}[\tan(C)/\sin(X)] \tag{35}$$

$$E(\text{elevation}) = \tan^{-1}\left(\frac{\cos(C)\cos(X) - 0.15126}{[\sin^2(C) + \cos^2(C)\sin^2(X)]^{0.5}}\right) \tag{36}$$

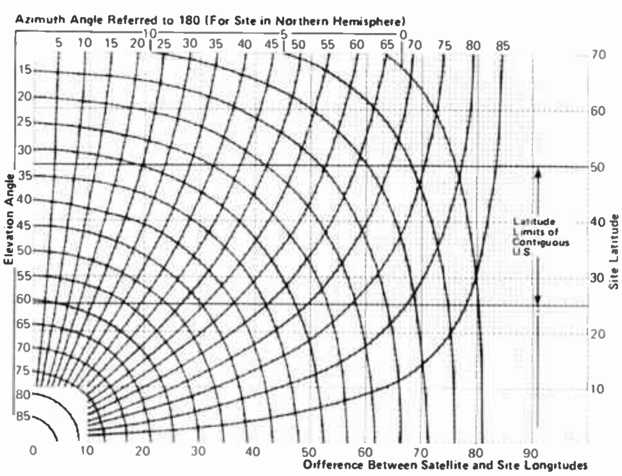


Figure 42. Universal azimuth-elevation look angles.

All of the other environmental conditions at a particular site should be addressed (such as effects of solar radiation, effects of lightning strikes, damage by salt water, acid rain, and pollution gases) in the planning and implementation of the earth station system.

Pointing and Tracking

The pointing and tracking accuracy are two very important considerations for an earth station antenna system. Pointing accuracy is defined as the precision with which an antenna can be held (for fixed position antenna) or steered under the specified operating conditions. The pointing error is a measure of pointing accuracy and is defined as the space angle difference between the command vector and the actual position of the antenna communication RF axis. Pointing error is usually specified to less than 0.2 of the half-power beamwidth (HPBW) of the antenna in the transmit frequency band. Tracking accuracy is the precision with which an antenna can track a source under specified operating conditions. The tracking error is a measure of tracking accuracy and is defined as the space angle difference between the communication RF axis of the antenna and the vector to the RF source. Tracking error is usually specified to be less than 0.1 of the HPBW. The sources of error shown in Table 6 should be considered in an overall budget or calculation of pointing accuracy and tracking accuracy:

TABLE 6.

Pointing and tracking error budget terms.

POINTING ERROR BUDGET	
Velocity lag	Breakaway friction
Wind up of gear train	Secant potentiometer
Angle encoder	Tachometer
Angle encoder coupling	Amplifier drift
Level	Amplifier bias
North alignment, initial zeroing of encoders	Motor cogging
AZ-EL axis orthogonality	Backlash
RF-EL axis orthogonality	Servo dead zone
Reflector alignment	Servo noise
Structural distortion	Axis wobble
Gravity	Radome diffraction
Ice	Boresight shift
Wind	vs. polarization
Thermal	vs. frequency
Acceleration	
Foundation displacement	
Acceleration lag	
TRACKING ERROR BUDGET	
Velocity lag	Wind torque (servo)
Tracking receiver	Null axis—beam axis alignment
Acceleration lag	Breakaway friction
Tachometer	Motor cogging
Amplifier drift	Backlash
Servo noise	

Many earth station systems operate in the point mode, that is, there is no requirement for automatic tracking of the satellite. This condition exists when the satellite orbital location is maintained within a small fraction of a degree (< 0.1°) and when the earth station antenna halfpower beamwidth is sufficiently broad (> 0.5°). Automatic tracking may become necessary as the antenna becomes large in terms of wavelengths (very narrow RF beam), or if the satellite is allowed to transverse an inclined orbit. The complexity of the tracking system is determined by the overall system accuracy requirements and the allowance in EIRP and G/T that is budgeted for impaired operation.

A hierarchy of pointing and tracking systems is as follows:

1. Initial fixed pointing is satisfactory (i.e., receive-only).
2. Repointing of the antenna is required to switch between various satellites or to correct for satellite motion.
3. Tracking is required to correct for satellite drift. Satellite position versus time is known, and program track is satisfactory.
4. Automatic tracking is necessary, but can be satisfied by a simple step-track system.
5. Full automatic tracking is necessary i.e., extended inclined orbits).

The simple step-track system is satisfactory for most satellite communication applications when automatic tracking is required. The step-track systems generate tracking information by moving the RF beam in several steps, comparing the signal level, deciding the proper direction to move for the next step, and then continuing this process until the RF signal is maximized. The step-track system uses a very low-frequency servo loop, and therefore will not track out such disturbances as wind. Step-track can be susceptible to fade conditions unless the sampling circuitry is preset to cut-off when a large signal loss is evident. Steptrack may also be augmented with a program track mode whereby the satellite movement is memorized and then followed by a memory command circuit.

The use of fully automatic tracking systems is typically used for TT&C earth stations or for those earth stations operating under extreme conditions with very narrow RF beamwidths. The automatic tracking configurations include conical scan, electronic beam scanning (i.e., Single Channel Monopulse, ESCAN), and three channel monopulse. The electronic scanning and three channel monopulse techniques offer the advantage of providing a data channel and a transmit channel without tracking modulation superimposed on the signals. This is not possible with the conical scanning technique.

Transmit Electronics

The transmit subsystem consists of equipment from baseband to the high power RF amplifier. These electronics include baseband processors for combined video and audio signals, modulators, upconverters, and high power RF amplifiers. Subcarrier modulators are used to insert various audio signals on the video signal prior to the wideband modulation of the signal. Each of the major components of the transmit electronics for a video service and an audio service are discussed below.

Video Services

Baseband processor. The incoming video signal is first processed by the baseband processing module. The video signal is pre-emphasized for either 525-line (NTSC) or 625-line (PAL/SECAM) operation and passed through a low-pass roofing filter. Preemphasis acts to improve the output video signal-to-noise ratio by compensating for the increase in noise density with frequency (triangular noise) which is characteristic of the receiver demodulator. (The preemphasis is removed by a deemphasis network after the receiver discriminator).

Energy dispersal modulation is applied to the incoming video signal. Satellite transmissions of video signals are processed in this manner to disperse the RF spectrum, thus preventing concentration of energy. This reduces interference with terrestrial microwave and other satellite links, and reduces intermodulation among the multiple carriers which exist in a real satellite. Energy dispersal modulation is applied using a triangular waveform with apexes located at the vertical intervals of the video signal.

Subcarrier modulator. Each subcarrier modulator inputs an audio signal and modulates that signal onto a carrier positioned in the range between 5.0 and 8.5 MHz. Typical subcarrier frequencies are 6.2 MHz and 6.8 MHz. Most video exciters synthesize the center-frequency of the subcarrier with a resolution (i.e. stepsize) of 10 kHz. Generally, a subcarrier modulator will allow selection of preemphasis. Frequency deviation is also typically adjustable between 50 kHz and 500 kHz peak.

Wideband modulator. The wideband modulator frequency-modulates the composite signal (i.e. baseband video and subcarriers) onto a 70 MHz IF carrier. The frequency deviation is adjustable up to 15 MHz peak. The modulation bandwidth of the output signal for NTSC video is approximately 36 MHz for full transponder application.

Upconverter. The upconverter converts the modulated 70 MHz IF signal to a frequency range compatible for satellite transmission. Generally, the signal is converted to C-band or Ku-band.

Video Exciters. Most commercial-use video signals transmitted by satellite are of two types: (1) a broadcast-quality commercial television signal, or (2) a compressed, digitally-encoded representation of a video signal.

Compressed digital video is typically used by private business television networks for training and internal communication on secure channels.

Two types of video uplink devices are described, the analog video exciter and the digital video codec.

Analog Video Exciter: A system that transmits conventional analog video, as used for broadcast-quality transmission, will use a video exciter. The video exciter inputs a set of baseband signals and outputs a single signal suitable for transmission over a satellite link. The center-frequency of the output signal generated by the video exciter is a frequency in the range of the frequency band in use (5.925 GHz to 6.425 GHz for C-band, 14.0 GHz to 14.5 GHz for Ku-band). The set of baseband signals typically consists of a video signal and one or more audio signals. In some cases, one of the audio signals is used to encode a data control channel. A block diagram of a typical video exciter is shown in Fig. 43.

Digital Video Codecs: A system used for transmission of compressed digital video uses a video codec (coder/decoder) to sample the applied analog video and audio waveforms, generating a digital representation of the input signals. Additionally, most codecs perform compression to reduce the bit-rate of the signal. At least two formats of digital video are represented as broadcast-quality. First, the D2 format encodes video at approximately 112 Mbps. This format is sometimes used for studio tape playback and recording of programming. Second, a 45 Mbps system is available for transmission of video over DS3 digital circuits. DS3 is a standard rate (45.3 Mbps) available from common carriers.

High-power amplifier. The high-power amplifier (HPA) amplifies the RF output signal from the upconverter to the required power level for transmission to the satellite. Amplifiers for satellite video applications are typically in the range from 1 watt to 3 kilowatts (kW). Amplifiers in the 1 watt to 10 watt range are available in solid-state configurations. Travelling wave tube (TWT) amplifiers are available in configurations up to approximately 750 watts. For power levels above 750 watts, klystron tube amplifiers are used.

Video uplink systems using U.S domestic C-band satellites usually employ 3.0 kW klyston HPA's. However, newer applications using higher-power Ku-band satellites typically use lower-power TWT amplifiers. Three hundred watt TWT amplifiers are commonly used in occasional video service applications such as education, private business networks, and satellite news gathering.

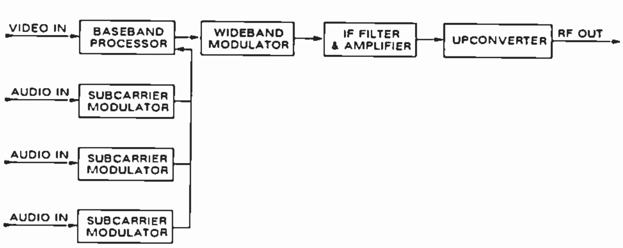


Figure 43. Typical video exciter block diagram.

The HPA usually contains bandpass filters to reject harmonics and power sampling circuits for monitoring the output transmit power and the reflected power from the antenna. Often, protection circuitry is added to turn off the HPA when the reflected power exceeds some predetermined level.

Automatic Transmit Identification System

With the increase in uplink activity in the 1980s, there has been a growing problem with interference between carriers. Most incidents are unintentional, but there have been reported cases of malicious interference. To counter the problem, the FCC introduced a requirement, effective March 1, 1990, that video uplinks must incorporate an automatic transmit identification system (ATIS) to identify the source of the transmission. The ATIS signal is a FM subcarrier positioned at 7.1 MHz. The ATIS subcarrier contains a message composed of international morse code characters to identify the source of the signal and provide a telephone number for communication with its operator. The message is repeated every 30 seconds and includes a unique 10 digit ID code that is unchangeable by the operator. The subcarrier frequency of 7.1 MHz was chosen since it is very close to the second harmonic of the color subcarrier and therefore not useable for any possible revenue source. The subcarrier injection level of −26 dB referenced to the unmodulated main carrier represents a reasonable compromise between ATIS system sensitivity, resistance to interference, and power taken from the main carrier. This injection level is approximately 0.05 of the normal level of a monaural TV associated audio subcarrier.

Audio Service

Audio signals are transmitted by satellite in both analog and digital form. Most of the domestic U.S. nationally-distributed audio material is delivered in digital format. In 1982, four of the leading radio networks in the United States (ABC, CBS, NBC and United Stations (formerly RKO)) made the switch from terrestrial to satellite distribution of network programming. The satellite distribution encodes program material in digital form at the source and distributes the information in that form.

Other networks (i.e., state and regional networks) transmit audio using single-channel-per-carrier (SCPC) systems. These systems frequency-modulate the applied audio signal onto a carrier and upconvert the signal to a satellite-compatible frequency-range.

Digital audio transmission. The digital audio system supports four types of signals: voice-grade, 7.5 kHz audio, 15 kHz audio and data. The 7.5 kHz and 15 kHz audio signals are sampled at 16 kHz and 32 kHz respectively. The signal is digitized with a 15-bit converter and u-law companded to 11 bits plus a parity bit. The voice-grade signal is digitized with a continuously-variable-slope-delta-modulation (CVSD) scheme producing a digital bit rate of 32 kbps. The resultant set of digital signals are multiplexed into a single T1 (1.544 Mbps) bit stream for transmission to the earth station. The T1 signals are routed to the earth station over redundant paths. The earth station receives the multiple, redundant T1 bit streams and demultiplexes the data into the original set of digital signal components.

These component digital signals are then multiplexed into a 7.68 MHz aggregate bit stream. The 7.68 Mbps data is forward-error-correction encoded, modulated onto a 70 MHz IF carrier, and upconverted/amplified to a C-band frequency for satellite distribution. For more information on the digital transmission of the digital information, the reader is referred to the previous section on digital transmission.

Typical bit rates for a 15 kHz audio channel using conventional pulse code modulation (PCM) is 512 kbps. Companding techniques reduce this bit rate to 384 kbps. The availability of low-cost, high-performance digital signal processing (DSP) integrated circuits have now resulted in more effective compression techniques for digital audio encoding. Techniques such as sub-band coding and frequency domain transforms are used to provide typical bit rates of 128 kbps for the same 15 kHz audio channel. Thus, DSP techniques have provided dramatic increases in spectrum efficiency.

Receive Electronics

The receive electronics are similar in scope to the above transmit subsystem but, of course, operate in the reverse order. First the incoming RF signal is filtered, amplified, downconverted (optional, depending on the frequency band of operation), and passed to the receiver where the signal is further downconverted, amplified, and demodulated to baseband.

Low-Noise Amplifiers/Low-Noise Converters

The first active signal processing of a downlinked satellite signal occurs at the low-noise amplifier (LNA) or low-noise block converter (LNB). Traditional C-band broadcast applications have used a LNA mounted at the antenna connected to the indoor electronics through a length of coaxial cable. Typical Ku-band systems use a LNB at the antenna that amplifies and down-converts the signal to L-band (950 MHz to 1450 MHz). The requirements of the LNA are as follows:

1. Provide high gain and low noise to establish high system G/T.
2. Provide transition from antenna waveguide to TEM coaxial cable. Since long waveguide runs are expensive, the LNA is designed to accept a waveguide input and provide a coaxial line output.
3. Provide adequate mechanical strength to permit mounting directly to the antenna waveguide and to allow connection of a long coaxial cable to the unit.
4. Provide RFI/EMI tight weatherproof housing for circuitry.

A block diagram of a typical LNA is shown in Fig. 44.

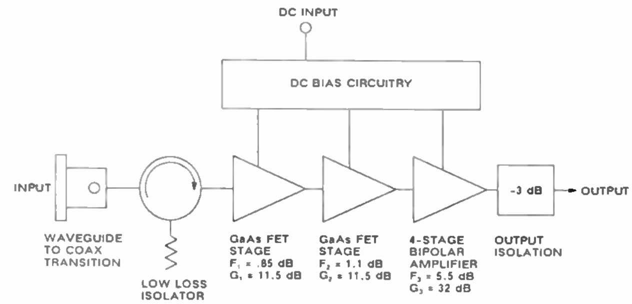


Figure 44. LNA block diagram.

Video Receivers

The video receiver takes the received satellite signal and produces a collection of baseband signals. The baseband signals are a video signal and one or more audio signals. Traditionally, broadcast-quality applications used C-band video receivers. However, with the increased use of Ku-band and decreasing costs for LNB's, many users have converted to L-band systems. A block diagram of a typical video receiver is shown in Fig. 45. The downconverter converts the input RF signal to an intermediate frequency, typically 70 MHz. The demodulator acquires the modulated video signal from the 70 MHz signal.

Downconverter. The downconverter converts the RF-input signal to an intermediate frequency (IF) prior to demodulation. This intermediate frequency is typically 70 MHz. The downconverter will provide either a single-input or multiple-inputs for multiple-antenna, multiple-polarization operation. Single-input downconverters require an external relay to select the appropriate polarization input. Modern video receivers typically provide up to four inputs to provide convenient use for two-antenna, dual-polarization operation.

The downconverter inputs a signal in range of 3700 MHz to 4200 MHz for C-band operation or in the range of 950 MHz to 1450 MHz (L-band) for C- and Ku-band operation. For the latter case, the LNB downconverts the input RF signal to L-band frequencies.

IF-filter/amplifier. The output of the downconverter is routed to the 70 MHz IF-filter/amplifier. The

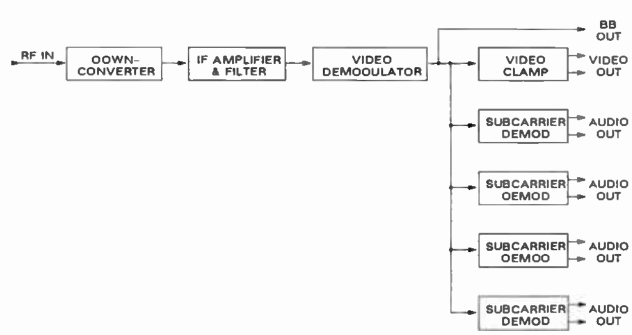


Figure 45. Typical video receiver block diagram.

signal is first bandpass filtered. In the early days of satellite video services the video receivers provided a single IF-filter bandwidth, however, most video receivers offered today provide multiple IF-filter bandwidths to provide convenient half-transponder as well as full-transponder operation. Some video receivers provide up to six IF-filter bandwidths. Six IF-filter bandwidths are economical because of the availability of surface-acoustic-wave (SAW) filters for this application.

Demodulator. The filtered, amplified IF-signal is fed to the demodulator for FM-demodulation. The output of the demodulator is the baseband video and multiple audio subcarriers.

Video processing. After demodulation, the video baseband contains the 30 Hz triangular energy dispersal waveform, which is removed by a circuit referred to as the clamp. A low-pass filter is applied to remove the audio subcarrier(s) from the video baseband signal.

Subcarrier demodulator. Video receivers generally provide up to four audio subcarrier demodulators and provide baseband audio (generally 600-ohm, balanced) outputs.

Digital demodulator/downconverter. Digital data is extracted from an intermediate frequency carrier by the demodulator section of a modem. For detailed information about the operation of the modem, refer to the section on Digital Transmission. The digital demodulator separates the transmitted data stream from the carrier. The transmitted data stream is then processed by the FEC-decoder; removing the error-correction bits, correcting detected errors, and outputting the received data and clock signals.

Protection Switching

With the exception of receive-only systems, most satellite transmission systems contain redundant subsystems to facilitate very high availability specifications. Satellite teleports advertise availability specifications as high as 99.995%. Protection switching implements automatic subsystem redundancy. The protection switch monitors one or more online subsystems for failure. Upon detecting a failure in the subsystem, the protection switch switches the inputs and outputs from the online unit to the backup unit, and given the same configuration as the failed online unit.

The configuration of the subsystems protected by a protection switch is referred to as *m:n*. *m* refers to the number of backups available. *n* refers to the number of online units monitored and protected by the protection switch. 1:1 refers to the simplest configuration: a single backup is available to replace a single online unit. A typical 1:1 protection system for a Ku-band uplink in shown in Fig. 46. A typical 1:1 C-band downlink is shown in Fig. 47. Operation of the 1:1 configuration is simplified in that the backup may be tuned to the same configuration as the online unit; switching to the backup merely requires switching the source of the input and output signals from the online unit to the backup. An example of a larger configuration is 2:6, where two backup units are available to protect six

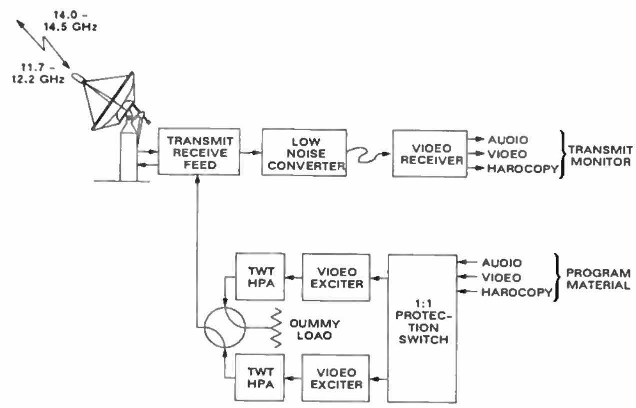


Figure 46. Redundant Ku uplink protection system.

online units. These larger configurations are used for large multiple-channel uplinks. The larger configurations become more cost-effective as the number of online channels increases.

Monitor And Control Systems

Monitor and control refers to systems used to monitor earth station components for failures, and provide manual and automatic control of the components. These systems are widely used for a variety of reasons. First, even though most earth station components provide front panel monitor and control functions, there are generally too many earth station components to monitor from the front panels of the respective components. A monitor and control system provides a single-point of monitoring and control for the operator, thus easing the operator workload, allowing the operator to handle more transmissions. Second, many earth stations are located remotely from the studio. A monitor and control system provides remote monitoring and control (i.e., operation) of the earth station. Remote operation is facilitated through the use of a low-speed data circuit between the earth station equipment and the monitor and control computer in the studio. This circuit is typically (1) a subcarrier on a microwave-radio channel, (2) a dial-up phone circuit with 1200 baud to 9600 baud modems, or (3) a dedicated EIA-422 hard-wire connection for distances less than 1,000 meters.

Earth Station Control Computer

Monitor and control systems are based upon the use of a general purpose computer executing applications

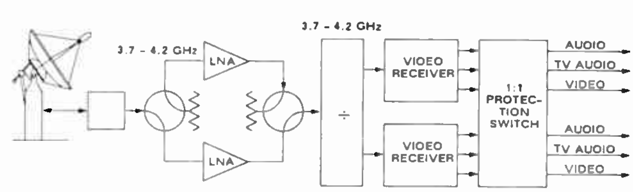


Figure 47. C-band redundant downlink protection system.

software designed for monitor and control of communications systems. This computer is referred to as the earth station controller.

Operator interface. Earth station control computers provide a video monitor for display and a keyboard for operator interaction. Some systems also provide pointing devices such as a mouse or trackball. The system displays earth station status in the form of a hierarchical, graphical display. The upper level display shows a high-level block diagram of the earth station. The operator may select more detailed displays of subsystems by selecting the subsystem symbol from the screen. Most systems display the earth station information on color displays. The earth station components/subsystems are coded with color to indicate state. Typical color conventions for earth stations are as follows: blinking red indicates an unacknowledged alarm; red i ndicates an acknowledged alarm; amber indicates ready or standby; green indicates online/OK/normal. In some cases, an audible alarm is provided. When a new alarm condition occurs, the audible alarm is turned on. The operator must acknowledge the alarm to turn off the audible alarm.

Earth station interface. The computer interfaces to the earth station components through a number of interfaces; including serial, contact closure, and other customer supplied interfaces.

1. *Serial Interfaces.* The most common interface is a serial ASCII-protocol-based interface. The interface is usually an asynchronous character-oriented scheme utilizing EIA-232C or EIA-422 signal levels. EIA-232C interfaces are used for short cable distances (computer to device) less than 5 m or for connection to a modem. EIA-422 interfaces may be used for cable lengths of up to 1000 m and are used in multi-drop mode, thus allowing many devices to share a single interface port.

   The protocols are usually ASCII-based since there is no satellite-communication standard, and most vendors use their own version of an interface protocol. Thus, earth station computers must offer a number of serial equipment interface ports and also support a variety of protocols on those ports.
2. *Contact Closure Interfaces.* The second type of interface commonly used is the contact closure interface. Older components typically provided contact closure interfaces and offered no serial interface functions. Many of these components are still in use today. In addition, there are components, such as waveguide switches and shelter alarms (i.e. intrusion, air conditioner, emergency generator, etc.), that only offer a contact closure interface. For example, a waveguide switch may provide two status points and two control points. Thus, earth station computers must provide some method of monitoring and controlling status and contact closure controls. Some vendors offer systems that connect the contact closure directly to the computer. Others offer general purpose interfaces that reside in the earth station and interface to the earth station computer through a serial interface.

   Status inputs to the earth station computer are usually optically-isolated. The earth station computer supplies the optical-isolator. The monitored device sinks current through the isolator to indicate one of two states. The alternate state is indicated by no current flow.

   Control outputs are of two types. The most flexible interface is the Form C output. The Form C output provides a common connection, and a normally-opened (NO) and normally-closed (NC) connection. The second type of control output is the open-collector output. This interface provides a connection to the collector of a transistor to sink current. One control state is enabled by thus sinking current from the controlled device through the transistor. The alternate state is enabled with the transistor off.
3. *Vendor Supplied Computer Interfaces.* A third type of earth station interface is a vendor supplied, device specific contact closure adaptor. Most TWT amplifiers provide contact closure-based remote monitor and control interfaces. Additionally, forward and reflected power indications are provided by a signal with voltage level proportional to power. Additionally, some TWT amplifiers require an analog current signal to control the attenuator. However, most TWT amplifier vendors offer an interface adaptor that converts the contact closures to a serial-protocol based EIA-232-C or EIA-485 interface. These adaptors reside in the rack with the TWT amplifiers and connect to the earth station computer through a serial connection.

CONCLUDING REMARKS

There are many additional aspects of the design, installation, operation and maintenance of an earth station antenna system that have not been discussed above. The site selection and preparation are, in particular, critical to the successful operation of the system as well as the foundation design. Details of this aspect of the earth station design should be reviewed with the help of qualified civil engineers and frequency coordination experts. The operations building and/or equipment houses should be in close proximity to the earth station, if possible, but remote operation is possible. The power requirements for the earth station should also be carefully planned to provide adequate power for the electronics; including the transmitter equipment and any power required for antenna deicing where applicable.

BIBLIOGRAPHY

Cook, J.H., Jr. & Hollis, S., Editors, "COMMUNICATIONS SYMPOSIUM '83 NOTEBOOK," Scientific-Atlanta, Inc., November, 1983, Atlanta, GA.

Pratt, T. & Bostian, C.W., SATELLITE COMMUNICATIONS, John Wiley & Sons, New York, N.Y.

Pritchard, W.L. & Sciulli, J.A., SATELLITE COMMUNICATION SYSTEMS ENGINEERING, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Martin, James, COMMUNICATIONS SATELLITE SYSTEMS, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Morgan, W.L. & Gordon, G.D., COMMUNICATIONS SATELLITE HANDBOOK, Wiley-Interscience Publications, John Wiley & Sons, Inc., New York, NY.

Freeman, R.L., REFERENCE MANUAL for TELECOMMUNICATIONS ENGINEERING, Wiley-Interscience Publications, John Wiley & Sons, Inc., New York, NY.

Jordan, Edward C., REFERENCE DATA for ENGINEERS: RADIO ELECTRONICS, COMPUTERS, and COMMUNICATIONS, Howard W. Sams & Co., Inc., Indianapolis, IN.

Inglis, Andrew F.,Editor, ELECTRONIC COMMUNICATIONS HANDBOOK, Chapters 2,5, 6, & 17. McGraw-Hill Book Co., New York, N.Y.

Johnson, R.C. & JASIK, H., ANTENNA ENGINEERING HANDBOOK, Chapter 36, "Earth Station Antennas," Cook, J.H., Jr., McGraw-Hill Book Co., New York, NY.

Ippolito, Louis J., Kaul, R.D., and Walace, R.G. Propagations Effects Handbook for Satellite Systems Design. NASA Reference Publication 1082, Dec. 1981.

4.5 Fiber Optic Transmission Systems*

C. Robert Paulson
AVP Communication, Westborough, Massachusetts

FIBER OPTIC TRANSMISSION SYSTEM PRINCIPLES

Communicating over a thin strand of glass fiber is literally analogous to communicating by making a flashlight blink. A practical forerunner of fiber optics was the light-signalling telegraph adopted by the world's navies after Edison and before Marconi. Technically, information transmission using laser technology is accomplished by regulating the flow of quantities of photons of light in an optical glass fiber, rather than the amplitude and frequency of electromagnetic waves in free space or the coulombs of electrons in a copper wire (see Fig. 1).

The electrical signals used to drive the input amplifier connected to a copper wire circuit, or to modulate a radio transmitter power amplifier or oscillator, are used instead as a control voltage to vary the amplitude of the light output (intensity) of a light source, a light emitting diode (LED) or injection laser diode (ILD), around a mid-brightness intensity level. The term *intensity* is preferred over the term *amplitude* to express the output signal from a light source. The source is never varied to full on or full off, because of the hysteresis (nonlinearity) characteristics common to most transduction processes (energy transfer from one form to another).

In order to transmit information, the modulated light signal is focused onto the end of a glass fiber. At the other end of the fiber, the attenuated modulated light is focused onto a photo or photon detector, which translates the light variations back into an electrical signal.

In fiber optic terminology light energy is expressed in terms of power in watts (or more commonly in terms of milliwatts) and light frequency is expressed in terms of wavelength instead of Hz.

The first practical applications of fiber optic technology for the transmission of communications signals occurred in the mid-1970s. By 1990, lightwave-conducting fiber cables were the "backbone" of many long distance telephone networks, carrying up to tens of thousands of digitized voice and computer digital data channels on each fiber.

Fiber optic cables are manufactured in a wide variety of combinations of fiber quantities and types, some including copper pairs for simultaneous electrical power distribution. Strength members in the cable may be either steel or nonconducting plastic for terminal isolation from different earth ground potentials. Outer sheaths range from materials meeting building codes for plenum installation to watertight steel housings for trans-oceanic cables.

In 1980, fiber cables were installed throughout Lake Placid, New York, in the first television broadcasting industry application of lightwave transmission technology. Each fiber carried a single 4.2 MHz NTSC video signal and a single audio channel. Dozens of these single circuits connected venue-located production trucks to the ABC Network production center. Since then there has been an ever increasing use of fiber optic transmission facilities by the broadcast industry, either inside the plant, or part of remote (outside broadcast) productions, or in common carrier and private circuits used for interconnecting broadcast facilities.

Fiber Optic Transmission Advantages

Bandwidth

The information carrying capacity of a carrier is directly related to its frequency. Light carrier frequencies are several orders of magnitude higher than the

* This chapter includes material prepared for the NAB *Engineering Handbook*, 7th Edition, by Pete Mountanos, then with the Grass Valley Group, Grass Valley, CA.

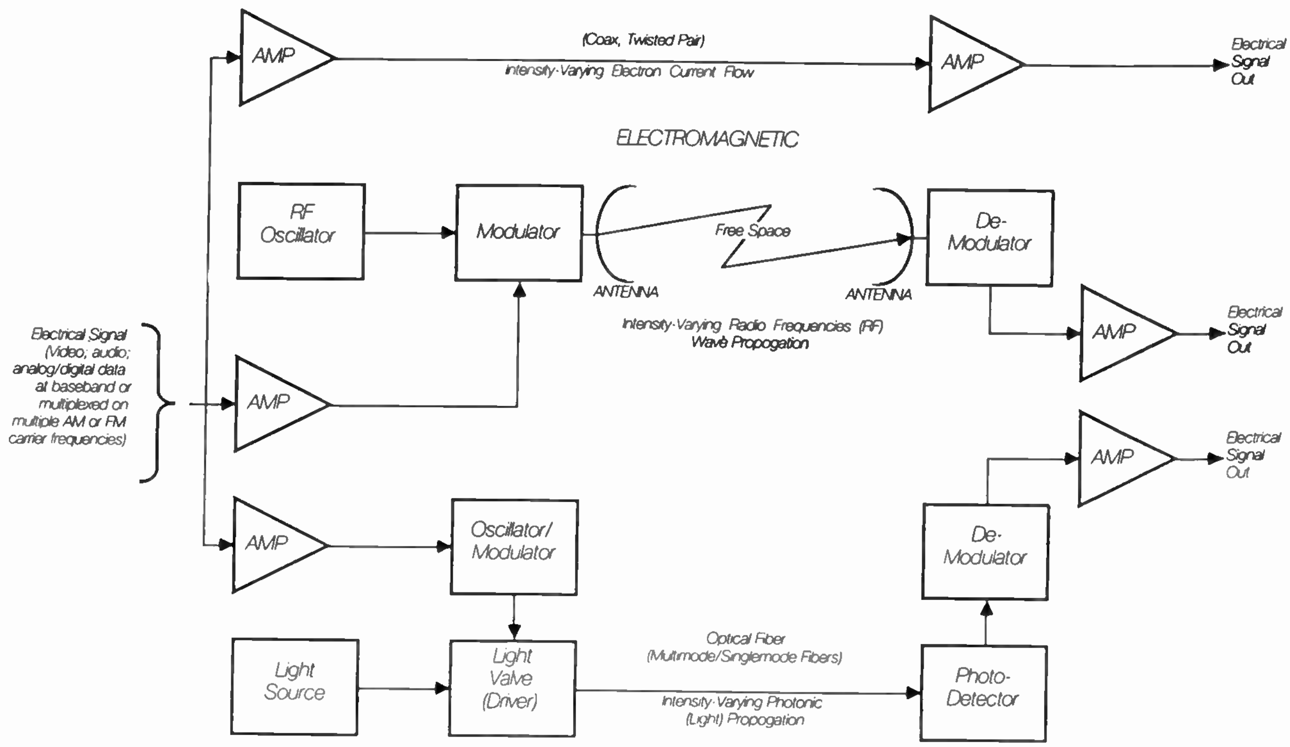


Figure 1. Alternative means for electrical signal transmission.

highest radio frequencies as shown in Fig. 3. Fiber optic transmission systems easily surpass the information carrying capacity of typical microwave radio and coaxial cable systems. Fibers can have a bandwidth in excess of several gigahertz which allows high speed transfer of most types of information. Multiplexing techniques allow many signals to be sent over a single fiber. Light carrier frequencies are expressed in wavelength (nanometers).

Low Loss

Fibers have substantially lower attenuation than coaxial cables and twisted copper pairs, and require no equalization. In premium cables the attenuation can be less than 0.5 dB/km at certain wavelengths while typical fibers have 2 to 3 dB/km.

Electrical Isolation

Fibers and their protective coatings are made of dielectric material, and the transmitters and receivers in each circuit are electrically isolated from each other. Isolation of the electrical grounds of each end is assured if the strength material (messenger) in the cable is also a dielectric. Lightwave transmission is free of spark hazards, and creates no electromagnetic interference (EMI). Lightwave transmission is impervious to contamination by EMI or radio frequency interference (RFI). An all-dielectric fiber cable may be installed in hazardous or toxic environments such as sewer pipes.

Size and Weight

An optical waveguide is less than the diameter of a human hair. Because copper cable is usually larger, stiffer and heavier than a fiber cable carrying the same quantity of signals, installation, ducting, and handling costs can be much lower for a fiber installation than for a similar coaxial system. An optical fiber cable may be the only alternative for circuit capacity expansion when ducts are full of copper. Many local telephone companies now install fiber on all new construction except for subscriber loops.

Transmission System Configurations

Fiber optic transmission circuits can be configured in a variety of ways (Fig. 4):

1. The signal at the interface may be either analog or digital.
2. The number of independent electrical signals which can be transmitted on a fiber may be one (simplex) or many (multiplexed).

Independent electrical signals may be combined (multiplexed) into one complex signal for optical transmission in various ways.

1. Frequency division multiplex (FDM) using analog AM and/or FM carriers.
2. Time division multiplex (TDM) which for digital signals means sampling the data at a high rate, converting the samples to digital codes, and interleaving the codes into time slots.

DECADE	ANALOG SYSTEMS	DIGITAL SYSTEMS
1950s	Development of gas laser light sources.	Early research in transmission of digitized sound signals.
1960s	First implementations of fiber optic transmission system principles.	Early efforts to develop multiplexed systems for digitized voice and data transmission.
1970s	First practical systems for analog signal transmission over fiber, through intensity (amplitude) modulation of light source by single baseband video signal combined with FM subcarriers for ancillary sound and data channels.	Common carrier industry implementation of DS() hierarchy of transmission services for digitized voice and data. First long-haul fiber optic transmission systems.
1980s	1984 - Introduction of single television channel per fiber (STVC/F) transmission systems with baseband video and multiplexed audio FM'd onto nominal 30 MHz carrier. 1984 - Introduction of short-distance multi-fiber RGB graphics systems. 1988 - Introduction of proprietary multiple television channel per fiber(MTVC/F) systems.	Rapid deployment of fiber in common carrier long haul circuits and IEC trunks. Sporadic utilization of DS3 services by broadcast industry. Utilization of sub-DS1 services in videoconferencing. First implementations of STVC/F and MTVC/F systems for full bandwidth NTSC signal transmission.
1990s	Rapid deployment of MTVC/F trunking and distribution networks in local cable systems. Continuing desultory sales of STVC/F systems in private networks. Sporadic broadcast industry utilization of MTVC/F Metropolitan Area Networks (MANs) for back haul of news and local remotes.	First implementations of B-ISDN and SONET systems for transmission of full bandwidth NTSC television signals and high resolution graphics and HDTV signals.
Early 21st Century	Analog transmission systems become historical curiosities.	User setup and use of digital common carrier networks for image transmission become as simple and cheap as DDD phone calls. (HAH!)

Figure 2. Chronology of development of fiber optic systems for video and ancillary sound and data transmission.

Transmission of multiple FDM and TDM signals onto a single fiber is accomplished by optical light wavelength division multiplexing (WDM) which is the coupling of several modulated light sources of discretely different wavelengths into a fiber by fusing the light source output "pigtails" together. The principles of WDM are explained in more detail later in this chapter.

Transmission Channel Options

Four transmission channel types can be created by combining several options. A channel denotes the complex electrical signal appearing at the modulator input and detector output of a single optical transmission circuit.

Single Analog Channel (SAC)

An example of a SAC involved one of the first uses of fiber in broadcasting which occurred in 1980 during the winter Olympics when "television signals" were transported from remote-site production trucks to the ABC Network broadcast center in Lake Placid, New

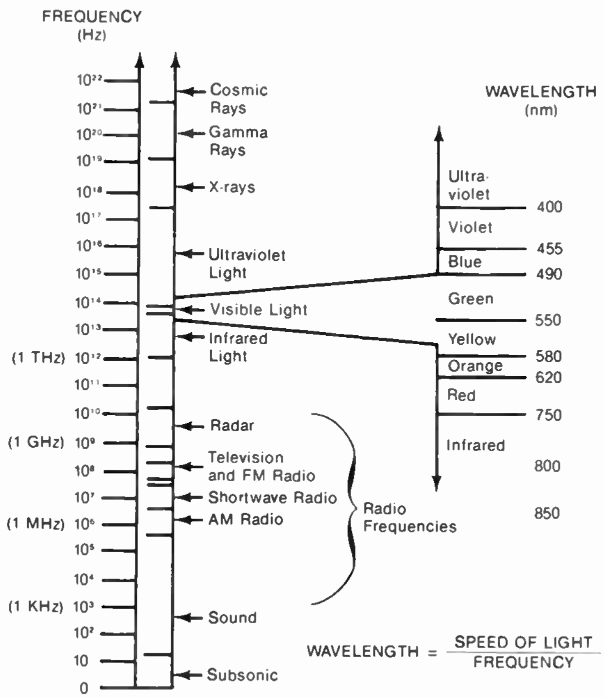


Figure 3. Electromagnetic spectrum.

ANALOG SIGNAL DOMAIN	DIGITAL SIGNAL DOMAIN
Composite Analog (NTSC, PAL, SECAM, PAL-M)	Composite Digital (D-2, D-3, ???)
Component Analog (RGB, Y/R-Y/B-Y, Y/C)	Component Digital (D-1 - Y/R-Y/B-Y, 8 bits, to R/G/B/Alpha, 10 bits)

Television video signal format options
(Plus encrypted versions of each)

Single signal group per fiber (FM/FM and AM/FM)	Single signal group per fiber (TDM)
Multiple signal groups per fiber (FDM/FM and FDM/AM)	Multiple signal groups per fiber (TDM/TDM)

Fiber Transmission Circuit Signal Group Options

Figure 4. Fiber transmission system design alternatives.

York, each on a single fiber. The television signal consisted of two baseband signals, 4.2 MHz NTSC video and its accompanying 15 kHz audio on a frequency modulated 15 MHz carrier resistively mixed with the amplitude (intensity) modulated video.

The resultant complex video signal voltage then intensity modulated a laser or light emitting diode (LED), which was then coupled into a glass fiber for transmission. Intensity modulation of a coherent light source results in the same transmission system noise shortcomings as are well known in AM transmission systems.

In 1984, higher performance FM transmission systems were introduced to fiber optics. The modulating signal consisted of baseband video (4.2 MHz bandwidth in NTSC systems, 5.5 MHz in PAL), plus one, two, or more audio and data channels modulated onto carriers in the spectrum from 5 to 20 MHz. The transmission carrier frequency was established in the 30 to 35 MHz spectrum. In some FM systems the transmitted signal consists of multiple modulated voice and data signals and television signals modulated on the same 70 MHz intermediate frequency (IF) often used on microwave transmission systems.

There are currently no interface standards for carrier and subcarrier frequencies, modulation indices, or other transmission parameters. As a result, transmitter and receiver terminals made by different manufacturers will work together only if specified to do so by system designers.

In television broadcasting applications of SAC systems, the transmitted television signal group may include two-channel stereo and auxiliary data channels on FM subcarriers above baseband video. In audio and recording industry applications, the signal can consist of a single or dozens of multiplexed analog or digital microphone or tape recorder output channels.

Fiber is well suited for handling component video for which identical characteristics of the transmission system is required for critical graphic and editing applications. The sheath for a dozen fibers is actually smaller than a single typical precision coaxial cable.

Multiple Analog Channel (MAC)

Multiple analog channels, each consisting of single signals or signal groups, may be transported on a single fiber by either frequency division or wavelength division multiplexing (FDM or WDM). Multiplexing in the electrical frequency domain is more common than in the light frequency (wavelength) domain because glass fibers have substantially different attenuation versus distance characteristics for each of the three most common optical wavelengths used in fiber-based transmission systems.

Some multiple analog channel systems use widely-separated carrier frequencies to transport component video (primary R, G, and B or Y, C_r, and C_b) components of an NTSC, HDTV, or other high-resolution, wide-screen color signal, or the analog outputs of a digital computer graphics system.

Single Digital Channel

At one time, the only circumstance that justified the installation of single analog channel systems within a television facility was the need to transport signals through high RFI environments. Digitizing the signal was not necessary for these relatively short runs to achieve noise-free transmission or to maintain signal quality.

Recently, the single digital channel option appears to be an economically justifiable approach for distributing video and audio groups inside broadcasting and post-production plants. Current generation signal processing and storage devices are now available with serializer chip options which move all signals through one high-speed input/output port. Digital routing of this signal may therefore be both economically and technically acceptable. Multiplexing the video and its accompanying audio and data signals into a single bit stream may offer significant cost savings as well as technical transparency, by use of a single level routing switcher with add/drop multiplexing.

Fiber-based, multiplexed digital routing appears to be an attractive alternative to coaxial cable in any plant dealing with high-resolution, wide-screen video and eight or more audio channels.

HIERARCHY	DIGITAL THROUGHPUT (bps)	VOICE CIRCUIT CAPACITY	TV CHANNEL CAPACITY	
			VIDEO COMPRESSION RATIO	PROGRAM AUDIO CHANNELS
Telephone Industry				
DDS-0 (T0)	56-64 k	1	2045:1	0
"Sub-T1"	384 k	--	298:1	--
"Sub T1"	768 k	--	149:1	--
DS-1 (T1)	1.544 M	24	74:1	0
DS-3 (T3)	44.734 M	672 (28 T1)	2.56:1	2
Television Industry				
D-2, D-3 (Composite)	140 M 4 audio chans	--	0	4
D-1 (Component)	216-273 M 4 audio chans	--	0	4

Figure 5. Telephone industry hierarchy of tariffed DS() transmission services, compared to television industry digital recording standards.

Multiple Digital Channel

Only a modest need for multiple digital channel systems is foreseen within a broadcast or post-production plant. As in the case of single channel per fiber digital systems, there appears to be limited technical and cost justification for digital transmission of multiple broadband video signals over a single fiber within a campus, or over a public or private network circuit less than a nominal 30 to 50 kilometers long. Digital routing of multiple signal groups requires reconfiguration of the cable network from a star to a ring. The configuration and its control are well known in the data processing industry, where the networks are called local area networks.

In the fiber-based public switched network, transportation of digitized broadband video/audio/data signal groups has become an attractive alternative to current terrestrial and satellite microwave services. The eventual availability of affordable, high-speed, broadband, user-configurable transmission networks will foster the use of multiple signal group fiber transmission in metropolitan areas as well as in long-haul applications. Technically, the delivered signal quality is expected to be better than that achievable in multi-signal-group analog systems, and there appears to be no practical distance limitation.

Video bandwidth compression. Two techniques for video bandwidth compression are now in use. Both digitize and compress the NTSC color video by a ratio between approximately 3:1 to over 2,000:1 as illustrated in Fig 5. One such technique creates a serial digital signal for a single television signal group, which can be interconnected to a telco modem or directly into a digital subscriber loop. Ancillary audio and data can be integrated with the video signal in the lower ratio compression systems. Audio and data must be transmitted over separate telephone circuits in the high ratio compression systems.

In the broadcasting industry, long-haul transmission of NTSC video/stereo audio signal groups over telco DS-3 (with a 2.75:1 bit rate compression) and satellite carriers are now available and routinely used, and accepted as good quality, by broadcasters and other television production organizations.

Broadband Digital Transmission Services

The Existing T and DS Series Hierarchy

Means for digital transmission of analog signals were originally developed by the Bell System decades ago to multiplex groups of voice bandwidth (30 to 3,000 Hz) signals in a TDM configuration onto twisted copper pairs in trunking cables. The TDM systems offered advantages over FDM systems in that substantially more channels could be carried on each pair, and the transmitted voice quality was dramatically improved by elimination of IMD and reduction in noise. Fiber optic transmission systems are ideally suited for digital wide-band transmission circuits.

The first level in the hierarchy of transmission speeds, offered by common carriers, 64 kbps, is synonymously referred to as T-0 (T-zero) and DS-0. This rate accommodates one 56 kbps digitized voice channel, which is created by sampling the voice signal at a 4 kHz rate and digitizing the sample to seven-bit resolution and some overhead housekeeping data.

Another level, popularly known as the Bell System T-1 Carrier (DS-1), has a digital signalling rate of 1.544 Mbps. It has a capacity of 24 voice channels (plus overhead) and can be used for a single program audio channel and stereo with newly developed digital compression schemes.

The only other widely tariffed level in this hierarchy available to broadcasters is DS-3 (T-3), with a signalling rate of 44.734 Mbps (commonly called 45 megabit). It has a capacity of 28 T-1 channels, for a total of 672 voice channels or one (compressed) television channel.

Interfaces to T-1 and T-3 transmission services can be provided within a broadcast plant by several competing common carriers.

T-3 transmissions were initially deemed "unacceptable" by most broadcasters until new digital compression algorithms became available. The visual accept-ability of any T-3 circuit transmission is established by the algorithm used to eliminate redundant bits from the digitized video signal. Different algorithms are used by various manufacturers. There is intense competition among codec developers, who keep their algorithms as closely guarded secrets. The future of T-3 transmission will depend largely upon the agreement of an algorithm interface standard which allows competing manufacturers codecs to "talk to each other."

Synchronous Optical Network (SONET)

In recognition of the wideband nature of optical transmission systems, the lowest level optical carrier (SONET OC-1) has a speed of 51.84 Mbps. Hardware already exists for converting a 45 Mbps T-3 signal into an STS-1 (Synchronous Transport Signal level 1) signal for transportation on the OC-1 carrier. Higher levels of the hierarchy of interest to broadcasters and HDTV users are OC-3 at 155.52 Mbps (accommodating a full bandwidth uncompressed digital NTSC video with 4-channel audio signal) and OC-12 at 622.08 Mbps (accommodating digitized high resolution images and accompanying sound with little or no compression).

FIBER OPTIC TECHNOLOGY PRINCIPLES

Lightwave Transmission Theory

The heart of the system is an optical fiber (also referred to as an optical waveguide) made of high-purity glass or plastic through which light from a solid-state laser or light-emitting diode (LED) is transmitted. Light, being a part of the electromagnetic spectrum, follows the same principles employed in designing microwave transmission systems and television and radio transmitters.

A basic transmission link consists of an optical transmitter terminal, a continuous length of fiber from a few meters to as long as 30 to 50 kilometers, and an optical receiver terminal (Fig. 6). The transmitter optical source output (light power of the order of a milliwatt) intensity or wavelength is controlled by a driver circuit, whose input is a varying voltage signal (modulation) which may be analog or digital, with frequencies ranging up to a gigabit/second (Gbps) and beyond. The modulating voltage may represent one or many individual signals. The transmitter also contains circuits to condition the signals as required for best end-to-end circuit performance, and a multiplexer which combines them into one electrical signal with the pre-emphasis characteristics required for optimum optical transmission circuit performance.

The receiver contains a light-sensitive solid-state detector (photo-diode) which converts the received light modulations back into an electrical signal, and output circuits which amplify, demodulate, and reconstruct the signal into its original form.

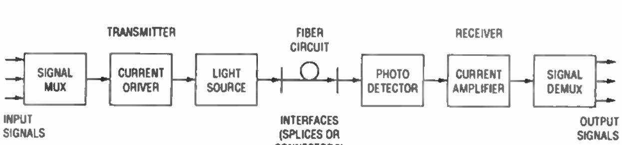


Figure 6. Basic fiber optic transmission link.

Light And Fiber Characteristics

Light is similar to radio waves, x-rays, and gamma rays in that it is a part of the electromagnetic spectrum shown in Fig. 3. In general, the frequencies of light used in fiber-optic transmission are in the region of 300–400 terahertz (10^{12} Hz or 1,000 GHz), or several orders of magnitude higher on the electromagnetic spectrum than the highest frequency radio waves. Light waves are more generally described in terms of wavelength rather than frequency. Certain wavelengths in the range of 800–1600 nanometers propagate most efficiently through the fibers that are currently available. The speed of light (300,000 km/sec) is the velocity of any electromagnetic energy in free space or a vacuum. Light travels slower in all other media, and different wavelengths travel at different speeds in any single medium. When an electromagnetic wave crosses a boundary from one medium to the next it changes speed which results in a change of path called *refraction*.

The index of refraction (n) is a dimensionless number that expresses the ratio of the velocity of light in free space (c) to its velocity (v) in a specific medium. Thus:

$$n = c/v$$

Refraction of a ray of light as it passes from one material to the next depends on the refractive index of the material. The three terms used to describe refraction as shown in Fig. 7 are (1) *normal* (the line perpendicular to the interface of the materials), (2) the *angle of incidence* (the angle between the incoming ray and normal), and (3) the *angle of refraction* (the angle between normal and the refracted ray).

When light passes from a high index of refraction medium to a lower one, the light is refracted toward the normal. As the angle of incidence increases, the angle of refraction approaches 90 degrees with normal. This is called the critical angle. If the angle is increased past the critical angle the light is totally internally reflected.

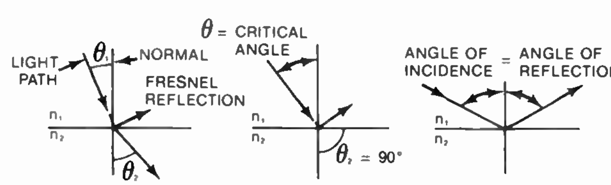


Figure 7. Lightwave refraction principles. When the refractive index of medium n_1 (the core) is less than that of medium n_2 (the cladding), light incident on the boundary at an angle less than the critical angle θ propagates through the boundary, but is refracted away from the normal. When the angle of incidence is greater than the critical, the light is totally reflected.

Optical Fibers

An optical fiber is made of glass (which may be doped with other materials to meet certain requirements) and has two concentric components as shown in Fig. 8. The inner core has a refractive index higher than the outer cladding. A light ray injected into the core at an angle will eventually hit the boundary between the core and the cladding. Rays arriving at an angle greater than the critical angle are reflected back into the core. These rays propagate (bounce their way) down the fiber until they emerge at the receiver terminal. Their time of arrival is obviously later than the ray which propagates down the center of the core.

Light striking the interface at less than the critical angle passes into the cladding and is lost. The total of the reflected rays constitutes the carrier of the information modulated on it. The exact characteristics of the bundle of light rays propagated in a fiber are determined by its size, construction, and composition. Maxwell's equations show that light does not travel randomly through a fiber; rather, it is channeled into modes. A mode is the path of one ray through a fiber.

Dispersion

Dispersion refers to the spreading of a light pulse as it travels down a fiber. A pulse measured at the output will be wider than it was at the input. This limits a fiber's bandwidth or information carrying capacity. Pulse rates must be slow enough that dispersion will not cause adjacent pulses to overlap. Consider the pulse shown in Fig. 9. Two trains of pulses, one faster than the other are injected into the same fiber. In both cases, the pulses are spread by dispersion. For the slower train, the interval between pulses is sufficient to allow each pulse to be distinguished. In the faster train, however, the interval is so short that individual pulses merge into one long indistinguishable pulse.

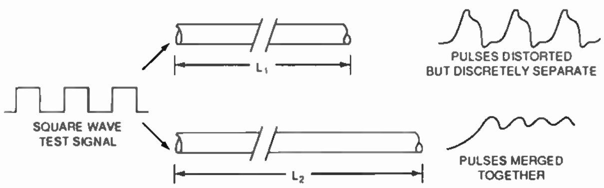


Figure 9. Lightwave pulse propagation principles. Modal and chromatic dispersion characteristics of a fiber determine the maximum frequency of modulating signal which can be recovered at the end of a fiber of any given length.

Dispersion is the limiting factor in determining a fiber's bandwidth. Modal dispersion results from differing path lengths of the multiple rays (modes) in a multimode fiber. Chromatic dispersion limits the bandwidth which can be propagated in a single mode fiber because no laser light source used in fiber transmission systems emits light at a precise single wavelength. A special class of long wavelength (distributed feedback) lasers with a very narrow output spectrum is available for use on single mode fibers of 50 kilometers or more.

This relationship between bandwidth and distance is uniquely determined in each fiber by the core diameter, step or index core grading and material purity. The figure of merit for any given fiber is its bandwidth-distance rating and specified in MHz-km. The product of the bandwidth to be transmitted and the length of the optical circuit must be less than the figure of merit number.

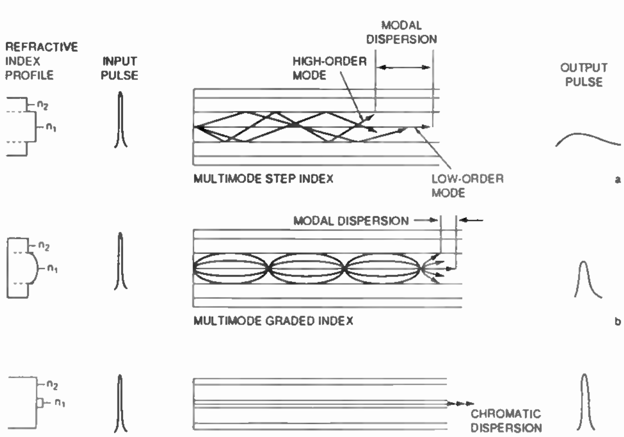


Figure 8. Optical fiber types. The core diameter and its refractive index characteristics determine the light propagation path or paths within the fiber core.

Types of Fibers

Of the many ways to classify fibers, the most informative is by refractive index profile and number of modes supported. The two main types of index profiles are *step-index* and *graded-index*. In a step-index fiber, the core has a uniform index with a sharp change at the boundary of the cladding. In a graded-index fiber, the core's index is not uniform. Instead it is highest at the center and decreases until it matches the cladding. There is no sharp break (Fig. 8b).

Step-index multimode fiber. (Fig. 8a) A multimode step-index fiber typically has a core diameter in the 50-1,000 micron range. This relatively large core permits many modes of propagation. Since light will reflect differently for different modes, the path of each ray is a different length. The lowest order mode travels down the center while higher order modes strike the core-cladding interface at angles near the critical angle. As a result, a narrow pulse of light spreads out as it travels through this type of fiber. This spreading is called *modal dispersion*. This type of fiber is lower in cost and can be used only for limited bandwidth applications.

Step-index single mode fiber. (Fig. 8c) Modal dispersion can be reduced by making the fiber core small, typically 5 to 10 microns (1/6 the diameter of a human hair). At this diameter, only one mode propagates efficiently. The small core size, however, makes fiber splicing quite difficult. Single mode propagation is the most suitable for high speed, long distance transmission. This type of fiber is more expensive but the advantages far outweigh the additional cost.

Graded-index multimode fiber. (Fig. 8b) A graded-index fiber also limits modal dispersion. The core is essentially a series of concentric rings, each with a lower refractive index. Since light travels faster in a lower index medium, light farther out from the axis travels faster. High-order modes have a faster average velocity than low-order modes, thus all modes tend to arrive at a point at nearly the same time. Rays of light are not sharply reflected by the core-cladding interface and instead are refracted successively by differing layers in the core. This fiber is a compromise between bandwidth, cost, and attenuation.

Attenuation

Attenuation means loss of power. During transit some of the light is absorbed into the fiber or scattered by impurities. Attenuation for an optical fiber is usually specified in decibels per kilometer (dB/km). For commercially available fibers, attenuation ranges from less than 0.5 dB/km for premium glass fibers to 1,000 db/km for large core plastic fibers. Emitted light is measured as power, thus 3 dB represents a doubling or halving of any given power level.

Attenuation and light wavelength are also uniquely related in fiber transmission systems (Fig. 9). Most fibers have a medium loss region in the 800 to 900 nanometer (nm) wavelength range (3 to 5 dB/km), a low loss region in the 1,150 to 1,350 nm range (0.6 to 1.5 dB/km), and a very low loss region (less than 0.5 dB/km) around 1,550 nm. Best performance can be achieved by careful balancing of fiber type, light source wavelength, and distance requirements.

Fiber Transmission System Components

Light Sources (Transmitters)

Light sources in fiber transmission systems are either a light emitting diode (LED) or an injection laser diode (ILD).

The LED is an incoherent source, compared to the ILD. It is characterized by output power levels well below one milliwatt, in the order of −10 to −30 dBm. LEDs operate at slower speeds, but also cost considerably less than ILDs. LEDs are quite suitable for applications requiring transmission over less than 10 km and modulation a bandwidth of less than 100 Mbps per second. LEDs operate at wavelengths of 840 to 1,300 nm.

An ILD light source requires more electronics to operate, but is more powerful, with an output of 0 dBm to 10 dBm (and higher for very specific applications). Upper power limits are a matter of device cost and permissible radiation levels. ILDs cost substantially more than LEDs. However, they operate in the low-attenuation 1,300 and 1,550 nm wavelength range. The combination of low attenuation and high modulation frequency limits make them cost-effective as sources in high capacity, long distance, transmission systems.

Selecting a light source for a fiber transmission system requires evaluation to insure that the modulation frequency limit is greater than the bandwidth to be transmitted, and enough power at the desired wavelength is available to satisfy the distance and signal-to-noise requirements. Another important consideration is that optical devices do not turn on with the same characteristics as customary electrical and RF devices. Fig. 10 and Fig. 11 show characteristics of typical lasers and LEDs.

Light Detectors (Receivers)

The light detector performs a complementary function to the light source, that is, it converts incident light energy back to electrical energy. Detectors in fiber transmission systems are positive intrinsic negative (PIN) or avalanche photo diode (APD) semiconductors. The limiting sensitivity (threshold) in detecting weak incoming signals determines link performance. Light detectors have considerable higher noise figures than standard RF devices, which is a major limiting factor in designing transmission systems.

In an ideal PIN diode, each incident photon creates an electron-hole pair in the semiconductor lattice which, in turn, sets one electron flowing in the external circuit. If the received light is weak, the generated current may not be strong enough to overcome noise inherent in the diode and receiver circuit. In such cases, it is desirable to increase the detector output before amplification by the receiver. Such gain is generated in an avalanche photo diode. A reverse bias adds several electron volts of energy to each liberated electron to provide amplification of the received signal, which is not linear with respect to the received light levels.

Repeaters

The maximum acceptable length of a fiber transmission circuit is dependent upon (1) the signal bandwidth to be transported, (2) the transmitter output power, (3) the attenuation of circuit fibers and interconnection

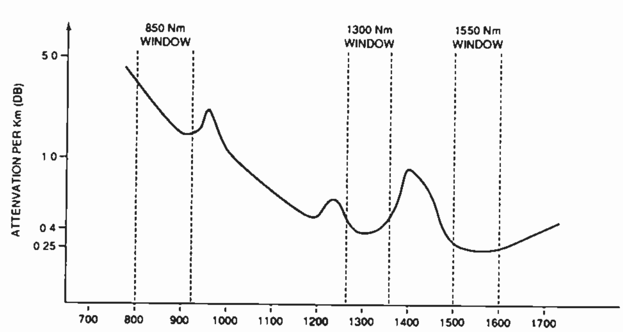


Figure 10. Fiber attenuation versus light wavelength characteristics. Attenuation has been reduced steadily in the last two decades through improved fiber drawing techniques and reduction in impurities. It has now approached the theoretical limits of silica-based glass at the 1,300 and 1,550 nm wavelengths.

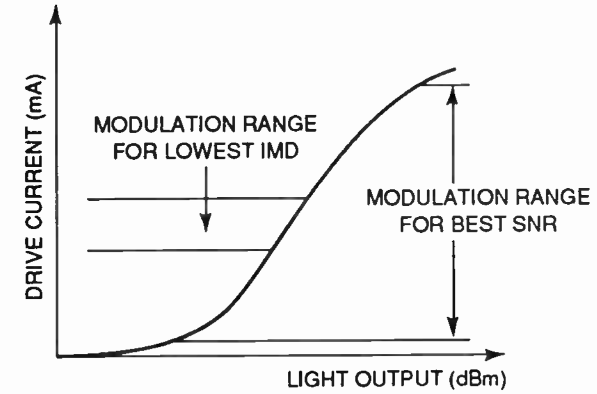


Figure 11. LED drive current/light output transfer characteristic. A restricted modulation index is necessary in analog signal modulation to limit harmonic and intermodulation distortion products.

components, (4) the modal and chromatic dispersion characteristics of the fiber, (5) the threshold sensitivity of the receiver, and (6) the end-to-end transmission circuit performance specifications set by the user. Analog circuits 50 kilometers long are now possible. Digital circuits with good error detection and correction coding schemes can operate for perhaps double this distance.

First generation analog and digital repeaters developed to extend transmission distances operate at electrical baseband. The repeater terminal is literally a receiver whose demodulator is hardwired to the modulator of a transmitter. Analog-domain repeaters of this type have a limit of four or five tandem sections and a transmission maximum of 150 to 200 km. Digital-domain repeaters (regenerators) have virtually no limit, as long as transmission circuit errors are detected and corrected at each regenerator prior to generating the new pulse.

Second generation repeaters increase the optical signal power in the light domain (Fig. 12). Optical amplifiers have no direct equivalent in electrical signal transmission (except a rough approximation to a parametric amplifier). They are nominal 8-meter lengths of Erbium-doped fiber inserted in the optical transmission path. The 10 to 50 mW output of a pump laser operating at 980 or 1,490 nm is coupled into the fiber via a passive coupler. Its interaction with the optical carrier results in the production of a new carrier modulated with the original information but operating at a nominal 1,550 nm wavelength.

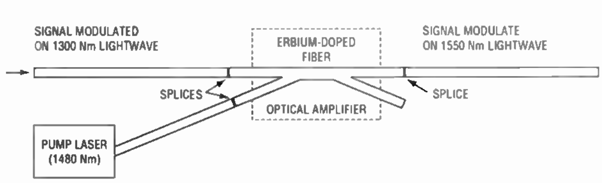


Figure 12. Block diagram of first-generation optical amplifier ("repeater").

This approach presents a problem because most of the fiber in the long-haul public network is optimized for 1,300 nm carrier transmission. The circuit performance payoff, light power increases of the order of 20 to 30 dB, has encouraged researchers to find solutions to this problem.

Switching

First generation fiber transmission circuit switchers operate at electrical baseband. Efforts to develop low-cost, efficient optical switchers are underway in telecommunications and computer industry advanced research laboratories.

The advantages of optical transmission and switching are not confined to broadband video image transmission. Optical domain interconnection of computer components is necessary to develop the next generation of super high-speed, large scale parallel processor number crunching engines.

Video And Audio Transmission Techniques

Modulation Techniques

Information can be modulated onto an optical carrier in three ways: (1) intensity or amplitude modulation (IM); (2) frequency modulation (FM); and (3) digital pulse code modulation (PCM).

Intensity modulation. Intensity modulation (IM) is the simplest form of modulation. The instantaneous amplitude of the input signal directly controls the output intensity of the light source. The peak white maximum amplitude of a video signal would result in the most output light and the tip of sync the least. As shown in Figs. 10 and 11, optical devices do not turn on in a linear fashion, which results in the introduction of intermodulation (IMD) into the video signal. Attenuating the input of the modulator to reduce distortion lowers signal-to-noise (SNR) which is directly related to the optical power. Further, since most optical devices do not completely turn-off (particularly lasers) there is always some residual signal (noise) present. AM systems therefore, while simple and relatively low cost, are not normally employed in applications where high performance is required.

Frequency modulation. Frequency modulation overcomes the SNR and IMD limitations of intensity modulation. FM is more complex and costly than AM, but transfers the problems of signal quality from the optical domain to the electrical domain. Ancillary audio and data signals can easily be multiplexed over the video baseband on separate FM subcarriers. The video/audio signal group is then frequency modulated on a carrier in the 30 MHz range or at the common 70 MHz IF, to produce a single signal with a fixed carrier amplitude which can use the linear portion of the light emitter output power.

Good SNR and IMD performance can be achieved through the use of a frequency modulated optical light source.

PCM or digital modulation. A fiber transmission system is most transparent to the signals being transported when it is operated in a pulse code modulation (PCM) mode. The wide bandwidth of optical circuits makes it especially suitable for transmission of information in the digital domain where extremely wide bandwidths are required compared to the same signal in the analog domain. Digital transmission is particularly suited to optical circuits because of the wide bandwidth and the stability of the propagation characteristics of the optical circuit components. Thus, digital modulation is the preferred choice of optical transmission system designers.

In evaluating the technical merits of transmitting television signal groups on fiber, the performance benchmark to which digitized video transmission must be compared is the analog transmission specifications in the EIA/TIA-250-C document (see Chapter 4.1, "TV Signal Transmission Standards"). The short-haul specifications call for 67 dB SNR, 1% differential gain and 2 degrees differential phase.

The SMPTE D-2 standard for digitizing an NTSC composite analog video signal has eight-bit resolution that results in a maximum of 56 dB SNR when the signal is converted back to analog. Nine-bit sampling yields 60 dB SNR, and ten-bit sampling provides 64 dB. The minimum bandwidth for transmitting a D-2 format digital signal (including four digitized 20 kHz program audio signals) is 140 Mbps. Transmitting a digital component video signal and eight program audio channels produces a rate of 300 Mbps.

Ten-bit sampling is required to meet the more stringent transmission standards for short-haul and intraplant circuits.

Multiplexing Techniques

Frequency Division Multiplexing (FDM)

The technique of summing multiple AM or FM carriers is widely used in coaxial cable distribution. However, the nonlinearity of intensity modulated optical devices result in substantial and often unacceptable intermodulation distortion in the delivered signal. Wide and selective spacing of the FDM carriers can significantly reduce this problem.

Wavelength Division Multiplexing (WDM)

This multiplexing technique reduces the number of optical fibers required to meet a specific transmission requirement. Two or more complete and independent optical transmission circuits operating at different optical wavelengths can be transported over a single fiber by combining them in a passive optical multiplexer. This is an assembly in which the glass pigtails from multiple optical transmitters operating at different wavelengths are fused together and spliced into the transporting fiber. Demultiplexing these optical signals at the receiver end of the circuit is accomplished in an opposite-oriented passive optical multiplexer. The pigtails are coupled into photodetectors through wavelength-selective optical filters.

A loose set of ground rules applies to the coupling of light sources at many different wavelengths. A given fiber might have two lightwave circuits of different wavelengths travelling in opposite directions. Three circuits can be transported without receiver optical filtering by using light sources from the 840, 1,300, and 1,550 nm wavelength windows. On the other hand, three discretely different wavelengths in the 1,300 nm range have been used in some component video transmission systems. The light sources were selected at nominal 1,280, 1,300, and 1,320 nm wavelengths, with sharply tuned optical filters providing received end separation. Audio signals were multiplexed on a separate optical carrier.

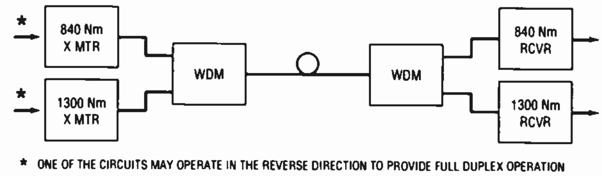


Figure 13. Wave Division Multiplexers (WDMs) are passive optical assemblies created by fusing optical fiber pigtails together.

Several considerations are important in designing WDM systems:

1. Adequate separation (measured in dB) of each optical channel is required because optical crosstalk between channels lowers the effective SNR of the link.
2. Passive multiplexers and optical filters produce some optical attenuation (1-5 db per device).
3. Optical light sources change output wavelength with changes in temperature so the tuned bandwidth of the WDM components must be wide enough to accommodate this drift.

Since satisfactory optical isolation is costly and technically difficult to achieve, WDM techniques are only appropriate for FM and digital systems whose carrier-to-noise characteristics can accommodate the power losses that occur in the WDM system.

System Design Considerations

System Performance Requirements

Designing a fiber transmission system begins with establishment of its end-to-end performance requirements. For video this can range from a very high performance entrance link to a lower performance, remotely located security camera.

Once the system performance of a link is defined, the optical loss budget must be calculated. Items in the loss budget begin with the transmission circuit length, multiplied by the attenuation per kilometer of the selected fiber at the transmission wavelength selected. Two other loss budget items are most important to include: (1) loss allowances for each splice and mechanical connection to be made over the length of the optical circuit, and (2) a safety factor of up to 6 dB to accommodate power losses resulting from optical device aging and splices that may be required during repairs of the fiber and for the possible future addition of WDM components.

These calculations define the output power from the optical transmitter and the receiver sensitivity needed to meet or exceed end-to-end signal transmission specifications.

Testing Optical Circuits

Two testing techniques can be used to evaluate the physical condition and losses in an existing fiber circuit. An optical time domain reflectometer (OTDR) may be used to measure attenuation and identify location of breaks and bad splices in a fiber circuit. The OTDR operates on the principle of transmitting pulses of light and measuring the return time of all back-scattered (reflected) light from each pulse (very much as a standard TDR is used for RF transmission components). This instrument provides a graphical representation of the condition of the fiber as a function of distance. It can literally pinpoint a break or flaw in a fiber to approximately a meter even over ranges of tens of kilometers.

By selecting a test light wavelength not in use in the fiber, the testing can actually be performed while the circuit is in use. This testing is important in the evaluation of circuits which are candidates for WDM capacity expansion. It is also useful prior to specifying new transmission systems using existing fiber. The testing will uncover routings, splices, and damage which may not be documented or evident in visual inspections.

An optical power meter can be used to measure the end-to-end attenuation of new construction, once the optical transmitter is on line, or routinely to check the condition of the fiber. Connecting the meter to the light source through an optical patch cord accounts for the first connector interface loss, and establishes the actual power to be coupled into the fiber circuit. Connecting it to the receiver end of the fiber verifies that the loss budget has been calculated accurately. A measured loss exceeding the calculated loss by more than one dB should initiate further tests to find the reason for the difference.

Fiber Cable Specifications

Specifying cable for new fiber circuits should be done with inputs from competing fiber cable vendors. Step-index multimode fibers may cost approximately the same as higher performance multimode or single mode fibers, but may not offer the same bandwidth and distance advantages. Larger core multimode fibers are easier to splice and attach connectors. In general, for circuit lengths of more than approximately 25 km, single mode fiber must be used. Selection of a cable type is defined by three practical system design considerations.

1. *Attenuation*. This is a straightforward mathematical exercise specified in terms of dB/km. If a cable has a specified attenuation of 3 dB/km for a given light wavelength, 4 kilometers will attenuate 12 dB of light power. Short wavelengths (800 to 900 nm) are used for short (nominal 5 to 10 km length) circuits.
2. *Bandwidth*. This consideration applies to multimode cable selection for use in FM and digital systems. An FM system should have sufficient bandwidth to pass the upper sidebands of the FM carrier. Optical devices in digital systems must have optical intensity rise and fall times which do not adversely affect the pulse shapes at the bit rate to be transmitted.
3. *Construction*. Specific applications require specific types of cable. For example, for outside use UV resistant jackets are required. Similarly, for direct burial, rodent-proof jackets are required. In general, cables for practically every application and environment are available.
4. *Connectors*. There are several standard connectors for fiber optic cables. The most popular are the ST and ST-II (ST and ST-II are registered trademarks of AT&T) which are constructed similarly to a BNC coaxial cable connector. While fiber connectors are readily available from several manufacturers, installing the connectors is not as simple or inexpensive as mounting a BNC connector on a video cable.

APPLICATIONS

Switched Fiber Systems for News Gathering

C&P Telephone in Washington, D.C. has a broadcaster-controlled switched fiber optic network service that makes the entire metropolitan Washington service area an extension of each participating news bureau's studio, with live on-air switching controlled by the newscast director. Its startup was not a direct reaction to broadcaster demand, however. Point-to-point fiber circuits were installed beginning in 1983, as the only practical solution to C&P problems of congested electronic news gathering (ENG) microwave paths and full coax cable ducts. Financial motivation for expanding and upgrading the copper coax plant with new technology fiber then appeared from two sources. First, their various rate-controlling public utilities commissions (PUCs) were pushing C&P for rate reductions, and were negative about approving rates based on installation of old technology 16 PSV video cable. Second, this cable could only be used for video signal transmission, whereas multi-strand fiber cable could carry a readily changeable mix of high-speed digitized voice/data traffic or analog video signals with multiplexed baseband voice and data signals.

The C&P fiber news gathering (FNG) service was born during the 1984 presidential election and 1985 inaugural activities, with transmissions of ABC and NBC originations from the Capitol to AT&T Long Lines. By the time of the Reagan/Gorbachev summit in December 1987, multiple fiber cables were in place from most of the active news sites to control rooms in the Washington News Bureaus of scores of domestic and international news organizations, patched through the C&P Radio and Television Center at 13th Street, NW.

The C&P solution to these constantly increasing backhaul circuit needs is a user-controlled broadband switched fiber plant now totalling about 1.7 million circuit miles of fiber in the C&P Metro service area. A minimum 12-fiber cable is pulled into every new installation. Examples include 40 circuits into the Capitol, 48 into the White House, 48 to the Washington International Teleport and up to 24 fibers each into the midtown area news bureaus. Each cable runs from an equipment rack in a toll operating center (TOC) to a termination panel on the customer's premises. From this panel the customer's coax cable and shielded twisted pairs carry the baseband video and audio and any ancillary data signals to or from switching equipment elsewhere in the facility.

Remote Studio Interconnection

KSL-TV Channel 5 produces television broadcasts originating in the Mormon Tabernacle in Tabernacle Square, Salt Lake City. KSL-TV provides equipment, transmission circuits, and people to transport live video and audio to a control room 2 1/2 blocks away. There it is edited and fed to the uplink or tape.

In addition to providing the video circuits, the station eliminated all of the many technical problems of transporting high quality channels this distance by installing a digital audio transmission system on two fibers in a cable connecting the two locations (Fig. 13). The system digitally integrates up to ten 20 Hz to 20 kHz program audio signals into a single 18.432 Mbps bit stream. Optional transmission circuit interface modules process this signal for transport on a standard 6 MHz bandwidth television channel over either coaxial cable or fiber. Fiber was the choice in this case because of the distance requirement.

Operation with up to 20-bit digitization provides well over 100 dB SNR, and the digitized audio signal is processed in the AES/EBU 32-bit digital format. The digital distribution channel is free of crosstalk.

Fiber Optic Circuit Designs

Fiber optic transmission systems are available in a variety of configurations. These include amplitude

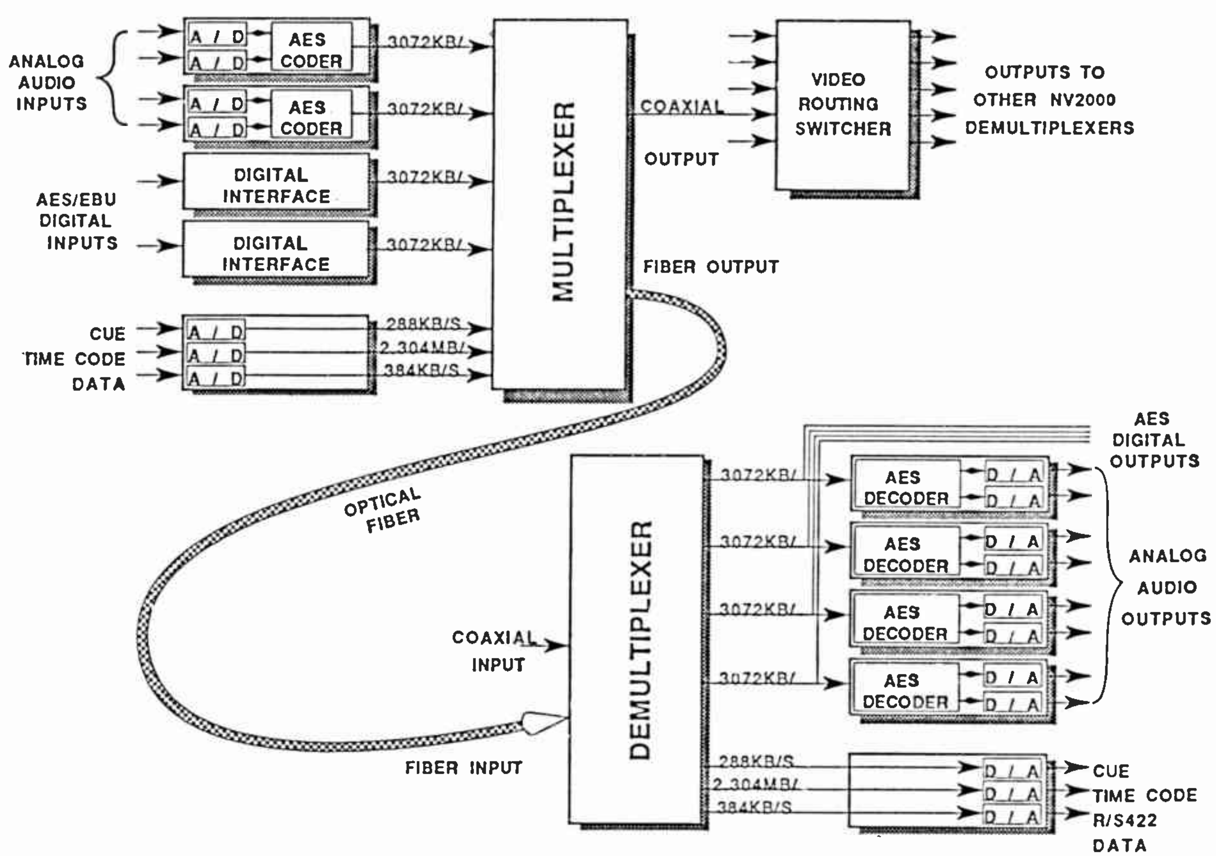


Figure 14. Block diagram of the NVision 2000 system which digitally multiplexes and modulates analog and digital program audio, and cue, time code and data channels, into a 6 MHz bandwidth or an 18.432 Mbps bit stream for transmission over coaxial cable or optical fiber.

(intensity) modulated links, frequency modulated links, multimode fiber, and single mode fiber. In general the cost is higher for an FM link and for single mode fiber. However, better video performance for a longer distance is obtained from the FM system. Single mode fiber provides the ability to handle a wider bandwidth, thus more multiplexed signals. By combining FM with single mode fibers a very high performance system can be designed.

In addition to the well known uses of fiber by telephone companies and other common carriers, applications for fiber for broadcasters and program producers range from simple one-channel video or audio links to multiple audio and video circuits and high-performance wide-band cable television trunk systems.

Simple Low-Cost System

A typical fiber optic system for a low-cost, single video channel application (e.g. remote camera to field van) with good video performance characteristics and a multimode optical fiber might be configured as follows:

Amplitude (intensity) modulated at 820 nm wavelength:

Optical transmitter output power	−12 dBm
Connector loss	− 1 dB
Fiber loss (2 dB/km) for 3 km	− 6 dB
Connector loss	− 1 dB
Received signal level	−20 dBm

(−23 dBm provides 56 dB SNR)

Thus, the received signal-to-noise level = 59 dB

For a simple, single channel, intensity modulated optical transmission system this is a good low-cost alternative to a bulky coaxial cable with equalized send and receive amplifiers and possibly a repeater.

More channels can be added with additional fibers (in the already small sheath) or by multiplexing multiple signals on a single fiber. The latter usually requires the use of single mode fibers in order to keep losses to a minimum.

High Performance System

The system described below easily achieves the performance criteria for a short-haul transmission system. An important advantage in using fiber is that there are no propagation losses (as in microwave transmission) other than the light attenuation in the fiber and components. Splices cause less than 1 dB

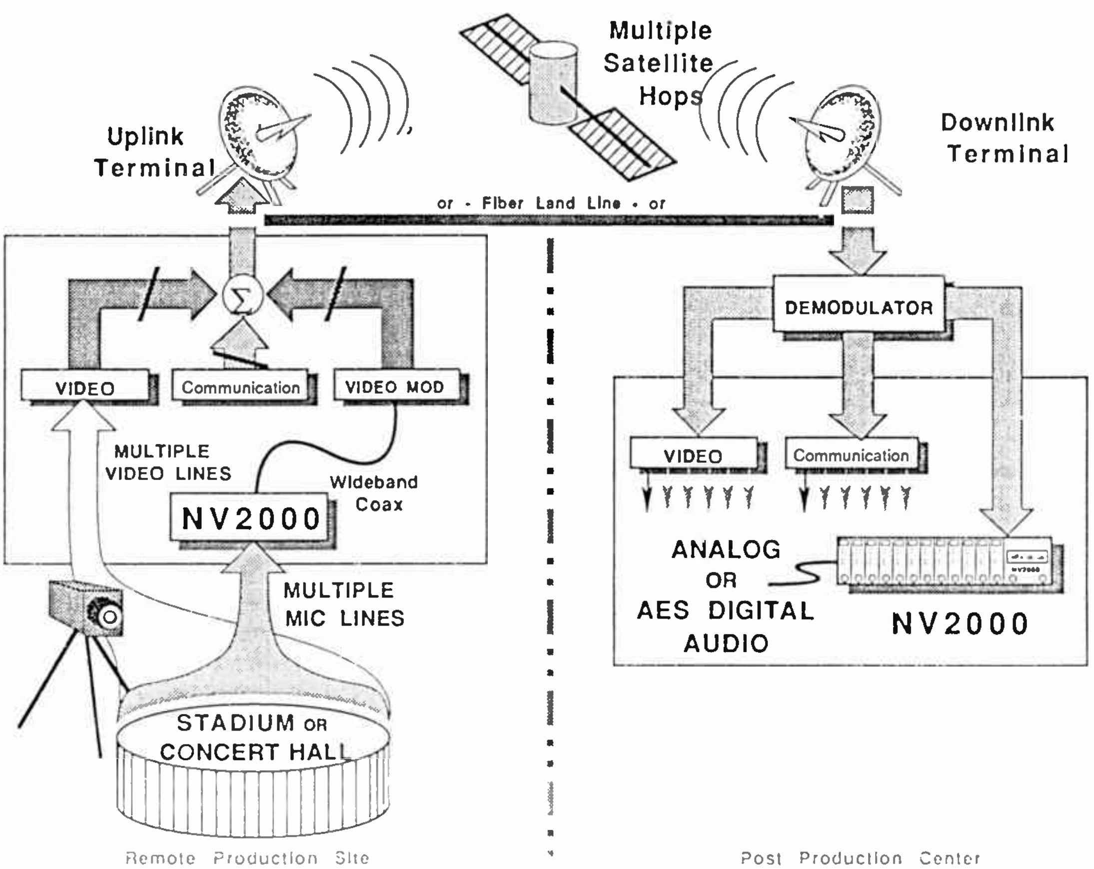


Figure 15. A fiber transmission system for remote video and/or audio production can now be configured to interconnect event sites and recording/editing facilities as figurative "studios without walls" spanning continents and oceans.

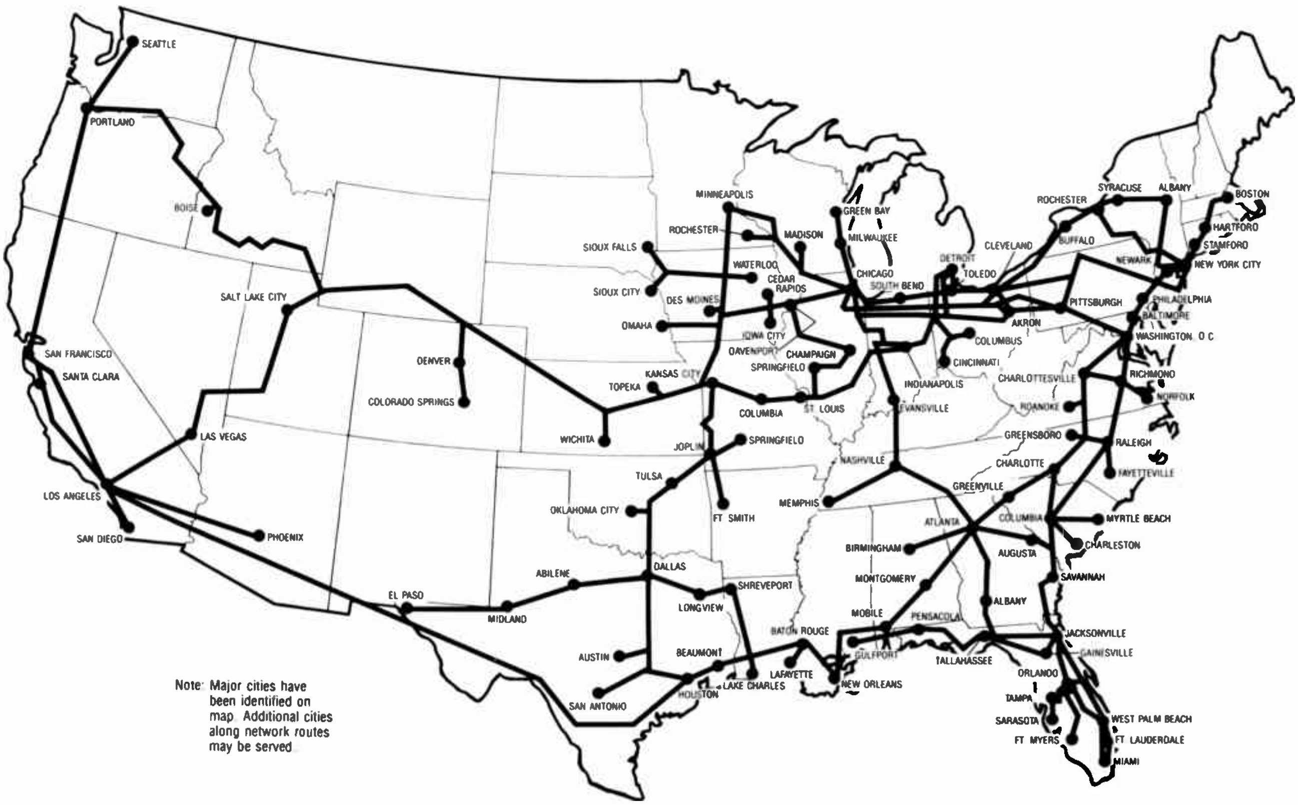


Figure 16. Vyvx NVN's national video network now interconnects 50 major television production and distribution centers from coast to coast with all-fiber circuits. NTSC video signal transmission is accomplished using advanced-design DS-3 "Excalibur" video compression codecs developed jointly by Vyvx and the Grass Valley Group (GVG). (Courtesy Wiltel [Vyvx NVN].)

each. Because of the wide bandwidth possible with fiber and no equalization or changes with temperature (as in coaxial cables) installation and maintenance of the system is kept to a minimum.

Frequency Modulated Link:

Optical transmitter power (10 chan.)	− 5 dBm
Connector loss	− 1 dB
Fiber loss (for 5 km @ 2 dB/km)	10 dB
Connector loss	− 1 dB
Received level	−17 dBm
(−30 dBm typically provides 55 dB video SNR—CCIR 5.0 MHz weighted)	
Received SNR (for each channel)	68 dB

A higher performance multi-channel system (e.g. 10 video channels and 20 audio channels) interconnecting two production facilities might be configured as follows:

1. Each FM 4.5 MHz video channel (much wider video bandwidth channels are available—up to 100 MHz) uses 3 MHz deviation and is modulated on a 70 MHz IF carrier. The audio circuits are either placed on subcarriers above the video signal or converted to digital and combined into a single channel modulating a single carrier. Several such audio and video 70 MHz carriers can be combined (after suitable frequency conversion) into a bandwidth of up to 500 or even 1 GHz.
2. As an option, the video and audio channels could also modulate individual optical transmitters at different wavelengths and optically combined. The optical output wavelength of either system is typically about 1300 nm (± 40 nm), for a single mode fiber, and the optical output power is typically about 0 to −5 dBm.

4.6 An Overview of Cable Television

Walter S. Ciciora, Ph.D.
American Television and Communications Corporation, Stamford, Connecticut

INTRODUCTION

Cable television is made possible by the technology of coaxial cable. Rigid coaxial cable has a solid aluminum outer tube and a center conductor of copper-clad aluminum. Flexible coaxial cable's outer conductor is a combination of metal foil and braided wire, with a copper-clad steel center conductor. The characteristic impedance of the coaxial cable used in cable television practice is 75 ohms. The well-known principles of transmission line theory apply fully to cable television technology.

The most important characteristics of coaxial cable are its ability to contain a separate frequency spectrum and respect the properties of that separate spectrum so that it behaves like over-the-air spectrum. This means that a television receiver connected to a cable signal will behave as it does when connected to an antenna. A television set owner can become a cable subscriber without an additional expenditure on consumer electronics equipment. The subscriber can also cancel the subscription and not be left with useless hardware. This ease of entry and exit from an optional video service is a fundamental part of cable's appeal to subscribers.

Since the cable spectrum is tightly sealed inside an aluminum environment (the coax cable), a properly installed and maintained cable system can use frequencies assigned for other purposes in the over-the-air environment. This usage takes place without causing interference to these other applications, or without having them cause interference to the cable service. New spectrum is created inside the cable by the re-use of spectrum. In some cable systems, dual cables bring two of these sealed spectra into the subscriber's home.

The principal negative characteristic of coaxial cable is its relatively high loss. Coaxial cable signal loss is a function of its diameter, dielectric construction, temperature, and operating frequency. A ball-park figure is 1 dB of loss per 100 feet. Half-inch diameter aluminum cable has 1 dB of attenuation per 100 feet at 181 MHz; at one-inch diameter, the attenuation drops to 0.59 dB per 100 feet. The attenuation of cable varies with the square root of the frequency. Thus, the attenuation at 216 MHz (within TV channel 13) is twice that of 54 MHz (within TV channel 2) since the frequency is four times as great. If channel 2 is attenuated 10 dB in 1,000 feet, channel 13 will be attenuated 20 dB. Fig. 1 demonstrates this relationship for 1,000 feet of half-inch aluminum cable.

CABLE NETWORK DESIGN

Since cable television is not a general-purpose communications mechanism, but rather a specialized system

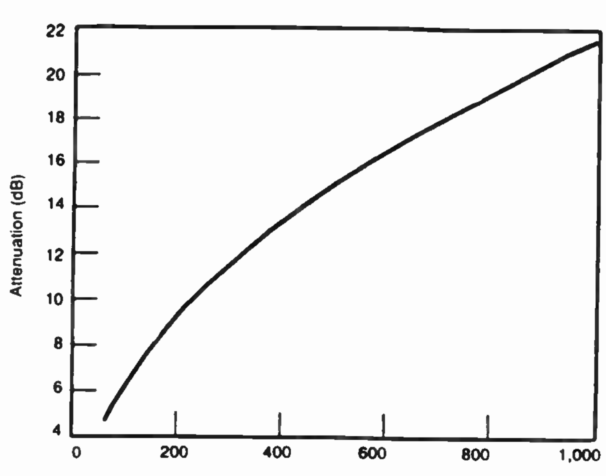


Figure 1. Coaxial cable attenuation versus frequency.

TABLE 1
Cable attenuation.

Velocity of Propagation = 82%

Cable *Type*	*RG-59*	*RG-6*	*.412"*	*.500"*	*.750"*	*1.000*
MHz	Attenuation (dB/100 m)					
5	2.55	2.00	0.75	0.56	0.39	0.33
30	6.27	4.94	1.90	1.38	0.95	0.79
50	8.42	6.65	2.53	1.94	1.36	1.08
300	20.00	16.00	5.94	4.89	3.44	2.58
450	24.60	19.80	7.47	6.07	4.40	3.58
550	27.20	21.90	8.37	6.81	4.90	4.03

Velocity of Propagation = 87%

Cable *Type*	*RG-59*	*RG-6*	*.412"*	*.500"*	*.625"*	*.750"*	*.875"*	*1.000*
MHz	Attenuation (dB/100 m)							
5	1.88	1.53	0.66	0.54	0.43	0.34	0.30	0.30
30	4.63	3.78	1.64	1.31	1.12	0.85	0.79	0.75
55	6.24	5.10	2.23	1.79	1.51	1.17	1.08	1.02
300	15.00	12.30	5.36	4.31	3.61	2.93	2.59	2.46
450	18.50	15.20	6.63	5.35	4.43	3.67	3.22	3.02
550	20.50	16.90	7.51	5.98	4.96	4.24	3.71	3.38

for transmitting numerous television channels in a sealed spectrum, the topology or layout of the network can be customized for maximum efficiency. The topology which has evolved over the years is called *tree-and-branch architecture*.

There are five major parts to a cable system: (1) the headend, (2) the trunk cable, (3) the distribution (or feeder) cable in the neighborhood, (4) the drop cable to the home and in-house wiring, and (5) the terminal equipment (consumer electronics).

Flexible coaxial cable is used to bring the signal to the terminal equipment in the home. In the simplest cases, the terminal equipment is the television set or VCR. If the TV or VCR does not tune all the channels of interest because it is not cable-compatible, a converter unit is placed between the cable and the TV or VCR tuner.

Broadcast channels 2 through 13 are not in a continuous band. Other radio services occupy the gaps. Cable can re-use these frequencies because its spectrum is self-contained within the coaxial environment. The cable converter has a high-quality broadband tuner and output circuitry which puts the desired cable channel on a low-band channel not occupied in the local off-the-air spectrum. Typically this is channel 2, 3, 4, or 5. The TV or VCR is tuned to this channel and behaves as a monitor. If programming of interest to the subscriber is scrambled, a descrambler is required. It is usually placed in the converter. Fig. 2 shows the cable frequency plan.

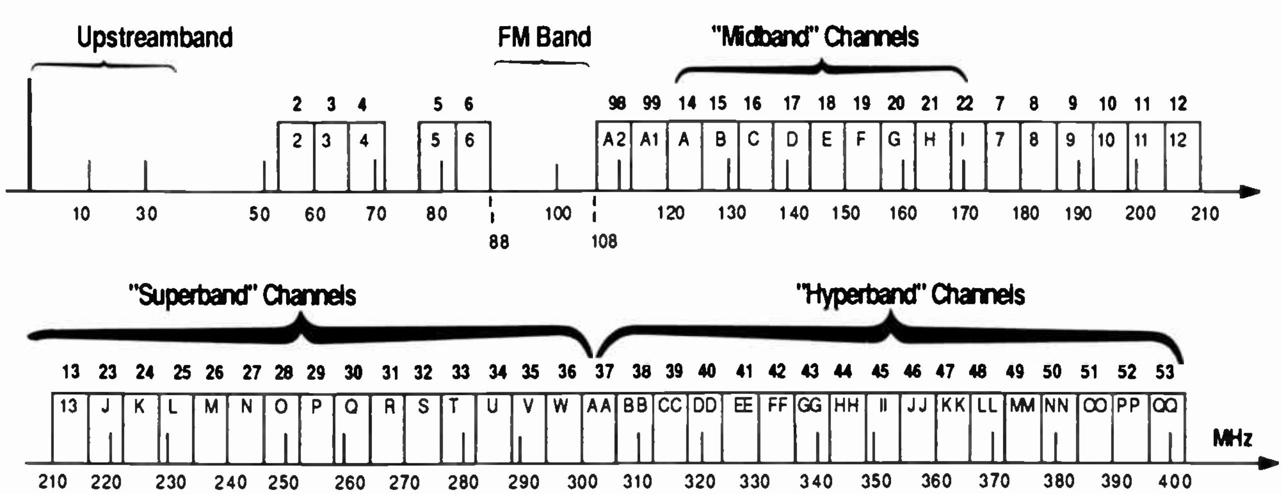


Figure 2. Frequency plan. Numbers above the rectangles are the new Electronics Industry Association (EIA) standard designations. Historical designations are inside the rectangles.

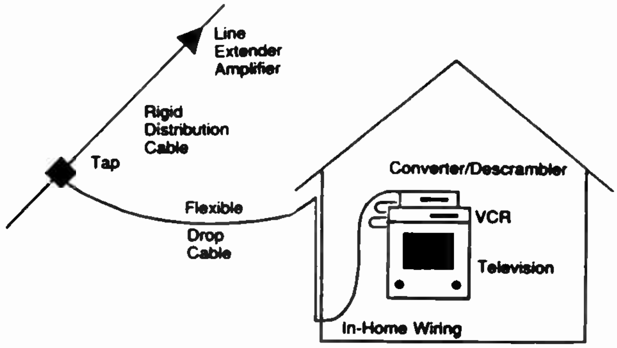


Figure 3. Terminal equipment and cable drop.

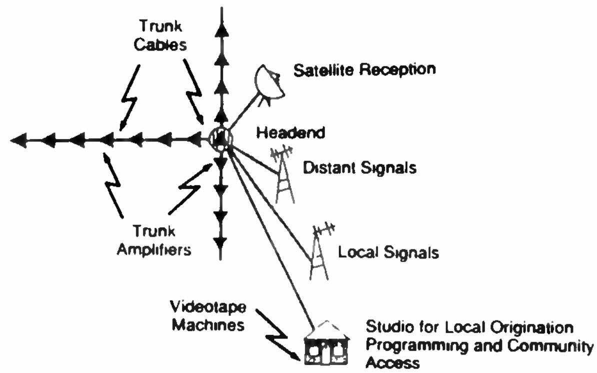


Figure 5. Cable system headend.

The home is connected to the cable system by an average of 125 feet of flexible drop cable. Fig. 3.

The distribution cable in the neighborhood runs past the homes of subscribers. This cable is tapped so that flexible drop cable can be connected to it and then routed to the residence. The distribution cable interfaces with the trunk cable through an amplifier, called a bridger amplifier, which increases the signal level for delivery to multiple homes. One or two specialized amplifiers called line extenders are included in each distribution cable. Approximately 40% of the system's cable footage is in the distribution portion of the plant and 45% is in the flexible drop cable to the home. See Fig. 4.

The trunk part of the cable system transports the signals to the neighborhood. Its primary goal is to cover distance while preserving the quality of the signal in a cost-effective manner. Broadband amplifiers are required about every 2,000 feet, depending on the bandwidth of the system. The maximum number of amplifiers which can be placed in a run, or cascade is limited by the build-up of noise and distortion. Twenty or 30 amplifiers may be cascaded in relatively high-bandwidth applications. Older cable systems with fewer channels may have as many as 50 or 60 amplifiers in cascade. Approximately 10% of a cable system's footage is in the trunk part of the system.

The headend is the origination point for signals in the cable system. It has parabolic or other appropriately shaped antennas for receiving satellite-delivered program signals, high-gain directional antennas for receiving distant TV broadcast signals, directional antennas for receiving local signals, machines for playback of taped programming and commercial insertion, and studios for local origination and community access programming.

Local origination is programming over which the cable operator has editorial control. It can range from occasional coverage of local events to a collection of programming almost indistinguishable from that of an independent broadcaster. Often mobile coverage of events is provided with microwave links back to the headend or back-feed of the signal up the cable system to the headend.

Community access is programming access for community groups mandated by the franchise. The cable system typically cannot exercise editorial control over quality or content of community access programming. See Fig. 5.

When the whole picture is assembled, the tree shape of the topology is evident. The trunk and its branches become visible. See Fig. 6.

SIGNAL QUALITY

Principal picture impairments can be divided into two categories: coherent and noncoherent. Coherent

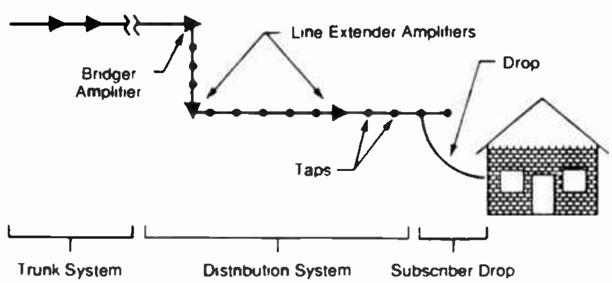


Figure 4. Distribution plant.

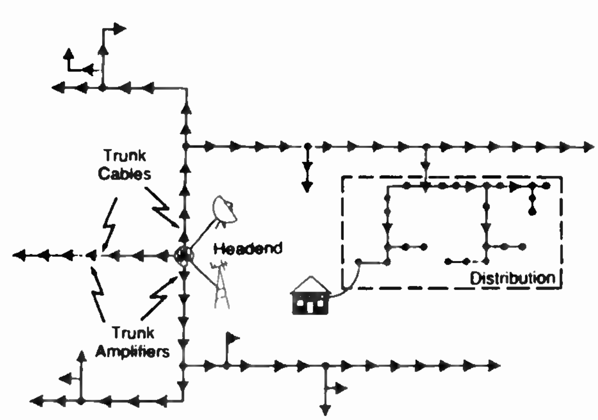


Figure 6. Tree-and-branch topology.

impairments result in a recognizable interfering pattern or picture. They tend to be more objectionable than noncoherent impairments of equal strength.

The principal noncoherent picture impairment is noise. Random noise behavior is a well understood part of general communications theory. The familiar Boltzmann relationship, noise figure concepts, etc., apply fully to cable television technology.

Noise levels are expressed in cable system practice as ratios of the video carrier to the noise in a television channel. This measure is called the carrier-to-noise ratio (CNR) and is given in decibels (dB). The target value for CNR is 45 dB to 46 dB. Noise in the picture, called snow, is just visible when CNR is 42 dB to 44 dB. Snow is objectionable at CNRs of 40 dB to 41 dB.

Coherent interference includes ingress of video signals into the cable system, reflections of the signal from transmission-line impedance discontinuities, cross-modulation of video, and cross-modulation of the carriers in the video signal. This latter phenomenon gives rise to patterns on the screen which are called beats. These patterns often look like moving diagonal bars or herringbones.

The evaluation of signal quality takes place on two planes: objective and subjective. In the objective arena, measurements of electrical parameters are used. These measurements are repeatable. Standardized, automated test equipment has been developed and accepted by the video industries. Fig. 7 lists the parameters usually considered important and the values of good, current practice. They are described in the remaining text.

The ultimate performance evaluation involves the subjective reaction of viewers. One example of the difficulties experienced is the fact that different frequencies of noise have differing levels of irritation. High-frequency noise tends to become invisible while low frequency noise creates large moving blobs which are highly objectionable. Subjective reactions to these phenomena are influenced by such factors as the age, sex, health, and attitude of the viewer. The nature of the video program, the characteristics of the viewing equipment, and the viewing conditions also impact the result.

Signal processing in the TV receiver changes the impact of signal impairments. Noise in the band of frequencies used to transmit color information is demodulated and converted into lower frequency, more objectionable noise. Noise in the synchronization part of the TV signal can cause the picture to break up entirely, resulting in much greater impairment than the same strength noise confined to other portions of the signal.

Parameter	Symbol	Value
Carrier / Noise (CNR)	C/N	46 dB
Composite Second-Order	CSO	−53 dB
Composite Triple-Beat	CTB	−53 dB
Signal Level at TV		0 dBmV

Figure 7. Signal quality target values.

In 1959, the Television Allocations Study Organization (TASO) studied the amount of noise, interference, and distortion viewers will tolerate in a TV picture. The results were expressed in a five-point scale with grades named excellent, fine, passable, marginal, and inferior. These are very old data. Work on advanced television systems (ATV) has stimulated an interest in upgrading the study to cover modern displays, video practices, and viewer tastes.

It is important to realize that the demand for signal quality is a function of time. Five to ten years ago, consumer electronics products were not capable of displaying the full resolution of the National Television Systems Committee (NTSC) signal. Gradually, these products improved until high-end models are capable of more performance than the NTSC signal can deliver. The Super VHS videotape system has greater resolution than broadcast NTSC. As time progresses, the level of performance of consumer electronics will continue to increase. As ATV and HDTV are introduced, still more demands will be made on cable system performance. The trend to larger screen sizes also makes video impairments more evident.

CABLE SYSTEM TRADE-OFFS

The experienced cable system designer has learned how to balance noise, nonlinear distortions, and cost to find a near optimal balance.

Signals in cable systems are measured in decibels relative to 1 mV (dBmV) across 75 ohms. Applying the well-known Boltzmann noise equation to 75 ohm cable systems yields an open-circuit voltage of 2.2 microvolts in 4 MHz at room temperature. When terminated in a matched load, the result is 1.1 microvolts; and the minimum room temperature noise in a perfect cable system is −59.17 dBmV.

Starting at the home, the objective is to deliver at least 0 dBmV, but no more than 10 dBmV to the terminal on the television receiver. Lower numbers produce snowy pictures and higher numbers overload the television receiver's tuner, resulting in cross modulation of the channels. If a converter or descrambler is used, its noise figure must be taken into account. There are two reasons for staying toward the low side of the signal range: cost, and the minimization of interference in the event of a signal leak caused by a faulty connector, damaged piece of cable, or defect in the television receiver. Low signal levels may cause poor pictures for the subscriber who insists on unauthorized splitting in the home to serve multiple receivers. Working our way back up the plant, we need a signal level of 10 dBmV to 15 dBmV at the tap to compensate for losses in the drop cable.

The design objectives of the distribution part of the cable system involve an adequate level of power not only to support the attenuation characteristics of the cable, but to allow energy to be diverted to subscribers' premises. Energy diverted to the subscriber is lost from the distribution cable. This loss is called *flat loss* because it is independent of frequency. Loss in the cable itself is a square-root function of frequency, and is therefore contrasted to flat loss. Because of flat losses, relatively high power levels are required in the distribution part of the plant, typically 48 dBmV at the input to the distribution plant. These levels force the amplifiers in the distribution part of the plant to reach into regions of their transfer characteristics which are slightly nonlinear. As a result, only one or two amplifiers, called line extenders, can be cascaded in the distribution part of the plant. These amplifiers are spaced 300 to 900 feet apart depending on the number of taps required by the density of homes.

Because the distribution part of the plant is operated at higher power levels, nonlinear effects become important. The television signal has three principal carriers: the video carrier, the audio carrier, and the color subcarrier. These concentrations of energy in the frequency domain give rise to a wide range of beats when passed through nonlinearities. To minimize these effects, the audio carrier is attenuated about 15 dB below the video carrier.

When cable systems only carried the 12 VHF channels, second-order distortions created spectrum products which fell out of the frequency band of interest. As channels were added to fill the spectrum from 54 MHz to as much as 650 MHz, second-order effects were minimized through the use of balanced, "push-pull" output circuits in amplifiers. The third-order component of the transfer characteristic dominates in many of these designs. Fig. 8 demonstrates the triple-beat phenomena. The total effect of all the carriers beating against each other gives rise to an interference called *composite triple-beat* (CTB). CTB is measured with a standard procedure involving 35-channel carriers. In a 35-channel cable system, about 10,000 beat products are created. Channel 11 suffers the most with 350 of these products falling in its video. Third-order distortions increase about 6 dB for each doubling of the number of amplifiers in cascade. A 1 dB reduction in amplifier output level will generally improve CTB

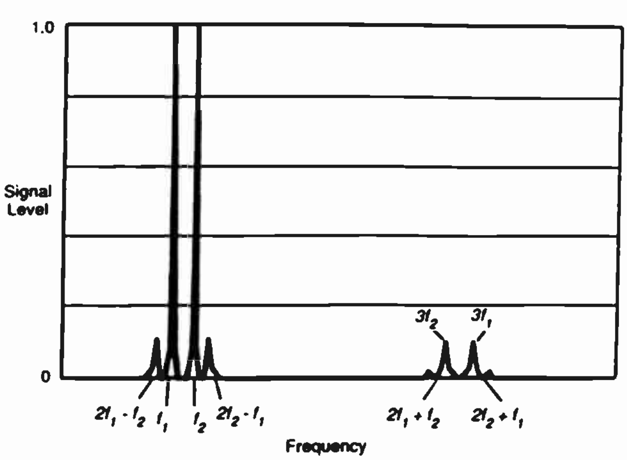


Figure 8. Triple-beat.

by 2 dB. If these products build to visible levels, diagonal lines will be seen moving through the picture. When these components fall in the part of the spectrum which conveys color information, spurious rainbows appear.

If we assign a design level of noise and nonlinear distortion at the subscriber's television receiver which is below the threshold of visibility, we can conceive of a budget of noise and distortion to be spent in the various parts of the system design. The distribution part of the system has relatively high powers and has used up most of the budget for nonlinear distortions. On the other hand, little of the noise budget has been consumed. It can be allocated to the trunk part of the system, which brings the signals into the neighborhood.

The design objective of the trunk part of the cable system is to move the signal over substantial distances with minimal degradation. Because distances are significant, lower-loss cables are used. One-inch and ¾-inch diameter cable is common in the trunk while ½-inch cable is found in the distribution portion. Signal levels in the trunk at an amplifier's output are 30 dBmV to 32 dBmV depending on the equipment used.

It has been determined through analysis, and confirmed through experience, that optimum noise performance is obtained when the signal is not allowed to be attenuated more than about 20 dB to 22 dB before being amplified again. Amplifiers are said to be spaced by 20 dB. The actual distance in feet is a function of maximum frequency carried and the cable's attenuation characteristic. Modern high-bandwidth cable systems have their amplifiers fewer feet apart than older systems with fewer channels. Since attenuation varies with frequency, the spectrum in coaxial cable develops a slope. This is compensated with equalization networks in the amplifier housings.

The attenuation of the cable is a function of temperature and aging of components. Modern amplifiers use a pilot signal to control automatic gain control (AGC) circuits. A second pilot signal, at a substantially different frequency than the first, allows the slope of the attenuation characteristic to be monitored and compensation to be introduced with automatic slope control (ASC) circuits. Thus, long cascades of amplifiers can, once properly set up, maintain their performance over practical ranges of temperature and component aging.

Since the signal is not repeatedly tapped off in the trunk part of the system, high power levels are not required to feed splitting losses. As a result, signal levels are lower than in the distribution portion of the plant. Typical levels are about 30 dBmV. For the most part, the amplifiers of the trunk are operated within their linear regions. The principal challenge of trunk design is keeping noise under control. Each doubling of the number of amplifiers in the cascade results in a 3 dB decrease in the CNR at the end of the cascade and a 6 dB increase in the amount of CTB.

If the noise at the end of the cascade is unacceptable, the choices are to employ lower noise amplifiers, shorter cascades, or a different technology, such as microwave links or fiber optic links.

SYSTEM CONFIGURATIONS AND TRENDS

Channel Carriage Capacity

Channel carriage capacity is based on radio frequency (RF) bandwidth. It is a useful characteristic for classifying cable systems. As shown in Fig. 9, there are three types of systems. Systems are categorized by their highest operating frequency. Downstream signals are transmitted to the customers' homes.

Small capacity cable systems operate in the 50 MHz to 220 MHz range with a bandwidth of 170 MHz. Twelve to 22 television channels are activated. These systems were constructed from the mid-1950s to the late-1970s. They account for approximately 10% of total plant mileage.

Referring to Fig. 10, a cable system configuration consists of: (1) the headend (the signal reception, origination, and modulation point), (2) main, coaxial trunk (or tree) cable, which runs through central streets in communities, (3) coaxial distribution (branch) cable to the customer's neighborhood, including distribution taps, (4) subscriber drops to the house, and (5) subscriber terminal equipment (television sets, converter/descramblers, VCRs, etc.). Distribution plant is sometimes called *feeder* plant. Programming comes to the headend by satellite signals, off-air signals from broadcast stations, and signals imported via terrestrial microwave. Signals originating from the headend are from a collocated studio facility, VCRs, character generators, or commercial insertion equipment.

Plant mileage is calculated using the combined miles of strand that support the coaxial cables in the air and the footage of trenches where cables are installed in the ground. There are more than 750,000 miles of plant in the 9,600+ U.S. cable systems. On average, the ratio of trunk footage to feeder footage is 1:3, with 75 homes per mile and 53% penetration. The average home drop is 125 feet.

Extension cables, or drops, interconnect main coaxial plant lines to customers' homes. They are not included in plant mileage. Drop cables are smaller in diameter than mainline coaxial cable plant lines. They interconnect between a power splitter, called a multitap directional coupler, and the customer's interface, usually a television set. The tap is located in the utility easement. In an average cable system, there will be 6,625 feet of drop cable per plant mile and 7,040 feet of hard-line coaxial cable. Put another way, 48% of the total plant is drop cable and associated F-connectors, which are the connectors used to connect coaxial cable to equipment. With the cable industry targeting subscriber penetration at 70%, in an average system 55% of the plant will be drop cable. About 45% of service calls are related to problems with the drop portion of the plant. About one-third of the drop-related service calls are caused by problems at the F-connectors. Cable systems replace approximately 30% of drops annually.

220 MHz systems built ten or more years ago are found in rural areas or in areas with clusters of small established communities. Some of these systems operate trunk lines running over 20 miles with 50 or more amplifiers in cascade. Total plant mileage for an average 220 MHz systems extends from 50 to 500 miles and services up to 15,000 cable customers. New construction of 220 MHz systems occurs only where there are small numbers of potential customers (no more than 300) and where plant mileage does not exceed 10 miles.

Medium capacity cable systems operate with upper frequencies at 270 MHz and 330 MHz, and total bandwidths of 220 MHz and 280 MHz, respectively. 270 MHz systems deliver 30 channels, while 330 MHz systems deliver 40 channels. Although new cable systems are seldom built with 40-channel capacity, plant extensions to existing 270 MHz, 300 MHz, and 330 MHz systems occur. Electronic upgrade is frequently employed to increase 270 MHz systems to 330 MHz. Some 220 MHz systems are upgrading to 300 MHz.

Medium capacity systems account for about 75% of total plant mileage. They serve a wide range of communities, from rural areas (populations of 5,000 to 50,000), to some of the largest systems built in the late-1970s. The San Antonio system, a medium capacity system, passes 420,000 homes, consists of 4,000 miles of plant with over 2,000 trunk amplifiers, and has in excess of 11,000 distribution amplifiers. The longest cascade in the system is 37 trunk amplifiers. There are 300 MHz systems with cascades of 45 or more trunk amplifiers.

Large capacity cable systems achieve high channel capacities through extended operating frequencies and through the installation of dual, collocated, coaxial

Bandwidth		Operating frequencies	# of channels
(RF range)			
Small:	170 MHz	50 MHz-220 MHz	12–22 (single coax)
Medium:	220 MHz	50 MHz-270 MHz	30 (single coax)
	280 MHz	50 MHz-330 MHz	40 (single coax)
Large:	350 MHz	50 MHz-400 MHz	54 (single coax)/ 108 (dual coax)
	400 MHz	50 MHz-450 MHz	60 (single coax)/ 120 (dual coax)
	500 MHz	50 MHz-550 MHz	80 (single coax)

Figure 9. Downstream signals: ranges of operating frequencies and channels.

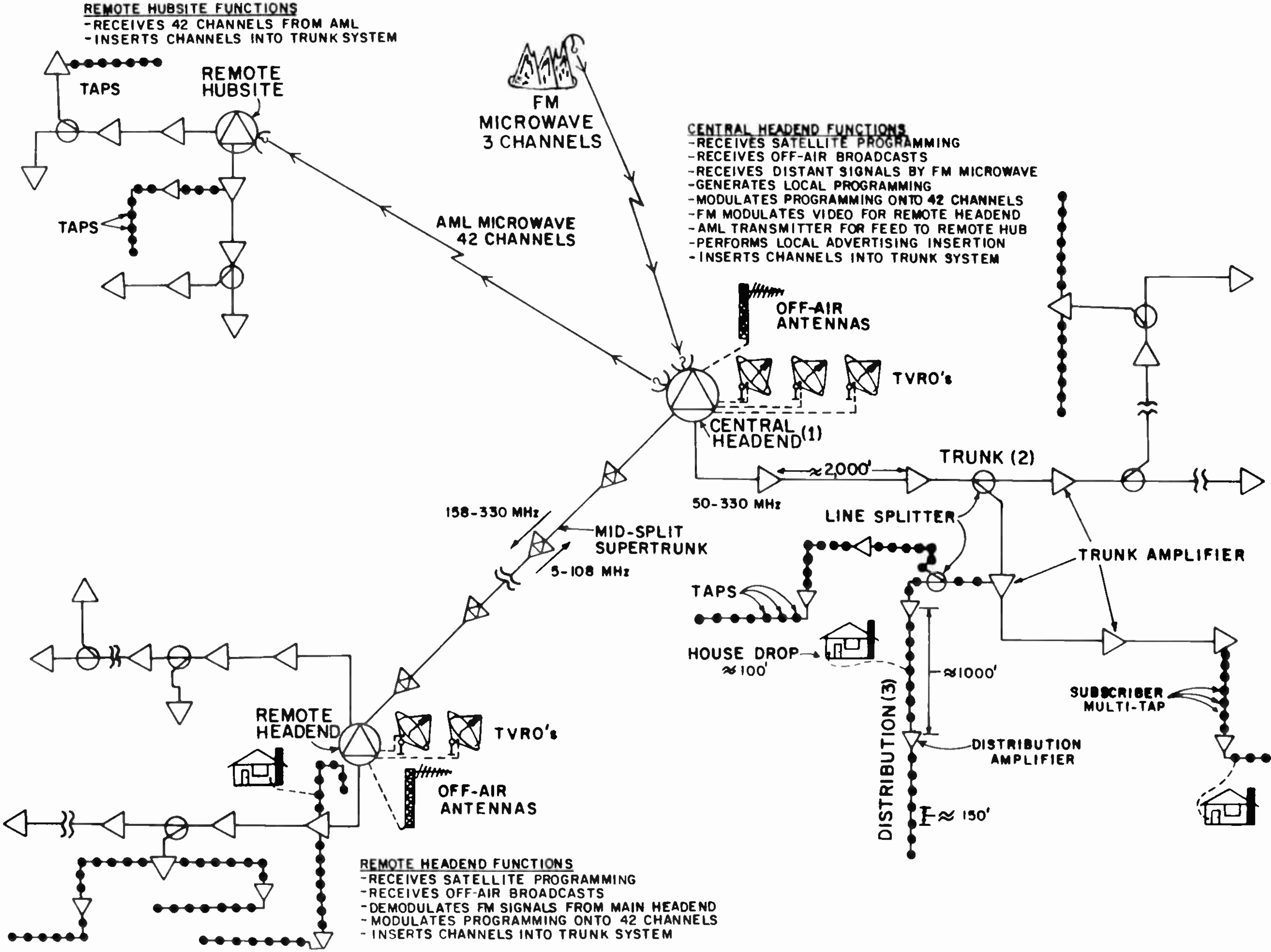


Figure 10. Sample cable television system architecture.

cable. Single coaxial cable systems range from 54-channel, 400 MHz to 80-channel, 550 MHz. With dual cable, it is not unusual to find 120 channels.

Large capacity cable systems account for about 15% of total cable plant miles. They are primarily high-tech systems designed for large urban areas previously not cabled. They serve 50,000 to 150,000 customers and consist of 400 to 2,000 miles of plant. They began construction in 1981. The earliest were 54-channel, 400 MHz systems. Then came 60-channel, 450 MHz systems. These were quickly followed by dual-cable, 400 MHz or 450 MHz systems with carriage of 108 to 120 channels. The dual-cable trend has tapered off and the remaining urban systems are being built with a single cable, 450 to 550 MHz.

Large capacity systems are designed, and some operate, as two-way cable plant. In addition to the downstream signals to the customers (50 MHz to upper band edge), upstream signals are carried from customers to the cable central headend, or hub node. They are transmitted using frequencies between 5 MHz and 30 MHz.

550 MHz systems average at least twice as many amplifiers per mile of cable plant as 220 MHz systems. If too many amplifiers are put in cascade, and 54 or more channels are transmitted, there would be objectionable distortion from the cascaded amplifiers. Technology has become available that reduces the number of required amplifiers to the number per mile in 300 MHz plant. In the 400+ MHz systems, amplifier cascades are kept to less than 30.

Industry Trends

When franchises come up for renewal, many civic authorities have required an increase in channel capacity: 220–270 MHz systems upgrade to 300–330 MHz. However, upgrades to 400–450 MHz are occurring in some urban U.S. markets. On the west coast, large bandwidth upgrades are occurring in systems with as few as 20,000 subscribers because of the competitive climate.

Advanced television (ATV) will put pressure on operators to expand plant bandwidth. The most talked about form of ATV is high definition television (HDTV).

The last few years have brought exciting trends employing new technologies. Fiber is now being installed to upgrade older systems and as part of rebuilds and new builds. The old trunk system of long cascades of amplifiers is now considered obsolete. Work on new amplifier technologies will allow a realization of cable's inherent bandwidth, which exceeds 1 GHz.

A bandwidth of 1 GHz contains 160 slots of 6 MHz. These can be allocated to NTSC, HDTV simulcast, and to new services. The most exciting potential lies with utilizing video compression technology to squeeze four or even five NTSC-like quality signals in a 6-MHz slot. This opens the door for hundreds of channels. "Near video on demand" becomes practical, i.e., the most popular movies could be repeated every few minutes to minimize the wait time before a movie starts. The average wait time could be made shorter than the trip to the video store, and the subscriber does not have to make a second trip to return the movie. They are assured that the movie is always in. A microprocessor can keep track of which channel to return to should the subscriber wish to take a break. It is possible to design systems that behave as if they are switched even though they remain more like a traditional cable tree-and-branch structure.

Channelization

There are three channelization plans to standardize the frequencies of channels. The first plan has evolved from the frequency assignments that the Federal Communications Commission (FCC) issued to VHF television broadcast stations. This plan is called the standard assignment plan.

The second channelization plan is achieved by phase locking the television channel carriers. It is called the IRC plan (Incrementally Related Carriers). The IRC plan was developed to minimize the effects of third-order distortions generated by repeated amplification of the television signals as they pass through the cable plant. As channel capacities increased beyond 36 channels, composite, third-order distortions became the limiting distortion.

The third channelization type is the HRC plan (Harmonically Related Carriers). It differs from the standard and IRC plan by lowering carrier frequencies by 1.25 MHz. With HRC, carriers are phase locked and fall on integer multiples of 6 MHz starting with channel 2 at 54 MHz. This plan was created to further reduce the visible impact of amplifier distortions.

If ATV signals are not transmitted at 6-MHz spacings, they can create distortions that will no longer be masked by the IRC or HRC channelization process in the remaining conventional channels. If these ATV channels are combined with conventional channels spaced at 6 MHz, distortion products will fall in the ATV channels in nonoptimal locations, resulting in degradation of the ATV picture.

The channelization plans were designed to reduce the visibility of distortion products by making their frequencies synchronous with the interfered carrier. Since carriers present in the downstream signal path add to the distortions, cable systems carry nonvideo carriers at a level that is 13–17 dB below the video carrier's amplitude. This drastically reduces distortion contributions. With broadcast TV channels carried on cable, special processing equipment is used that reduces the aural carrier. This amplitude reduction does not significantly affect the audio signal-to-noise ratio (SNR) quality of monaural television sound. However, this lower level created SNR problems at the end of the cable system for higher bandwidth signals such as FM stereo. When stereo television audio was developed, careful attention was focused on encoding techniques that would promote SNR immunity of the difference channel.

FM radio services are carried at an amplitude that is 15–17 dB below channel 6's video carrier level. The services are carried on cable in the FM band slot of 88–108 MHz. In an IRC channel plan, channel 6's aural carrier falls at 89.75 MHz, which reduces the available FM band to 90–108 MHz.

Low speed data carriers are transmitted in the FM band or in the guard band between channels 4 and 5 in a standard frequency plan. The amplitude of these carriers is at least 15 dB below the closest video carrier level.

SIGNAL TRANSPORTATION SYSTEMS

Transportation systems were developed to deliver high quality signals from the central headend point to remote headend or hub locations where cable signals are injected into cable trunking systems. (See Fig. 10.) The increase in channel capacity and subsequent need to decrease amplifier cascades directly affected the development of transportation systems. Urban franchises cover large areas, yet wish to transmit downstream signals from a common headend point. Other motivating factors include increased local programming originating from the cable system's studio, the requirement to deliver city government programming originating from municipal locations, a new business opportunity for the insertion of local commercials to satellite-delivered services, and the development of pay-per-view (PPV) programming.

Several transportation methods have become popular over the last ten years: amplitude-modulated microwave link (AML), frequency-modulated microwave link (FML), frequency-modulated coaxial link (FMCL), amplitude-modulated coaxial supertrunks, and fiber interconnects.

AML and FML Microwave Links

The most popular mode of signal transportation is the AML microwave system. AML allows the delivery of the entire downstream cable spectrum through the air to reception points located eight to ten miles away from the microwave transmit site. An AML microwave transmitter provides adequate power to deliver signals to eight reception hubs. The advantage of AML is that the receiver simply performs a wide-band, block downconversion. The entire downstream frequency spectrum (i.e., 50–400 + MHz) is reproduced at an appropriate level for direct insertion into the cable plant, without the need for individual channel frequency retranslation or combining. The only equipment required at an AML reception site is the microwave receive antenna and a microwave receiver. These can be mounted to a small tower or telephone pole.

AML upconverts each television channel and combines it at its respective microwave frequency using a complex waveguide system. A transmitter retrofit, to support transmission of ATV of wider bandwidth, would be difficult and expensive. Adding additional NTSC channels costs around $13,000 per channel. In some cases, each channel added lowers the maximum power per channel. This may either shorten path lengths or increase the probability of signal degradation during rain fades.

One of the main uses of FML is to cover distances not feasible with AML or where video SNR in excess of 56 dB is required. FML is a single-channel, frequency-modulated microwave transmission system. FML occupies significantly more bandwidth than amplitude modulated systems, from 12.5 MHz to 25 MHz per channel, depending on path length. Fewer channels are available for use. Since the transmitter is single channel, it could be more readily modified to accept a wider ATV signal than AML. Since this transmission system delivers FM signals, demodulation, vestigial sideband amplitude remodulation (VSB-AM), and frequency translation must occur before insertion into the cable plant. The FML system is often used to deliver specialized programming (i.e., local origination) to a remote headend.

Frequency Modulated Coaxial Trunk

FMCL is an adaptation of FM microwave designed for transmission over coaxial cable plant. There are FM modulators for each television channel at the originating end and FM demodulators at the remote site. VSB-AM modulation must occur before these channels can be inserted into the cable that provides direct service to subscribers. The primary use of FMCL is for studio quality transmission of specialized programming from the origination site to the headend. Since it involves single-channel modulation and operates in a sealed coaxial environment, differing modulation bandwidths may be used.

Amplitude-Modulated Coaxial Supertrunk

The amplitude-modulated supertrunk transportation system can be as simple as a conventional coaxial trunk transporting VSB-AM signals between two headends. Or the supertrunk can link a primary headend directly to the beginning of the normal cable distribution trunk. The television channels can be combined at their final frequency assignments for direct insertion into the distribution trunk. Or they can be grouped to fit into a reduced coaxial bandwidth (i.e., 5–108 MHz) for transportation on a special service coaxial trunk. An example of this trunk is a mid-split system, which is a bi-directional trunk. As an example, the mid-split trunk can use 158–330 MHz downstream and 5–108 MHz upstream.

Another variation of AM supertrunk is split-band trunking. The spectrum from 50 MHz to 450 MHz is split into two groups of 30 television channels each and inserted on two side-by-side trunks that connect two sites. The signals are equally loaded onto two trunk cables to improve television signal quality. This is achieved because the amount of distortions produced are directly related to the number of channels carried on each trunk cable.

Whenever signals are not transported on their final frequency assignment, frequency translation devices, called *channel processors*, are used to shift the television signals to their final frequency. AM supertrunks can be similar to cable distribution trunks. The difference is that the AM supertrunk will use new technology trunk amplifiers, such as feedforward. These transport with significantly less distortions and better CNR.

Fiber Interconnect

Fiber optic transportation systems are now frequently used. Analog video fiber technology is preferred for this application, although occasionally cable systems use digital video. Initially, the signals were carried using FM modulation. Analog fiber networks carry six to 12 video signals per fiber when frequency modulation techniques are used. The focus of current research is to optimize fiber optic transmitters and receivers. The goal is the broadband carriage of 40–60 NTSC, amplitude-modulated video channels over a single fiber on path lengths up to 20 kilometers, approximately 12.5 miles.

Examples of cable system applications for analog FM video fiber optic transportation systems are: studio-quality video from a remote television receive only (TVRO) earth station, a broadcaster feed from a local television studio to a headend, and transmission of a group of channels to which local advertising has been added from a central point to several headend locations.

An analog AM fiber optic system called *fiber backbone* was recently developed at American Television and Communications Corporation (ATC) to deliver a broadband group of channels with higher quality than is available through conventional coaxial trunks. Lowering costs and enhancing reliability are also objectives. This technology allows distribution of 30 to 40 combined RF channels to nodes located along the trunk lines.

The initial application of this technology was in Orlando, Florida, in 1988 as a backup for microwave links. Heavy rain causes these links to fade or suffer complete blockage. Automatic equipment switches to the fiber link. The AM fiber technology is now commonly used instead of microwave in many applications.

Digital fiber links are used when video signals are partially transported in a digital common carrier network. The video interfaces used operate at a DS-3 rate of 45 Mbps, which routes through common carrier points-of-presence and switching networks.

FREQUENCY BAND USAGE REGULATION

Frequencies Under Regulation

FCC Rules and Regulations govern the downstream cable frequencies that overlap with the over-the-air frequencies used by the Federal Aviation Administration (FAA). These frequencies are in the 108 MHz to 137 MHz and 225 MHz to 400 MHz bands. They are used by the FAA for aeronautical voice communications and navigational information. Since cable plant is not a perfectly sealed system, the FCC and the FAA want to maintain a frequency separation between signals carried on cable and frequencies used by airports near the cable system boundaries. In 400 MHz systems, over 30 channels are affected by the FCC Rules on frequency offset and related operating conditions.

Effects of the New FCC Rules

The maximum, unregulated, carrier power level rule has been reassessed and changed. The previous limit of 1×10^{-5} Watts (28.75 dBmV) has been raised to 1×10^{-4} Watts (38.75 dBmV). Carriers with power levels below 38.75 dBmV are not required to follow the frequency separation and stability criteria. Carriers within ±50 KHz of 156.8 MHz, ±50 KHz of 243 MHz, or ±100 KHz of 121.5 MHz, which are emergency distress frequencies, must be operated at levels no greater than 28.75 dBmV at any point in the cable system.

INCREASING CHANNEL CAPACITY

There are several ways to increase channel capacity. If the actual cable is in good condition, channel capacity is upgraded by modifying or replacing the trunk and distribution amplifiers. If the cable has seriously deteriorated, the cable plant must be completely rebuilt.

Upgrades (Retrofitting) and Rebuilds

An upgrade is defined as a plant rehabilitation process that results in the exchange or modification of amplifiers and passive devices (such as line splitters, directional couplers, and customer multitaps). A simple upgrade requires new amplifier circuit units called hybrids. A full upgrade replaces all devices in the system. In an upgrade project, most of the cable is retained. The goals of an upgrade project include increasing the plant's channel capacity and system expansion to outlying geographic areas. New amplifier technology such as feedforward and/or power doubling circuitry and advances in amplifier performance have greatly enhanced the technical and financial viability of upgrades. Upgrades are often the least expensive solution to providing expanded service.

A rebuild is the most expensive solution to providing upgraded service. In a rebuild, the outside plants is replaced. Customer drops are replaced on an as-needed basis. The strand that supported the old cable is occasionally retained. A rebuild requires a minimum of system downtime, since both old and new plants are active for a period of time. This allows the customer's drop to be switched directly from the old system to the new.

Once the plant has been rebuilt or upgraded, customers are provided newer converters with additional capabilities. The displaced units are moved to other systems or used for the basic tier of service.

SYSTEM DISTORTION AND SYSTEM MAINTENANCE

Constraints on the design and implementation of cable systems are imposed by each device used to transport or otherwise process the television signal. Each active device adds small distortions and noise to the signal. Even passive devices contribute noise. The distortions and noise compound so that with each additional device the signal becomes less perfect.

Any nonlinear device, even bi-metallic junctions, cause distortions. The primary contributors are the slight nonlinearities of amplifiers. Because the amplifiers are connected in cascade, nonlinear contribution from each device in the cascade cumulatively degrades the signal.

Noise in any electronic system can come from many sources. The major source is the random thermal movement of electrons in resistive components. For a cable system at 20°C or 68°F, the thermal noise voltage in a 4 MHz bandwidth will be 1.1 microvolts or −59.1 dBmV. This is the minimum noise level, or noise floor. Noise contributions from amplifiers add on a power basis, with the noise level increasing 3 dB for each doubling of the number of identical amplifiers in cascade. Eventually, the noise will increase to objectionable levels. The difference between the RF peak level and the noise level is measured to quantify the degree of interference of the noise power. The power levels in watts are compared as a ratio. This is called the signal-to-noise ratio, or SNR. In a cable system, the apparent effect of noise is its interference with the video portion of the TV channel. This level is compared to the video carrier and called the carrier-to-noise ratio (CNR).

As the CNR value decreases, the interference of noise with the signal becomes visible as a random fuzziness, called snow, that can overwhelm the picture resolution and contrast. The point where the picture becomes objectionably noisy to viewers is approximately at a CNR = 40 dB. In well-designed systems, the CNR is maintained at 46 dB. While an increase in signal level would improve the CNR, unfortunately, there can be no level increase without increases in distortions.

The distortion products of solid-state devices used in cable amplifiers are a function of the output levels and bandwidths. The higher the signal level, the greater the distortion products produced. Modern amplifiers use balanced configurations which almost completely cancel the distortion caused by the squared term of the amplifier's transfer characteristic. The dominant remaining distortions are called triple beats. They are caused by the cubed term. Because distortion products add on a voltage basis, the composite triple-beat (CTB) to carrier ratio decreases by 6 dB for each doubling of the number of amplifiers in cascade, whereas the CNR decreases by 3 dB for each doubling.

The operating parameters of amplifiers determine the number which can be cascaded and, hence, the distance which can be covered.

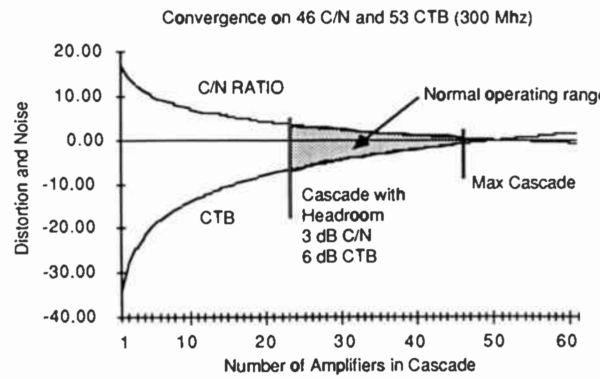


Figure 11. Distortions in a cascade.

A cable system's operating limits are defined in terms of its noise floor and distortion ceiling. In Fig. 11, the noise floor and distortion ceiling are presented as a function of the number of amplifiers in cascade for a system at 300 MHz. The diagram shows that a cascade of 46 trunk amplifiers is possible while realizing a 46 dB CNR and a 53 dB CTB. However, other operating realities dictate that substantially more headroom for both distortion and noise be incorporated into the design. The other factors to consider include: change in cable attenuation and noise with temperature, automatic gain control (AGC), and automatic slope control (ASC) tolerances, system frequency response, accuracy of field test equipment, designs anticipating ATV, and maintenance probabilities. In the example cited, allowing 1 dB of AGC/ASC change, a 3 dB peak-to-valley and 2 dB of test equipment uncertainty results in a 6 dB tolerance. Limiting distortion, CTB, implies that the cascade should only be half the length predicted in the chart, or 23 amplifiers.

The foregoing discussion is applicable only to the trunk portion of the system. As signal levels are increased in the distribution sections, additional allowances must be made in the system design. As a rule of thumb, CNR is determined primarily by the conditions of trunk operation and signal-to-distortion ratio (SDR) primarily by the conditions of distribution operation.

Two other factors limit the geography of a cable system. Cable attenuation rises with increasing frequency. More equal gain amplifiers are required to transmit the signal a given distance at higher frequencies. But noise limits the maximum number of amplifiers used. The second factor is that amplifier distortion is a function of channel loading: the more channels carried, the greater the distortions.

To obtain optimum cascade length, AGC/ASC tolerance, accurate alignment, calibrated test equipment, and well-founded system maintenance programs are of paramount importance. Maintenance programs are designed to ensure that system alignment is kept within acceptable limits. Where trunk lines carry signals through cascades of up to 40 trunk amplifiers, it is crucial that each amplifier have a flat amplitude-versus-frequency response. The additive effects of even minimal amplitude response variations in each amplifier creates significant system flatness problems at the end of long cascades.

Maintaining Amplitude-Versus-Frequency Response

A maintenance program objective should be to achieve a system amplitude-versus-frequency response of less than [(N/10)+1] dB peak-to-valley. This will minimize the degradation of CNR which occurs in channels that fall in the valleys of the system response. Here N is the number of trunk amplifiers in the cascade. It is imperative that the optimum response is maintained at each amplifier. The NCTA's recommended practice specifies that there will be no more than 3 dB difference in adjacent channel video carrier amplitudes provided to the customer. All channels must fall within a 12 dB overall passband response window.

The common method of evaluating a system's frequency response is a sweep generator that injects a rapidly swept carrier over the system's passband. The sweep generator is set to sweep from 50 MHz to the upper system passband frequency with a duration time as short as 1 millisecond. The field sweep receiver is then synchronized to the generator. The portable receiver provides a display of the system's response at each amplifier as the maintenance technician progresses through the trunk cascade.

Another tool for checking system flatness is the spectrum analyzer. To determine the overall system response, the individual video carrier amplitude of each television channel is measured. Because the carrier amplitudes are usually adjusted for a sloped amplifier output to minimize distortion products, more interpretation is needed with a spectrum analyzer. The sweep system allows the amplifiers to be adjusted for the flattest response.

Excessive response variations can cause additional distortions since some carriers will now exceed the amplitude at which the system's amplifiers were designed to operate. To keep these response variations from becoming excessive, amplifier manufacturers make a response control device known as a mop-up circuit. It is installed at periodic locations in trunk amplifiers throughout the cascade. These mop-up circuits are tunable filters adjusted to remove small peaks (less than 1.5 dB) caused by the amplifiers or cable. Incorrect use of these devices to solve defective equipment problems (i.e., bad cable sections, splices, or line passives) can cause impairment to a video channel by changing its in-band frequency response or its chrominance-to-luminance delay characteristics.

Group Delay Through the Cable Plant

Trunk amplifiers with bi-directional capability exhibit group delay as a result of the band-splitting diplex filters. The visible effect of the filtering is a loss of resolution in the picture. This will be a concern for ATV. The diplex filters, which are high pass and low pass filters with a 40 MHz crossover frequency, are part of the trunk amplifier's circuitry.

Channels 2, 3, and 4 suffer from the repeated effects of the filtering. Other locations where filtering occurs are apartment complexes, hotels/motels, or hospitals, where channels are deleted from the spectrum by special bandstop filters. Locally originated channels are inserted in the deleted portions of the spectrum.

Filtering occurs at the headend or hub in connection with channel processing equipment or the channel modulators. The effects of these filters are taken into account when the headend equipment is designed. There is only one of these devices per channel. Thus, the delay effects of this equipment rarely create problems. However, with some configurations of signal transportation systems, additional single-channel or multiple-channel filtering may take place and cause delays at hubs.

System Reflections

Signal reflections occur throughout the cable plant and are called *micro-reflections*. They are caused by the individual slight errors in impedance match. The severity of the mismatch is measured by the magnitude of the return-loss ratio. The larger the return loss, the better. Mismatches include connectors, splices, and even damage to the cable itself. The example in Fig. 12 is a splice installed in a trunk line approximately 150 feet past an amplifier. The splice, which may only have a 12 dB return loss, reflects signals back upstream that have only been attenuated by 13.5 dB (1.5 dB in cable attenuation plus the return loss of the splice). The reflected signals then arrive back at the output of the amplifier attenuated by a total of 15 dB (1.5 dB additional cable loss for the upstream trip plus the previous 13.5 dB). The signals are now reflected by the amplifiers output mismatch: a return loss of 16 dB is common. At this point, the reflected signal has an amplitude that is 31 dB below the primary signal and delayed by the round-trip propagation through 300 feet of cable, which takes about 350 nanoseconds. The signal is horizontally delayed approximately 1/7-inch on a 27-inch television set. This is not enough to become a visible ghost or second image. However, depending on the relative phases of the RF carriers of the primary and reflected signal, the visual effect may be enough to cause a softening of a previously well-defined

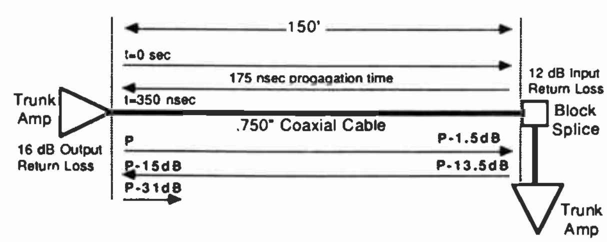


Figure 12. Example of system reflection.

luminance level transition. With repeated mismatches, the crispness of the pictures may be noticeably reduced. This softening effect is easily seen on displays of character-generated text pages.

Phase Noise

Phase noise is added to the original signal through modulation and frequency conversion processes. A significant amount of phase noise must be added to the video carrier before generated impairments become perceptible. Narrow-band phase noise (measured 20 kHz from the video carrier) in a TV channel produces variations in the luminance and chrominance levels that appear as an extremely grainy pattern within the picture. The perceptibility level of phase noise on the video carrier is 53 dB below the carrier at 20 kHz. If the frequency conversion or modulation processes are operating close to specification, phase noise impairments should not be perceptible on the customer's TV unless the converter/descrambler is malfunctioning or is of poor quality.

Amplifier Distortions and Their Effects

New amplifier technology based on feedforward and power-doubling techniques increases power levels with fewer distortions. However, additional sources of minutely delayed signals have been created. The signal delays produced in these amplifiers have similar end results in picture degradation as the delayed signals generated by reflected signals in the cable plant. But they are caused by a different mechanism. These amplifiers use parallel amplification technology. The signals are split, separately amplified, and then recombined.

With a feedforward amplifier, the signals are purposely processed with delay lines. If the propagation time is not identical through each of the amplifiers' parallel circuits, signals will be recombined that are delayed by different amounts of time. In most circumstances, the amount of differential delay is small and will not produce a visible ghost, but it may cause loss of picture crispness. Since the hybrids used in these amplifiers are normally provided in matched pairs or in a single hybrid package, these delays are only a problem when the hybrids are not replaced as a matched set.

In systems that carry more than 30 channels, CTB is the limiting distortion. However, cross-modulation (X-MOD) distortion, which is often the limiting factor in systems with less than 30 channels, can reappear as the controlling factor in dictating system design. The HRC and IRC channelization plans previously described were developed to minimize the visible degradation in picture quality that is caused by CTB.

X-MOD is one of the easiest distortions to identify visually. Moderate cross modulation appears as horizontal and vertical synchronizing bars that move across the screen. In severe cases, the video of multiple channels is visible in the background.

Moderate CTB is the most misleading distortion since it appears as slightly noisy pictures. Most technicians conclude that there are low signal levels and CNR problems. CTB becomes visible as amplifier operating levels exceed the design parameters. Once CTB reaches a severe stage, it becomes more readily identifiable because it causes considerable streaking in the picture.

Composite second-order beats (CSO) can become a limiting factor in systems that carry 60 or more channels and use the HRC or IRC channelization plans. This distortion appears as a fuzzy herringbone pattern on the television screen. The CSO beats fall approximately 0.75 MHz and 1.25 MHz above the video carrier in a television channel. An IRC channelization will frequency lock these beats together, while increasing their amplitude relative to the carrier level.

Hum modulation caused by the 60 Hz amplifier powering is identified by a characteristic horizontal bar that rolls through the picture. If the hum modulation is caused by the lack of ripple filtering in the amplifier power supply, it will appear as two equally spaced horizontal bars that roll through the picture.

Frequency Bands Affected by Radio Frequency Interference

Discrete beat products can be difficult to identify by the displayed picture impairment. Radio frequency interference that leaks into the cable system from nearby RF transmitters causes spurious carriers to fall in the cable spectrum. Common sources of signal leakage are cracked cables and poor quality connections. When either of these situations happen, strong off-air television and FM radio broadcast signals interfere.

If television stations are carried at the same frequency on cable as broadcast and the headend channel processing equipment is phase locked to the off-air signal, the effects of this interference will be ghosting. The ghost appears before (to the left of) the cable signal, since propagation time through the air is less than through cable. If the signals are not phase locked together, lines and beats appear in the picture.

Under HRC or IRC channelization plans, the headend channel processing and modulating equipment are already locked to a reference oscillator. It is not possible to lock the unit to both the cable channelization plan and the off-air signal. In the IRC plan it may be desirable to unlock a channel from the reference source and lock it to the off-air station. However, the group phase locking advantage is lost, and several dB of distortion resistance on that channel is sacrificed. With HRC systems, video carrier frequencies fall 1.25 MHz lower than their off-air channel's counterpart. It is impossible to carry any channel on the off-air assignment unless the upper adjacent channel slot is not used. With the exception of the new 550 MHz cable systems, the interference from these sources is limited to the VHF channels 2–13 and the FM broadcast spectrum of 88–108 MHz.

Often consumer electronics hardware causes interference from off-air signals. If the internal shielding of the equipment is inadequate, the internal circuits will directly pick up the signal. This phenomenon is called DPU for direct pick-up interference. This is the original motivation for cable converters. Early set-top boxes tuned no more channels than the TV set, but they protected against DPU by incorporating superior shielding and connecting to the TV set through a channel not occupied off-air.

DPU can be misleading. When the subscriber switches to an antenna, he might receive better pictures than from the cable connection. He concludes that his TV receiver is operating correctly and the cable system is faulty. The only convincing argument is a demonstration with a receiver which does not suffer from DPU. Viacom Cable has measured off-air field intensities of eight volts per meter. The German specification for immunity to DPU is four volts per meter. The U.S. has no such specification. However, TV receivers sold in the U.S. are built to comply with the Canadian specification of one-tenth of a volt per meter. This is inadequate. VCR tuners are generally inferior to TV tuners because the VCR market is even more price competitive. The Electronic Industries Association (EIA) and NCTA Joint Engineering Committee is studying this issue under its work on IS-23, the Interim Standard on the RF cable interface.

The second most likely sources of radio frequency interference are created by business band radios, paging systems, and amateur radio operators. These signals leak into the cable system and interfere with cable channels 18 through 22 and channels 23 and 24 (145–175 MHz and 220–225 MHz). It is easy to determine that these signals are caused by an RF transmitter because of the duration of the interference and, sometimes, by hearing the broadcast audio. Since the signals are broadcast intermittently, it is almost impossible to determine the exact location(s) of ingress. Cable systems that operate above 450 MHz may find severe forms of interference. They are subjected to high-power UHF television stations, mobile radio units and repeaters, as well as a group of amateur radio operators signals in the top ten to 12 channels. The extreme variation of shortwave signals in time and intensity makes location of the point(s) of infiltration of these signals difficult.

The upstream 5–30 MHz spectrum is a problem for operators who have two-way cable systems. There are many sources of interference and these signals accumulate upstream. In a two-way plant, a single leak in the system can make that portion of the upstream spectrum unusable throughout the entire plant; whereas in the downstream spectrum a leak may only affect a single customer's reception.

SIGNAL SECURITY SYSTEMS

Means of securing services from unauthorized viewership of individual channels range from simple filtering schemes to remote controlled converter/descramblers. The filtering method is the commonly used method of signal security and is the least expensive.

Trapping Systems

There are two types of filtering or trapping schemes: positive trapping and negative trapping. In the positive trapping method, an interfering jamming carrier(s) is inserted into the video channel at the headend. If the customer subscribes to the secured service, a positive trap is installed at the customer's house to remove the interfering carrier. The positive trapping scheme is the least expensive means of securing a channel where less than half the customers subscribe.

A drawback to positive trap technology is its defeatability by customers who obtain their own filters through theft, illegal purchase, or construction. Another drawback is the loss of resolution in the secured channel's video caused by the filter's effect in the center of the video passband. Preemphasis is added at the headend to correct for the filter's response, but loss of picture content in the 2–3 MHz region of the baseband video signal remains. New positive trap schemes take advantage of sharper surface acoustic wave (SAW) or crystal filter technology. This approach allows the interfering carriers to be positioned a few kilohertz away from the secured channel's video carrier in contrast to 2.25–2.75 MHz with the conventional approach. Illegal construction of these filters is much more difficult. The new technology has the promise of achieving greater signal security and a higher quality picture for the authorized customer, but at a higher price.

Negative trapping removes signals from the cable drop to the customer's home. The trap is needed for customers who do not subscribe. This is the least expensive means of securing a channel when over half the customers subscribe. The negative trap is ideal. There is no picture degradation of the secured channel because the trap is not in the line for customers who take the service. A drawback occurs for customers who do not subscribe to the secured service but want to view adjacent channels. These customers may find a slightly degraded picture on the adjacent channels due to the filter trapping out more than just the secured channel. This problem becomes more significant at higher frequencies, due to the higher Q (efficiency) required of the filter circuitry. From a security standpoint, it is necessary for the customer to remove the negative trap from the line to receive an unauthorized service. Maintaining signal security in negative trapped systems depends upon ensuring that the traps remain in the drop lines.

Scrambling and Addressability

There are two classes of scrambling technologies: (1) RF synchronization suppression systems, and (2) baseband scrambling systems.

The concept of addressability should be considered separately from the scrambling method. Nonaddressable converter/ descramblers are programmed via internal jumpers or a semiconductor memory chip called a *PROM* (programmable read only memory) to decode the authorized channels. These boxes' authorizations must be physically changed by the cable operator. Addressable converters are controlled by a computer-generated data signal originating at the headend either in the vertical blanking interval (VBI) or by an RF carrier. This signal remotely configures the viewing capabilities of the converter. Impulse-pay-per-view (IPPV) technology is supported by addressable converter/descrambler systems.

RF Synchronization Suppression Systems

Converter-based scrambling systems that perform encoding and decoding of a secured channel in an RF format comprise the commonly used scrambling technology. There are two basic RF scrambling formats. The more common is known as gated or pulsed synchronization suppression. With this method, the horizontal synchronizing pulses (and with some manufacturers, the vertical synchronization pulses) are suppressed by 6 dB and/or 10 dB. This is done in the channel's video modulator at the intermediate frequency (IF). The descrambling process in the converter/descrambler occurs at its channel output frequency. This is accomplished by restoring the RF carrier level to its original point during the horizontal synchronization period. Variations of this method pseudorandomly change the depth of suppression from 6 dB to 10 dB, or only randomly perform suppression.

The other popular format is known as sine-wave synchronization suppression. In this format, the scrambling effect is achieved by modulating the video carrier signal with a sine wave. This causes the synchronization to be suppressed, as well as changing the characteristic content of the basic video information. This encoding process is performed at the IF frequency in the channel's video modulator at the headend. The decoding process in the converter/descrambler is accomplished by modulating the secured channel with a sine wave of the same amplitude and frequency as in the headend, but whose phase is reversed.

A phase-modulated RF scrambling technique based on precision matching of SAW filters constructed on the same substrate has been introduced. This low-cost system is extending operators' interest in RF scrambling techniques for use within addressable plants.

Baseband Scrambling Systems

Baseband converter/descrambler technology provides a more secure scrambling technology for delivering video services. The encoding format is a combination of random or pseudorandom synchronization suppression and/or video inversion. Because the encoding and decoding are performed at baseband, these converter/descramblers are complex and expensive.

Maintenance of the system's video quality is an ongoing issue. The encoders are modified video processing amplifiers. They provide controls to uniquely adjust different facets of the video signal. The potential for set-up error in the encoder, in addition to the tight tolerances that must be maintained in the decoders, has presented challenges to the cable operator.

The addition of digitized and encrypted audio replacing the synchronization intervals is a recent extension to baseband video scrambling. While there are issues of Broadcast Television Systems Committee (BTSC) stereo compatibility, providers of these products claim that the combination of video scrambling with audio encryption provides the highest degree of security which can be afforded by the industry.

Off-Premises Systems

The off-premises approach is compatible with recent industry trends to become more consumer electronics friendly and to remove security-sensitive electronics from the customer's home. This method controls the signals at the pole rather than at a decoder in the home. This increases consumer electronics compatibility since authorized signals are present in a descrambled format on the customer's drop. Customers with cable-compatible equipment can connect directly to the cable drop without the need for converter/descramblers. This allows the use of all VCR and TV features.

Signal security and control in the off-premises devices take different forms. Nearly all off-premises devices are addressable. Specific or all channels are controlled remotely. While there were several attempts to take modified addressable converter/descramblers and enclose them on the pole, this approach was not successful. The system was costly and less consumer friendly than having the converter/descrambler in the home because it delivered only a single channel at a time to the customer's equipment.

Interdiction technology involves a scheme similar to that of positive trap technology. In this format, the pay television channels to be secured are transported through the cable plant in the clear (not scrambled). The security is generated on the pole at the subscriber module by adding interference carrier(s) to the unauthorized channels. An electronic switch is incorporated allowing signals to be turned off. In addition to being consumer electronics friendly, this method of security does not degrade the picture quality on an authorized channel. For these reasons, the industry feels that this signal control technology may offer promise.

THE SIGNAL AND THE CUSTOMER'S EQUIPMENT

Signal Splitting at the Customer's Premises

The common devices at the cable drop to the customer's home are grounding safety devices called, *ground blocks*, and a two-way signal splitter that sometimes has a built-in grounding terminal.

Some systems use ground blocks or two-way splitters that incorporate a high-pass filter. These filters are used in the two-way plant to minimize RF ingress into the cable plant in the 5–30 MHz reverse (upstream) spectrum. These filters have a low enough crossover frequency not to cause group delay in the downstream video channels.

Splitters or ground blocks should have little effect on picture quality, provided there is adequate signal to handle the splitter's loss. The signal strength may be below specifications due to an excessively long drop or more activated cable outlets in the house than the cable design anticipated. To compensate, some systems use an AC-powered drop amplifier. These amplifiers can create problems such as a reduced CNR, or increased distortions.

Consumer electronics switching devices, designed to allow convenient control and routing of signals between customer products and cable systems' converters/descramblers, have built-in amplification stages to overcome the losses associated with the internal splitters. These amplifiers add distortions or noise. When cable systems were designed, consumer electronics switching devices were not taken into account because they did not exist.

Signal splitting in VCRs can be a problem. To compensate for recording SNR deficiencies, inexpensive VCRs sometimes split the signal unequally between the by-pass route and the VCR tuner. This gives the VCR a stronger signal than the TV receiver to improve VCR performance. In addition, this strategy reduces the quality of the signal routed to the TV. When it is compared with VCR playback, the disparity in performance is reduced.

Consumer Electronics Compatibility

Cable systems are becoming more consumer friendly by trapping popular secured services. More cable-ready television sets are appearing in the customer's home. With VCRs that are also cable ready, some of the interface issues are becoming simpler. Some television sets have built-in A/B selector switches and video input ports that allow signal source switching to be performed through the television's remote control. Up to three A/B switches, two splitters, and two converter/descramblers have been wired into configurations allowing the consumer to watch and record the programming desired. Even with this configuration, consumers lose the ability to preprogram the VCR to record more than one channel.

Hopefully, the days of complex interfaces will soon be over. A positive step in this direction is the development of the EIA IS-15 Multiport TV Interface. This connection system is oriented toward supporting external, low cost, descramblers. If a descrambler is connected to the TV set or VCR via this technique, the user regains use of advanced features precluded by converters.

FIBER IN CABLE TELEVISION PRACTICE

In the typical cable system, long cascades of amplifiers build up noise and limit bandwidth. While the cable is often capable of transmitting one gigahertz (GHz) and more, the amplifiers are limited to 650 MHz or so. In

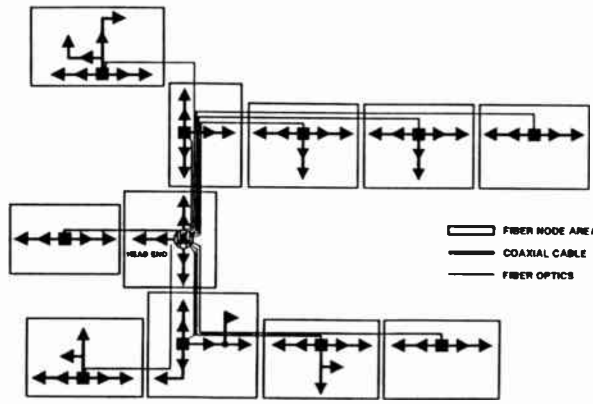


Figure 13. Cable systems with fiber optic backbone trunk.

addition, the long cascades are a serious reliability problem.

The fiber backbone approach breaks the cable system into a multitude of much smaller cable systems with amplifier cascades limited to four to six amplifiers. Each of these small cable systems is fed with a fiber link to the headend. (See Fig. 13.) The advantages include significantly lowered vulnerability to amplifier outages, reduced bandwidth restrictions and noise build-up due to amplifiers in series, and greatly reduced ingress. The latter effect makes two-way cable practical. A major attraction of the fiber backbone is that its implementation cost is low. Fiber is *over-lashed* onto the existing trunk plant. The existing cable is broken into segments and used for the small-scale cable systems. Some of the amplifiers are reversed in direction. Nothing is wasted.

In the typical cable system, 14% of the cable footage is in the trunk, 40% of the footage is in the tapped distribution, and 45% is in the drops, with 1% used for other purposes. The cost effectiveness of the fiber-backbone technique stems from the fact that it involves only 10% of the plant footage, since less than the total trunk footage is involved in the upgrade. In some design studies, the cost of implementing this upgrade came to less than $25 per subscriber. The fiber backbone effectively cures most of the ills of the trunk part

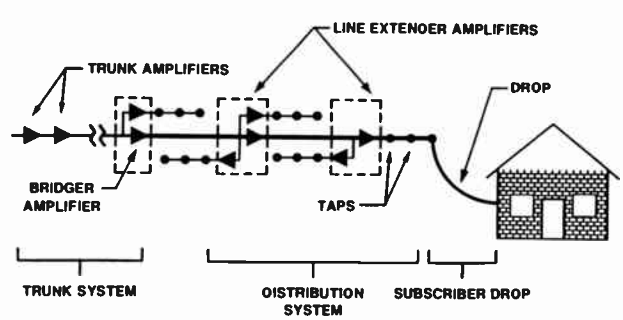


Figure 14. "Super" distribution plant.

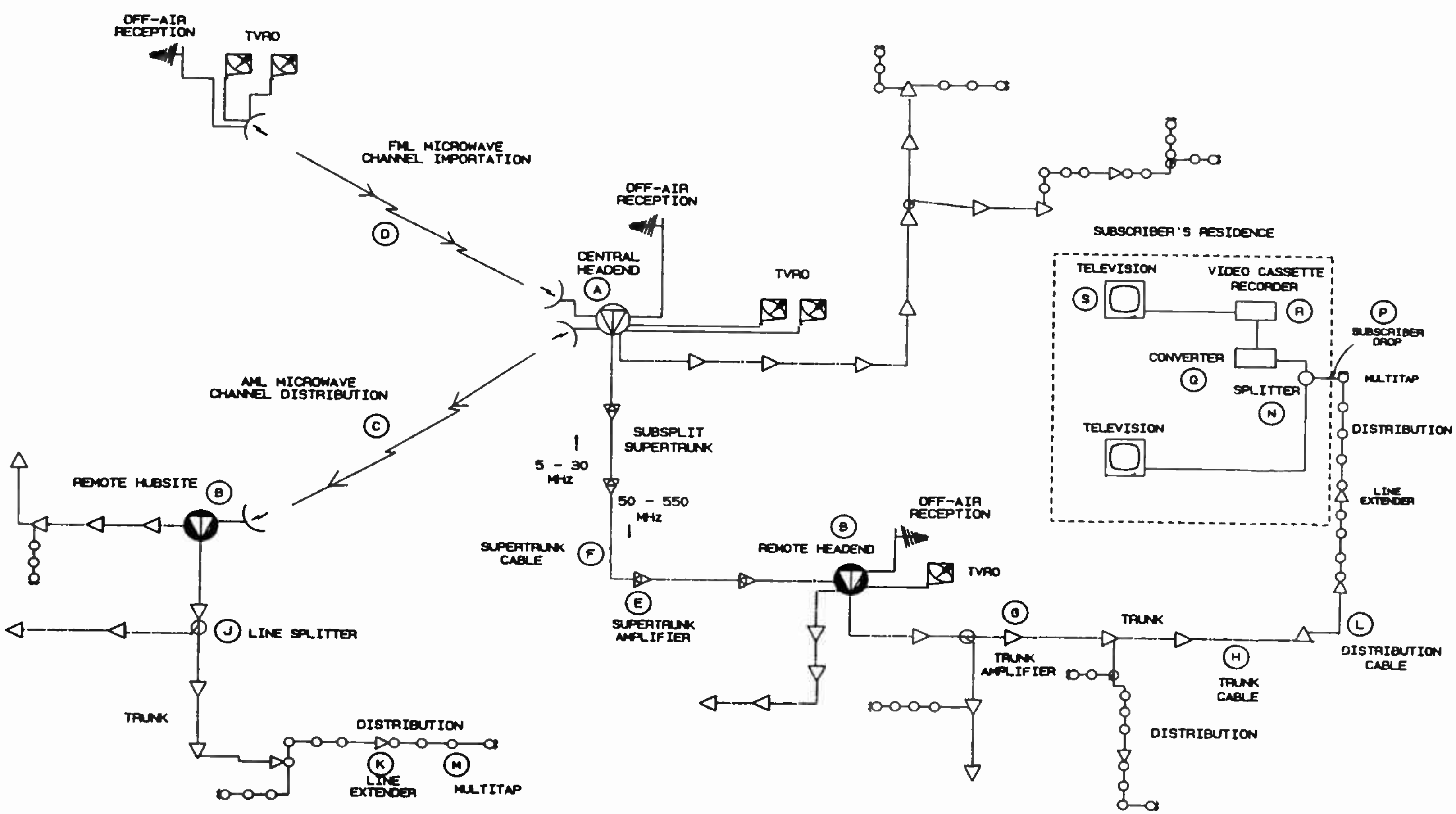


Figure 15. Generic cable system.

of the plant. This improves NTSC delivery and is one way to prepare the trunk plant for HDTV.

If we now turn our attention to the distribution plant, we find that we must run higher signal levels to support the tapping of energy to serve drops to customers. The higher signal levels mean that we begin to reach into nonlinear areas of the amplifier's operating characteristic. Nonlinear distortion builds up. In addition, the taps are not perfectly impedance-matched to the cable. Consequently, the signal is reflected back and forth between the taps resulting in a smearing of the picture. This phenomena is called *microreflection* for two reasons. The strength of the reflection is low and the time delay of the reflection is short.

Rogers Cable of Canada has suggested the answer to these difficulties. They have called their technique Super Distribution. In one form of its implementation, line-extender amplifiers are structured to have up to three hybrid amplifier chips. One feeds the next line extender amplifier, one feeds half of the taps back to the previous line extender, and the third feeds half of the taps to the next line extender. The existing tapped feeder cable is cut in half between line extenders. New, untapped cable is over-lashed to connect line extenders. The consequences are that the signal level between amplifiers is lower since that cable is not tapped. The signal level on the tapped runs is lower because they are shorter. Nonlinearities are reduced, and they do not build up as in the previous structure. Also, signal leakage may be less of a problem because of the lower signal levels. In addition, the number of taps in series in any cable is drastically reduced, thereby reducing the amount of microreflections experienced by any one subscriber. This technique effectively cures the ills of the distribution portion of the plant. Rogers estimates the cost of this upgrade to be less than $25 per subscriber. See Fig. 14.

With fewer amplifiers in series, the constraints on their design and operation are reduced. Higher bandwidths become practical.

Recently, an extension of the fiber backbone approach, which takes fiber farther into the plant, was announced by ATC. In one version, passive splitters are added to the fiber run from the headend. Then shorter runs into the neighborhood bring the optical plant closer to the home. No trunk amplifiers or trunk cable is used. Even fewer amplifiers are interposed between the subscriber and the headend. In another implementation, low cost lasers are used as repeaters to feed the branches at the end of the trunk run. The potential of optical amplifiers promises to yield further evolution of this technique.

BIBLIOGRAPHY

The Optimum Gain for a CATV Line Amplifier, K.A. Simons, Proceedings of the IEEE, Vol. 58, No. 7, July 1970.

NCTA, Technical Papers, 38 consecutive editions, National Cable Television Association, Washington, DC.

NCTA Recommended Practices, National Cable Television Association, Washington, DC.

Getting The Picture, Stephen B. Weinstein, IEEE Press.

Color Television, Edited by Theodore S. Rzeszewski, IEEE Press.

Television Technology Today, Edited by Theodore S. Rzeszewski, IEEE Press.

Cable Communication, Baldwin, Thomse, and McVoy, Prentice-Hall, 1983.

Cable Television, John E. Cunningham, Howard W. Sams.

Cable Television Technology, James N. Slater, John Wiley.

Cable Television Technology, Deschler, Kenneth T., McGraw-Hill, 1986.

Technical Handbook for CATV Systems, K. Simons, General Instrument.

IEEE Transactions on Cable (no longer active).

IEEE Transactions on Consumer Electronics.

Information, Transmission, Modulation, and Noise, Mischa Schwartz, McGraw-Hill.

5.1 Microphones

Greg Silsby
Greg Silsby & Associates, Akron, Ohio

INTRODUCTION

The effectiveness of a broadcast station's purchase and use of microphones can be improved by an understanding of the underlying principles of the different types of microphones. This chapter reviews the basic types, techniques employed in their construction, and how the different characteristics can be used in various applications for best results.

Microphones will be described in terms of their generating element (transducer) types:

- Carbon
- Ribbon
- Dynamic
- Moving Coil
- Condenser

Next is information concerning performance differences, including frequency response and polar patterns. The section on polar patterns leads into a discussion of various influences on microphone working distance followed by explanations of sensitivity and impedance. The chapter continues with a review of problems and solutions, including:

- Voiding and avoiding noise problems
- Acoustic phase interference with multiple microphones
- Acoustic phase interference with the single microphone
- Barrier miking technique

A section on the various stereo miking techniques places particular emphasis on preserving monaural broadcast compatibility. Finally, suggestions are made regarding microphone care and maintenance.

TRANSDUCER TYPES

Any discussion of microphones should begin with some understanding of what a microphone is and what it does. A microphone is a transducer: a device which, when activated by energy from one system, converts that energy to another form, while supplying it to another system. In the microphone, acoustical energy (sound waves impinging on the diaphragm) is converted to a varying voltage that is the electrical analog of the sound. The method by which the microphone converts acoustical energy to electrical energy is one of the ways by which microphones are differentiated.

Carbon Microphones

The earliest transducer principle used to construct a microphone was variable resistance. This is the principle by which the carbon element operates. The carbon microphone contains carbon granules which are compressed and decompressed by the movement of the microphone's diaphragm as it responds to air pressure changes. As a DC current is passed through the granules, the changes in density result in a variable resistance to the current flow. The resistance change bears the characteristics of both the amplitude and frequency of the sound waves that act upon the diaphragm.

While carbon microphones offer excellent durability, they do require a power supply. More serious deficiencies include their high distortion and limited frequency response range. Their common application in more recent years has been in telephones and other communications areas, such as military and aviation. Even in these fields, carbon microphones are being replaced by specialized dynamic and condenser systems.

Dynamic Ribbon Microphones

The ribbon, or velocity microphone, utilizes a very thin, corrugated metallic foil ribbon suspended within the flux field of a permanent magnet (Fig. 1). The ribbon microphone is also referred to as a velocity microphone because it responds to air particle velocity at the ribbon. The term *pressure gradient* is also used

Figure 1. Construction of a ribbon, or velocity, microphone. (Courtesy of Shure Brothers Incorporated.)

to denote that the pressure exerted on the ribbon is in proportion to the difference between the pressures present on each side. While the ribbon's ends are held in place, the rest of it is allowed to move freely back and forth in a sympathetically-induced mechanical recreation of the amplitude and frequency of the sound presented to it. The ribbon's movement causes it to cut the lines of flux between the magnet's poles, inducing a small AC voltage onto the ribbon. The leads from the ribbon's ends connect it to a step-up transformer which converts the extremely low impedance of the ribbon (approximately 1 ohm) to a usable figure which might lie between 50 ohms and 500 ohms.

While the inherent directional characteristic of the ribbon microphone is bi-directional (figure-eight polar pattern), other patterns have been derived to allow much more flexibility in their application.

Ribbon microphones have historically been noted not only for their potential for delivering a very warm sound, due to their extended low frequency sensitivity, but for their susceptibility to damage from air blasts. Blowing across the ribbon, coughing into it, or subjecting the ribbon microphone to the wind in location recording could easily stretch or break the ribbon. Even rapid panning on a studio boom has caused ribbon failure. Newer designs, however, have provided considerable improvements in durability and a lower failure rate.

Dynamic Moving Coil Microphones

Although the ribbon microphone is a type of dynamic microphone, in common usage, the term dynamic microphone usually refers to a microphone with a moving coil.

The dynamic microphone has a diaphragm with its back attached to a voice coil (Fig. 2). This extremely light-weight coil of wire is suspended in a magnetic field supplied by a permanent magnet structure. The ends of this voice coil are brought out to stronger leads which connect either to a transformer or the microphone's output connector.

Sound waves acting upon the diaphragm cause it to move back and forth in sympathy. The attached voice coil is then forced to cut the lines of flux in the magnetic field causing a small AC voltage to appear at the ends of the leads. This signal should closely emulate the sound waves in frequency and amplitude. The dynamic element is sometimes referred to as a generator or motor mechanism. It is, by design, very similar to a speaker. Loudspeakers are often used as both the speaker and the mike in an intercom.

The diaphragm design is crucial to good dynamic microphone performance. It must be highly compliant to allow effortless excursion at all frequencies of interest. In addition, this movement must be accomplished with maximum linearity and a minimum of break-up modes. Nonlinearities of diaphragm movement would include, for example, any tendency for a rocking motion to be set up which, at some frequencies, would allow one side of the voice coil to be travelling downward while the other side is moving upward. Such rocking results in-phase cancellation and a dip in the frequency response. Break-up modes occur when a portion of the diaphragm resonates independently of the rest of the surface. Again, phase cancellation results and, with it, response anomalies occur.

The design and construction of a high quality dynamic microphone suitable for broadcast use is a blending of science and art. As with other transducers, much of its design may be modeled by the use of equivalent electrical circuits. The process is similar to the computer-assisted design of loudspeakers using

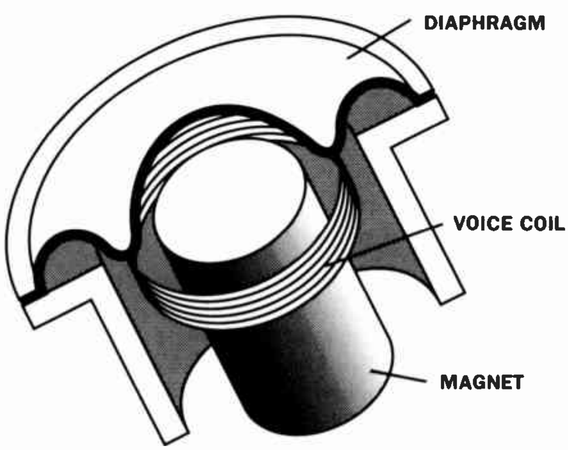


Figure 2. Dynamic moving coil element. (Courtesy of Shure Brothers Incorporated.)

Thiel parameters. In the microphone, however, the seemingly infinite variety of complexities in certain parts of its design is so great that the artistry of the microphone engineer is constantly called upon. In the design process, the goals set before the microphone design engineer are sometimes all but impossible to achieve within the same product. As is true in other areas of engineering, design trade-offs are numerous and the laws of physics tend to win in the end.

Size plays an important role in the performance of the dynamic microphone. Small dynamic motor mechanisms tend to be very inefficient. Their acoustic sensitivity is low while their mechanical sensitivity is high by comparison. The result may be a poor system signal-to-noise ratio (SNR), and a lot of handling or noise transmitted through the mike stand. Internal shock-mount systems may be used to reduce the mechanical excitation but the design goal of small size may then be defeated.

Small size usually means sacrificed low-frequency response in dynamic microphones. This is not to say that large dynamic microphones will always have an extended low frequency response. Very small dynamics will almost certainly, though, be lacking in low frequency output.

Still another physical characteristic of the dynamic microphone that affects its performance is the mass of the diaphragm/voice coil assembly. The greater the mass, the more limited will be the high frequency response. Common design practice includes the use of Helmholtz resonators immediately in front of the diaphragm. These create peaks and effectively extend high frequency response beyond the normal limits of the system (Fig. 3).

Most broadcast-quality dynamic low-Z microphones exhibit an impedance that is a function of the turns and gauge of their voice coil wire. Some older, more

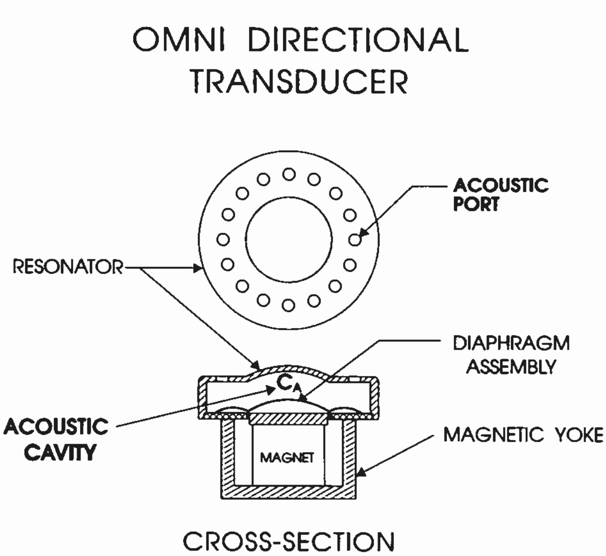


Figure 3. Cutaway drawing of a dynamic element showing helmholtz resonators. (Courtesy of Electro-Voice, Inc.)

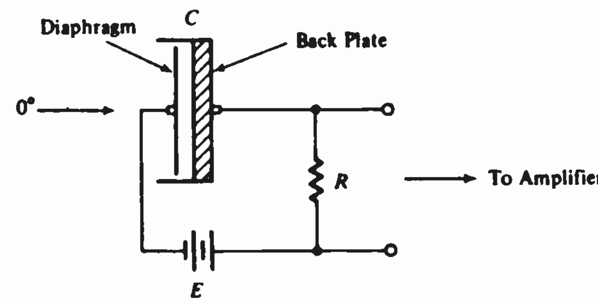


Figure 4. Conventional capacitor microphone system.

public address-oriented microphones employ a transformer within their housing to correct for design tradeoffs in their voice coil. The transformer adds to the microphone's cost and may also restrict performance if it is not a high quality unit, limiting the frequency response and possibly increasing distortion. But properly designed dynamic microphones can be the most rugged of the high-quality transducer types; some have truly become legendary for their ability to provide high quality broadcast audio with virtual bullet-proof construction.

Condenser Microphones

In the condenser microphone, a capacitor forms the generating element (Fig. 4). One side of the capacitor is the diaphragm; the other is the fixed backplate. Air between these two plates acts as a dielectric. The capacitor, of course, must possess a positive electrical charge on one plate and a negative charge on the other. The conventional or discrete condenser gets this polarizing or bias voltage from an external DC power supply. In older condenser microphones, separate leads were required to deliver DC from the power supply to the microphone. Today, *phantom power* is used in most conventional condenser systems to deliver the required DC voltage to the microphone over the same conductors and their shield that are used to carry the audio signal. Upon activation of the power supply, a positive voltage is quickly built up on the rear surface of the diaphragm. This causes an electrical current to flow through the resistor until the surface of the backplate finally receives a negative charge of equal value.

As sound waves (air pressure changes) strike the diaphragm, causing it to move back and forth, the distance between the two plates rapidly increases and decreases. This causes proportional changes in the capacity of the condenser to hold a charge. The result is an AC current flow in the resistor and a voltage across the resistor that corresponds to the excursion of the diaphragm. While this voltage effectively represents the output voltage of the microphone, the source impedance is far too high to be carried for any distance over microphone cable. This output signal, then, is presented to an impedance converter (often an FET) inside the microphone. Power for the impedance con-

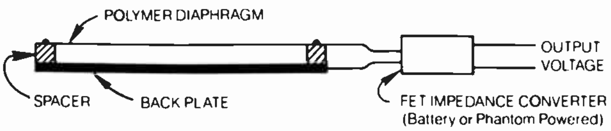


Figure 5. Electret condenser element. (Courtesy of Audio-Technica U.S., Inc.)

verter is derived from the same source which provides the polarizing voltage for the element. The impedance converter delivers a low impedance output that can be fed down long microphone cables with minimal loss.

Electret Condenser

The electret condenser microphone utilizes a material which has the ability to hold a charge applied during the manufacturing process. Most high-quality electrets apply this material to the fixed back-plate of the capacitor (Fig. 5). Some designs employ a charged diaphragm instead, but pay several performance penalties in doing so. Lowering the weight of the diaphragm by moving the electret material to the backplate results in lower handling noise, extended frequency response and improved transient response.

Although the electret functions much like the discrete condenser, but produces its output voltage without the need for an external high-voltage DC supply, an impedance converter is still required. The low voltage needed to power it, however, may be derived from internal or external batteries or an external AC-powered supply.

Phantom Power For Condenser Mikes

Phantom power, or simplex power, provides one means for remotely powering condenser microphones. Phantom power requirements of various microphones may fall anywhere between 9 V and 48 V. External supplies are most often designed to deliver 9 V, 12 V, 18 V, 24 V, 30 V, or 48 V. While many electrets will operate on any of these supplies, some modern discrete condenser designs require 48 V. The phantom supply voltage in nonelectret condenser microphones is often stepped-up by an internal circuit to provide a sufficient capacitor-polarizing charge for good signal-to-noise figures.

In a phantom power circuit, the plus side of the DC supply is applied equally to both of the signal conducting leads of a balanced microphone line. This may be done by means of either a build-out of matched precision resistors, or via a center-tapped transformer. In each case, the return path is the shield. In the microphone, the DC may be similarly tapped via the resistor or center-tapped transformer method to provide the power it needs. The DC is prevented from appearing at the impedance converter output by DC-blocking capacitors or the internal center-tapped transformer.

If a balanced-output dynamic microphone is connected to a line with phantom power present, performance should not be altered; nor should damage occur to the dynamic element. The voice coil or output transformer winding connects across the two signal leads and should see no potential difference between them. Because there is no connection between either lead and the shield (the DC return path), there is no circuit. If an unbalanced dynamic microphone is connected to a phantom supply, the DC will pass through the voice coil and probably destroy it.

A less common powering system called "A-B" or "T" powering is not compatible with dynamic or phantom-powered microphones. A-B power puts the positive side of the DC on one signal lead and the negative on the other. This will damage even a balanced output dynamic microphone.

FREQUENCY RESPONSE

Often one of the first specifications considered on a microphone data sheet is the frequency response range or limits. The required low and high-frequency limits may depend upon the nature of the sound source that is being miked, the medium by which the signal is to be stored and transmitted and the environment in which the miking is to be done.

Often, unfortunately, more attention is given to the response limits than to how the microphone actually sounds in its intended application. Nonlinearities of the response often contribute more to the listener's subjective impression of sound transmitted by the microphone than do the response limits. A specification that reads "Frequency Response 40–18,000 Hz," by itself, says little about a microphone's actual sound in use. Add to that some limits, like ±3.0 dB, and one's knowledge of the microphone, while improved, is still extremely lacking. The shape of the response and polar resonse characteristics will tend to contribute to the character or personality of the microphone's sound.

Response nonlinearities can create acoustic feedback in sound reinforcement, nasality, poor intelligibility, excessive sibilance, muffled sound or any of a variety of other acoustic problems. On the other hand, a microphone's response may be deliberately, carefully tailored by the design engineer to solve problems rather than create them. A rolled-off low frequency response and a rising high frequency response may be employed in a microphone that is intended for use at a considerable distance (Fig. 6). This tailoring can reduce the effects of unwanted low frequency information, such as traffic noise or the rumble of air handling systems, while boosting the high frequencies that are normally attenuated with distance.

Some microphones that are intended to be worn on the body exhibit a response that compensates for the chest cavity resonance which they tend to pick up in a lavalier mounting position.

A rolled-off low end response may help considerably in attenuating handling or stand-born noise as well as wind noise or the breath blasts of plosives in speech.

Unfortunately, published specifications can serve as only one guide in understanding a microphone. The

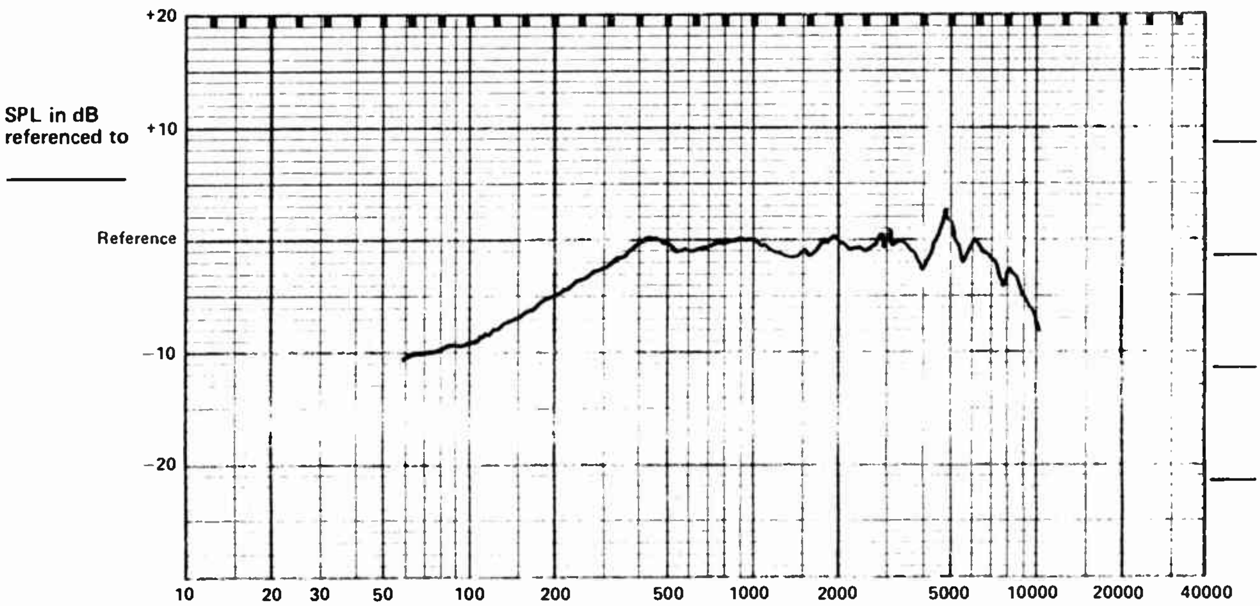


Figure 6. Frequency response curve of a shotgun microphone tailored for distant miking.

extent to which they serve any function at all is, of course, dependent upon the credibility of the source. This may help to explain why a very poor sounding $15 microphone might show limited specifications that appear superior to an excellent sounding $150 microphone.

Certainly, a response curve drawn with the microphone directly facing the sound source (on axis response) will give us a clearer picture of the instrument than will a statement of its response limits. Even curves from well-respected manufacturers, however, may be difficult to compare due to the variety of test procedures and standards. For example, the frequency response chart is an X-Y graph that compares decibel output to frequency. If that chart is compressed vertically or stretched horizontally, the response curve will automatically appear more linear. Frequency response plots for the same microphone may also vary greatly when run on the same equipment, with the same graph paper. Such variables as the pen speed, damping or even the direction of the tone sweep (low to high or high to low) may result in vastly different curves.

Directional microphones present still another possible anomaly in the testing process. These microphones exhibit a phenomenon known as *proximity effect* which results in a bass-boosted output when close-miking a small sound source. Proximity effect is neither good nor bad; its value is dependant upon the intended application. Various designs will differ in the amount of proximity effect that is possible to attain. A single response curve of a directional microphone, tested at some particular (or worse, unknown) distance may be of little value to someone who wishes to use it at another. Ideally, directional microphone data should include close and distant curves (Fig. 7).

POLAR PATTERNS

As difficult as it might seem, at times, to pick up a desired audio signal, the real problem more often lies in eliminating unwanted sounds. Microphones with various directional patterns are often enlisted to improve the ratio of desired signal to ambient noise or other unwanted sounds.

For most miking applications, a suitable on-axis curve alone may not be satisfactory. Ambient noise, "leakage" from other instruments in a band or orchestra, room reverberation, and feedback potential from PA floor monitors are some of the reasons why it is important to know the off-axis response of a microphone. The best view of the microphone's off-axis response is obtained by examining several different polar plots. These should be drawn at low, mid-band, and high frequencies. Overlaid, these plots should reveal how well the microphone maintains its directionality at each frequency.

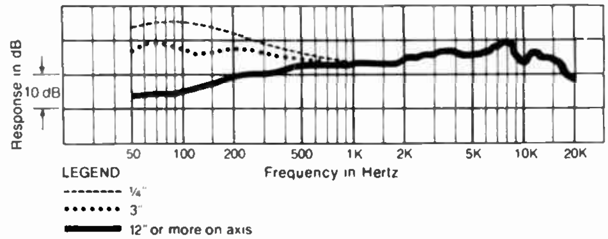
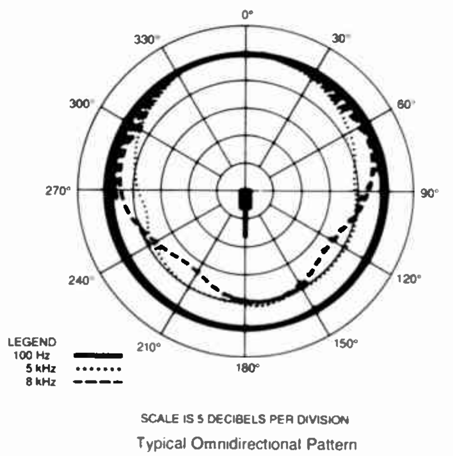


Figure 7. Influence of proximity effect on a directional microphone response. (Courtesy of Audio-Technica U.S., Inc.)



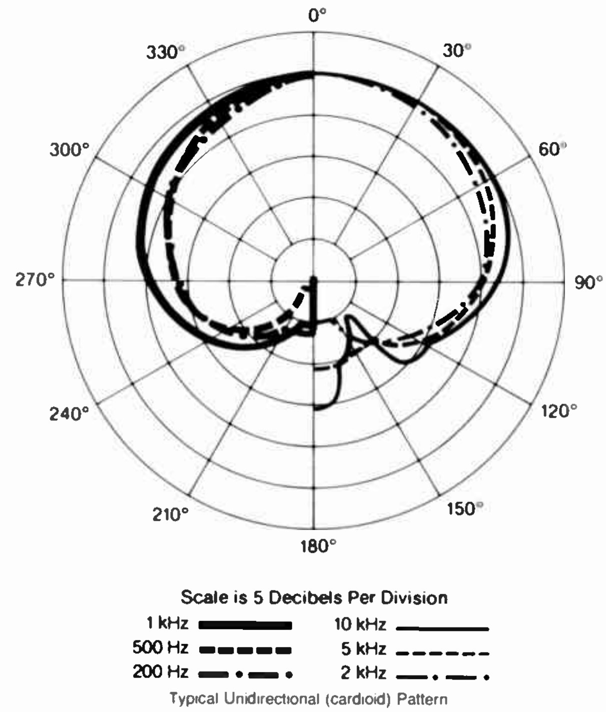


Figure 8. Polar patterns drawn at several frequencies. (Courtesy of Audio-Technica U.S., Inc.)

There are several broad classes based on fundamental polar patterns, to which most microphones' directional characteristics conform, to some extent or another. These include:

- Omnidirectional
- Bidirectional
- Cardioid
- Supercardioid
- Hypercardioid

The accompanying chart shows the relative data for each of the patterns (Fig. 9). The data, however, is taken from mathematical models representing the perfect polar characteristic for each example. Actual microphone polars may vary from near perfection to close resemblance. In the real world, the microphone design engineer must go beyond math equations to accomplish a desired axial response and sensitivity while maintaining polar uniformity. It is truly a blending of art and science.

Polar Scaling and Dynamic Range

Care should be taken when reading polar patterns to observe several variables in the way that they are represented. These variables can drastically alter one's perception of the mike's directionality if they are not examined closely.

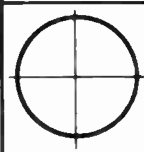
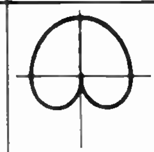
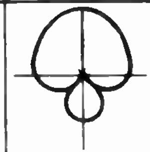
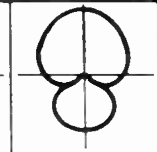
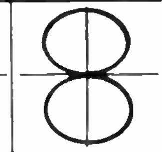
First, check to see whether the scaling is logarithmic or linear. A log scale (the most commonly used) will show a fairly modest inward curve of the cardioid pattern at 90°, indicating a 6 dB drop in level. The linear scale polar for the same microphone will show a polar pattern that appears much more directional. The outside circle of the linear polar represents 100% while the center of the circle equals 0 output. Since a 6 dB loss is equal to a 50% drop in voltage, the polar curve at 90° sweeps in to half the distance between the outside of the circle and its center.

Second, determine the graduations between concentric circles. Are the darker lines 5 dB or 10 dB apart? Finally, take note of the dynamic range of the polar pattern. This may be determined by counting graduation lines inward from the point where the polar crosses 0°, in 5 dB or 10 dB steps (as marked) to the smallest inner circle. Polars may be found in most any range, with 25 dB, 30 dB, and 40 dB all being common. These differences will also alter the shape of a polar pattern.

Keep in mind that the polar pattern represents a cross-sectional, two-dimensional diagram of a three-dimensional function. The 131° arc, for example, that is described by the 3 dB down points on either side of the axis of the cardioid mike can really best be thought of as a conical area within which the mike is virtually uniformly sensitive. This area is often referred to as the mike's angle of acceptance or included angle.

Omni Observations

The omnidirectional microphone is the easiest type to make. It consists of a diaphragm and generating element backed by a totally sealed case. When placed in a sound pressure field, the perfect omni disregards the direction of the sound's origin. A positive pressure

CHARACTERISTIC	OMNI-DIRECTIONAL	CARDIOID	SUPER-CARDIOID	HYPER-CARDIOID	BIDIRECTIONAL
Polar response pattern					
Polar equation	1	$.5 + .5 \cos\theta$	$.375 + .625 \cos\theta$	$.25 + .75 \cos\theta$	$\cos\theta$
Pickup ARC 3 dB down (1)	—	131°	115°	105°	90°
Pickup ARC 6dB down	—	180°	156°	141°	120°
Relative output at 90° dB	0	−6	−8.6	−12	−∞
Relative output at 180° dB	0	−∞	−11.7	−6	0
Angle at which output = 0	—	180°	126°	110°	90°
Random energy efficiency (REE)	1 0dB	.333 −4.8dB	.268 −5.7dB (2)	.250 −6.0dB (3)	.333 −4.8dB
Distance factor (DF)	1	1.7	1.9	2	1.7

NOTE:
1 = Drawn shaded on polar pattern
2 = Maximum front-to-total random energy efficiency for a first order cardioid
3 = Minimum random energy efficiency for a first order cardioid

Figure 9. Microphone polar patterns and their characteristics.

(air expanding) at the diaphragm, for example, causes the diaphragm to move inward regardless of the sound's point of origin (Fig. 12). Such a microphone may also be referred to as a *pressure microphone*.

So it would be with the perfect omni. Most omnidirectional microphones, though, are not truly omnidirectional. The case of the mike represents a barrier to higher frequencies arriving from off-axis. Due to this *case effect*, most omnis are increasingly directional as the frequencies get higher. Likewise, the smaller the omni, the more truly omnidirectional it may be. In addition to case effect, energy arriving at the diaphragm from on-axis is reinforced at those frequencies to which the size of this baffle area is significant. This *baffle effect* causes a rise in the microphone's high frequency output, but only with respect to energy arriving on axis.

Cardioid Comments

Directional microphones employ a damped porting system in their element design which allows sound waves to act upon the rear of the diaphragm as well as the front. The design introduces varying amounts of phase shift for sound arriving from off-axis, resulting in cancellation. The rear entry ports of most directional microphones are spaced at a single distance or "D" from the diaphragm (Fig. 13). Multiple port systems are also available and are designed to reduce proximity effect.

Care should be taken when using any directional microphone not to obstruct the ports with the hand, clothing, gaffer's tape, logo flags, etc. Covering even some of the ports may result in serious degradation of the microphone's directional characteristics and overall sound quality.

The chart in Fig. 9 shows how the various patterns should relate to the reference omni in their ability to reject unwanted energy arriving from various points off-axis. A sound source which delivers 60 dB sound pressure level (SPL) to a cardioid on-axis from one foot away will sound to the microphone as if it has dropped in level to 54 dB (or been moved to two feet away) when the microphone is simply rotated to position the sound source at 90° off-axis. Here a properly designed cardioid emulates well its mathematical model. At 180° off-axis, however, the cardioid can't live up to its equation. The chart indicates 0 output. In reality, well designed cardioids are capable of significant cancellation at 180° but, typically, only on the order of 20 dB or better.

Still, a 20 dB drop in the level of some unwanted noise may be enjoyed simply by turning the cardioid

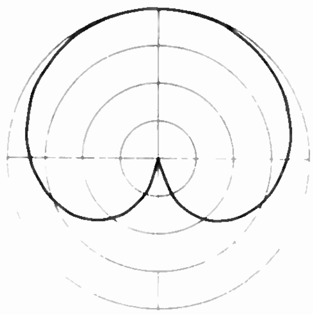


Figure 10A. Cardioid log scale polar. Scale is 10 dB per division.

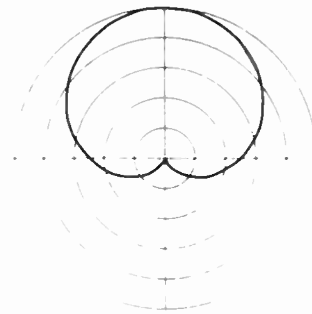


Figure 10B. Cardioid linear scale polar.

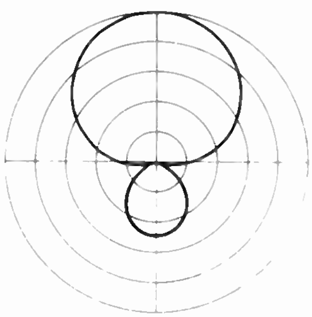


Figure 11A. Hypercardioid linear scale polar. Dynamic range is 50 dB.

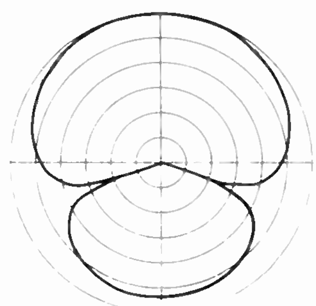


Figure 11B. Hypercardioid log scale polar. Scale is 10 dB per division. Dynamic range is 40 dB.

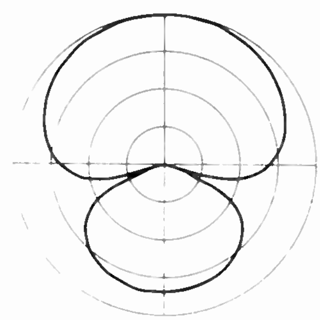


Figure 11C. Hypercardioid log scale polar. Scale is 10 dB per division. Dynamic range is 60 dB.

to face directly away from the noise. That's equivalent to moving the sound source to ten times its actual distance from the microphone.

The 180° response curves (back curve) of many cardioid microphones show their tendency to more closely resemble omnis at both the low and high frequencies. The much more impressive cancellation in the midrange offers just the kind of temptation that sometimes causes a manufacturer to release a data sheet that shows one polar pattern only, and that at some unknown frequency. One of the most beneficial performance advantages one should look for in a well-designed microphone is off-axis linearity. (Fig. 14)

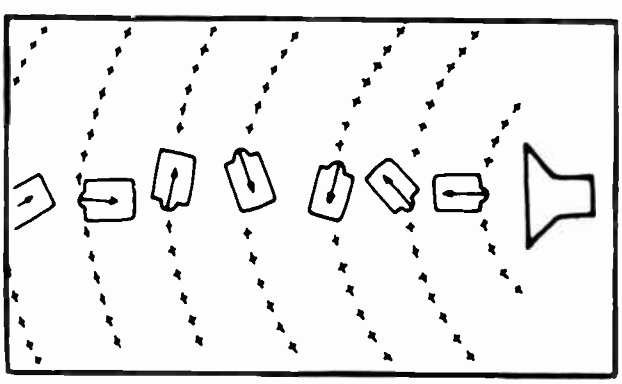


Figure 12. Omnidirectional microphone principle.

Omni vs. Cardioid

Compared to the omnidirectional microphone, the cardioid has several characteristics that should be noted:

1. The cardioid reduces the pick-up of ambient noise and reverberant energy. A look at the comparison chart shows that the random energy efficiency (REE) of the cardioid is 0.333 (compared to a REE of 1 for the omni). The random energy efficiency is a measurement which compares a microphone's sensitivity to random (or reverberant) energy to its on-axis sensitivity. While this shows the cardioid to be one third as sensitive to random ambient noise as the omni, remember that discrete sound sources positioned at the null of the cardioid will be attenuated to a much greater extent. This will prove true outdoors where sounds arrive at the microphone directly with minimal reflections. Indoors, the advantage of the cardioid's deep rear null is only appreciated when the microphone is situated within critical distance of the offending sound source. Up to critical distance, the direct sound is greater in intensity than the reflected energy; after that point, the two remain approximately equal.

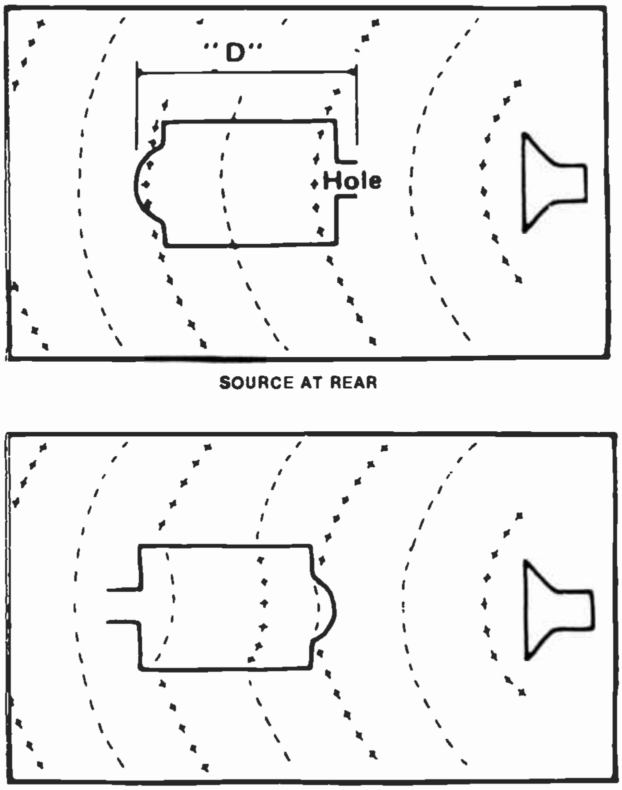


Figure 13. Single-D cardioid microphone operating principle.

2. The cardioid is more susceptible to the problems of "P-pop" (the blast of plosives such as "P" and "T" in speech), wind noise and handling or mechanical noise.
3. The cardioid is more complex to design and construct. Due to this fact, expect to pay considerably more, sometimes, for a cardioid microphone than for an omni of seemingly equal audio quality. The cardioid's more complex construction results in the omni often being the more rugged of the two as well.
4. The cardioid offers greater resistance to feedback in most sound reinforcement applications. This results from its lower REE and is further aided by proximity effect (see characteristic 6 below).
5. The cardioid increases the effective working distance. The comparison chart lists a distance factor (DF) of 1.7 for the cardioid, meaning that a cardioid microphone has a working distance advantage over the omni of 1.7:1. This is calculated on the assumption of a perfect cardioid, in a totally diffuse noise field. An ideal cardioid, then, could be used at a distance of 1.7 times that of the omni for a given ratio of desired, on-axis signal to ambient noise.

   In the real world, nonlinear polar response and the inability of cardioid microphones to achieve total cancellation at their null would seem to reduce greatly the cardioid's working distance advantage. It should be noted that in actual use, however, increasing working distance often has more to do with attenuating a single, offending noise source than with overcoming a diffuse noise field. Directing the deep null of a good cardioid microphone at a whirring still-camera motor drive at a press conference may offer more than a 1.7:1 working distance advantage over an omni.
6. The cardioid exhibits proximity effect. While some designs are quite low in proximity effect, all exhibit some bass-boost phenomenon when used close. Although this may be considered an enhancement in many close-miking applications, care should be taken, however, to avoid input overload or loss of intelligibility that may result from excessive proximity effect.

Other Patterns

An examination of Fig. 9 will quickly show how the three other polar patterns compare to the omni and the cardioid. The hyper-cardioid, for example, combines a tight acceptance angle with superior side rejection and offers the lowest REE. The result is a good pattern for distant miking (often required on a boom mount). Most short shotgun microphones approximate a hyper-cardioid pattern.

Notice that the bidirectional pattern offers the best side rejection, but with no advantage over the cardioid in REE. Bidirectional microphones are typically ribbons or dual-diaphragm condensers. They are excellent for picking up two sound sources in music or dialogue miking with no phase problems.

Compared to the omni, each of the directional patterns exhibits all six of the characteristics described above.

WORKING DISTANCE

Sometimes it is not sufficient to reduce ambient noise merely by using a polar pattern that offers the lowest REE. In very noisy environments (in an aircraft, on a battlefield, in a factory, or at a sporting event) it may be desirable to differentiate between close sound (an announcer, for example) and distant sound.

Microphones which offer considerable proximity effect may be used to advantage in these situations. Close-mike the talent and roll-off the low end as needed to flatten the response. In extreme situations, a noise-canceling (differential) microphone may be required.

Due to the special design (rear ports and back damping systems) of the differential microphone, sound arriving from a distance strikes both sides of the diaphragm with equal intensity and in-phase. A positive pressure on the front, for example, would encounter a positive pressure on the rear of the diaphragm, causing the signal to be canceled.

A combination of inverse square law and port damping causes sound which originates very close to the

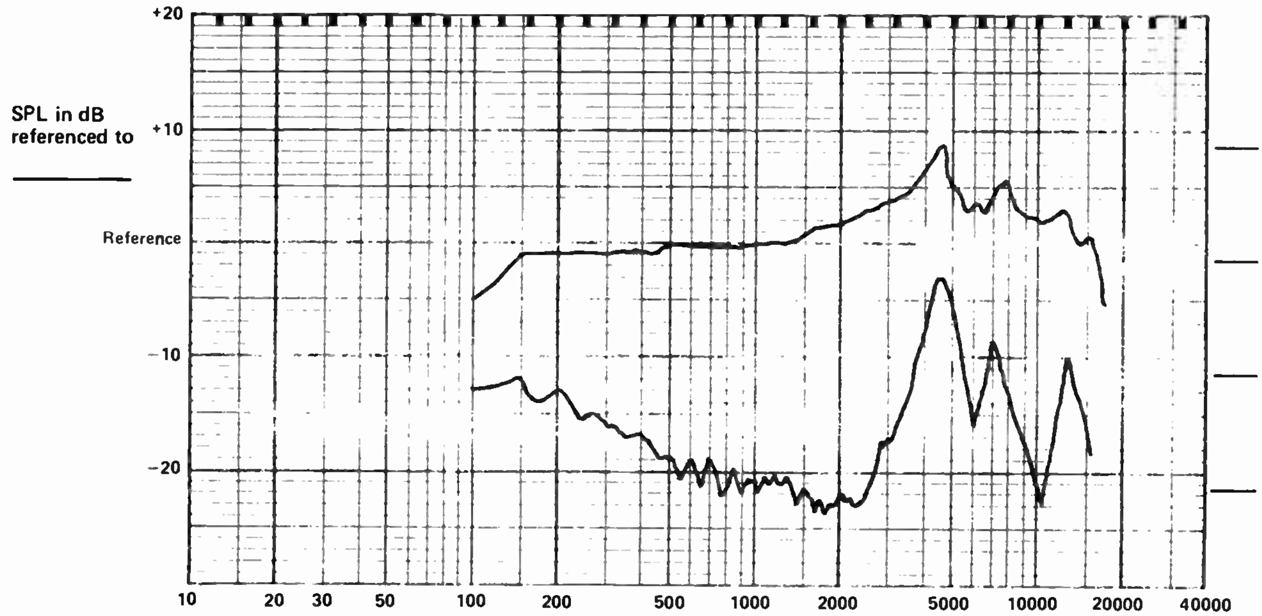


Figure 14A. Front and back curves of a typical cardioid hand-held vocal microphone.

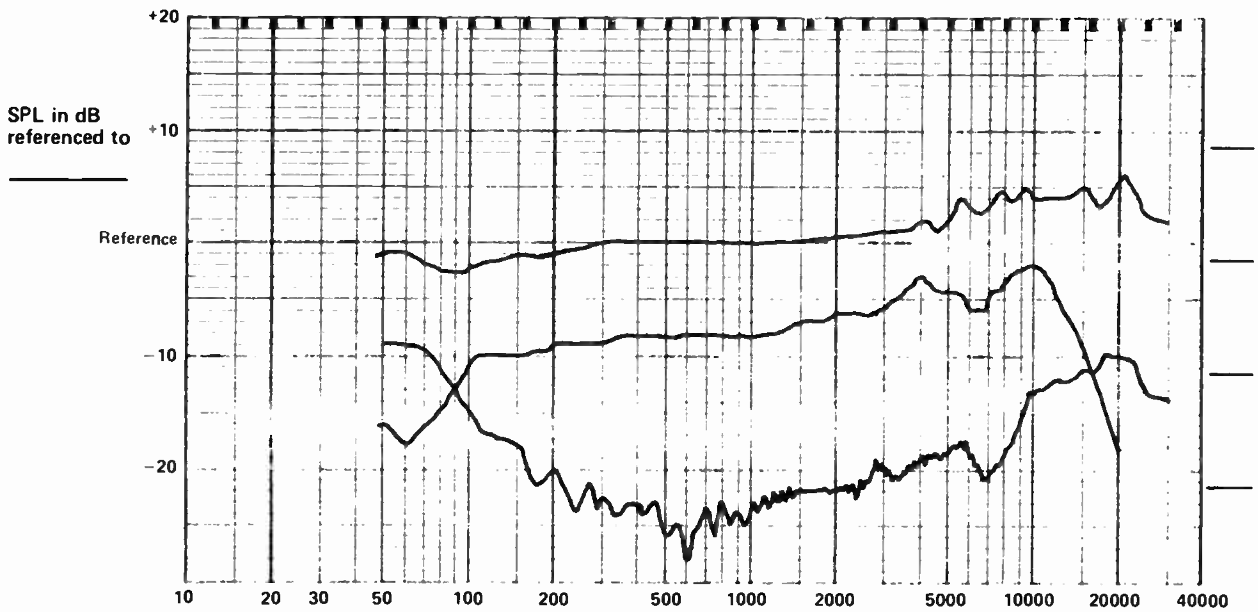


Figure 14B. Front, side and back curves of a high-quality cardioid condenser microphone.

front of the noise-canceling microphone to be far lower in intensity and to exhibit some phase error by the time it arrives at the rear of the diaphragm. The result is diaphragm movement and maximum output. The noise-canceling microphone is able, therefore, to differentiate between close and distant sound sources. The audio quality of such systems normally limits them to voice communication applications.

Inverse Square Law

The easiest, and certainly the cheapest, way to limit the apparent working distance of a microphone is by positioning the microphone very close to the sound source. Inverse square law shows that a decrease in the distance between the microphone and the sound source of one half increases sound intensity (power per unit are) at the microphone by a factor of four and

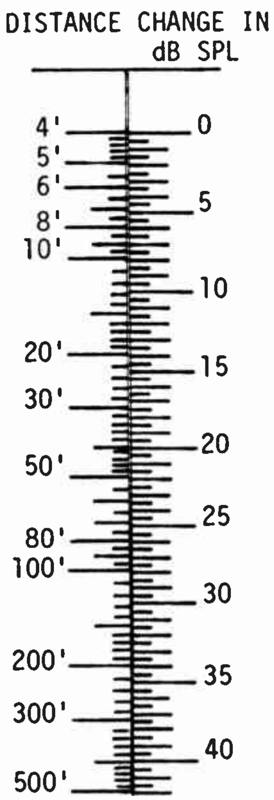


Figure 15. Inverse square law.

the SPL by a factor of two, or 6 dB (Fig. 15). As the input sensitivity control of the mixer or recorder is lowered to compensate for the additional 6 dB available from the now-closer sound source, the microphone, in effect, becomes less sensitive to distant sounds.

Headset Microphones

Headset microphones provide benefits gained from always being a fixed distance from the talker's mouth. Background noise is reduced (due to inverse square law) and levels remain consistent. Omni, cardioid, and differential elements are available in headset systems. Cardioids offer the best combination of ambient noise suppression and acceptable broadcast quality.

Going For The Distance

In some distant-miking applications, the effective maximum working distance may be determined by the electronic signal-to-noise ratio of both the microphone and subsequent amplifiers. For example, the selection of an ideal boom mike for picking up dialogue in a very quiet environment, with no reverberation problems, may have little to do with polar patterns. Instead, one might look for a mike with high output and extremely low self-noise. Most often, though, one is concerned with both electronic signal-to-noise and signal-to-ambient noise ratios.

Shotgun Microphones

Effective working distances beyond those afforded by cardioid, supercardioid, or hypercardioid systems may be realized through the use of a shotgun or line microphone. The line microphone uses a slotted interference tube ahead of the element to provide a high degree of cancellation at the sides. Sound waves arriving on-axis are essentially unaffected by the tube. Sound arriving from slightly off-axis, however, is forced to turn and travel down the tube to the element. This results in numerous out-of-phase conditions being set up in the tube, with cancellation increasing as the microphone is rotated to 90°.

The length of the tube will determine the low frequency limit to which this cancellation can be effective. All things being equal, longer tubes are more likely to maintain their cancellation at lower frequencies than are short tubes. Of course, all things are seldom equal. Newer line mike designs from several manufacturers provide superior pattern control using shorter interference tubes than those required by older standards. The new generation of shorter, lighter products are much easier to handle in both fishpole and studio boom applications.

Some shotgun microphones are much less uniform in off-axis response than simpler, more conventional designs (Fig. 16). Even with their multi-lobed polar patterns, however, their very narrow acceptance angle can often save the day.

Shotguns work best outdoors and in controlled acoustic environments such as well-designed studios. They do not function properly near reflective surfaces. Distant miking down a hallway will not be assisted greatly by the use of a shotgun microphone. A shotgun microphone pointed toward an open window will see all sounds originating on the other side of that window as if they are coming from a source the size of the

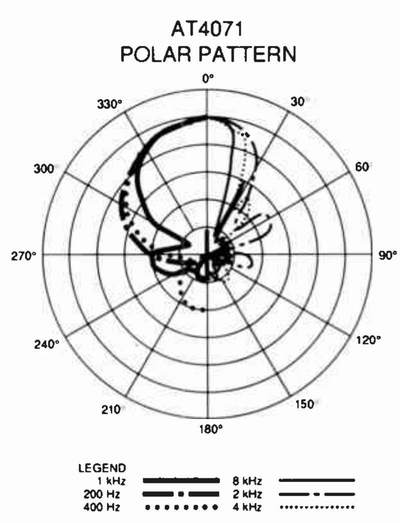


Figure 16. Shotgun polars. (Courtesy of Audio-Technica U.S., Inc.)

window. No directional advantage will be realized other than to reduce, perhaps, the level of some nearby off-axis sounds originating from the microphone side of the window.

Acoustic Gain Devices

Shotgun microphones increase working distance by throwing away off-axis energy, thereby narrowing the acceptance angle. Some devices actually increase working distance by providing acoustical gain. Increased acoustical gain means less electrical gain is required and, therefore, less electronic noise is encountered. Providing that a satisfactory signal-to-ambient-noise ratio may be obtained, the improved signal-to-electrical-noise ratio derived from the use of an acoustical gain device will increase the effective working distance of the overall system.

The most commonly used acoustic gain device is the parabola or parabolic reflector (Fig. 17). The parabolic reflector is shaped so that sound is reflected onto a focal point a short distance in front of the center-point of the dish. An omnidirectional microphone placed at this point receives multiple reflected sound waves in-phase which add together to produce significant gain. The response of such systems is very ragged and limited. Low frequency response is extended as the dish diameter is increased. While totally unacceptable for most broadcast applications, the audio quality achieved with the parabolic microphone is often deemed adequate for sound effects pick-up such as at sporting events. In some situations, acceptable quality means the best level of performance possible through practical or affordable means.

A second type of acoustic gain device is the horn. Low cost re-entrant horns are often used for talk-back in paging systems. They can be quite directional and are extremely sensitive. Their audio quality as a microphone, however, is no better than as a loudspeaker. Horns are sometimes used as microphones in surveillance applications where natural sound is secondary to sensitivity and intelligibility. Installed on the side of a building, the small horn is virtually as inconspicuous as a light fixture and is seldom thought of as a microphone. Some horns are built with 45 ohm voice coils, providing higher output to microphone inputs.

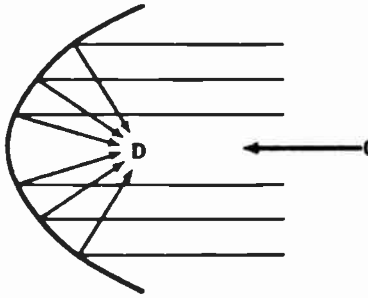


Figure 17. Parabola principle.

Acoustic gain is also realized by using a microphone in the very close vicinity of a large hard, reflective surface. Omnidirectional microphones may be flush-mounted into the barrier, facing out. In this position the microphone is in a half-space environment, or looking into only half the world. The output, for sound arriving on-axis, is increased by 6 dB. As the sound source is rotated off-axis, however, the microphone output drops. At 90° off-axis, the output is down 6 dB, or equal to the omni in free space. The resulting polar pattern resembles a cross-section of a cardioid cut through the microphone at 90°.

See the section "Acoustic Phase Interference" later in this chapter for further details regarding surface mounting.

Frequency Response and Distant Miking

Distant miking may result in noticeable or even dramatic changes in the spectrum of the sound being recorded. High frequencies, attenuated by the air, may require boosting to restore a normal sound. Similarly, high-pass (low-cut) filters may prove helpful in reducing low frequency room reverberation or background noise, thereby extending the useful working distance. If possible, any boost in equalization (EQ) should be done in the recording process to minimize noise that would be amplified in post-production equalization. Microphones with rising high frequency response or a rolled-off low end will reduce the need for EQ.

SENSITIVITY RATINGS

Certainly, one of the greatest sources of confusion on a microphone specification sheet has to be the sensitivity rating. Several different rating systems are currently in use.

Table 1 includes some figures to keep in mind when trying to interpret any of the ratings.

Open Circuit Output Voltage

Microphones are often specified as having a particular output voltage when looking into an open circuit. In most modern equipment, microphone inputs are at least ten times the measured impedance of the microphone and may be regarded as an open circuit. Specifications may be given as an actual output voltage or as decibels below one volt at a sound pressure level of 74 dB (1 dyne/cm^2 or 0.1 Pa). These ratings are referred to as the *open circuit output voltage rating* or *open circuit sensitivity*.

TABLE 1

10 dynes/cm^2	=	+94 dB SPL
	=	10 microbar
	=	1 pascal (Pa)
	=	1 newton/cm^2

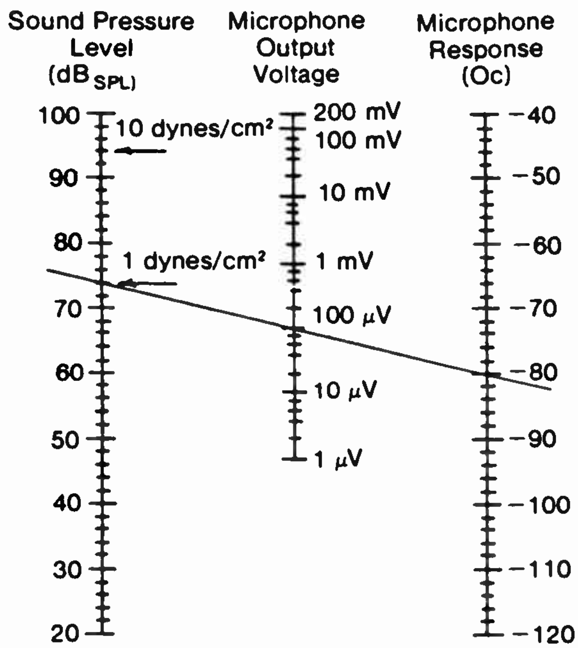


Figure 18. Nomograph: open circuit voltage rating.

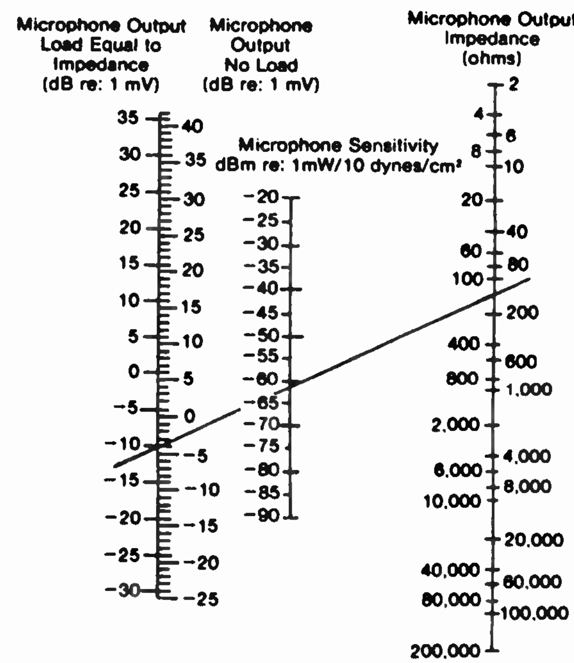


Figure 19. Nomograph: equivalent power rating.

The open circuit sensitivity may be expressed in dB by means of the following formula:

$$V_{OC} = 20 \log E_O - SPL + 74$$

where:

V_{OC} = Open circuit voltage in dB (ref 1 V/0.1 Pa)
E_O = Microphone output in volts
SPL = Actual SPL at the microphone

Power Level

A microphone's sensitivity may also be specified in terms of its output power level. This equivalent power level rating takes into consideration the open circuit rating and either the actual impedance of the microphone or the rated impedance. Specifications given would be in dBm (or just dB) referenced to 0 dB = 1 mW/10 dynes/cm² or 0 dB = 1 mW/Pa.

Calculating the power level rating may be done this way:

$$P_E = V_{OC} - 10 \log_{10} Z + 44 \text{ dB}$$

where:

P_E = Equivalent power level
Z = Impedance of the microphone

EIA Sensitivity Rating

This is one of those specifications that is sometimes specified but hardly ever used. The formula for determining EIA sensitivity is as follows:

$$ESR = V_{OC} - 10 \log_{10} R_{MR} - 50 \text{ dB}$$

where:

ESR = EIA sensitivity rating
V_{OC} = Open circuit voltage in dB
R_{MR} = Center value of the nominal impedance range

For R_{MR}, one may use the following table:

TABLE 2

Table for determining R_{MR}.

RANGE (In Ohms)	CENTER VALUE (In Ohms)
20—80	38
80—300	150
300—1,250	600
1,250—4,500	2,400
4,500—20,000	9,600
20,000—70,000	40,000

Line-Level Microphones

Line-level microphones designed for remote use incorporate a microphone, preamplifier, limiter, and power supply in one hand-held package. Such products were once built by both Shure Brothers and Electro-Voice but are no longer available from either. Line level may be required to get the signal above the level of induced noise (see the section "Hum Rejection"), to overcome resistive losses in long cable runs or merely to level-match to a system that does not provide a microphone preamp at its input. Applications of line-level microphones may include operation into

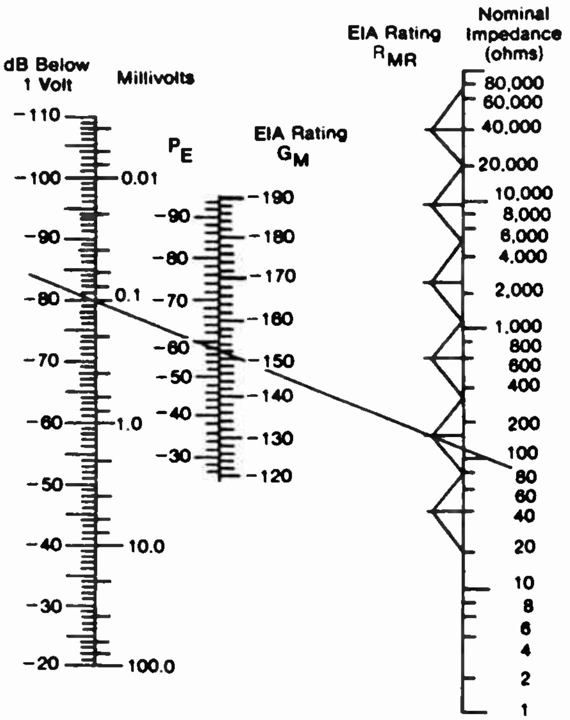


Figure 20. Nomograph: EIA sensitivity rating.

extremely long lines, into unshielded twisted-pair cable, phone lines, or into line-level inputs such as those often found on microwave transmitters. (Also see 6.1–3 on Telco connection.)

A line-level microphone is sometimes preferred over a separate microphone and preamp for simplicity of operation. A separate preamp like the one shown in Fig. 21, however, allows use of a variety of microphones and offers several powering, control, and connection advantages.

OUTPUT IMPEDANCE

The impedance (Z) of a microphone is a measurement of its AC resistance looking back into the transducer. Broadcast microphones should be low impedance, ranging typically from 50 ohms to 600 ohms. Dynamic moving-coil microphones achieve their low impedance by either a low-Z voice coil winding or a transformer. Condenser microphones use an impedance converter circuit to step-down the capacitor's high-Z output.

Low impedance offers the advantages of low susceptibility to hum and electrical noise pick-up and the ability to run relatively long lines with minimal loss of level or high-frequencies.

Unlike matching power amplifier impedances to speaker systems which may be desired for best power transfer, microphones like to see load impedances on the order of ten times their internal impedance or even higher. This assures maximum voltage transfer. A microphone that looks like a resistive source of 150 ohms, looking into a load resistor of 150 ohms, for example, will suffer a 6 dB voltage drop compared to an open-circuit connection.

DYNAMIC RANGE

The difference between a microphone's own self-noise and the maximum SPL it can handle is its dynamic range. In many field applications, ambient noise provides sufficient masking to make the self-noise spec of minor interest. The importance of this specification increases, of course, as greater working distances are demanded or ambient noise levels are lowered.

The impedance converter of condenser microphones, like any active circuit, will create some noise. The degree to which it does so will vary greatly from one design to another. The impedance converter design also determines the headroom, or maximum SPL that the condenser microphone can handle. A maximum SPL of as much as 141 dB is achieved in several high-quality condenser microphones. Such levels may never be encountered by lavalieres worn at a news desk or seldom by shotgun microphones that are more often attempting to pick up weak audio signals.

Dynamic microphones contribute virtually no self noise. When greatly amplified, only the noise of the thermal agitation of air molecules is detected. While this is very low in level, the dynamic microphone does not automatically win the award for first choice in a low-noise system. Because the output level of the dynamic is often lower than that of a condenser system, the user may end up working into the noise floor of the preamplifier in order to provide sufficient system gain.

Some new dynamic microphones employ powerful rare-earth magnets to increase the efficiency of their

Figure 21. Microphone preamplifier. (Courtesy of Shure Brothers Incorporated.)

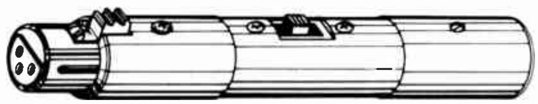


Figure 22. Microphone in-line attenuator. (Courtesy of Audio-Technica U.S., Inc.)

motor mechanisms. Their higher output, while still not as high as many condenser microphones, can provide a considerable S/N advantage over earlier designs.

Properly designed dynamic microphones are practically impossible to drive to audible distortion. The distortion heard when a dynamic microphone is subjected to the lips-touching proximity of a very loud rock and roll vocalist is usually the clipping of the electronics following the microphone. Outputs of one volt or more may actually be delivered in such applications as rock and roll music or hog calls.

Preamp or amplifier clipping may be avoided by padding the microphone output or adjusting the trim (gain adjustment) of the mixer. Be aware that many mixers offer adjustment only after a gain stage or transformer, either of which may distort before any control is possible. In-line attenuators, or pads, are commercially available that allow the selection of 10 dB, 20 dB, or 30 dB of attenuation. These plug directly into the microphone line. Before using any in-line device, it is a good idea to verify that it is compatible with the powering system being used for condenser microphones.

DURABILITY AND RELIABILITY

Modern design and manufacturing techniques have provided the broadcaster with a variety of very durable microphone types. There are a few concepts that generally apply to a microphone design's potential durability:

1. The moving coil dynamic system should offer more inherent durability than condenser or ribbon designs, although both of the latter systems have been employed recently in products that exhibit impressive durability. Also, if dropped on concrete, a lighter condenser mike hits the surface with less momentum than does its heavier dynamic counterpart.
2. Omnidirectional microphones are more easily built to be durable than are more complex directional designs.
3. Internally shock-mounted microphones may present a compromise in durability relative to rigidly mounted, nested designs.

Reliability is influenced by a variety of factors including design, construction, and quality control as well as the care that is taken in the use, transport, and storage of the microphone. Further discussion of this may be found in the last section of this chapter.

VOIDING AND AVOIDING NOISE PROBLEMS

This chapter has dealt already with improving the ratio of desired signal to ambient noise. Other unwanted signals include wind noise, "P-pop," mechanical or handling noise, AC hum and radio frequency interference (RFI). The reduction or elimination of each of these can be handled both through microphone design and user technique.

RFI problems can usually be traced to a point in the low level circuitry at which the signal leads are unbalanced, Hi-Z, or both. Condenser microphones, for example, may sometimes be sensitive to RFI at or around their impedance converter. In such an event, the manufacturer should be consulted for low-pass modifications or information.

P-pops may be reduced through several approaches:

1. Use an omni if possible. Directional microphones are much more prone to P-pop problems than are omnidirectional ones.
2. Position the microphone out of the area of the breath blast. In an announce application, speak across the mike: on-axis but at a slight angle to the diaphragm. Stand-ups and hand-held interview miking should be done with the microphone capsule below the axis of the mouth.
3. Use a pop filter. This is often the same as the manufacturer's windscreen. Avoid using foam filters that are not designed for the particular microphone model. If such a requirement arises, test the combination carefully for frequency response and directional characteristics before putting it into service. Windscreen/pop-filter foam is specially designed, reticulated foam which comes in a variety of densities. Even very acceptable open-cell foam may be too thick for use on some microphones. Nonreticulated foam (such as nerf balls) will roll-off high frequency response and alter the polar patterns of directional microphones.
4. Fashion a pop filter. For radio and other off-camera miking, a piece of silk may be suspended a short distance in front of the microphone diaphragm. This is easily accomplished using a frame made of small, wooden hoops—the type used for embroidery. The outer of these concentric hoops expands for removal from the other ring. Surround the inner hoop with nylon hose material, pulling the fabric to one side where it may be sewn together with thread, and the excess cut off. Replace the outer ring and tighten it with the screw mechanism supplied. Attach the pop filter to an arm that will keep it positioned in front of the microphone. Adjust the distance between the two for best performance.
5. Use a high-pass filter. Most of the disturbing explosive energy of a P-pop is very low in frequency. Try using a very abrupt high-pass (low-cut) filter in the microphone line. Rolling this energy off before it gets to the board or recorder

input will further reduce distortion in the audible range.

Wind noise may be dealt with similarly to P-pop:

1. Omnis are preferred for low wind noise.
2. Use a properly designed windscreen. Most are made of reticulated foam. Superior results with shotgun microphones can be attained by using a well-engineered fabric/mesh cylindrical screen which provides an air space between the material and the microphone (Fig. 23). To handle severe cases (gale-force winds) special fur-like socks are available to wrap around the tubular windscreen. While this will result in some performance trade-offs, recordings made under such conditions are typically not intended for critical listening. Windscreens must cover all openings to the element: front and rear.
3. Use a high-pass filter.
4. A microphone with a limited low frequency response will help minimize wind noise. Extended-response condenser microphones can produce very high outputs of infrasonic energy when panned or when air around them is moved by air handling systems, etc. The result may be preamp overload or undesirable compressor or limiter action. Again, windscreens and or high-pass filters may solve these problems.

HANDLING OR MECHANICAL NOISE

The problem of mechanical, nonacoustic noise is one that plagues the user whether the microphone is hand held, body worn or hardware mounted. The reduction of a transducer's sensitivity to such noise, or the improvement of its acoustic-to-mechanical-noise sensitivity ratio, starts with the basic element design.

First, omnidirectional microphones are lower in mechanical noise than comparable directional systems. Second, condenser microphones, because of their inherently lower diaphragm mass, are superior in this respect to dynamics. Also, as the size of a dynamic system is reduced, such as in dynamic lavalieres or some small hand-held systems, the acoustic sensitivity tends to go down while the mechanical-noise sensitivity may remain fairly constant.

When the resonant frequency of the microphone's mechanical sensitivity is low, a high-pass filter may be used.

Microphone elements are, of course, often internally shock-mounted by the manufacturer to avoid the transfer of noise from the case to the element. Lowest noise is achieved through the combining of omni or even omni condenser systems with internal shock mounts.

External shock mounts are often employed in stand or boom-mounted microphone applications. Properly designed shock mounts allow excursion on-axis, or perpendicular to the diaphragm plane. Excellent mechanical isolation may result from using an external shock mount on an internally shock-mounted microphone. Provided that the resonances of the two systems are dissimilar, they will stagger, increasing the effective isolation.

Another method of reducing mechanical noise is to raise the resonant frequency of the mechanical drive system. An example would be that of bracing wooden platforms, tables or lecterns to eliminate the very audible, drum-like sound produced when they are struck. The use of very high-density materials for microphone support systems will result in a higher resonant frequency. A mike stand set onto concrete or into sand gains advantage from this density.

Mechanical noise transfer to the diaphragm may also be reduced through decoupling the diaphragm from tensile forces, converting them to lateral forces. This may be illustrated as follows:

Select a microphone that has some noticeable handling noise problems and plug it into a talk-out system, raising the gain until the handling noise is evident. Now hold the microphone face up (diaphragm horizontal), with the cable hanging straight down, and tap on the cable. This should produce a low thump. Next rotate the microphone 90° so that the cable is hanging at a right angle to the microphone axis. Tap the cable again and the thump should be all but gone.

Figure 23. Zeppelin-type windscreens for shotgun microphones. (Courtesy of Audio-Technica U.S., Inc.)

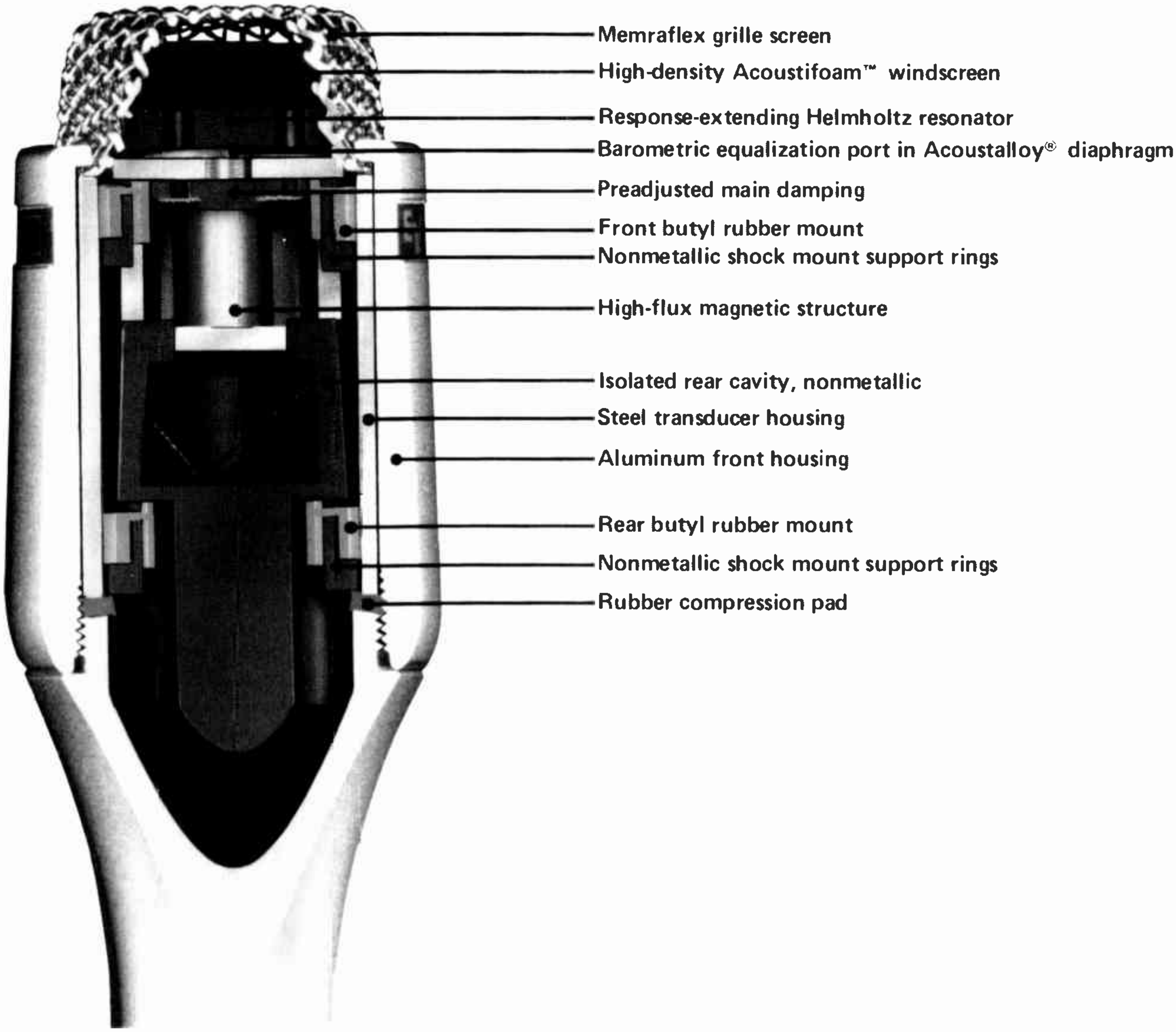


Figure 24. Cutaway of shock-mounted omni. (Courtesy of Electro-Voice, Inc.)

The mechanical drive that takes place longitudinally on the cable is no longer easily transmitted to the diaphragm in the direction of its compliance. Output, therefore is greatly reduced. This principal may be applied to custom hardware designs and the dressing of cables as they enter stand or boom-mounted microphones. A loop of cable or a small coiled cord lowers mechanical noise transfer by this method. Direct interconnection of boom cabling and a shock-mounted microphone can create a direct mechanical short between the two, by-passing much of the shock mount's effectiveness. A coiled cord may be employed between these two points to eliminate this direct path.

AC HUM REJECTION

Microphones may also be sensitive to noise induced in them and their cables by electromagnetic or electrostatic radiation. This may be the result of proximity to power transformers, fluorescent ballasts, high-voltage AC lines, SCR dimmers, etc. The following points should be considered in attempting to avoid or eliminate induced AC hum:

1. Insure that lines are balanced low impedance. The higher the impedance of the microphone, the greater will be the voltage of the electrostatically induced hum. The balanced line insures that nearly equal hum will be induced on each conductor. Little differential is seen, then, at the amplifier input, resulting in common mode rejection of the hum. Pins 2 and 3 of the 3-pin microphone connector should carry the signal, with pin 2 most often "high". This would mean that positive pressure on the diaphragm produces a positive

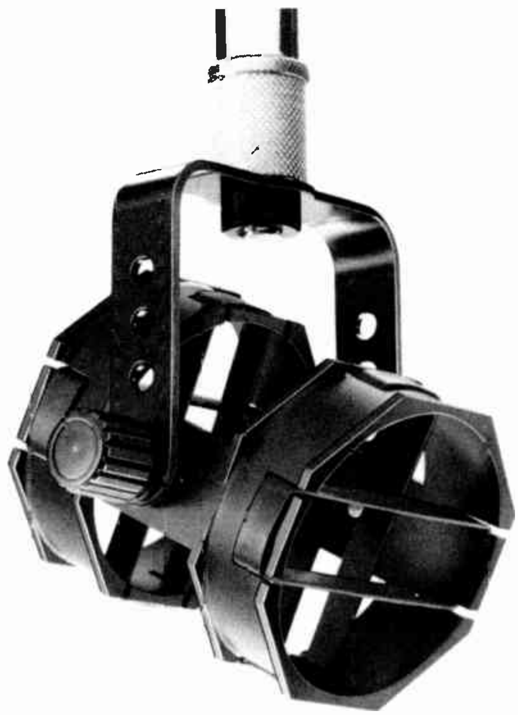


Figure 25. An external shock mount. (Courtesy of Audio-Technica U.S., Inc.)

voltage on pin 2. The shield connects to pin 1 (ground).

2. Route cables with caution. If possible, do not run low-level cables near high-voltage lines, avoiding especially long parallel runs of adjacent microphone and AC power cables. When such cables must cross, they should do so at right angles. If more than one AC line must cross mike lines, separate the AC lines so that they cross at different points.
3. Use twisted pair cable. Leads should be twisted inside the microphone and out. The virtually identical positioning this provides for the two conductors within the hum field, and the fact that they are out-of-phase with each other, will further increase induced hum cancellation.
4. Use well-shielded cable. Installed cable may use a foil shield as flexibility and low flex memory is not a factor. Good stage and field cable, though, normally has a braided shield. Some of the cables which offer the best combinations of flexibility and good shielding use conductive cloth or conductive vinyl under the braided shield.
5. Follow good grounding practices, avoiding ground loops.
6. In general, dynamic microphones are much more sensitive to induced hum than are condenser microphones. The voice coil can be a very effective inductor. Hum-buck coils are employed in some designs which lower electromagnetic hum sensitivity by about 20 dB. The hum-buck coil is wrapped around the outside of the motor mechanism and wired in series with the voice coil but out-of-phase. When both coils are placed into an electro-magnetic field, equal energy is induced onto each. Because they are out-of-phase with each other, the offending signal is cancelled.
7. Transformers located within microphones should be avoided if electromagnetic hum is a possible problem. Some transformers are constructed with hum-bucking characteristics, however, to greatly reduce their hum induction potential.
8. In severe problem situations, operation at line level rather than microphone level may be required.

ACOUSTIC PHASE INTERFERENCE

Another miking problem relates not to sounds that are added to the output, such as popping or hum, but to portions of the spectrum that are greatly attenuated. This change in the microphone's apparent frequency response is due to acoustic phase cancellation. Acoustic phase cancellation may occur when two or more microphones are mixed, or even when a single microphone is subjected to an overdose of reflected sound. Figs. 26A–26E show the severe phase cancellation problems that can result from several typical miking situations.

Although sound arriving at each microphone is identical, originating at the same source, it arrives at the two microphones by paths of varying lengths. This causes a difference in the arrival times and results in-phase cancellation of certain frequencies. The curves given for each of these examples are fast fourier transform (FFT) derived displays of the actual frequency response of the two microphones combined, with respect to a sound source positioned as shown. The FFT analyzer and its companion microprocessor were used to compare the combined output of two matched, calibrated microphones to the output of one of the two microphones by itself. If no phase cancellation occurred, no trace variations would appear on the X/Y plot, as its plot would be a straight line.

It does not take much study of the response charts to see that, no matter which way the microphones and sound source are oriented, the summed response of the two mikes is awful. These experiments reveal quite qraphically what the ear often perceives as a comb-filter or notch-filter effect that sweeps up and down in frequency (and even changes Q) as the variables D1, D2, and D3 change with the movement of the microphones or sound source. In more subjective terms, the resultant sound may be described as hollow, as if the sound is being forced through an empty cardboard tube.

Unfortunately, situations that cause acoustic phase cancellation arise quite frequently. One classic example occurs with a pair of microphones on a podium, spaced apart to provide on-mike coverage as the

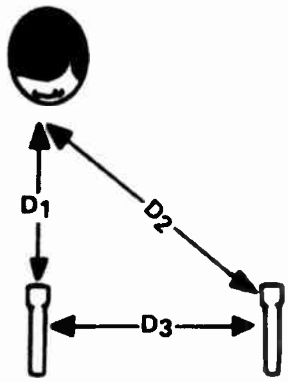


Figure 26A. Phase cancellation with multiple microphones.

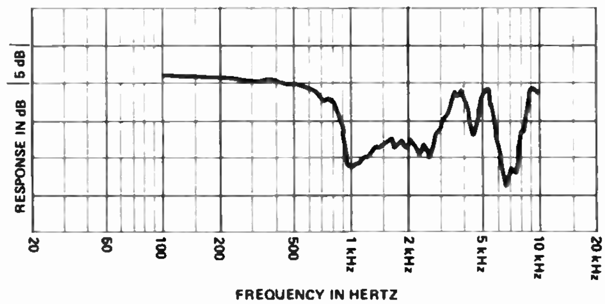


Figure 26B. D_1 = 12″, D_2 = 21.6″, D_3 = 18″

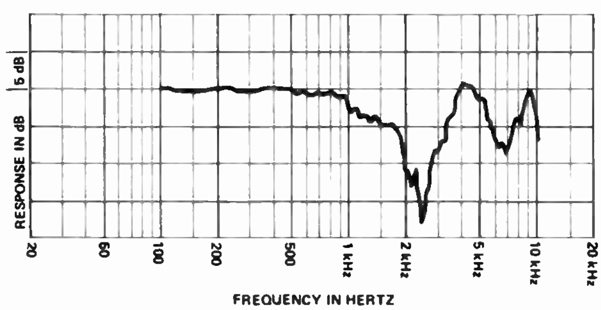


Figure 26C. D_1 = 18″, D_2 = 21.6″, D_3 = 12″

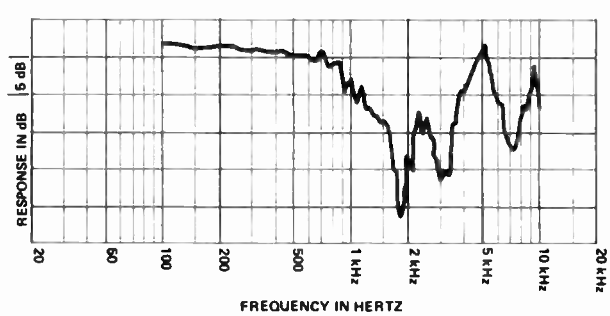


Figure 26D. D_1 = 24″, D_2 = 30″, D_3 = 18″

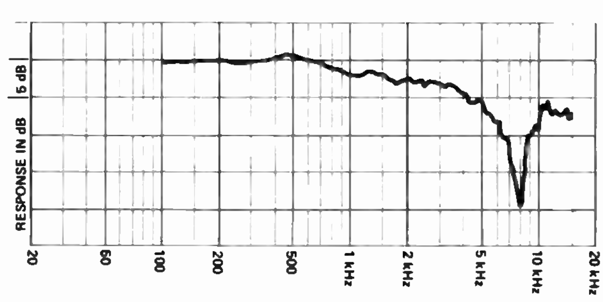


Figure 26E. D_1 = 5.6′, D_2 = 6′, D_3 = 2′

speaker turns his head to address all of the audience in front. The curves shown in Figs. 26B and 26D are typical of the problems caused by this approach. If the output of these two microphones is summed and fed simultaneously to a house sound system, there probably will be gain-before-feedback problems as well. Add the insistent creeping oscillation of a system on the brink of exceeding unity gain to the already intolerable frequency response of the microphone pair, and you'll have the quality of sound that makes the audience check the credits to see who handled the audio.

The simplest solution to the problems caused by this spaced-pair podium miking technique is to use one microphone only, placing it in front of the person speaking and toward the center of the podium. Fig. 27 shows three microphones immediately adjacent to each other. Two (or three) microphones are often employed in this manner when redundancy miking is desired for critical applications. When two are used, place one above the other. Normally, only one of these microphones would be open at a time. The second is strictly a back-up. Sometimes multiple mikes are used to feed separate systems, such as for house PA, government agencies, and broadcast. Each may still be used as a back-up.

Adjacent pairs of cardioid microphones may at times be angled inwardly with their axes crossed and their diaphragms closely spaced. This may be done in order to broaden the acceptance angle of the two mikes while still maintaining some cancellation at the rear. The microphones' close proximity allows their diaphragms to occupy nearly the same point in space, thus reducing sonic time path differences. This ensures that negligible phase cancellation will occur should their output be summed. The same formation is often used as a two-mike stereo pickup technique and has the added benefit of good mono compatibility.

Obviously, there are many times when the outputs of two or more open microphones must be mixed. How, then, to avoid phase cancellation? Phase cancellation problems occur when identical signals, at the same or nearly the same amplitude, are allowed to combine at

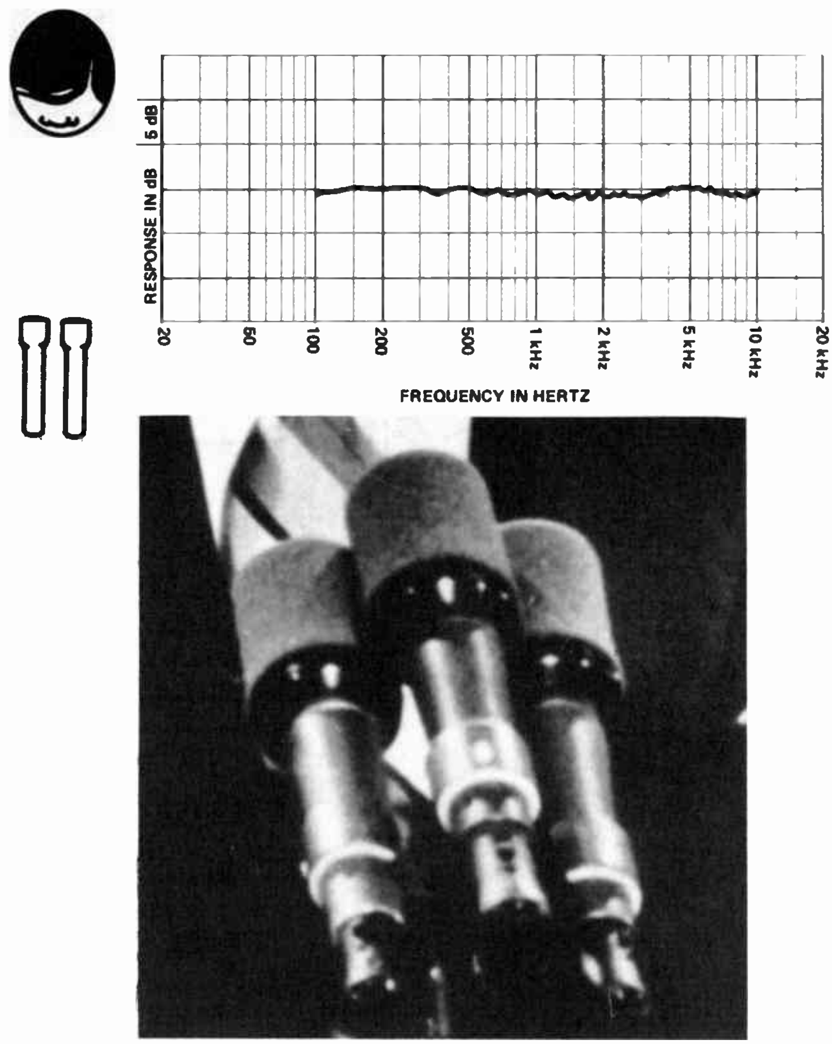


Figure 27. Redundancy miking. (Courtesy of Shure Brothers Incorporated.)

something other than zero phase angle. The result of reducing phase angle error by placing the two pickup elements close together has been shown. It is also possible to influence the amplitude difference between the two combined out-of-phase signals through careful microphone placement. A good rule of thumb to follow in microphone placement is called the *3:1 Ratio Rule*.

To eliminate the problems demonstrated in Figs. 26A–26E by employing the 3:1 Ratio Rule, D3 must always be at least three times D1. Fig. 28 shows examples of both the violation and enlistment of the 3:1 Ratio Rule. Subjective tests have shown that an amplitude difference of at least 9 dB between the two signals will reduce phase cancellation to an inaudible level. The 3:1 Ratio Rule is a means by which this 9 dB minimum difference may be quickly approximated in most multiple-microphone setups.

The amplitude variance desired may also be achieved by judicious use of the mixer's gain or fader controls. Only microphones in actual use should be opened to their normal operating levels; others should be lowered in level or preferably off. Attentive monitoring on an accurate control room speaker system or headphones will reveal audible phase problems, especially with relatively simple sound sources such as speaking voices or most solo instruments.

Acoustic Phase Cancellation with a Single Microphone

Acoustic phase cancellation can also occur in a single microphone system when reflected energy from a nearby barrier such as a music stand, podium, table or floor is introduced at the microphone's diaphragm at a sound pressure level within 9 dB of the direct

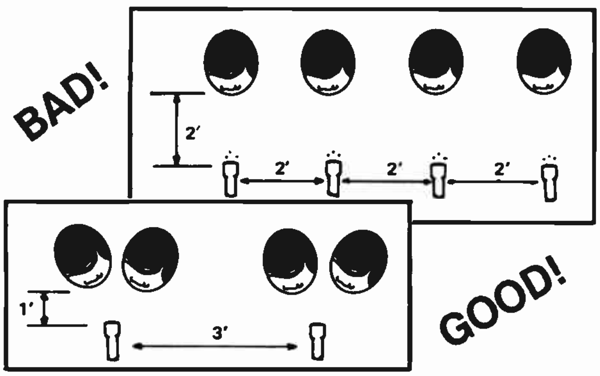


Figure 28. Obeying and violating the 3:1 ratio rule.

sound. Such problems may be avoided by four different means:

1. Increase the reflected path length.
2. Shorten the direct sound path length.
3. Reduce the reflectivity of the barrier. It may be possible to cover the barrier with an acoustically absorptive material or construct it of an acoustically transparent material. For example, use an acoustically-transparent, visually-opaque screen in chroma keying to eliminate reflections into a weather person's lavalier.
4. Position the microphone so close to the barrier surface that the direct and reflected sounds arrive at virtually the same time, causing them to be in-phase.

The latter method also results in higher microphone output because of the additive effect of the two in-phase signals. As discussed in the section on acoustic gain devices, the microphone is operating in (or nearly in) a half-space environment. This barrier (boundary layer) miking technique may be employed with omnidirectional or directional elements (Fig. 29). The omni may be recessed flush into the surface facing out for best performance. Some of the earliest use of the near-barrier technique was with the Electro-Voice Mike Mouse, an acoustically transparent foam holder which positions the mike (usually a cardioid) inside, very close to the barrier, with the diaphragm perpendicular to the surface.

The results of barrier miking will vary because of the following influences:

1. The size of the barrier. (The barrier must be large to support low frequency response.)
2. The size of the diaphragm. (Very small diaphragms may be positioned extremely near the barrier, resulting in less high-frequency cancellation.)
3. The reflectivity and resonant characteristics of the barrier.
4. Ambient noise or reverberant energy problems.
5. Reflections from other nearby reflective surfaces.

MICROPHONE POLARITY REVERSAL

Phase cancellation will also occur if the outputs of two microphones, positioned in the same sound field, are combined with their polarities reversed (i.e., pins 2 and 3 are reverse wired at one end of one cable). The frequencies canceled, and the degree to which cancellation will occur, will depend upon how far apart the microphones are spaced, how closely matched their frequency responses are, and the relative levels of the two mixed signals. It should be noted that the terms *phase* and *polarity* do not mean the same thing. *Phase* refers to a difference in the realtive timing of two signals. *Polarity* refers to the wiring of a microphone or connectors in its circuit and, when reversed, results in a simple shift of 180° in the phase of the signal. Having noted this distinction, in common usage, the terms *in-phase* and *out-of-phase* are often used to refer to matters of polarity. Most microphones will be wired

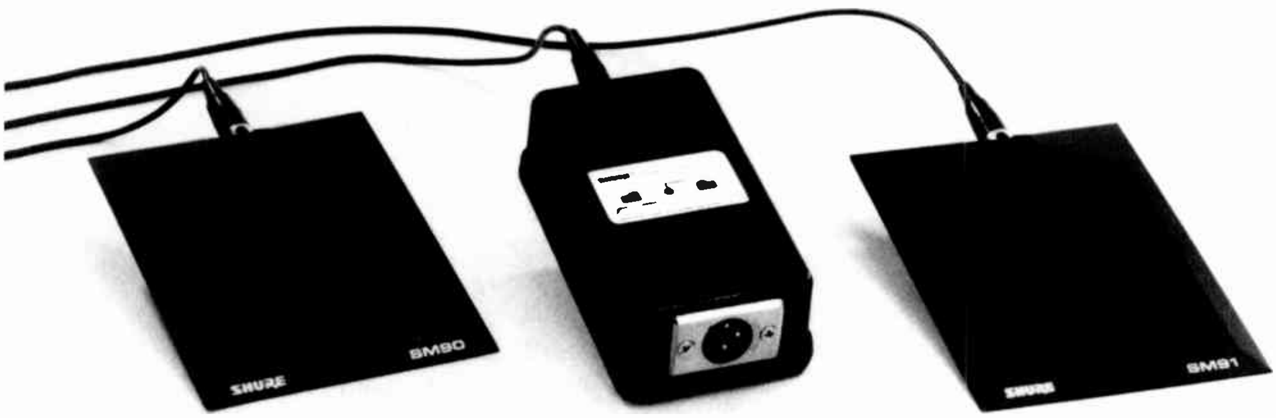


Figure 29. Cardioid and omni microphones designed for surface mounting. (Courtesy of Shure Brothers Incorporated.)

to what is sometimes called the "RCA pin count" which is:

PIN #1 = SHIELD
PIN #2 = HIGH
PIN #3 = LOW

This conforms to IEC standards 268–12 and 268–4. Refer also to EIA-221 which states that the in-phase terminal should be the red (or other than black) conductor and that the out-of-phase terminal shall be the black conductor. The terms "in-phase terminal" or "high" (pin #2) indicate the terminal that has a positive voltage present when a positive pressure is applied to the microphone diaphragm.

Checking Polarity

While there are commercially available devices which use a pulse generator to check for polarity reversal, microphones and their cables may also be checked by simply bringing them together and summing their output while speaking into them from a foot or so away. Two mikes which are "in-phase" will deliver a higher output under such a test; if they are reversed in polarity, the output should drop noticeably.

Polarity Reversal as a Tool

While inadvertent polarity reversal in a mike line can result in some very bad audio, deliberate polarity reversal is sometimes employed as a problem solver.

Reducing Background Noise

A pair of mikes may be reversed in polarity to reduce the pick up of ambient noise. This technique is sometimes employed with two mikes in fixed locations, such as in a press box at a sporting event. If these mikes are brought together, a noise-canceling or *differential* microphone is created. The talker must now speak into one of the mikes only, virtually in a lip-touching position. Due to inverse-square law, the amplitude of the voice at that microphone will be much greater than at the other, resulting in reasonable output level. Distant sound will be picked up equally well, however, by both elements, and canceled.

STEREO MICROPHONE TECHNIQUES FOR BROADCASTING

As a general rule, in producing stereo audio for broadcast, dialogue should be placed in the center channel, or perceived center, with only music and effects in stereo. Exceptions would be where spatial effects are important to the broadcast material.

Spaced-Pair Microphones

It is of utmost importance to provide high quality stereo audio without compromising mono audio quality. The importantance of maintaining compatibility with monaural receivers (or stereo receivers operating in the monaural mode) normally excludes the use of spaced-pair microphone techniques involving omnidirectional or cardioid microphones. Spaced microphones depend upon a combination of amplitude and timing (phase) differences to provide stereo separation. They do not sum well for mono as the very phase differences which aid in separation result in multiple comb filter effects in the mono mix.

Coincident Microphone Techniques

Coincident microphone techniques utilize two microphones whose diaphragms are placed as near to the same point in space as possible. They offer the potential for good stereo without adversely effecting the mono signal with the phase anomalies introduced by spaced microphones. Coincident microphones depend only upon amplitude differences for stereo separation and imaging and provide excellent mono compatibility. There are several coincident microphone schemes including X − Y, M − S, and Blumlein.

X − Y Microphones

The simplest of the coincident techniques is called "X − Y," which crosses two directional microphones so that their patterns meet at their 3 dB-down points (Fig. 30). The two microphones should be positioned so that one capsule is directly above the other, with the capsules on the same vertical axis. This minimizes any reflection or shadowing of high frequencies that each might contribute to sound arriving on the horizontal plane (Fig. 31).

An *ideal* cardioid microphone would have an acceptance angle of 131° and so would be 3 dB down at 65.5° off-axis. If a pair of coincident cardioid capsules are rotated apart so as to overlap their polars at 65.5° off each mike's axis, the resulting angle between the microphones' axes is 131°.

If the angle is too great, sound sources at the center of the stereo image are placed farther off axis of each microphone and are thereby attenuated, making them sound as if they are farther away. Similarly, too narrow

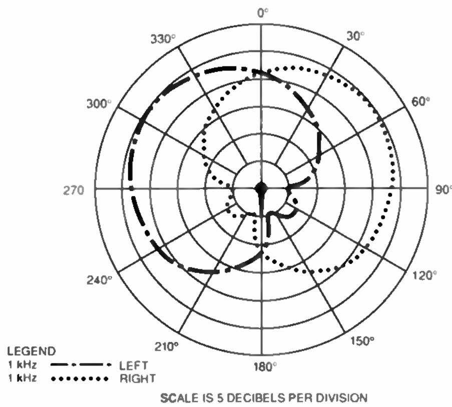


Figure 30. X-Y pattern orientation. (Courtesy of Audio-Technica U.S., Inc.)

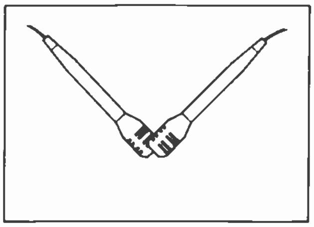


Figure 31. X-Y capsule orientation. (Courtesy of Audio-Technica U.S., Inc.)

an angle results in near-center sources sounding louder, or appearing to be closer. Crossing the patterns to overlap at their 3 dB-down points insures that sound arriving from the center of the stereo stage will be summed such as to provide uniform sensitivity from left through center to right.

While a 131° angle would be correct for *ideal* cardioid capsules, the optimum angle for *real* microphones will likely be somewhat less. It is very beneficial for the microphones used in X – Y to have uniform polar patterns. Because the polars of most cardioid microphones tend to collapse at higher frequencies, it is frequently recommended that X – Y positioning be 90°. This narrow spacing, however, often results in too much overlap of the polars. Stereo separation suffers and center-channel information tends to be brought forward of where it should lie in the stereo image.

The optimum angle for many cardioid elements will be approximately 120°. Experimentation and a thorough knowledge of the polars of the microphones you choose to use will help you get the best X – Y results.

Even highly directional shotgun microphones may be used successfully in X – Y, particularly if you are careful to select some of the newer models that have greatly improved polar uniformity. Remember to cross the microphones at the elements, not the ends of the microphones.

One can make a quick check, outdoors, of the angle adjustment of an X – Y pair. Sum the outputs of the two microphones into a mono audio monitor and provide equal gain for each. Feed pink noise into a small powered speaker about five to six feet in front of the shotgun microphones. Rotate the X – Y pair horizontally at their capsules so that you can monitor the pink noise from the far left channel to far right. Pay particular attention to the amplitude at center. There should be a smooth transition from left to right channel. If there appears to be a hole in the middle, the angle is too great. If the noise seems suddenly closer at the center, try increasing the angle.

Several X – Y stereo microphones are available that integrate two directional elements into one housing (Fig. 32). These greatly simplify microphone placement. While easy to use, they should be handled with some measure of intelligence. It may seem obvious that in most cases the axes of the left and right microphones should be near horizontal. Some X – Y microphones, however, hide their capsules in round housings or windscreens that do not permit a quick visual indication of just what is horizontal.

It is not a good practice to use an X – Y microphone for a close-up announcer or reporter application. Even slight side-to-side head movements can cause the voice to shift dramatically from one channel to the other. A good example of a bad application would be to use an X – Y microphone vertically, hand-held, for a stand-up.

Mid-Side Coincident Microphones

The most versatile of the coincident microphone types for stereo broadcasting is the M – S or "Mid – Side" microphone (Fig. 33). The M – S microphone is a combination of a *Mid* microphone, typically a cardioid or hypercardioid, and a bidirectional, *Side*, microphone. The capsules of the two are placed as close together as possible. A matrixing network combines their outputs and decodes them as left and right channel information (Fig. 34). The information derived from the matrix is nearly identical to that delivered by an X – Y pair, but with some important control advantages.

Sound originating from directly on-axis of the M – S mike will be picked up by the Mid element and delivered equally to left and right channels through the matrix. The Mid mike, which faces the left, with its

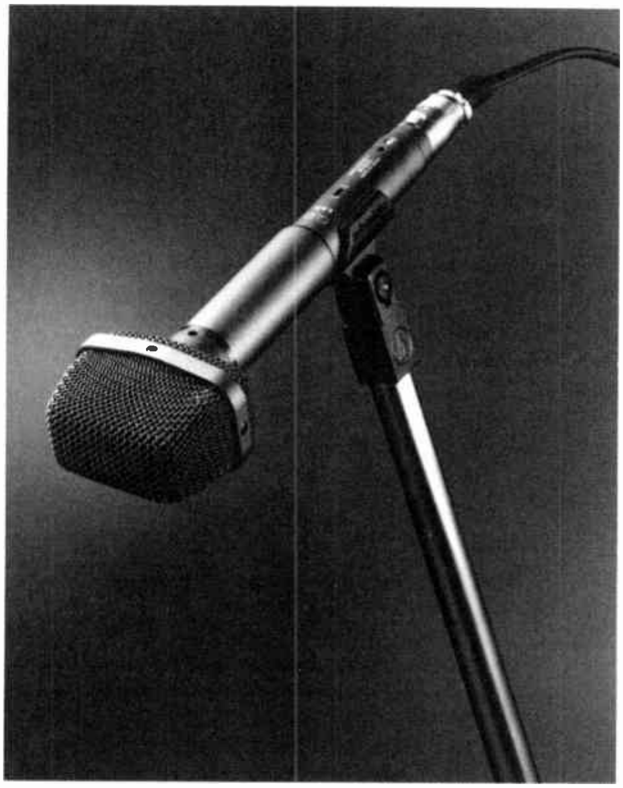


Figure 32. An X-Y microphone. (Courtesy of Audio-Technica U.S., Inc.)

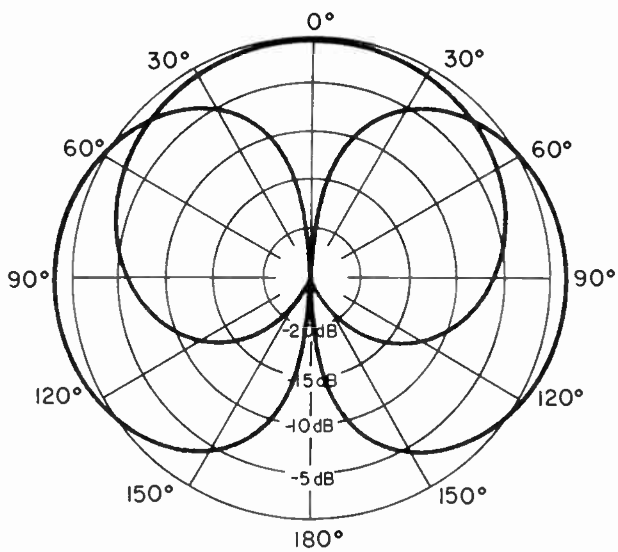


Figure 33. M-S pattern orientation. (Courtesy of Shure Brothers Incorporated.)

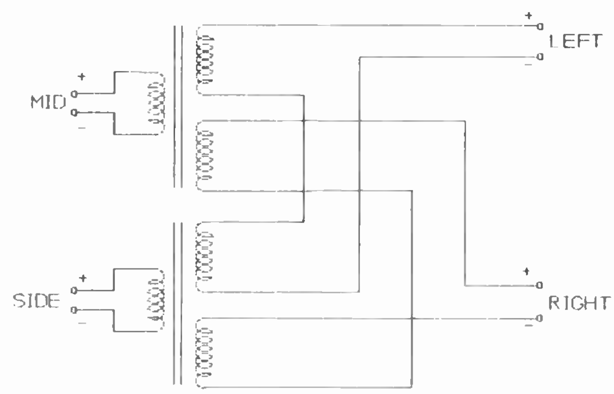


Figure 34. A passive M-S matrix.

rear lobe facing right, is insensitive to sound arriving from the center of the stereo stage as the sound is arriving at 90° off-axis, where the null is deepest. It is, of course, sensitive to sound arriving from each side.

This is part of the process by which the M – S microphone derives directional cues. Sounds arriving from the left are picked up by the Mid and Side elements and, because they are in-phase, are summed and sent to the left channel. Because the rear of the Side element is out-of-phase with the Mid mike, their sum cannot be used to produce right channel information. Instead, an inverted-polarity version of the Side microphone output is mixed with the Mid mike and delivered to the right channel. This processing happens in the sum-and-difference matrix according to the equations:

$$\text{Left} = \text{Mid} + \text{Side}$$

$$\text{Right} = \text{Mid} - \text{Side}$$

Commercially available M – S systems offer well-matched capsules, easy operation, and considerable control flexibility. The model in Fig. 35 incorporates the matrix system and offers a choice of outputs: Mid and Side, or stereo. It also features a three-postion control over the degree of stereo "spread" and stereo ambience pickup.

An M – S pair may be constructed using Mid and Side mikes from the station's collection plus a matrixing system. Several M – S matrices or "decoders" have been introduced for such use (Fig. 36). A mixer may also be used for deriving L/R information from the M – S pair as shown in Fig. 37.

The M – S technique offers several control capabilities. The first two of these controls may be exercised in either production or post production.

1. Adjusting the relative levels of the M and the S signals will narrow or broaden the perceived stereo image. This may be done in the field using the M – S mike's matrix system or you may opt to record the outputs of the M and S capsules on separate tracks, saving the matrixing of them for post production. Matrixing in post will allow the audio perspective to be adjusted to make sense with the video.
2. Panning the M signal off-center may be done to deliberately shift the stereo image. For example, crowd noise at a sporting event may be shifted to appear more closely balanced left and right of a mike position without moving either the mike or several thousand fans.
3. Substituting various patterns, from omni to hypercardioid, for the mid microphone, will affect the apparent mike-to-sound source distance as well as the signal-to-ambient or reverberant-noise ratio.

Blumlein Miking

The Blumlein technique employs crossed bidirectional elements and, like the M – S, responds to amplitude differences to achieve stereo separation. The stereo sound achieved by this approach can be very natural and mono integrity is well maintained. The Blumlein is more sensitive, however, to ambient noise and reverberation than the M – S and placement is critical.

CARE AND FEEDING

Microphones require a certain amount of care in their handling and storage. Here are some basic factors to consider and tips on microphone care. Misuse, or even some attempts to service or clean the microphone, could affect some manufacturers' warranties. When in doubt, ask or return mikes to the manufacturer's recommended service organization for maintenance.

1. Use a windscreen or pop filter to protect the microphone if it is to be subjected to air-borne contaminates such as dust or smoke.

Figure 35. M-S microphone with internal matrix. (Courtesy of Shure Brothers Incorporated.)

2. A foam windscreen will also protect a mike from exposure to rain or snow for a surprising period of time. Over time, the cells will fill with water resulting in high-frequency loss and level drop. The foam may be quickly squeezed to reduce the moisture content or a dry screen substituted as required.
3. Foam windscreens will accumulate deposits of dust and other contaminates. The result will be a deterioration of frequency response and, perhaps, even altered polar response. Foam may be cleaned with soap and water. Rinse well to remove all residue. Nondetergent soaps work well.
4. Many mikes may be carefully opened to remove a foam pop filter and sometimes a cloth insert. Do so only in a very clean environment. These filters should be cleaned as above.
5. Avoid allowing dynamic microphones to be set on work benches or other areas where metal particles or metallic dust may be attracted to their internal magnetic structures. Very small particles can work their way onto the diaphragm and alter the response greatly. In some cases, the dynamic mike can be opened to reveal the diaphragm for examination. Metallic particles may be VERY carefully removed onto the magnetized tip of a screwdriver. The screwdriver shaft should be steadied on the edge of the mike case and the tip very carefully lowered to attract particles which would likely be held immediately above the voice coil gap.

Figure 36. An M-S matrix decoder. (Courtesy of Audio Services Corporation.)

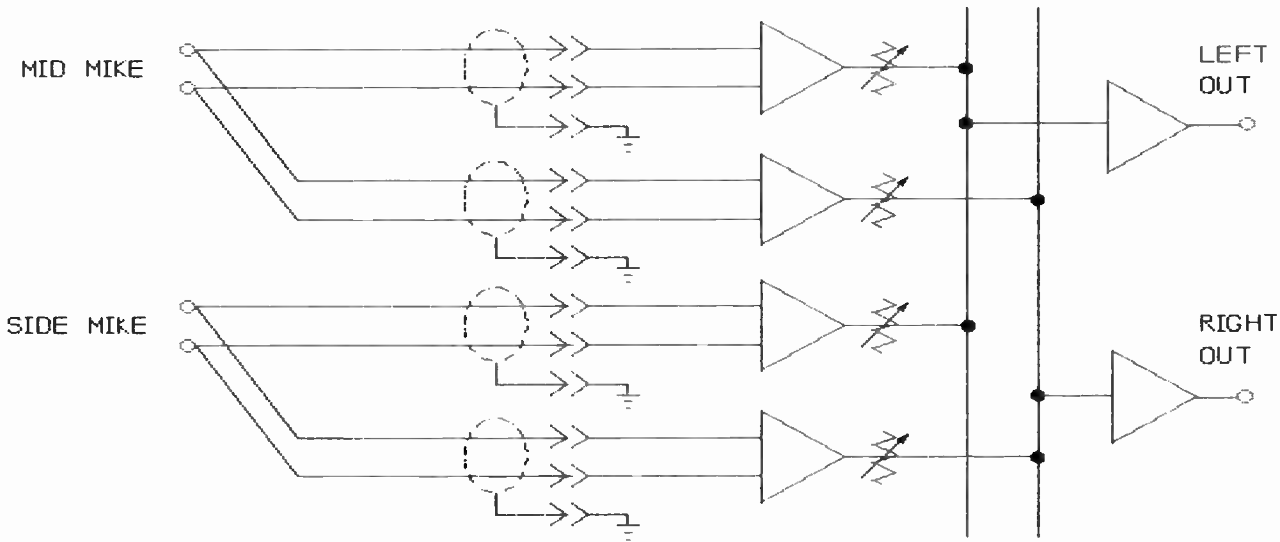


Figure 37. Mixer used as an M-S matrix.

6. Avoid subjecting electret condenser microphones to high temperatures. This means do not store in the trunk or glove compartment of a car left in the sun on a hot day. Also avoid leaving the electret mike on a boom very close to hot lights. The result may be a loss of charge on the capacitor element and a drop in level.
7. Avoid moisture with all mikes but especially with condenser microphones.
8. If given a choice between using mercury or alkaline batteries to power a microphone, remember that mercury cells die much more suddenly than alkalines. The gradual drop in level with an alkaline can be a life saver. Mercury batteries also drop in output level in cold weather. Furthermore, mercury batteries may give off a gas that can corrode the contacts.
9. Avoid unnecessary mechanical shocks. Store in clean, padded enclosures.
10. Moving a condenser microphone from a cold environment to a warm one may cause noise problems from condensation.
11. Avoid moisture in cables and connectors, particularly where phantom power is being used.

BIBLIOGRAPHY

Howard M. Tremaine, *Audio Cyclopedia*, Howard W. Sams & Co., 1974.

John Eargle, *Sound Recording*, Van Nostrand Reinhold Co., 1976.

Lou Burroughs, *Microphones: Design and Application*, Sagamore Publishing Company, Inc., 1974.

James Long, "Layman's Guide to Microphone Specifications," *Audio Magazine*, August, 1969.

James Long, *The Microphone Handbook*, John Eargle, Elar Publishing Company, Inc.

John M. Woram, *The Recording Studio Handbook*, Elar Publishing Company, Inc., 1982.

Jon R. Sank, "Microphones," *Journal of the Audio Engineering Society*, July/August 1985, Vol. 33, No. 7/8. Harry Olson, *Modern Sound Reproduction*, Van Nostrand Reinhold Co.

Don and Carolyn Davis, *Sound System Engineering*, Howard W. Sams & Co.

"Microphones: An Anthology of Articles on Microphones," *Journal of the Audio Engineering Society*.

5.2 The Broadcast Television Camera System

Peter Gloeggler
Sony Corporation, Montvale, New Jersey

OVERVIEW

This chapter presents the fundamental concepts embodied in a high performance color television camera intended for broadcast and program production use. The criteria for selecting the most appropriate camera type, imager type and lens for a given application are provided. The chapter also includes a practical guide for quick performance verification before camera use.

INTRODUCTION

Television, with its array of complex electronic equipment, may at first glance appear to be radically different from any technology that has preceded it. On closer examination this new medium reveals itself to be quite similar to other imaging systems with which we are much more familiar. Color television, color printing, and color photography share the *tri-stimulus* concept of making a colored image. In each of these systems, a combination of three basic colors is used to create the wide spectrum of colors each of these mediums is capable of producing. Fig. 1 shows a plot of the three NTSC color primaries plotted on the CIE chromaticity diagram. The full range of colors contained within the red, green, and blue (RGB) triangle on this diagram can be reproduced by the National Television System Committee (NTSC) color television system.

In its most basic form, all the color television systems currently in use separate the incoming light into three separate images, corresponding to the red, green, and blue spectra of the original scene.

Pick-up tubes or solid-state imagers convert the three optical images into three electrical signals that correspond to these images. These signals are then transmitted to another location for display. The signals are frequently combined into a single composite signal, using NTSC, PAL Phase Alternation Line, or SECAM (SEquential Couleur Avec Memoire, for sequential color with memory) encoding before transmission. At the receiver (or monitor) these composite signals are separated again into the red, green and blue components and applied to a color display device, most commonly a cathode-ray tube (CRT).

In the case of the CRT, an array of red, green, and blue phosphors on the screen are excited by electron beams corresponding to the red, green, and blue signals, and create the final color image.

Fig. 1 shows the R, G, B color primaries of the NTSC color system plotted on the CIE chromaticity

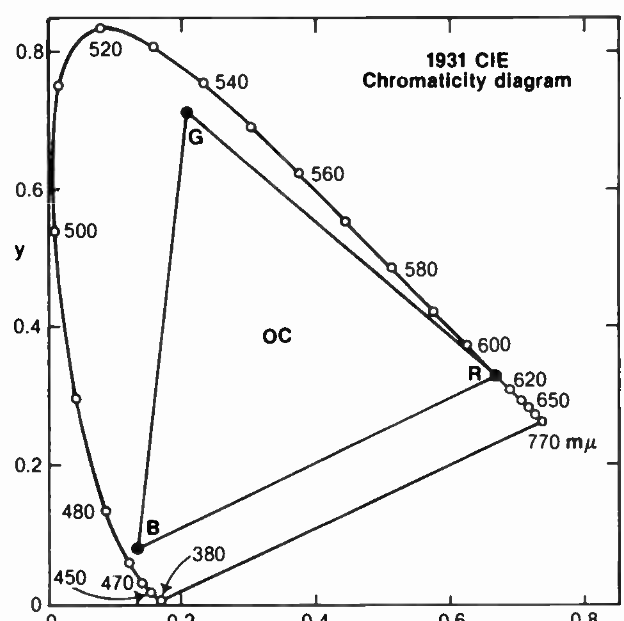


Figure 1. CIE chromaticity diagram showing the NTSC color primaries. (Courtesy of Wiley Publishing.)

diagram. The wide range of colors that can be reproduced by the NTSC television system is defined by the triangle formed by the R, G, B color primaries.

THE BASIC CAMERA SYSTEM

Television became a practical reality with the invention of two electronic devices that use an electron beam in a vacuum envelope to overcome the limitations of the electro-mechanical devices used in earlier experiments with television. Both devices use an electron beam, albeit to opposite purposes. The first device, the *cathode ray tube* (CRT), uses an electron beam to scan a phosphor layer which emits light in direct proportion to the intensity of the electron beam striking it. Since the intensity of the electron beam can in turn be controlled by an electrical signal applied to the CRT, this electrical signal can be used to create an optical image.

The second device, the *image pick-up tube* also uses an electron beam, but in this case the electron beam is used to scan an image focused on the tube and produce an electrical signal that is proportional to the optical image. The image pick-up tube and its solid state successor the charge-coupled device (CCD) imager constitute the heart of a television camera.

The TV camera is only the starting point in the long process of bringing a television image to the viewer, but the TV camera is one of the most exciting and challenging parts of the TV system. Creating a modern TV camera requires the latest technology from a wide variety of disciplines: optics, integrated circuit technology, microprocessor control and manufacturing technology.

TV cameras, from the largest cameras used in the studio, to small cameras used for electronic news gathering (ENG), share many common elements.

Video cameras, typically contain these major parts:

1. A lens to capture the scene.
2. Color separation optics to separate the incoming light into three "primary" color images: red, green, and blue.
3. Three pick-up devices that convert the red, green, and blue images into three electrical signals: the RGB video signals.
4. Video processing circuitry that performs a variety of processing steps to:
   A. Correct for errors in the lens and optical system.
   B. Correct for the color temperature of the scene illumination.

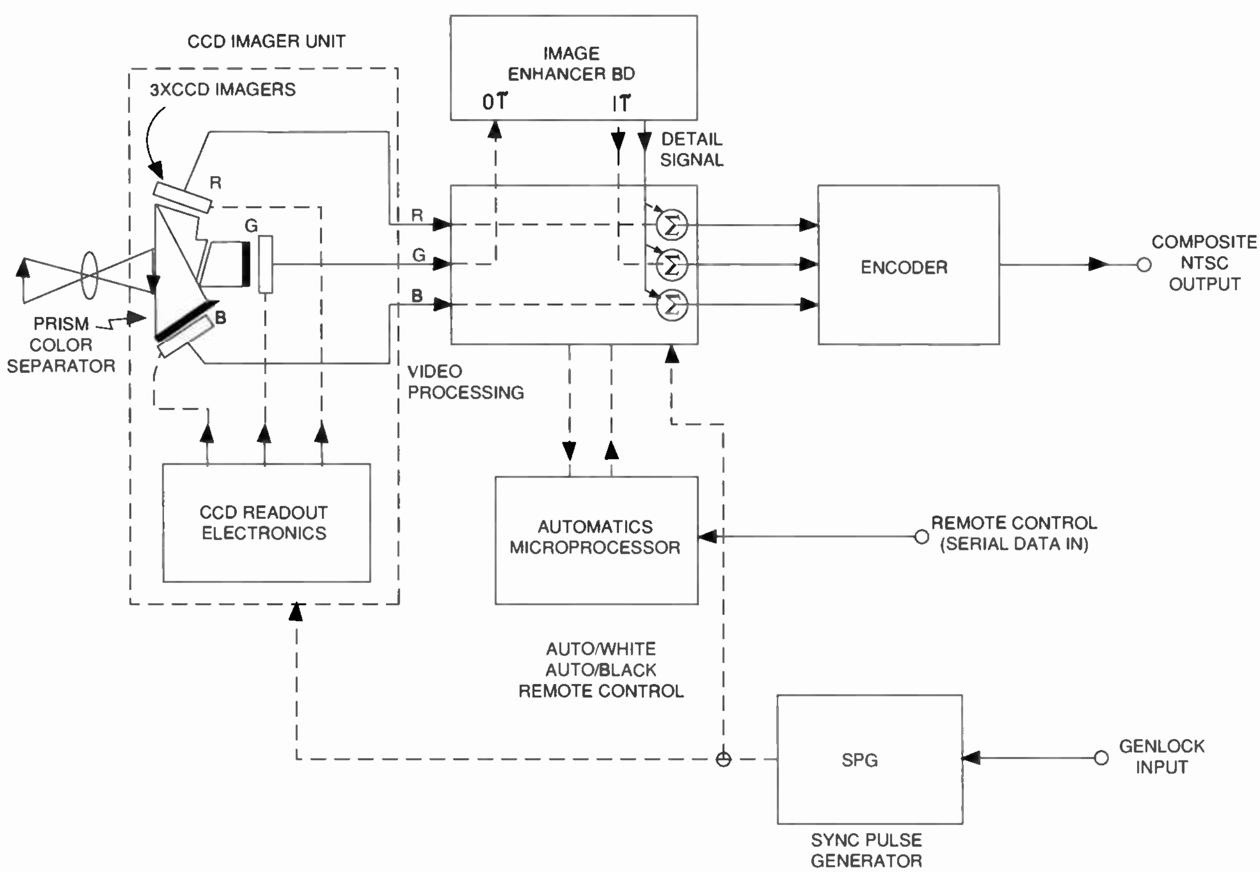


Figure 2. Basic block diagram of a TV camera.

C. Compress the dynamic range.
D. Increase the apparent sharpness of the image by the use of image enhancement.
E. Introduce "gamma" precorrection.

5. Encoding system: circuitry that encodes the basic RGB camera signals into a composite signal conforming to the NTSC color encoding scheme.
6. Timing signal generator: circuitry that generates the waveforms necessary for the internal operation of the camera.
7. Power supply: circuitry that generates the various voltages required to operate the various parts of the camera.

DETAILS OF THE CAMERA SYSTEM

To illustrate the complex processes that occur in a TV camera, we will follow the rays of light entering the lens, conversion of the light to an electrical signal, and the processing of the signal to the final output signal delivered by the camera.

THE OPTICAL SYSTEM

The optical system of the TV camera is used to form a precise image of the original scene on the surface of the imaging devices.

The optical system of the camera consists of:

1. A lens to capture the image.
2. Optical filters to condition the image.
3. The color separation system to separate the incoming light into the three primary color components.

Clearly the quality of the optical system is critical to the performance of the overall camera. In the world of optics, there is usually a close correlation between cost and performance, so it is not surprising to find a somewhat lower level of performance in the optical system of lower cost cameras and lenses.

Brief descriptions of the optical components used in a camera are given here. (For further information, see Refs. 1, 3.)

The Lens

With the exception of the early RCA studio cameras that used a turret of fixed lenses, the zoom lens is the universal type of lens used with television cameras. Zoom lenses are available at a wide range of prices and performance levels, up to and including performance levels required for use with high definition television (HDTV) systems. Since the zoom lens provides such a high level of performance, only a few TV lenses with fixed focal length, also called prime lenses, have been developed for use in special applications.

At first look, the requirements for a lens intended for use with a TV camera appear to be the same as lenses intended for use with a film camera. Indeed, film type lenses are routinely used with black and white cameras and single imager color TV cameras. A markedly different lens is required for use with a professional color TV camera. The back focal distance (i.e., the distance from the end of the lens to the plane where the image is formed) is increased significantly in these lenses, to allow the insertion of the prism color separation system between the lens and the imagers. For this reason, film style lenses cannot be used with prism color television cameras without sophisticated relay optics. Such relay optic adapters are available and typically used with extreme wide angle and extreme telephoto film lenses for applications that cannot be achieved with available TV zoom lenses.

Another factor which has influenced the design of lenses for prism cameras is the changeover from tube imagers to solid-state imagers. Solid-state imagers impose much stricter requirements on the design of the lens than tube-type cameras. For example, significantly more chromatic aberration can be accepted in a lens used with tubes than with solid-state imagers. With tubes, lateral chromatic aberration in the lens (a change in the relative size of the red, green, and blue images) was easily corrected by a shift in the scanning signal applied to the three tubes. Similarly, longitudinal chromatic aberration (the red, green, and blue images occur at different distance from the lens) was corrected via the camera back focus adjustment; a part of the standard set-up of a tube camera. Since older lenses did not have to meet the strict requirements of lenses to be used with CCDs, it is usually not possible to use lenses designed for pick-up tubes with a CCD camera and obtain optimum performance. Lenses that have been designed to meet the strict requirements of CCD imagers are generally designated as CCD-capable lenses. Lenses being sold for TV camera use should meet the strict requirements for CCD imagers, however, one should not assume this is the case.

(For more details on lens parameters see Appendix B.)

Optical Filters

Several optical filters are used to achieve a high level of performance found in a modern camera:

- Color correction and neutral density filters
- Infrared filter
- Quarter-wavelength filter
- Anti-aliasing filter

Color Correction and Neutral Density Filters

The human eye adapts to changes in the characteristics of the surrounding illumination. Thus, the color of an object observed by the eye will be perceived as the same in daylight or with incandescent illumination. Cameras do not adapt to illumination and the color of an object will differ for each color of the illuminant if appropriate adjustments are not made in the camera. Color cameras are designed to operate with lighting of a specific *color temperature* (see Appendix C for a more detailed discussion of the concept of color temperature), most commonly 3,200° Kelvin; the lighting provided by common halogen studio lights. However, the camera must also have the ability to operate properly in other environments. A combination of optical color correction filters and electrical adjustments of the relative gain of the camera's red, green, and blue processing channels, are used to reproduce the object correctly under different illumination. Outdoor scenes challenge the camera with a wide range of illumination, rich in reds early in the day and late in the afternoon, and dominated by blue in the middle part of the day. In terms of color temperature, the TV camera must be able to achieve *white balance* for illumination with a color temperature of less than 3,000° Kelvin for warm incandescent lighting, to color temperatures as high as 10,000° Kelvin on a bright, wintry day outdoors. Outdoor illumination is frequently much brighter than indoor lighting, and for this reason a neutral density filter is commonly used in conjunction with a daylight correction filter, to keep the lens iris around the middle of its operating range. Color cameras commonly include a four position filter wheel that allows convenient selection of the most appropriate combination of neutral density and color correction filters, for a given scene illumination. Although many cameras have the ability electronically to achieve white balance even when an incorrect filter is used, the signal-to-noise ratio and the ability to handle pictures with very high contrast will be impaired.

Infrared Filter

To avoid incorrect color reproduction from unwanted response of the CCD or pick-up tube imager to infrared light (wavelengths below the visible spectrum), cameras normally include an infrared filter in the optical path to limit the response of the camera to the visible spectrum.

Quarter-Wavelength Filter

The dichroic coatings used in the prism color separation system respond differently to light of different polarizations. To avoid a change in colorimetry with a change in polarization in the scene, a quarter-wavelength filter is introduced before the prism in high performance cameras. The quarter-wavelength filter alters the plane polarization of the incoming light to circular polarization. Lower cost cameras may not include this filter due to its high cost.

Anti-Aliasing Filter

A filter quite new to the design of TV cameras, the anti-aliasing filter has the difficult task of removing the extremely fine picture detail that can produce objectionable aliasing in a solid-state imager, but have relatively little effect on picture detail within the resolving power of the imager. The quality of this filter is the key element that controls the compromise between useful resolving power and curtailment of the aliasing phenomenon in the imager.

Different from the smooth homogeneous imaging layer used in the pick-up tube, the new CCD solid-state imagers use an orthogonal, two-dimensional array of separate discrete sensors. Each sensor element develops a charge that is directly proportional to the light level impinging on this specific sensor element. The CCD array is therefore a collection of discrete electronic samples of the original optical image. Recognizing the CCD imager as a sampled data system, it becomes clear that the Nyquist sampling theorem applies. That is, errors will result when an input is sampled at a rate that is less than twice the frequency of the highest frequency in the input. Here the frequencies are spatial rather than temporal, but the results are the spatial equivalent.

To avoid the aliasing phenomenon in the electrical output of the imager, an optical low-pass filter must be added to the camera to remove fine image detail above the spatial Nyquist limit. The optical low-pass filter is similar in design to a multi-element star filter, and one of the signatures of this optical filter, commonly found in almost all CCD cameras, is star filter-like rays emanating from a strong highlight. Two-axis optical low-pass filters are most frequently used in industrial grade CCD cameras, while three-axis filters are normally provided in broadcast grade cameras, since aliasing tends to be much lower with the more complex three-axis filter.

The Color Separation System

A color camera contains an optical system that separates the incoming light into three component colors. The type and quality of the color separation system used determines not only the faithfulness with which the camera reproduces colors, it also determines the physical arrangement of the imagers and thereby the design of the camera.

Current color cameras use one of the following designs:

- A single-imager with color separation achieved by color filters positioned over the individual pixels
- Two-imagers, one for luminance and the other for color reproduction
- Three-imagers with prism-type optics for color separation

Sensitivity and the ability to accurately reproduce colors, particularly when luminance values approach black or near peak white, are compromised in the two-imager cameras. Resolution, sensitivity, and color reproduction near black and peak white luminance values are compromised in the single-imager camera. For this reason applications of one an two-imager cameras tend to be limited to consumer and industrial applications, while prism cameras are usually the choice for broadcast and program production.

The RGB primary color separation system uses red, green, and blue optical filters in front of the CCD pixels.

The complementary type color separation system uses yellow, cyan, and magenta optical filters in front of the CCD pixels.

The prism-type color separation system has become the dominant system for use in high-performance

TABLE 1
Relative performance of color separation systems.
(Performance in a camera using ⅔ inch, f/5.6, 768 pixel CCD.)

Color Separation System	Number of Imagers Requires	Resolution	Sensitivity at 2000 Lux (Full Output)	Color Fidelity	Cost
Three-Imager Cameras:					
Prism Type	3	Good ˜700TVL	Good ˜f/5.6	Good	High
Two Imager Cameras:					
Luma/Chroma Type	2	Good ˜700TVL	Fair ˜f/4.0	Fair*	Medium
Single Imager Cameras:					
Color Stripe: RGB Primary	1	Poor ˜330TVL	Poor ˜f/2.8	Fair*	Low
Color Stripe: Complementary	1	Fair ˜330TVL	Fair ˜f/4.0	Poor*	Low
Mosaic-Type: Primary Color	1	Fair ˜440TVL	Poor ˜f/2.8	Fair*	Low
Mosaic-Type: Complementary	1	Fair ˜440TVL	Fair ˜f/4.0	Poor*	Low

*The color stripe/mosaic method of color separation tends to be ineffective for very dark colors (luminance < 5 IRE) and very bright colors (luminance > 95 IRE).

cameras by virtue of its compactness and high performance. An illustration of the three types of color separation systems and a more detailed description of the prism-type system is provided in Appendix A, "Color Separation Systems."

IMAGING DEVICES

The imaging device is the transducer that converts the optical image into an electrical analog of the image. Cameras used in television broadcasting are almost entirely color cameras and use one, two, or three devices, either tubes or solid-state devices.

Image Diagonal

Image size is one of the more important factors that determine the overall performance of a camera. High-performance cameras suitable for broadcast and production commonly use ⅔-inch or ½-inch format CCD imagers. The ⅔-inch image format has established itself in portable tube cameras as the optimum compromise between camera size, weight, and performance and this format is quickly displacing 25 mm and 30 mm pick-up tubes, even in cameras designed for the very demanding requirements of studio use. Fortunately, the same ⅔-inch image format is also an excellent compromise between performance and cost for solid-state imagers. Limitations on minimum linewidth and with integrated circuit technology make it difficult to shrink the size of the imager and maintain the same level of performance. For a given type of CCD imager, ⅔-inch format cameras, therefore, tend to offer a higher level of performance than cameras using ½-inch imagers. In addition, a theoretical analysis indicates that the design requirements for the lens and the optical system become more severe as the image diagonal is reduced.

Sensitivity is one more specification of the camera that is strongly related to image diagonal. The prism color separation system has a basic limitation on the rate of convergence of the light rays that pass through it (the rate of convergence is limited to f/1.4 in the commonly used "three-piece" prism system). As a result, it is possible to use a lens with almost twice the light gathering power (larger objective diameter) with a ⅔-inch format camera as compared to a ½-inch format camera. As a negative, the cost of the imagers and the overall size of the camera tends to increase when CCDs with larger image diagonal are used. It is not surprising, then, to find that ⅔-inch imagers are the choice for the highest cost cameras, while ½-inch cameras predominate among the lower cost cameras offered by a manufacturer.

The Camera Pick-Up Tube

The pick-up tube was one of the key inventions that made the television system, as we know it today, possible. In the pick-up tube, the optical image focused on the photosensitive layer is scanned by an electron beam, and an electrical signal directly proportional to the optical image is provided at the output of the pick-up tube.

In a series of steps, the early pick-up tubes evolved from the Image Iconoscope (demonstrated by Dr. Zworykin in 1933 and frequently described as the most important invention in the development of television cameras), to the Image Orthicon, the Vidicon, and finally to the Plumbicon® and Saticon® pick-up tubes. Refined over a period of many years, the Plumbicon® and Saticon® achieved a level of performance that has allowed these two devices to reign as the unchallenged imagers for high performance cameras over a period of approximately 20 years. Of the two devices, the Plumbicon® is considered to have somewhat more desirable characteristics overall. Short- and long-term image retention, in particular, tend to be better and, for this reason, the Plumbicon® has become the most commonly used imager in broadcast cameras. Saticons® are generally less costly than Plumbicons® and are more commonly found in lower cost cameras. There are clear exceptions to the lower performance image of the Saticon®. Recently, high performance Saticons® have been developed; currently, both Saticons® and Plumbicons® are competing for the most demanding application: the imager for high definition TV cameras.

In 1970, Willard Boll and George Smith of Bell Labs announced the discovery of the charge coupled device, or CCD, and with this development, the dream of a solid state imager for use in a television camera took a big step toward reality. After almost 20 years of development, the CCD imager has now been improved and is displacing the pick-up tube in almost all cameras, with the exception of high-definition television (HDTV) cameras. This chapter, therefore, concentrates on the details of the new solid-state imagers and limits the discussion of cameras using pick-up tubes.

Pick-up tubes have some strong advantages and a somewhat longer list of strong disadvantages.

The Advantages of Pick-up Tubes Versus CCD Imagers

1. Resolution of pick-up tubes can be excellent, exceeding 1,000 TV lines in tubes intended for high definition cameras.
2. Dynamic range of pick-up tubes can be excellent when supported by external circuitry that senses the presence of a highlight, and automatically increases the beam current for the duration of the highlight (only). Highlights f/3.5 to f/4 above 100 units of video, or the equivalent of about 1,000 IRE are reproduced with a minimum of artifacts. Unfortunately, the benefit of this large dynamic range is tempered by the appearance of severe artifacts when the highlights exceed the maximum beam current of the tube by even a small amount.
3. The scanning waveforms applied to the three pick-up tubes can be manipulated to correct for unavoidable aberrations in the lens; even real-time correction of lens errors with zoom position is possible.
4. The output of the pick-up tube is a smooth analog representation of the optical image. The aliasing phenomenon and the stepping phenomenon on diagonal lines of solid state imagers is avoided.

The Disadvantages of Pick-up Tubes Versus CCD Imagers

1. Larger size, higher weight, and higher power consumption of pick-up tubes are severe limitations for portable camera use.
2. The pick-up tube has a limited lifetime and must be replaced on a routine basis at considerable expense.
3. The characteristics of pick-up tubes change continuously with use, necessitating routine re-adjustment of the internal camera circuitry to maintain optimum performance.
4. Pick-up tubes can be severely damaged by pointing the camera into the sun or other extreme highlight even momentarily.
5. Highlights in excess of the f/3.5 to f/4 dynamic range created severe artifacts known as "blooming" and "comet tailing," in earlier pick-up tubes. Recent pick-up tube designs are claiming reduction of these artifacts to be a non-issue in practical applications.
6. Cameras using pick-up tubes can sustain permanent damage from high ambient temperatures. Damaging temperatures are easily reached when the camera is used in direct sunlight for long periods or if the camera is stored in a car exposed to the direct sun.

The CCD Solid-State Imager

Ever since the introduction of the first professional CCD camera by RCA at the National Association of Broadcasters convention in 1983, the imagination of the broadcaster has been sparked by the concept of a new kind of camera, one that would eliminate the intensive care and maintenance required to sustain the performance of a pick-up tube camera.

Another three years would pass before the dream would become at least a partial reality. At NAB '86, a new solid-state camera, the Sony BVP-5, was introduced with a docking video recorder (VTR) for electronic news gathering (ENG) use. The Sony BVP-5 went on to become the first CCD camera with widespread use in a broadcast environment.

Gone was the notion of the perfect camera; the considerable benefits of the solid-state imager were tempered by a whole new array of terms to describe the imperfections and foibles of these new kind of devices. Strange terms like frame transfer, interline transfer and frame interline transfer, pixels, aliasing, vertical smear, fixed pattern noise, spatial offset, and others entered the vocabulary and needed to be understood to decipher the competing claims of this new technology.

It is useful to compare the capability of these devices against the well-known characteristics of the Plumbicon® and Saticon® pick-up tubes, as a frame of reference.

The Advantages of CCD Imagers Versus Pick-up Tubes

1. Cameras using solid-state imagers are smaller, lighter, and consume less power than cameras of comparable performance using pick-up tubes.
2. CCD imagers do not change characteristics or wear out with use, unlike pick-up tubes.
3. CCD cameras are not damaged by strong highlights or high ambient temperatures.
4. Short- and long-term image burn-in and "comet tailing," objectionable in pick-up tubes, are more easily avoided.
5. CCD camera resolution and registration accuracy are uniform over the whole screen, compared to significant fall-off with pick-up tubes.
6. No centering or registration adjustments are required with CCD cameras.
7. A whole new concept, the "electronic shutter" to reduce blur in fast motion is possible with the CCD imager.
8. Dynamic resolution, (i.e., resolution with the subject or the camera in motion) is significantly better with the CCD imager than with the pick-up tube.

The Disadvantages of CCD Imagers Versus Pick-up Tubes

1. "Aliasing," caused by the limited number of discrete picture elements in the CCD, can cause objectionable artifacts with the presence of very fine details in the picture.
2. Diagonal lines have a stepping effect due to the limited number of picture elements in the CCD.
3. The dark noise of one picture element differs slightly from the noise level of the next pixel, resulting in an objectionable phenomenon called "fixed pattern noise" when gain boost is used.
4. Registration of the imagers is fixed in a CCD camera. Longitudinal and lateral chromatic aberrations in the lens optical system are much more difficult to correct in CCD cameras.
5. The "vertical smear" phenomenon, a line from top to bottom when an excessive highlight is encountered, is only totally eliminated in the frame transfer CCD imager.
6. Dynamic picture distortion is an artifact that introduces a bowing effect in a vertical structure as the camera is panned quickly.

The Basics of the CCD Imager

For all the wonders promised by this new solid-state era, it must be remembered that the CCD is, first and foremost, an analog device, just like the pick-up tube. It thus possesses a wide range of analog characteristics, some labeled with the same familiar terms used to describe the characteristics of pick-up tubes, others labeled with terminology specific to this new technology.

The CCD is radically different in concept from its photoconductive counterparts. In place of the amorphous layer in the pick-up tube that is scanned by an electron beam, the CCD consists of a large number of discrete picture elements (pixels) arranged in an orthogonal, two-dimensional array. Each pixel accumulates an electronic charge directly proportional to the unique light level stimulating it. The CCD array, then, is a large collection of discrete electronic samples of the optical image.

Unlike the smooth, continuous output of the pick-up tube as it is scanned by the electron beam, the output signal of the CCD imager is a series of discrete samples, generated as the information contained in the array is read pixel by pixel. The resulting output signal of the CCD imager is very much like the analog output signal of a pick-up tube if it were sampled at a corresponding rate.

Similarly, the CCD behaves as other sampled data systems. For instance, the Nyquist sampling theorem mandates that the input to a sampled data system must be bandwidth limited at one half the sampling frequency to avoid aliasing errors in the sampled data. The nature of the aliasing experienced with a solid-state imager and the need for an anti-aliasing filter may be analyzed on this basis.

In the solid-state imager, the sampling takes place in the optical domain, and for that reason, the bandwidth limitation prescribed by the Nyquist theorem must also take place in the optical domain. This requires use of an optical filter that will attenuate the high-frequency energy impinging on the imager that can cause aliasing, without attenuating the useful in-band frequencies. Difficult trade-offs must be considered in the design of this filter. A more effective optical filter reduces aliasing, but also reduces the ability to reproduce fine detail in the picture.

Fig. 3 shows the theoretical output of a CCD imager that has the form of Sin X/X; the response of the optical pre-filter to reduce aliasing, and the typical response of a Plumbicon® pick-up tube for comparison.

The Nyquist theorem now also makes it clear that an increase in the number of pixels in the imager is equivalent to an increase in the sampling frequency of a sampled data system. An increase in the resolution of the overall system is thereby possible.

The concept of sampling also explains the number of horizontal elements in current CCDs, which feature 768 horizontal pixels. A quick calculation shows that this seemingly random number of 768 samples per active horizontal line corresponds to four samples

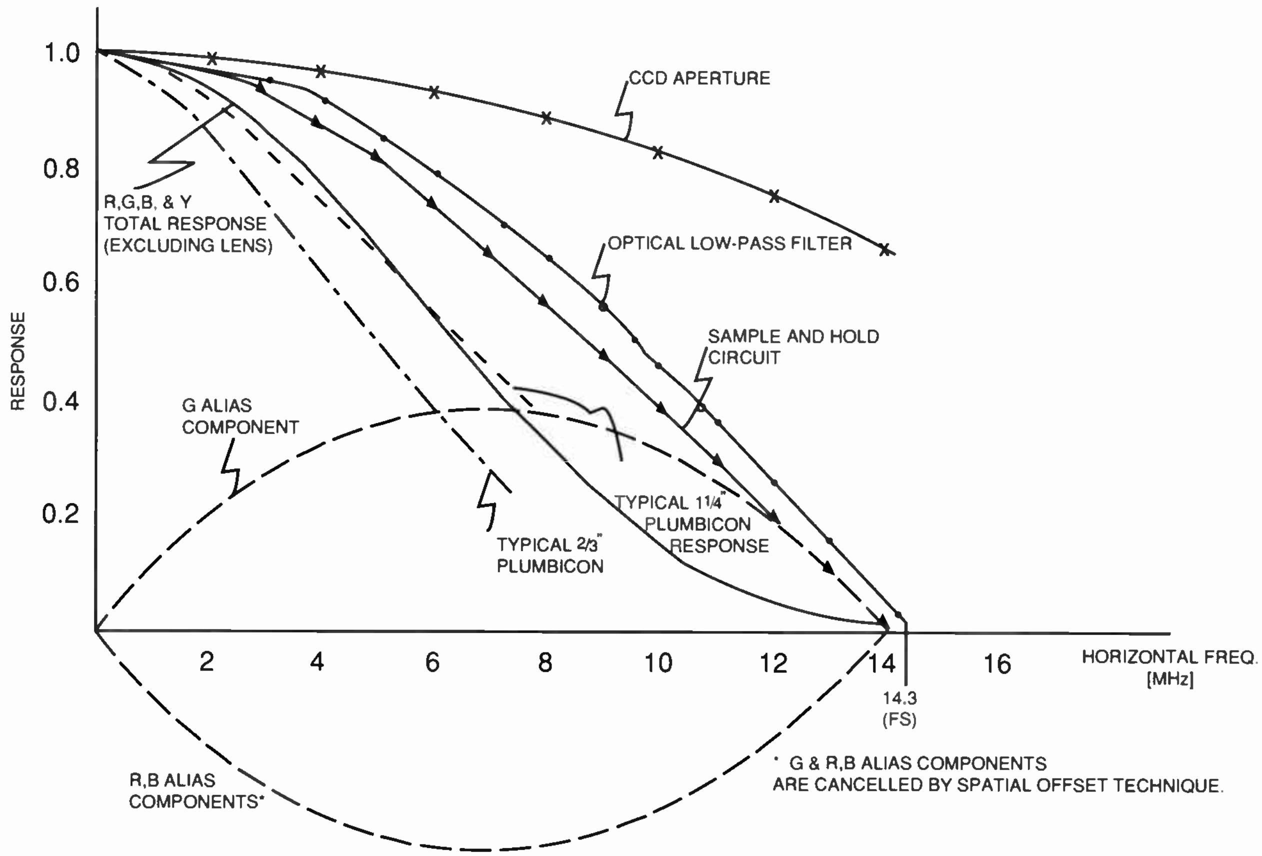


Figure 3. Typical MTF of 3 chip CCD camera compared with typical responses of plumbicon tubes. Alias components in the G and the R and B signals cancel each other due to spatial offset.

per color subcarrier cycle, better known as 4 × f_{sc} sampling:

$$4 \times f_{sc} = 4 \times 3.579545\ MHz$$

and

$$14.318 \times 10^6 \frac{samples}{sec} \times 53.6 \times 10^{-6} \frac{sec}{|active|} = 768 \frac{sample}{line}$$

Increasing the number of pixels in the imager has the potential to increase the resolution and lower the aliasing phenomenon. However, technical difficulties arise when the sampling frequency is not an exact multiple of the subcarrier frequency. For this reason, using a CCD imager with only a small increase above 768 horizontal pixels may not yield an overall improvement.

CCD Transfer Mechanism

The idea of creating a solid-state camera imager was considerably advanced with the development of a new kind of integrated circuit, the CCD. This device has two inherent properties that are the basic ingredients of a successful imager:

1. The individual cells of the CCD have the ability to accumulate charge in direct proportion to the amount of incident light.
2. A charge in one cell of the CCD can be readily shifted to the next cell by the application of a suitable shift pulse.

The frame transfer (FT) CCD. A frame transfer CCD array has two separate sections: (1) an array of imaging cells to capture the image, and (2) an equally large array of cells carefully protected from light to store the image. A means of transferring the charges from the imaging section to the storage section is also needed.

The basic diagram of such a device is shown in Fig. 4. During the active field period, the cells of the imaging sensor array accumulate charges representative of the optical image focused on it. The information in the sensor array is shifted to the lower storage array at a high speed during the next vertical interval. The now empty imager array will be charged with the succeeding image during the next active field period while, at the same time, picture information in the storage section is read to create the camera output signal.

The name given to this basic imager design is the

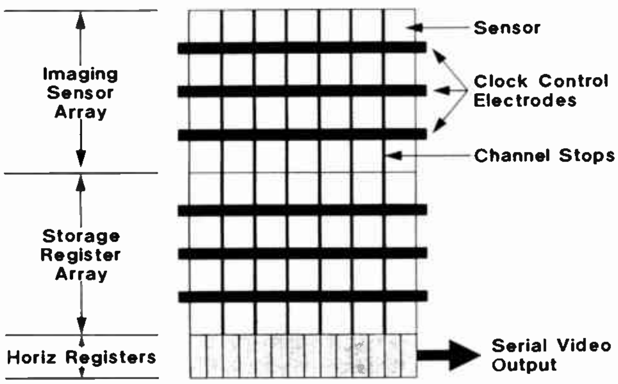


Figure 4. Frame-transfer CCD structure.

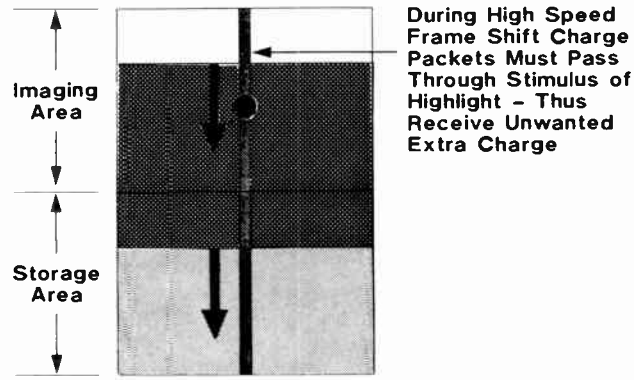


Figure 5. Mechanism of vertical smear in the frame-transfer CCD.

frame-transfer CCD. An important characteristic of the FT CCD is its optical efficiency. Each picture element is large and virtually contiguous with the next element to capture virtually all of the incoming light. The FT CCD also offers a very efficient mechanism to transfer the charges from the imaging section to the storage section. The imaging elements function both as imagers and as effective shift registers to transfer the contents of the imager section into the storage section.

The weakness in the basic FT imager is the contamination that can result as the charges are shifted from the imaging array to the storage array. The sensors are still being stimulated by the continuously impinging optical image during the short time the content of the imager section is transferred to the storage section. Erroneous signals at levels approximately 50 dB below normal signal level may result, levels clearly not acceptable in a professional camera. The name *vertical smear* has been applied to describe this phenomenon.

A mechanical shutter interposed between the lens and the prism provides a solution to this problem. The shutter is synchronized with the vertical blanking interval to totally block the incoming light during the transfer process and thereby avoid the contamination that can otherwise occur. The use of an FT CCD imager with a mechanical shutter form a viable imaging system for a broadcast camera. The FT CCD was introduced to the broadcast industry by RCA in 1983 and continues to be used in current broadcast cameras.

The interline-transfer (IT) CCD. A completely different approach to the problem of moving the charge packets from the imaging array to the storage array is taken in the interline-transfer CCD. Storage elements

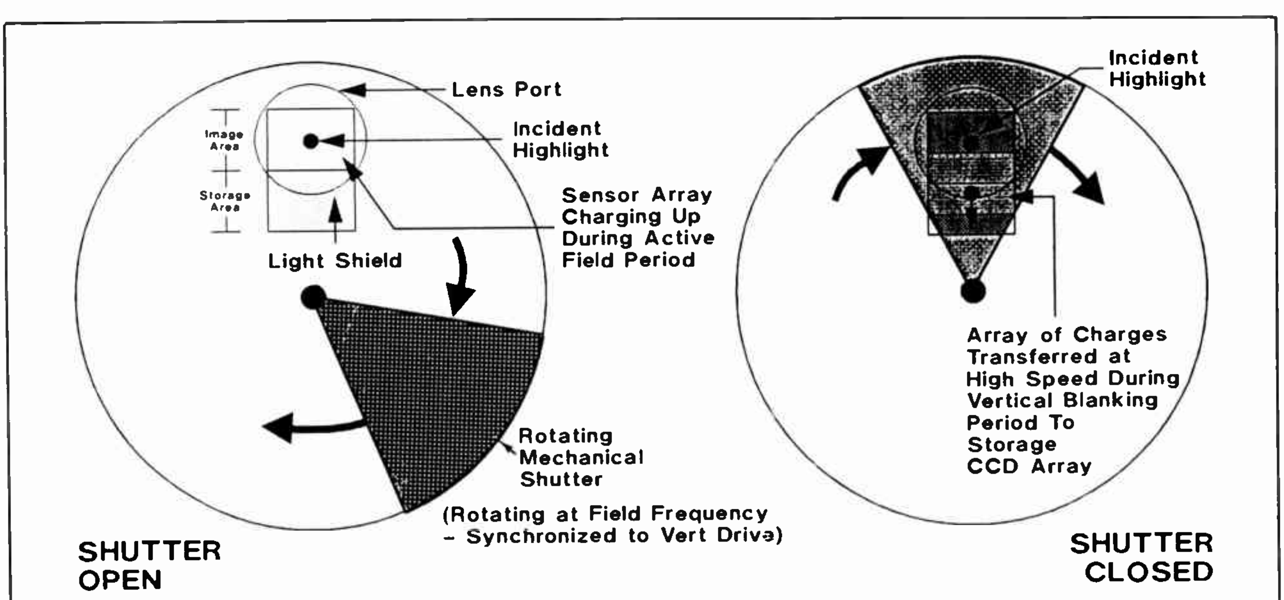
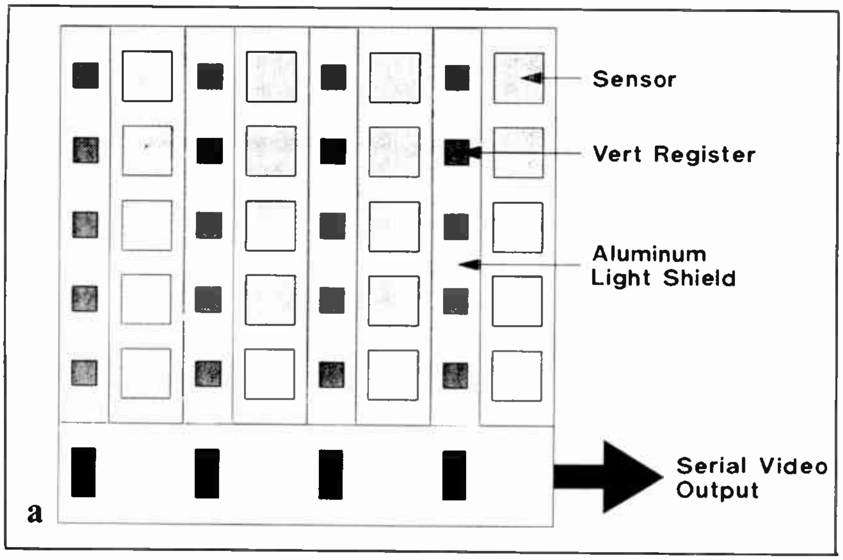


Figure 6. Removal of vertical smear in the frame-transfer CCD by means of a rotating mechanical shutter.



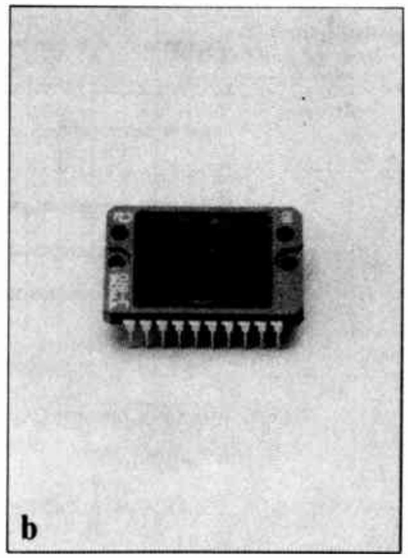


Figure 7. Structure of the interline-transfer CCD.

protected from light are located next to each of the imaging elements. A simplified structure of the IT CCD is shown in Fig. 7.

During the active field period, the sensor array again accumulates charges proportional to the image focused on it. However, during the next vertical interval, these charges are quickly transferred to the adjacent light-protected storage elements called the interline storage register. During the succeeding active field, the now empty sensor array charges with the next image, while the previous "image" is read from the interline storage register to become the video output of the camera. Comparing the IT CCD to the FT CCD imager, we note that the optical sensing area had to be sacrificed in the IT CCD to make room for the interline storage registers. This reduces sensitivity, which can, however, be recovered by other means. The use of the separate interline storage register now makes it possible to transfer the charge packets out of the sensor array without worrying about incident light directly contaminating these charges; but there is still one more problem to contend with: the intense highlight. When an intense highlight, such as the filament of a car headlight, is focused on the sensing element of the imager, a small amount of this light is reflected within the CCD and will be able to add charges to the immediately adjacent interline storage register. This means the vertical smear phenomenon is still present, albeit at a much lower level than with the FT CCD; approximately 85 dB below normal output. While not

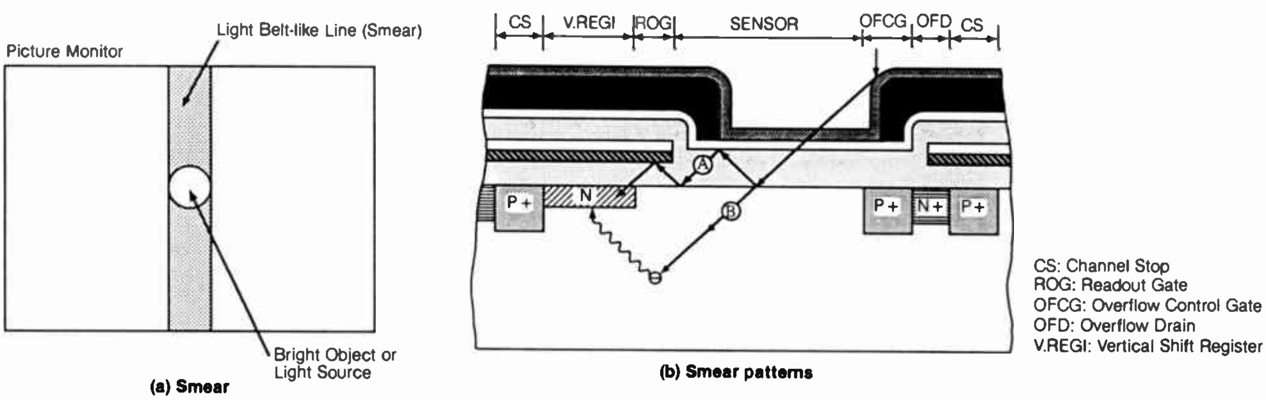


Figure 8. Vertical smear in an interline-transfer CCD. This characteristic of CCD cameras is observed when a very bright object appears in the scene. A vertical streak appears above and below the object or light source and is most commonly observed when the headlights of an oncoming car are shot with a CCD camera at night. It is caused by direct leakage of the incoming light into the vertical storage register path "A" or via electrons generated through the photoelectric effect, path "B" into the vertical storage register.

perfect, the IT imager achieves the long-sought goal of a true solid-state imager with performance suitable for many broadcast applications.

Note that there is currently no industry agreement on a method of measurement for the Vertical Smear CCD artifact. The method used by some manufacturers uses a square highlight at the center of the screen with an area equal to 10% of the active picture area. Based on the threshold of 2 IRE units as the level below which an artifact is not considered objectionable in the active picture area, a CCD camera with 85 dB smear rejection can be exposed to a highlight about 5½ f-stops above 100 IRE units of video before the 2 IRE unit smear level is reached.

Frame interline transfer (FIT) CCD Imager. The FIT imager combines the best features of the IT and FT structure and virtually eliminates the vertical smear phenomenon. It is, however, a more complex and hence more expensive semiconductor. (See Fig. 9 for the basic structure of a FIT CCD.)

In the FIT CCD, the charge packets from the sensor elements are first transferred to the interline storage register, and while still within the vertical interval, these charges are quickly transferred to a frame storage array totally protected from light. Since the charges are present in the interline register only for a very short time, the amount of contamination from an intense highlight is extremely small. In the FIT structure, the vertical smear phenomenon has been suppressed to an insignificant level of 120 dB or more below the main signal level.

Spatial Offsetting

Different from the FT CCD imager that has almost contiguous sensing elements, the presence of the interline storage register in the IT and FIT CCDs reduces the area available to the active sensor element to a fraction of the overall pixel area. A reduced sensing area carries the penalty of reduced sensitivity, but allows the use of spatial offsetting, an extremely powerful method, to increase the effective luminance resolution of the camera. When spatial offsetting is used, the red and blue CCDs are each bonded to their respective optical ports with extreme precision such that they are physically displaced by one-half pixel distance horizontally from the green CCD. (See Fig. 10.)

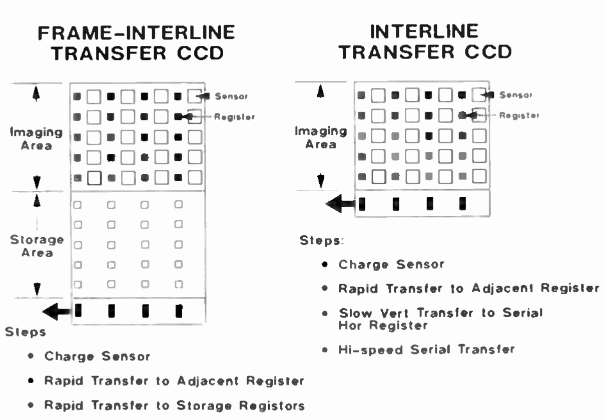


Figure 9. Steps in the transfer of pixel data for FIT and IT CCDs.

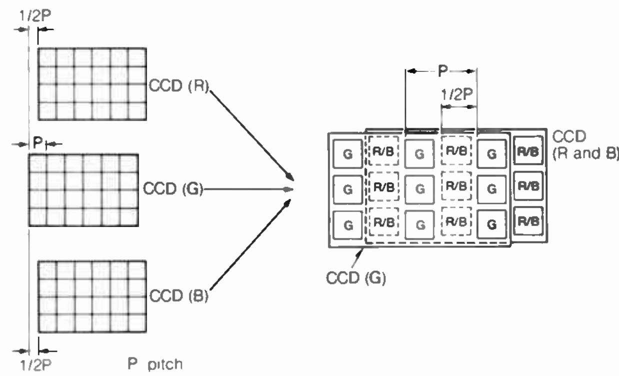


Figure 10. Spatial offsetting used in CCD cameras.

Spatial offsetting does not increase the resolution of the individual red, green, and blue channels, but spatial offsetting significantly increases the effective number of samples in the luminance channel, and thereby increases the effective resolution of this signal.

An example of the benefit derived from using spatial offsetting techniques is as follows:

A camera using three 768 pixel imagers that are co-sited (no spatial offset) achieves a limiting luminance resolution of 560 TV lines. When spatial offsetting is used in a camera using three 768 pixel imagers, a limiting resolution of 700 TV lines is achieved for the luminance (Y) channel.

A black triangle on a white background is used to illustrate the effectiveness of spatial offsetting to enhance the effective resolution of the luminance channel. In Fig. 11, the resulting image projected on the green imager is shown in close-up A, and the corresponding image projected on the red and blue imagers is shown in B. The quantities of charges corresponding to the input signal stored in the individual sensors of each CCD can be found in the waveforms labeled (1) through (7). Thus, the output of the green CCD can be visualized as shown in illustration A, and the outputs of the red and blue CCDs can be visualized as shown in illustration B. The luminance signal is derived by linear summing of the red, green, and blue signals and is shown in illustration C. The diagonal line in the original scene is reproduced with relatively coarse steps in the individual red, green, and blue channels, whereas the diagonal is reproduced with much finer steps in the luminance output, a clear indication of the enhancement in resolution that occurs in the luminance channel when spatial offsetting is used.

On-Chip Lens Array (OCL)

A relatively new word in the vocabulary of CCD terms, the *on-chip lens*, or OCL construction, repre-

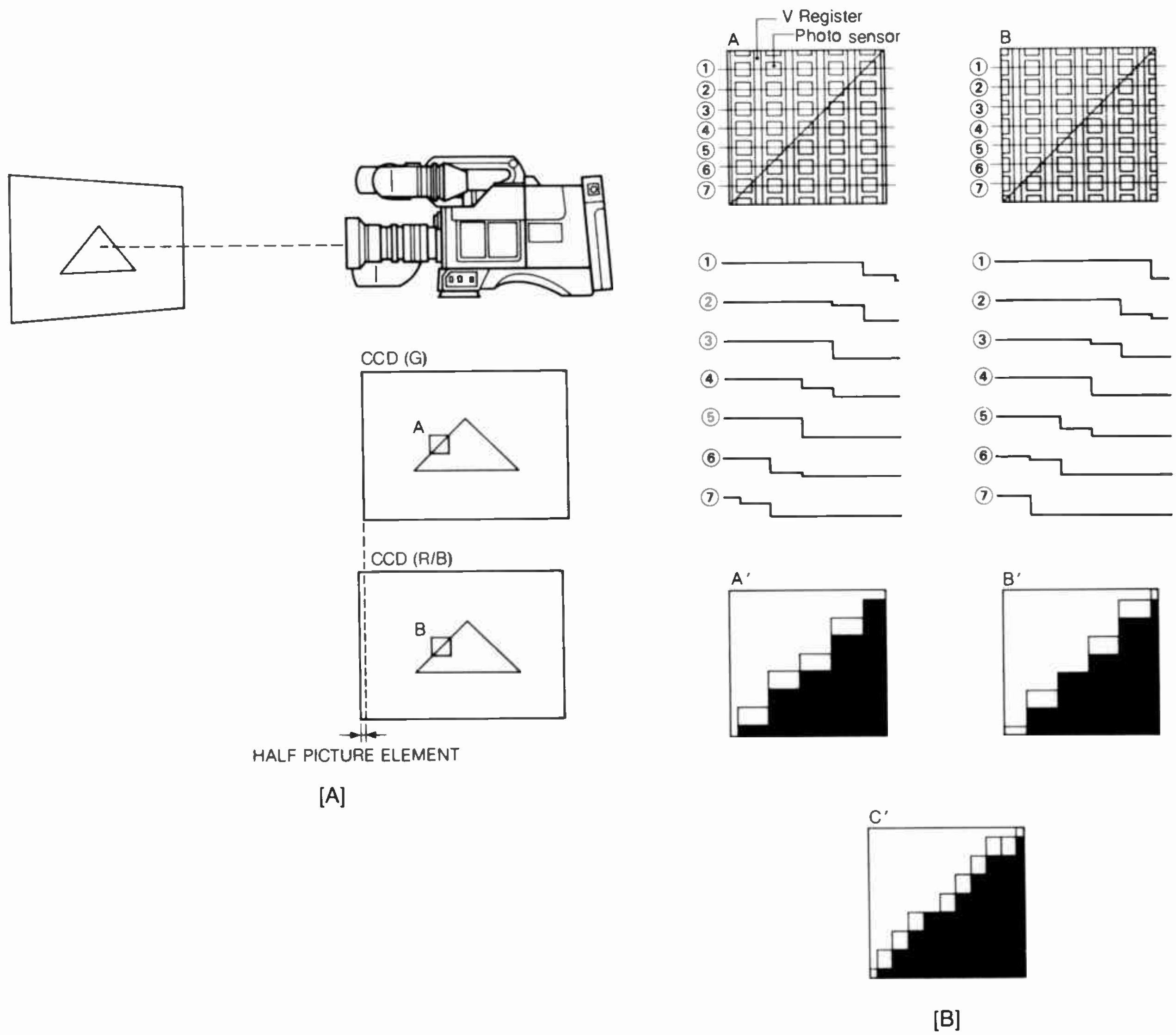


Figure 11. Enhancement of diagonal resolution by spatial offset of the green image with respect to the red and blue images.

sents a major advance in CCD technology. Sensitivity of $2/3$-inch IT CCD and FIT CCD cameras is increased from f/5.6 to a remarkable f/8.0 at 2,000 lux. This represents a threefold improvement in sensitivity compared to a $2/3$-inch Plumbicon® imager rated at f/4.5 and a four-fold increase in sensitivity compared to a camera equipped with Saticons® rated f/4.0 at 2,000 lux.

In an IT and FIT CCD imager, the area assigned to one pixel must be shared by the active sensor, the adjacent interline storage register and the necessary interconnect wiring. As a result, a significant part of the incoming light falls on an area of the pixel that does not contribute to accumulation of charge. With the OCL construction, the sensitivity of IT and FIT CCDs is dramatically increased by carefully positioning an array of microscopic lenses, one lens per pixel, over the imager to gather light over a much larger area and focusing it onto the photosensitive area of the pixel. By recovering light previously lost, almost one f-stop improvement in sensitivity is achieved. In addition to an increase in sensitivity, the OCL construction also reduces the amount of oblique rays that cause "smear" in a CCD and thereby dramatically improves the highlight handling capability of the imager. A slight increase in the diffusion of light is the only negative characteristic of the OCL construction and is far outweighed by the positive benefits.

CAMERA SIGNAL PROCESSING

CCD vs. Tube Cameras—Cameras using solid-state CCD imagers are very similar to cameras using tube

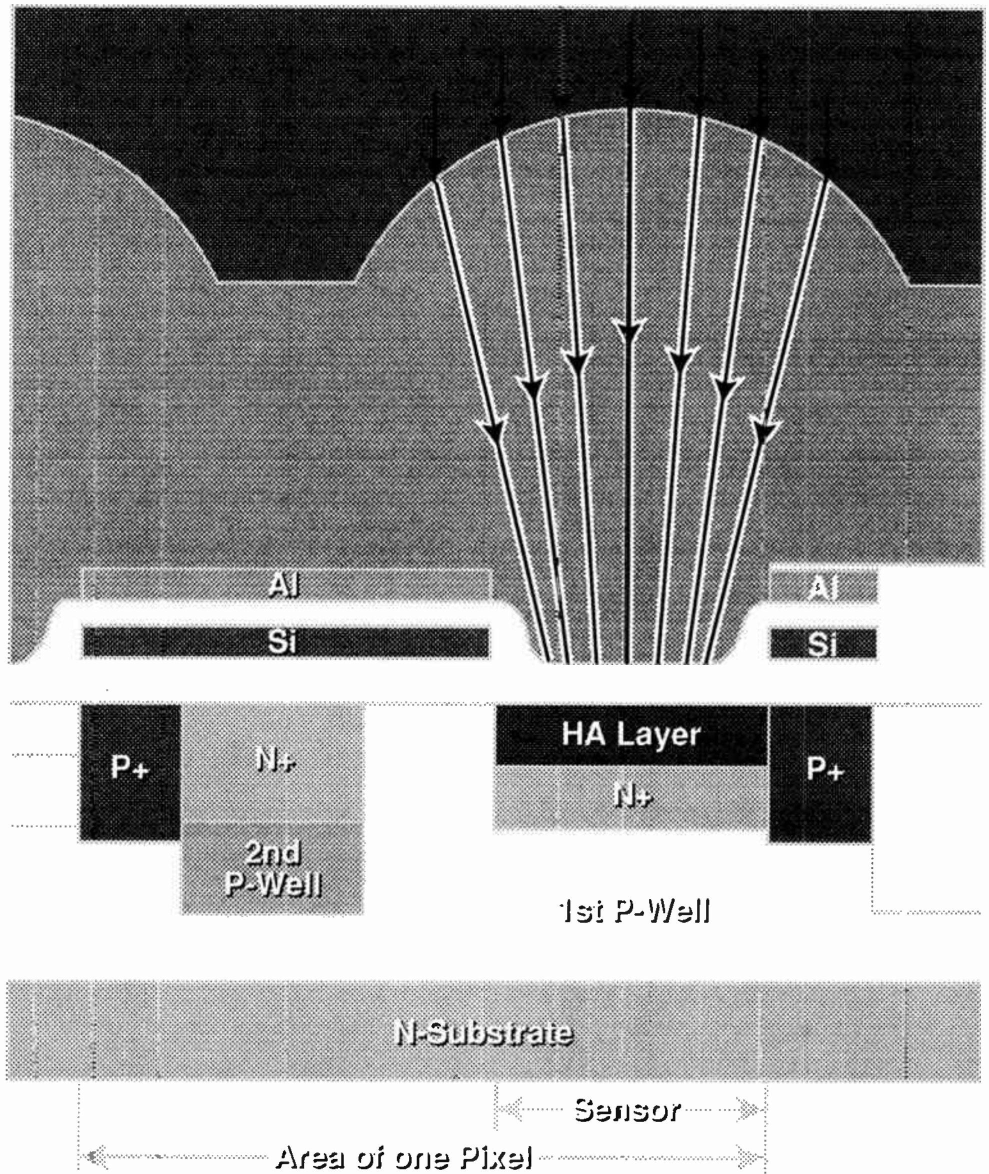


Figure 12. On-chip lens construction (courtesy Sony). The lens collects light that would otherwise fall on insensitive areas and concentrates it onto the imaging part of the pixel.

imagers. In the tube camera, deflection circuitry is used to scan the image. In the CCD camera, a clocking system is used to read the charges proportional to the intensity of the optical image. In both cases, an analog signal of similar characteristics is provided at the output of the pre-amplifier. The actual camera signal processing from this point is essentially the the same for both types of cameras.

This chapter focuses mainly on CCD cameras. Appendix D details some of the special considerations applicable for a pick-up tube camera. The basic block diagram of the camera signal processing stages are shown in Fig. 13.

As with tube cameras, the quality of the pre-amplifier used in a solid-state imager is critical to the performance of the imager, but in this case the pre-amplifier is part of the imager integrated circuit itself.

Signal Processing Stages

The signal processing requirements for a camera with CCD imagers are basically the same as the requirements of a tube camera. For both types of imagers, the red, green, and blue output signals of the pre-amp circuitry are essentially the same. These signals are processed in three separate but parallel paths in the signal processing stages.

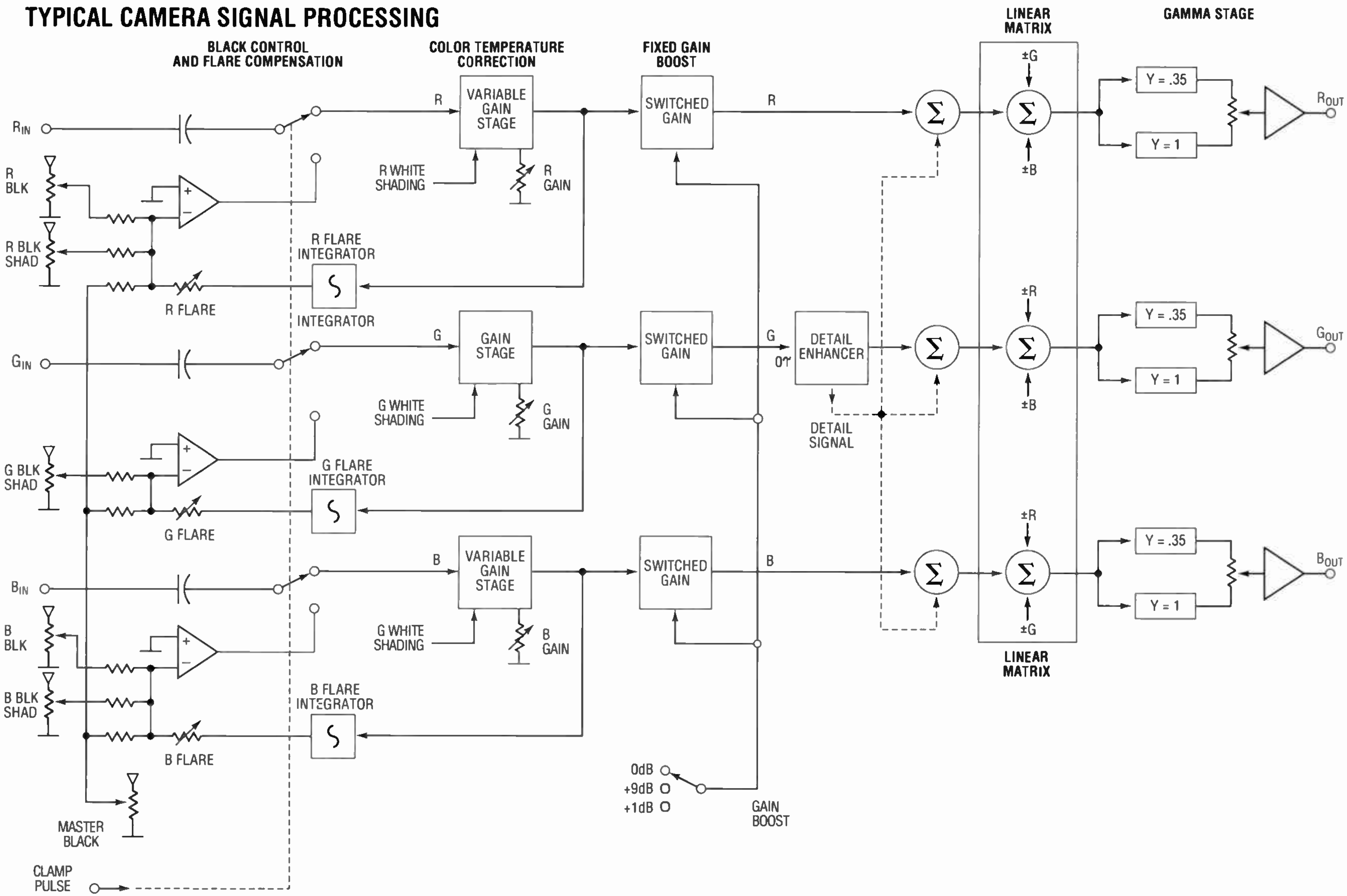


Figure 13. Basic camera signal processing.

The camera processing circuitry has four major functions:

1. Adapt to the normal variations in the scene illumination and content.
2. Correct for unavoidable errors in the lens and optical system.
3. Correct for the limited dynamic range and other limitations of the video recorder, transmission system, and display devices.
4. Introduce intentional pre-emphasis, such as gamma correction, to correct for a known nonlinearity in the cathode-ray tube.

The camera processing stages implement these specific tasks:

Black Level/Auto Black Set Circuitry: The residual signal level from a CCD imager corresponds to the absence of all light or pure black. This black current, which is a consequence of the electron mobility (a quality common to all semiconductor devices), increases with increasing temperature. For this reason, a small section of the CCD imager is masked to protect it from light. The black current reference is determined from masked pixels and used to maintain accurately a constant camera black level with variations in temperature. The overall camera black level can be adjusted using the master black control. The individual red/black level and blue/black level controls can be used for manual black balance, or the "Auto Black" circuit, normally provided in the camera, can be used to set the camera black balance automatically.

Shading: Shading correction circuitry compensates for errors in the lens, color separation optics, and imaging devices. Both black and white shading correction controls are normally provided in a professional camera. One example of a shading error is the nonuniform light transmission of a lens. Light transmission of a lens tends to be strongest at the center and falls off towards the edges. White shading (parabolic) can be used to reduce this error, significantly.

Flare Correction: The flare correction circuitry provides an approximate correction for the scattering of peripheral rays in the various parts of the optical system. When flare occurs, a small amount of light is scattered into areas of the picture that should be totally devoid of light, resulting in a general rise of the picture's black level. The amount of scattered light (i.e., rise in black level), tends to increase as the overall amount of light in the scene increases. In the flare correction circuit, an integrator is used to measure the total amount of light in the scene, as an indication of the amount of scattered light. A portion of this integral is then introduced into the camera's black level circuits to restore black areas of the picture back to true black (i.e., capped-level black).

Auto White and Color Temperature Correction: This circuitry is used to correct for the natural color temperature variations of the scene. The red and blue processing channels are equipped with a voltage-controlled gain stage that makes it possible to adjust the camera white balance using the red and blue gain controls. An *auto white* circuit, that establishes a white balance automatically, is normally also provided in the camera.

Gain Boost: This circuitry increases the effective sensitivity of the camera, at the expense of creating a noisier picture, by introducing at least three fixed electronic gain steps, typically totalling +9 dB or +18 dB, into the processing circuitry.

Linear Matrix: The primary function of the linear matrix is to correct for deficiencies in the color separation process. The spectral characteristics of the optical system do not provide negative lobes prescribed by the ideal spectral characteristics for the color separation process. A matrix with negative and positive coefficients is used in the signal processing stages to correct for the missing lobes.

If an adjustable linear matrix instead of a fixed matrix is provided in the camera, it is possible to match the colorimetry of two dissimilar cameras by judicious adjustment of this matrix. *Image Enhancement (Contour Correction or Detail Correction):* This circuitry increases the subjective crispness of the television picture. A detail signal is used to outline transitions in the image, resulting in an apparent increase in the crispness of the picture. A similar outlining of transitions does not occur in film. The closest parallel is in graphic arts, where the apparent crispness of a graphic is increased by outlining the transitions of text or drawings to make them stand out. The dramatic improvement in subjective picture quality achieved by the detail correction circuit is clearly shown by turning the circuit on and off in a working camera. High-end cameras generate enhancement for both horizontal and vertical transitions, and this type of image enhancer is called a *two-line enhancer*. To provide enhancement for both horizontal and vertical edges, it is necessary to delay the signal from the imager for two full horizontal line periods. Extensive circuitry, including

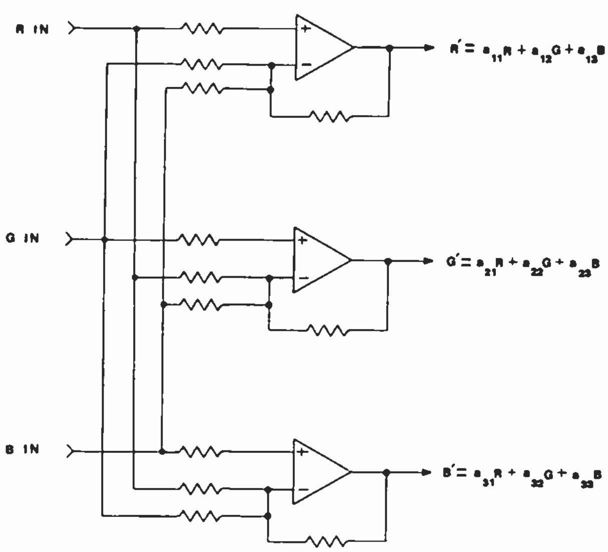


Figure 14. Basic diagram for the linear matrix circuit.

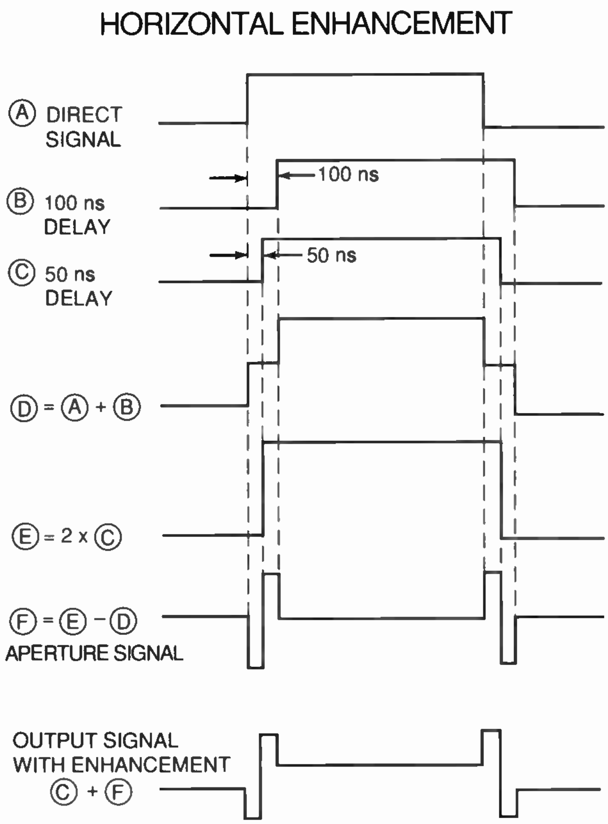


Figure 15. Generation of the horizontal enhancement signal.

two expensive glass delay lines, is required; hence the designation two-line enhancer. Less costly cameras frequently provide enhancement for the horizontal transitions only. This type of enhancer is typically called a *one-line enhancer* since it avoids the high cost of the horizontal delay lines.

The detail circuit generates the detail signal, and it is one of the most complex circuits in the camera.

Gamma Correction: This circuit precorrects for the nonlinearity of the CRT used in a TV set or monitor by boosting the signals near black and compressing the signal in the white part of the picture—the exact inverse of the nonlinearity of the CRT. The overall result of applying a signal with gamma correction to a CRT is a faithful reproduction of the gray scale in the original scene.

Automatic Knee, Knee, and White Clip: Outdoor scenes routinely encompass contrast ratios as high as 1,000:1. One of the most difficult tasks of the TV camera is to capture the important content of the scene, while compressing the unimportant details to fit into the limited 40:1 contrast range from camera to the home TV set, that is supported by the overall TV system.

Significant progress has been made in the ability of the TV camera to handle the large dynamic range of outdoor scenes. The hard white clip at 105 IRE implemented in early TV cameras resulted in a loss of significant picture information in the highlight areas. The next advance in handling was the *knee circuit*: a soft bend or knee starting at about 96 IRE introduced into the video output response, while the hard knee was moved up, to about 108 IRE. At least one additional f-stop of highlight detail was compressed in the range from 96 IRE to 108 IRE. The most recent advance in camera highlight handling is the *automatic knee circuit* also called dynamic contrast control (DCC) or *dynamic knee*. With this circuit, the onset of the soft knee is progressively shifted as low as 85 IRE with the existence of strong highlights. The equivalent of up to 600 IRE of picture information in the highlight area is now compressed in the range from 85 IRE to 108 IRE of video (i.e., the range of video that can be recorded by a video tape recorder, transmitted reproduced by a display device).

Auto-Iris and Zebra Circuits: In a production environment, significant time and effort are expended to

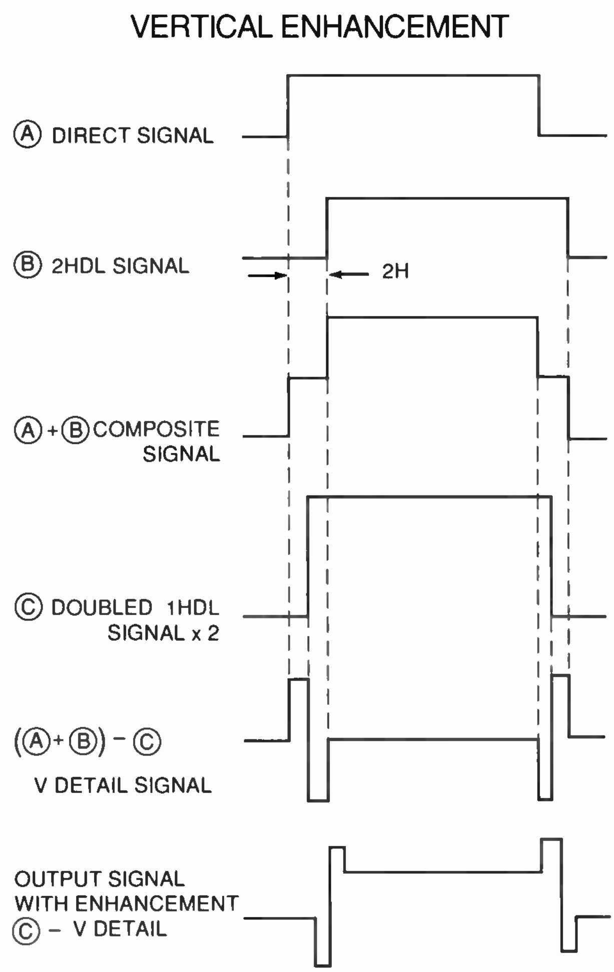


Figure 16. Generation of the vertical enhancement signal.

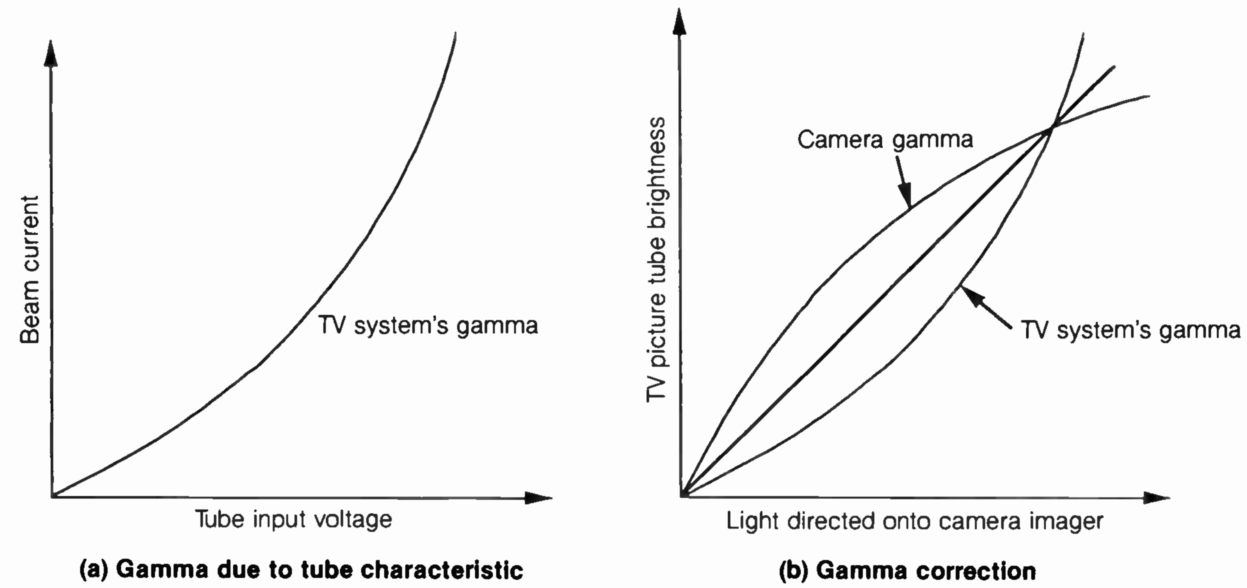


Figure 17. Gamma correction: intentional distortion of the camera output signal to correct for nonlinearity in cathode ray display tubes.

adjust the lighting and iris opening for the desired artistic effect. Manual adjustment of the iris with the support of a waveform and picture monitor is most commonly used in this application. The *auto-iris* circuit in combination with the *zebra* indication are the tools provided by the camera for ENG and field production, where it is not practical to use a waveform and picture monitor. The basic auto-iris circuit evaluates the video level, compares this level to a reference level (typically called the auto iris set point), then sends a correction signal to the lens to open or close the iris as appropriate to achieve proper exposure. Various designs and degrees of sophistication are used to improve the performance of the auto-iris circuit in difficult lighting situations. The method used to evaluate the video level is the key to improved performance. Some of the improvements to the basic auto-iris circuit are:

Peak/average picture evaluation: The peak level of the video signal is detected by one circuit; the average video level is detected by a second circuit; a mix of the two circuits (peak/average adjustment) is then used for improved iris control.

Weighting and *masking* are also commonly used to improve the performance of an auto-iris circuit; since most of the important action takes place at the central part of the picture, correct exposure at the center is most important. Without weighting or masking, a bright light source at the periphery of the picture, such as the sky, may cause the iris to close and provide improper exposure at the center. When bright highlights at the edges of the picture are excluded from the picture evaluation by masking, or the video level at the center of the picture is weighted more heavily than video at the periphery, performance of the auto iris circuit with difficult scene content is significantly improved.

Despite the various refinements in picture evaluation, the resulting auto-iris setting is not satisfactory with some difficult scenes. The zebra circuit can then be of significant help to find the proper manual-iris setting by superimposing diagonal stripes on bright areas in the viewfinder image. There are two common operating modes for the camera is zebra circuit. In one mode, it is adjusted to identify in the viewfinder all video at the 60 IRE to 70 IRE level (i.e., proper

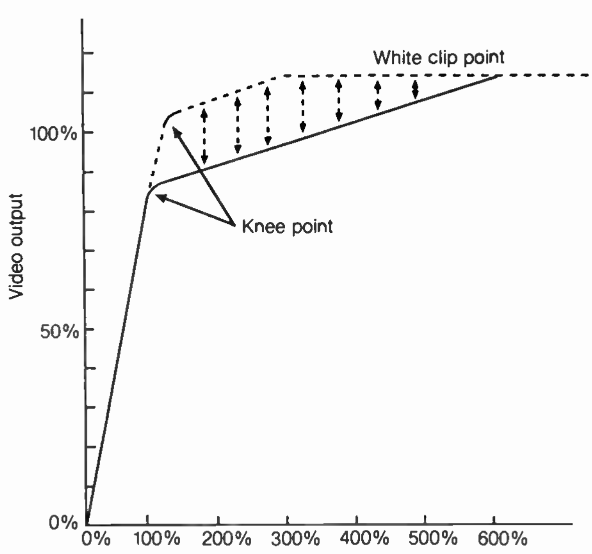


Figure 18. Auto-knee dynamic contrast control.

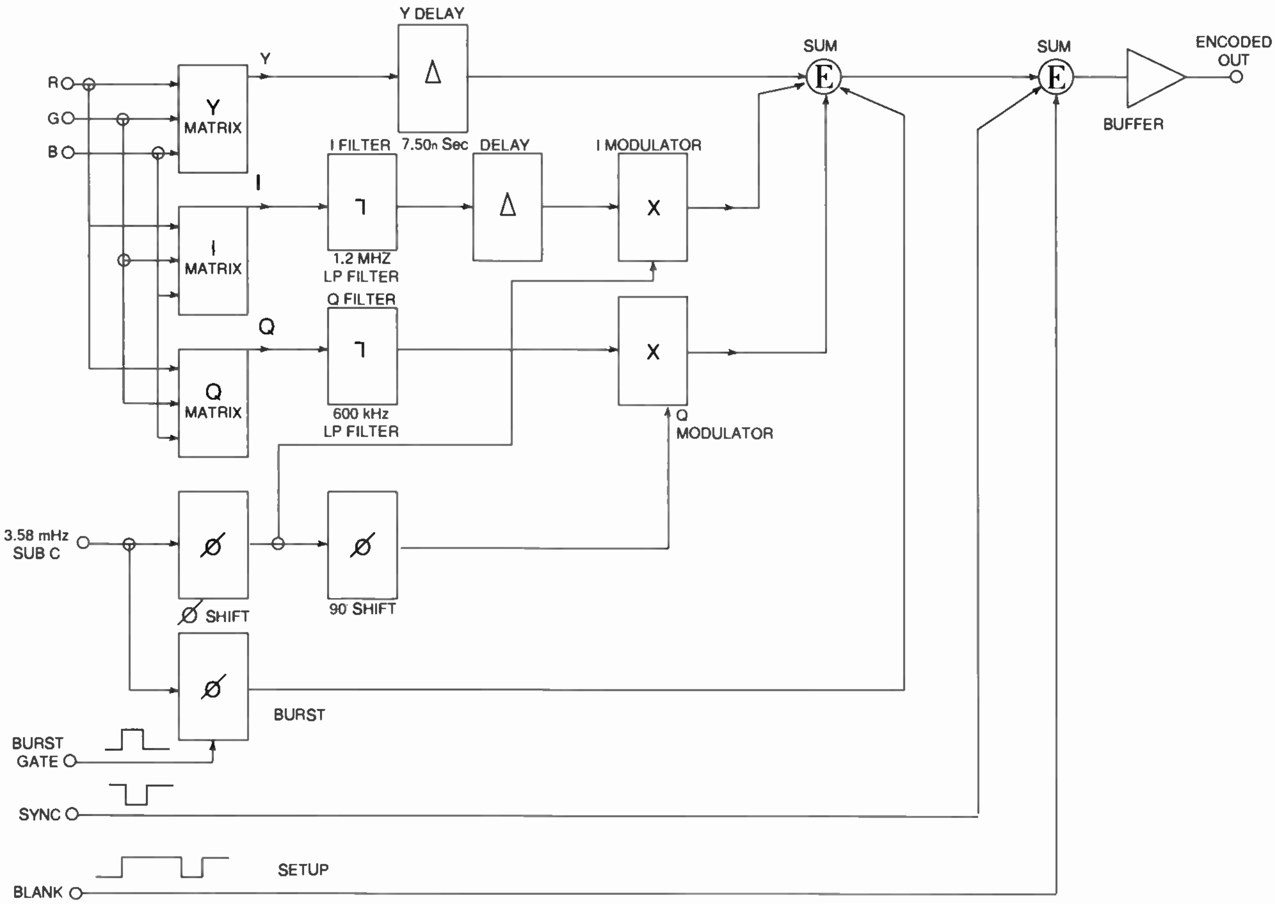


Figure 19. Simplified block diagram of an NTSC encoder.

exposure for facial features). The other operating mode of the zebra circuit is to identify all video that is at 100 IRE and above (i.e., all video that is going into the white clippers). In the latter mode, the iris is adjusted until the brightest part of the scene that still carries useful information is just below the threshold of the characteristic diagonal zebra stripes.

Encoder: For some applications of a TV camera the *component* video signals in the form of R/G/B or Y/R-Y/B-Y are used directly. For other applications it is desirable to encode the red, green, and blue signals generated by the camera into a *composite* color signal. In the U.S. the camera signals are normally encoded to the NTSC standard.

The encoder used in a high performance TV camera differs from other commonly used encoders in having exceptionally wide-band luminance response reaching as high as 10 MHz. Fig. 19 shows a simplified block diagram of an encoder.

The RGB output of the camera processing is applied to the luminance, or Y matrix, and the I and Q color matrixes. The equations of these three matrixes are:

$$Y = 0.30R + 0.59G + 0.11B$$
$$I = 0.60R - 0.28G - 0.32B$$
$$Q = 0.21R - 0.52G + 0.31B$$

The output of the Y matrix is delayed by about 750 nanoseconds to match the inherent delay in the I and Q chroma channels. The I and Q signals are processed by 1.2 Mhz and 0.6 MHz low-pass filters respectively. The Q filter inroduces a delay of about 750 nanoseconds which introduces delay into both signals. While the I signal filtering is less severe and has correspondingly less delay. An additional delay or a special filter design is used to equalize the delay of the I and Q channels. The bandwidth-limited I and Q signals are then applied to the I and Q modulators where the I and Q information modulates two 3.58 MHz subcarriers in quadrature (phase shifted by 90°). The output of the I and Q modulators, the 3.58 MHz burst, sync, and set-up are added to the Y signal to create the composite NTSC output.

Normally a color bar generator is included in the encoder to record a reference signal on the header of a tape and to aid in the correct set-up of a monitor.

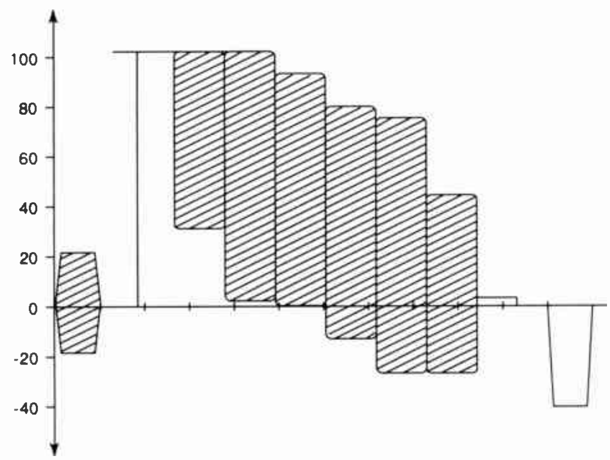


Figure 20. Encoder output waveform.

The typical output waveform of an NTSC encoder is shown in Fig. 20, with a color bar signal. The phase relationship of the I and Q axis vs. burst is shown in Fig. 21.

Sync/Timing Generator: This is the master timing system which provides the various timing waveforms required by the camera, in accordance with the television standard used. In the U.S., broadcast cameras are normally designed so that the output signal meets the requirements of the Electronic Industries Association (EIA) recommended standard RS-170A.

The timing waveforms required by the camera circuits are normally derived by digital countdown from a crystal oscillator operating at four times subcarrier frequency ($4f_{sc}$) or 14.3 MHz. A temperature compensated crystal oscillator is used to maintain the strict frequency requirements of the RS-170A standard when the camera is used in a self-contained mode. For multi-camera operation it is necessary to synchronize the cameras to each other. For this purpose, the camera sync/timing generator is provided with gen-lock capability. The sync/timing generator has the ability to phase-lock to an external composite video signal applied to the camera gen-lock input.

Power Supply: The power supply voltages required to operate a solid-state camera are quite low, and no special precautions need to be observed when working on the inside of a CCD camera. The accelerating voltages used in the camera viewfinder are the exception, of course. Tube cameras, on the other hand, require substantial accelerating voltages; and appropriate precautions should be observed when working on any viewfinder or tube camera.

Virtually all modern cameras use a switching regulator that converts the incoming voltage to the various voltages required by the camera. The universally accepted power supply range for portable cameras is 10.8 to 17.4 VDC; and by the nature of a switching regulator, the camera presents a constant power load to a battery or an external power supply.

THE DIGITAL CAMERA

Digital video processing was first introduced into videotape recording, where the benefits were most dramatic and cost-effective. In other types of studio equipment, digital processing was too expensive for the benefits it provided. As a result, digital islands in an analog environment represented the most cost-effective system configuration. With time, the performance of digital circuits has improved dramatically, while costs have been reduced significantly. It is increasingly cost-effective to use digital circuitry in production equipment, including cameras.

The standard imager for cameras is currently the CCD, a thoroughly analog device despite the fact that the information from the imager is read out in discrete packets. The dynamic range of a CCD or a pick-up tube imager, is quite large. It is not uncommon for the early processing stages of an analog camera to process faithfully signals as high as 600 IRE of video. Digital processing with a very high number of bits-per-sample is required to handle this large dynamic range while resolving fine shades of luminance difference. For this reason, current digital processing cameras use analog circuitry in the early stages of the camera signal processing, and limit digital implementation to the latter stages of the image processing. Despite the advances in digital processing circuitry, power consumption and weight of current digital processing cameras are larger than equivalent analog cameras, and intensive efforts are under way to improve these aspects.

Since studio cameras are inherently large, an increase in weight and power consumption is not an impediment to digital implementation. The limited bandwidth available with triax cable from the camera head to the camera control unit (CCU) is the real obstacle. Triax cable has become the industry standard for studio cameras by virtue of its demonstrated ruggedness and reliability in the most severe field production applications. A multi-conductor copper/

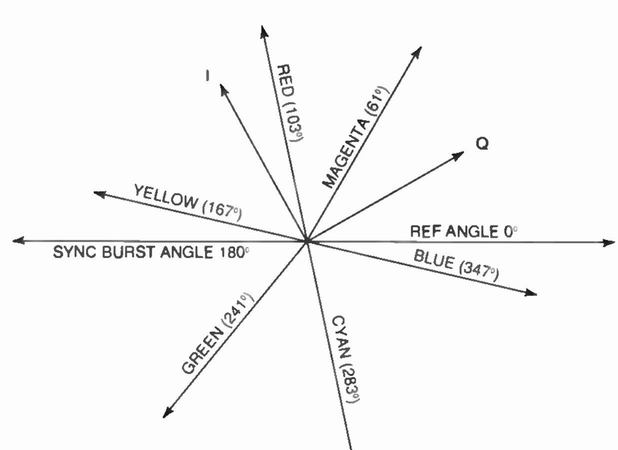


Figure 21. Subcarrier phase relationship in the NTSC encoder.

fiber-optic cable suitable to provide power to the camera head, and provide the wide bandwidth required for digital video transmission, must demonstrate ruggedness and reliability that is similar to triax to become accepted by the user.

Until the current analog imagers are displaced by true digital imagers, studio and portable cameras will remain hybrid analog/digital devices. Recent advances in the performance of digital circuitry now make it likely that within the next several years, cameras that include digital processing will equal or exceed the performance of the current analog cameras adding one more key element to the full digital studio.

BASIC CAMERA PERFORMANCE VERIFICATION

With tube cameras, it was almost mandatory to fine tune the camera before a major shoot to achieve optimum performance. The intrinsic stability of the CCD imager and the stability of the circuitry used in current CCD cameras now make it possible to operate the camera for several months without internal readjustment. Physical damage to the camera or lens, in use or transport, is probably the most frequent cause for a loss of performance in a current CCD camera. With careful handling, the probability of malfunction is very small. It is nevertheless prudent to schedule a quick check-out of the camera before the start of a major shoot, when the high cost of talent and other aspects of the production are considered.

The following items are appropriate for inclusion in such a quick check-out procedure. If the test results show a significant deviation from the manufacturers specifications or from the data previously obtained for the same test, a more thorough examination of the camera, as prescribed in the camera service manual, is then indicated. Several basic tests with expected data are described. Adjustments appropriate to CCD cameras are described.

Visual Inspection and Mechanical Check

Visually inspect the camera and lens for any evidence of physical damage as a clue to the possibility of more serious internal damage. Carefully operate the lens adjustments: manual and servo zoom, focus, and manual iris. If there is evidence of binding or a rough spot in any of these adjustments, physical damage that may affect the optical performance of the lens must be suspected. Inspect the front and rear lens elements, clean if necessary using pure alcohol and soft, lint-free wipes. Fingerprints in particular should be removed as quickly as possible to avoid harm to the optical coating of the lens elements. Note that the lens manufacturers generally discourage the use of silicon-impregnated wipes for cleaning high quality optics.

Confirmation of the Camera Encoder

A properly adjusted encoder is particularly useful since it provides a convenient window to look inside the camera and confirm proper operation of the remaining circuitry. Encoder set-up is particularly easy to confirm since the color bar generator, normally provided in a professional camera, provides a convenient self-test of the encoder. To confirm proper operation of the camera encoder:

1. Apply the camera (encoder) output signal to a waveform monitor (WFM), vectorscope, and picture monitor (a high- resolution black and white monitor with 800 television lines (TVL) or higher resolution is recommended).

   It is important to terminate the WFM, vectorscope, and picture monitor using a discrete 75 ohm termination. The preferred tolerance for this termination is $\pm 0.1\%$ and no worse than $\pm 1\%$. Internal terminations should not be used unless they have been tested and, if necessary, replaced with terminations within the recommended tolerance.
2. Select the color bar mode on the camera. Confirm on the vectorscope, the burst and I and Q vectors are of the correct phase and amplitude. Confirm that all of the color bar vectors fall within the tolerance boxes on the vectorscope.

If all of the above vectors are within tolerance, correct operation and adjustment of the encoder is confirmed.

Confirmation of Auto Black

1. Activate the auto black circuitry of the camera.
2. Confirm that the lens is capped during this adjustment. Confirm the character display in the viewfinder which indicates the auto black adjustment has been successfully executed.
3. Select the 0 dB, +9 dB, and +18 dB gain settings in sequence, and confirm that the black level adjustment is correct for all three gain settings. This is most easily confirmed with the vectorscope: the output signal should be a dot at the center of the display with no shift in position as the gain is switched. The only change should be an increase in noise at the higher gain settings.

Adjusting the Lens Back-Focus

This adjustment trims the lens to the specific optical dimensions of the camera. Whenever a new lens is put on a camera, it is necessary to make this adjustment. Some lenses use a screwdriver lock, while others use a knurled knob to lock the lens back-focus adjustment in place. Accidental misadjustment in use has been known to occur, and it is therefore recommended to confirm this adjustment.

To adjust the lens back-focus:

Locate a Siemens Star Chart (Fig. 22) at least ten feet from the camera. Place the chart in a location with low lighting such that the lens iris is wide open. Using the high resolution picture monitor:

1. Adjust the zoom lens for full close-up and adjust for best focus using the focus ring on the front of the lens.

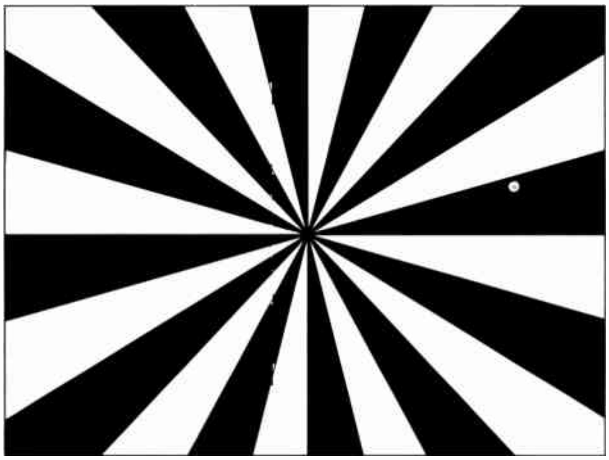


Figure 22. Siemans star chart.

2. Adjust the zoom for maximum wide angle position and adjust for optimum focus using the lens back-focus adjustment. Repeat both steps several times. Securely lock the lens back- focus adjustment in place. Confirm that the lens stays in focus over the full zoom range.

Confirmation of Black Shading

1. Cap the lens. Raise the master black level to about 10 IRE to 12 IRE to avoid any clipping.
2. Using the waveform monitor in the vertical display mode, then in the horizontal display mode, if the black level is a thin horizontal line, there is no black shading in any of the three color components.
3. Restore the black level to its proper position.

Adjusting the Detail Circuit

Using the 11-step gray scale chart, confirm the amplitude of the detail signal as required for the application. A stronger detail signal is typically required for a lower performance recorder such as U-Matic®, and less for a higher performance recorder such as a BetacamSP®.

Optional Tests

Camera System Resolution

If the camera system is capable of resolving fine detail close to the limiting resolution specified by the manufacturer, there is a strong assurance that the lens, camera optics, and overall camera signal processing are working correctly.

Use a suitable chart with resolution wedges or a chart with a multi-burst pattern and confirm that the overall resolution of the camera system is close to the manufacturers specification.

White Shading

To confirm white shading:

1. Set up a uniformly-lit white test chart.
2. Using the waveform monitor, open the lens to obtain about 70 IRE units of video (confirm the iris is in the range of f/4.0 to f/5.6, adjust the lighting if necessary), confirm that there is a minimum of vertical, then horizontal, shading.
3. Correct using the camera horizontal and vertical white shading controls.

Flare

The camera flare correction circuitry provides an approximate correction for flare or scattering of peripheral rays in the various parts of the optical system.

To confirm the adjustment of the flare correction circuit:

1. Frame an 11-step gray scale chart that includes a very low reflectance strip of black velvet added to the chart.
2. Adjust the iris from fully closed until the white chip is at 100 IRE units of video. The flare compensation circuitry is adjusted correctly if there is almost no rise in the black level of the velvet strip as the white level is increased to 100 IRE units; and only a small rise in the black level, with no change in hue, when the iris is opened one more f-stop beyond the 100 IRE units point.
3. Adjust the R, G, and B flare controls as defined in the camera service manual if the flare correction is not adjusted correctly.

Linear Matrix

When it is necessary to use two dissimilar camera models in a multi-camera shoot and either of the two models provides an adjustable linear matrix, it is possible to use the variable matrix to obtain a better colorimetry match between cameras. Specific matrix parameters and adjustments (if any) will be found in the camera service manuals.

CAMERA SELECTION

This section is intended to give some practical guidelines as to the types of camera to consider for a given application.

Studio Remote Applications

Triax cameras are now slowly displacing multi-core cameras in the studio. The small size, ruggedness, and low cost of triax cable, plus the ease of repair, provide a strong advantage over cameras using multi-core cable.

A *studio camera* system is the system of choice:

- When there can be no compromise in picture quality. The ability to use studio lenses with superior optical performance is a key factor.
- When a teleprompter is required for coaching the talent and both power and the signals for the teleprompter must be carried in the camera triax cable.
- When extended zoom capability (well above the 18 mm × 10 mm capability of portable lenses) is

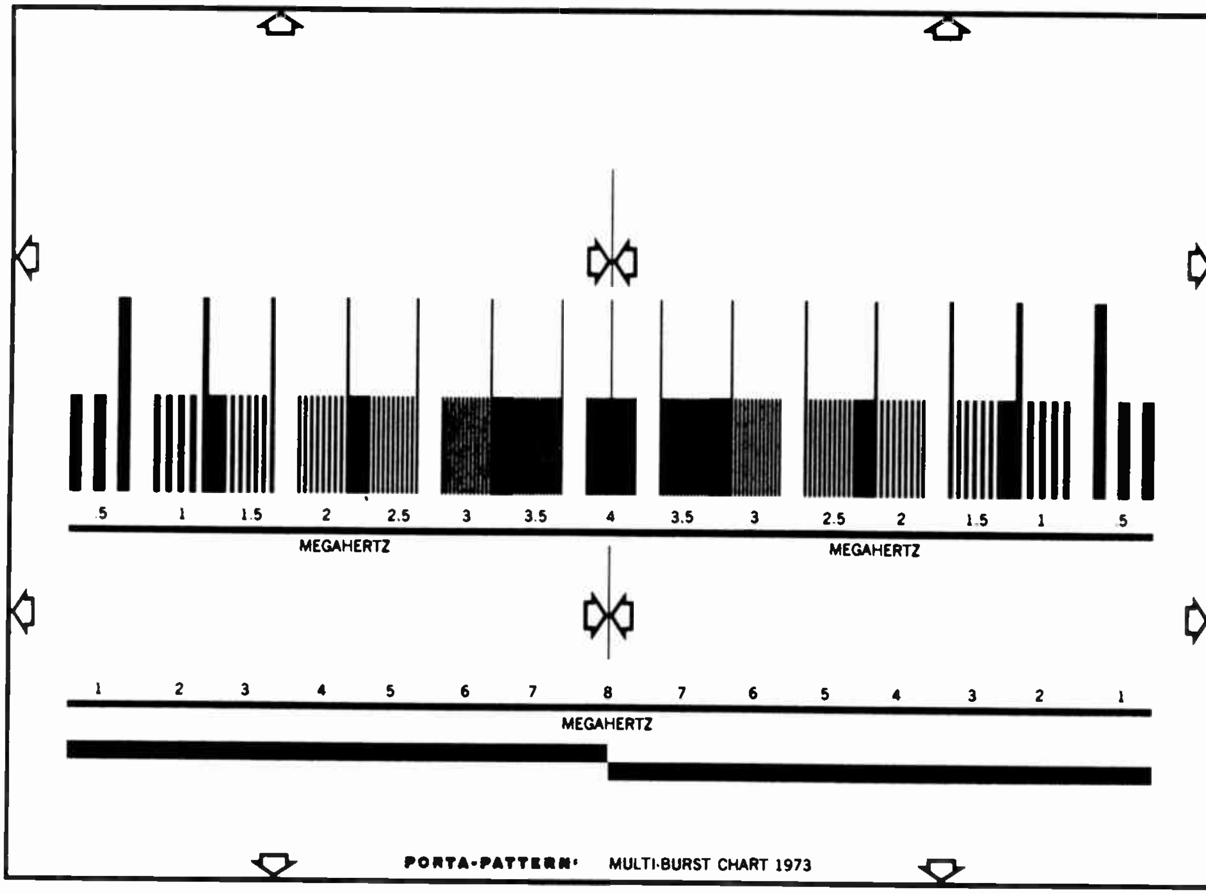


Figure 23. Multiburst chart (10 MHz).

required, which can only be achieved with studio lenses.

- When extensive remote control capabilities are required. Automatic set-up of the camera system is typically provided in these cameras.

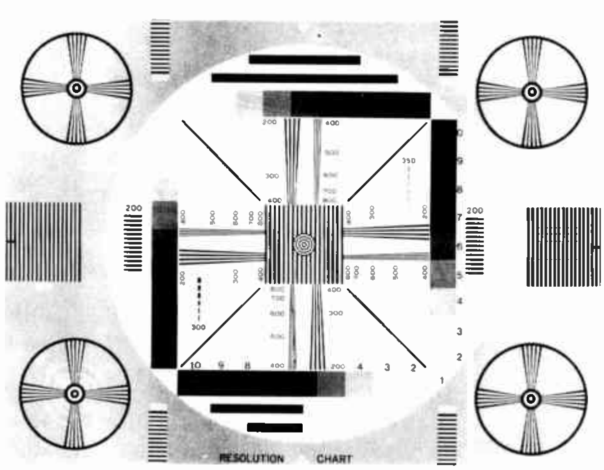


Figure 24. Resolution wedge chart.

A *portable camera* is indicated when:

- Mobility and handheld operation is important.
- The "reach" of available ENG style lenses is adequate.
- The lighter weight of the portable camera head is critical such as with automated pedestals.
- Teleprompter and multiple channels of intercom are not mandatory, or use of additional cables is not objectionable.

Electronic Field Production (EFP)

Field production has seen a gradual change from the use of a self-contained camera and a separate VTR to the use of a portable camera that is capable of accepting a dockable VTR, creating a two-piece camcorder. With the camcorder, creative freedom is enhanced significantly by the ability to shoot on a tripod or shoulder, as desired. The dockable camera can also be configured with a suitable adapter and operated with a camera control unit (CCU) for convenient remote operation of the camera system.

The one-piece camcorder, that integrates the camera and recorder in one unit, is now becoming the camcorder of choice for field production use due to its outstanding ease of use and high performance. Frequently used in conjunction with a handheld remote control unit, the one-piece camcorder now combines convenience with artistic control.

Electronic News Gathering (ENG)

For the ENG cameraperson, weight, size, and ease of use are the critical issues, particularly with the trend to reduce ENG crew size. The one-piece camcorder optimizes size, weight, and balance to provide the most ergonomic camera for ENG use. For these reasons, the one-piece camcorder is now the virtual standard camera for ENG use.

Camera Specifications

A TV camera performs the complex task of creating an electronic image of a real scene, ranging from scenes with extreme highlights, to scenes with large dynamic range that must be compressed to fit within the capability of the TV system, and to scenes with marginal illumination. Defining the performance of a camera in a complete but concise set of specifications may be impossible, but camera manufacturers need to provide more complete specifications than are currently provided. It is not unusual to find a low-cost camera with virtually the same published specifications as a camera costing significantly more. Actual day-to-day performance, on the other hand, will probably show the more expensive camera to be far superior in handling difficult lighting situations. For this reason the published camera specifications are no more than a basic guide for cameras to consider for actual evaluation.

It is no longer necessary to limit the choice of camera to the one with the best picture quality since virtually all current professional CCD cameras make high quality images. Such factors as ease of use, cost, and operational features can now frequently be the deciding factors in choosing one camera over another for a specific application.

In an actual camera evaluation, it is typically not necessary to spend a great deal of effort to confirm the specifications provided by the manufacturer. Instead it is recommended to expose the camera or cameras considered for purchase to the most difficult shooting situations that they are likely to be exposed to in the intended application(s), and then make the final choice based on the overall advantages of one camera over the others.

APPENDIX A—COLOR SEPARATION SYSTEMS

A color camera contains an optical system that separates the incoming light into its three component colors. Currently, the three types of color cameras, in descending order of performance, are:

- The "three-imager" design with prism type color separation.
- The "two-imager" design with a beam splitter and color- stripe color separation for the chrominance imager.
- The "single-imager" design using alternate optical filters over the individual pixels of the imager to achieve color separation.

Dichroic mirrors have also been used for color separation in some portable and studio cameras[2]. The dichroic mirror system is not used in current cameras because there is significant loss of light as compared with the prism system.

The prism color-separation system has become the dominant system for use in broadcast cameras and will therefore be described in more detail.

Color Separation System for Three-Imager Color Cameras

In front of the prism, several color correction and neutral density filters are mounted on a filter wheel to provide large-scale compensation levels for the color temperature and intensity of the incoming light. The infrared filter, quarter-wavelength filter, and anti-aliasing filter described earlier are also located here. The beam-splitting prism system makes ingenious use of selective reflection at dichroic layer surfaces to separate the incoming light into the red, green, and blue color components. Total reflection, at untreated surfaces, is also used within the prism system.

A dichroic layer is formed by vacuum evaporation of 10 to 20 layers with alternating high and low refractive indices. Proper choice of the material and thicknesses can give the dichroic layer the property of reflecting only one color and passing other colors.

The spectral characteristics of the prism are the determining factor in the quality of color reproduction achieved by the camera. The trimming filters on the output of the prism are used to shape the rising and falling edges of the spectral characteristics to improve the reproduction of neutral colors. Fig. A-2 shows an example of the spectral characteristics of a complete prism color-separation system. The spectral character-

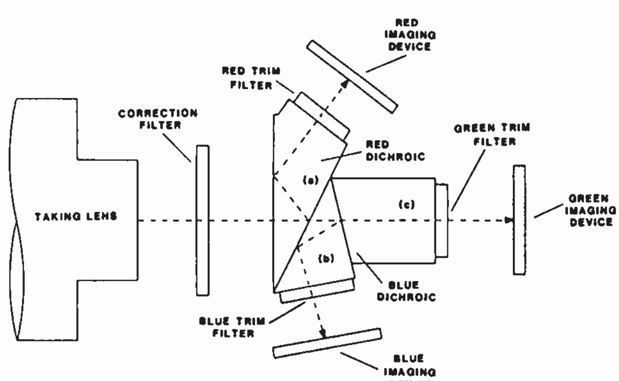


Figure A-1. Three imager prism-type separation system. The color-selective characteristic of the blue-reflecting and red-reflecting dichroic coating divert the blue and red components of the image to the respective blue and red imagers. The remaining component, green, passes straight through to the green imager. The imagers, themselves, have no special color characteristics.

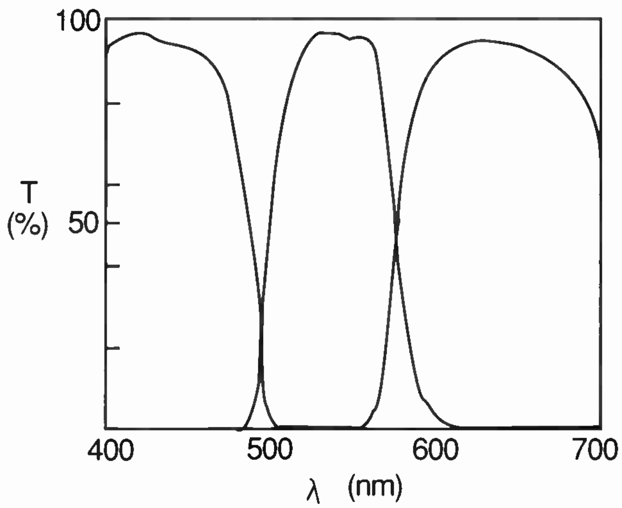


Figure A-2. Spectral characteristics of an entire color separation system.

istics of the complete optical system are called the taking characteristics of the camera. The ideal taking characteristics are shown in Fig. A-3. Note that the ideal taking characteristics include negative lobes, shown dotted in this figure. The optical system cannot generate the required negative lobes (i.e., cannot subtract light); however, subtraction of the video signals is possible in the camera processing stages. One of the functions of the linear matrix circuitry is to correct for the missing negative lobes in the optical system.

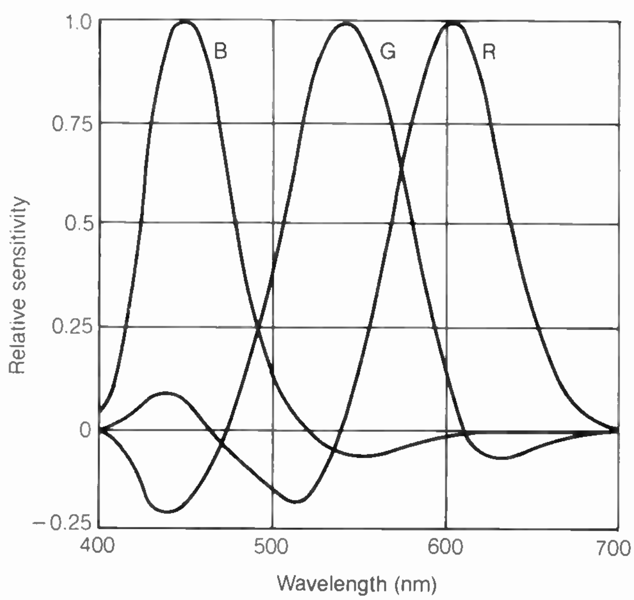


Figure A-3. Ideal taking characteristics. (Courtesy Wiley Publishing.)

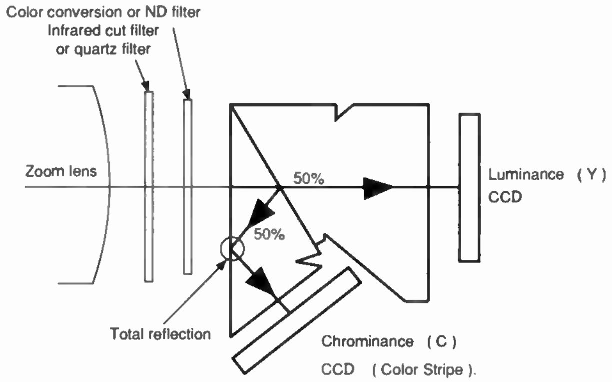


Figure A-4. Two imager color separation system.

Color Separation System for Two-Imager Color Cameras

The incoming light is applied to a partially silvered mirror (or prism surface) to divert half the incoming light. The straight-through portion goes to the luminance (Y) CCD imager. The reflected portion goes to the chrominance channel CCD imager. Color separation is achieved on the chrominance imager by the use of alternate red and blue optical filters on the pixels of the CCD. The luminance resolution of the two-imager system is high. Color performance is only a little better than the single-imager system.

Color Separation System for Single-Imager Color Cameras

In the classic single-imager camera, the red, green, and blue optical filters are deposited on the surface of the CCD in a mosaic or stripe filter pattern.

TABLE 1

Relative performance of a mosaic vs. a stripe filter.

MOSAIC FILTER	STRIPE FILTER
Higher Cost	Lower cost
Higher H resolution	Lower H resolution
Higher sensitivity	Lower sentivity
Complementary color filter	Primary color filters
Poorer colorimetry	Better colorimetry

In some cases an alternate pattern of color filters is used to achieve a different combination of characteristics.

APPENDIX B—LENS CHARACTERISTICS

A zoom lens is typically used with a television camera. The zoom lens has the following characteristics:

1. The focal length can be continuously varied over a range. The action of varying the focal length is called zooming.

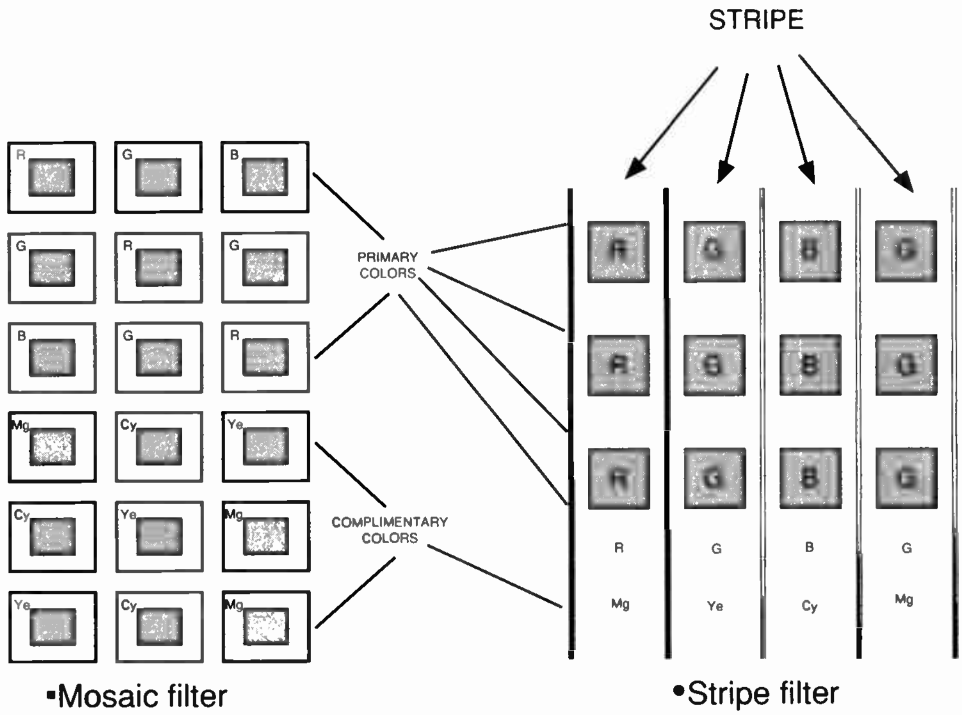


Figure A-5. Single imager color separation system.

2. The image position does not change as the lens is zoomed (focal length of the lens is changed). The ratio of the maximum focal length divided by the minimum focal length is called the zoom ratio.

A common type of zoom lens is illustrated in Fig. B-1. In this type of zoom lens, a lens group called the variator moves along the optical axis to vary the focal length, while another lens group called the compensator is moving in a complementary manner to maintain the image in the same plane. Precisely machined cams are used to effect the complex movement of the variator and compensator elements.

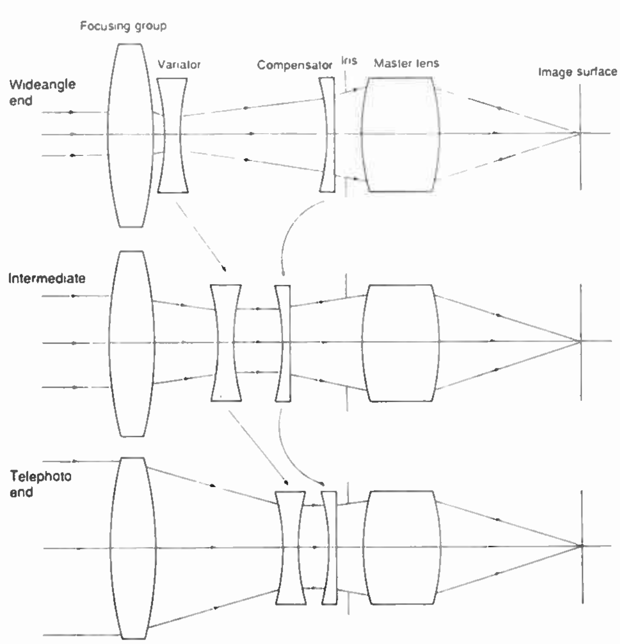


Figure B-1. Common type of zoom lens.

Focal Length

If parallel rays of light pass through a (convex) lens, they will converge to one point on the optical axis, called the focal point of the lens. The focal length of a lens is the distance from the center of the lens to the focal point. Practical camera lenses consist of several convex and concave lenses grouped together to minimize aberrations, however, this combination of lenses functions essentially the same as a single convex lens of negligible thickness (and no aberrations) also called a thin lens. The focal length of a complex lens is the distance from the center of this equivalent thin lens to the focal point. The focal length is the most fundamental parameter of a lens and directly defines the angle of view of the lens.

Angle of View

The angle of view of a lens is directly related to the image size and the focal length, as shown. Currently, the standard television system uses an aspect ratio of four units horizontal to three units vertical (4:3); therefore, the angle of view is greater in the horizontal than the vertical direction.

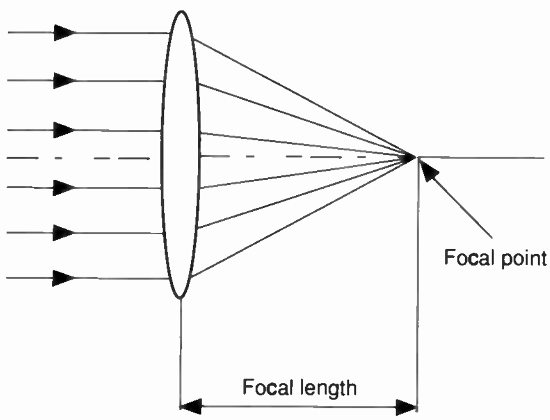
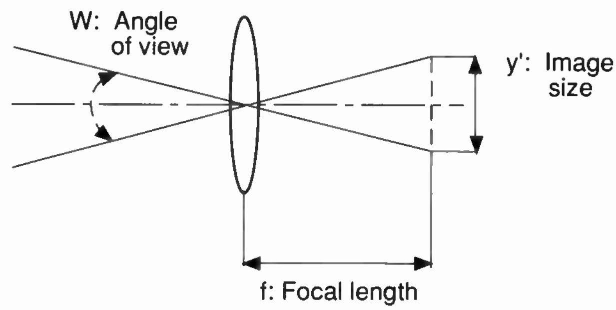


Figure B-2. Focal length of a single lens or "thin" lens.



Angle of view: $w = 2\tan^{-1}\frac{y'}{2f}$

Figure B-4. Calculating the angle of view.

Field of View
The field of view is the size of an object that fills the image format.

$$y = 2\,l\tan\frac{w}{2} \quad or \quad y = y'\frac{l}{f}$$

where: y = Object dimension
l = Object distance
w = Angle of view
y' = Image size
f = Focal length

For example, if the object being viewed is 1,000 feet away and the lens has a vertical angle of view of 2°, then

Vert. Field of View = 2 × 1,000 × tan(2°/2)

Vert. Field of View = 35 feet

Horiz. Field of View – 35 × 4/3 (aspect ratio) = 46.7 feet

Relative Aperture (f-number or f-stop)
This characterizes the speed or light gathering power of the lens. The relative aperture is defined as:

$$F\text{-}number = \frac{f}{D}$$

where: f = Focal length
D = Effective aperature diameter

For example, the typical sensitivity of a CCD type ENG camera is 200 lux at f/5.6 (≅186 foot candles).

f-number	Illumination Sensitivity
f/5.6	2,000 lux
f/4.0	1,000 lux
f/2.8	500 lux
f/2.0	250 lux
f/1.4	125 lux

A smaller f-number means a more sensitive lens. The f-number is also closely related to the depth of field.

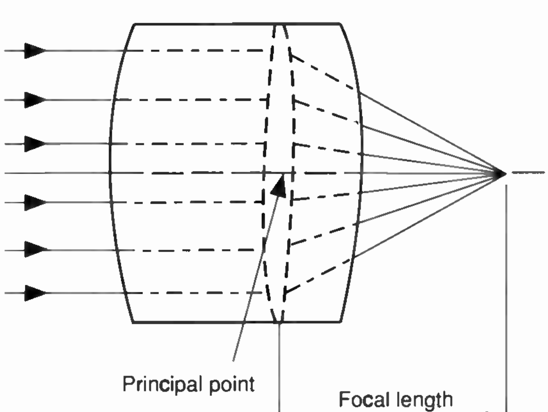


Figure B-3. Focal length of a compound lens.

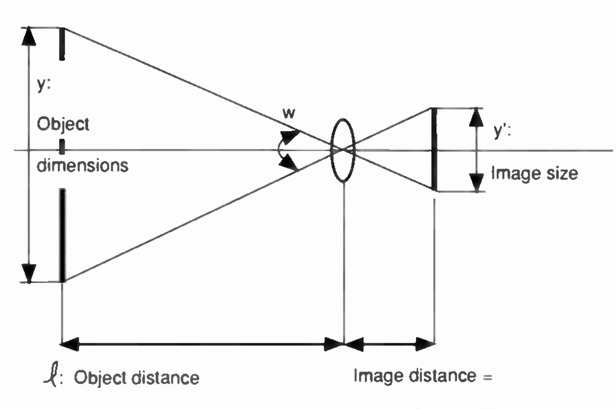


Figure B-5. Field of view.

Depth of Field and Depth of Focus

Depth of field is defined as the difference between the maximum and minimum distances to a subject such that the image is within an allowable amount of blurring. (See Fig B-6.) The human eye cannot detect blurring below a certain size called the circle of confusion (the permissible circle of confusion becomes smaller as the resolving power of the system is increased). Correspondingly the distance the image plane can be moved within the allowable circle of confusion is defined as the depth of focus. In CCD cameras the CCD imagers are precisely fixed to the color separation prism and it becomes critical that the optical system be compensated such that the R, G, and B images are well within the depth of focus, a difficult requirement that is now imposed on lenses intended to be used with CCDs.

The smaller the f-number of a lens, the shallower the depth of focus and therefore the smaller the depth of field. In other words the depth of field can be controlled by adjusting the iris (aperture) of the lens in conjunction with a suitable ND filter to provide correct exposure. The depth of field characteristically becomes greater both as the focal length of the lens becomes shorter, and as the distance between the lens and the subject becomes longer.

Ramping

Ramping or f-drop is the term applied to the drop in sensitivity (increased f-number) typically experienced in a zoom lens at the upper end of its focal length range. It is common for lens manufacturers to allow a certain amount of ramping in portable lenses to reduce the size and weight of a zoom lens. Studio lenses, however, frequently employ a large enough focusing group to avoid the ramping effect.

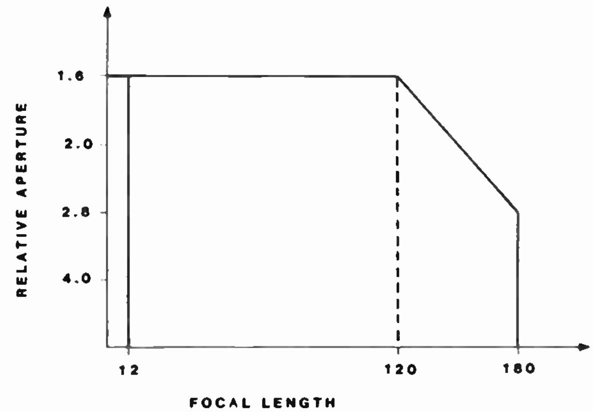


Figure B-7. Lens ramping.

When the ramping effect is present in a lens, the iris must be readjusted to maintain full video level at the extreme zoom positions. In the example of Fig. B-7, the lens has a relative aperture of f/1.6 through a zoom range of 12 mm to 120 mm, but ramps down to f/2.8 at its maximum focal length of 180 mm.

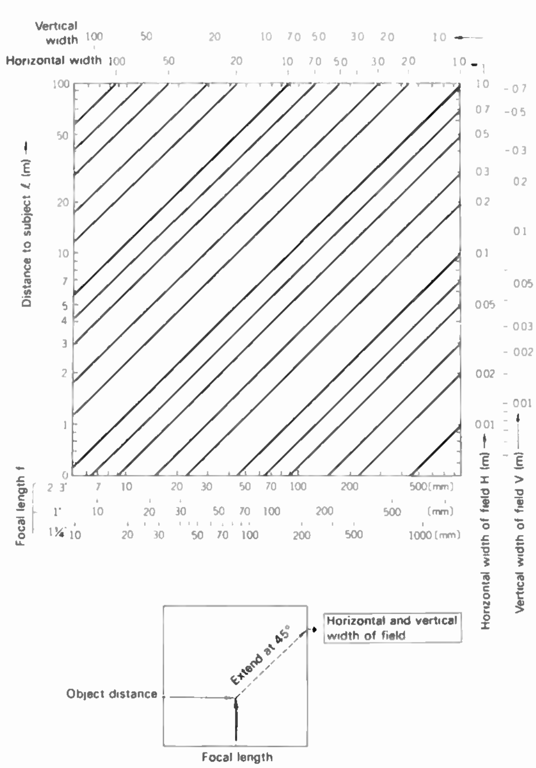


Figure B-6. Depth of focus and depth of field.

Modulation Transfer Function (MTF)

MTF defines the resolution capability of a lens. MTF is typically measured by using a black to white sine wave pattern. The peak-to-peak amplitude of a low frequency sine wave (0.5 MHz) is used as the 100% reference; the peak-to-peak amplitude for other frequencies defines the MTF of the lens for those frequencies. Most lenses exhibit their best MTF when operated in the middle of their aperture range. For example, a lens with an aperture range of f/1.6 to f/16 will probably gives its best MTF at about f/4.0. If the lens is further stopped down towards f/16, some small reduction in modulation depth may result due to the light being refracted as it passes through such a small aperture. However, as the lens is opened towards its maximum aperture of f/1.6, a more severe loss of modulation depth is likely to occur as the spherical aberrations become more significant.

In addition, the MTF of the lens tends to be best at the center of the picture and fall off toward the edges of the picture. Fig. B-8 illustrates the MTF characteristics of a typical zoom lens. Note that the MTF becomes worse as the iris is opened. This difference is more exaggerated in the corners, which illustrates the importance of considering the f-stop when evaluating the performance of a lens.

Uniformity of Field (Vignetting)

This is a measure of the response to a white field over the full picture area. The typical uniformity of

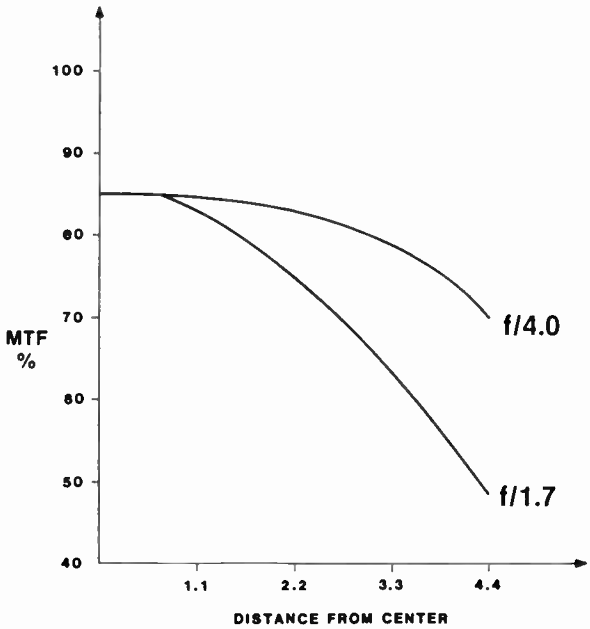


Figure B-8. Typical MTF characteristic for ⅔ inch lens (average for all focal lengths).

field characteristic of a zoom lens is shown in Fig. B-9, which illustrates that the response is usually greatest in the center and least in the corners. The non-uniformity in the corners becomes more pronounced at wide aperture settings and varies with focal length. High quality TV cameras frequently contain a white shading circuit that increases gain toward the edges and provides first order correction for this effect.

Longitudinal Chromatic Aberrations and Lateral Chromatic Aberrations

The refraction of optical glass varies with the wavelength of the light. This causes the lens to form different images for different colors. See Fig. B-10. Chromatic aberration is classified in two types: longitudinal chromatic aberration, where the position of the image plane varies according to color, and lateral chromatic aberration, in which the size of the image on a given image plane changes according to color.

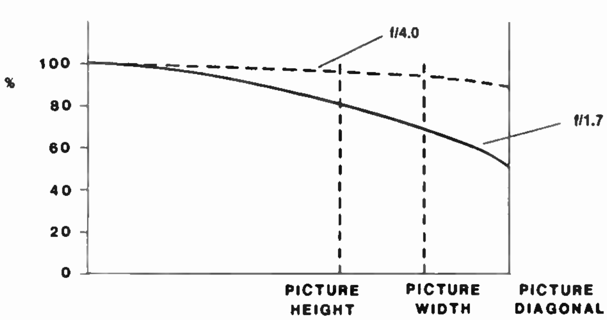


Figure B-9. Uniformity of field.

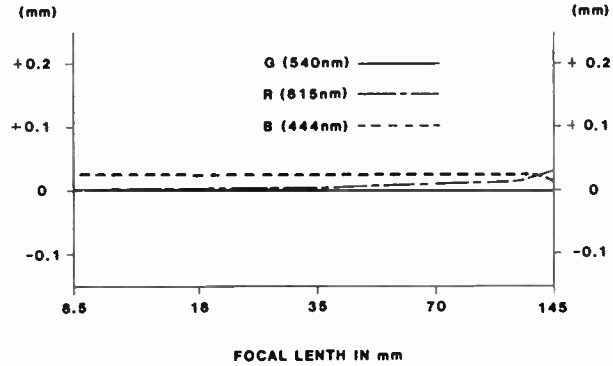


Figure B-10. Longitudinal chromatic aberration.

Although the lens designer makes a great effort to achieve consistency, the focal length of a zoom lens tends to vary slightly with color throughout its zoom range. This means that there will be some small differences between color channels, in both focus and image size, as the zoom is operated.

Geometric Distortion

All lenses exhibit some degree of geometric distortion, a symmetrical error (shown in Fig. B-11) affecting

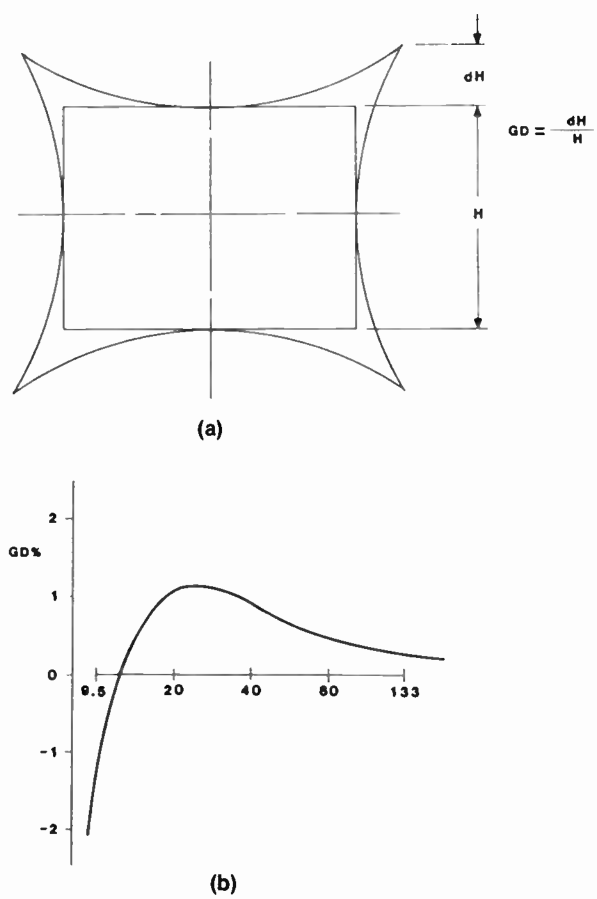


Figure B-11. Geometric distortion.

the red, green, and blue images equally. In the case of the zoom lens, the distortion varies with focal length such that the negative (barrel) distortion occurs at short focal lengths, and positive (pincushion) distortion occurs at long focal lengths. For camera measurements and adjustment, it is advisable to use the middle of the zoom range where this error is minimized.

Minimum Object Distance (MOD)

MOD is the closest focussing distance of a lens. Wide angle lenses generally have the shortest MOD, and telephoto lenses the longest. When a lens is chosen for a specific application it is critical to insure that the MOD is adequate.

APPENDIX C—COLOR TEMPERATURE

The illumination from an incandescent lighting fixture is frequently described as a warm source of light by cinematographers, and indeed, when analyzed with a spectra device that measures the relative intensity of the color components, red wavelengths described as a warm color, predominate. Daylight illumination on the other hand is frequently described as being cool and the same spectral analysis shows blue wavelengths, usually described as a cool color, to predominate. The concept of color temperature makes it possible to describe the character of the illumination in a much more precise manner than the vague descriptions above. The concept of color temperature also makes it possible to easily determine the proper color correction filters to use with a specific source of illumination.

Color temperature is defined as the visible spectrum emitted by a black body that has been heated to specific temperature in degrees Kelvin. The color of an illuminant is determined only by its temperature and is almost the same for all substances at a given temperature. For example a quartz halogen studio light is said to have a color temperature of 3,200°, since the visible spectrum of the studio light closely resembles the spectrum emitted by a black body heated to 3,200° Kelvin. Daylight is said to have a nominal color temperature of 5,600°, since its spectrum provides a close match to the spectrum of a black body heated to a temperature of 5,600° Kelvin.

The light emitted by a black body that has been heated, is a smooth, continuous spectrum that covers the complete visible spectrum. Lighting sources such as metal vapor lights and fluorescent lights, typically have a discontinuous spectrum of strong spectral lines. Nevertheless, a color temperature designation is frequently assigned to such sources, to give an overall indication of the character of the light as perceived by the eye. These lighting sources, frequently used in stadiums and arenas because of their efficiency, must be used with great caution to achieve acceptable color reproduction with a color TV system.

The relative energy versus wavelength at various color temperatures is given in Fig. C-1. For example, consider an illuminant of approximately 2,900° Kelvin

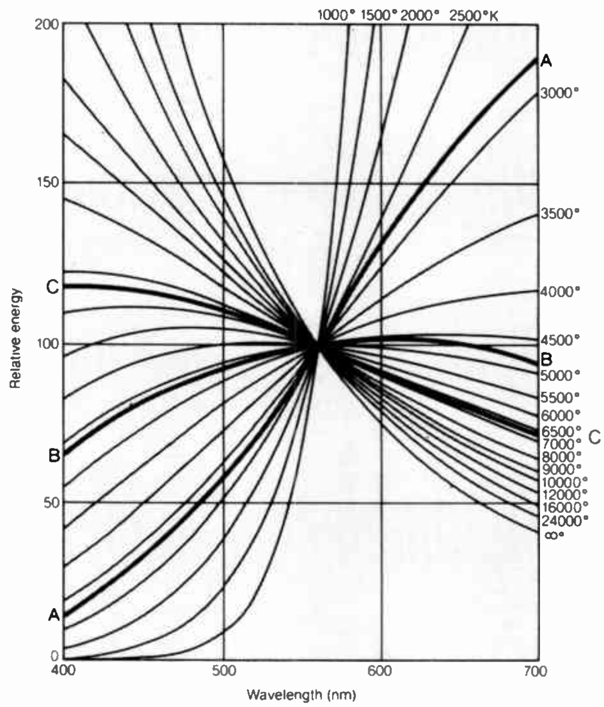


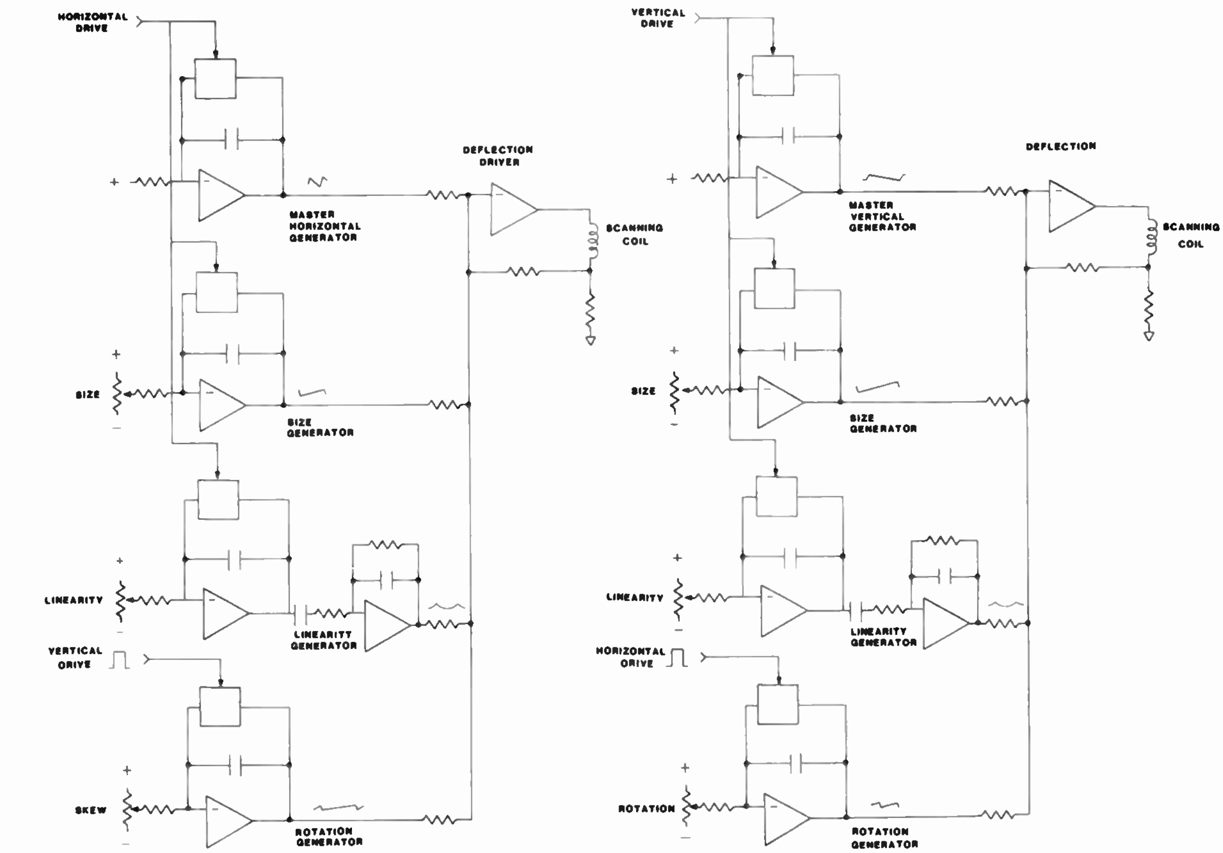
Figure C-1. Spectral distribution at each color temperature.

(curve A). The magnitude of the long wavelengths which correspond to red (600–700 nm wavelength) are large while the magnitude of the of short wavelengths coresponding to blue (400–500 nm wavelength) are very small. This relationship reverses for high color temperatures.

APPENDIX D—SPECIAL CONSIDERATIONS APPLICABLE TO PICK-UP TUBE CAMERAS

Camera Pre-Amplifier

The output signal from a Plumbicon® or Saticon® pick-up tube is very fragile. The output impedance is very high and the signal level is very low. Capacitive loading on a source with high output impedance causes a frequency roll-off that is directly proportional to the capacitance. For this reason the first stage of the camera pre-amplifier, a low input capacitance FET, is mounted in the deflection yoke, within inches of the pick-up tube to minimize capacitance and pick-up of noise. The magnitude of the output signal from the pick-up tube is generally 180 to 200 nanoamperes for 100 units of video from the green tube in ⅔-inch tubes and about 300 nanoamperes for the 25 mm and 30 mm tubes used in studio cameras. The signal levels from the red and blue tubes are even lower than from the green tube. The pre-amp circuitry forms an operational amplifier with a very high loop gain that accurately converts the signal current to a precise output voltage.



(a) SIMPLIFIED HORIZONTAL DEFLECTION BLOCK DIAGRAM

(b) SIMPLIFIED VERTICAL DEFLECTION BLOCK DIAGRAM

Figure D-1. Simplified block diagrams.

The Deflection System

The deflection system is employed in a tube camera to provide the necessary scanning waveform to the yoke scanning coils. A simplified deflection system is shown in Fig. D-1.

In this case, the vertical and horizontal master sawtooth generators produce the main scanning waveforms; dc voltage controlled sawtooth generators provide the registration functions of size, skew, and rotation; and parabolic generators provide linearity adjustment.

Usually the drivers function as current sources with sufficient feedback to insure the scanning current through the deflection coil is a linear function of the input voltage waveform. In Fig. D-1, the deflection current is given by:

$$I(hor) = \frac{V(master) + V(size) + V(linearity) + V(skew)}{R}$$

$$I(vert) = \frac{V(master) + V(size) + V(linearity) + V(skew)}{R}$$

APPENDIX E—11-STEP GRAY SCALE

An 11-step gray scale chart is highly recommended for use in adjusting a camera. The 11-step chart incorporates a true gamma of 0.45 as well as a 40:1 contrast ratio. It is highly recommended not to use the traditional EIA 9-step logarithmic gray scale chart for camera verification and adjustment for the following reasons:

1. All the instruction manuals provided with current cameras provide waveforms and voltages appropriate for an 11-step chart. It is extremely difficult to relate these waveforms and voltages to a nine-step chart.
2. The EIA chart uses logarithmic steps that do not mathematically fit the gamma curve of the camera.
3. The nine-step chart covers only a 20:1 contrast ratio, inadequate to test a modern TV camera that is capable of at least a 40:1 contrast ratio.

###

The author wishes to express his sincere thanks to Mr. Larry Thorpe and all the members of the Sony team that have contributed valuable material and assistance in preparing this chapter.

REFERENCES

1. Benson, B. Whitaker, J. *TV Engineering*, McGraw-Hill, 1986.
2. *NAB Engineering Handbook*, 7th Edition, Washington, DC: National Association of Broadcasters, 1985.
3. *TV Optics—The Canon Guidebook of Optics for Television Systems*, Kawasaki, Japan: Broadcast Equipment Center, Canon, Inc., 1986. (Contact a local office of Canon USA for availability.)

5.3 Video Signal Switching, Timing, and Distribution

Jay Kuca
Grass Valley Group, Grass Valley, California

INTRODUCTION

The video signal contains a large amount of information. In addition to basic picture parameters such as brightness, hue, and saturation, the video signal includes horizontal, vertical, and color timing information. And recently, more and more ancillary information has been added, primarily in the vertical blanking interval (VBI): vertical interval test signals (VITS), vertical interval reference signal (VIRS), teletext, closed captions, time of day and time code information, and source identification data.

Technological advances in the last thirty years have made it possible to pack all of this information into substantially the same bandwidth originally allocated for the monochrome television picture alone. But along with the increased information density has come a need to maintain a very high level of performance along the entire signal path. Therefore, careful attention must be paid to the design of any modern video system if optimum results are to be achieved.

Whenever video signals are distributed, levels, frequency and phase response, and other transmission parameters must be held to extremely tight tolerances. Whenever video signals are switched, the switching system must be designed so that the integrity of the timing as well as the transmission parameters is observed. And whenever video signals are combined, whether by mixing, keying, or special effects generation, the relative timing relationships between the signals must be carefully maintained.

THE VIDEO SIGNAL

Fig. 1 is an oscillogram of a typical video signal. The peak-to-peak amplitude of the signal is 140 IRE units, or 1 volt. This is the standard level for signal distribution in all professional television facilities. The signal is

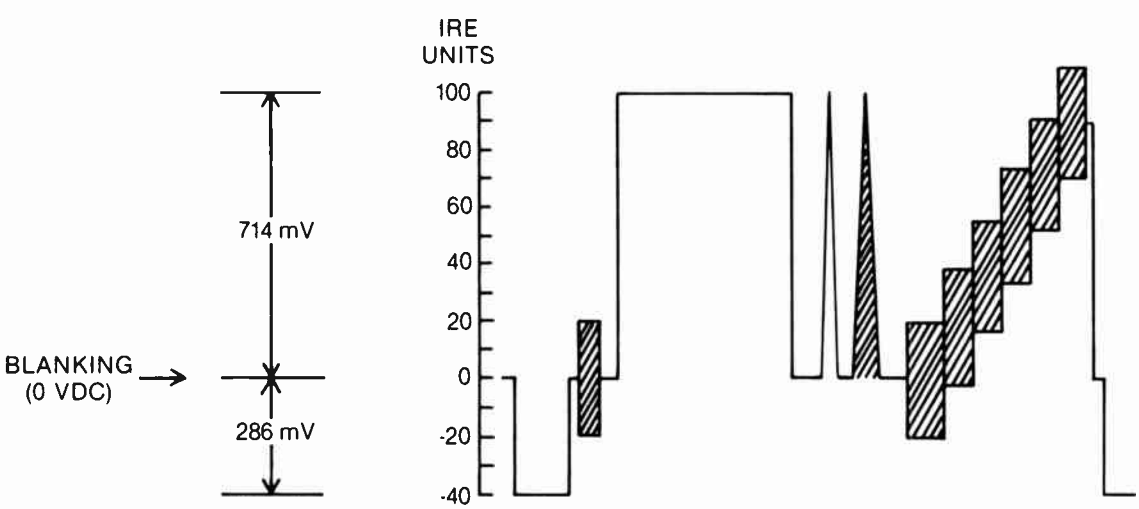


Figure 1. Oscillogram of a typical video signal.

comprised of + 100 IRE units (714 mV) of picture and − 40 IRE units (286 mV) of sync.

Even though video is an AC signal, proper operation of most equipment depends upon a fixed DC component. By convention, blanking (0 IRE) is assumed to be at ground potential (0 VDC). This reference is lost, of course, in portions of the system which are AC coupled, but when the signal is clamped or DC restored, standard practice dictates re-establishing blanking at ground level.

The bandwidth of a studio quality NTSC video system is determined primarily by the state of the art and economic factors. And while it can be argued that the transmission system cuts off any information above 4.2 MHz, it is good engineering practice to strive for response which is flat to 8 MHz to 10 MHz in the studio. Future high-definition or enhanced-definition services may require even wider bandwidths.

Timing integrity is ensured through the use of a master sync pulse generator. In nearly all installations, timing for all signals in the plant is derived from this single reference. In most cases, the master sync generator is driven from a high stability crystal oscillator. Occasionally, a precision rubidium or cesium "atomic" reference oscillator is used. Fig. 2 illustrates the relationships between the horizontal, vertical, and color timing signals.

INTERCONNECTION CONVENTIONS

All studio equipment is designed to be interconnected with coaxial cable with a nominal characteristic impedance of 75 ohms. In the simplest case, a point-to-point connection between two pieces of equipment, a continuous length of cable is driven from a 75 ohm source and terminated in a 75 ohm load. See Fig. 3.

When it is necessary to distribute the signal from a single source to more than one destination, two possible approaches exist. The first is entirely passive in nature: one end of the cable is driven from the 75 ohm source and, instead of terminating the far end in a 75 ohm load, a "loop-through" connection (high impedance bridging) is made to the first piece of equipment. The loop-through is carried on to the next piece of equipment, and so on. A 75 ohm termination is placed on the loop-through connection on the last piece of equipment. This approach, shown in Fig. 4, will work well provided the loop-through inputs are high impedance compensated and the cable lengths are kept relatively short. The number of loop-throughs should also be kept as small as possible. If this is not done, frequency response errors and, in severe cases, signal reflections are likely.

The second and most desirable approach is to distribute video signals using an "active" device such as a distribution amplifier or routing switcher. This effectively results in the equipment being interconnected in the point-to-point manner described above.

Nearly all modern equipment uses BNC connectors for video input and output connections. Some older equipment still in service may use "UHF" (PL-259/SO-239) connectors. Occasionally, space limitations dictate the use of a subminiature connector such as a "BSM" series. However, the BNC is by far the most common connector in use.

A wide variety of cable types are available which are suitable for high quality video interconnect. One of the most popular is Beldén 8281, a double-shielded 75 ohm precision coaxial cable designed specifically for video use. 8281 provides a good balance between loss and physical size. Its double shielding helps reduce the likelihood of stray signal pickup. Where space limitations, increased flexibility, cost, or other factors suggest the use of a smaller diameter cable than 8281, an "RG/59-type" cable is the usual choice.

POTENTIAL SYSTEM PROBLEMS

Most problems in television system design are a consequence of the fact that individual pieces of equipment must be interconnected by coaxial cable. All cable exhibits high-frequency losses which may result in the impairment of picture detail, especially on long runs. Stray signals, such as 60 Hz hum or noise may be coupled into cables. And since signals take a finite amount of time to travel through cables, time delays must be considered.

Fortunately, most of the problems are not new, and solutions have been developed over the years. For example, equalizers which compensate for high-frequency cable losses can be found in a variety of modern video equipment. Quite often, a differential amplifier is used in the input stage to cancel the effects of common mode hum and noise; in some equipment, a clamp stage may be provided for hum suppression. Modern timing concepts have led to the development of a wealth of useful tools: pulse and subcarrier regenerators, slave sync generators, remote phasing sync generators, isophasing amplifiers, and frame synchronizers.

The tools exist to solve nearly any problem in video system design. The key is to anticipate the problems in new designs and to diagnose the problems as they develop in existing systems.

VIDEO SYSTEM DESIGN

Video system designs can be classified as falling into one of two broad categories: *hardwired* or *configurable*. A hardwired system is one which is dedicated to a single function: a small ENG edit bay, for example. A configurable system is one which is designed to perform multiple functions, perhaps simultaneously: for example, a centralized VTR room which is used for on-air tape playback, commercial production, and dubbing.

Most systems were hardwired prior to the advent of video tape recorders because program production and presentation was done "live," in real time. VTRs made off-line production and program assembly possible.

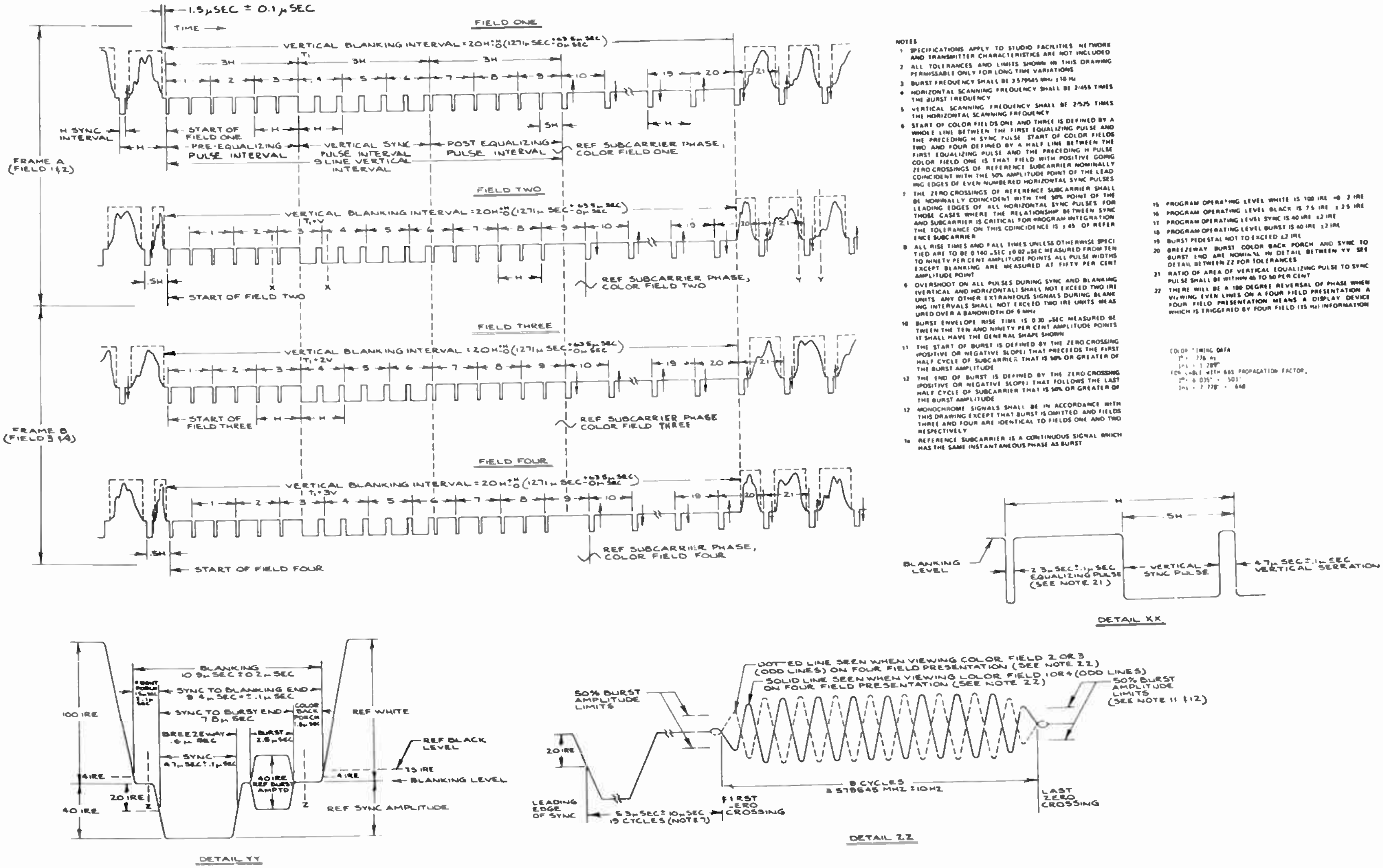


Figure 2. NTSC timing parameters. (Courtesy of The Grass Valley Group) An updated version of the EIA Tentative Standard RS 170 A.

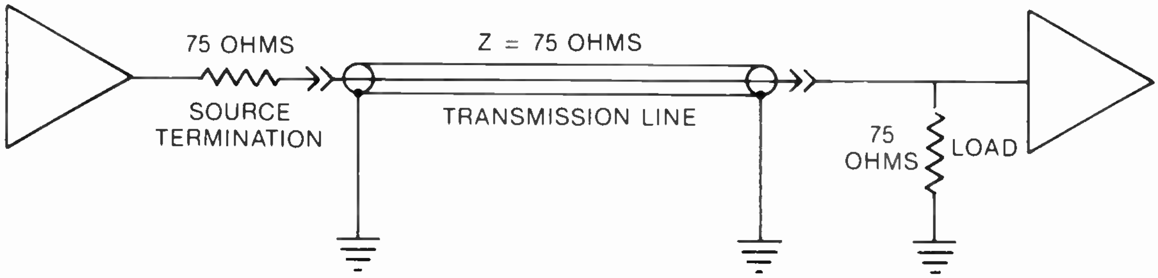


Figure 3. Point-to-point video connection.

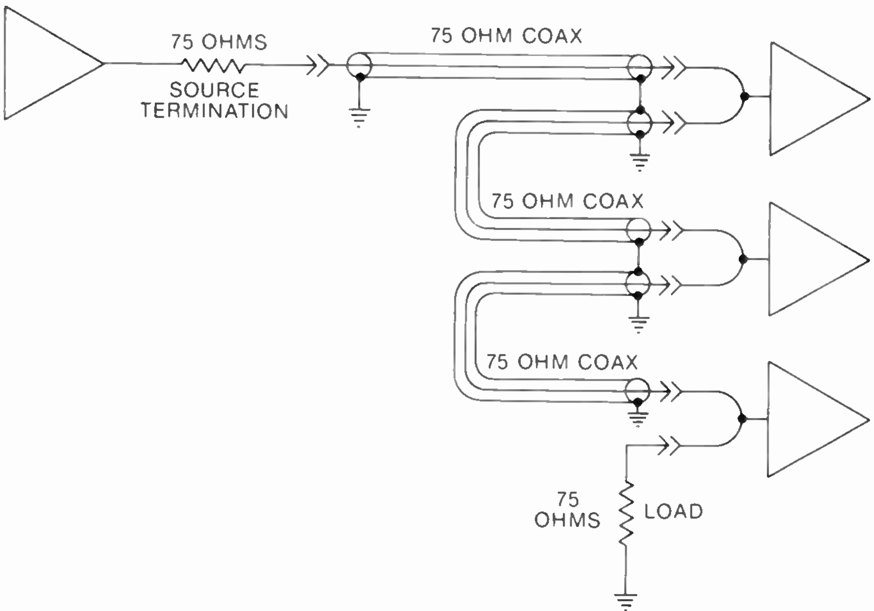


Figure 4. Loop-through video connection.

But the high cost of the early recorders made it necessary to design the system to ensure optimum equipment utilization. This was accomplished by making the system configurable.

The hardwired approach is fairly self-explanatory. Signals are routed point-to-point, using distribution amplifiers where a single source must feed more than one destination. Quite often, patch panels with "normalling" (through-connection) jacks are interspersed throughout the system, thus providing a certain degree of configurability.

A truly configurable system is generally built around a routing switcher. In the limit, all signal sources in the plant appear as inputs to the switcher and all signal destinations in the plant are driven from a routing switcher output bus. Such a high degree of flexibility is rarely required, however. As a result, most video systems are partially hardwired and partially switched.

Source timing is an important consideration in any video system. As a general rule, the more flexible the system is, the more difficult timing becomes. In a hardwired system, for example, timing signals can be hardwired along with the video signals. But in a routing switcher based system, configurations may be possible which create timing problems. The most common of these is the problem of studio switcher delay.

It is standard practice to designate a point in the system as a "zero timing point." The timing of all sources in the plant is adjusted so that they are in time at this point, which is usually a patch panel or routing switcher output. Timing will be correct so long as the path from the source to that point is fixed. But when the system is flexible enough to allow the source to go through one or more studio switchers enroute to the zero timing point, the source timing will be upset by the cumulative delay through the studio(s). If one attempts to treat that source as fully timed (by mixing or wiping between it and another source, for example), the problem will be immediately apparent.

One way around the studio delay problem is to backtime all sources feeding the studio by an amount equal to the studio delay. This can be done quite easily if each source is equipped with its own sync generator. It is important to note, however, that if the source must feed other points in the system in addition to the studio, it may not be properly timed for those points.

The best way around the studio delay problem is to put frame synchronizers on the outputs of all studios. In that way, timing is guaranteed, regardless of how complex the path becomes. Where such an approach is not economically feasible, it is essential to analyze and fully understand the limitations of the system.

SYSTEM DESIGN APPROACH

More often than not, video system design is approached in a haphazard rather than systematic way. This is often due to the pressures of attempting to provide the latest technology to the end user while keeping a station on the air. Even though it is a rare occasion that one is able to start from scratch, any system design project stands to benefit from up-front planning.

The task begins with a survey of the facilities. What signal sources are there? What destinations must they feed? What is the physical layout of the plant? What is its size and where will the equipment be located? What allowances have to be made for future growth?

The next step is to analyze user needs. Each operating area should be examined in light of what signals are needed, how frequently they will be used, and what level of quality is necessary. It is important to be realistic when asking these questions. For example, it might be nice to have a seldom used test signal available on demand in a particular operating area; but is it really necessary?

Frequency of use will help to determine whether a signal should be hardwired, switched, or patchable. Frequently used signals should be hardwired in order to eliminate possible operational errors. Signals which are used occasionally can be switched, if necessary. Seldom used signals can be handled with manual patching.

The desired signal quality is an important factor. Does the signal have to be timed? What about equalization, clamping, or other forms of signal conditioning? Unnecessary expense can be avoided if the answers to these questions are carefully considered.

When the research is complete, the planning phase can begin. During the planning phase, the details of the design will begin to take shape. The physical realities of the system will come to light: equipment layout, interconnect, electrical power requirements, etc. Notes gathered during the research phase will give way to sketches, which will eventually give way to final drawings.

The importance of accurate system documentation cannot be overemphasized. Good system-level drawings will prove to be invaluable when troubleshooting or when making changes. And the documentation will be much easier to follow if wire and rack numbers are used.

After the final "as built" drawings are completed, it is a good idea to produce a simplified drawing of the signal flow through each operational area. These can be posted, perhaps along with information on how to deal with "emergency" situations.

VIDEO DISTRIBUTION AMPLIFIERS

The distribution amplifier (DA) is one of the most elementary building blocks in a video system. Its primary function is to allow a single 75 ohm video source to feed multiple 75 ohm loads with no apparent loss. Quite often, other functions such as cable equalization are included in the DA's circuitry.

Even though DA functions and designs vary, it is quite easy to construct a generalized model. As shown in Fig. 5, the typical DA consists of an input amplifier followed by a conditioning stage (gain, equalization, delay, etc.), followed by an output amplifier.

The input amplifier is generally designed with a rather high input impedance, on the order of several tens of kilohms. This allows the driving signal to be looped through (bridged across) the input without being

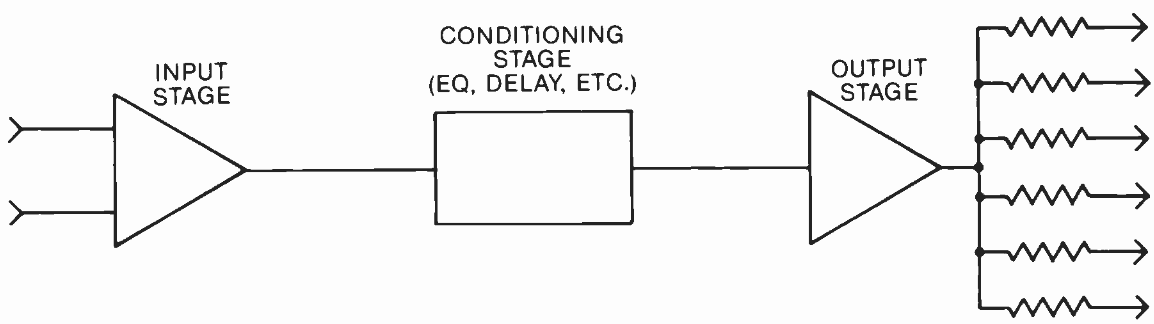


Figure 5. Simplified block diagram of a video DA.

loaded down. Quite often, a complex network is used at the input connection to cancel out any reactive components which might affect the loop-through. The input stage may be either single ended or differential, the latter being more common on high quality DAs.

A differential input stage can be very useful in combating one of the most serious problems encountered in unbalanced signal transmission: common mode hum and noise. In larger facilities, common mode hum due to ground loops may run as high as several volts. Even in small plants with excellent ground systems, it is possible to have several tens of millivolts of difference in ground potential between different pieces of equipment. A differential amplifier will cancel any common mode signals while amplifying the (desired) differential mode signal. Common mode rejection ratios (CMRR) of 50 dB to 60 dB are typical of most modern differential input DAs.

The conditioning stage may be a wideband gain stage, an active equalizer stage, or a delay amplifier, depending upon what additional function the DA performs. If distribution is the only function, the conditioning stage will be omitted.

The output stage is an amplifier with very low output impedance. It is followed by source termination resistors which ensure that the cable "sees" 75 ohms at the sending end as well as the load end. Output-to-output isolation is guaranteed by the low impedance of the output stage.

Cable equalization is often required in video systems. This is due to the high frequency losses in coaxial cable. Belden 8281 has approximately 0.8 dB of loss per 100 feet at 10 MHz. Since runs of 100 feet or more are not uncommon in most facilities, achieving "flat" system response makes equalization of these losses necessary. The approach taken by most manufacturers is to construct an equalizer network whose response very closely approximates the inverse of the cable loss curve. When this equalization is applied to the cable/DA system, the result is flat overall response.

Cable equalizers must be designed for a specific type and length of cable. When a variable equalization control is provided, it usually serves to adjust the amount of eq applied. Because cable loss is a complex function, the equalization curve is only "right" for one cable length. Slight response errors will be noted at other length settings.

Cable losses are often equalized at the load end of the cable in short or moderate length runs. The reason for this is simple: if a "flat" signal is applied to the input of the amplifier, the high frequency boost of the equalizer stage will be added to the signal, and the output stage will have to track these larger than normal amplitude high frequencies. In cases where the amount of equalization is more than a few dB, the output stage will be unable to deliver the necessary current into its low impedance load. If, however, the DA is located at the load end of the cable, the cable losses have attenuated the high frequencies at the input of the amplifier. The high frequency boost of the equalizer stage restores the flat response characteristic to the signal, and the output amplifier does not have to deal with larger than normal swings at high frequencies.

In cases of extremely long cable runs, cable losses may be so great that post-equalizing alone would significantly raise the noise floor of the signal due to the large amplification needed. In this situation, a moderate amount of pre-equalization at the sending end is applied in conjunction with post-equalization at the receiving end to maintain an adequate signal-to-noise ratio overall.

Another function which is often required in video systems is clamping. Clamping is used primarily to reduce low frequency hum and tilt, but may also be used to restore the DC component of a video signal which has passed through one or more AC coupled devices. Clamps operate on a line-by-line basis and, as such, are fast acting. They can be triggered by noise impulses and should therefore not be used on noisy signals.

DC restoration may be used in lieu of clamping to reduce the DC signal bounce due to multiple AC couplings. Because of the slow time constants employed in DC restorers, they are not affected by noise; but by the same token, they are not useful in eliminating hum.

Delay DAs are often used in place of coaxial cable or other passive delay elements. Delays of up to a microsecond or more are possible on modules no larger than a conventional DA. Many delay DAs include circuits to compensate for delay line response errors; they produce better results than a passive delay line and DA in combination. Some delay DAs also include other functions such as cable equalization.

PULSE AND SUBCARRIER DISTRIBUTION AMPLIFIERS

Despite the increasing use of color black (also known as "black burst") as a synchronization reference signal, separate pulses and subcarrier are used in many systems as timing references. Pulses (sync, blanking, H-drive, V-drive) are generally distributed as 2 to 4 volt negative-going signals. Subcarrier is generally distributed as a 1 to 2 volt peak-to-peak signal.

Most video DAs will not accommodate signals larger than 2 volts peak-to-peak (p-p). So, while most video DAs would be suitable for subcarrier distribution, the typical video DA will not handle the aforementioned pulse signals.

DAs designed specifically for pulse distribution are optimized for large negative signal swings. They will not function well, if at all, as video amplifiers.

Pulse DAs fall into one of two general categories: regenerative and linear. Regenerative DAs include circuitry for pulse regeneration; typically edge detectors driving one-shots. Timing of the one-shots is usually adjustable, allowing the relative delay of the amplifier to be varied. Delays of up to several microseconds are common. Due to the inherent delay of the pulse regeneration circuitry, these amplifiers always exhibit some value of minimum delay.

Regenerative subcarrier DAs are also available. These amplifiers generally feature adjustable circuits to shift the phase of the subcarrier. The adjustment range is usually slightly more than 360 degrees.

When pulses and subcarrier are distributed through regenerative DAs, there is a good chance that the phase relation between the color subcarrier and horizontal sync pulses (SC/H) will not be maintained throughout the system. This is because it is possible to adjust sync and subcarrier timing independently. For this reason, linear (nonregenerative) pulse and subcarrier DAs should be used in SC/H phased installations. Specially designed linear pulse DAs are available, and most high quality video DAs can be used as linear subcarrier DAs.

DISTRIBUTION AMPLIFIER SELECTION

As a signal traverses a typical path through a typical system, it may pass through ten or more DAs. Therefore, the performance of a single amplifier must be, for many parameters, an order of magnitude better than the system specification. Table 1 lists the performance specifications of a representative unit, the Grass Valley Group 8502.

TABLE 1

Typical video DA performance specifications.

Frequency Response	± 0.025 dB to 5 MHz − 0.2 dB at 10 MHz
Differential Phase	<0.1 degree
Differential Gain	<0.1%
T-Pulse to Bar	<1.0%
Tilt	<0.5%
Chrominance/Luiminance Delay	<10 ns
Hum and Noise	>70 dB below 1 volt p-p
Common Mode Rejection	>60 dB
Input Return Loss	>40 dB to 5 MHz
Output Return Loss	>40 dB to 5 MHz
Output Isolation	>40 dB to 5 MHz

Selecting the right DA for a particular application involves determining the desired level of performance and the features and functions required. Since most manufacturers offer a range of models which all fit into the same mounting tray, it is wise to look at the sum total of all requirements and then choose the manufacturer who can best accommodate those requirements.

DAs fall into the "install it and forget it" category: low maintenance and long service are important factors. When selecting distribution amplifiers, consider the quality of the components which the manufacturer has used: will they hold up for ten or fifteen years? Is the design a conservative one, or are devices being run at their limits? Thermal stress will not only lead to equipment failure, it will cause drift as well.

VIDEO SIGNAL SWITCHING

Video signals can be switched by a variety of methods. Simple monitor selectors may be nothing more than passive mechanical switches. In some cases, electromechanical devices such as relays are employed. Most modern switching systems employ solid state switching devices.

Fig. 6 is a simplified block diagram of a video switching system. Input amplifiers buffer the input signals and provide drive to the switching elements (crosspoints). The crosspoints are bussed in horizontal rows (output buses) to an output amplifier. If the switcher has more than one output bus, the buffered input signals are bussed in vertical columns (input buses) which run across the output rows.

Not shown in the block diagram is the crosspoint control logic. In single-bus switchers with ten inputs or less, control is generally accomplished using a single wire per crosspoint. In single-bus switchers with more than ten inputs, an encoding scheme such as BCD is generally used to reduce the number of wires required. Larger multiple bus switchers generally use some sort of serial control scheme.

Single-Bus Switchers

Single-bus switchers run the gamut from simple to elegant. For noncritical applications such as picture monitor input selection, a mechanical switcher may be adequate. Most mechanical designs are entirely passive; that is, they have no input or output amplifiers. As such, they have terminating inputs and require dedicated (nonloop-through) input feeds.

A slightly more sophisticated design adds an output amplifier only. Switching is still done using mechanical switches, but the bus is high impedance to allow looping inputs. This approach makes it possible to "stack" units to provide an economical multiple input and multiple output configuration. In terms of performance, this type of switcher is no better than the

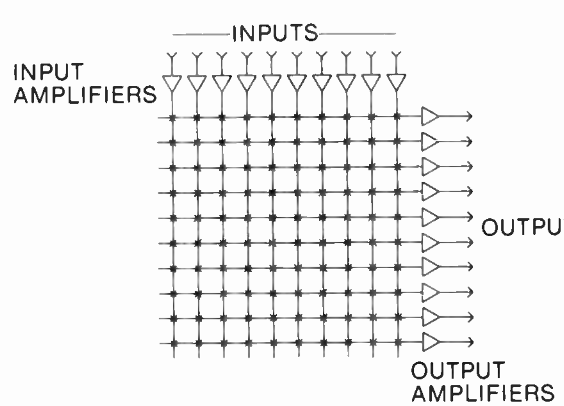


Figure 6. Simplified block diagram of a video switching system.

passive type. If the input loop-throughs are not compensated, its performance may actually be worse.

Any single-bus switcher which is in a program path should be designed more along the lines of Fig. 6. In other words, it should have input buffer amplifiers, solid state crosspoints, and an output amplifier. If it is being used for "live" switching, the buffer amplifiers should clamp or DC restore the input signals to provide bounce-free switching. If the unit is used in an actual program loop, the design should permit servicing with minimum disruption; that is, all active circuitry should be mounted on plug-in printed circuit boards.

Single-bus switchers can be combined to make multiple-bus switchers, but in most cases the result will cost more and lack the flexibility and performance of a true multiple-bus design.

Multiple-Bus Switchers

There are two basic architectures for multiple-bus switchers. An output-oriented switcher is one with all buses physically separated in the switcher electronics. A matrix switcher is the other in which the circuitry for several output buses is packaged together.

Output-oriented switchers are generally larger and more expensive than comparable matrix switchers. The reason for this is that there is more duplication of circuitry and more interconnections required. The chief advantage of an output-oriented switcher over a matrix switcher is that circuit cards can be removed from the output oriented switcher without disturbing multiple signal paths. This is an important consideration in large facilities such as network switching centers.

Matrix switchers are smaller and less expensive than comparable output oriented switchers. They offer several other advantages as well: reduced power consumption, slightly higher reliability due to fewer mechanical interconnections, and shorter signal path. Matrix switchers are generally designed around a relatively small square or rectangular building block, making input and output expansion in bite-sized increments a simple matter.

Input signal distribution in both types of switchers can be handled using either looping techniques or distribution amplifiers. The cost objectives of most matrix designs suggests the use of looping, while the large physical size of the output-oriented switcher makes the distribution amplifier approach almost a necessity.

As is the case with single-bus switchers, if "live" switches are made on a multiple-bus switcher, the input signals must be clamped or DC restored to ensure transient-free switches. This signal conditioning may be done external to the switching system, but better overall performance will result if it is done in the switching system. The reason for this is that a common precision DC reference can be used for all inputs if the clamping or DC restoration is done inside the system.

Switch timing is also important if glitches are to be avoided. Most multiple-bus switchers have a sync input which is used as a reference for a master trigger pulse generator. The trigger is set to occur during the vertical interval, thus ensuring glitch-free switching between synchronous sources.

Timing "scatter" is an important parameter for any switching system which will be used in a studio environment. It is the measure of the delay difference between paths in the switcher. In a system which uses looping techniques to distribute the input signals, there will be a fixed amount of delay between output bus groups. But on a given bus, there should be very little path length difference from input to input; less than ± 1 degree of subcarrier is a typical figure. In a system which uses DAs to distribute the input signals, it should be possible to achieve the same timing accuracy from any input to any output.

With respect to the use of routing switchers in video systems, two basic schools of thought prevail: the master grid approach, and the cellular approach. With the master grid approach, one large switching system is used with all switched sources in the plant appearing as inputs. This provides maximum flexibility, but at considerable expense. All sources are not needed at all destinations, so many crosspoints in the system are "wasted."

In the cellular approach, the plant is subdivided into functional areas (studios, master control, VTR/telecine pool, news editing, etc.) and a small switching system placed in each area. Each system has only those area sources it requires plus a few "master" sources common to the other areas (color black, color bars, etc.). To accommodate occasional special setups, tie lines may be run between the systems.

The cellular approach can be considerably more cost effective than the master grid approach. But the incidence of special setups must be carefully considered. Since multiple switching operations may be required in the case of specials, the chance for error increases significantly. This must be weighed against the additional cost of the larger matrix required for a master grid system.

Switching System Control

Switching system control has advanced considerably over the past ten years or so. Most multiple bus switchers now have microprocessor-based control systems. Control panels communicate with the switcher over twisted pairs or coaxial. Sources may be selected using their familiar names (e.g., "BLACK," "BARS," or "VTR 5") instead of input numbers. Simultaneous input/output selections, or "salvo" switches, can be easily made.

A recent development in control system capability has been made possible through the availability of a semiconductor device known as an electrically erasable programmable read-only memory (EEPROM). This device can be programmed, erased, and reprogrammed just like an ordinary PROM. But, unlike a PROM, the EEPROM does not require a special programming device; more importantly, it can be erased electrically.

TABLE 2

Typical video switching system performance specifications.

Frequency Response	±0.1 dB to 5 MHz, −0.5 dB at 8 MHz
Differential Phase	<0.1 degree
Differential Gain	<0.1%
2T-Pulse to Bar	<0.25%
Tilt	<0.5%
Crosstalk	>60 dB to 5 MHz
Hum and Noise	>70 dB below 1 volt p-p
Input Return Loss	>40 dB to 5 MHz
Output Return Loss	>40 dB to 5 MHz
Switching Transients	<30 mV
Timing Scatter	± degree maximum, input to input on any one putput bus

in circuit. System dependent information, such as source name to input lookup tables, can be stored and reprogrammed by the user.

Representative performance specifications which should be expected of a modern routing switcher are listed in Table 2.

VIDEO SIGNAL TIMING

Timing requirements stem from the fact that a video signal is a serial stream of analog information which is used to electronically reconstruct a picture. Horizontal timing information is used to trace out each line of the picture on the face of a display tube in exact synchronism with the camera tube beam which scanned the original scene. Vertical timing information is used to signal the end of each "still" picture in the frame-by-frame reconstruction process. Color timing information in the form of a brief burst of subcarrier on each line is used to ensure that the color at each point on the screen is accurately reproduced.

Timing in the television studio is the responsibility of a ubiquitous black box called a sync generator. In the early days, a monochrome sync generator filled a six-foot equipment rack. Today, color units which are several orders of magnitude more accurate and stable can be packaged on a single printed circuit card.

The size and complexity of the early generators prompted the master generator approach to plant timing. Since it would have been impractical to include a sync generator with each source, a set of signals was defined which would provide the information necessary to generate and process a television picture. The responsibility for generating these signals fell to the master sync generator. These signals were then distributed to each piece of equipment requiring them. They included sync and blanking (used mainly in video circuits), and horizontal and vertical drives (used to drive the sweep circuits in cameras).

The emergence of color meant that the 3.58 MHz subcarrier had to be generated by the same master generator, and distributed throughout the plant. In addition, another pulse called burst flag was needed to gate the color burst. Thus, six separate signals were required just to keep everything locked to the same time reference.

The advent of solid state devices led to a dramatic reduction in the size of the master sync generator. At the same time, integrated circuit logic made it possible to easily derive horizontal and vertical drives and burst flag from sync. Before long, timing input requirements dwindled to sync, blanking, and subcarrier on most equipment.

The capabilities of large scale integrated circuit (LSI) technology has made it possible to reduce most of the circuitry required for pulse generation to a single chip. For this reason, most modern equipment has its own sync generator built in. In addition to making stand-alone operation possible, this reduces the input timing reference requirements to a single-wire signal such as color black.

Fig. 7 is a simplified functional block diagram of a sync generator. A precision oscillator is used to drive a set of complex logic which generates all of the required timing signals. Buffer amplifiers convert the logic levels to the required 4 volt negative-going pulses and 2 volt peak-to-peak subcarrier.

The FCC requires that subcarrier frequency be maintained within 5 Hz of 3.579545 MHz. Therefore, an oven or temperature-compensated crystal oscillator is generally employed.

Most modern sync generators utilize a single LSI chip as the heart of the pulse generation logic. Numerous off-the-shelf and custom parts exist which have 90% or more of the logic required. External parts are often needed to provide additional features such as variable pulse width.

All IC logic families produce pulses with very fast rise times. So, in addition to providing the proper impedance and voltage levels, the output buffer stages often include shaping or filtering circuits to control the rise times of the output pulses. Controlled rise times help to ensure the reliable and predictable operation of pulse detection and regeneration circuits in the equipment being driven by the sync generator.

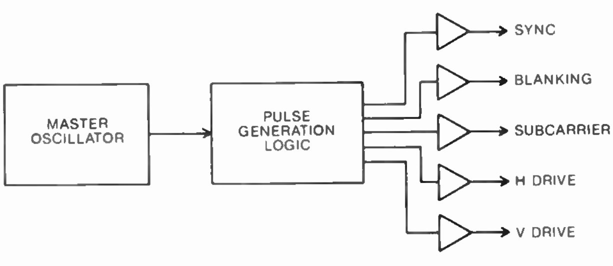


Figure 7. Simplified block diagram of basic sync generator.

GENLOCK

A considerable amount of additional circuitry is required to enable a sync generator to *genlock* to a composite video signal such as color black. Fig. 8 shows the basic functional blocks: an input amplifier, subcarrier regenerator, sync separator, and video presence detector.

The input locking signal may contain common mode hum or noise which could upset the operation of the circuits which follow. For this reason, a differential input amplifier is generally used, followed by a clamp.

Since the locking signal may not always be present, the genlock circuitry includes a video presence detector. The output of this detector is a logic level signal which tells the sync generator whether to select the "free run" or "locked" mode of operation.

The buffered and clamped input signal is low-pass filtered and then fed to a sync separator which puts out sync and burst gate. The burst gate, along with high-pass filtered video, is used to regenerate the 3.58 MHz subcarrier impressed on the input locking signal. The separated sync and subcarrier drive phase-locked loops in the generator. Adjustments on these locking circuits permit the generator phase to be varied with respect to the phase of the input signal.

The generator may be genlocked to the input signal in one of two basic ways: by independent locking of sync and subcarrier, or by dependent locking of sync and subcarrier. With the independent method, the generator's sync and subcarrier locking circuits are continuously referenced to the input signal's sync and subcarrier. With the dependent method, lock is initiated in the same manner as the independent method. But once full genlock is achieved, the input signal's subcarrier is used as a reference for both loops.

An independent locking generator will track the SC/H phase of the input signal. It will also lock to a nonstandard color signal (one in which the subcarrier and sync are not phase locked at all).

A dependent locking generator will maintain correctly SC/H phased outputs regardless of the SC/H phase of the input signal. It will also lock more reliably to signals with occasional noise hits or with sync-only timebase error.

Genlocking the master generator to a nonsynchronous source such as network has given way to the use of frame synchronizers. However, genlocking is still used in "master/slave" sync systems where individual sources or groups of sources are driven off "slave" generators which are locked to the plant's "master" generator. And, as was previously mentioned, many of today's cameras and VTRs contain their own (genlockable) sync generators.

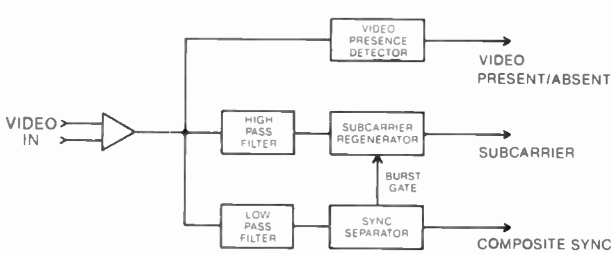


Figure 8. Simplified block diagram of basic sync generator genlock circuits.

As a result, it is not uncommon to find situations where three or more sync generators are cascaded. This makes it very important that the first generator in the line (usually the plant's master generator) have very good timebase stability. It is equally important that all slave generators exhibit very low "jitter." A figure of 5 ns or less is typical for timebase stability, while 2 ns or less is a reasonable jitter spec.

Where genlocking to an outside source such as network is done on a regular basis, a feature called "protected genlock" is highly desirable. With protected genlock, the local generator is locked to the outside source using the independent locking method. Any horizontal timing jumps on the locking signal will not disturb the generator; it will remain frequency locked to the locking signal's subcarrier. If the input signal disappears or becomes monochrome, the protected genlock logic will instruct the generator to make a smooth switch to the free run mode.

TIMING CONCEPTS IN THE MODERN STUDIO

In the early days of color, the system designer's attention became focused on the importance of color timing. While timing errors as large as a hundred nanoseconds could be tolerated in monochrome systems, color systems demanded accuracy two orders of magnitude better. The vectorscope took its place alongside the waveform monitor as an indispensable measurement tool.

As monochrome systems were converted to color, these new timing tolerances had to be addressed. The required degree of precision was not difficult to attain if reasonable care was exercised.

Unfortunately, the importance of the interrelationships between sync and subcarrier, especially SC/H (Subcarrier to Horizontal) phase, were not fully appreciated at the time. It wasn't until color videotape recording and editing came along that these issues came to light.

When the NTSC color system was devised, the frequency of the color subcarrier was chosen to be an odd multiple of half the line scanning frequency. This was done so that the subcarrier peaks would occur in alternating positions on adjacent scan lines, thus minimizing the visibility of the subcarrier on monochrome receivers.

The multiplier which was finally chosen was 455/2, or 227.5; this is also the number of cycles of subcarrier per scan line. Since there are 525 lines per frame, there are 227.5 times 525 or 119437.5 cycles of subcarrier per frame. The extra half-cycle of subcarrier means that it takes two frames (or four fields) for the color sequence to repeat.

When a videotape is edited, the machine must ensure that the continuity of this sequence is preserved. If it is not, there will be a noticeable horizontal shift in the picture at the edit point. Many machines incorporate *color framers* to prevent this from occurring. They operate by matching the color framing of the signal recorded on the tape with that of the house reference.

Reliable operation of these color framers is dependent upon the maintenance of consistent SC/H phase throughout the plant. Furthermore, when tapes which have been recorded in another facility are edited, the SC/H phase of the signal from the tape must match the local plant's SC/H phase. For this reason, an industry-wide standard has been established which defines correct, or zero, SC/H phase as the coincidence of the 50% point on the leading edge of horizontal sync and the zero-crossing of subcarrier. (See Fig. 2.)

Establishing and maintaining correct SC/H phase begins with having a means to measure it. Most modern waveform monitors provide this capability as illustrated in Fig. 9.

Maintaining SC/H phase depends upon careful selection of equipment and periodic measurements of actual performance. Sync generators with low SC/H phase drift are key elements in any SC/H phased plant. It is equally important to avoid system design techniques which might lead to SC/H phase problems (such as the use of regenerative pulse or subcarrier distribution amplifiers).

SC/H phase should be checked as routinely as video levels or color phase.

TIMING TOOLS

Numerous approaches to plant timing exist, but they all fall into one of two general categories: (1) adjusting the relative phase of the source timing reference, or (2) delaying the source video. In general, it is always better to adjust the timing signal. This is because even the best video delay systems introduce a certain amount of signal distortion. Obviously, video delay is not a viable solution where a source must be advanced in time with respect to other sources.

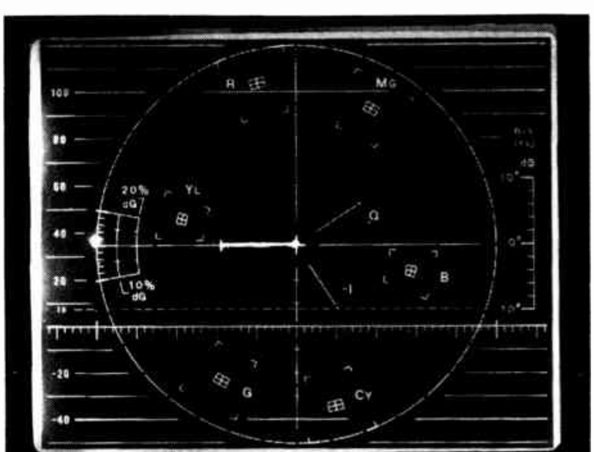


Figure 9. SC/H phase measurement on the Tektronix model 1750 waveform and vector monitor shows the relation between the burst phase vector and the "sync dot" as either an absolute value or relative to an external reference. Phase may be read out from the graticule or by adjusting and reading calibrated controls. (Courtesy Tektronix, Inc.)

Video delay lines are often used when a single source must feed different timing points. This is quite often the case in a system where a source must appear as a direct input to a studio production switcher, but the same source may be selected on a routing switcher which, in turn, feeds the production switcher. This is illustrated in Fig. 10.

In general, it is better to use a delay DA than a passive delay line/video DA combination. This is because a properly designed delay DA will include circuits which compensate for response errors in the delay elements. Where passive delays are used, they should be of the highest possible quality and the shortest electrical length required to get the job done.

Other devices which can be useful in plant timing are regenerative pulse and subcarrier DAs. They should be applied with caution, however, since they can create system SC/H phasing problems.

GUIDE TO PLANT TIMING

At first glance, timing a television plant may appear a formidable task. Like anything else, though, the task can be whittled down to size through a systematic method comprised of analysis, synthesis, and execution.

Analysis entails the careful examination of the timing requirements in each individual operating area of the plant. This information, along with a clear understanding of the capabilities of the equipment, can be used to synthesize a plan for overall plant timing.

The demands placed on most plants are constantly changing and ever-increasing; today's special setup may be standard operating procedure tomorrow. It is important to realize, therefore, that the plan which has been put together so carefully may change. If this happens, it may be necessary to add or replace equipment or perhaps even alter the system design.

Studio Timing

The studio (or edit suite) is an excellent place to begin a discussion on system timing, since it is the place where proper source-to-source timing relationships are most important. This is due to the fact that the studio is the primary area of the plant where multiple sources are combined. If visible picture jumps and color shifts are to be avoided when making mix, wipe, or key transitions, all studio sources must be carefully timed.

Timing the inputs to a studio switcher is a very simple, straightforward process:

1. Identify those sources which do not have timing adjustments or whose timing can only be adjusted over a limited range. These will more than likely

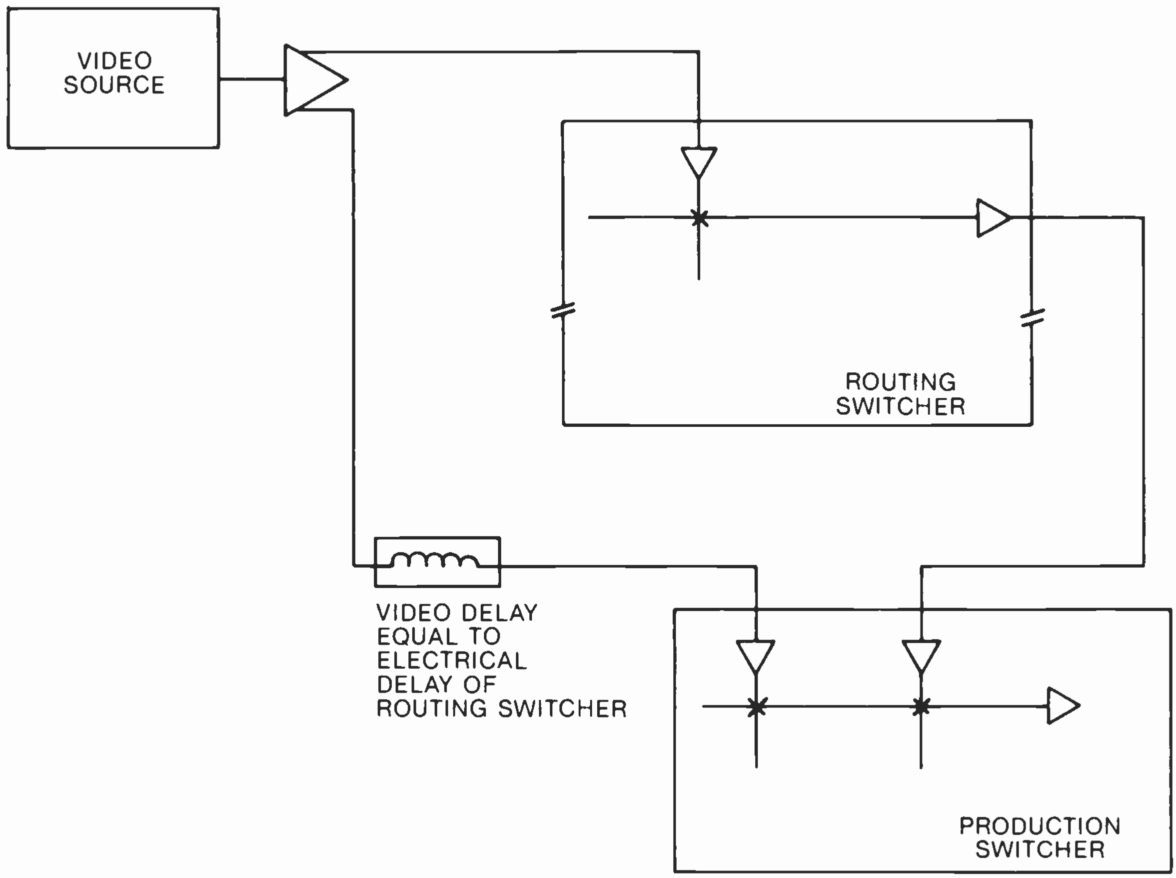


Figure 10. Timing a single video source to two different timing points.

include the black and background generators in the production switcher itself and may also include test signal generators, character generators, and older studio cameras or film chain cameras.

2. Using delay DAs or delay lines, match the timing of all of the sources identified in item 1 above to the "latest" in the group of fixed sources. Designate one of these sources as the reference source for setting the timing of all of the adjustable sources. Note that if there is only one fixed source, it should be used as the reference source.
3. Adjust the timing of each of the adjustable sources to match the timing of the reference source.

All timing adjustments should be made by selecting each source, in turn, on the production switcher. An externally referenced waveform monitor and vectorscope connected to the output of the production switcher provide the necessary monitoring capability. Timebase and gain magnification should be used on the test instruments to ensure the accuracy of timing adjustments.

The output SC/H phase of most equipment will change when timing adjustments are made. If the plant is to be SC/H phased, it is imperative that this parameter be checked whenever source timing is adjusted. On equipment which uses separate pulses and subcarrier for timing, care should be taken to maintain consistent phase relationship between the input sync and subcarrier once the output SC/H phase of the source is established.

If the plant is equipped with multiple studios, the procedures outlined above should be repeated in each studio. On those sources which feed more than one studio, extreme care should be taken to ensure that cable lengths are made equal. If this is not done, it will be impossible to achieve proper timing in all locations.

Timing in Master Control

In establishing overall system timing, master control may be regarded as a studio and can therefore be timed using the above procedure.

One timing problem which is fairly common in master control is that of studio delay. It generally shows up as a slight timing disturbance when master control switches from a given program source to that same program source through a studio switcher (a common occurrence going into or coming out of a live program such as news). An easy solution to the problem is to resynchronize the output of each studio (see "Cascading Studios," below).

Timing Through an Assignment Switcher

If routing switcher buses are used as input preselectors for a studio switcher, the propagation delay through the routing switcher must be taken into account when timing the plant. This is usually accomplished by setting up both direct and delayed paths for any sources which feed a production switcher both directly and through a routing switcher (see Fig. 10).

In plants where the routing switcher is used for source assignment, it is a good idea to designate the output of the routing switcher as the zero timing point for the plant. If this is done, a bus on the switcher can be designated as the plant "quality control point"—one location (with one set of accurately calibrated test instruments) where the signal levels and timing relationships of all sources in the plant can be observed and compared. The generalized timing procedure for studio switchers can be used to time sources into the plant routing switcher.

Cascading Studios

Occasionally, it may be necessary in some plants to cascade studios. In order to preserve overall system timing, the delay through each studio must be taken into account. One approach to keeping everything in time is to time all sources into the first studio "early" by an amount equal to the studio delay. However, if more than two studios are involved and timing into a routing switcher and master control is a factor, alteration of source timing may not be practical.

A far better approach is to resynchronize the output of each studio using a frame synchronizer. The timing of each studio can then be set to match that of all other sources in the plant.

A word of caution is in order, however: since nearly a frame of video delay is required to resynchronize a source which is only delayed as little as one hundred nanoseconds or so, the video from the studio when passed through a frame synchronizer will be out of sync with the audio by approximately 33 ms. While this amount of time difference might not easily be perceived by most individuals, the cumulative effects of several passes probably would be noticeable. Therefore, it is a good idea to insert an audio delay line in the program output from the studio to preserve accurate lip sync.

NON-NTSC SYSTEMS

Until recently, it was a safe assumption that all video signals throughout a broadcast facility are 1 volt p-p standard NTSC signals. However, the increasing popularity of component analog video recorders among broadcasters represents a potential challenge to this assumption.

Conventional one-inch type C and ¾-inch U-matic video recorders record an NTSC signal on the tape. One-half inch broadcast machines (Betacam, Betacam SP, and MII) record a component analog video (CAV) signal on the tape. These new video recorders have NTSC inputs and outputs and are designed to operate in an NTSC environment, but CAV inputs and outputs are also available.

The chief advantage afforded by this equipment is very high picture quality in a compact, cassette-based format. First generation quality approaching that of one-inch is possible. The quality of successive generations is good and can be improved considerably if dubbing (and any attendant picture manipulation or processing) is done in the component domain.

In order to extract the most quality from these newer tape formats, it is best to avoid repeated conversions to and from NTSC. This can be accomplished quite easily, even in an existing all-NTSC plant, by creating CAV "islands."

COMPONENT ANALOG VIDEO (CAV)

CAV camera/recorder units are used extensively in ENG (electronic news gathering) and EFP (electronic field production) applications. They are lightweight, relatively low cost, and produce very high quality pictures. Source material recorded in the field is almost always edited before being used on-air. Often, digital effects, captions, and other graphics are added to the finished product. This makes the ENG/EFP edit suite a good model for a CAV island.

The rules of good system design outlined earlier in this chapter apply generally to component systems. There are a few differences, all related to the nature of the component signal itself. NTSC is a "one wire" system; CAV is a "three wire" system.

All color television signals originate in component form, usually as red, green, and blue. The NTSC system was developed to allow these color components to be transmitted as a composite signal over a single transmission channel. Since a single wire is more convenient to deal with than three, NTSC has, over the years, become a distribution standard as well as a transmission standard. In small islands, the inconvenience of three wires versus one wire is a small price to pay for the enhanced picture quality which CAV offers.

In designing a CAV island, the choice must be made as to which component set the island will be designed around: RGB or Y, R-Y, B-Y. There are pros and cons associated with both, but most system designers seem to lean toward Y, R-Y, B-Y. This is because the Y signal provides a convenient feed for monochrome preview monitors and because the ½-inch CAV recorders utilize the Y, R-Y, B-Y component set.

In the Y, R-Y, B-Y component set, the Y channel carries the luminance information. The Y channel signal is nominally 1 volt and looks very much like a conventional NTSC signal, except that it is monochrome and therefore contains no subcarrier or burst. The other two channels, called the color difference channels, contain the color information. They are quite different in appearance from a conventional video signal. First of all, they do not contain sync. Second, they are bipolar signals with a nominal level of ±350 mV.

CAV Signal Distribution

When distributing CAV signals, a number of their unique characteristics should be considered. First of all, because they are monochrome signals, overall source timing is not as critical as with NTSC. So, rather than being concerned with timing accuracy to within a nanosecond, the system designer can think in terms of ten or so nanoseconds. However, interchannel timing, timing differences between the components, should be maintained as tightly as possible in order to prevent picture impairments due to differences in luma/color timing.

Second, the bipolar characteristics of the color difference signals must be preserved throughout the system. Blanking level must be maintained as close to 0 VDC as possible. But because the color difference signals do not contain sync, clamps or DC restorers cannot be used. Instead, it is best to use straight DC coupling. Many DAs and switching systems offer this capability, usually by setting straps or jumpers.

Third, interchannel gains should be closely matched and maintained as tightly as possible to preserve color balance.

Interchannel gain errors or DC shifts will result in color errors in the final picture. These errors can be avoided through careful setup and maintenance of system levels and through the selection of distribution and switching equipment with good gain and DC level stability.

Format Conversion

Regardless of which component set is selected for the design of the island, some interformat conversion will undoubtedly be required. For example, assume that a CAV edit suite is designed around the Y, R-Y, B-Y component set. Integrating an RGB character generator into such a system will require interformat conversion from RGB to Y, R-Y, B-Y. This is a relatively simple process which can be accomplished using a variety of commercially available products.

In some instances, conversion in the other direction, from Y, R-Y, B-Y to RGB, might also be required. Fortunately, this is also a relatively simple process.

Conversion between the CAV and NTSC environments may also be a consideration. In a CAV edit suite, for example, it may be necessary to use or produce NTSC signals. This can be accomplished quite easily by means of the tape medium. But if "real-time" NTSC interface is anticipated, decoding and encoding equipment will be required.

Commercially available NTSC decoders and encoders span a wide price and capability range. In choosing among the alternatives, the system designer must carefully evaluate the level of performance required.

There are basically two types of decoders: notch filter and comb filter. Notch filter decoders are relatively inexpensive and are fairly good at removing residual subcarrier from the luminance information. The basic characteristics of a notch filter does cause some loss of horizontal resolution in the luminance channel. Comb filter decoders are more expensive but provide excellent suppression of residual subcarrier. Luminance horizontal resolution is not usually affected, but apparent vertical resolution can be impaired, especially in areas of the picture where there are sharp horizontal transitions.

A third type of decoder, the so-called adaptive comb filter decoder, combines the best characteristics of both types of filters. Adaptive decoders contain circuitry which analyzes picture content and dynamically selects the appropriate filter characteristic. As might be expected, adaptive decoders are relatively expensive.

A similar range of choices exists with respect to NTSC encoders. Inexpensive units employ simple filters, while more expensive units employ more complex filtering techniques. The quality of the encoded picture varies directly with the cost of the box. The old adage, "You get what you pay for," definitely applies.

A fairly recent development in decoding and encoding technology is the application of digital technology to these processes. The Emphasys® decoding and encoding systems from The Grass Valley Group are 100% digital implementations. Hybridized digital/analog products are available from several other manufacturers. All of these products represent the state-of-the-art in decoders and encoders.

Other CAV System Considerations

Unfortunately, there is no single standard for CAV. One area of considerable difference is the precise signal levels for the Y, R-Y, B-Y component set. At the time of this writing, there are three standards: SMPTE/EBU, Betacam, and MII. The differences between these standards are relatively small, and fall within the adjustment range of all affected equipment within the CAV system. A thorough discussion of these standards can be found in the publication, "Solving the Component Puzzle," which is available on request from Tektronix, Inc.

Timing within a CAV island is relatively straightforward. If the island is to be timed into a large plant, there are certain things which must be taken into account. Chief among these is decoder delay. A simple notch filter decoder will have an electrical length of several hundred nanoseconds. An adaptive decoder may be as long as several horizontal lines. If decoded feeds are to be in time with sources local to the island, the entire timing reference for the island must be delayed (with respect to plant zero time) by an amount equal to the decoder delay. This can be accomplished quite easily through the use of a dedicated sync generator for the island; this generator would be locked to plant sync and adjusted to provide the correct amount of delay.

If the output of a CAV island is to be timed into the plant, the island sync generator could be adjusted to advance the timing within the island by an amount equal to the delay through the encoder placed on the output of the island. If the island is to be fully timed into the plant, on both the input and output side, the island output would have to be re-synchronized using a frame synchronizer.

Test and monitoring in the component domain requires specialized equipment. Conventional NTSC equipment is not suitable. Tektronix, Magni, and several other manufacturers offer a range of equipment specifically tailored for CAV applications.

Fig. 11 depicts a representative CAV edit suite and illustrates some of the concepts covered in the text.

DIGITAL SYSTEMS

Digital video equipment has been around for many years. Electronic character generators, first introduced in 1970, were among the first devices to use digital technology. The first video product to employ digital circuits for signal processing was the timebase corrector, introduced in 1974. The first digital frame synchronizer followed in 1975, and the first digital video effects system in 1976.

This equipment used digital technology for one of two reasons: the superiority of a digital solution over analog (time base correctors, for example), or the inherent impracticality of an analog solution (frame synchronizers, for example). All of these early devices were intended to be used in an analog environment and therefore had no digital input or output (I/O).

As the variety of commercially available digital video equipment increased, so did the interest in digitally interconnected video systems. The push for digital I/O shifted into high gear in the mid-1980s with the impending availability of the digital video tape recorder (DVTR). In 1987, when the first commercially-available digital recorder was introduced, many manufacturers had already incorporated digital I/O into their products.

Digital Video Standards

The majority of digital video equipment available today is equipped with both analog and digital inputs and outputs. Digital interconnectability has been made possible by a set of internationally agreed upon standards, many of which have been prompted by the standardization of digital video recording formats.

Digital video equipment exists for both component and composite systems. The first digital video recording format, the "D-1" format, records a component signal on the tape. Most high quality digital video effects and electronic graphics systems employ component processing. Therefore, if these devices are equipped with digital input and output, it is likely to be component digital input and output. Component signals which have been recorded digitally suffer virtually no degradation from generational losses or encoding/decoding artifacts. This is why component digital is the format of choice for critical graphics work or high-end post-production.

The high cost of component DVTRs have made them prohibitively expensive for use in many applications: program and spot playback in broadcast facilities, for example. The need for a more economical digital video tape format prompted the development of the "D-2," or composite digital format. Because of its high performance-to-cost ratio, composite digital is the digital format of choice for most broadcast applications.

Interconnection Conventions

Until recently, parallel interconnect was the only method available to the digital video system designer. SMPTE proposed standard T14.223 specifies a scheme which uses a separate pair of wires for each data bit and a separate clock. The connector called out in the standard is a 25 pin "D" connector. The same electrical and mechanical specifications apply to both component and composite systems. (The signal format is, of course, not the same for the two systems.)

There are numerous drawbacks to the parallel interconnect method. First of all, the cable is considerably more expensive than coax, both in terms of the cost of the cable itself, and the cost of applying connectors to it. Second, the cable and connectors are more bulky than coax. Finally, the distance over which parallel data can be reliably sent is limited to 100 meters maximum.

Digital video equipment is beginning to appear equipped with serial input and output. SMPTE proposed standard T14.224 specifies the mechanical and electrical characteristics of this interface as well as the formatting of the serial data stream. The connector called out in the standard is the familiar BNC. The data rate for serialized composite digital signals (525/60 NTSC standard) is 143 Mbits/sec. The data rate for component digital is 270 Mbits/sec.

Integrating older equipment with parallel input and output into a serial digital system may be accomplished using external conversion devices. Stand-alone serializers and deserializers are available from a number of different manufacturers.

Standard high quality video cable may be used for serial digital interconnects. Because of the high data rates involved, a low-loss double-shielded cable is recommended. Cable runs of up to 300 meters are possible with low-loss cable.

Impairments in Serial Digital Systems

Digital video signals are more robust than analog signals and are, therefore, less susceptible to hum, noise, crosstalk, and other such impairments. Gain variations, frequency response errors, and drifts are analog problems which do not affect digital signals until they becauseｓ severe.

Cable equalization is still required, but it is of a much different sort than that required in analog systems. Analog cable equalizers must be designed to ensure flat response from nearly DC to 8 MHz or 10 MHz. There is very little low frequency information in a serial digital signal, so the cable equalizer can be designed to compensate primarily for high frequency losses. Flatness of response is not important. And since the data stream can be (and usually is) regenerated, gain adjustments are not necessary. In practice, it should

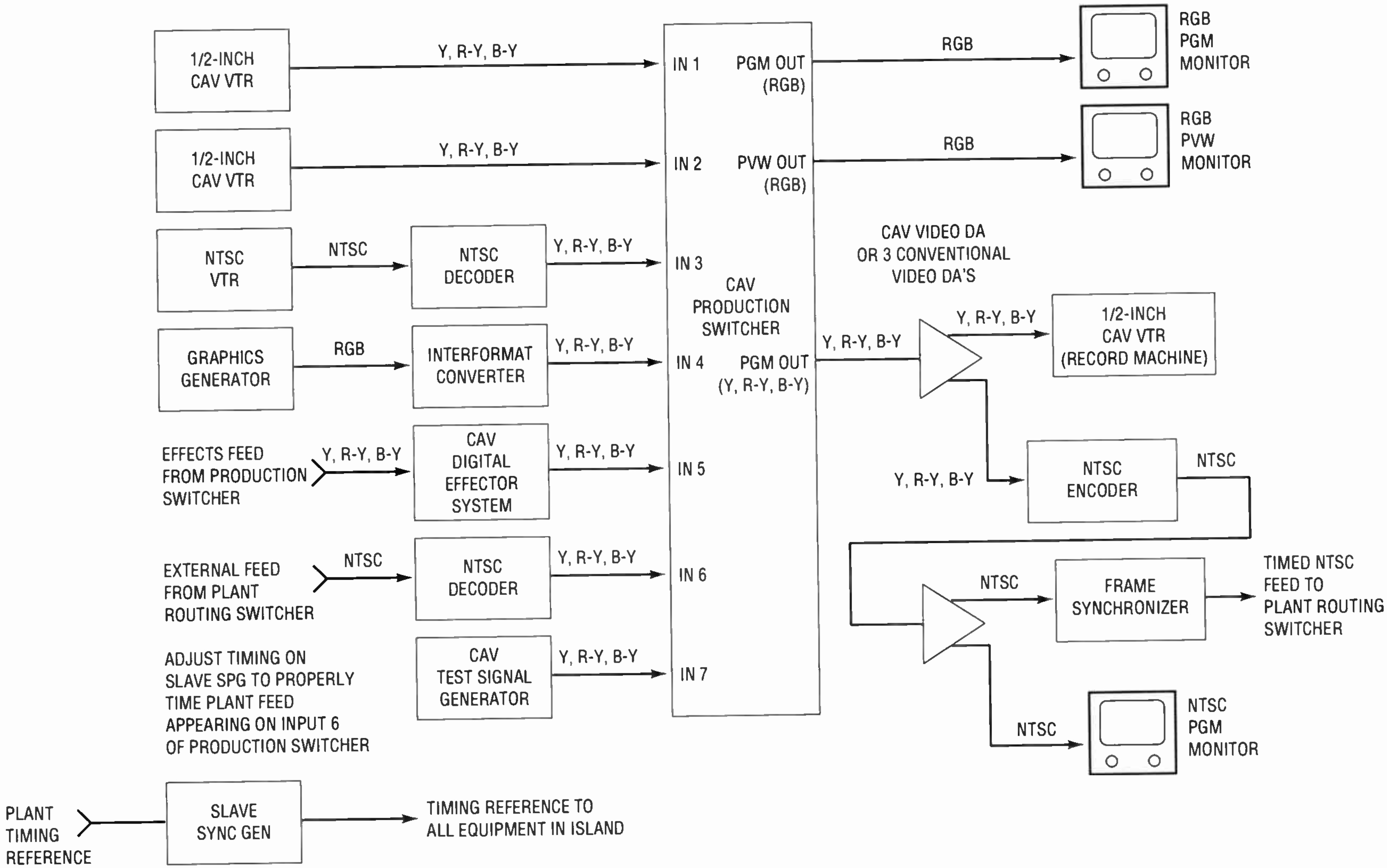


Figure 11. A typical CAV "island" within an NTSC plant.

be possible to design a digital system in which all cable equalization and gain adjustments are "automatic."

One problem which may occur in serial digital systems, and is unique to the digital world, is clock jitter. The serial digital signal contains an embedded clock signal which must be extracted to recover the data. As the signal passes through various devices in the system, the data stream may be subjected to repeated regeneration. This has the potential of adding significant jitter to the signal, making clock recovery difficult if not impossible. The solution to this problem is to deserialize the signal then reserialize the parallel data using a stable clock.

Timing Considerations

Precise timing is a necessity in analog video systems, particularly where multiple signals arrive at a single point where they will be combined; at the input of a production switcher, for example. This is typically not a requirement in digital systems. Most digital switchers and video effects devices have an input timing window within which all signals are expected to fall. Manual and, in some cases, automatic adjustment is provided to time the signal precisely for the internal processes which the device performs.

Digital signal processing involves longer time delays than analog. As an example, a large analog production switcher has an input to output delay of 500 to 600 nanoseconds. In comparison, its digital equivalent has a delay of a line (63.5 microseconds) or more. In order to deal with delays of this magnitude, the systems designer will need correspondingly longer delay lines.

Digital video effects devices typically exhibit through-delays of either a field or a frame. Delays of this magnitude can create lip sync problems with companion audio, not unlike that which broadcasters learned about with the advent of the frame synchronizer as noted above. The designer should take care to ensure that companion audio signals are delayed to match any video signals which are subjected to very long processing delays.

The Transition to Digital

Eventually, large scale television systems comprised entirely of digital equipment will become commonplace. But it will take many years for the existing installed base of analog equipment to be completely replaced.

The challenge to the system designer is to make the transition to digital smoothly. Each new project, whether it be the ground-up design of a new island or a rebuild of the entire plant, should be viewed in the context of how to link it to the digital future.

BIBLIOGRAPHY

Benson, K. Blair, ed. The Television Engineering Handbook. New York: McGraw-Hill, 1986.

Electronic Industries Association. EIA Standard RS-170, Electrical Performance Standards—Monochrome Television Studio Facilities. Washington, D.C.: EIA, 1957.

Electronic Industries Association. EIA Standard RS-250B, Electrical Performance Standards for Television Relay Facilities. Washington, D.C.: EIA, 1976.

The Grass Valley Group. NTSC Studio Timing: Principles and Applications. Grass Valley, CA: GVG, 1986.

Horn, John. Solving the Component Puzzle. Beaverton, OR: Tektronix, Inc., 1987.

Independent Broadcasting Authority. IBA Technical Review. Vol. 13, Standards for Local Television and Local Radio Stations. Winchester, Hampshire, England: IBA, 1980.

The Institute of Electrical and Electronics Engineers. IEEE Standard 511-1974, Trial Use Standard on Video Signal Transmission Measurement of Linear Waveform Distortion. New York: IEEE, 1974.

5.4 Magnetic Recording Media: Principles And Current State Of The Art

Part I

Michael P. Sharrock
3M Company, St. Paul, Minnesota

INTRODUCTION

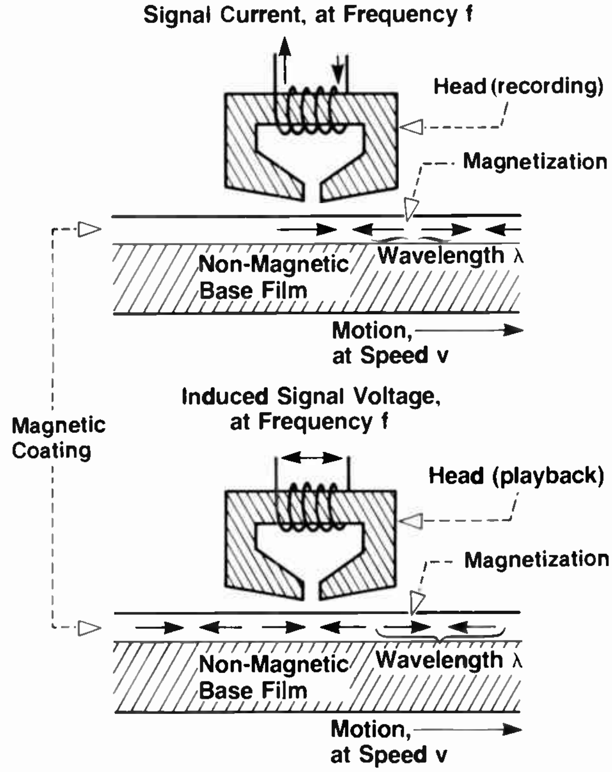
The broadcast of audio and video programming is greatly facilitated by the ability to store and recover information, and sometimes erase or alter it. Despite the availability of photography and other means of storing images, sounds, and symbols, magnetic recording has enjoyed a long pre-eminence in these applications because of a number of features:

1. Easy recording by an economical transducer (head).
2. Stable long-term storage.
3. Simple playback process with good signal-to-noise ratios, possibly using the same head as for recording.
4. Easy erasure, editing or updating, especially important in the broadcast studio environment.
5. Relatively high surface information density.
6. Thin flexible media, which can be rolled up (unlike phonograph records or video disks) for extremely high volume information density.
7. Inexpensive media. High-quality magnetic video cassette tape is available at a lower cost per square inch than the transparent tape used to mend books!

This chapter provides a brief overview of the basic physical principles involved in magnetic recording, some of the leading types of recording materials, and the directions of current research and development. More thorough discussions of various aspects of recording theory and practice are available. [Refs. 1–4]

BASIC PRINCIPLES

Audio and video signals are handled as electrical quantities (voltages and currents). The physical means of converting between these and a recorded magnetic pattern are shown in Fig. 1. The changing flow of electrical current in the recording head gives rise to a changing magnetic field which imposes a pattern of magnetization on the recording medium (i.e., tape) as it moves past the head. During playback the reverse



$$\lambda = \frac{v}{f}$$

Figure 1. The elements of magnetic recording and the relationship between signal frequency, transport speed, and recorded wavelength.

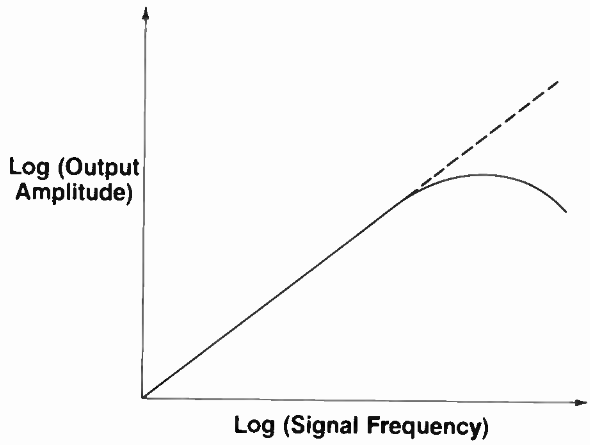


Figure 2. Logarithmic plot of output amplitude vs. signal frequency, for sinusoidal magnetic recording at constant magnetization intensity and depth. The slope of the linear rise is 6 dB per octave.

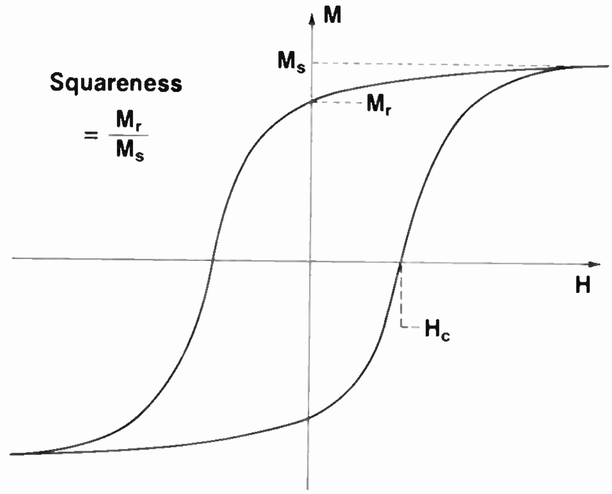


Figure 3. Plot of magnetization intensity M vs. applied magnetic field H, for a typical magnetic recording material.

process occurs, as the changing magnetic field experienced by the head, due to the passing recorded magnetic surface, induces a signal voltage. Some recording systems use the same head for recording and playback; in others, a specialized head is used for each of these functions.

The inductive nature of the playback head dictates that only a change in the magnetic field experienced by the head leads to an output signal. Further, a faster change produces a greater output amplitude. A consequence of this is a frequency dependence in the output amplitude. Fig. 2 shows this dependence, for sinusoidal signals recorded at a constant depth and magnetic amplitude in the medium. The output is proportional to frequency over a large range; the fall-off at high frequencies (high recording densities, short recorded wavelengths) is due to high-frequency signal loss from a number of sources, which will be discussed later. Some of the frequency dependence of the output is largely compensated for in many applications by the use of a frequency-dependent electronic gain; this process is called equalization.

Properties of Magnetic Media

The performance of a magnetic recording system depends, to a large degree, upon the magnetic properties of the material used to make the recording medium. Some of the most significant properties are explained in Fig. 3, which shows a plot (called a hysteresis loop) obtained by placing a sample of the material in a varying magnetic field (H) and measuring its magnetization intensity (M). The points at which the plot crosses the vertical M axis define the quantity known as the remanent magnetization intensity (M_r), which describes how much magnetization intensity the material is able to retain in zero field after being magnetized by a saturating field. The quantity usually specified (in the CGS system) is $4pM_r$, also designated as B_r, which is called the retentivity. The ratio of M_r to the saturated magnetization intensity (M_s) is called the squareness; it helps to define the shape of the loop and therefore the recording properties of the material. The points at which the loop crosses the horizontal H axis define the intrinsic coercivity, often referred to simply as the coercivity (H_c). This is the field needed to reverse half of the magnetization after saturation in one direction, resulting in zero net magnetization. The steepness of the plot as it goes from one direction of magnetization to the opposite relates to the ability of the material to record a signal with sensitivity and precision. A quantitative measure of this steepness is the switching-field distribution (SFD), described in Fig. 4.

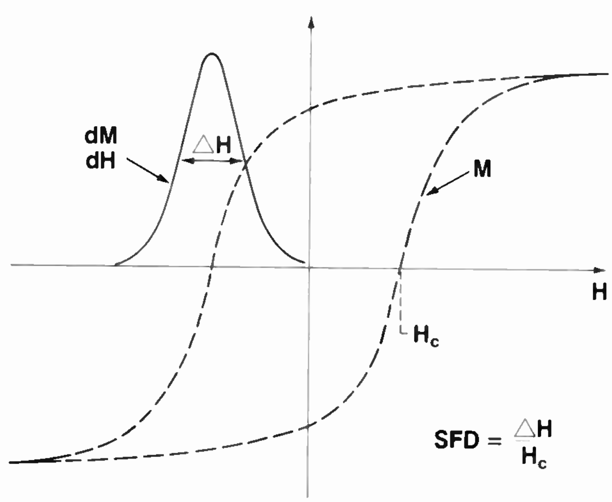


Figure 4. Plots of magnetization intensity M (Broken line) and first derivative dM/dH (Solid line) vs. applied field H. Width DH, used in the definition of the switching-field distribution (SFD), is measured at half the peak height.

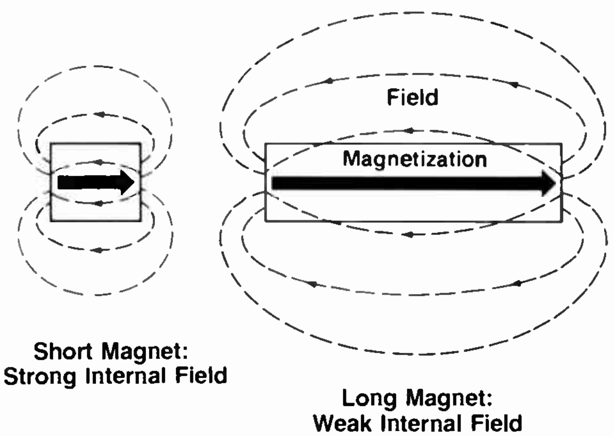


Figure 5. Internal demagnetization fields for magnets of different shapes.

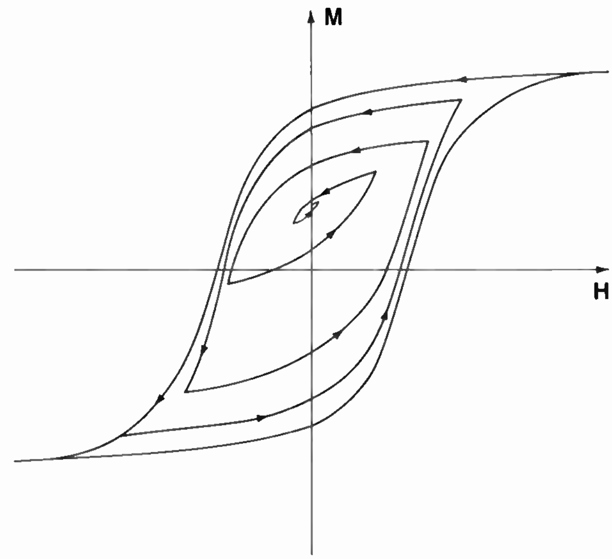


Figure 6. The principle of ac-bias. As a specific small area on the tape moves away from the recording head gap, it is subject to a high-frequency bias field of decreasing amplitude. This causes its magnetization intensity to vary as shown by the arrows, and to be left in a final state proportional to the signal field upon which the high-frequency field is superimposed.

The coercivity is the measure of a material's ability to resist magnetization changes due to an opposing magnetic field. As such, it is very important to record signals at high densities; that is, with short spacings between transitions. This relationship between coercivity and high-density recording capability can be understood with the help of Figs. 1 and 5. Magnetic recording in most cases produces regions in the surface layer of the tape or disk where the magnetization direction is aligned with the direction of relative head-medium motion, as depicted in Fig. 1. These regions behave like the magnets shown in Fig. 5; each generates an internal field that tends to oppose the magnetization that gives rise to it. The internal field depends upon the shape of the magnetized object, as shown, and is proportional in strength to the magnetization intensity in the object. The recorded waveform is therefore increasingly prone to self-demagnetization as the distance between transitions becomes smaller. Of course, the shape can be made more unfavorable to self-demagnetization by decreasing the depth of the recorded layer, but this will tend to reduce the signal strength. The demagnetizing effect of the internal field can actually be assisted by the recording head, because of the rapid reversal of its field as transitions are recorded with short intervals between them. This so-called "writing loss" accentuates the need for an adequate coercivity value when recording at high densities.

Linearity and Nonlinear Systems

The response of a magnetic recording material to an applied field is clearly nonlinear (see Fig. 3). In many applications, it is desirable to remove this nonlinearity so that the recording magnetization can be made proportional to the signal current. The needed linearization of the response is accomplished through the technique known as AC bias, in which the desired signal current is added to an alternating current of much higher frequency and amplitude before being applied to the recording head. The process and the result are illustrated in Fig. 6; more complete discussions are available elsewhere [Refs. 2,5]. The recording material is taken through a number of progressively smaller hysteresis loops as it leaves the vicinity of the recording head gap and is left in a state in which the retained magnetization intensity is very nearly proportional to the signal current. The use of AC bias is important in applications where the waveform read back must be an accurate reproduction of the input waveform; analog audio recording is an example.

In some applications, the recorded information is contained in the timing of events (such as the zero-crossing points) of the signal, rather than in the shape or amplitude of its waveform. Two of these applications, which do not usually utilize AC bias, are digital recording and frequency modulation (FM) recording. Digital recording is very simple in concept, although its applications can be very sophisticated. Digital recording seeks only to convey reliably the strings of binary digits (bits), 0's and 1's, that make up the information used in computers and other data-handling devices. Fig. 7 shows a simple scheme for doing this. In practice, more complex procedures are used; some additional magnetic transitions are used for timing purposes, and some convey information that is redundant but needed for the detection and correction of errors. Errors may occur through excessive random noise or through defects in the surface of the tape or disk.

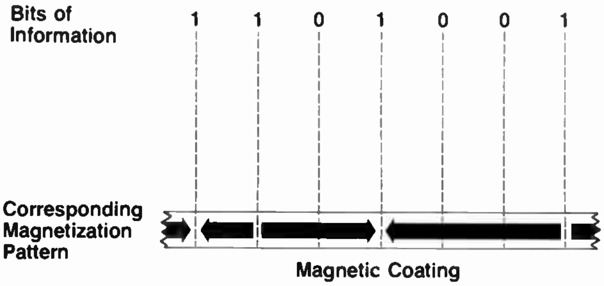


Figure 7. A simple form of digital recording, in which 1's are encoded as magnetic transitions and 0's as the absence of transitions.

Digital Recording

Digital recording, in addition to its obvious computer applications, is being used increasingly to replace analog recording techniques in audio and video applications [Refs. 6–8]. The original signal is measured at precise intervals and the measurements expressed as a series of numbers that are then recorded digitally like any other data. The recovery of the signal consists of reading back the series of numbers (again, at precise intervals), converting them into values of an analog quantity (i.e., voltage), and smoothing these discrete values into a waveform.

The first design consideration of a digital recorder is the frequency of sampling, which must be at least twice the highest frequency present in the signal to be recorded, according to the Nyquist theorem. Thus, for a high-fidelity audio signal containing frequencies up to a 20 kHz, sampling might be done at a rate between 40,000 and 50,000 times per second. Next, the dynamic range of the final reproduction is determined by the precision of each conversion to digital form. If a final dynamic range of 60 dB (an amplitude ratio of 1,000) is required, then the binary code for each sampling must be able to express integer numbers from 0 to 1,000. This requires 10 bits per sampled value, since $2^{10} = 1{,}024$. Similarly, a dynamic range of 90 dB requires a minimum of 15 bits per sampled value. The dynamic range is essentially the signal-to-noise ratio of the reconstructed signal, since the steps between adjacent numerical values provide the only noise inherent in the final output.

The advantages of a properly designed digital audio system include an excellent dynamic range, and an essentially perfect time base (absence of wow and flutter). In both audio and video recording, digital technology provides the ability to do repeated editing or dubbing operations without signal degradation. The tenth-generation reproduction made on a digital recorder is virtually indistinguishable from the original. This resistance to degradation of the recorded signal is intrinsic to digital technology and results from the fact that the information is encoded only in the timing of transitions. As long as these can be reliably detected and are not excessively shifted in time, substantial degradation of the amplitude or shape of the digital waveform can occur with no effect on the result.

Digital magnetic recording, while not requiring accurate reproduction of a waveform, makes great demands upon the recording system because of the drastically increased bandwidth. If an audio signal of 20 kHz bandwidth is to be reproduced with the modest dynamic range of 60 dB, frequencies up to at least 225 kHz are needed to accommodate the audio signal's digitally encoded representation. An actual studio recorder might operate beyond 600 kHz, depending upon sampling rate, dynamic range, and the digital code used. Such frequencies dictate very short intervals between the magnetic transitions on the recording surface if tape usage is not to become prohibitive. On the other hand, the signal-to-noise ratio of the digital waveform that is read back need not be as great as that required in an analog recording. In general, the lower the signal-to-noise ratio of the digital waveform, the higher the probability of misread bits (bit errors), but these can be detected and either corrected or concealed through the use of encoding schemes that involve a certain amount of redundancy in the recorded information. These schemes can overcome not only isolated single-bit errors but also the effects of physical defects in the medium's surface, which can delete a number of adjacent bits. [Ref. 9] Error correction techniques make possible the use of very high densities in digital recording, provided that the heads and media can operate with the required spatial resolution.

Frequency Modulation Recording

Another mode of magnetic recording that does not require linear waveform reproduction is frequency-modulation (FM). This is primarily used in video recording [Ref. 10], but can also be applied to audio recording and is in fact used to give high-fidelity sound in some video cassette recorders. The principle of FM is that a carrier wave is generated whose frequency varies with the value of the signal to be transmitted (Fig. 8). It is this frequency-modulated carrier that is recorded on the magnetic medium. As in digital

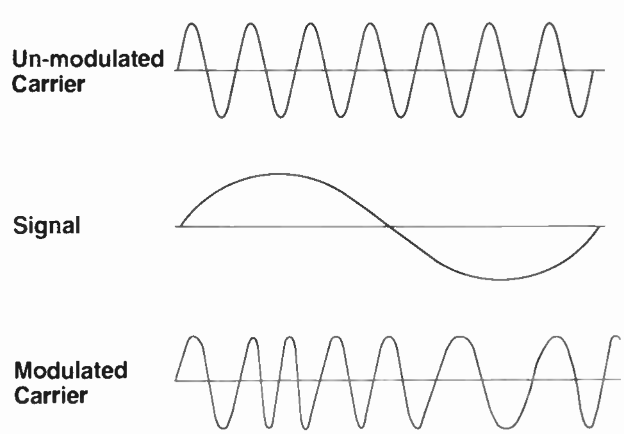


Figure 8. The principle of frequency modulation (FM).

recording, the information is contained in the intervals between the zero-crossings of the recorded waveform; the amplitude is not significant as long as the signal-to-noise is adequate for reliable detection. Unlike digital recording, FM recording contains an analog relationship between frequency and the value being transmitted. The benefits of FM, as compared to direct analog recording, are two-fold. One is the insensitivity to amplitude variations in the recorded waveform. The second, which is crucial to magnetic video recording, is that FM reduces the ratio of the highest to the lowest frequencies that need to be recorded; this is shown schematically in Fig. 9. Suppose an engineer desires to record a video luminance (brightness) signal that contains frequencies from 30 Hz to 4.5 MHz. The upper and lower frequency limits have a ratio of 150,000, or 17 octaves. (By comparison, a full-fidelity audio system is required only to reproduce frequencies from 20 Hz to 20 kHz, a ratio of 1,000, or 10 octaves.) A frequency span of 17 octaves is impractical in magnetic recording. The 6-dB-per-octave frequency dependence (Fig. 2) would require a dynamic range of more than 100 dB for direct video recording, too much for successful equalization. Other difficulties would also arise from the large ratio of longest to shortest recorded wavelength. Frequency modulation, as indicated in Fig. 9, reduces the ratio of highest to lowest frequency; this ratio after modulation is as low as one octave in some video recording systems. Details of

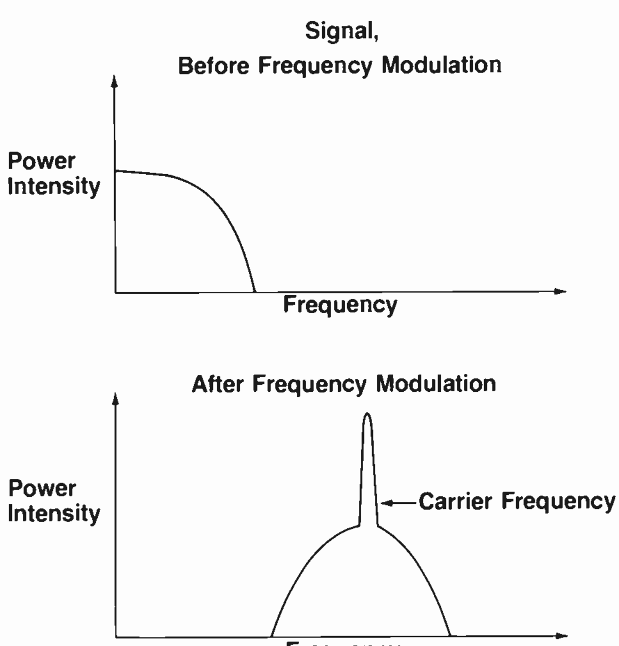


Figure 9. In frequency-modulation (FM) the information is shifted to higher frequencies, but the ratio between highest and lowest frequencies is greatly reduced. FM schemes actually used in video recording do not use all of the sideband frequencies; see pp. 396–401 in reference 2.

video recording formats are included in Chapter 5.6, "Video tape Recording" of this *Handbook*.

The frequencies involved in video reproduction are much higher than those in audio. Since there are practical lower limits on the recorded wavelength, this means that much higher head-to-tape speeds are needed. To attain these speeds, a video recorder uses heads mounted on a spinning drum. The tape is transported around part of this drum at a speed much lower than that at which the heads are moving across the tape surface. In most currently used video recorders, the tape wraps helically around the drum; they are therefore referred to as "helical" recorders. The various systems in current use employ different FM carrier frequencies as well as different schemes for conveying the color information, which in practice has a lower resolution and therefore requires a smaller bandwidth than that of the luminance information. In most current systems the color (hue and saturation) information is conveyed by amplitude and phase modulation of a special color subcarrier (at 3.58 MHz in the U.S.A.). In the one inch, open-reel tape format commonly used in broadcasting (C-format), the modulated color subcarrier is simply added to the luminance signal and the resulting sum is the signal used to frequency modulate the video carrier. This is called composite analog video. In ¾-inch U-matic systems and in consumer cassette systems, the color subcarrier (with its amplitude and phase information) is heterodyned down to a lower frequency (0.6 to 0.7 MHz, depending on the system), which is then recorded directly, using the frequency-modulated luminance carrier for AC bias. Some ½-inch cassette systems for professional applications utilize FM recording of both luminance and color information with a separate head, tape track, and carrier frequency for each. This achieves broadcast-quality reproduction using cassettes similar to those used in consumer systems, but at the cost of higher tape consumption (shorter playing times) than in the consumer systems.

Digital video makes great demands on recording density; these demands will be reviewed briefly here in order to place in context some of the advances that have occurred. In the format known as D-1, a component digital system, the luminance signal is sampled at a rate of 13.5 million times per second and each of the two chrominance (color) signals at half this rate [Ref. 11]. Each sampled value is digitized with eight bits. The resulting data rate is approximately 2×10^8 bits/second. The magnetic recording or recovery of this information will therefore involve about 2×10^8 magnetic transitions per second. (The exact ratio of bits to transitions depends upon the choice of encoding scheme, but is generally close to one.) Two heads read or write simultaneously to accommodate this data rate. The D-1 format is referred to as a 4:2:2 system; the numbers give the relative sampling rates of the luminance signal and the two chrominance signals, and is a world technical standard by virtue of CCIR Recommendation 601.

The second currently used digital video format, called D-2, operates with a composite signal, like the C-format analog system [Refs. 12,13]. The luminance (brightness) signal and the 3.58 MHz color subcarrier, which conveys chrominance (color) information by both amplitude and phase modulation, are added together. The resulting composite signal is then sampled at a rate of 14.318 million times per second, again with eight bits per sample. (The sampling rate is four times the color subcarrier frequency.) The resulting data stream is approximately 1×10^8 bits/second, about half that of the D-1 system. The D-2 system is simpler and cheaper than the D-1. The relative advantages of composite and component recording, relating to compatibility with other equipment and signal processing ability, are beyond the scope of this *Handbook* chapter.

High Definition Television

Recent years have seen strong interest in various forms of high definition television (HDTV). These involve the use of about 1,100 or 1,200 horizontal lines per picture, twice the number used in conventional television, to give enhanced vertical resolution. Horizontal resolution is comparably enhanced, with the result that the information to be conveyed per second, and therefore the required bandwidth, is increased about four-fold. Fortunately, along with HDTV have come powerful techniques for data compression. These involve compromising the ability to simultaneously convey high definition and rapid motion in the picture, and thus allow HDTV to be encoded digitally with a data rate on the order of 1×10^8 bits/second, similar to that of the normal definition, uncompressed D-2 format [Ref. 14]. A system has been developed, however, to record the full, uncompressed data stream from a 4:2:2 sampling of a component HDTV signal [Ref. 15]. Its overall rate is 1.2×10^9 bits/second, divided into eight channels of 1.5×10^8 bits/second each. During recording or replay operations, eight heads are thus operating simultaneously.

The importance of error correction, mentioned earlier, to the success of digital audio and video recording must be stressed. Sophisticated codes reduce the final error rates to levels that are orders of magnitude smaller than those that would occur in the data as read directly from the tape. These coding techniques allow information to be stored at much higher densities than could be used if the recording and playback processes were required to be essentially error-free. The price that must be paid for this (besides additional electronic complexity) is the recording of additional bits to be used in detecting and correcting errors. With modern coding techniques this price is relatively modest; the D-2 system has about a 15% overhead for error correction [Ref. 13].

MAGNETIC RECORDING MEDIA

Most magnetic recording media in use today are manufacturered by dispersing small magnetic particles in an organic binder and coating the resulting "paint" on a support material. See Fig. 10A. Each particle remains uniformly magnetized; its magnetization can change in direction but not in magnitude. (This is a result of the very small size.) Each particle generally has two opposite stable directions of magnetization, aligned with a so-called "preferred" axis. The collective effect of particles having their magnetization switched between these two directions is a change in the overall magnetization of the medium. The particles are sufficiently small that there are many of them in each magnetized region of the recorded pattern; this provides the needed spatial resolution and signal-to-noise ratio. The particles can be of various types, according to the intended application, and are discussed in more detail elsewhere [Refs. 16, 17]. In addition to the magnetic particles, the coating contains polymers for strength, lubricants to reduce friction and wear, and surfactants to aid in dispersing the particles. Nonmagnetic particulate components are often included in order to impart a controlled degree of abrasivity (for head-cleaning purposes), adequate opacity for optical tape position sensing, and sufficient electrical conductivity to prevent static charge accumulation.

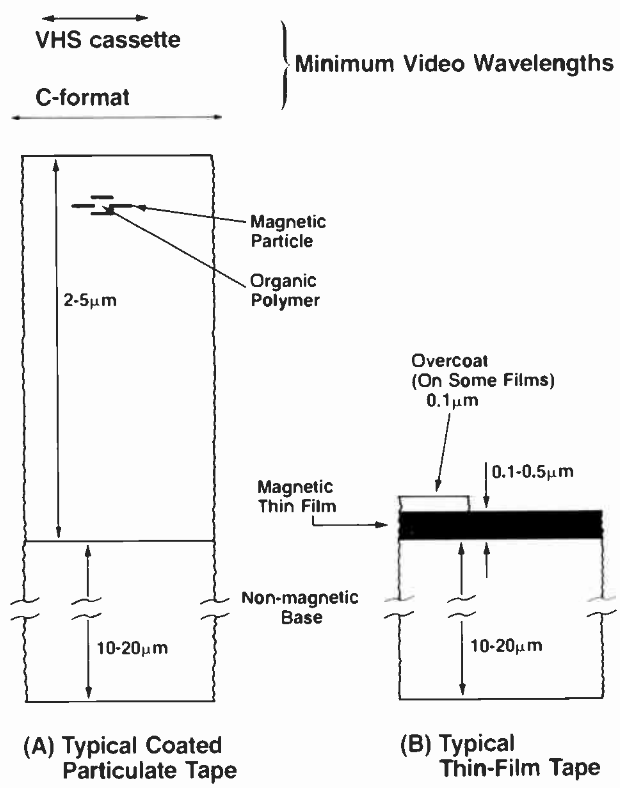


Figure 10. Cross sections of a coated particulate recording tape and thin-film tape. The particles are drawn approximately to scale with the wavelengths, but their spacing and uniformity have been exaggerated. Only a few have been drawn, although they actually fill the magnetic coating layer, occupying about half of its volume.

Another possible construction for a magnetic recording medium is shown in Fig. 10B. Like the dispersion-coated type, it also uses a nonmagnetic base. The magnetic substance is in the form of a thin metallic film [Ref. 18]. The film is made up of small crystals whose magnetization direction can be changed; these function similarly to the particles used in the dispersion coating. Thin films offer magnetization intensities far surpassing those available in the coatings of dispersed particles, which compensates for the films' lesser thickness. The thinness of the films is in itself an advantage in that the magnetic material is all in very close proximity to the record or playback head. This intimate contact with the head is extremely important to high-density recording. These attractive features make thin-film recording materials an exciting area of research and development and possible candidates for future products. They are at present used commercially in rigid disks (for computer data storage) and in a high-bandwidth version of the 8-mm video cassette tape (known as Hi8) [Ref. 19]. On the negative side, metallic thin films are subject to some concerns regarding their durability, chemical stability, and economical manufacture. Many of them are coated with a thin nonmagnetic overlayer for protection and lubrication. This layer, of course, tends to diminish the closeness of contact with the head. Metallic thin films are made by chemical plating or by some form of vacuum deposition and usually contain cobalt as a major constituent. The need for the special equipment used to carry out these processes is an additional barrier to the widespread commercial use of thin-film recording media. Furthermore, the thin-film products must compete with a particulate technology that not only benefits from 40 years of experience in manufacturing and application but continues to improve.

Types Of Material In General Use

For the foreseeable future, economic and practical considerations dictate that coatings of dispersed particles will be predominant in audio and video tapes. The earliest commercially available audio tapes, made in the late-1940s, used iron oxide particles. The most important oxide, $\gamma\text{-}Fe_2O_3$, has been greatly improved as to particle shape and size since its introduction, and it retains an important role in audio, video, and data-recording tapes today. It is also a commonly used material for the disks, both flexible and rigid, used in digital data storage for computers. The relatively low coercivities (250 to 400 Oe) iron oxides, however, proved to be a serious limitation as recording densities increased.

The requirement for higher coercivities was first met in the mid-1960s with the introduction of chromium dioxide (CrO_2). Like iron oxide, this material derives its coercivity and squareness properties, and therefore its ability to retain magnetization, primarily from the needle-like shape of the particles. Digital computer tapes were the first products to be made with CrO_2. Shortly after, it was applied in audio cassettes and is today used in audio, video, and data applications. An unusual feature of CrO_2 is its very low Curie point, about 250°F. This is the temperature above which the magnetization of a material vanishes, to reappear (possibly with a different direction) upon subsequent cooling. The very low value for CrO_2 permits its use in thermomagnetic duplication (TMD) [Ref. 20]. In this process, the information is copied from a master tape, having a much higher Curie temperature, to the CrO_2 tape by heating while their surfaces are pressed together. Chromium dioxide is the only particulate material known to be useful in practice for thermomagnetic copying. Contact duplication processes, of which thermomagnetic copying is one, are of economic importance because they permit copying of high-density information at much higher speeds than are practical in direct head recording.

At the present time, the most widely used materials for applications requiring coercivities in excess of 400 Oe are iron oxide particles modified by the addition of cobalt. The interaction of the cobalt ions with the iron oxide structure provides additional resistance to the switching of the magnetization direction, and thus increases the coercivity. Values in the range of 500 to 600 Oe for audio cassette tapes, 600 to 700 Oe for most video cassette tapes, and 800 to 900 Oe formats (S-VHS video cassettes; D-1 digital videotapes) are readily achieved. Many processes exist for adding the cobalt to the particles. Most in use today leave the cobalt segregated near the surface of the particles. Such particles, called *cobalt-surface-doped, cobalt-adsorbed, cobalt-epitaxial*, or *cobalt-ferrite epitaxial* are now the predominant materials used in video cassettes and open-reel videotapes and are finding increased use in high-density computer diskettes.

Tapes made with particles of metal (iron or iron alloys) have much higher remanent magnetization intensity values than are possible in tapes that use oxide particles, and accordingly can give much higher signal levels. This property is especially valuable in very compact formats, and metal particles are used in some audio cassette tapes, in 8-mm video cassette tapes, and for digital audio and video recording tapes (DAT and D-2). Metal-particle tapes typically have coercivity values around 1,500 Oe. Such high values are needed to resist self-demagnetization, so that high-density signals can be successfully recorded at the high magnetization intensity levels offered by this material. Because of their high surface-to-volume ratios (common to all small particles) and their elemental metallic nature, metal particles made for recording must be stabilized against corrosion by a controlled oxidation of their surfaces. This process is called passivation.

Fig. 11 summarizes the properties of various particulate recording materials. The particles discussed thus far are acicular, or needle-shaped, each having its magnetically preferred axis along its length. In the manufacture of tape, these particles are usually oriented with their preferred axes approximately collinear with the tape length. This is done by the application of a magnetic field while the coated dispersion is still liquid. The particle orientation is then approximately

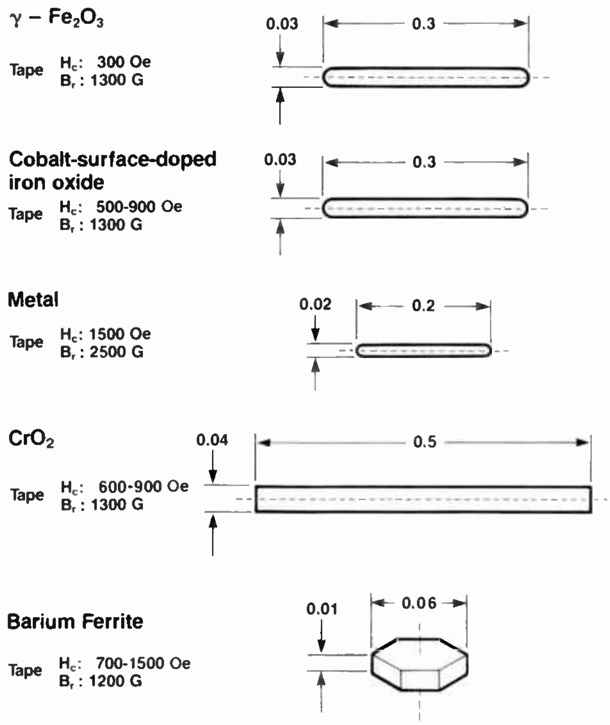


Figure 11. Various particles used in magnetic recording. The preferred axes of magnetization are shown by broken lines. Some typical, approximate dimensions (in mm) and magnetic parameters (in longitudinally oriented tape form) are indicated. Shapes are somewhat idealized.

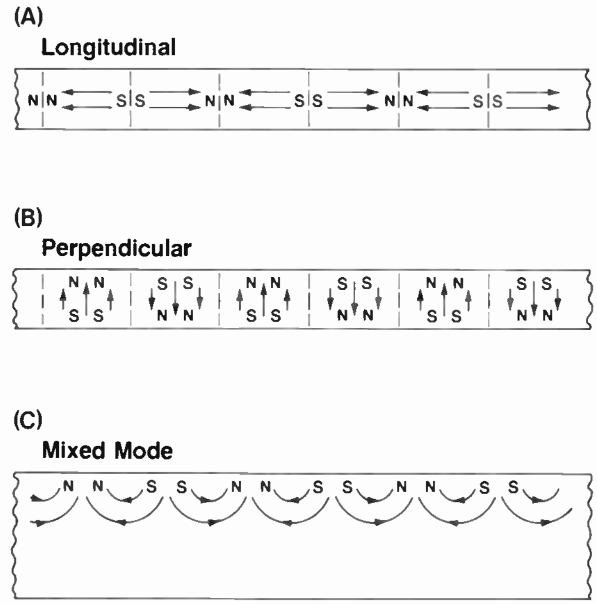


Figure 12. Possible magnetization patterns in recording media.

along the direction of relative head-tape motion; this is true even in helical formats, because the head track makes only a small angle with the tape edge. This orientation increases the remanent magnetization intensity and decreases the SFD (Figs. 3 and 4) as measured in the longitudinal direction. Tapes made in this way are designed to optimize the recording efficiency for longitudinal magnetization (Figs. 11 and 12A).

Barium ferrite is unique among the particulate recording materials; the particles are in the form of plates, each having its magnetically preferred axis perpendicular to the flat faces (Fig. 11). This material has attracted considerable attention for its potential use in making a perpendicular recording medium [Ref. 21]. It was thought that the particles could be relatively easily oriented flat in the plane of the coating; their preferred axes would then support recorded magnetization in the perpendicular direction (Fig. 12B). Acicular particles, in contrast, are known to be very difficult to orient standing on end in the coating so as to favor this direction. Successful perpendicular-recording media have in fact been made with barium ferrite [Ref. 22], but much of the development in recent years has involved the conventional longitudinal magnetic orientation [Refs. 23, 24]. Tapes and disks have also been made with the barium ferrite particles having nearly random orientation [Refs. 25, 26]; these could support complex magnetization patterns having both longitudinal and perpendicular components (Fig. 12C).

Barium ferrite has not succeeded in displacing acicular particles from any current tape applications. It is used, however, in a new format of flexible (floppy) disks for high-density data storage [Ref. 25]. Barium ferrite, oriented to favor perpendicular magnetization, may find use in prerecorded digital audio tape (DAT) cassettes [Refs. 27, 28]. These would be recorded by a type of contact duplication where an alternating magnetic field is used to reproduce the information in the copy tape, rather than the elevated temperature employed in TMD. Barium ferrite and metal are both more costly particulate materials than the iron oxides and CrO_2 and are used only if their properties are needed in a given application. Barium ferrite lacks the high magnetization intensity of the metallic recording particles. It has, however, some unusual magnetic properties that include very narrow switching-field distribution [Refs. 29, 30].

Manufacturing Considerations

Whatever the composition of the particles in a recording tape or disk, their size is extremely important, since it determines the number of particles contained in each magnetized region of the recorded signal. The greater this number, the greater will be the maximum signal-to-noise ratio that the recording medium can provide, since each particle contributes a pulse of magnetization to the playback head [Ref. 31]. A great many small pulses clearly will combine to make a more accurate, less noisy signal than a smaller number of large pulses (even though the total density of magnetic material present may be the same), and the effect of random fluctuations in particle size or placement will be less.

One of the most important and difficult arts in the manufacture of magnetic recording media is the dispersion of particles in the coating. If appreciable amounts of the particles are clumped together, the recording performance suffers. The noise properties of the signal are dominated to some extent by the clusters of particles, which are of course much larger than the individual particles themselves. A great deal of attention must therefore be paid to the formulation and milling of the dispersion, and to its coating and drying.

In the search for magnetic materials that give higher performance, care must be taken to avoid various undesirable features. For example, the trend to reduce particle size, and thus enhance signal-to-noise ratios, must not be carried to the point where the particles are so small that they begin to show the effects of magnetic instability. In particular, the presence of excessively small particles can lead to the unwanted acquiring of a signal in one layer of tape on a reel as a result of the fields due to a strong signal on an adjacent layer. This phenomenon is often referred to as *print-through* or *pre-* and *post-echo*, and is important only in (AC bias) audio recording [Ref. 32]. In some applications, especially audio, thorough erasure is an important concern and one that places an upper limit on the coercivity of the recording medium. Efficient recording and erasure also require that the material's switching-field distribution (Fig. 4) be relatively narrow. This means that the fields required to reverse the magnetization of the individual particles should be tightly clustered around the coercivity (which is essentially the median switching field). The achievement of a narrow SFD requires that particle size and shape distributions be adequately narrow.

Although this discussion has focussed on magnetic properties, one must not underestimate the importance of the nonmagnetic components of recording media. The base film, the binder polymer, and the various lubricants and other additives in the coating are crucial to strength, durability, and freedom from undesirable levels of friction. The manufacture of recording media requires at least as much capability in organic chemistry as in magnetic materials. The ability to reliably produce coatings with smooth surfaces, substantially free of defects, is also an absolute necessity.

PROGRESS TOWARD HIGHER RECORDING DENSITIES

The state of the art in magnetic recording has clearly advanced in recent years. Fig. 13 shows the great gains in practical recording density that have occurred in the video area. Efforts in recent years, however, have been aimed less at increasing the compactness of formats (continuing the slope of Fig. 13) than at enhancing the capabilities of the recording systems in other respects. Two examples are furnished by the development of the S-VHS and high-band 8mm (Hi8) video cassette systems, which are modifications of the previously existing VHS and 8mm systems. In both cases, the goal was improved picture quality through increased bandwidth, with the physical format essentially unchanged.

A major motivation for enhanced recording density has been the desire for digital formats, in both audio and video recording. As was mentioned earlier, a digital format will always require the writing or reading of many more magnetic transitions per second of programming material than does a comparable analog format. This increase is most strikingly seen in video recording, because the information content is much higher to begin with than in audio. Table 1 gives approximate recording densities and tape consumption rates for a few common video formats, some of which have been discussed above. The top row of numbers illustrates the order-of-magnitude increase in transition rate required by changing from analog to digital format, and also the additional increase imposed by uncompressed digital HDTV. The next three rows illustrate the increases in magnetic recording densities that have been achieved; the systems are listed in approximate

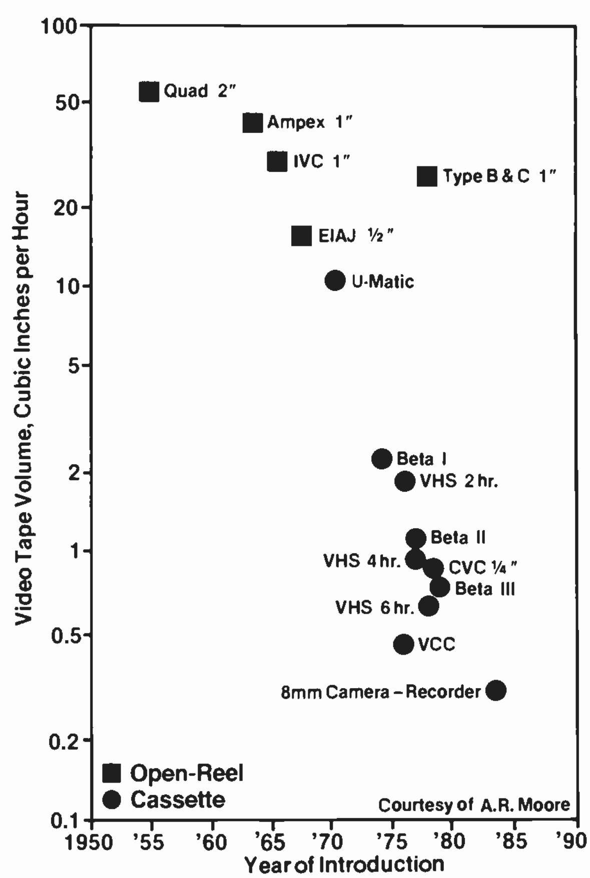


Figure 13. Tape volume required per hour of playback, plotted against year of introduction, for various video formats. This plot was supplied by A.R. Moore of 3M.

TABLE 1
Approximate transition rates, recording densities, and tape consumption for various video formats.

System	C-Format One-Inch	VHS Cassette (2-Hour)	D-1	D-2	Uncompressed HDTV	High-band 8-mm (2-hour)
	Analog	Analog	Digital	Digital	Digital	Analog
Material	Co-oxide[a]	Co-oxide	Co-oxide	MP[b]	MP	MP or thin film
Maximum magnetic transition rate (10^7/sec)	2	0.9	~20	~10	~120	1.4
Maximum transition density along track (10^3/inch)	20	38	56	60	74	103
Track density (10^2/inch)	1.4	4.4	5.6	6.5	6.9	12
Maximum area transition density (10^7/square inch)	0.3	1.7	3.1	3.9	5.1	12
Tape consumption (square inches/sec)[c]	7	0.5	6.4	2.6	24	0.11

[a] Iron oxide particles with added cobalt.
[b] Metal particles
[c] Excludes area of tape not occupied by video information, which would be a small fraction of total area.

order of development. The last row of numbers shows the tape consumption that results from the requirements and the densities. Note that increased recording densities permit the D-2 digital composite recorder to have a lower consumption than the C-format analog composite recorder. Both are broadcast-quality studio recorders, with the D-2 system offering the additional advantage of multiple-generation copying without degradation of the final output signal.

Track Density

The achievement of higher information densities on the surface of magnetic recording media requires progress in many areas: magnetic, mechanical, and electronic developments are all involved. The track density (inverse of track pitch, the perpendicular distance between tracks) is in some respect simpler to approach than the linear density along the track. Decreasing the width of a recorded track means that, all else being equal, fewer magnetic particles pass under the head per second. Again with all else equal, the recovered signal amplitude and the contribution of the medium-to-noise power are both proportional to the track width [Ref. 31]. This noise contribution is due to the discrete, particulate nature of the medium and is analogous to electronic shot noise. If it is the predominant noise source, which is the case in many audio and video systems, the root-mean-square noise amplitude is proportional to the square root of the track width; the resulting signal-to-noise amplitude ratio then also has this dependence. A decrease in track width must be compensated for if some aspect of performance is not to be degraded. A more strongly magnetized material can increase signal amplitude; reduced particle size can decrease the noise level; more sophisticated signal processing can make the final output (especially in digital systems) more tolerant of errors due to lower signal-to-noise ratio. Other aspects of reducing track width are mechanical; heads must be made narrower and track-following must be made more precise.

The increase of the density of magnetic transitions along a track involves more complicated phenomena, which depend on nearly every aspect of the tape, the head, and their interaction. The coercivity is an extremely important tape property. This parameter has been increased from about 300 Oe, in γ-Fe_2O_3 audio tapes, to more than 1,500 Oe, in current metal-particle tapes, in order to resist the self-demagnetization effect as transitions are made closer together. The later stages of this trend have required corresponding developments in heads. Tapes having coercivity values appreciably above 800 Oe cannot be satisfactorily recorded or erased by the inexpensive, durable heads of ferrite composition found in, for example, a VHS cassette recorder or a C-format studio video recorder. Metallic heads, capable of producing stronger magnetic fields, were introduced for the earliest applications of metal particles, in audio cassette recording. These metal heads lack the wear resistance of ferrite, however, and are unsuitable for use in video recorders, which have high head-to-tape speeds. To combine metal-particle capability with good wear properties, heads have been developed that have a ferrite structure with metal in the gap regions. These are called *metal-in-gap* (MIG), and in some cases *tilted sendust sputtered* (TSS) heads [Refs. 33, 34]. Together with such advanced heads, the high magnetization intensities and commensurately high coercivity values of metal particle (MP) tapes have made possible the high recording densities shown in the three columns on the right of Table 1. It remains to be seen whether barium ferrite or other new materials will find use in advanced video formats. The appearance of a thin-film tape as one of two alternative media for the high-band 8mm system (Table 1) is noteworthy [Ref. 19]; thin-film tapes have not been used in any professional audio or video formats as yet.

Particle Orientation and Size

The benefits of using a medium that favors the perpendicular direction of magnetization rather than the longitudinal (Fig. 12) have been studied and debated. Audio and video tapes are generally made to strongly favor the longitudinal direction, by means of particle orientation. Perpendicular recording has for the most part been implemented by the use of specialized thin-film materials [Ref. 35]; see, however, previous comments about barium ferrite. In principle, perpendicular recording could avoid the increase of self-demagnetizing effects with increasing transition density. In fact, the reverse would seem to occur as the magnetized regions become narrower with increasing density (Figs. 5 and 12). In practice, however, the deep, narrow recorded zones that would be needed to take advantage of this effect would be very difficult for an actual head to create. Perpendicular and longitudinal recording are therefore likely to be found essentially equivalent in their high-density capabilities [Refs. 36, 37].

Whatever the head design and the recording material's composition, the particle size is an extremely important parameter to system performance at high information densities. The need to have a large number of particles in each magnetized region of the recorded pattern (between adjacent transitions), in order to assure an adequate signal-to-noise ratio, has been discussed. As linear density increases, however, another consideration becomes important: the particle length (measured along the track) may become a limiting factor on the ability to record sharp, closely spaced transitions [Ref. 38]. For example, in the uncompressed HDTV system of Table 1, the minimum spacing between transitions is 0.345 mm [Ref. 15]. In this format, the use of acicular particles of 0.3 mm length (a size common in many analog videotapes) would cause some loss of signal output, even though the 37 mm track width and the particle width of about 0.03 mm would ensure that there are a large number of particles per magnetized region (between transitions). Particles having lengths of 0.2 mm would be more appropriate to the digital HDTV recorder. As was mentioned earlier, the need to reduce particle size as recording density increases must be balanced against the need for adequate magnetic stability [Ref. 38].

Head-To-Tape Spacing

A factor of enormous importance in all recording at high linear densities is the effective spacing between the head and the magnetic coating. This follows from the fact that the resolution with which the head can record and read back deteriorates rapidly as spacing increases. The effect on the recording process has been estimated as a loss of 44 dB per wavelength of spacing, for a sinusoidal signal [Ref. 39], and on the playback process as a loss of 55 dB per wavelength [Ref. 40]. This total of nearly 100 dB per wavelength means, for instance, that for sinusoidal recording at 50,000 transitions per inch (corresponding to a wavelength of 40 microinches or about one micrometer) every microinch of spacing will cause about 2.5 dB of overall loss in the reproduced signal. Obviously, variations in the spacing will amplitude-modulate the signal with the same sensitivity, and therefore add noise components to the signal. Thus, media surface smoothness may well be the ultimate limiting factor of magnetic recording density, although further development of heads and media are needed in order to reach this limit. Progress in this direction will demand continued careful attention to durability and friction, because of the increasingly close contact between head and recording surface.

In addition to smoothness, the freedom from flaws and contaminants is crucial, if information drop-outs, or momentary losses of signal, are to be minimized. As was mentioned earlier, techniques of error detection and correction (or concealment) allow satisfactory performance even if imperfections exist in the recording medium. Even a highly sophisticated error-correction system, however, can be overwhelmed by defects that are excessive in number or size. As the recording density increases, the size of the flaw needed to cause a given system failure (uncorrectable data loss, visible video defect, audible audio defect) decreases. Thus, the development of advanced magnetic recording media entails the search for the means of producing highly perfect surfaces, as well as ideal magnetic properties.

ACKNOWLEDGEMENT

The author wishes to thank C. A. Newman for help in preparing the manuscript.

REFERENCES

1. J. C. Mallinson, The Foundations of Magnetic Recording, San Diego, CA: Academic Press, 1987.
2. F. Jorgenson, The Complete Handbook of Magnetic Recording, TAB BOOKS, Inc., Blue Ridge Summit, Pennsylvania, 1981.
3. C. D. Mee, Magnetic Recording, New York, NY: McGraw-Hill, 1987.
4. E. N. Bertram, "Fundamentals of the Magnetic Recording Process," Proceedings of the IEEE, Vol. 74, pp. 1949–1512, Nov. 1986.
5. P. Vogelgesang, "Another Look at Recording Bias," Modern Recording & Music, pp. 36–43, Dec. 1980.

6. B. Blesser, "Digital Recording," db, The Sound Engineering Magazine, a series beginning with July, 1980. (The July, August and September issues give a good introduction to the basics.)
7. J. P. Watney, "Technical Choices for a Video Recorder," in Television Image Quality, J. G. Friedman, Ed., Society of Motion Picture and Television Engineers, Scarsdale, New York, 1984, pp. 127–136.
8. P. A. Dare and R. Katsumi, "Rotating Digital Audio Tape (R-DAT): A Format Overview," SMPTE Journal, Vol. 96, pp. 943–948, Oct. 1987.
9. J. R. Watkinson, "Error Correction Techniques in Digital Audio," Fifth International Conference on Video and Data Recording (April 2–5, 1984), Institution of Electrical and Radio Engineers, Proc. No. 59, 1984.
10. H. Kybett, "Video Tape Recorders," Howard W. Sams & Co., Inc., Indianapolis, 1978.
11. C. Chan and T. Eguchi, "Product Implementation of the 4:2:2 Component Digital Format," SMPTE Journal, Vol. 96, pp. 949–958, Oct. 1987.
12. E. Engberg, R. Brush, M. Lemoine, S. Magnusson, F. Morrison, D. Rodel, D. Ryan and J. Watney, "The Composite Digital Format and Its Applications," SMPTE Journal, Vol. 96, pp. 934–942, Oct. 1987.
13. R. Brush, "Design Considerations for the D-2 NTSC Composite DVTR," SMPTE Journal, Vol. 97, 182–193, March 1988.
14. S. B. Luitjens, "Magnetic Recording Trends: Media Developments and Future (Video) Recording Systems," IEEE Transactions On Magnetics , Vol. 26, pp. 6–11, Jan. 1990.
15. M. Umemote, Y. Eto, H. Katayama, and N. Ohwada, "Record and Playback Systems for 1.2 Gb/s HDTV Digital VTR," Proceedings of the Eighth International Conference on Video, Audio, and Data Recording (April 24–26, 1990), Institution of Electrical Engineers Conference Publication No. 319, pp. 54–59.
16. M. P. Sharrock, "Particulate Recording Media: A Review," IEEE Transactions on Magnetics, Vol. 25, pp. 4374–4389, Nov. 1989.
17. M. P. Sharrock, "Particulate Recording Media," MRS Bulletin, XV, pp. 53–61, March 1990.
18. J. H. Judy, "Thin Film Recording Media," MRS Bulletin, XV, pp. 63–72, March 1990.
19. K. Chiba, K. Sato, Y. Ebine, and T. Sasaki, "Metal Evaporated Tape for High Band 8-mm Video System," IEEE Transactions on Consumer Electronics, Vol. 35, pp. 421–427, Aug. 1989.
20. G. R. Cole, L. C. Bancroft, M. P. Chouinard, and J. C. McCloud, "Thermomagnetic Duplication of Chromium Dioxide Video Tape," IEEE Transactions on Magnetics, MAG-20, pp. 19–23, Jan. 1984.
21. O. Kubo, T. Ido, and H. Yokoyama, "Properties of Ba-Ferrite Particles for Perpendicular Magnetic Recording Media," IEEE Transactions on Magnetics, Vol. 18, pp. 1122–1124, Nov. 1982.
22. T. Fujiwara, "Magnetic Properties and Recording Characteristics of Barium Ferrite Media," IEEE Transactions on Magnetics, Vol. 23, pp. 3125–3130, Sept. 1987.
23. J. Takahashi, K. Itoh, and S. Ogawa, "Longitudinal Recording on a Rigid Disk Using Ba-ferrite Powder," IEEE Transactions on Magnetics, Vol. 22, pp. 713–715, Sept. 1986.
24. E. F. Wollack, R. M. Erkkila and T. C. Mitchell, "The Effect of Coercivity on Short Wavelength Output for Barium Ferrite Particulate Tapes," IEEE Transactions on Magnetics, Vol. 25, pp. 4054–4056, Sept. 1989.
25. K. Yamamori, T. Suzuki and T. Fujiwara, "High Density Recording Characteristics for Ba-ferrite Flexible Disks," IEEE Transactions on Magnetics, Vol. 22, pp. 1188–1190, Sept. 1986.
26. T. Nagaki, R. Chubachi, K. Odaka, Y. Shinpuku, T. Iijima, and K. Horikawa, "Magnetic Contact Duplication for R-DAT with Barium Ferrite Media," IEEE Transactions on Magnetics, Vol. 22, pp. 1191–1193, Sept. 1986.
27. T. Suzuki, T. Ito, M. Isshiki and N. Saito, "Barium Ferrite Tape for DAT Magnetic Contact Duplication," IEEE Transactions on Magnetics, Vol. 25, pp. 4060–4062, Sept. 1989.
28. Y. Okazaki, M. Noda, K. Hara and K. Ogisu, "Barium Ferrite Tape for R-DAT Magnetic Contact Duplication," IEEE Transactions on Magnetics, Vol. 25, pp. 4057–4059, Sept. 1989.
29. D. E. Speliotis, "Distinctive Characteristics of Barium Ferrite Media," IEEE Transactions on Magnetics, Vol. 23, 3143–3145, Sept. 1987.
30. M. P. Sharrock, "Magnetic Remanence Properties of Particulate Recording Media: Comparison of Barium Ferrite with Acicular Materials," IEEE Transactions on Magnetics, Vol. 24, 2856–2858, Nov. 1988.
31. J. C. Mallinson, "On Extremely High Density Digital Recording," IEEE Transactions on Magnetics, Vol. 10, pp. 368–373, June 1974.
32. H. N. Bertram, M. K. Stafford, and D. R. Mills, "The Print-Through Phenomenon," Journal Audio Engineering Society, Vol. 28, pp. 690–705, Oct. 1980.
33. F. Jeffers, "High-Density Magnetic Recording Heads," Proceedings IEEE, 74, pp. 1540–1556, Nov. 1986.
34. T. Kobayashi, M. Kubota, H. Satoh, T. Kumura, K. Yamauchi, and S. Takahashi, "A Tilted Sendust Sputtered Ferrite Video Head," IEEE Transactions on Magnetics, Vol. 21, pp. 1536–1538, Sept. 1985.
35. K. Ouchi and S. Iwasaki, "Perpendicular Magnetization Structure of Co-Cr Films," IEEE Transactions on Magnetics, Vol. 18, pp. 1110–1112, Nov. 1982.
36. D. J. Bromley, "A Comparison of Vertical and Longitudinal Magnetic Recording Based on Analytic Models," IEEE Transactions on Magnetics, Vol. 19, pp.2239–2244, Sept. 1983.

37. J. C. Mallinson and H. N. Bertram, "A Theoretical and Experimental Comparison of the Longitudinal and Vertical Modes of Magnetic Recording," IEEE Transactions on Magnetics, Vol. 20, pp. 461–467, May 1984.
38. M. P. Sharrock, "Time-Dependent Magnetic Phenomena and Particle-Size Effects in Recording Media," IEEE Transactions on Magnetics, Vol. 26, pp.193–197, Jan. 1990.
39. H. N. Bertram and R. Niedermeyer, "The Effect of Spacing on Demagnetization in Magnetic Recording," IEEE Transactions on Magnetics, Vol. 18, pp. 1206–1208, November 1982.
40. R. L. Wallace, Jr., "The Reproduction of Magnetically Recorded Signals," Bell System Technical Journal, Vol. 30, pp. 1145–1173, Oct. 1951.

5.4 Magnetic Recording Media

Part II

Ronald C. Reph
3M Company, St. Paul, Minnesota
and
Norman C. Ritter, (Retired)
3M Company, St. Paul, Minnesota

INTRODUCTION

Magnetic tape recording plays an important role in many information storage applications. People involved with recording and storage should be aware of the correct conditions and procedures that will ensure the best overall results. The purpose of this section is to provide some guidelines in achieving excellent tape performance and tape life. Although this information is targeted for the videotape recording field, it is applicable to other magnetic recording areas as well.

Magnetic recording has proven to be a very versatile and reliable form of information storage. The electromagnetic media used in recording tapes are capable of retaining signals for an indefinite period of time. The recorded signals will remain virtually unchanged for decades unless they are intentionally erased. Problems encountered with magnetic recordings are mostly physical in nature and result from improper handling and storage conditions. These problems can be caused by malfunction of recorders, improper operator care, or poor environmental conditions in either operations or storage. The following information is intended to help achieve optimum tape performance by elaborating on key issues with as little rhetoric as possible so as to simplify the readers' future reference to parts of this chapter.

THE RECORDING AND OPERATIONS AREA ENVIRONMENT

Debris from the environment can be the major cause of dropouts, scratches, machine part wear, and reduced head or tape life. As recording density increases, these factors become increasingly important and may often be overlooked or taken for granted in many operations.

1. The tape and machine rooms should approach "clean room" conditions, such as those found in mainframe computer installations. This degree of air filtration, though ideal, is not realistically achievable in many video tape operations. It is recommended the air filtration used must at least meet the efficiency rating of 90% based on the National Bureau of Standards Dust Spot Efficiency Test - Atmospheric Dust. Qualified air conditioning and air filtration firms are aware of these requirements. The airflow system should be designed to maintain a positive pressure in the recording area. Positive pressure prevents dust particles from floating into the room from other locations. The duct size and placement should allow the necessary amount of air flow to maintain the proper temperature and humidity control, without having air velocities sufficiently high to blow settled dust and debris around the room or on the tape transport area.
2. The temperature should be controlled at 70° F plus or minus five degrees, and the relative humidity at 50% plus or minus 20%. Without this control, the risk of head clogging, striction, and high headwear increases. The old cliche is still true: "What is comfortable for humans is a good environment for magnetic tape."
3. To avoid contamination of both tape and machines, smoking, eating, and drinking should be restricted to areas away from the recording locations.
4. Special attention must be given to the floors in or near the recording area, as they are a major source of debris due to pedestrian traffic. Cement flooring must be sealed and tile floors should not be waxed. Both should be mopped clean on a routine basis. The use of indoor-outdoor type carpeting that contains anti-static material is acceptable, depending on the needs of the operation. Carpeting helps to reduce room noise from equipment and affords an atmosphere conducive to good operator practices. The use of industrial vacuum cleaners

is recommended on carpeting or floors, and should be done on a regular basis. The exhaust from these units should be located outside the recording area.

5. Since people are a major source of debris, the recording area should be located such that it is not a normal passageway to other parts of the operation. This not only improves the cleanliness of the area but avoids continual distraction of the tape machine operator.
6. During times of construction in any area of the facility, the frequency of cleaning should be increased substantially. The construction area should be sealed off with plastic sheeting as well as possible. Cement dust seems to penetrate everything and can create very costly and time-consuming equipment repairs.
7. The tops of equipment and other exposed surfaces should be cleaned on a periodic basis with an electronic vacuum with static grounding and a fine-particle filter. Using a vacuum of this type will remove the small particles and capture them rather than spread them to another area. The amount of dust and debris found during this routine is an indication of the amount of dust and debris that may have found its way onto the tapes and into the recording equipment.

OPERATING PRACTICE RECOMMENDATIONS FOR ALL FORMATS

Videotape operations are usually involved with both reel-to-reel and cassette-based systems.

The factors affecting tape life and reliability can be related to the tapes' exposure to environmental conditions both on and off the equipment, who or what is handling it, and the tape path condition. The cassette approach virtually eliminates the human handling of the tape, and reduces the exposure to atmospheric contamination; however, the mechanical threading mechanisms require attention as does the condition of the tape path. The recommended environmental conditions apply to both type systems. Before proceeding further, some comments on the potential effects of debris may help to put things in perspective.

People involved in any videotape recording operation, at some points, may ask whether all the emphasis put on cleaning, handling, and environmental control really accomplishes anything; that is, in other terms, "familiarity breeds contempt." Fig. 1 gives a good picture of the relative sizes of what is being dealt with. Not only is this debris capable of causing large dropouts in the area where they are located on the tape; but some of the particles, when wound into a tape pack,

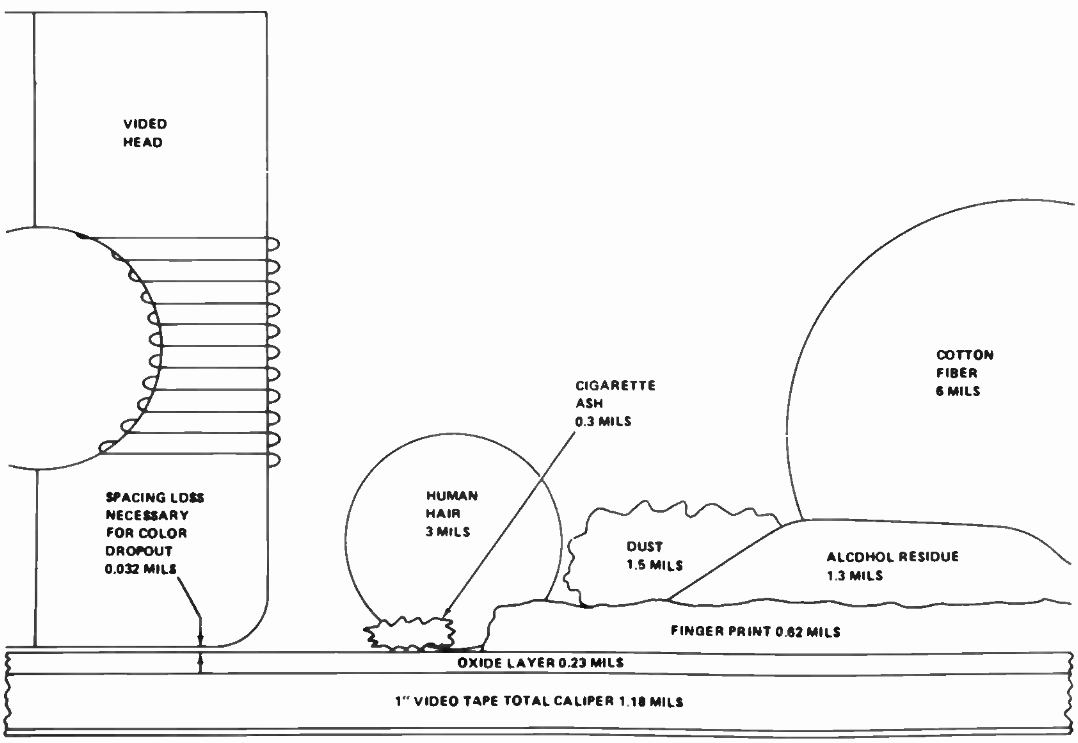


Figure 1. Debris perspective on 1" videotape.

may cause impressions into adjacent layers which in turn will results in additional drop-out activity.

Wallace's law on signal loss due to head to tape spacing theoretically indicates that compacted debris contained in a one pint jar, if equally distributed, would be sufficient to put every videotape every produced out of manufacturers' dropout specifications.

Good housekeeping and operations practices are very important to a successful videotape operation. The following are some operational recommendations that apply to the handling of reel-to-reel machines, the cleaning should be done before each tape is threaded onto the machine. With cassette systems, cleaning the tape path on a daily, weekly, or monthly basis is usually adequate, depending on the format and operation.

1. The recommended cleaning fluids are Freon TF (a DuPont trademark), Gensolve D (an Allied Chemical trademark), isopropyl alcohol, or other cleaners recommended by the machine manufacturer. As always with any chemical, please refer to the Material Safety Data Sheet for the safety, health, and environmental information. The cleaning should be accomplished by applying the liquid to a lint-free cloth, such as Texwipe (trademark), and gently rubbing the tape path surfaces with the dampened material. During this procedure extreme care should be exercised in cleaning the video heads. They can be easily broken or damaged with excessive pressure. Since the video heads will not withstand strong scrubbing actions, compared to the other transport areas during cleaning, the stronger solvents such as isopropyl alcohol, or machine manufacturers recommended solvents (usually Xylene type), are the best choices for removing debris build up on them. They tend to help soften material buildup in addition to acting as a wash.
2. The cleaning of the capstan and the capstan pinch roller are especially important for two reasons. Dirt or debris build up will cause impressions in the tape on each revolution that can cause dropouts through a long period into the roll. The accumulation of material may also reduce the frictional pressures on the tape which are needed to maintain proper linear tape speed control. One word of caution: the material used in the pinch roller can be adversely effected by some cleaning solutions. The machine manufacturers recommendations should be followed for cleaning the pinch roller.
3. Cleaners, other than Freon TF or Gensolve D, are relatively slow in evaporating, so allow ample time for the transport to dry before threading the tape. This usually means about 30 seconds.
4. The video drum (scanner assembly) on helical machines have tape edge guiding surfaces that may accumulate debris. These surfaces need extra cleaning attention to ensure smooth tape handling and good interchange with other recordings.
5. The condition of the pinch roller and its engagement with the capstan are very important in preventing creasing or edge damage to tapes. After much usage the pinch roller may become barrel-shaped, nonuniform, hardened, or misaligned with the capstan. A straight-edge applied against the pinch roller, along with proper lighting, is one method of determining the flatness of its surface. One align check, proven to be helpful, is watching the vertical direction the tape travels when the pinch roller engages with the capstan. During repeated stops and starts, as the pinch roller is engaged, the tape should not show signs of edge buckling on adjacent guides, if these parts are properly functioning.
6. The wear on capstans, pinch rollers, audio stacks, erase stacks, fixed guides (either front or backside), fixed drum assemblies, and roller guides that don't turn are key elements of the helical transport to consider if striction or runability problems occur during editing or other modes of operation. As the stationary components wear, the land area the tapes pass over increases. This increased surface area and a worn or polished drum all create additional drain on the tape movement through the transport. The required replacement frequency of these parts will vary with the type of video tape operation, but for machines dedicated to heavy editing service, these fixed parts will require replacement more often. As a guide, they should be considered for replacement after approximately 1,400 hours of use. An exception is the pinch roller which may require replacement on a more frequent basis.
7. The measurement and adjustment of the capstan pinch roller pressure against the capstan should be done after the machine has been running for approximately one hour.
8. The tape edge guiding surfaces in the transport should be periodically inspected for wear. If grooves are visible, the guides should be replaced.
9. The tension gauges used for making machine tension adjustments must be checked for proper calibration. The calibration check should be done with the same piece of tape that will be used during the machine tension measurement and adjustment. Tape thickness and other properties will affect the accuracy. Also, the calibration should be done with weights that will check the low and high ends of the scale.
10. The tapes and cassettes should be kept in their containers when not being used.
11. Cardboard fragments from opening the master shipping cartons is a common source of debris. This unpacking should be done outside the operation area.
12. Before attempting to use tapes that have just been exposed to environments considerably different than the operating environment, they should be given approximately 24 hours to stabilize in the operating environment.
13. The use of commercially available videotape cleaners on 2-inch, 1-inch, and 3/4-inch formats

has proven to be effective in extending tape life on old and used tapes, if these cleaning units are checked for proper alignment, tape tensions, and guide wear.

14. Reel-to-reel tapes and cassettes should be given a full length rewind to achieve a uniform and even wind before going to storage or preparing for shipment.

OPERATING PRACTICE RECOMMENDATIONS SPECIFIC FOR REEL-TO-REEL FORMATS

1. The tape reel is specially designed for transporting magnetic tape. The reel should always be handled by the hub, which is the strongest portion of the reel. (See Fig. 2.) The reel flanges are designed to protect the tape edges, not to guide the tape. A reel should never be carried by the flanges. Handling the tape by the reel flanges or dropping an unprotected reel can bend the flanges. If the tape is rubbing against the flange of the reel, either the flanges are bent, or the reel pedestal or guides require adjustment. Avoid squeezing the flanges of a reel when putting it on or taking it off the transport.
2. The take-up reels should be cleaned at the start of each day. The tape winding surface of the hub and the inside surfaces of the flanges are the main areas of the reel needing this attention.
3. The tape should not be touched or handled except at the ends, which is necessary for thread up on reel-to-reel machines.
4. The sudden stoppage of a spinning reel of tape should be avoided in order to prevent the possibility of interlayer slippage that results in cinching or windowing of the tape. (See Fig. 3.) Newer machines have a built-in control that slows the spinning tape motions before stopping or reversing wind modes. Machines without this feature require the operator to slow the winding action down before stopping or changing wind directions. The brake tension adjustments on the machine should also be checked according to the machine manufacturer's operating manual to minimize problems of this type.
5. The outer wrap of reel-to-reel tape, when not being used, should be taped down with hold down tab material designed for this use. Proper hold down tab material will not leave adhesive residue on the tape after removal.
6. When threading the reel-to-reel machines, do not let the end of the tape touch the floor or table tops,

Figure 2. Proper handling of a videotape reel.

Figure 3. Cinched tape. Note the complete fold over of one tape strand within the corrugated area.

in order to avoid pick-up of dust and debris. Also, any wrinkled or creased ends of the tape should be cut off. Do not fold the tape on the take-up reel when threading.

OPERATING PRACTICE RECOMMENDATIONS SPECIFIC FOR CASSETTE FORMATS

Cassettes have the desirable characteristic of eliminating possible damage from human handling and open exposure to some adverse environmental conditions. However, the cassettes have built-in hardware requirements that may cause problems. The following recommendations will help avoid these potential problems.

1. The construction of a video cassette requires that the tape be spliced to a leader at both ends of the tape. Each splice results in an extra thickness at that point in the overall tape and leader package. The cassette construction also requires that the lead be firmly fastened to the take-up hub at one end and to the supply hub at the other; the method of attachment is called *staking*. The staking results in a disruption of the smooth tape winding surface at both hubs. When a tape pack is allowed to lie wound-up on the irregular surfaces created by the splices and staking areas, the tape will acquire impressions that extend through several layers of the tape pack. The number of layers affected is dependent upon the tape thickness, the tightness of wind, the size of the irregularities, and the length of time the tape is wound onto these areas. The picture impairment that results may be described as a horizontal noise bar that rolls through the picture on a "once-around" basis. Because the tail end of the tape must be exposed to these imperfections, the tape manufacturer provides extra tape beyond that required for normal program time. This extra length avoids the need for using the layers at the tail end of the tape, which has the impressions from the splice and hub staking hardware. When fully rewound, as in the case of a new cassette, the front end of the tape pack is not exposed to these irregularities at the take-up hub end. Therefore, it is recommended that all cassettes be fully rewound before they are ejected from the machine.
2. If one minute of test signals are recorded ahead of the start of program, which is a recommended practice, the problem of head end staking and splice impressions that may occur when a tape is purposely or inadvertently not rewound after recording or playback are not likely to cause picture impairment reaching into the beginning of the program area.
3. A cassette should never be unloaded from the machine at a point that is in the program area and then left in that condition for an extended period of time. This recommendation is based on the fact that this area of tape will be tightly wrapped around the guide posts in the cassette. The resulting pressure against the guides causes the tape to acquire impressions from the guides. These impressions, resulting from what the manufacturers call a "cold flow" effect (the tape takes on a "set"), may cause the same kind of picture impairment as previously described for the staking and splice impressions.
4. With cassette machines as a major factor in the airing of commercials, a word of caution is appropriate for those systems that provide multiple use of commercials on one cassette. When a cassette tape has had repeated playback in a specific area of the tape, and then the tape is to be rerecorded over the areas that extend beyond the original message area, it is recommended that the tape be evaluated before being put back into use. The reason: it is possible that each end of the repeated play area on the tape has been exposed to previous stresses or debris build-up that will cause picture impairments when a new recording is made over those areas. It should also be recognized, however, that some formats have sufficient built-in error concealment or correction that will minimize the effects from these type of problems.

STORAGE ENVIRONMENT

1. Temperature and humidity should be controlled within the limits previously specified for the operations area; namely 70° F, ±5°, and a relative humidity of 50%, ±20%. This has been adequate for most operations and library needs; however, if archival storage times are being considered, it is recommended that the relative humidity be kept at 45% or lower and the temperature below 70° F.
2. The tape reel or cassettes should be kept in their original containers and stored on end to assure that the tape is being supported by the hub.
3. The air to this area should also be filtered to the same degree as previously recommended for the operations areas.
4. When tapes are removed from the library, the containers should be inspected for accumulation of dust or debris and wiped clean, if necessary, before being taken to the operations area.
5. The storage environment must be controlled, but it should be pointed out that most tapes are designed to withstand relatively short periods of extreme environments, such as those encountered during shipment; namely, −30° F to 120° F.

ACCIDENTAL ERASURE

Because magnetic tape has the advantage of being erasable and then reusable, the question regarding the threat of accidental erasure arises. The following information should alleviate those fears as the chances of accidental erasure are extremely remote.

1. The coercivities of videotapes are in the range of 300 oersteds to over 1,500 oersteds. Therefore, magnetic field strengths below 50 oersteds have little effect on the recorded tape. Most sources of magnetic fields are well below this level. As an example, the earth's magnetic field is approximately 0.6 oersteds. The field from an electric hand drill under full load is about 10 oersteds at the surface of the drill case.
2. The strength of a magnetic field drops dramatically as the distance from a typical field source increases. Therefore, the mere spacing provided by a shipping container separating magnetic fields offers considerable protection.
3. Permanent magnets used in holding various objects to steel surfaces can have field intensities of about 1,500 oersteds at the surface of their pole tips. These would present an erasability problem only when held extremely close to the reel of tape itself.

5.5 Audio Recording

Part I: Analog Magnetic Recording

James R. Carpenter
Broadcast Electronics, Quincy, Illinois

INTRODUCTION

The goal of any audio recording system is to store and reproduce sound as faithfully as possible. Analog magnetic tape is the storage medium on which audio signals are represented as variations in the magnetization level along the tape. As with any analog storage system, performance is limited by the maximum undistorted output level on one hand and noise level on the other.

The purpose of this first part of the chapter is to (1) provide some background on the development and principles of magnetic recording, and (2) discuss the operation and maintenance of magnetic audio recorders.

The second part of the chapter focuses on digital recording, including compact discs and digital audio recorders. Substantial portions of the practical material in the first part can be applied to the new technologies in the second part.

AN HISTORIC PERSPECTIVE

In 1897, Danish inventor Valdemar Poulsen invented a rudimentary magnetic wire recorder, which he called the "Telegraphone." Unable to market the concept successfully, he sold his Telegraphone patents in 1905.

In subsequent years, organizations in several countries slowly improved magnetic recording: progressing from steel wire and steel tape, to paper and celluloid tape. Even with significant improvements made in tape manufacturing, primarily by BASF and AEG, the audio quality of magnetic recording was very poor. Also, the poor fidelity caused by DC bias magnetic recording stalled the commercial acceptance of tape recording systems.

The German radio network, Reichs Rundfunk Gesellschaft (RRG), purchased several AEG Magnetophon recorders in the late 1930s to determine if audio recorder performance could be elevated to broadcast levels. While investigating an improved record amplifier circuit he designed, RRG engineer Walter Weber found performance incredibly improved. Low noise, good frequency response, and low distortion: true high fidelity from magnetic tape. Investigation revealed that his amplifier circuit was oscillating at a high frequency. Thus, in 1939, Walter Weber stumbled upon AC bias of magnetic tape quite by accident. In the same year, Marvin Camras independently discovered AC bias while working on improving wire recorders at the Armour Research Institute in Chicago. AC bias was employed on a Poulsen Telegraphone by U.S. Navy researchers Carpenter and Carlson sometime in 1921, but the project was dropped, and it took another 20 years for AC bias to be "rediscovered."

Late in 1947, the tiny Ampex corporation revolutionized the American recording industry by introducing their Model 200, the first truly commercially successful high-fidelity tape recorder. Since then, magnetic recording quality has steadily improved with better head design, tape formulations, and advanced electronic design.

TAPE CHARACTERISTICS AND TYPES

All magnetic tapes for audio recording are composed of magnetic particles that are attached to a plastic backing material with an adhesive binder. Variations in physical properties of the magnetic material, binder, and plastic backing result in the wide variety of tape types currently manufactured. The magnetic material itself largely determines tape electrical performance. The oxide or metal particles used in magnetic tape are actually microscopic needle-shaped bar magnets about 0.5 microns long and 0.05 microns wide. (A micron is 1×10^{-6} meters.) Each of these tiny particles produces a magnetic field of constant intensity. These particles can be thought of as binary switches which can exist in only two magnetic states. When an external magnetic field of opposite polarity is applied to an individual particle, that particle's magnetic polarity will switch at a certain field intensity. The magnetic field intensity which needs to be applied to a particle to make it switch states is called the *coercivity* of the particle. Coercivity is measured in oerstads and abbreviated H.

Not all particles in a batch of magnetic material have the same coercivity. The familiar bell shaped curve in Fig. 1A shows the overall distribution of particle coercivity in a magnetic tape.

A tape is considered erased when the magnetic particles are randomly oriented so there is no net polarization. That is, there is no external magnetic field produced by the tape, because all its tiny internal magnets are opposing each other. When an increasing external field is applied to an erased tape, nothing happens at first because the field is not strong enough to switch any of the particles. As the external field intensity increases, particles on the low end of the coercivity distribution curve start to switch. More and more particles switch as field intensity increases. When all of the particles are switched, increasing the external field intensity does not change magnetization. The tape is then said to be *saturated*.

The magnetic field retained by the tape after saturation is called the *retentivity*. Retentivity is measured in Gauss and abbreviated B. The higher the retentivity of a tape, the higher its output. When a certain width of tape is specified, retentivity is expressed as *remnance* or *fluxivity*, and is measured in nanowebers per meter (nW/m). This usage is the most common. There are two reasons why this is important.

First, the magnetization curves (or B/H curve) shown in Figs. 1B and 1C are very nonlinear, which means that something must be done in the recording process to reduce distortion and produce a linear output. That something is AC bias, which will be covered later. Second, different tapes have different values of coercivity and retentivity as listed in Fig. 1D, which affects how they can be used. Tapes with high coercivity require higher magnetic fields for recording and erasing. These tapes are called high-bias tapes. High-bias tapes cause compatibility problems with older equipment incapable of producing the large magnetic field strengths these tapes require. High-coercivity tapes are less subject to a phenomenon called self-erasure, which degrades high frequency response.

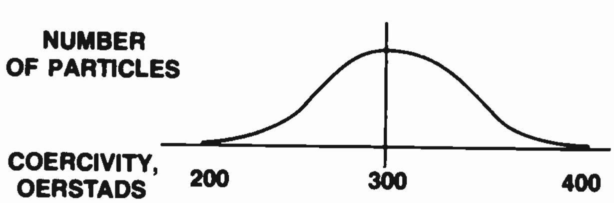


Figure 1A. Coercivity distribution of magnetic particles in a ferric oxide tape (Gaussian distribution).

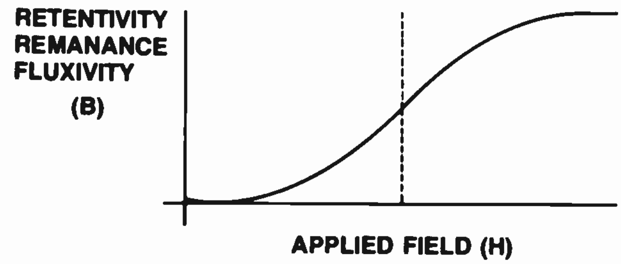


Figure 1B. Magnetization curve of a sample ferric tape.

The compact cassette, with its slow tape speed and relatively poor high-frequency performance, led to the development of chromium dioxide and metal particle tapes, which have higher coercivity than conventional ferric oxide tapes. These new tape types have not been a problem in consumer applications, as models and features are changed with the seasons. Broadcast equipment on the other hand lasts much longer, and most installations prefer to standardize on one type of tape so that recorders do not have to be constantly re-equalized and re-aligned. Improvements in audio mastering tape for broadcast and recording applications have been directed toward increasing retentivity, which results in higher output level and thus lower noise. High-coercivity tapes have become more popular as more consumer and professional tape machines are updated to handle increased magnetic field intensities.

Tapes used in reel-to-reel and cartridge machines are usually ferric oxide types with little or no cobalt doping. The cassette format generally uses chromium dioxide or doped ferric oxide tapes, with a trend toward the metal particle tapes for high quality recording.

Caution should be exercised concerning the use of chromium dioxide tapes for archival storage of program material. The low Curie temperature (see Fig. 1D) of

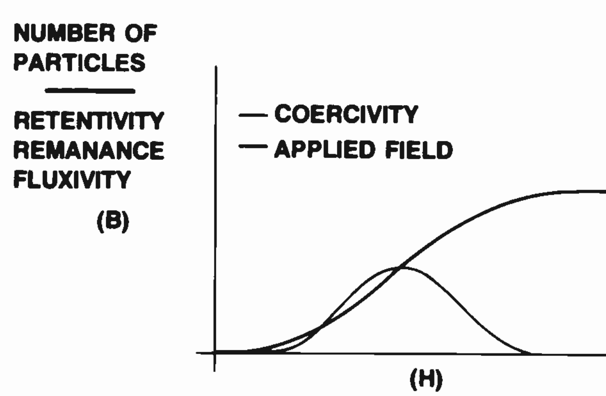


Figure 1C. DC magnetization curve of the tape is derived by summing all the contributions of the individual particle coercivities and retentivities. More correctly, the DC magnetization curve of a tape is the integral (area sum) of the coercivity distribution of the individual particles.

TAPE TYPE	RETENTIVITY	COERCIVITY	CURIE POINT
Ferric Oxide	1200 Gauss	300 Oerstads	470 C
Ferric Oxide Cobalt Doped	1500 Gauss	600 Oerstads	470 C
Chromium Dioxide	1500 Gauss	1500 Oerstads	130 C
Metal Particle	3000 Gauss	1100 Oerstads	770 C

Figure 1D. Performance of common magnetic tape materials.

this magnetic material means that this type of tape may be more affected by heat. A controlled temperature environment is important for storage of all magnetic recording materials.

In summary, the overall performance of magnetic recording is limited by the tape itself. Maximum output level is limited by tape retentivity (fluxivity) after saturation, and noise is limited by the magnetic particle size. Particle size cannot be arbitrarily reduced to improve noise. Thermal energy would tend to randomize the orientation of very small particles, effectively erasing the tape. Also, electrical noise generated in the tape heads and reproducing circuitry is just a few dB below the noise level of modern tapes. Tape development has focused on increasing retentivity to increase output level, thus improving signal-to-noise ratio. High-coercivity tapes improve high frequency response by reducing self-erasure.

Recording System Performance

For any given type of tape, the physical track width recorded on the tape and the speed at which it is pulled across the heads will determine the best possible signal-to-noise ratio. Signal-to-noise performance improves approximately 3 dB for every doubling of the recorded area. Increasing tape speed and increasing track width improve signal-to-noise ratio. Wide track tape running fast will outperform narrow track tape running slowly.

The chart in Fig. 2 depicts several tape types and their relative best-case performance. This chart will help evaluate different tape systems to determine which one suits a particular recording need. Maximum level is considered to be the point where total harmonic distortion reaches 3%.

THE NEED FOR BIAS

The need for bias is best explained by looking again at Fig. 1B, which shows the magnetization characteristics of the tape. The curve shown is very nonlinear, indicating the severe distortion that would be produced if some type of bias were not used. DC bias would reduce distortion by operating the tape in the linear portion of the transfer characteristic, but the limit to this technique seems to be about 5% total harmonic distortion (THD). As described above, it was accidentally discovered that imposing a high-frequency AC bias signal on the audio waveform improved performance tremendously.

AC bias is a sine wave frequency about five to ten times the highest frequency to be recorded, and the level is several times the peak amplitude of the audio signal to be recorded. With no audio signal applied, the bias waveform alternately flips the polarity of the

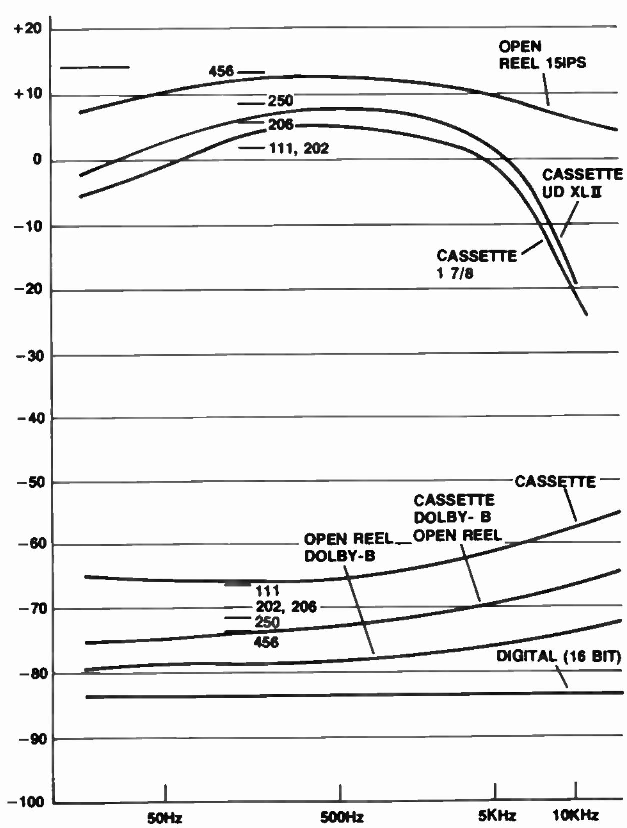


Figure 2. Maximum output level and noise level vs. frequency for common tapes and formats.

tiny magnetic particles as they pass the record head. The field strength decays with distance from the trailing edge of the record head gap, resulting in a random distribution of magnetic particle orientation and no net magnetization on the tape. When audio is applied, the bias waveform is offset by the audio signal amplitude, producing a net magnetization of the tape at any given point. When the tape is played back, the audio will be recovered via those magnetization changes, which vary in frequency and intensity according to the original input signal.

Optimum Bias

Tape types have different bias requirements depending on their coercivity. If the bias field is too small, the tape is said to be underbiased. If the bias field is too large, the tape is said to be overbiased. When bias is reduced below its optimum level (underbias), distortion rises dramatically, noise increases, and high-frequency response has noticeable distortion. When bias is increased beyond its optimum value (overbias), overall sensitivity drops, high frequency response is degraded, noise level drops, and distortion increases slightly. Fig. 3 illustrates the bias conditions. Optimum bias can be determined by using a distortion analyzer. Bias should be adjusted for minimum harmonic distortion in the range of 1 kHz to 6 kHz. Since bias affects high frequency response, equalization must be adjusted after determining optimum bias. When setting up machines, *it is safer to set bias slightly higher than the optimum* for the tape used to make measurements. Then, normal variations in tape composition from run to run will not cause an underbias condition, resulting in excessive distortion.

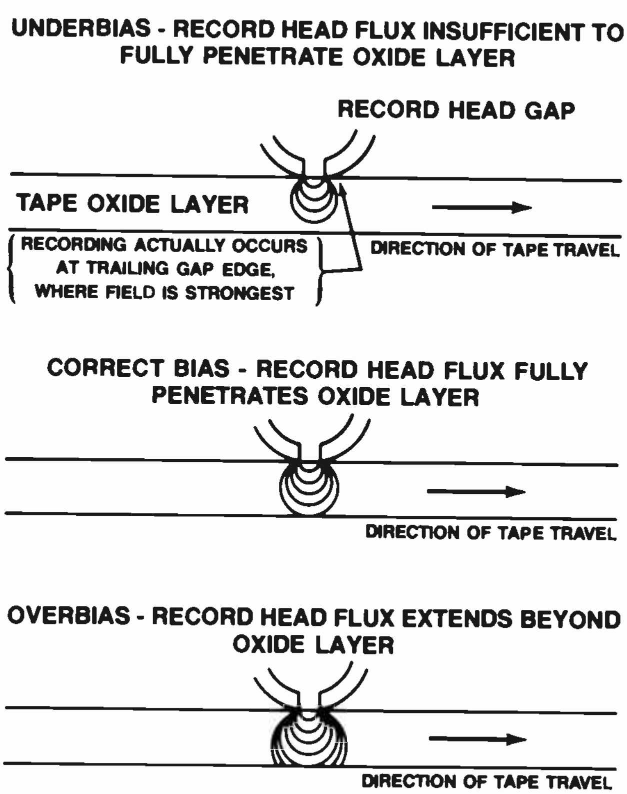


Figure 3. Bias field penetration.

Self Erasure

The magnetic field at the trailing edge of the record gap diminishes rapidly with distance. However, at very short recorded wavelengths (high frequencies, slow tape speeds), the field is often strong enough to cause partial erasure of information already past the gap. This effect results in high frequency loss and increased distortion. High bias (high coercivity) tapes are less susceptible to self erasure. This means that less equalization is required during recording, improving signal-to-noise ratio and high frequency response. For example, the compact cassette format specifies 70 μsec equalization for chromium dioxide tapes, and 120 μsec for ferric oxide tapes. Chromium tapes require less high frequency equalization on playback, improving signal-to-noise ratio.

THE PLAYBACK PROCESS

Electrical and mechanical design of the playback head is much more critical than the record head. Since all recording is done on the trailing edge of the record head gap, the gap width is not as important as for a reproduce head. The only critical requirement is that the combination of bias and audio waveforms do not saturate the magnetic material in the record head.

In playback, gap width is critical because it determines the highest frequency that can be reproduced. Making the gap width extremely small reduces the sensitivity of the head. All tape heads have electrical resistance and magnetic reluctance, which produce thermal noise. As the gap width is reduced, output level will drop and signal-to-noise ratio will be degraded past the point where the tape itself is the limiting factor. When the recorded wavelength is equal to the width of the playback head gap, the *net* magnetic field inside the gap will be zero and there will be no output voltage.

Output voltage begins dropping long before the recorded wavelength is equal to the playback-head gap width. This effect is called *gap loss*. A good rule of thumb is that the playback head gap should be from ¼ to ⅒ the shortest recorded wavelength, resulting in a gap loss of 0.9 dB and 0.14 dB, respectively. The following formula may be used to calculate the gap loss knowing the gap width and recorded wavelength.

$$\textit{Gap loss}\ (\mathrm{dB}) = 20 \log \frac{\sin(180\, g/l)}{\pi\, g/l}$$

where:
g = Gap length in inches or microns
l = Recorded wavelength in inches or microns

Recorded wavelength is determined by the following relationship:

$$Wavelength\,(W) = \frac{Velocity\,(V)}{Frequency\,(F)}$$

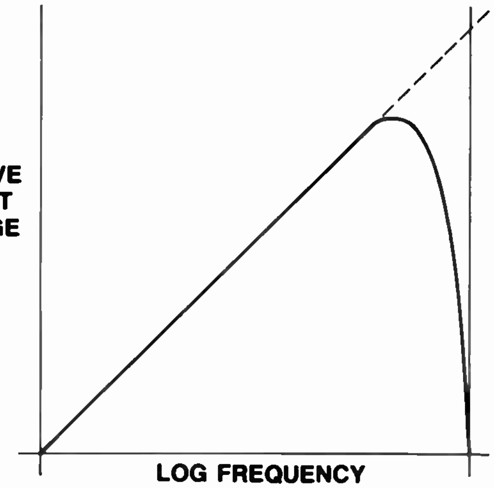
For this application, tape speed is V and audio frequency is F. Thus,

$$Recorded\ wavelength\,(in/cy) = \frac{Tape\ speed\,(in/sec)}{Audio\ frequency\,(cy/sec)}$$

Wavelength drops as audio frequency increases or tape speed decreases. The wavelength recorded on tape, for an audio frequency of 15 kHz and different tape speeds, is shown below (multiply inches by 2.54×10^4 to obtain microns).

TAPE SPEED	RECORDED WAVELENGTH FOR 15 KHZ	
(IPS)	(Inches)	(Microns)
1.875	0.000125	3.175
3.750	0.000250	6.350
7.500	0.000500	12.700
15.000	0.001000	25.400
30.000	0.002000	50.800

Therefore, head gaps must be made smaller as tape speed drops, leading to inevitable compromises in design that can result in poor high-frequency performance at low tape speeds.



OUTPUT VOLTAGE FROM IDEAL PLAYBACK HEAD FOR CONSTANT RECORD FLUXIVITY. DOTTED LINE IS THEORETICAL RESPONSE, SOLID LINE TAKES INTO ACCOUNT GAP LOSS.

Figure 4. Playback head response.

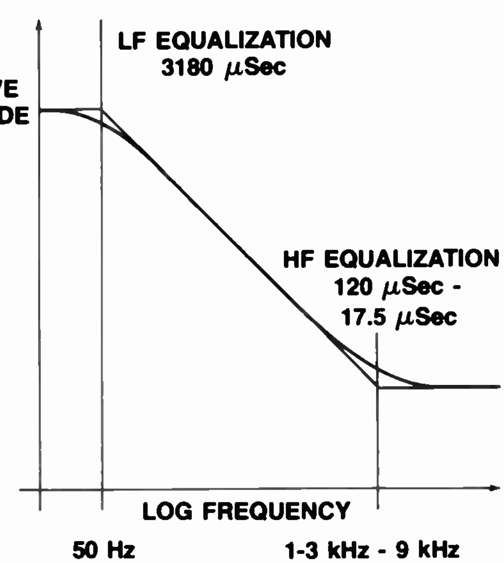
EQUALIZATION

The output voltage of a perfect playback head is directly proportional to the rate of flux change in its gap. The flux change is caused by the changing magnetization of the tape. For constant recorded flux intensity, the rate of change of flux increases with increasing frequency, producing an output voltage that increases linearly with increasing frequency. This rising response, along with the high frequency roll-off due to gap loss, is shown in Fig. 4.

In order to obtain flat playback response, a complementary equalization curve must be employed in the playback amplifier to boost the low frequencies. A playback equalization curve with a 50 Hz turnover frequency is widely used. This is also referred to as the 3180 μsec curve. If the record and playback heads were perfect, and in perfect contact with an ideal tape, no other equalization would be necessary.

Tape self-erasure, head electrical resistance, inductance, and magnetic reluctance, along with imperfect head-to-tape contact, result in the need to provide additional record and playback equalization.

Equalization is specified as a playback response curve for the total head, amplifier, and equalizer system. High frequencies are effectively boosted in playback by adding a high frequency RC time constant to the normal 3180 μsec (50 Hz) RC time constant. What really happens is the roll-off provided by the 3180 μsec time constant is stopped by this high frequency RC time constant. The equalization frequency varies from 9 kHz to 1.3 kHz according to tape speed as shown in Fig. 5A. Lower tape speeds require

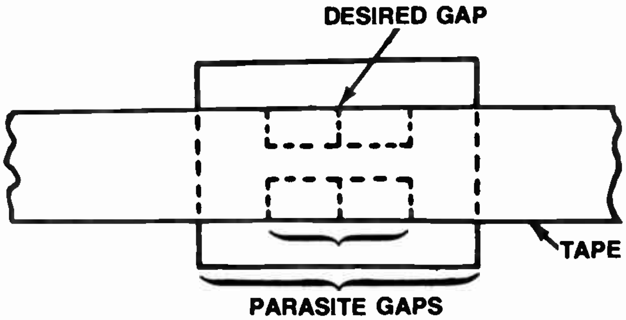


PLAYBACK EQUALIZATION CURVES SHOWING 3180 μSec LF TIME CONSTANT. HF TIME CONSTANT VARIES WITH TAPE SPEED, FORMAT, AND TAPE TYPE.

Figure 5A. Ideal playback equalization curves.

SPEED	U.S.		EUROPE	
1/4" open reel	LF	HF	LF	HF
30.00 ips	3180	17.5	3180	35.0
15.00 ips	3180	50.0	3180	35.0
7.500 ips	3180	50.0	3180	70.0
3.750 ips	3180	90.0		
1.875 ips	3180	90.0		
Cassette format				
1.875 ips	3180	120.0	ferric oxide tapes	
1.875 ips	3180	70.0	chromium ioxide and metal tapes	

Figure 5B. Magnetic tape playback equalization time constants in microseconds.



POLE PIECE AND CASE LENGTH PRODUCE PARASITIC MAGNETIC GAPS WHICH CAUSE LOW FREQUENCY CONTOUR EFFECTS, COMMONLY CALLED HEAD BUMPS.

Figure 6. Playback head contour effects.

more equalization and so have longer time constants (equalization starts at a lower frequency).

These system time constants and the equalization curves they represent are all referenced to constant fluxivity versus frequency in the playback head gap. These values are for gamma ferric oxide tapes. Some of these values may no longer be used, since newer high-energy tapes do not require as much high-frequency equalization as ferric tapes. The 3180 μsec LF time constant is widely used, although it is not a recognized standard. Consult the tape electrical data sheet, available from the tape manufacturer, for suggested equalization settings. *Because of the variations in tape formulations, establishing a standard brand and type for use throughout the station or recording facility will simplify setup and maintenance and give more uniform results.*

Since equalization has been standardized somewhat, it can be generally ignored, except when setting up and checking the performance of a recorder-reproducer. In order to accomplish this task properly, the recorder-reproducer must be properly aligned mechanically, degaussed and adjusted for flat playback response using a standard test tape in good condition. Once the playback has been correctly equalized, the recorder equalizer is then adjusted to produce flat playback response. This is easily done with three-head machines, as the play head can be monitored as the record head is equalized.

When using cassette equipment, the tape type must be noted because of the different equalization curves employed. A chromium tape (70 μsec) played back on a machine set up for ferric tape (120 μsec) will sound overly bright due to the resulting 50 μsec overall preemphasis (3 dB boost at 3 kHz rising to 10 dB boost at 10 kHz). Conversely a ferric tape played on a machine set up for chromium tapes will sound dull, being 3 dB down at 3 kHz and 10 dB down at 10 kHz. Some cassette players have automatic playback equalization switching keyed by a cutout in the tape housing.

Playback Head Contour Effects

There is a limitation on the lowest frequency that can be reproduced by a playback head. As frequency drops, the recorded wavelength increases and becomes large compared to the length of the entire playback head structure. The playback head structure, including the pole pieces and the case itself, becomes a parasitic gap as illustrated in Fig. 6. Low frequency flux cuts this parasitic gap and produces an output that adds to or subtracts from the desired gap output, depending on frequency. The result is a series of low frequency response dips and peaks called head bumps. The pole piece length can be made very large, moving the effect out of the audio frequency range.

Contour effects are most pronounced on tape machines which have insufficient space for heads with long pole pieces. Professional reel-to-reel tape machines running at high speeds have pole pieces with faces in excess of one inch long. Special contour reproduce heads have been developed by several manufacturers for cartridge machines. These heads, available since the late 1970s, largely overcome contour effects.

MECHANICAL ALIGNMENT

Correct mechanical alignment of magnetic recorders is absolutely necessary for proper electrical performance. A mechanical alignment should be performed or at least checked to insure that the tape contacts the head in the correct orientation. Head orientation is specified in terms of height, *zenith*, *azimuth*, and *meridian*. Figs. 7A, 7B, and 7C show each type of physical orientation.

Azimuth is the most critical of all the mechanical adjustments, and the most difficult to hold correct in day-to-day operations. When record azimuth is correctly set, the tape will be magnetized in vertical bars exactly perpendicular to the direction of travel. At high frequencies and slow tape speeds, the width of these vertical bars is extremely narrow. Any departure from perfect azimuth in recording or playback will

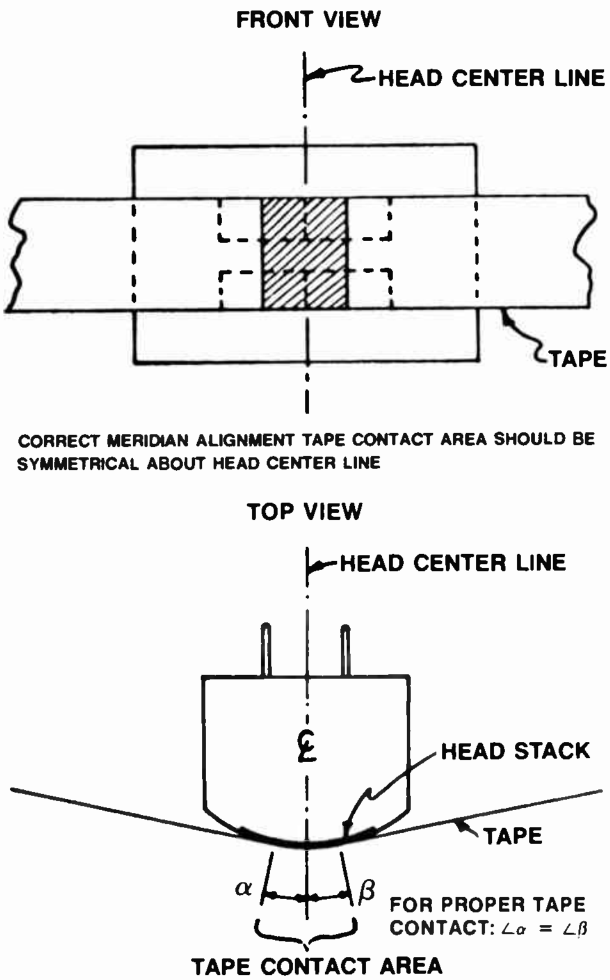


Figure 7A. Mechanical alignment—meridian adjustment.

cause the tape head gap to cut across one or more of these magnetized bars, as shown in Fig. 7D, resulting in reduced high-frequency output. This azimuth loss effect increases with increasing track width and decreasing wavelength. It can be calculated using the following formula:

$$Azimuth\ Loss = 20 \log \frac{\sin 180\, w \tan(a/l)}{\pi\, w \tan(a/l)}$$

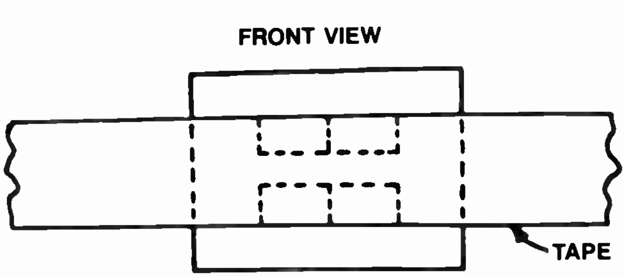


Figure 7B. Mechanical alignment—height adjustment.

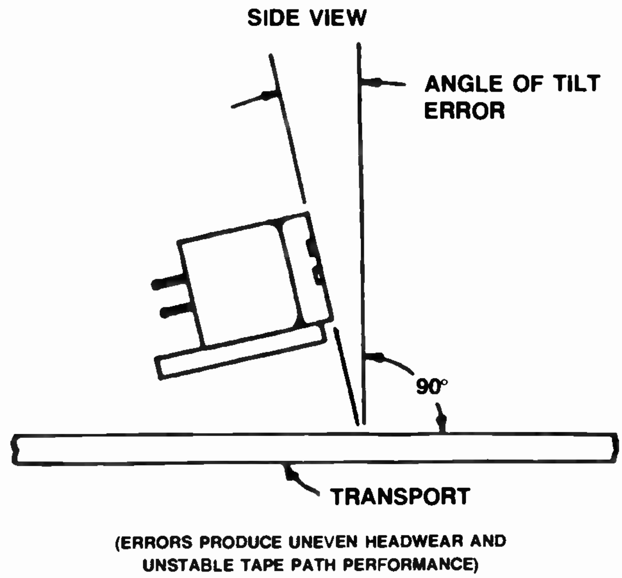


Figure 7C. Mechanical alignment—zenith adjustment.

where:
w = Track width in inches
l = Recorded wavelength in inches
a = Azimuth error in degrees

Azimuth errors of 0.5° and 0.1° will cause the following response roll-offs at 15 kHz:

Tape Format and Speed	*0.5°*	*0.1°*
1.875 ips stereo cassette (track width 0.0235 in)	15.1 dB	1.6 dB
3.75 ips ¼ in 2 track (track width 0.075 in)	18.9 dB	4.3 dB
3.75 ips 1.4 in 4 track (track width 0.043 in)	13.5 dB	1.3 dB
7.5 ips ¼ in 2 track (track width 0.075 in)	13.9 dB	1.0 dB
7.5 ips ¼ 4 track (track width 0.043 in)	10.5 dB	0.32 dB
15 ips ¼ in 2 track (track width 0.075 in)	7.3 dB	0.25 dB
15 ips ¼ in 4 track (track width 0.043 in)	2.1 dB	0.08 dB

There is no substitute for high tape speed when good high-frequency performance is desired. The relatively good performance of the cassette format is due to its narrow track width. Most quality recorders can hold azimuth error angle to less than 0.1°.

In stereo tape machines, azimuth error produces an additional high frequency effect called *mono sum loss*. Azimuth error in a stereo machine produces a slight time delay between the two channels. This time delay causes a corresponding phase shift between audio in the left and right channels. If the phase shift reaches 180 degrees at some frequency, and the two channels

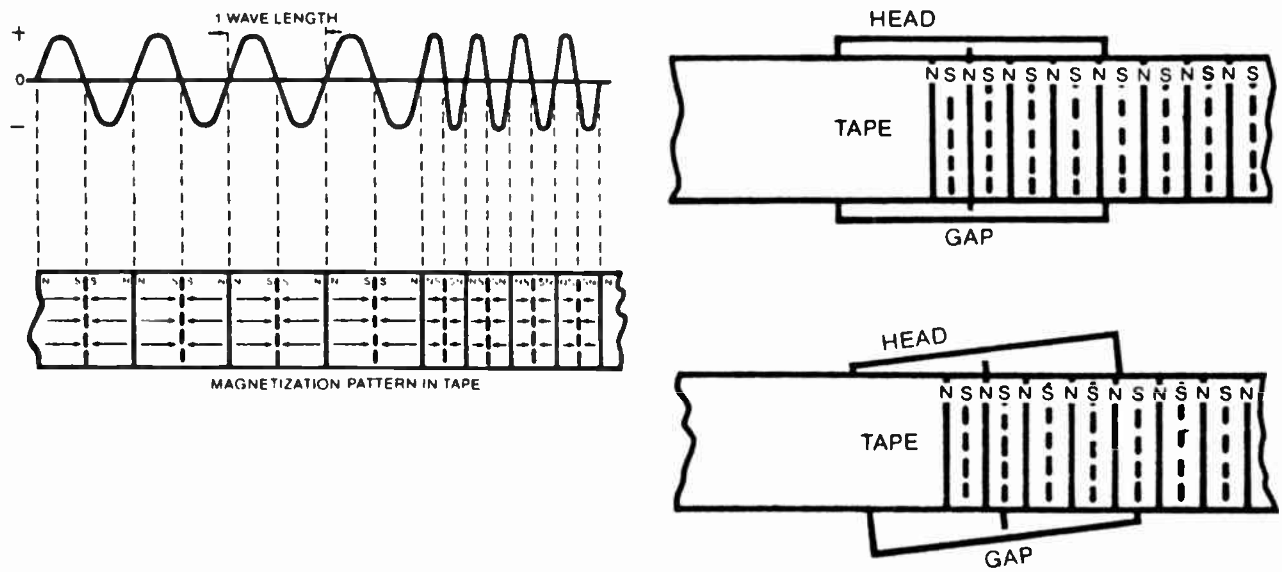


Figure 7D. Mechanical alignment—incorrect azimuth can cause the head gap to overlap enough of a high frequency cycle, or cycles, to result in some degree of cancellation.

are electrically summed to mono, total cancellation will occur, creating a response notch.

SYSTEM PERFORMANCE SUMMARY

Tape performance is determined by the tape type itself, the track width, and the speed. High bias (high coercivity) tapes are less susceptible to self erasure than ferric oxide tapes, and so provide better high-frequency performance. In addition, they have higher output due to higher retentivity (remnance, fluxivity), which results in better signal-to-noise ratios. All other parameters being equal, doubling the track width will improve signal-to-noise ratio 3 dB. Increasing track width, however, results in greater azimuth loss given a fixed azimuth error angle. Increasing speed improves high frequency performance by reducing azimuth loss and gap loss. Increasing speed degrades low frequency performance due to head contour effects.

Improving System Performance

Analog tape performance has steadily improved over the years, but the signal-to-noise ratio problem (tape hiss) is still a concern. Increasing the recorded level on the tape to improve the signal-to-noise ratio usually leads to overload, especially on high-frequency content program material that results in a harsh sibilant sound. Several *noise reduction* systems have been developed to improve tape system noise performance and are available in two configurations. The most effective is the encode-decode system that processes the audio signal during recording and playback. A single-ended noise reduction system operates on playback only, eliminating the need for encoding. Dolby™ and dBx™ are the trade names of encode-decode systems.

Encode-Decode Noise Reduction Systems

The Dolby noise reduction system was the first widely available electronic system designed specifically to reduce tape noise. The basic principle is to boost the high-frequency content of low passages during recording. On playback, a complementary high-frequency reduction is employed to restore correct frequency response. The end result is a 10 dB improvement in signal-to-noise ratio (A system).

There are four different Dolby systems for specific tape noise reduction applications. DolbySR™ is the newest noise reduction system. Its goal is to tailor the signal to the overload characteristics of the medium. This system has characteristics common with the other Dolby noise reduction systems, but takes them a step farther. Dolby A is a four-band system designed primarily for use in recording studios. Dolby B™ is a two-band system that has found universal acceptance in consumer recording equipment due to its simple setup and good performance. Dolby C™ is similar to, but provides more correction than, Dolby B.

The dBx noise reduction system uses complementary compression and expansion to improve tape signal-to-noise ratio. Signals are compressed 2:1 before recording. This means that a 20 dB input level change is transformed into a 10 dB level change on the tape. During playback, the recorded signal is passed through a 2:1 expander. This means that the 10 dB level change on the tape is transformed back to the original 20 dB level change. In addition to complementary compression and expansion, the dBx system uses preemphasis and deemphasis to further improve signal-to-noise ratio. Theoretically, a recorder with 60 dB dynamic range could be used to make recordings with an effective 120 dB dynamic range.

Single-Ended Noise Reduction Systems

Single-ended tape noise reduction systems, which are used only during playback, rely on the principle of auditory masking to reduce noise. High-frequency noise and noise when no signal is present are the most audible forms of tape noise. Since tape noise is directly proportional to bandwidth, single-ended noise reduction systems reduce playback bandwidth during quiet passages when no high-frequency program information is present. DNR™ and Dynafex™ are the two most popular single-ended noise reduction systems currently in use. The Dynafex system also uses additional amplitude expansion, an audio processing technique, to reduce playback noise.

Modulation Noise

All tape noise reduction techniques are limited by a phenomenon called *modulation noise*. The surface of magnetic tape is not perfectly smooth. Irregularities in the tape surface produce a noise which is related to the signal. If one were to record a single tone and then filter out that tone on playback, one would notice that playback noise increases with increasing recorded level. This modulation noise is usually masked by the signal, and is not noticed. Tape noise reduction systems make modulation noise more noticeable. In general, the deleterious effects increase with increasing noise reduction. Additionally, scrape flutter components produced by idler wheels, tape stiction, tape guides, reel flanges, and even pressure pads can contribute to modulation noise components.

THE FUTURE OF ANALOG RECORDING

Sales of prerecorded compact tape cassettes exceeded sales of LP records in 1982, about 15 years after the cassette became a popular storage medium. Compact discs now rival the sales of cassettes. The low-end rotary-head digital audio tape (DAT) machine, while popular with professional users, does not yet enjoy wide consumer acceptance due to its high cost. The higher cost of full-featured digital audio tape recorders, suitable for broadcast and recording applications, is currently holding back their universal acceptance. A look back to Fig. 2 will reveal that present tape formulations, when properly used, combined with noise reduction systems can provide analog tape systems with a signal-to-noise ratio within a few dB of the 16 bit (96 dB) digital audio recording process. Since the best broadcast channel signal-to-noise ratios are presently about 80 dB, analog tape will continue to provide excellent cost-effective service for many more years.

TAPE MACHINE ELECTRONICS

Although the tape machine is a complicated electromechanical system, the main electrical functions of any tape machine can be broken down into four main areas:

- Bias
- Equalization
- Control
- Interfacing

Bias

A large AC bias current is mixed with the signal to be recorded on the tape to permit low-distortion recording (discussed earlier). The bias signal is an ultrasonic (60 kHz to 200 kHz), low-distortion sinewave, usually generated by a digital divider and low-pass filter in modern equipment, or a cross-coupled transistor oscillator in older equipment. The bias signal is then mixed with the audio signal, either actively or passively, and applied to the record head. In passive mixing, a parallel resonant circuit must be employed to isolate the bias signal from the remainder of the audio circuitry. With active bias mixing, the isolating property of an op-amp summing junction, as shown in Fig. 8, is used to add the bias signal, thus reducing or eliminating the need for a bias trap.

The adjustment of bias current is usually performed with a potentiometer in series with the bias line. There are several designs which vary the bias current according to the high-frequency content in the program material, in order to squeeze the last possible amount of head-room from the tape in use. Since high-frequency audio material also acts as a bias signal, the ultrasonic bias can be reduced in proportion to the amount of high-frequency audio material present. This allows the high frequencies to be recorded at a higher level than would be possible with a fixed bias system and reduces the possibility of tape saturation.

Equalization

Equalization is the electrical alteration of the amplitude versus frequency characteristic to compensate for changes in frequency response due to various physical factors. In the absence of equalization circuitry, the response of a tape machine would be the inverted "U" caused by the combination of severe bass and treble loss inherent in the response of magnetic tape heads, as shown earlier in Fig. 4. The factors responsible include the increasing output of magnetic playback head with increasing frequency (6 dB/octave) low-

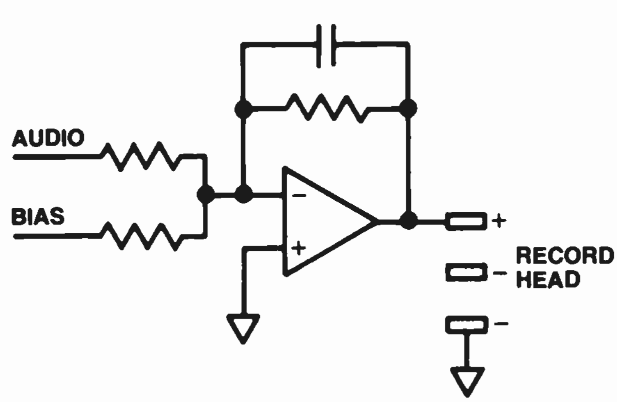


Figure 8. Active bias summing.

frequency "contour" effects, and tape self-demagnetization. The equalization circuits of a tape machine create the necessary compensating bass boost and treble reduction (shown in Fig. 5) to achieve flat audio response at the machine's output. Since the output voltage of a typical head is only 1 millivolt at 1 kHz, the equalization stages also provide some of all of the 60 dB of gain necessary to drive the output circuitry. The low head output voltage also requires that a low noise, high speed circuit be used as the equalization amplifier.

A typical circuit used to generate this playback response curve (NAB standard curve) for a cartridge machine is shown in Fig. 9. The curves are usually defined in terms of the time constants (or turnover frequencies) of the filters which would result in the desired amplitude response. The relationship between time constants and turnover frequencies is:

$$F = \frac{1}{2\pi T}$$

where:
F = Frequency in Hz
T = Time in seconds

For example, 50 microseconds (50×10^{-6} sec) is equivalent to 3180 Hz. In Fig. 9, the R3/C1 values set the low-frequency turnover and R2/C1 sets the high-frequency turnover. Both the high and low-frequency turnovers allow a small adjustment range to compensate for variations in head performance. The gain of this circuit is about 40 dB at 1 kHz.

Control

The internal control of tape machines is often accomplished with digital logic in either discrete or microprocessor form, instead of the mechanical or relay logic found in older tape machines. The amount of control included is dependent on the sophistication of the machine and the intended applications. All tape machines require the following control logic functions:

- PLAY
- STOP
- RECORD

The following functions are included in reel-to-reel and cassette machines and some cartridge machines:

- FAST FORWARD
- REWIND
- TAPE SPEED SELECTION
- EQUALIZATION/BIAS SELECTION
- NOISE REDUCTION SELECTION

Interfacing

Interfacing tape machines to the rest of a station or recording facility can be divided into two basic areas: audio and remote control.

Audio

The nominal audio input level is in the range of 1 volt (0 dBm or 1 mW is 0.744 volts at 600 ohms and +8 dBm is 1.95 volts). Tape recorders normally accept a large range of audio input amplitudes, usually ±20 dB, without overload and with excellent common mode rejection. Two types of recorder input circuits used are (1) the transformer isolated circuit shown in Fig. 10, and (2) the balanced differential instrumentation amplifier (op-amp) circuit shown in Fig. 11. With the low cost of integrated circuits, an excellent transformer coupled circuit can be more expensive to produce than an excellent balanced differential circuit. Also, at high input levels, low-frequency distortion can be a problem in transformer coupled circuits.

Three types of instrumentation amplifier circuits are frequently used as audio input circuits. As in most other applications, the most expensive three op-amp circuit has the best performance. In the circuit shown in Fig. 11A, the input impedance is low and unequal, the *common mode rejection* (CMR) is low, and is dependent on the balance of the source's output impedance and the matching of all four resistors. Also, two resistors must be changed to vary the gain.

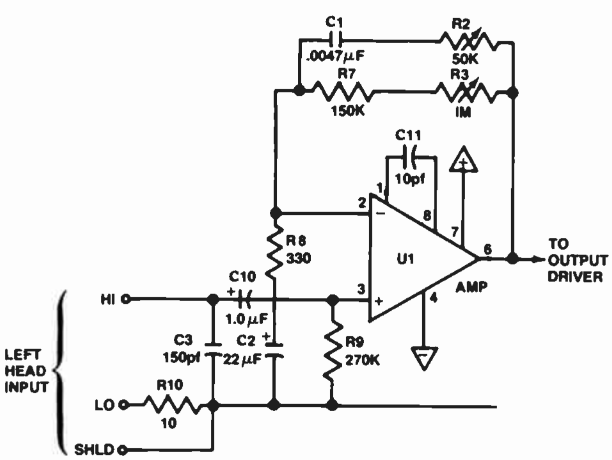


Figure 9. NAB playback equalizer circuit.

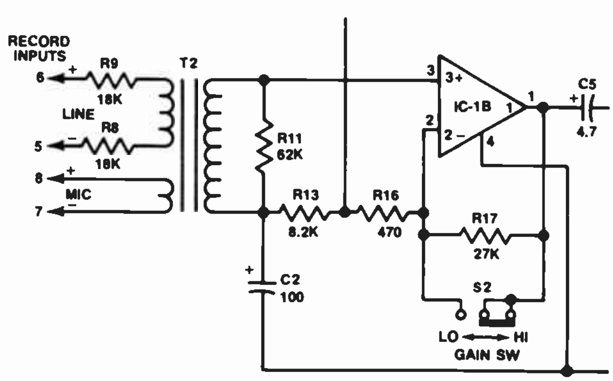
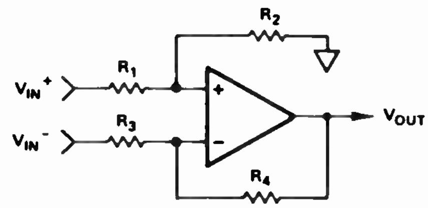


Figure 10. Transformer coupled input circuit.



$$V_{OUT} = V_{IN}^{+}\left(\frac{R_2}{R_1+R_2}\right)\left(\frac{R_3+R_4}{R_3}\right) - V_{IN}^{-}\left(\frac{R_4}{R_3}\right)$$

Figure 11A. Simple differential input circuit.

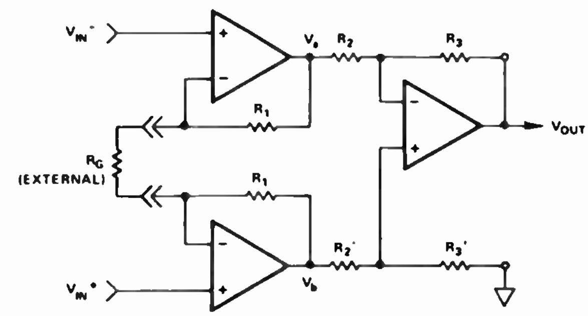


Figure 11C. "Classic" three op-amp differential input circuit.

The circuit shown in Fig. 11B overcomes most of the drawbacks of the first circuit. Input resistance is high, CMR is dependent only upon the resistor match, and gain can be set using only one resistor. The major disadvantage is that the common mode voltage input range is gain dependent and can be very poor at high gains.

The best choice is shown in Fig. 11C, where the input impedance is high, the CMR is dependent only on the ratio of R2, R2', R3, R3', and not on R1, R1', making the CMR independent of gain.

The output circuit of a professional tape machine must be able to drive long, balanced, 600 ohm or 150 ohm audio lines at levels of 0 to +8 dBm plus peaks with low distortion, reasonable headroom (10 dB to 14 dB) and good DC isolation. These performance criteria require the use of some sort of high quality balanced output drive circuits. The output circuit choices are similar to those for the input circuits, either transformer coupled as shown in Fig. 12, or electronically balanced and floating circuits as shown in Fig. 13. The transformer coupled output offers very good CMR and very good DC isolation with disadvantages in size, cost, and greater low frequency distortion.

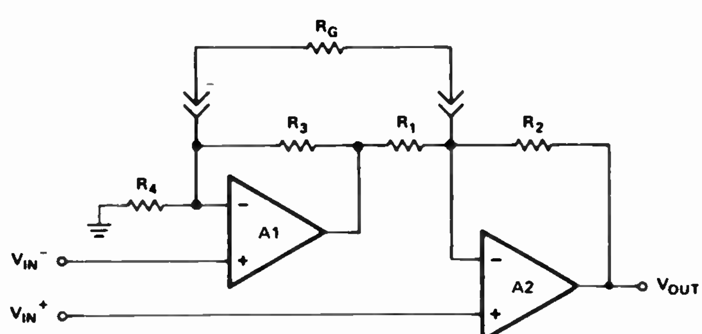
There are three choices in electronically balanced output configurations shown in Fig. 13. A circuit with the virtues of simplicity and one resistor gain adjustment, but not truly balanced, is shown in Fig. 13A. Grounding the positive output terminal drops the output 6 dB, but grounding the negative output terminal reduces the output to nothing. The circuit in Fig. 13B loses 6 dB of output if either output terminal is grounded. The circuit in Fig. 13C is a balanced and floating design, which is an electronic approximation of a transformer. This design includes cross-coupled feedback, which keeps the output level constant even if one output terminal is shorted.

Remote Control

The remote control interface circuits of most professional tape machines is a simple open-collector active pull-down transistor circuit, with the appropriate pull-up supply voltage available at the remote connector of the machine, as shown in Fig. 14. Some microprocessor controlled machines have serial interfaces (usually RS-232) as well. Although barrier strips and "Jones" connectors were once the connectors of choice, most new machine remote interfaces use the 25-pin "D" connectors made inexpensive and widely available by the personal computer industry. Some machines have dry relay contacts available for other interface controls.

For some machines with digital control logic, special optically isolated remote control interfaces may be necessary to prevent crosstalk between control circuit and to avoid ground loops.

One of the problems in using consumer-type cassette tape machines in professional applications is the lack of machines with easily implemented remote controls.



$$\text{IF } \frac{R_2}{R_1} = \frac{R_4}{R_3}, \quad V_{IN} = V_{IN}^{+} - V_{IN}^{-}$$

$$\frac{V_{OUT}}{V_{IN}} = 1 + \frac{R_2}{R_1} + \frac{2R_2}{R_G}$$

Figure 11B. Two op-amp differential input circuit.

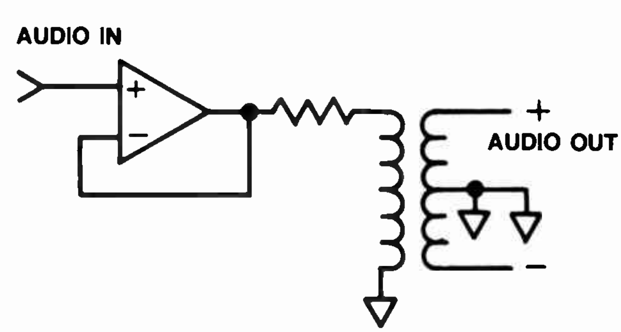


Figure 12. Transformer coupled output circuit.

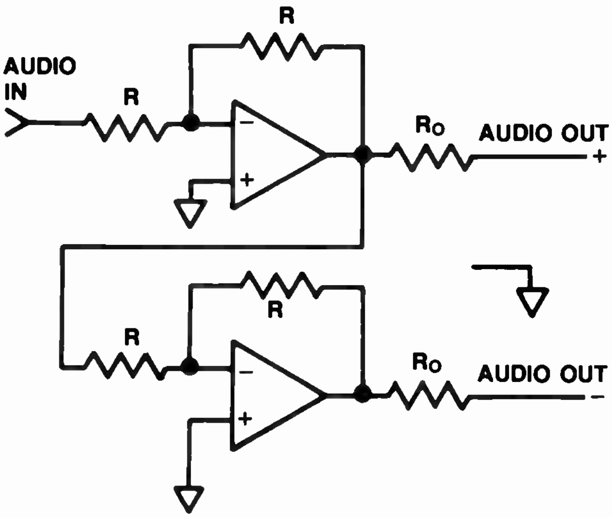


Figure 13A. Simple op-amp line output circuit.

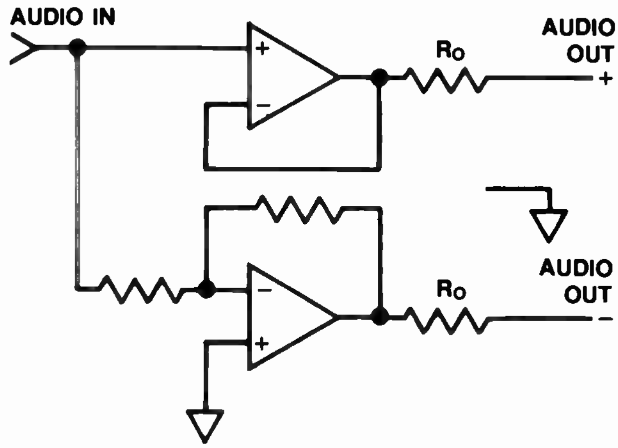


Figure 13B. Balanced (not floating) line output circuit.

The consumer units usually use a hand held infrared remote, which requires circuit surgery to use the simple contact closure remote controls available from consoles and other professional equipment.

TAPE MACHINE PERFORMANCE COMPARISON

The following material is presented in order to give an idea of the relative performance level of each format presented. It is representative of the median performance level of professional machines gleaned from industry test reports of machines in each category.

1. SIGNAL-TO-NOISE (3% THD Reference Level, No Noise Reduction)
   Cassette 55 dB
   Reel-to-reel 65 dB
   Cartridge 60 dB
2. DISTORTION (THD Referenced To Normal Operating Level)
   Cassette 1.0%
   Reel-to-reel 0.5%
   Cartridge 1.5%
3. FREQUENCY RESPONSE (±3 dB)
   Cassette (1⅞ IPS) 20 Hz to 20 kHz (@ 20 dB below reference)*
   Reel-to-reel (7.5 IPS) 20 Hz to 22 kHz (@ 10 dB below reference)
   Cartridge (7.5 IPS) 30 Hz to 18 kHz (@ 10 dB below reference)

   *The small tape area and slow speed of cassette machines mean that they are more susceptible to tape saturation, so cassette machine frequency response drops rapidly at higher levels such as 10 dB below reference.
4. HEADROOM
   Cassette 3 to 6 dB (above 200 nWb/m)
   Reel-to-reel 10 to 14 dB (above 320 nWb/m)
   Cartridge 10 to 14 dB (above 250 nWb/m)
5. WOW & FLUTTER (DIN WTD PEAK)
   Cassette 0.05%
   Reel-to-reel 0.04%
   Cartridge 0.1%

TAPE RECORDING REFERENCE LEVELS

Each tape format has its own particular "official" *reference level* or *flux level*. The table below lists the reference levels for several of the more popular levels, the difference between them, and the formula for calculating the difference. These can be used when evaluating the specifications of tape machines when the flux level is tied to the signal-to-noise ratio. By adding or subtracting the dB difference in reference levels between two different sets of specifications, a more accurate specification comparison can be made.

FLUX LEVEL *(nWb/m)*	DIFFERENCE *(in dB)*	REFERENCE *SOURCE*
160	0.0	NAB 1975 Reference Tape
185	1.3	Ampex Reference Tape
200	1.9	MRL Reference Tape
250	3.9	Elevated Level #1
320	6.0	Elevated Level #2

$$\textit{FLUX LEVEL (dB difference)} = 20 \log \frac{\textit{New Level (nWb/m)}}{\textit{Ref Level (nWb/m)}}$$

Example: Compare a cart machine with a signal-to-noise ratio of 65 dB referenced to 400 nWb/m to another cartridge machine with a signal-to-noise ratio of 57 dB at 160 nWb/m.

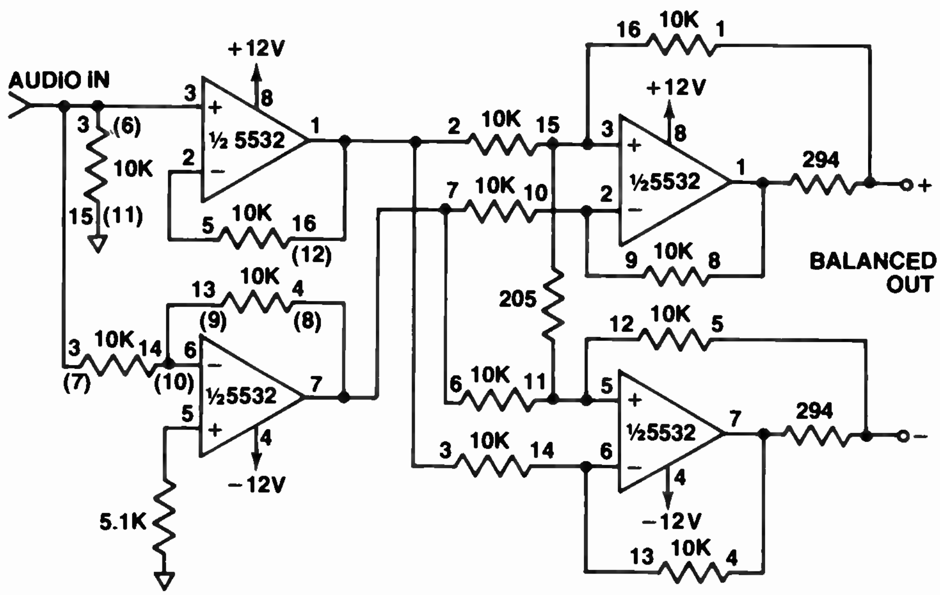


Figure 13C. Balanced and floating line output circuit.

Difference (dB) = 20 log (400 ÷ 160) = 7.96 dB
Add the 57 dB and 8 dB and obtain 65 dB.

This calculation shows the two machines have virtually the same signal-to-noise ratio performance.

AUDIO TAPE MACHINE MAINTENANCE

This section covers general maintenance procedures and the specific maintenance areas of heads, pressure rollers, motors, and electronics. There is also a short discussion on troubleshooting and some general considerations. All tape machines have one thing in common; they cannot perform to their designed specifications if they are not properly maintained. Often the maximum performance of a machine is determined by the level and frequency of care rather than design or construction limitations.

General Maintenance Procedures

A good method for keeping tape machines in good working order, especially in a broadcasting environment where many such machines are involved, is a scheduled maintenance system. The design and implementation of a good maintenance schedule can prevent most common types of problems from affecting the on-air signal.

One such maintenance system identifies each individual machine by serial number, location in the

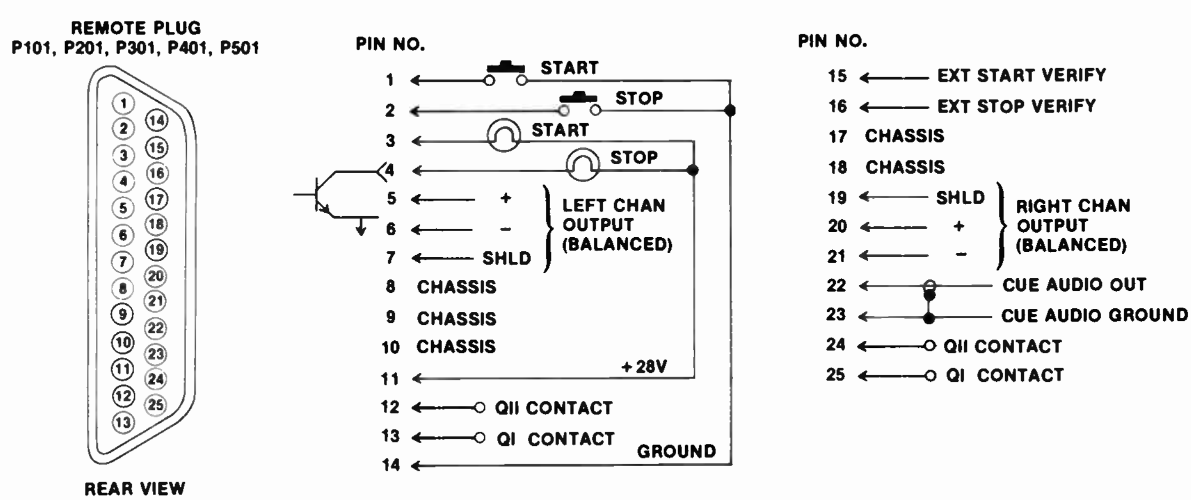


Figure 14. Typical remote control interface circuit.

system, and date of original installation. Since maintenance level varies with the area of use, it is essential to know the location of each machine at all times. This tracking of individual units allows rotating machines with spares, and it gives an accurate record for insurance and warranty information. The location record for each unit should also include its problem history record. This problem history can provide information on machines that would otherwise be lost. This information can help point out trends such as the need for increased maintenance in a certain area, shortcomings in equipment and installation design, and environment problems.

The most important information that can be provided by the equipment history card is the cost and frequency of repairs. This helps in deciding when it is no longer economical to continue to repair a tape machine.

There are many possible methods for logistically handling both the maintenance schedule and the problem history. However, with the advent of business and personal computers, the job can be simplified. It is now possible to set up a system so that at the beginning of each day, or each week, a printout is produced of the work scheduled for that given day or week. This allows efficient scheduling in any size installation.

Placing a Tape Machine In Service

The first step in setting up a maintenance system is establishing initial conditions: design specifications and the results of initial tests by the manufacturer and purchaser. For example, when a new machine is received, a test data sheet is usually included with or bound into the service manual. The data sheet will show results of final tests conducted by the manufacturer prior to shipment, such as:

- Signal-to-noise ratio for a given flux level
- Frequency response at one or more flux or record levels
- Nominal operating levels
- Wow and flutter and measurement standard
- Distortion in percent for a given flux or record level and frequency
- Crosstalk between channels at a given frequency and record level
- Verification of functions such as tone sensing and start and stop times

These data should be checked to the extent that time and available test equipment permit. Then the new machine should be re-equalized for the type of tape that will be used and any new measurements recorded.

This overview of maintenance systems is intentionally broad. While a program for maintenance is essential, the specific system must be tailored to fit the individual installation. *A maintenance system must be set up so that it is easy to work with and convenient to use, otherwise it will not be followed.* In other words, use something that works.

Daily Maintenance

Daily maintenance tasks are those that must be performed routinely to insure continued good performance and prevent failures not related to mechanical wear or electronic failure. These are some of those tasks.

Head Cleaning

This is probably the most often performed maintenance task and one that is given possibly the least consideration. It is very easy to provide a cotton swab and a bottle of alcohol to the machine operators and tell them to clean the heads at the beginning of each shift. However, consideration should be given to several of the items involved in the head cleaning process.

Head cleaning fluid. There are nearly as many head cleaning fluids available as there are heads to clean. Everyone has a favorite, and the manufacturer usually recommends a cleaner. Here are a few things to consider when using head cleaners:

1. Safety for the user. Some liquids, such as carbon tetrachloride, which might be used as head cleaners, are dangerous to breathe. The more common alcohol cleaners are effective and safer.
2. Safety for the tape, heads, and other hardware. Strong solvents can dissolve materials used in the construction of heads, rollers, and other hardware. The tape itself may be vulnerable to solvent remaining after the cleaning process. Furthermore, a liquid may be spilled or oversprayed and cause damage to other parts of the equipment.

Lacking specific recommendations from the equipment manufacturer, the safest cleaners are those made with *ethyl alcohol* (ethanol), *isopropyl alcohol*, or *freon*. These substances evaporate quickly and leave no residue. Do not use rubbing alcohol as it may contain some substances which will not evaporate.

Remember that the more effectively a solvent dissolves the oxide build-up on heads, the more effectively it may dissolve tape and other materials.

The cleaning utensil. Cotton swabs, available at drugstores, are the most commonly used tools for cleaning heads. They provide the user with extra reach (longer than fingers) so that hard-to-reach heads can be cleaned more easily and thoroughly. However, cotton swabs have two main drawbacks; they tend to shed the cotton that is twisted onto the ends of the stick, and the surface that can be put in contact with the head at any one time is quite small. If swabs are used, the heads should be carefully inspected for cleanliness before putting the machine back in service.

Another common head cleaning utensil is a clean, soft, lint-free cloth. This method allows pressure to be applied as needed to remove dirt build-up. There are also drawbacks to using a cloth. It is not always possible to get a finger past the head cover and onto the heads in some machines. Extending the reach by wrapping the cloth on the end of a screw driver or pencil eraser is not a good idea because of the damage that could be done to the head assembly. The other problem is possible excessive pressure on the heads. This may cause misalignment of the head assembly. Finally, tools such as screw drivers tend to become magnetized and should not be placed close to heads and other tape contact surfaces.

Tape Path Cleaning

Another daily maintenance chore that goes with head cleaning is the cleaning of the rest of the tape path. Pressure rollers, capstan shafts, idler wheels, and guides should also be cleaned with the same diligence as the heads. While they are not in the audio path itself, they contribute to the overall performance of the tape machine by affecting such parameters as wow and flutter and pulling torque.

Head and Tape Path Demagnetization

A third task that should be performed daily is demagnetization, or degaussing, which must be done with care. Not only can heads become magnetized through normal use, but improper demagnetization can leave residual magnetism in the head which will result in not only poor audio performance from machines, but partial erasure of high frequencies on the tape.

A variety of head demagnetizers is available from various sources. However, consider how the degausser will be used prior to purchase. Tape heads have a wide range of hardness which affects the retentivity of the material. The harder heads require a much more powerful demagnetizer. An under-powered demagnetizer, no matter how many passes are made at the head, will not demagnetize an extended-life head. The other prime factor in buying head degaussers is the accessibility of the heads. Degaussers in cartridge form, do an excellent job on cart machine heads, but are not suitable for any other application. Be sure the demagnetizer is powerful enough and will fit the units it is intended to service.

In addition to selecting the proper maintenance tool, it is no less important to use it properly. If operators are to perform the degaussing of heads, they should be properly trained. In general, the procedure should include the following steps:

1. Turn off power to the tape machine. The degausser induces strong magnetic fields which could severely overload sensitive circuitry.
2. Turn on or plug in the degausser while holding it at least three feet from the tape machine.
3. Slowly wipe the tip of the degausser across the heads, guides, rollers and other components in the tape path. Avoid touching any tape contact surface unless the tip is protected with a clean, soft cover material.
4. Slowly rotate the degausser while removing it away from the tape head. The degausser should be at least three feet from the heads before turning it off.

In order to verify that the heads have been completely degaussed, take a blank tape and check the machine's signal-to-noise figure. If it shows an increase in the signal-to-noise over the normal level, repeat the demagnetizing process. Just as in the head cleaning process, all other metallic components in the tape path should be demagnetized regularly.

These daily maintenance tasks should be performed without fail. Postponing or poorly executing these tasks can only result in poor performance. Be diligent!

Weekly Maintenance

Weekly maintenance should consist mainly of a brief physical inspection of the unit for obvious problems such as overheating, noisy motor bearings, excessive debris from tape shedding on the deck or tape guides. In other words, make a point to look at the machine on a regular basis to verify that it is still operating normally. Also talk to the operators. They can help head off failures by reporting if a machine is acting abnormally.

Monthly or Quarterly Maintenance

This section is given a dual heading because of the variation in usage of tape equipment at every installation. The frequency of the checks outlined here may vary. It will depend on how the machine is used, the environment it is in, and the reliability of the individual unit. Some units may require that these checks be made more frequently than others, but this is something that will not be evident until service history records have been developed for all machines. Start by using a relatively short interval; and if no corrections are required, the interval can be lengthened. *The primary purpose of a regular maintenance program is to prevent failures, not fix them.* Do not allow the service interval to become too long. Also, most of these checks require that the machine be taken out of service, so it would be advisable to stagger the schedule.

General Cleaning

Each unit should be carefully dusted by using either a soft-bristled brush or low pressure air. Covers should be removed during this process, where it is convenient, so that debris is free to fall clear of the machine. This is also a good opportunity to clean the tape path thoroughly and demagnetize all the areas that are difficult for the operators to get to.

Visual inspection of the mechanical parts should be made to check for head wear, deck wear in cart machines, guidance wear, and any other mechanical parts that are subject to fatigue and constant use.

Motors should be checked for bearing noise. While small amounts of bearing noise may not be a sign of eminent failure, they should be considered a warning signal. Listening to the mechanical sounds of a machine can often tell as much as more sophisticated tests for wow, flutter, and torque.

Some tape machines need periodic lubrication of certain bearings or moveable parts. Use a light machine oil or the manufacturers recommended lubrication. Wipe off any excess lubrication so as not to collect dust.

Routine Performance Check

Aside from any obvious faults that may have been detected by operators or a visual inspection, there are a few quick checks which will determine if the unit is performing up to standard. First, the results of routine performance tests should be kept on file with the history card for each machine. Most manufacturers will provide test data sheets that can serve as a standard and a guide for the test information. Data taken at each performance check should be compared back to the original. Second, test tapes should be those recommended by the manufacturer and adjustments should be made in accordance with the specific equipment manual.

Azimuth Adjustment and Phase Response

The significance of these adjustments was discussed earlier in this section, but they are essential to good tape machine performance. If the azimuth adjustment is unstable or hard to align, look for some mechanical problem in the head mounting assembly or the tape guidance system. The important thing is to get a good "feel" for how the machine reacts to adjustments. This is the best way to obtain an accurate setting.

Playback Frequency Response

A quick check of playback response, using a standard reference tape or suitable in-house reference recording, will provide an early warning of the beginning of head wear and drift in response and output levels. A worn playback head will begin to lose the high-end response. A certain amount of this can be compensated for by re-equalizing the playback electronics.

Record Frequency Response

Proper record frequency response is essential to maintaining high-quality performance. A quick check of this specification, by making a test recording after adjusting for proper playback performance, will reveal if the unit needs an alignment and will provide information concerning wear of the record head.

Record/Playback Distortion and Noise

These two parameters will give more clues as to the overall performance of the tape machine electronics than any other measurements. Poor signal-to-noise figures or high distortion can be the indicators of many more severe problems that can be prevented if action is taken early. Take the time to investigate the symptoms in increased noise and distortion when they begin to appear.

Wow and Flutter and Tape Tension

Measurements of *wow*, *flutter* and tape tension or tape-reel motor torque can indicate impending failure of motors or bearings, or a need for lubrication. The equipment to measure wow and flutter is expensive and usually not found in stations, but it can be rented for periodic checks. Listening to a tape with a continuous tone will give a good qualitative check of wow and flutter.

There are standard gauges that can be used to measure tape tension and tape-reel motor torque. A periodic test of tension and torque will insure that good tape shuttling performance is maintained.

For cart machines, a quick test of wow, flutter and torque can be made using a tone recorded on the longest cartridge available, thus placing the greatest load on the mechanical system.

Although the five performance checks described above represent a fairly complete testing of a tape machine, the tests can be performed in less than 30 minutes per machine in most cases. The time spent on these tests is well worth it.

Semi-Annual Maintenance

Maintenance performed at this interval should be an in-depth version of the monthly or quarterly tasks. Complete specification verification should be done along with a thorough mechanical analysis. It should be determined that all motors have sufficient torque, both starting and pulling, that solenoids, brakes, and other electromechanical devices all operate smoothly, and that all mechanical operation is quiet.

Again, the timing of these procedures is dependent on how the machine is used. For machines that receive relatively light duty, this thorough examination would not be required as often as a machine in heavy duty "on-air" operations. Preventative maintenance performed less frequently than on a semi-annual basis provides little protection against failures between checks.

TAPE AND TAPE CARTRIDGES

One of the major keys to good tape machine performance is consistency in all areas, including the tape and tape cartridges. Once a certain type of tape and one certain brand of tape cartridge has been chosen, stay with them. Optimize the machines to that type of tape and use it throughout the system. Research thoroughly the decision to change to a different kind of tape and tape cart to insure it meets the needs of the facility or is compatible with the tape machines on which the tape will be used.

Some facilities use different tape in different areas. For example, mastering quality tape may be used in a production facility but a lower grade of tape is used in news. Be sure that tape and tape machines are not intermixed to avoid problems with equalization, tape handling, and operator expertise.

TEST TAPES AND ALIGNMENT TOOLS

Test tapes are an essential part of the equipment needed for proper care of tape machines. The manufacturer usually recommends a certain test tape for their machine. However, if more than one brand of cart machine or reel-to-reel machine is in use in the facility, by all means use only one test tape, and not a separate one for each brand of machine.

The care of test tapes is important, and a few basic rules should be followed. Aside from the normal care considerations given when the tape is not being used, always demagnetize all heads and tape guides before playing the tape. Also, check that no stray bias or DC voltage is present on the heads before using. The normal useful life of a test tape with proper care should be several hundred passes, after which it should be replaced just as would be done with any worn-out tool.

Aside from normal hand tools, some manufacturers recommend certain head, motor, and guidance alignment tools. These tools should be kept together and used only on the machines for which they were designed.

SUMMARY

The general tape machine maintenance procedures described above are those that can and should be performed on a regular basis. When a good preventative maintenance program and system of records is set up and maintained, it will help make the engineer's job much easier, and provide time for dealing with emergencies that are not related to the tape equipment.

There are two key factors to remember in summing up this section on tape machine maintenance. First, be diligent in the preventative maintenance program and most problems will never get on the air or into a final production. Second, be consistent with maintenance and operational procedures and all machines will perform the same in the news room, the control room, the remote van, or in the production studio.

Instituting and following a good maintenance program will result in obtaining the maximum life and performance out of every tape machine.

5.5 Audio Recording
Part II: Digital Audio Technology

Ken C. Pohlmann
University of Miami, Coral Gables, Florida

INTRODUCTION

Since its first appearance in professional recording studios in 1971 in the guise of digital delay lines, digital audio technology has fundamentally changed the nature of the audio and video recording and broadcast industries. The high-quality signal processing, lossless generation copying of program material, flexible editing capabilities, and other attributes now rival or surpass other analog technologies. The changes engendered by digital audio extend to all aspects of audio in broadcasting. Digital signal processing, high density storage formats, data reduction techniques, and digital audio broadcasting technologies are transforming the recording and broadcast industries.

AUDIO DIGITIZATION

Using the principles of discrete time sampling and quantization, a sampled signal can be processed, transmitted, or stored, and through conversion can reconstruct an accurate representation of the original analog signal.

Discrete Time Sampling

An analog waveform, such as an acoustic pressure function in air, exists continuously in time over a continuously variable amplitude range. Such an analog function may be discrete time sampled; moreover, the sample points can be used to reconstruct the original analog waveform. This digitization of audio forms the basis for the encoding and decoding of the audio signals in any digital audio format.

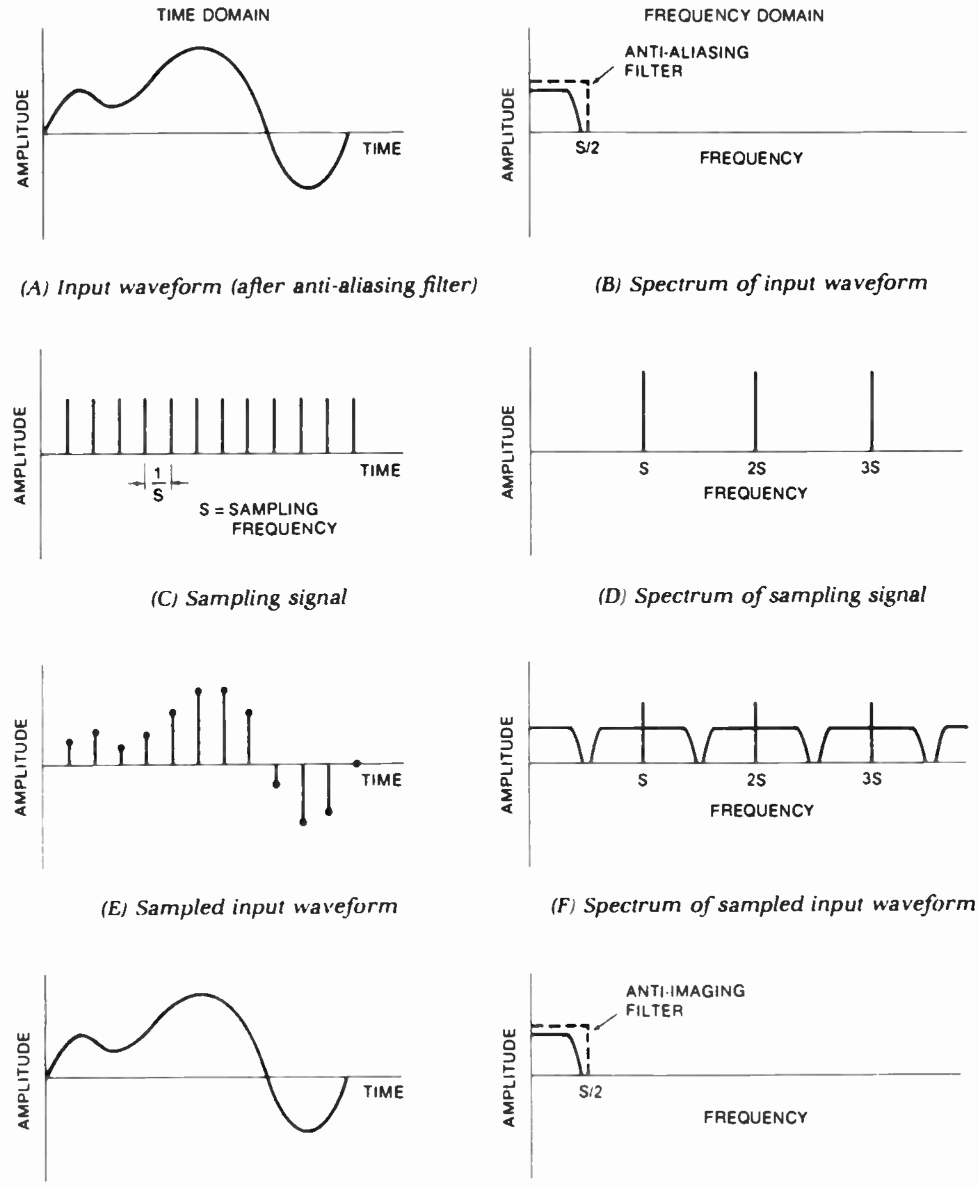
The Nyquist theorem states that given correct, bandlimited conditions, sampling can be a lossless process. However the relationship between sampling frequency, and audio frequencies must be observed. The Nyquist theorem defines the relationship: if the sampling frequency is at least twice the highest audio frequency, complete waveform reconstruction can be accomplished.

The choice of sampling frequency determines the frequency response of the digitization system; S samples per second are needed to represent a waveform with a bandwidth of S/2 Hz. As the sampled frequencies become higher, there will be fewer samples per period. At the theoretical limiting case of critical sampling, at an audio frequency of half the sampling frequency, there will be two samples per period. A low-pass filter is placed at the output of every audio digitization system to remove all frequencies above the half-sampling frequency. This is required because sampling, through modulation, generates new frequencies above the audio band. The output filter removes all spectra above the half-sampling frequency. This is summarized in Fig. 1.

By definition, audio samples contain all the information needed to provide complete reconstruction. However bandlimiting criteria must be strictly observed; a too-high frequency would not be properly encoded, and would create a kind of distortion called aliasing. An input frequency higher than the half-sampling frequency would cause the digitization system to alias. If S is the sampling frequency and F is a frequency higher than half the sampling rate, then new frequencies are also created at S ± F, 2S ± F, 3S ± F, etc. An input low-pass filter will prevent aliasing if its cutoff frequency equals the half-sampling frequency. To achieve a maximum audio bandwidth for a given sampling rate, filters with a very sharp cutoff characteristic, so-called brickwall filters, are employed in either the analog or digital domain.

Amplitude Quantizing

The amplitude of each sample yields a number which represents the analog value at that instant. By



(A) Input waveform (after anti-aliasing filter)
(B) Spectrum of input waveform
(C) Sampling signal
(D) Spectrum of sampling signal
(E) Sampled input waveform
(F) Spectrum of sampled input waveform
(G) Output waveform (after anti-imaging filter)
(H) Spectrum of output waveform

Figure 1. Summary of discrete-time sampling, shown in the time and frequency domains. (From Pohlmann, *Principles of Digital Audio*.)

definition, an analog waveform has an infinite number of amplitude values, however quantization selects from a finite number of digital values. Thus after sampling, the analog staircase signal is rounded to a numerical value closest to the analog value. The difference between the original values of the signal and values after quantization appears as error.

The number of quantization steps available is determined by the length of the data word in bits; the number of bits in a quantizer determines resolution. Sixteen bits yield $2^{16} = 65{,}536$ increments. Every added bit doubles the number of increments, hence the magnitude of the error is smaller. The accuracy of a quantizing system provides an important performance specification. In the worse case, there will be an error of one half the least significant bit of the quantization word. The ratio of maximum expressible amplitude to error determines the signal-to-error (S/E) ratio of the digitization system. The S/E relationship can be expressed in terms of word length as S/E (dB) $= 6.02n + 1.76$, where "n" is the number of bits.

Although a 16-bit system would yield a theoretical signal-to-error ratio of 98 dB, as the signal amplitude decreases, the relative error increases. Consider the example of a signal with amplitude on the order of one quantization step. The signal value crosses back and forth across the threshold, resulting in a square wave signal from the quantizer. Dither suppresses such quantization error. Dither is a low amplitude analog noise added to the input analog signal (similarly, digital dither must be employed in the context of digital computation when rounding occurs).

When dither is added to a signal with amplitude on the order of a quantization step, the result is duty-cycle modulation which preserves the information of the original signal. The average value of the quantized signal can move continuously between two steps, thus the incremental effect of quantization has been alleviated. Audibly, the result is the original waveform, with added noise. That is more desirable than the clipped quantization waveform. With dither, the resolution of a digitization system is below the least significant bit.

The recording section of a pulse code modulation (PCM) system, shown in Fig. 2A, consists of input amplifiers, a dither generator, input (anti-aliasing) low-pass filters, sample-and-hold circuits, analog-to-digital converters, a multiplexer, digital processing circuits for error correction and modulation, and a storage medium such as digital tape. The reproduction section, shown in Fig. 2B, contains processing circuits for demodulation and error correction, a demultiplexer, digital-to-analog converters, output sample-and-hold circuits, output (anti-imaging) low-pass filters, and

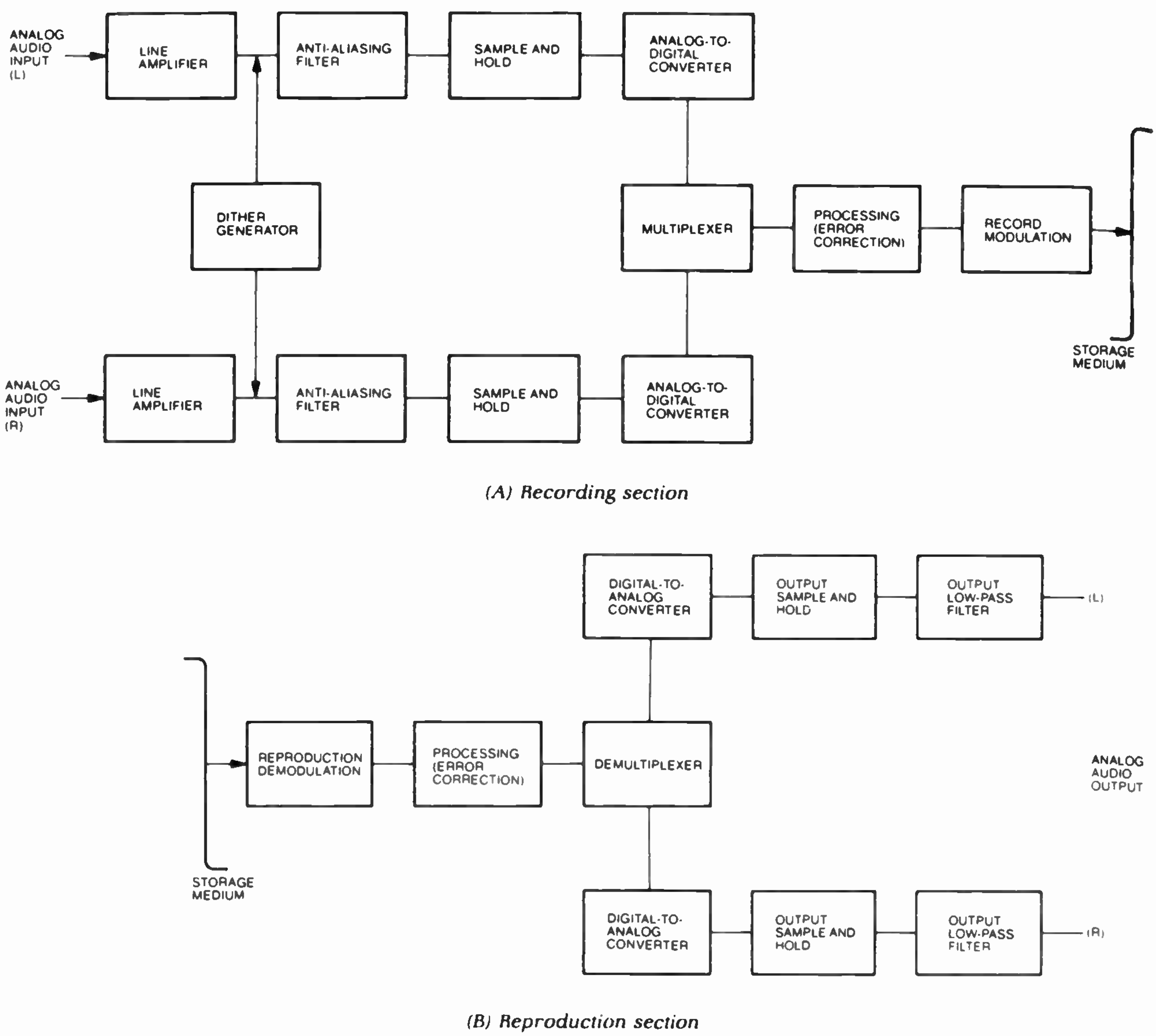


Figure 2. Block diagram of the recording (A) and reproduction (B) sections of a linear PCM system. (From Pohlmann, *Principles of Digital Audio*.)

output amplifiers. In most contemporary designs, digital filters are used in both the input and output stages. The output section forms the basis for a compact disc player.

COMPACT DISC FORMAT

The *compact disc* (CD) format was developed to store up to 74 minutes of stereo digital audio program material of 16-bit PCM data sampled at 44.1 kHz. Total user capacity is over 650 Mb. In addition, for successful storage, error correction, synchronization, modulation, and subcoding are required.

Compact Disc Physical Design

The diameter of a compact disc is 120 mm, its center hole diameter is 15 mm, and its thickness is 1.2 mm. Data are recorded in an area 35.5 mm wide. It is bounded by a lead-in area and a lead-out area, which contain nonaudio subcode data used to control the player's operation. The disc is constructed with a transparent polycarbonate substrate. Data are represented by pits which are impressed on the top of the substrate. The pit surface is covered with a thin metal (typically aluminum) layer 50 mm to 100 mm thick, and a plastic layer 10 mm to 30 mm thick. A label 5 mm thick is printed on top. Disc physical characteristics are shown in Fig. 3.

Pits are configured in a continuous spiral from the inner circumference to the outer. The pit construction of the disc is diffraction-limited; the dimensions are as small as permitted by the wave nature of light at the wavelength of the readout laser. A pit is about 0.5 mm wide. The track pitch is 1.6 mm. There are a maximum of 20,188 revolutions across the disc's data area.

The disc rotates with a constant linear velocity (CLV) in which a uniform relative velocity is maintained between the disc and the pick-up. To accomplish this, the rotation speed of the disc varies depending on the radial position of the pick-up. The disc rotates at a speed of about 8 rev/s when the pick-up is reading the inner circumference, and as the pick-up moves outward, the rotational speed gradually decreases to about 3.5 rev/s. The player reads frame synchronization words from the data and adjusts the speed to maintain a constant data rate.

The CD standard permits a maximum of 74 minutes, 33 seconds of audio playing time on a disc. However by modifying encoding specifications such as track pitch and linear velocity, it is possible to manufacture discs with over 80 minutes of music. Although the linear velocity of the pit track on a given disc is constant, it can vary from 1.2 m/s to 1.4 m/s, depending on disc playing time. All audio compact discs and players must be manufactured according to the Red Book, the CD standards document authored by Philips and Sony.

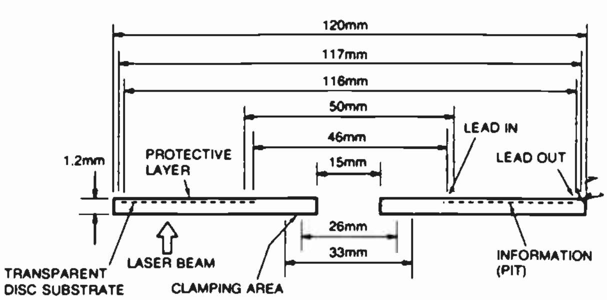


Figure 3. Compact disc physical specifications. (From Pohlmann, *The Compact Disc*.)

Compact Disc Encoding

Compact disc encoding is the process of placing audio and other data in a frame format suitable for storage on the disc. The information contained in a CD frame prior to modulation consists of a 27-bit sync word, 8-bit subcode, 192 data bits and 64 parity bits. The input audio bit rate is 1.41×10^6 bps. Following encoding, the channel bit rate is 4.3218×10^6 bps. Premastered digital audio data are typically stored on a ¾-inch U-matic video transport via a digital audio processor with a 44.1 kHz sampling rate and 16-bit linear quantization.

A frame is encoded with six 32-bit PCM audio sampling periods, alternating left and right channel 16-bit samples. Each 32-bit sampling period is divided to yield four 8-bit audio symbols. The CD system employs two error correction techniques: interleaving to distribute errors, and parity to correct them. The standardized error correction algorithm used is the Cross Interleave Reed-Solomon Code (CIRC), developed specifically for the compact disc system. It uses two correction codes and three interleaving stages. With error correction, over 200 errors per second can be completely corrected.

Subcode Data

Following CIRC encoding, an 8-bit subcode symbol is added to each frame. The eight subcode bits (designated as P,Q,R,S,T,U,V, and W) are used as eight independent channels. Only the P or Q bits are required in the audio format; the other six bits are available for video or other information as defined by the CD+G/M (Graphics/MIDI) format. The CD player collects subcode symbols from 98 consecutive frames to form a subcode block with eight 98-bit words; blocks are output at a 75 Hz rate. A subcode block contains its own synchronization word, instruction and data, commands and parity. An example of P and Q data are shown in Fig. 4.

The P channel contains a flag bit that can be used to identify disc data areas. Most players use information in the more comprehensive Q channel. The Q channel contains four types of information: control, address, data, and cyclic redundancy check code (CRCC) for subcode error detection. The control bits specify several playback conditions: the number

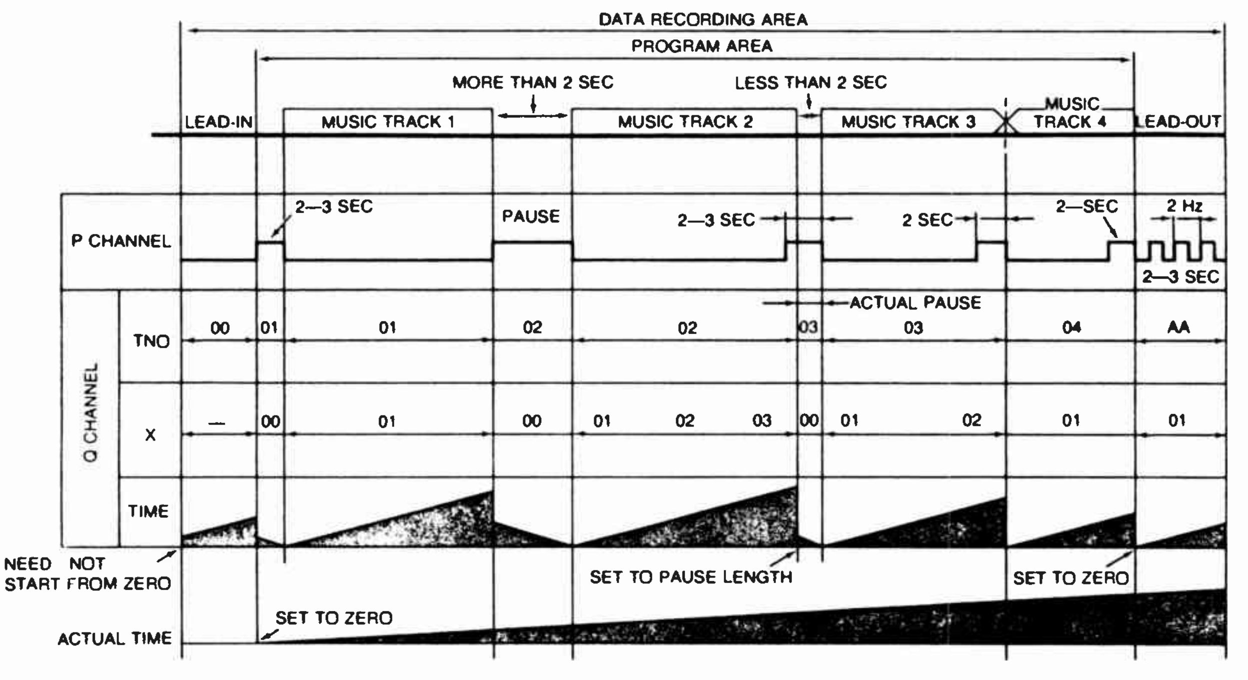


Figure 4. Typical subcode contents of the P and Q channels. (From Pohlmann, *The Compact Disc*.)

of audio channels (two/four); preemphasis (on/off); and digital copy prohibited (yes/no). The address information consists of four bits designating three modes for the Q data bits. Mode 1 data are contained in the table of contents (TOC) which is read during disc initialization. The TOC stores data indicating the number of music selections as a track number, and the starting points of the tracks in disc running time. In the program and lead-out areas, Mode 1 contains track numbers, indices within a track, track time, and disc time. The optional Mode 2 contains the catalog number of the disc. The optional Mode 3 contains a country code, the owner code, the year of recording, and a serial number.

EFM Encoding and Frame Assembly

The audio, parity, and subcode data are modulated using EFM (*eight-to-fourteen modulation*) in which symbols of eight data bits are assigned an arbitrary word of fourteen channel bits. By choosing 14-bit words with a low number and known rate of transitions, greater data density can be achieved. Each 14-bit word is linked by three merging bits. The 8-bit input symbols require 256 different 14-bit code patterns. To achieve pits of controlled length, only those patterns are used in which more than two but less than ten 0s appear continuously. Two other patterns are used for subcode synchronization words. The selection of EFM bit patterns defines the physical relationship of the pit dimensions. The channel stream comprises a collection of nine pits and nine lands that range from 3T to 11T in length where T is one period. A 3T pit ranges in length from 0.833 to 0.972 mm and an 11T pit ranges in length from 3.054 mm to 3.560 mm, depending on pit track linear velocity. Each pit edge, whether leading or trailing, is a 1, and all increments in between, whether inside or outside a pit, are 0s, as shown in Fig. 5.

The start of a frame is marked with a 24-bit synchronization pattern, plus three merging bits. The total number of channel bits per frame after encoding is 588, comprised of: 24 synchronization bits, 336 ($12 \times 2 \times 14$) data bits, 112 ($4 \times 2 \times 14$) error correction bits, 14 subcode bits, and 102 (34×3) merging bits.

Data Readout

CD pick-ups use an *AlGaAs* (aluminum gallium arsenide) semiconductor laser generating laser light with a 780 nm wavelength. The beam passes through the substrate, is focused on the metalized pit surface and is reflected back. Because the disc data surface is physically separated from the reading side of the substrate, dust and surface damage on the substrate do not lie in the focal plane of the reading laser beam and hence their effect is minimized. The polycarbonate substrate has a refractive index of 1.55; because of the bending of the beam from the change in refractive index, thickness of the substrate, and the numerical aperture (0.45) of the laser pick-up's lens, the size of the laser spot is reduced from approximately 0.8 mm on the disc surface to approximately 1.7 mm at the pit surface. The laser spot on the data surface is an Airy function with a bright central spot and successively

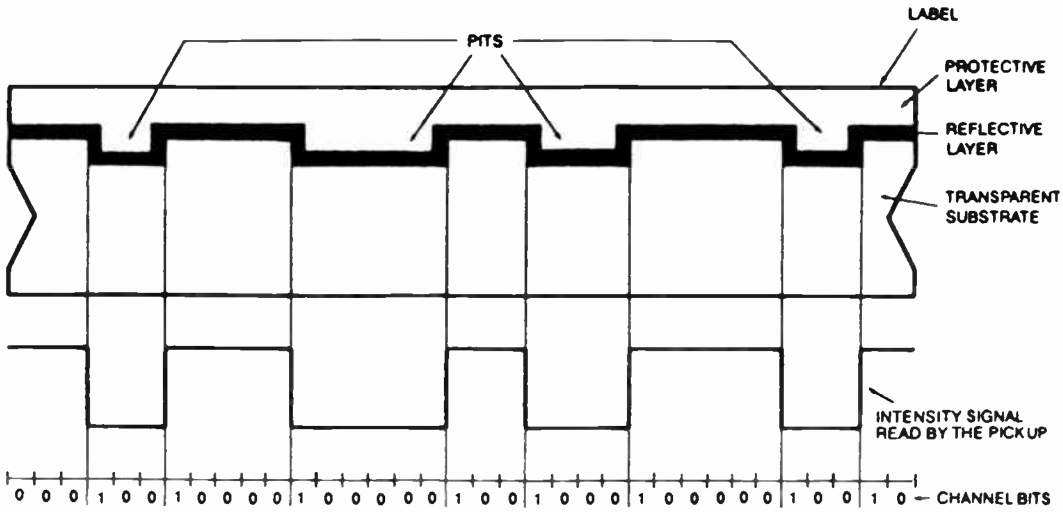


Figure 5. Channel bits as represented by the pit structure. (From Pohlmann, *Principles of Digital Audio.*)

darker rings, and spot dimensions are quoted as half-power levels.

When viewed from the laser's perspective, the pits appear as bumps with height between 0.11 mm to 0.13 mm. This dimension is slightly less than the laser beam's wavelength in polycarbonate of 500 nm. The height of the bumps is thus approximately ¼ of the laser's wavelength in the substrate. Light striking land travels a distance one-half wavelength longer than light striking a bump, as shown in Fig. 6. This creates an out-of-phase condition between the part of the beam reflected from the bump, and the part reflected from the surrounding land. The beam thus undergoes destructive interference, resulting in cancellation. Optically, if the CD pit surface is considered as a two-dimensional reflective grating, the focused laser beam diffracts into higher orders, resulting in interference. The disc surface data thus modulates the intensity of the reflected light beam. In this way the data physically encoded on the disc are recovered by the laser.

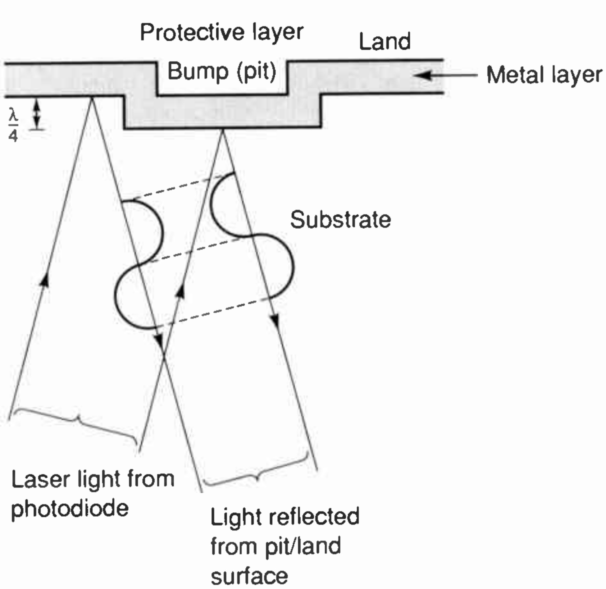


Figure 6. A pit causes cancellation through destructive interference.

Data Decoding

A compact disc player's data path, shown in Fig. 7, directs the modulated light from the pick-up through a series of processing circuits, ultimately yielding a stereo analog signal. Data decoding follows a procedure which essentially duplicates, in reverse order, the encoding process. The pick-up's photodiode array and its processing circuits output EFM data as a high frequency signal. The first data to be extracted from the signal are synchronization words. This information is used to synchronize the 33 symbols of channel information in each frame, and a synchronization pulse is generated to aid in locating the zero crossing of the EFM signal.

The EFM signal is demodulated so that 17-bit EFM words again become 8 bits. A memory is used to buffer the effect of disc rotational wow and flutter. Following EFM demodulation, data are sent to a CIRC decoder for de-interleaving, and error detection and correction. The CIRC decoder accepts one frame of 32 8-bit symbols: 24 audio symbols and 8 parity symbols. One frame of 24 8-bit symbols are output. Parity from two Reed-Solomon decoders is utilized. The first error correction decoder corrects random errors and detects burst errors and flags them. The second decoder primarily corrects burst errors, as well as random errors that the first decoder was unable to correct. Error concealment algorithms employing interpolation and muting circuits follow CIRC decoding.

In most cases, the digital audio data are converted to a stereo analog signal. This reconstruction process

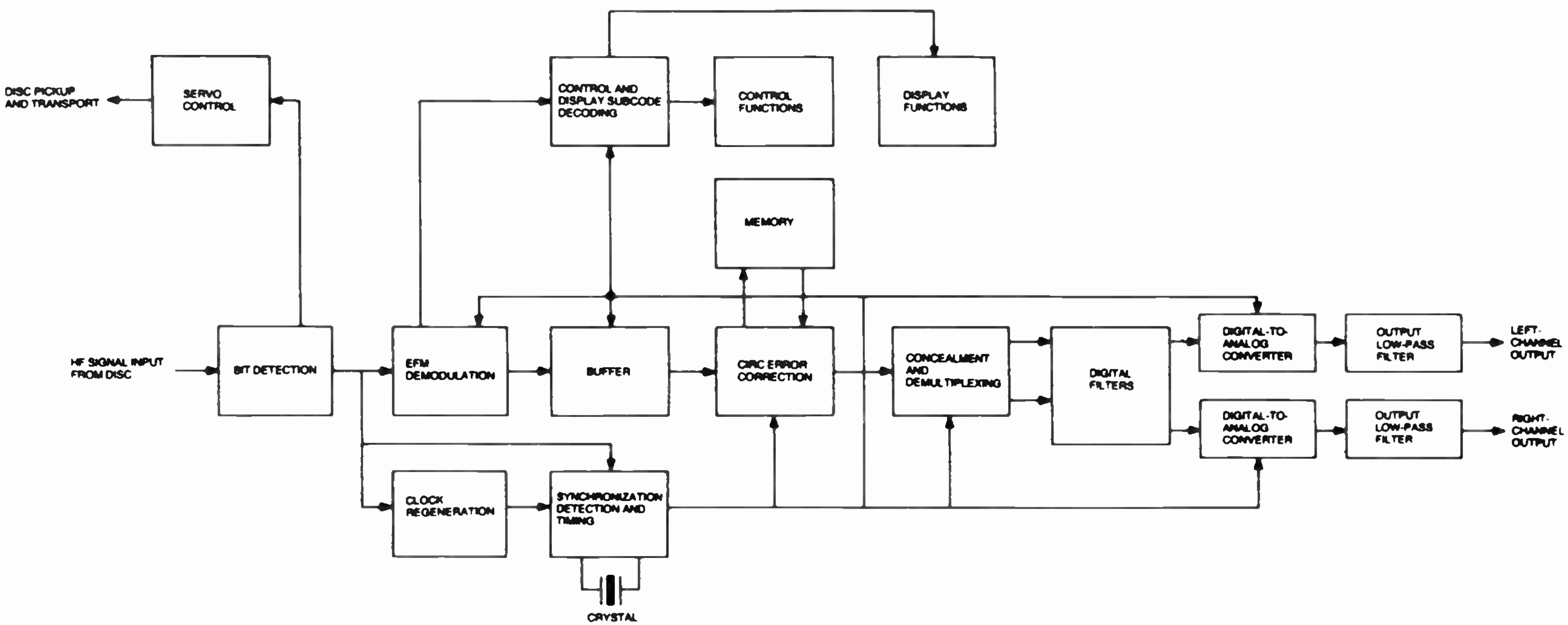


Figure 7. Block diagram of a CD player with digital filtering. (From Pohlmann, *The Compact Disc*.)

requires one or two digital-to-analog (D/A) converters, and low-pass filters to suppress high- frequency image components. Rather than use an analog brickwall filter after the signal has been converted to analog form, the digitized signal is processed before D/A conversion using an oversampling digital filter. An oversampling filter uses samples from the disc as input, then computes interpolation samples, digitally implementing the response of an analog filter.

A finite impulse response (FIR) transversal filter is used in most CD players. Resampling is used to increase the sample rate; for example, in a four-times oversampling filter, three zero values are inserted for every data value output from the disc. This increases the data rate from 44.1 kHz to 176.4 kHz. Interpolation is used to generate the values of intermediate sample points, for example, three intermediate samples for each original sample. These samples are computed using coefficients derived from a low-pass filter response.

The spectrum of the oversampled output waveform contains image spectra placed at multiples of the oversampling rate; for example, in a four-times oversampled signal, the first image is centered at 176.4 kHz. Because the audio baseband and sidebands are separated, a low-order analog filter can be used to remove the images, without causing phase shift or other artifacts common to high-order analog brickwall filters.

Traditionally, D/A conversion is performed with a multi-bit PCM converter. In theory, a 16-bit converter could perfectly process the 16-bit signal from the disc. However because of inaccuracies in converters, 18-bit D/A converters are often used because they can more accurately represent the signal. Alternatively, low-bit (sometimes called 1-bit) D/A converters can be used. They minimize many problems inherent in multi-bit converters such as low-level nonlinearity and zero-cross distortion. Low-bit systems employ very high oversampling rates, noise shaping, and low-bit conversion.

Also present in the audio output stage of every CD player is an audio deemphasis circuit. Some CDs are encoded with audio preemphasis characteristic with time constants of 15 and 50 microseconds. Upon playback, deemphasis is automatically carried out, resulting in an improvement in signal-to-noise ratio.

RECORDABLE CD-R FORMAT

With a CD-R (or CD-WO) write-once optical disc recorder, the user may record data until the disc capacity is filled. Recorded CD-R discs are playable on conventional CD players. A block diagram of a CD-R recorder is shown in Fig. 8. An encoder circuit accepts an input PCM signal and performs CIRC error correction encoding, EFM modulating, and other coding, and directs the data stream to the recorder. The recorder accepts audio data and records up to 74 minutes in real time. Some CD-R recorders combine a complete read/write unit, converters, encoder, decoder, and subcode generator in a chassis the size of a CD player. In addition to audio data, a complete subcode table is written in the disc table of contents, and appropriate flags are placed across the playing surface.

Write-once media is manufactured similarly to conventional playback-only discs. As with regular CDs, they employ a substrate, a reflective layer, and a protective top layer. Sandwiched between the substrate and reflective layer, however, is a recording layer comprised of an organic dye. Together with the reflective layer, it provides a typical in-groove reflectivity of 70% or more. Unlike playback-only CDs, a pregrooved spiral track is used to guide the recording laser along the spiral track; this greatly simplifies recorder hardware design and ensures disc compatibility. Shelf life of the media is said to be 10 years or more at 25° centigrade and 65% relative humidity. However the dye used in these discs is vulnerable to sunlight thus discs should not be exposed to bright sun over a long period.

The CD-R format is defined in the Orange Book standard authored by Philips and Sony. In CD recorders adhering to the Orange Book I standard, a disc must be recorded in one pass; start-stop recording is not permitted. In recorders adhering to the Orange Book II standard, recording may be stopped and started. In many players, tracks may be recorded at different times and replayed but because the disc lacks the final table of contents, it can be played only on a CD-R recorder. When the entire disc is recorded the interim TOC data are transferred to a final table of contents, and the disc may be played in any CD audio player. The PMA (program memory area) located at the inner portion of the disc contains the interim TOC record of the recorded tracks. In addition, discs contain a PCA (power calibration area); this allows recorders to automatically make test recordings to determine optimum laser power for recording. Some recorders exceed the Orange Book II standard; they generate an interim TOC which allows partially recorded discs to be played on playback-only CD players.

CD-R recorders are useful because they eliminate the need to create an edited master tape prior to CD recording. If a passage is not wanted, it can be marked prior to writing the final TOC so that the recorder will not play it back. For example, dead air during a live performance can be marked so it is deleted whenever the disc is played back. The data physically continues to exist on the disc, however.

ERASABLE CD-E FORMAT

Compact disc systems that provide for both recording and erasing are often known as CD-E systems. Erasable optical systems permit data to be written, read, erased, and written again. Several recordable/erasable optical media have been introduced, most notably, magneto-optical (MO) media. Magneto-optical recording technology combines magnetic recording and laser optics, utilizing the record/erase benefits of mag-

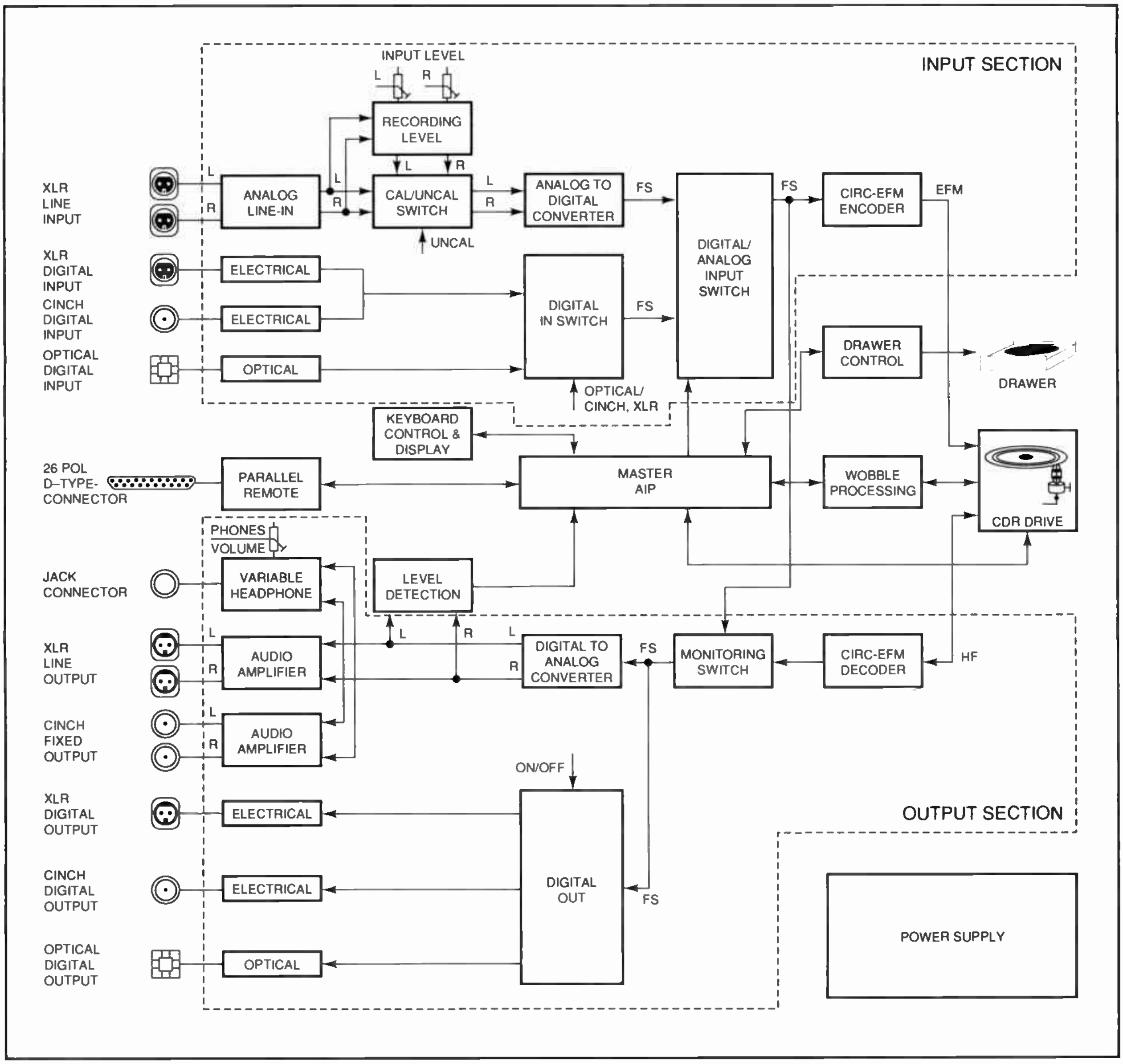


Figure 8. Block diagram of a CD-R recorder. (From Pohlmann, *The Compact Disc*.)

netic materials with the high density and contactless pick-up of optical materials.

With magneto-optics, a magnetic field is used to record data, but the applied magnetic field is much weaker than conventional recording fields. It is not strong enough to orient the magnetic particles. However the coercivity of the particles sharply decreases as they are heated to their Curie temperature. A laser beam focused through an objective lens heats a spot of magnetic material, and only the particles in that spot are affected by the magnetic field from the recording coil, as shown in Fig. 9A. After the laser pulse is withdrawn, the temperature decreases and the orientation of the magnetic layer records the data. In this way, the laser beam creates a small recorded spot thus increasing recording density.

The Kerr effect may be used to read data; it describes the slight rotation of the plane of polarization of polarized light as it reflects from a magnetized material. The rotation of the plane of polarization of light reflected from the reverse-oriented regions differs from that reflected from unreversed regions, as shown in Fig. 9B. To read the disc, a low-powered laser is focused on the data surface, and the angle of rotation of reflected light is monitored, thus recovering data from the laser light. To erase data, a magnetic field is

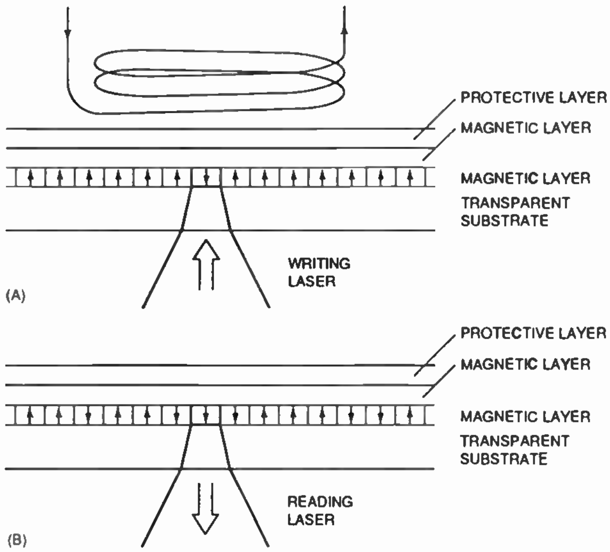


Figure 9. Magneto-optical recording (A) and playback (B).

applied to the disc, along with the laser heating spot. Tests indicate that magneto-optical media could be erased/recorded over 10 million times and would retain data for over 10 years.

DIGITAL AUDIO TAPE FORMAT

The rotary-head digital audio tape (R-DAT or DAT) format was originally designed as a consumer medium to replace the analog cassette. However the format has found wider application as a low-cost professional digital recording system.

Format Specifications

The DAT format supports four record/playback modes and two playback-only modes. The standard record/playback, and both playback-only modes, called "wide" and "normal", are implemented on every DAT recorder. The standard mode offers 16-bit linear quantization and 48 kHz sampling rate. Both playback-only modes use a 44.1 kHz sampling rate, for user- and pre-recorded tapes. Three other record/playback modes, called Options 1, 2, and 3, all use 32 kHz sampling rates. Option 1 provides 2-hour recording time with 16-bit linear quantization. Option 2 provides 4 hours of recording time with 12-bit nonlinear quantization. Option 3 provides four channel recording and playback, also using 12-bit nonlinear quantization. These specifications are summarized in Fig. 10.

The user can write and erase non-audio information into the subcode area: start ID indicating the beginning of a selection, skip ID to skip over a selection, and program number indicating selection order. This subcode data permits rapid search and other functions. Although subcode data are recorded onto the tape in the helical scan track along with the audio signal, it is treated independently and can be re-written without altering the audio program, and entered either during recording or playback. With the ID codes entered into the subcode area, desired points on the tape such as the beginning of selections can be searched for at high speed by detecting each ID code. During playback, if the skip ID is marked, playback is skipped to the point at which the next start ID is marked, and playback begins again.

In the DAT format, the recorded area is distinguished from a blank section of tape with no recorded signal, even if the recorded area does not contain an audio signal. Unlike blank areas, the track format is always

MODE / ITEM	DAT (REC/PB MODE)				PRERECORDED TAPE(PB ONLY)	
	STANDARD	OPTION 1	OPTION 2	OPTION 3	NORMAL TRACK	WIDE TRACK
CHANNEL NUMBER [CH]	2	2	2	42	2	2
SAMPLING FREQUENCY [kHz]	48	32	32	32	44.1	
QUANTIZATION BIT NUMBER [BIT]	16 (LINEAR)	16 (LINEAR)	12 (NONLINEAR)	12 (NONLINEAR)	16 (LINEAR)	16 (LINEAR)
LINEAR RECORDING DENSITY [kBPI]	61.0	61.0			61.0	61.1
SURFACE RECORDING DENSITY [MBPI²]	114	114			114	76
TRANSMISSION RATE [MBPS]	2.46	2.46	1.23	2.46	2.46	
SUBCODE CAPACITY [KBPS]	273.1	273.1	136.5	273.1	273.1	
MODULATION SYSTEM	8–10 MODULATION					
CORRECTION SYSTEM	DOUBLE REED-SOLOMON CODE					
TRACKING SYSTEM	AREA SHARING ATF					
CASSETTE SIZE [mm]	73 × 54 × 10.5					
RECORDING TIME [MIN]	120	120	240	120	120	80
TAPE WIDTH [mm]	3.81					
TAPE TYPE	METAL POWER					OXIDE TAPE
TAPE THICKNESS [μm]	13 ± 1 μ					
TAPE SPEED [mm/s]	8.15	8.15	4.075	8.15	8.15	12.225
TRACK PITCH [μm]	13.591				13.591	20.41
TRACK ANGLE	6°22'59.5"					6°23'29.4"
STANDARD DRUM SPECIFICATIONS	ø30 90° WRAP					
DRUM ROTATIONS [rpm]	2000		1000	2000	2000	
RELATIVE SPEED [m/s]	3.133		1.567	3.133	3.133	3.129
HEAD AZIMUTH ANGLE	±20°					

Figure 10. DAT standard specifications. (From Pohlmann, *Principles of Digital Audio.*)

encoded on the tape even if no signal is present. If these sections are mixed on a tape, search operations may be slowed. Hence, blank sections should be avoided. A consumer DAT deck with an interface meeting the specifications of the Sony Philips digital interface format (SPDIF) will identify when data have been recorded with a copy-inhibit SCMS (Serial Copy Management System) flag in the subcode (ID6 in the main ID in the main data area) and will not digitally copy that recording. In other words, SCMS permits first generation digital copying, but not second generation copying. Analog copying is not inhibited.

DAT Recorder Design

From a hardware point of view, a DAT recorder utilizes many of the same elements as a CD-R recorder: A/D and D/A converters, modulators and demodulators, error correction encoding and decoding. Audio input is received in digital form, or is converted to digital by an A/D converter. Error correction code is added and interleaving is performed. As with any helical scan system, time compression must be used to separate the continuous input analog signal into segments prior to recording, then rejoin them upon playback with time expansion to form a continuous audio output signal. Subcode information is added to the bit stream, and it undergoes eight-to-ten (8/10) modulation. This signal is recorded via a recording amplifier and rotary transformer.

In the playback process the rotary head generates the record waveform. Track finding signals are derived from the tape and used to automatically adjust tracking. Eight-to-ten demodulation takes place and subcode data are separated and used for operator and servo control. A memory permits de-interleaving as well as time expansion and elimination of wow and flutter. Error correction is accomplished in the context of de-interleaving. Finally, the audio signal is output as a digital signal, or through D/A converters as an analog signal.

The DAT rotary head permits slow linear tape speed while achieving high bandwidth. Each track is discontinuously recorded as the tape runs past the tilted head drum spinning rapidly in the same direction as tape travel. The results are diagonal tracks at an angle of slightly more than six° from the tape edge, as shown in Fig. 11. Despite the slow linear tape speed

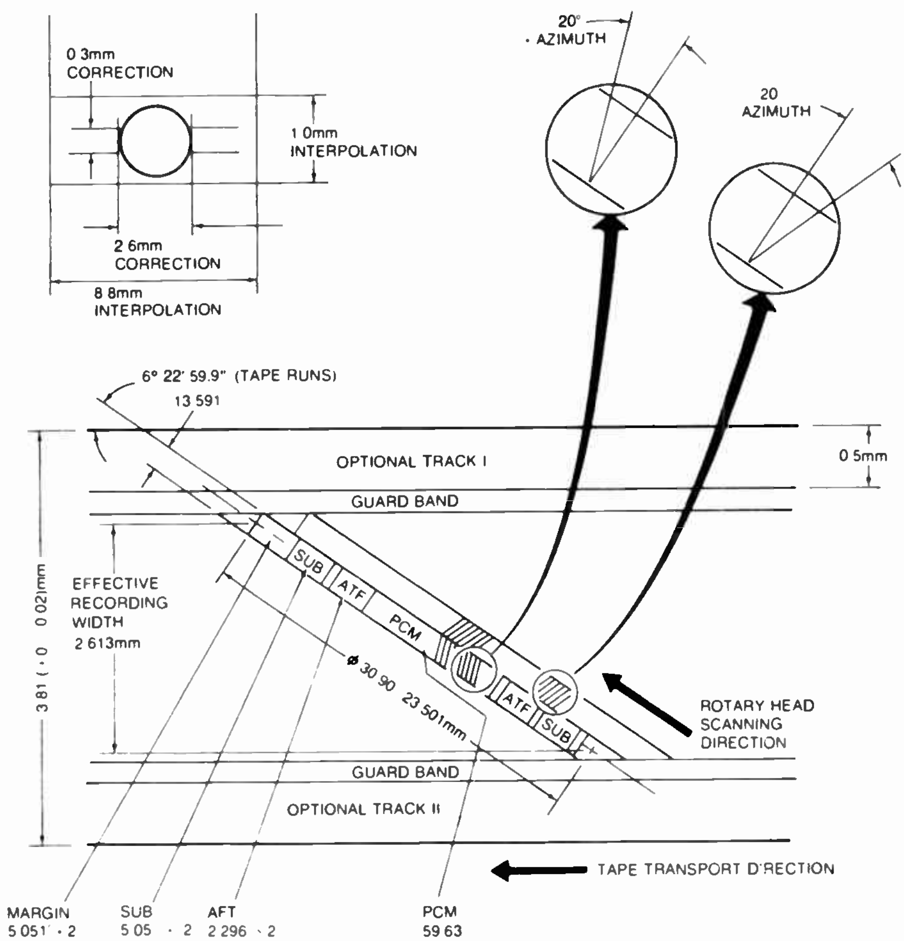


Figure 11. DAT track configuration. (From Pohlmann, *Principles of Digital Audio.*)

of 8.15 millimeters per second (¼ inch per second), a high relative tape-to-head speed of about 3 m/sec (120 inches per second) is obtained. A DAT rotating drum (typically 30 millimeters in diameter) rotates at 2,000 rpm, typically has two heads placed 180° apart, and a tape wrap of only 90°. Four head designs provide direct read after write, so the recorded signal can be monitored.

Azimuth recording, sometimes referred to as guardbandless recording, is used in which the drum's two heads are angled differently from each other with respect to the tape; this creates two track types, sometimes referred to as A and B, with differing azimuth angles between successively-recorded tracks. This ±20° azimuth angle means that the A head will read an adjacent B track at an attenuated level due to phase cancellation. This reduces crosstalk between adjacent tracks, eliminates the need for a guardband between tracks, and promotes high density recording. Erasure is accomplished by overwriting new data to tape such that successive tracks partially write over previous tracks. Thus the head gaps (20.4 microns) are approximately 50% wider than the tracks (13.59 microns) recorded to tape.

The length of each track is 23.501 mm. Each bit of data occupies 0.67 microns, with an overall recording data density of 114 megabits per square inch. With a sampling rate of 48 kHz and 16-bit quantization, the audio data rate for two channels is 1.536 megabits per second. However, error correction encoding adds extra information amounting to about 60% of the original, increasing the data rate to about 2.46 megabits per second. Subcode raises the overall data rate to 2.77 megabits per second.

The primary types of data recorded on each track are PCM audio, subcode, and ATF (automatic track finding) patterns. Each data (or sync) block contains a sync byte, ID code byte, block address code byte, parity byte, and 32 data bytes. In total, there are 288 bits per data block; following 8/10 modulation, this is increased to 360 channel bits. Four 8-bit bytes are used for sync and addressing. The ID code contains information on preemphasis, sampling frequency, quantization level, tape speed, copy-inhibit flag, channel number, etc. Subcode data are used primarily for program timing and selection numbering. The subcode capacity is 273.1 kilobits per second. The parity byte is the exclusive or sum of the ID and block address bytes, and is used to error correct them.

Since the tape is always in contact with the rotating heads during record, playback, and search modes, tape wear necessitates sophisticated error correction. DAT is thus designed to correct random and burst errors. Random errors are caused by crosstalk from an adjacent track, traces of an imperfectly-erased signal, or mechanical instability. Burst errors occur from dropouts caused by dust, scratches on the tape or by head clogging with dirt.

To facilitate error correction, each data track is split into halves, between left and right channels. In addition, data for each channel is interleaved into even and odd data blocks, one for each head; half of each channel's samples are recorded by each head. All of the data are encoded with a doubly-encoded Reed-Solomon error correction code. The error correction system can correct any dropout error up to 2.6 millimeters in diameter, or a stripe 0.3 millimeter high. Dropouts up to 8.8 millimeters long and 1.0 millimeter high can be concealed with interpolation.

SERIAL INTERFACING

Most professional digital audio devices employ an output protocol using the joint Audio Engineering Society (AES) and European Broadcasting Union (EBU) serial transmission format for digital audio data. It is known as the AES/EBU or AES3 format and is specified in the ANSI S4.40–1985 standard. Manufacturers of consumer electronics have adopted a derivative transmission format which has been standardized by the International Electrotechnical Commission (IEC), and is commonly referred to as the SPDIF (Sony/Philips Digital Interface Format). It is specified in the IEC Report TC 84/WG11 and IEC Publication 958.

The AES/EBU digital audio format transmits and receives left and right channel data using one digital cable. The transmission rate corresponds exactly to the source sampling frequency. One frame consists of two subframes, labelled A (left channel) and B (right channel), each with 32 bits. Each subframe contains data for one audio channel. The first four bits are used for synchronization and identifying preambles. The next 24 bits carry audio data, with the MSB transmitted last; 16-bit audio data leaves 4 bits at zero; the first 4 bits in the field are set aside for auxiliary audio or other data, as shown in Fig. 12.

The last four bits form a control field with the V, U, C, and P bits. The validity (V) bit indicates if the previous audio sample is error-free. The user (U) bit can be used to form a block of user data associated with the audio channel. The channel (C) status bit is used to form a data block; for each channel, one block is formed from the channel status bit contained in 192 successive frames. The parity (P) bit is used to provide even parity for each subframe. The AES/EBU standard specifies that data are transmitted over twisted pair conductors, with 3-pin XLR connectors.

In the IEC or SPDIF serial format used in consumer digital audio equipment, the first bit of the channel status byte is set to 0 to signify a consumer interface, and a different channel status specification is used. In addition, when interfacing CD data, provision is made to transmit the subcode data in the user bit channel. The consumer format also contains provision for SCMS (Serial Copy Management System) in the channel status bits. In particular, bit 2 is used to flag copyrighted material, and bit 15 distinguishes between original and copied material. In the IEC or SPDIF format, video coaxial cable with phono plugs or fiber optic cables can be used to convey data.

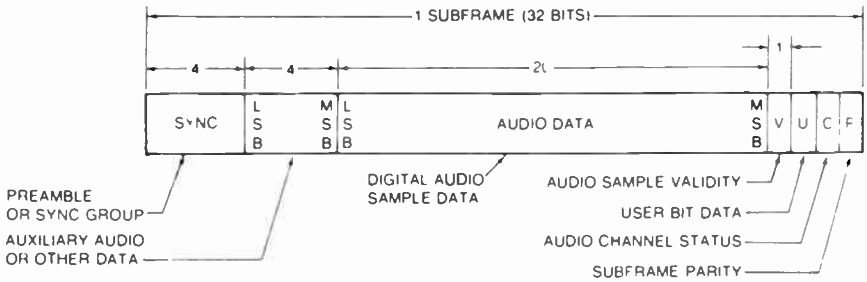


Figure 12. AES/EBU subframe format. (From Pohlmann, *Principles of Digital Audio.*)

Digital interconnection formats such as AES3 and SPDIF allow audio data to be transferred from one device to another without any generation loss whatsoever. For example, audio data from a DAT recording could be conveyed to a hard disk editing system, processed there, and returned to DAT without transmission error. However it is important to only connect AES/EBU outputs to AES/EBU inputs, and IEC/SPDIF outputs to IEC/SPDIF inputs. AES/EBU and IEC/SPDIF interfaces should not be interconnected. Transmitted data could be invalid, or lead to improper machine operation. Because consumer DAT recorders employ the SPDIF interface, and contain SCMS circuits, they should not be used for professional applications.

DIGITAL SIGNAL PROCESSING

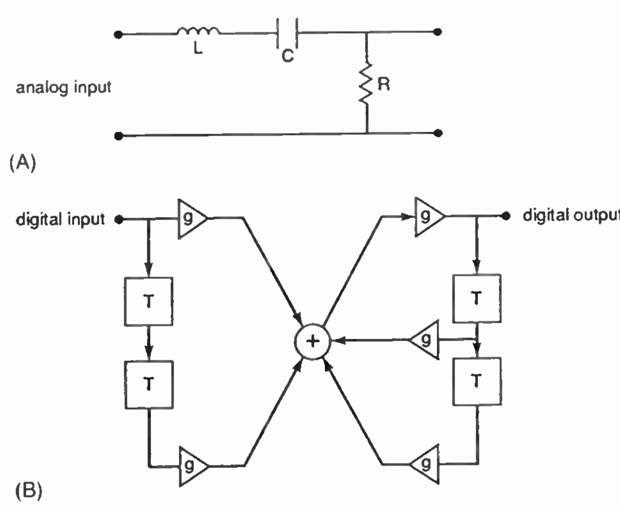
Digital signal processing (DSP) has improved the performance of many existing audio functions such as equalization and dynamic range compression, and permits new functions such as ambience processing, dynamic noise cancellation, and time alignment. DSP is a technology used to analyze, manipulate, or generate signals in the digital domain. It uses the same principles as any digitization system however instead of a storage medium such as CD or DAT, it is a processing method.

DSP Applications and Design

DSP employs technology similar to that used in computers and microprocessor systems, however, there is an important distinction. A regular computer processes data, whereas a DSP system processes signals. It is accurate to say that an audio DSP system is in reality a computer dedicated to the processing of audio signals.

Some audio functions DSP can perform include: error correction, multiplexing, sample rate conversion, speech and music synthesis, data compression, filtering, adaptive equalization, dynamic range compression and expansion, crossovers, reverberation, ambience processing, time alignment, acoustic noise cancellation, mixing and editing, and acoustic analysis. Some DSP functions are embedded within other applications; for example, the error correction systems and oversampling filters found in CD players are examples of digital signal processing. In other applications the user has control over the DSP functions.

Digital processing is more precise, repeatable, and can perform operations that are impossible with analog techniques. Noise and distortion can be much lower with DSP, thus audio fidelity is much higher. In addition, whereas analog circuits age, lose calibration, and are susceptible to damage in harsh environments, DSP circuits do not age, cannot lose calibration, and are much more robust. However DSP technology is an expensive technology to develop. Hardware engineers must design the circuit or employ a DSP chip, and software engineers must write appropriate programs. Special concerns must be addressed when writing the code needed to process the signal. For example, if a number is simply truncated without



MPY	X0,Y1,A	X: (R0) +, X0	Y: (R4) +, Y0
MAC	X0,Y0,A	X: (R0) +, X0	Y: (R5) +, Y0
MAC	X0,Y0,A	X: (R0) +, Y0	Y: (R5) +, Y0
MAC	X0,Y0,A	X: (R0) +, X0	Y1,Y: (R4) +
MOVE		A,X1	A,Y: (R5)

(C)

Figure 13. A band-pass filter represented by an analog circuit (A), digital signal processing circuit (B), and digital signal processing instructions (C). (Courtesy Motorola.)

regard to its value, a significant error could occur, and the error would be compounded as many calculations take place, each using truncated results. The resulting numerical error would be manifested as distortion in the output signal. Thus all computations on the audio signal must be highly accurate. This requires long wordlengths; DSP chips employ digital words that are 32 bits in length, or longer.

In addition, even simple DSP operations may require several intermediate calculations and complex operations may require hundreds of operations. To accomplish this, the hardware must execute the steps very quickly. Because all computation must be accomplished in real time, that is, within the span of one sample period, the processing speed of the system is crucial. A DSP chip must often process 20 or 30 million instructions per second. This allows it to run complete software programs on every audio sample as it passes through the chip.

DSP products are more complicated than similar analog circuits, but DSP possesses an inherent advantage over analog technology; it is programmable. Using software, many complicated functions can be performed entirely with coded instructions. Fig. 13A shows a band-pass filter using conventional analog components. Fig. 13B shows the same filter, represented as a DSP circuit. It employs the three basic DSP operators of delay, addition, and multiplication. However this DSP circuit may be realized in software terms. Fig. 13C shows an example of the computer code (Motorola DSP56001) needed to perform band-pass filtering with a DSP chip. There are many advantages to this software implementation. Whereas hardware circuits would require new hardware components and new circuit design to change their processing tasks, the software implementation could be changed by altering parameters in the code. Moreover the program could be written so different parameters could be employed based on user control.

As noted, digital signal processing can be used in lieu of most conventional analog processing circuits. The advantages of DSP are particularly apparent when various applications such as recording, mixing, equalization, and editing are combined in a workstation. For example, a Macintosh personal computer, combined with a DSP hardware card, hard disk drive, and appropriate software, and a DAT or CD recorder forms a complete post-production system. Such a system allows comprehensive signal manipulation including ability to cut, paste, copy, replace, reverse, trim, invert, fade-in, fade-out, smooth, loop, mix, change gain and pitch, crossfade, and equalize. The integrated nature of such a workstation, its low cost, and high-processing fidelity make it clearly superior to analog techniques.

DSP is used to perform data reduction, sometimes known as data compression; this is an important aspect of digital audio broadcasting (DAB) technology, as well as new storage formats such as digital compact cassette (DCC) and mini disc (MD). A high-quality data reduction algorithm can eliminate perhaps 75% of the audio data from a recording, yet provide a signal fidelity that audibly is virtually indistinguishable from the original. However, whereas recordings of linear data can be cloned, *copies of reduced data recordings will show deterioration with successive passes through the compression and decompression algorithms.*

Data reduction techniques have many important applications. In the case of DAB, it would be impractical to transmit digital audio signals in a linear PCM format because the bandwidth requirements would be too great. Instead, DAB must employ data compression to reduce the broadcast data by ratios of four to one or more. For example, the Eureka 147 broadcast system uses MUSICAM data compression in which sub-band ADPCM (adaptive delta pulse code modulation) is employed. Dolby's AC-2 system uses transform coding of 256 spectral bands. Such systems can compress a standard sample rate 16-bit signal to a 256, 128, 96, or 64 kps signal. New data reduction techniques, along with other new DSP technologies, will provide further benefits in digital audio processing, storage, and transmission for broadcasters.

BIBLIOGRAPHY

Audio Engineering Society, "AES Recommended Practice for Digital Audio Engineering—Serial Transmission Format for Linearly Represented Digital Audio Data," *Journal of the Audio Engineering Society*, vol. 33, no. 12, December, 1985.

Carasso, M.G., Peck, J.B.H., Sinjou, J.P., "The Compact Disc Digital Audio System," *Philips Technical Review*, vol. 40, no. 6, 1982.

EBU (European Broadcasting Union). "Specification of the Digital Audio Interface." EBU Doc. Tech., 3250.

IEC (International Electrotechnical Commission). "Draft Standard for a Digital Audio Interface." IEC Report TC 84/WG11, November 1, 1986.

Pohlmann, Ken C. *The Compact Disc*, Second Edition. AR Editions, Madison, WI, 1992.

Pohlmann, Ken C., *Principles of Digital Audio*, Second Edition. Howard W. Sams and Co., Carmel, IN, 1989.

Pohlmann, Ken C., editor, *Advanced Digital Audio*. Howard W. Sams and Co., Carmel, IN, 1991.

5.6 Videotape Recording

Stanley N. Baron
National Broadcasting Company, Inc., New York, New York

INTRODUCTION

Abstract

This chapter provides an overview of video recording technology. A brief history, a definition of basic terms, and the technology employed is presented first. This is followed by sections explaining the principles and characteristics of each of the tape formats widely used in professional applications. Sections on maintenance and system features complete the chapter.

History

The development in the 1930s and 1940s of electronic television led to attempts to record the electronic signal. Early developments were based on audio recording technology, which produced acceptable, although limited results. These first longitudinal machines required large reels of tape and very high tape speed.

In the spring of 1956, a practical videotape recorder was introduced by Ampex. It dramatically changed the operational and production methodology of the television broadcast industry. The culmination of over five years of development effort, the system was intended to replace the "hot kine" film process used to record television programs for time delay rebroadcast and to allow for the editing of television programs. These first machines boasted a 4 MHz video bandwidth, audio bandwidth to 10 KHz and video peak-to-peak signal-to-noise ratio of 30 dB.[1]

These first recorders were based on a quadruplex recording technique, recording on open reels of two-inch wide tape. These units were neither portable nor able to withstand hazardous environments. Since that time, the television industry has strived to improve on this technology by developing smaller, more reliable, and more portable recorders that use less tape, cost less to purchase, operate, and maintain, and can operate in the environments necessary to accommodate a wide variety of applications.

The quadruplex recording technique was superseded by helical scan format technology. Helical scan format techniques had been investigated during the development of the videotape recorder at Ampex but progress was impeded by two problems. First, the video track in a helical recorder is at a small angle with respect to the direction of motion; and any instability in tape movement produces very large time base errors on the order of several microseconds. These errors made display on monitors and adherence to FCC timebase and frequency stability transmission tolerances difficult to meet. The quadruplex format, however, tracks at nearly right angles to the tape motion and only small amounts of longitudinal speed errors are introduced into the video signal. These small errors can be corrected using very limited time base correction in the form of an electronically variable analog delay line. With the introduction of wider range time base correctors (TBCs) and improvements in materials, the simpler mechanical requirements and lower costs associated with the helical scan technology could be applied to professional video products.

Professional helical scan formats were introduced in Europe (Bosch), Japan (Sony), and the United States (Ampex) at approximately the same time, all based on one inch wide oxide tape. The Sony and Ampex recorders were very similar. The tape was almost completely wrapped around the scanner and recorded one television field per revolution of the scanner. Scanner diameter, tape speed, and other characteristics were similar. The small gap where the wrap was incomplete resulted in the loss of approximately 10 lines of the vertical interval. Sony employed an extra set of heads on the scanner to record the missing lines; Ampex chose to ignore them. The Society of Motion Picture and Television Engineers (SMPTE) formed an engineering committee to attempt the creation of a single standard based on the Ampex and Sony systems. The result of that effort was an agreement on a single format designated as the one inch C format. Within a short period the C format achieved acceptance as the professional production standard of choice.[2]

The ¾-inch U format cassette based system was conceived originally as a lower cost, limited bandwidth, nonbroadcast system. The heterodyned signal was sufficiently stable to lock a monitor but could not be broadcast. The advent of low cost TBCs and portable versions of the U format system made the format practical for broadcast news operations. Its obvious advantages over film with respect to turnaround time, program length, and general adaptability to news gathering made it the *de facto* electronic news gathering (ENG) standard of its time.

The introduction of the ½-inch VHS and Beta cassette tape formats was directed at the consumer market and featured long playing times at the sacrifice of professional performance levels. Subsequent improvements in tape technology offered an opportunity to provide ½-inch tape formats meeting various levels of performance for specific professional applications. Four different ½-inch analog component formats were introduced in the 1982–1987 time period. The four systems, in the order of introduction, are the M format, Betacam format, M-II format, and BetacamSP format. They led to the introduction of a new generation of light weight, small systems, and the introduction of the handheld camcorder. The S-VHS system was designed as an enhanced consumer system offering 400 lines of horizontal resolution in the luminance channel. Application of this lower cost format in professional applications is constrained by limited multiple generation operation. Equipment with professional features is now being offered.

The introduction of digital technology in professional broadcasting applications was inevitable. The international agreement on a component digital video interface based on a luminance sampling frequency of 13.5 MHz for both 525 line and 625 line systems, as documented in CCIR Recommendation 601, provided the basis for video data recording in the digital domain. Manufacturers offer digital systems in both the Recommendation 601 component (Y, R-Y, B-Y) and NTSC composite recording formats. The component digital videotape format (D-1) system makes multigeneration copies of previously unobtainable quality. The composite format (the 19 mm, D-2 and ½-inch, D-3) systems provide improved multigeneration capabilities when compared to analog techniques, being only limited by the quality of the external NTSC decoding/encoding circuitry applied in those post-production applications requiring a return to the component parameters. The composite format systems provide improved recording efficiency when compared to the D-1 component format system and are finding application as replacement for one-inch C format systems in general broadcast applications. The evolution of digital recording systems is following the path of the analog systems with each succeeding generation offering more recording time per volume of tape.

QUADRUPLEX RECORDERS

The quadruplex recording system required several years to come to maturation and eventually achieved great success. The evolution of tape recording technology made the introduction of the professional helical tape recorders possible and the quadruplex systems were superseded.

There are still some quadruplex recorders in service throughout the world. The Sixth Edition of the *NAB Engineering Handbook* contained a complete description of the format and is recommended for those who require information.

HELICAL SCAN RECORDERS—DEFINITIONS AND CHARACTERISTICS

The following basic definitions apply to the helical-scan technology. The definitions are adapted from terms defined by the SMPTE.[3,4,5]

DRUM: A cylindrical column around which the tape is at least partially wrapped in such a manner as to permit rotating pole tips protruding radially from the column to form the head-to-tape interface of a videotape recording system.

SCANNER: A mechanical assembly containing a drum, rotating pole tips, and tape-guiding elements used to record and reproduce videotape recordings.

LOWER DRUM: That part of a drum in a helical-scan videotape recording system which contacts the reference edge of the tape and usually includes tape guiding elements.

EFFECTIVE DIAMETER: The effective diameter is the diameter at the surface of tape wrapped around the drum which includes the drum diameter and the air film between the drum and the tape.

HELIX ANGLE: The angle formed between the path of the rotating pole tips and the tape reference edge-guiding system on the scanner of a helical-scan videotape recording system.

TRACK ANGLE: The angle of a video record with respect to the reference edge of the tape in a helical-scan videotape recording system.

TRACK LENGTH: Length of the actual recorded track with the tape under normal tension.

TRACK PITCH: The centerline-to-centerline distance between two adjacent tracks measured perpendicularly to the tracks.

TRACK WIDTH: The width of the recorded track.

WRITING SPEED: The relative video pole tip to tape speed.

WRAP ANGLE: An angle that is positioned with its apex on the center of drum rotation and subtended by the line of contact between the drum and the reference edge of the tape.

VERTICAL INTERVAL DROPOUT: That part of a television field not recorded by the video pole tip due to the tape being wrapped around the drum less than 360 degrees.

CENTER SPAN TENSION: A calculated value of the tape tension at a point midway between the tape entrance and exit guides of the scanner in the digital audio and video recording system.

The following basic definitions apply to the digital recording technology. The definitions are adapted from terms defined by the SMPTE.[6]

PROGRAM AREAS: That part of the tape on which is digitally recorded the program video and audio signals.

PROGRAM AREA TRACK PATTERN: The arrangement of video and audio sectors on helical-scan tracks within the program area.

VIDEO AND AUDIO SECTORS: A sector is a structured sequence of data which incorporates the video or audio data and appropriate synchronizing and identification patterns, so that the video or audio data can be recovered from tape and identified for subsequent processing.

PREAMBLE: A preamble consists of a runup sequence, a sync pattern, an identification pattern, and some fill data.

RUNUP SEQUENCE: A runup sequence consists of a sequential bit pattern chosen to facilitate the locking of data-extraction circuits.

SYNC PATTERN: A sync pattern consists of two or more consecutive bytes whose bit pattern is chosen to be a robust indication of the start of the sync block.

IDENTIFICATION PATTERN: An identification pattern consists of two to four consecutive bytes providing a unique address of the position of the sync block within two to four frames of recorded data. It may be coded to remove direct current and provide error protection.

FILL DATA: Fill data consists of a few bytes of a fixed pattern which is designed to provide a minimum separation on tape between a runup sequence, a sync pattern, and the first sync block. Fill data may also be recorded in the edit gap between sectors in the track.

SYNC BLOCK: A sync block consists of a sync pattern followed by an identification pattern followed by inner code blocks.

INNER CODE BLOCK: A inner code block consists of a number of bytes of video data, audio data, or outer code check data, followed by a number of inner code check data. A D-2 or D-3 inner code block may include an identification pattern.

POSTAMBLE: A postamble consists of a sync pattern followed by an identification pattern and, possibly, some fill data.

The Signal System

Basic Principles

Video recorders all utilize the same basic principles. The technology is based on the principle that if the current through a solenoid is varied, the strength of the effective magnet and its associated field is varied. The magnetism of any ferrous material moved through the field of the solenoid also will be varied. In videotape recorders, the record head is the solenoid and the ferrous material on the tape is affected as it is moved past the head (Fig. 1).

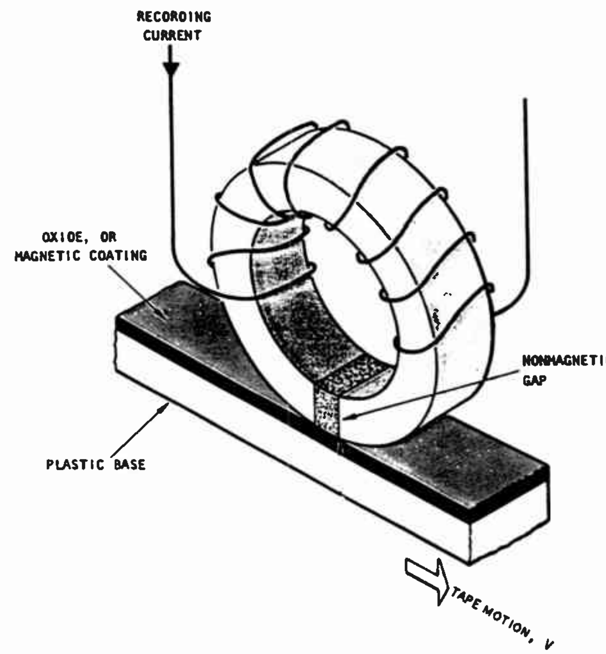


Figure 1. Magnetic recording process.

Frequency Modulation

In analog recording systems, (quadruplex, C format, U format, and the ½-inch analog systems) the varying amplitude video signal is converted to a frequency modulated (FM) signal for recording on the tape. This is accomplished by deviating a carrier frequency in response to the amplitude of the video (or audio) signal.

An FM signal can be amplified to a level sufficient to saturate the tape during recording, ensuring good signal strength. The demodulation of the FM signal is insensitive to amplitude variations and performs well in the presence of noise, since in the recovery (playback) cycle, changes in the carrier's amplitude are essentially removed by limiting, and the original video signal is recovered by detecting zero crossings.[7]

During playback, the tape passes the heads and the magnetism of the ferrous material on the tape generates a current in the head coils. The voltage produced is given by:

$$v = kNR$$

where: k is a constant
N is the number of turns of wire in the coils
R is the rate of change of flux ($d\phi/dt$)

The output voltage is, therefore, directly proportional to the rate at which the magnetic flux is changing. If the frequency of the signal recorded on the tape is doubled, then the rate at which the flux changes is doubled, and the output voltage is doubled (the output of the signal rising at 6 dB per octave).

In conventional NTSC systems, the video signal would range from about 30 Hz to 4.5 MHz, a span of approximately 18 octaves. This is almost an impossible range to accommodate. However, if the video signal is modulated onto a high frequency carrier such as a 5 MHz carrier, then the same range of video signals can be accommodated in a range from 5 MHz to 10 MHz which represents a single octave.

Digital Recording Systems

In digital recording systems, the video signal is presented as a sequence of digital words, each word consisting of bits, each bit having a value of either "1" or "0". In the component digital video recording system (D-1), NRZ channel coding is used. In the 19 mm composite digital video recording system (D-2), Miller-square channel coding is used. In the ½-inch composite digital video recording system (D-3), an 8-14 modulation system is used. An explanation of each encoding method is found in the "Principles and Characteristics" section of each for the system descriptions. For all systems, during the time interval of a recorded data "1", the polarity of the resulting flux is such that the north pole of the magnetic domain points in the direction of head motion; and during the interval of recorded data "0", the polarity of the resulting flux is such that the south pole of the magnetic domain points in the direction of head motion.

Recovery of the digital information consists of using the polarity of flux to induce a current in the read head whose direction is appropriately interpreted as either a logic "1" or a logic "0".

Recording and Playback Systems

Analog Recording Systems

The analog recording systems all work in a similar manner, since they are based on frequency modulation technology. Fig. 2 shows a very basic block diagram for the C Format system, but the basic principles are applicable to all systems.

The input video signal is terminated and amplified. It is at this first stage that any clamping and gain adjustment occurs. A sync separator provides the necessary pulses to meet servo subsystem requirements. In C format systems, the signal is fed to a burst amplifier to provide the 6 dB gain in burst signal with respect to video and sync required by SMPTE RP 86. The resulting signal is preemphasized as required within each system, frequency modulated, amplified, and applied to the tape head.

During playback, the signal from the heads is coupled to preamplifiers mounted as closely to the heads as possible to minimize noise. On most professional helical recorders, the preamplifiers are mounted on the

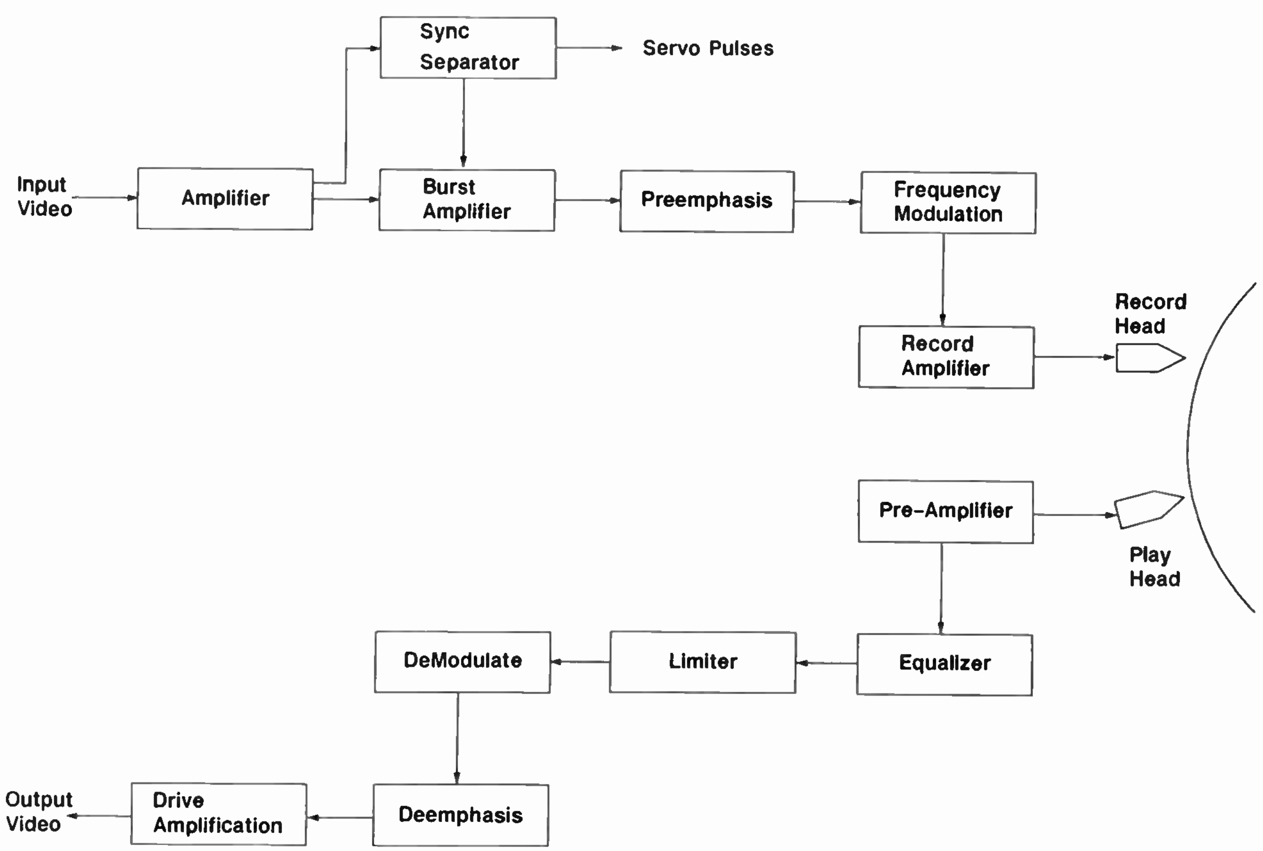


Figure 2. FM recording process.

scanner mechanism and rotate with the heads. The amplified signal is then equalized to correct for system frequency response characteristics producing a flat response. The equalized signal is processed through a limiting amplifier and demodulated. After demodulation, deemphasis and amplification produce a final reconstructed video signal.

Various techniques and circuits have been developed to record and recover the video signal. Variations in techniques and circuits are dependent both on system requirements and manufacturer implementation. The user should understand the characteristics and limitations of the system being used to determine its appropriateness with respect to any given application.[8]

Digital Recording Systems

In digital recording systems, the television field is considered an orderly set of elements arranged by rows (lines of the field) and columns (picture element samples or pixels within each line). This orderliness occurs when the sampling frequency is an even multiple of the line frequency. Such sampling frequencies produce an orthogonal sampling structure which means that samples repeat on a line, field, and frame basis. Ancillary data contained within the blanking portions of the field can be considered in a similar manner. Each pixel or ancillary data sample location can be identified by a pair of integers (i,j) where i represents the row or line number (thereby indicating the vertical position) and j identifies the column or horizontal location. In composite systems (D-2 and D-3 formats), each element represents the luminance and chrominance video information contained within the sample (Sij). In component systems (D-1) each element has associated with it luminance (Yij) and two cosited chrominance values ($CBij$, $CRij$) where $CBij$ and $CRij$ designate scaled B-Y and R-Y components respectively.

In the D-1 component system, the luminance sampling frequency is 13.5 MHz (858 times NTSC line frequency) and the chrominance sampling frequency is at 6.75 MHz (13.5 MHz ÷ 2). In the composite systems, the video sampling frequency is 14.318 MHz (4 times color subcarrier = 910 times the NTSC line rate).[9,10]

In all of the digital systems, the digital video signals are in the form of eight-bit words, according to the structure and temporal organization provided for in the designated sampling system. The input video data words are remapped according to a specific redistribution, the data word shuffled, and various forms of error correction and concealment coding added. Audio, ancillary data, and synchronization information are handled in a similar manner. The scheme used for each system varies and a more detailed description can be found in the "Principles and Characteristics" section for each tape format.

Digital System Error Protection Strategy

Various methods can be used to reduce the effect of data errors on the quality of replayed video and audio signals. The combination of methods used in any one system is the error protection strategy for that particular system. Error protection techniques include error correction, error concealment, and source precoding.

Error correction uses mathematically related check data, recorded with the audio and video data, to determine the value and location of the error and thereby enable correction of data errors. BCH (Bose-Chaudhuri-Hocquenguem) codes constitute a class of cyclic codes which employ simple decoding algorithms while exhibiting powerful error-correcting properties. BCH codes can be implemented using polynomial generator circuits in which the data stream is fed through a multistage shift register whose output produces the resultant and is also fed back, multiplied by coefficients and added into the shift register at specified points. Reed-Solomon (RS) codes are a subclass of BCH codes which can be designed to correct a specified number of errors. An RS code of class (N,k) is understood to have a code length of N symbols, k message symbols, and N-k parity check symbols. All three digital VCR tape formats described in this chapter (D-1, D-2, and D-3) utilize Reed-Solomon coding.[11]

Error concealment uses information interpolated from adjacent samples as an estimate of the value of the data words which are in error and replaces the incorrect data word with the estimated value.

Source precoding remaps the data words so that for the most probable distribution of data errors there is a reduction in the peak error produced in a video sample. Source precoding is used in the type D-1 and the ½-inch digital formats and is not used in the type D-2 format.

Tape Transports

Helical Recorders

The purpose of a tape transport is to move the tape past the magnetic record and playback heads with precise control of position, tension, and velocity. Helical-scan recorders can be viewed as having a three-dimensional tape path with the tape entering the drum from one plane of travel and exiting from another. Helical-scan recorders employ what is known as an omega wrap. Viewed from above, the tape path in the drum area resembles the Greek capital letter omega (Ω) (Fig. 3). In this pattern, the entry and exit planes are inclined with respect to the drum surface but in

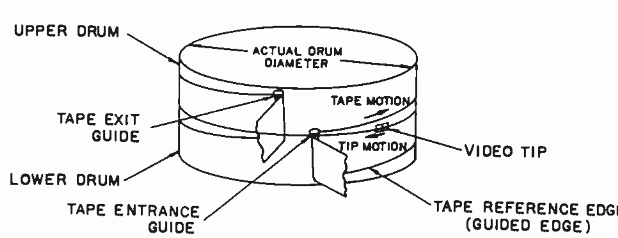


Figure 3. Helical scanner.

opposite directions. A series of tilted guides is used to return the tape from one plane to another. A 90-degree wrap on the exit and entry guides followed by a twist equal to the helix angle brings the tape paths into parallel configuration so that the reels of tape or the spools within the cassette can be coplaner. Maintaining proper position, tension, and velocity is a complex problem requiring very sophisticated servo control.

Capstan and Scanner Servos

Precise speed control is provided by the capstan and scanner servos. During the recording process, they are referenced to the synchronizing information in the video signal to be recorded. Capstan speed directly controls the longitudinal tape speed and is related to the vertical synchronizing signal. A control track signal is derived from the video synchronizing signal and is recorded on the tape. Scanner position affects the rotational position of the record and play heads. The scanner generates a pulse whose position with respect to the video record head is known and this pulse is compared with the horizontal and the vertical synchronizing pulse. The results of the comparison are used to adjust the scanner position appropriately. During playback, the servo system compares the recovered control track signal with an external reference sync (or if reference sync is not available, an internal oscillator) and adjusts the speed and phase of the capstan motor. The recovered vertical and horizontal synchronization pulses off the tape are used in a similar manner to control the scanner position.

Tape Tension Servos

Tape tension around the scanner during record and playback must be the same to avoid errors in the recovered pictures. This is accomplished by servoing both the supply and takeup reel motors during record and playback. Sensors translate the position of the tension arms near the reels into an error signal that is used to drive the reel motors. Tape tension should be checked on a periodic basis or when interchangeability problems arise.[12]

Video Head Position Servo

The transverse motion of the head relative to tape on helical-scan recorders during nonstandard playback modes such as slow-motion resembles a saw-tooth waveform. It follows the track at the scanning rate vertical interval in which no data are recorded (equal to 10 lines in C format machines). The tracking system must continuously respond to these changes without affecting the picture. Ampex developed a system based on moving a head relative to the rotating portion of the scanner and the tape. They called this system Automatic Scan Tracking (AST). Others refer to the system as a dynamic tracking system.

Dynamic tracking uses a piezoelectric element on which the head is mounted (Fig. 4). A bi-morph bender element is used because it has low mass, fast response, and is capable of large deflections and, further, can be designed to be rigid in directions other than the

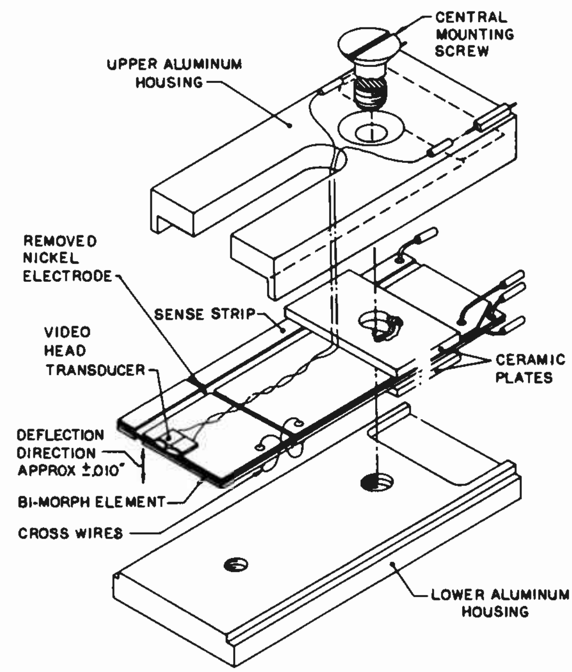


Figure 4. Ampex AST video playback head.

direction of deflection. The position of this head subsystem is controlled by a closed loop servo system with a deflection range of several video tracks. This range enables the servo to compensate for tape speed tracking errors in slow motion and still frame operation.[13]

ONE-INCH C FORMAT ANALOG SYSTEM

Introduction

Early in the introduction of one-inch helical-scan recorders,the Society of Motion Picture and Television Engineers (SMPTE) formed an engineering committee to see if a common format could be found upon which consensus could be reached. After about one year of effort, the SMPTE committee drafted an acceptable document. This second generation professional videotape recording system soon became the workhorse of the professional television industry. While it took almost 25 years to build a world population of 15,000 quadruplex recorders, it took less than one-third that time to populate more than double that number in one-inch helical scan machines with the type C format outnumbering the type B by about a 7:1 ratio.[14]

Principles and Characteristics

The C format system employs a helical-scan architecture with the tape unspooling from the supply reel, and looping around a tension arm that senses the tape as it leaves the supply reel (Fig. 5). The tape then passes over a full width erase head, past a guide and is presented to the scanner drum in such a manner that the heads slant across the tape to form part of a helix.

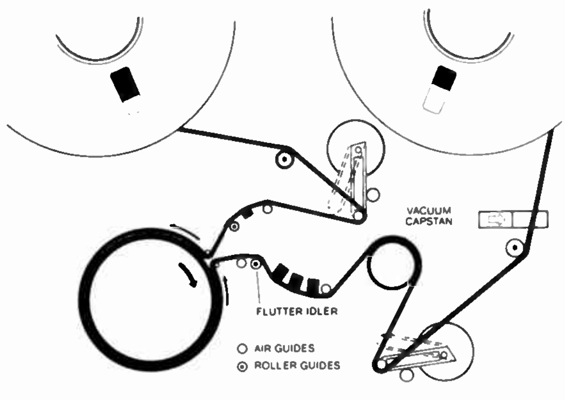


Figure 5. Tape path of Ampex VPR-3.

The C format system provides for up to three sets of heads, placed 120 degrees apart on the scanner. The tape is wrapped around the scanner for slightly more than 340 degrees (Fig. 6). Each set of video heads consists of two groups of heads, one for the video channel and one for recording the missing ten lines of sync. Each group consists of an erase head, a record head, and a playback head. The C format defines the sync head group as optional, but if not used, dummy heads must be installed to minimize velocity errors by making sure that all machines present the same basic mechanical configuration. A common mechanical configuration is required to optimize the interchange of videotapes between machines. Errors would arise if the second, optional set of heads were not replaced by dummies since a video head traveling across the tape produces a bow wave which is reflected back into the path of the following head. If the wave were present during recording on one machine but absent during playback on another machine, substantial errors would result.

The scanner rotates at approximately 3600 rpm, recording one TV field per revolution. The pattern of the video and sync record locations on the tape is shown in Fig. 7. The C format system also provides for four longitudinal tracks, three for audio (Audio 1, 2, and 3), and one control track. The video tracks do not interfere with the audio tracks and can be separately edited. Audio 3 has wide-band capability to accommodate time code.

During recording, the amplitude of the incoming video signal is converted to an fm modulated signal with the following reference levels:[15]

Peak-white: 10.00 ± 0.05 MHz
Blanking: 7.90 ± 0.05 MHz
Sync-tip: 7.06 MHz nominal

Further, the amplitude of the current applied to the record head windings is decreased with increasing frequency.

Applicable Standards Documents

1. SMPTE/ANSI V98.18M—American National Standard for Video Recording—1 inch Type C Helical Scan—Basic System and Transport Geometry Parameters.
2. SMPTE/ANSI V98.19M—American National Standard for Video Recording—1 inch Type C—Records.
3. SMPTE/ANSI 20M—American National Standard for Video Recording—1 inch Type C Recorders and Reproducers—Frequency Response and Reference Level.
4. SMPTE/ANSI 24M—American National Standard for Video Recording—1 inch Reel Dimensions.
5. SMPTE/ANSI 25M—American National Standard for Video Recording—1 inch Magnetic Recording Tape.

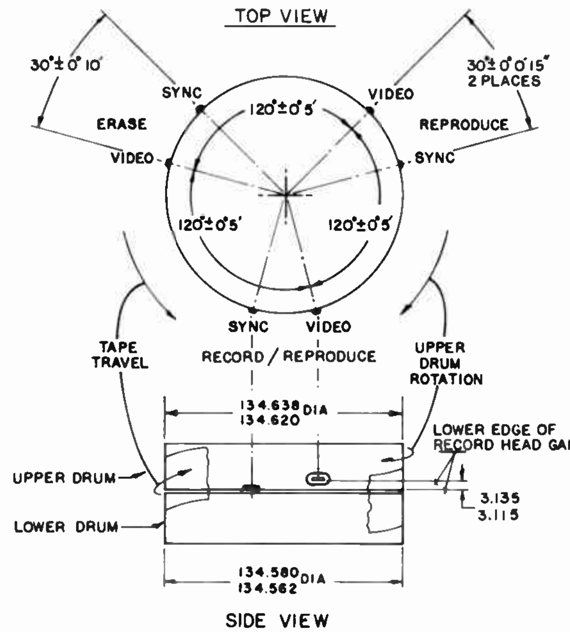
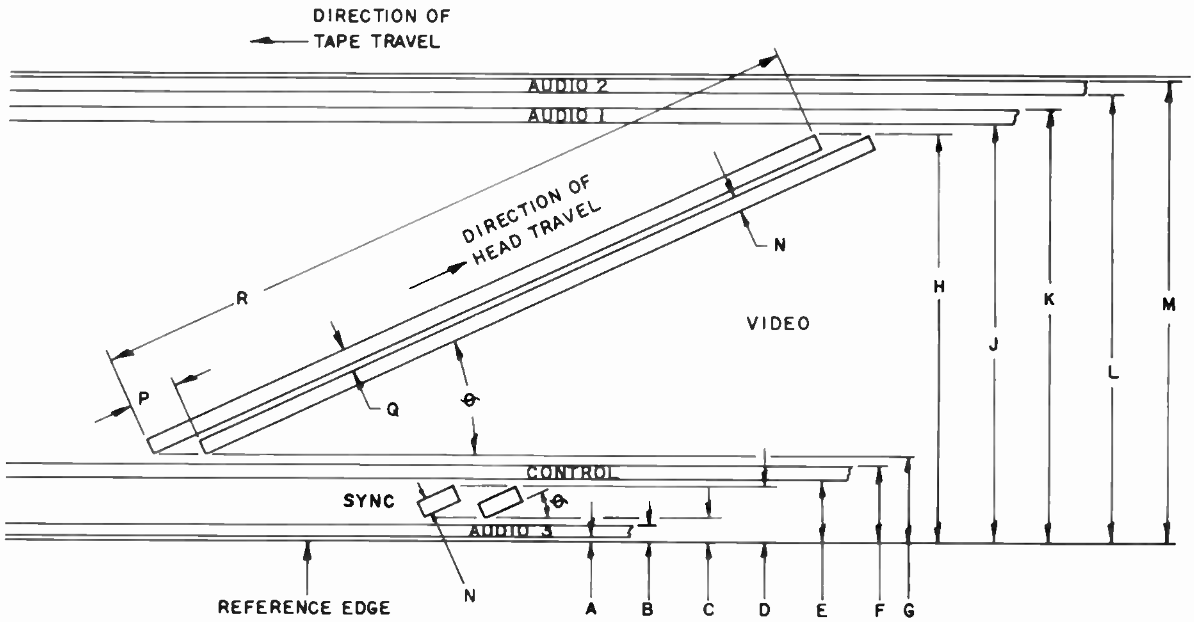


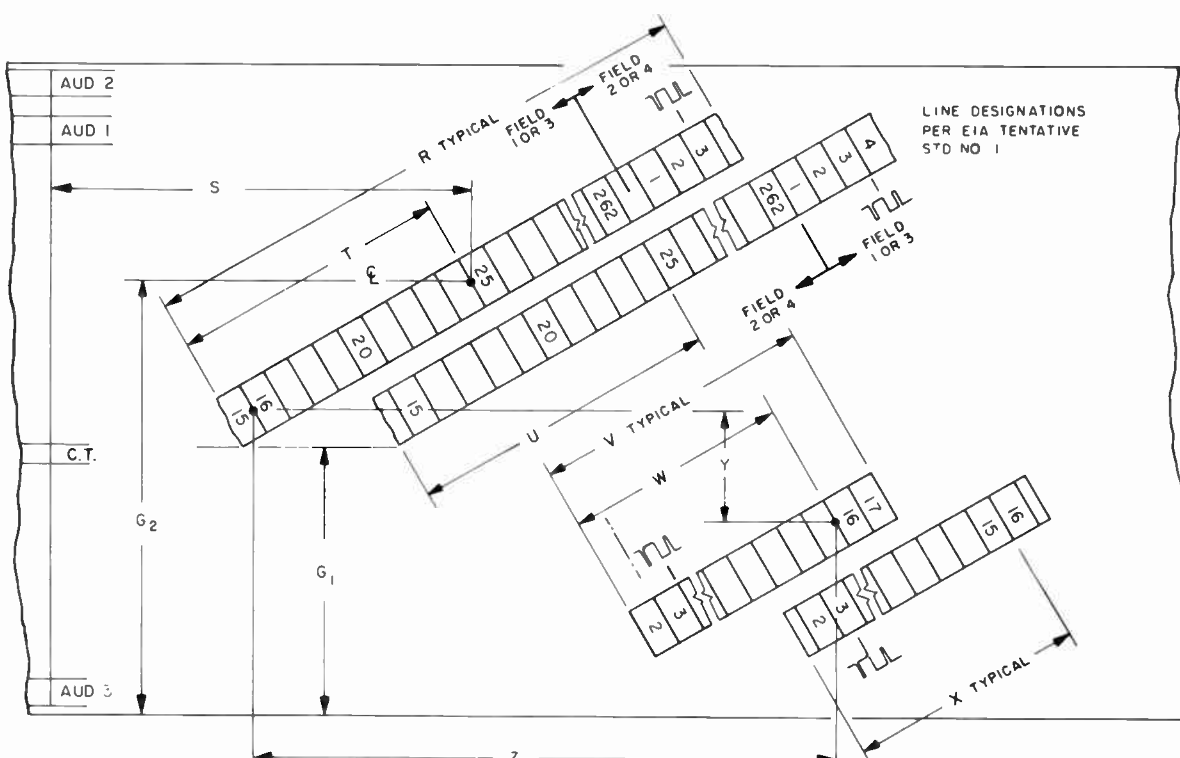
Figure 6. Pole tip locations and drum dimensions—C format system.

TABLE 1
Typical performance specifications.

Parameter	Value[16]
VIDEO	
Bandwidth	4.2 MHz
SNR	47 dB
Diff Gain	<4%
Diff Phase	<4%
Y-C Delay	<20 ns.
AUDIO	
Bandwidth	50—18kHz
Response	±2 dB
SNR	56 dB (Audio 1 & 2) 54 dB (Audio 3)



Record Location and Dimensions



Video and Sync Record Location

Figure 7. Video, audio, sync and control track locations—C format system.

Record Locations and Dimensions

	Dimensions	Millimeters Minimum	Millimeters Maximum	Inches Minimum	Inches Maximum
A	Audio 3 lower edge	0.000	0.200	0.0000	0.0079
B	Audio 3 upper edge	0.775	1.025	0.0305	0.0404
C	Sync track lower edge	1.385	1.445	0.0545	0.0569
D	Sync track upper edge	2.680	2.740	0.1055	0.1079
E	Control track lower edge	2.870	3.130	0.1130	0.1232
F	Control track upper edge	3.430	3.770	0.1350	0.1484
G_1	Video track lower edge	3.860	3.920	0.1520	0.1543
G_2	Video line 25 start	4.650	4.710	0.1831	0.1854
H	Video track upper edge	22.355	22.475	0.8801	0.8848
J	Audio 1 lower edge	22.700	22.900	0.8937	0.9016
K	Audio 1 upper edge	23.475	23.725	0.9242	0.9341
L	Audio 2 lower edge	24.275	24.525	0.9557	0.9656
M	Audio 2 upper edge	25.100	25.300	0.9882	0.9961
N	Video and sync track width	0.125	0.135	0.0049	0.0053
P	Video offset	4.067 (2.5H) ref		0.1601 ref	
Q	Video track pitch	0.1823 ref		0.00718 ref	
R	Video track length	410.764 (252.5H) ref		16.1718 ref	
S	Control track head distance	116.23	117.03	4.567	4.607
T	Vertical phase odd field	16.270 ref (10.0H)		0.6406 ref	
U	Vertical phase even field	17.080 ref (10.5H)		0.6724 ref	
V	Sync track length	25.620 (15.75H)	26.420 (16.25H)	1.0087	1.0402
W	Vertical phase odd sync field	22.360 (13.75H)	23.170 (14.25H)	0.8803	0.9122
X	Vertical phase even sync field	23.170 (14.25H)	23.980 (14.75H)	0.9122	0.9441
Y	Vertical head offset	1.529 ref		0.0602 ref	
Z	Horizontal head offset	35.350 ref		1.3917 ref	
θ	Track angle	2°34′ ref			

Figure 7. Video, audio, sync and control track locations—C format systems (Cont.)

6. SMPTE/ANSI 26M—American National Standard for Video Recording—1 inch Helical Scan Recorders—Raw Stock for Reference Tapes.
7. SMPTE/ANSI V98.27M—American National Standard for Video Recording—1 inch Type C Reference Recorders—Basic System and Transport Geometry Parameters.
8. SMPTE/ANSI V98.28M—American National Standard for Video Recording—1 inch Type C Reference Tapes—Records.
9. SMPTE RP85—Recommended Practice: Tracking Control Record for 1 inch Type C Helical Scan Videotape Recording.
10. SMPTE RP86—Recommended Practice: Video Record Parameters for 1 inch Type C Helical Scan Videotape Recording.
11. SMPTE RP93—Recommended Practice: Requirements for Recording American National Standards Time and Control Code on 1 inch Types B and C Helical Scan Videotape Recorders.
12. SMPTE RP121—Recommended Practice: Tape Dropout Specifications for 1 inch Types B and C Videotape Recorder/Reproducer.
13. SMPTE RP142—Recommended Practice: Stereo Audio Track Allocations and Identification of Noise Reduction for Videotape Recording.
14. SMPTE RP148—Recommended Practice: Relative Polarity of Stereo Audio Signals.

¾-INCH FORMAT ANALOG SYSTEMS

Introduction

The U format analog videotape system was developed in the late 1960s as a joint effort of the Matsushita Electric Industrial Co. (U-Vision), Victor Company of Japan (U-Tape), and Sony Corporation (U-Matic). The system employs a ¾-inch cassette and records the video information using an analog component color-under recording format.

The U format system bandwidth characteristics are less than the full bandwidth of the NTSC system, but the U format ease of use and equipment size and weight characteristics made it a desirable system for most electronic newsgathering (ENG) applications and many electronic field production (EFP) applications.

As technology has improved, the U format continues to be improved to its limits. There are, however, definite constraints to those limits. The color-under technique allows slower tape and drum speeds by mixing the color information to a lower frequency and bandlimiting the luminance information. The restricted luminance bandwidth limits detail and can cause ringing on abrupt transitions. Because the chroma is recorded as an AM signal, it is susceptible to off-tape amplitude variations (unlike FM). Because the color signal is separately filtered and processed, the chrominance-to-luminance delay can be difficult to maintain and the resulting bandlimiting can cause ringing on large color transitions.

Principles and Characteristics

A helical scan architecture is used in the ¾-inch U format machines, in that the tape unspools from the cassette supply reel and is presented to the drum in such a manner that the heads slant across the tape to form part of a helix.

Two sets of video heads are placed 180 degrees apart on the scanner, and the tape is wrapped around the scanner for slightly more than 180 degrees (Fig. 8). Each set of video heads consists of two heads, one for the channel record/playback and one for the channel erase. The scanner rotates at exactly 29.97 revolutions per second (approximately 1,800 rpm) recording one TV frame per revolution and one TV field per 180 degree scan.

The ¾-inch U-format provides, in addition to the two video channels, four longitudinal tracks; two for audio, one for control track, and one for address track (Fig. 9). The video tracks do not interfere with the address track and can be independently edited.

The U format system provides an overlap period during which the same information is recorded by the head preparing to leave the tape and the head which is initiating its pass. This overlap also exists in playback. Electronic switching is used to control the changeover to produce a continuous TV signal.

The switching position takes place in the vertical interval before the leading edge of vertical sync as shown in Fig. 10. The RF output extends past the switching point by approximately three horizontal lines to provide three horizontal lines of overlap.

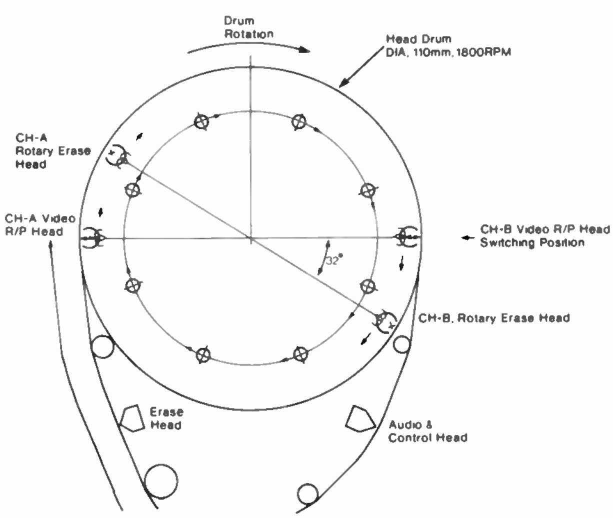


Figure 8. Pole tip locations and drum dimensions—U format system.

Color-Under Recording

In the color-under record system, the composite color video signal is first separated into its luminance (Y) and chrominance (C) components.

The luminance (Y) component is frequency-modulated for recording with the following reference levels:[17]

Peak White: 5.4 ± 0.1 MHz
Sync Tip: 3.8 ± 0.2 MHz

The chrominance (C) component is heterodyned with a fixed frequency continuous signal and down-converted to the region of 600 to 800 kHz depending upon the system used. The frequency selected for the down conversion is chosen to minimize the color-dot-crawl effects. The down-converted subcarrier (f_s) is related to the line rate (f_h) in the following manner:

$$f_s = \tfrac{1}{4}(2n + 1)\, f_h$$

For U format systems,

$$f_s = 688.374 \text{ kHz}$$
$$n = 87$$

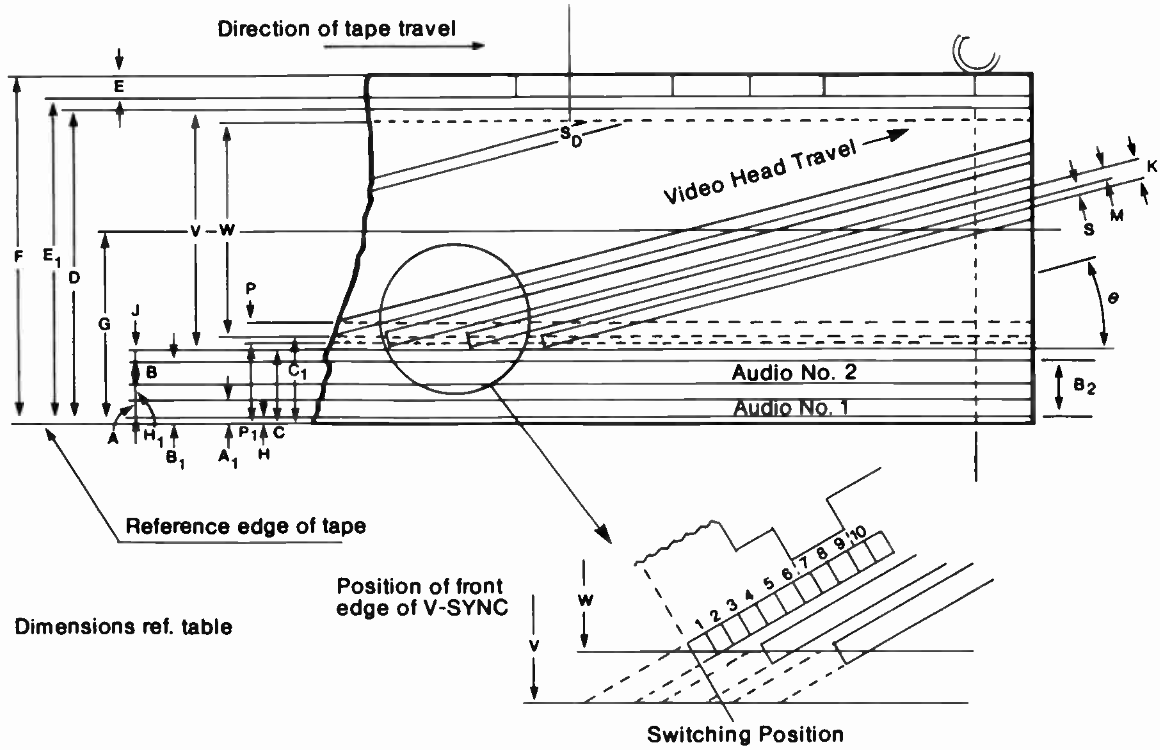
After down-conversion, the chrominance signal is superimposed on the Y-FM signal with the Y-FM carrier acting as a bias (Figs. 11 and 12). The Y-FM signal is recorded at an optimum level and the C-signal is recorded at a level designed to obtain maximum linearity. This is generally 10 to 15 dB below the level of the FM carrier.

High frequency nonlinear preemphasis is applied to the Y-channel to increase detail information prior to recording. A deemphasis noise reducing circuit is applied during playback.

During playback the chrominance signal is extracted via a low-pass filter and the Y-FM signal is demodulated. The chrominance subcarrier signal is then up-converted to 3.58 MHz by a process that compensates for the time-base error; it is then mixed with the luminance signal to produce an NTSC composite video signal.

Applicable Standards Documents

1. SMPTE/ANSI 21M—American National Standard for Video Recording—¾ inch Type E—Records.
2. SMPTE/ANSI 22M—American National Standard for Video Recording—¾ inch Type E—Cassette.
3. SMPTE/ANSI 32M—American National Standard for Television Analog Recording—¾ inch Type E—Small Video Cassette.
4. SMPTE RP87:—Recommended Practice: Reference Carrier Frequencies, Pre-emphasis Characteristic and Audio and Control Signals for ¾ inch type E Helical Scan Videotape Cassette Recording.
5. SMPTE RP148:—Recommended Practice: Relative Polarity of Audio Signals.



	Dimensions	Millimeters		Inches	
A	Audio No. 1 width	0.80	± 0.05	0.0315	±0.0020
A_1	Audio No. 1 reference	1.00	nom	0.0394	nom
B	Audio No. 2 width	0.80	± 0.05	0.0315	±0.0020
B_1	Audio No. 2 reference	2.50	nom	0.0984	nom
B_2	Audio track total width	2.30	± 0.08	0.0906	±0.0031
C	Video area lower limit	2.70	min	0.1063	min
C_1	Video effective area lower limit	3.05	min	0.1201	min
D	Video area upper limit	18.20	max	0.7165	max
E	Control track width	0.60	nom	0.0236	nom
E_1	Control track reference	18.40	+ 0.28 — 0.18	0.7244	+ 0.0110 — 0.0071
F	Tape width	19.00	± 0.03	0.7480	±0.0012
G	Video track center from reference edge	10.45	± 0.05	0.4114	±0.0020
H	Audio guard band to tape edge	0.2	± 0.1	0.008	±0.004
H_1	Audio-to-audio guard band	0.7	nom	0.028	nom
J	Audio-to-video guard band	0.2	nom	0.008	nom
K	Video track pitch (calculated)	0.137	nom	0.00539	nom
L	Audio and control head position from end of 180° scan	74.0	± 0.5	2.913	±0.020
M	Video track width	0.085	± 0.007	0.00335	±0.00028
P*	Address track width	0.50	± 0.05	0.0197	±0.0020
P_1	Address track lower limit	2.90	± 0.15	0.1142	±0.0059
S	Video guard band width	0.052	nom	0.00205	nom
V	Video width	15.5	nom	0.610	nom
W	Video effective width	14.80	± 0.01	0.5827	±0.0004
θ	Video track angle, moving tape	4° 57' 33.2"			
	stationary tape	4° 54' 49.1"			

Figure 9. Video, audio and control track locations—U format system.

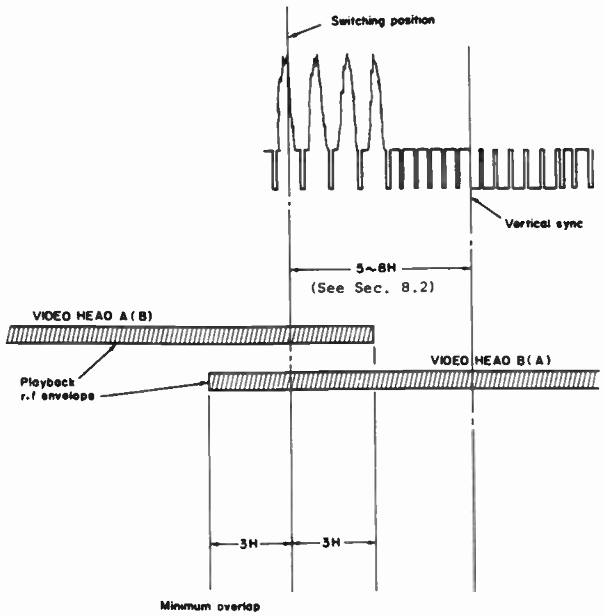


Figure 10. Video switching locations—U format system.

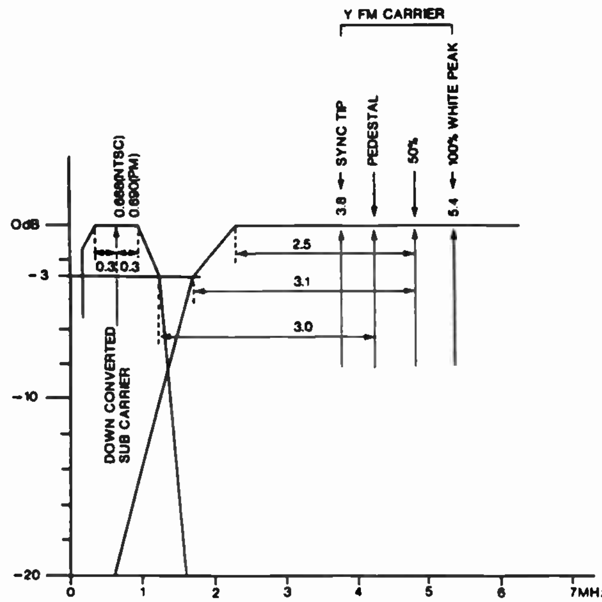


Figure 12. Frequency spectrum—U format system.

½ INCH FORMAT ANALOG SYSTEMS

Introduction

Four different ½-inch analog videotape systems were developed in the 1982–1987 period to meet various levels of performance for specific professional applications. All four systems employ various size ½-inch cassettes. The four systems, in the order of introduction, are the M format, Betacam format, M-II format, and BetacamSP format.

All four are based on a helical scan videotape recording system. All four employ component recording in that they record the luminance (Y) and baseband chrominance (C) signals for one television field on separate tracks with each pass of the recording head. Assigning separate tracks for the luminance and chrominance video permits the system to have a wider baseband recording capability, but with a sacrifice in tape usage.

Differences in the record locations and dimensions, tape speed, oxide formulation, and the method of encoding the chrominance channel, among others, produce various levels of performance among the four systems. These differences in implementation and the resulting differences in performance are described in the following sections.

The M format and Betacam format system bandwidth characteristics are less than the full bandwidth of the NTSC system, but the ease of use and equipment size, leading to a lightweight, single unit camrecorder, make them desirable systems for most electronic newsgathering applications and many electronic field production

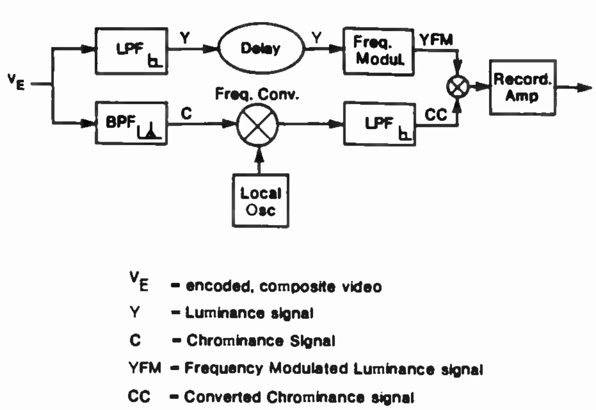


Figure 11. Color-under recording.

TABLE 2

Typical performance specifications.

VIDEO[18]

Parameter	Format	Value
Input Signal	NTSC Composite	1.0 V p-p
Horizontal	Conventional Systems	260 lines (3.2 MHz)
Resolution	SP Mode	340 lines (4.2 MHz)
SNR		>47 dB
Record/Play Time	Standard Cassette	60 min.

AUDIO

Parameter	Mode	Value
Frequency Response		50 Hz to 15 kHz
SNR	Conventional	>50 dB
	SP	>52 dB
Wow & Flutter	Conventional	<0.20% rms
	SP	<0.18% rms

applications. The second generation M-II and BetacamSP formats, based on a new metal particle tape formulation, provided full NTSC bandwidth capability in light weight, small size equipment similar to the first generation.

Component recording offers the following advantages over composite recording:

- Editing without being constrained by color framing considerations
- No color misregistration from envelope delay problems inherent in composite systems
- Lack of subcarrier moire patterns

The disadvantages are:

- The need to maintain signals in component format throughout the editing process to avoid subjecting video to continual encoding and decoding if the full benefits available from the format are to be achieved
- Chrominance/luminance delay variations
- Less efficient use of tape

Principles and Characteristics

A helical scan architecture is employed in all four ½ inch format machines, in that the tape unspools from the cassette supply reel and is presented to the drum in such a manner that the heads slant across the tape to form part of a helix.

Two sets of video heads are placed 180 degrees apart on the scanner and the tape is wrapped around the scanner for slightly more than 180 degrees (Fig. 13). Each set of video heads consists of two heads, one for the luminance (Y) channel and one for the chrominance (C) baseband channel with the two channels (Y,C) recorded simultaneously. An additional set of heads is employed for flying erase. The scanner rotates at exactly 29.97 revolutions per second (approximately 1800 rpm) recording one TV frame per revolution and one TV field per 180 degrees scan.

All four systems provide, in addition to the two video channels, four longitudinal tracks: two for audio, one for control track, and one for address track (Fig. 14). The video tracks do not interfere with the address or longitudinal address tracks and can be independently edited.

The differences in performance are attributable to, among other factors, the variations in implementation such as drum diameter, frequency spectrum, chrominance coding, writing speed and tape speed, the details of which are provided in the following tables and Fig. 15.

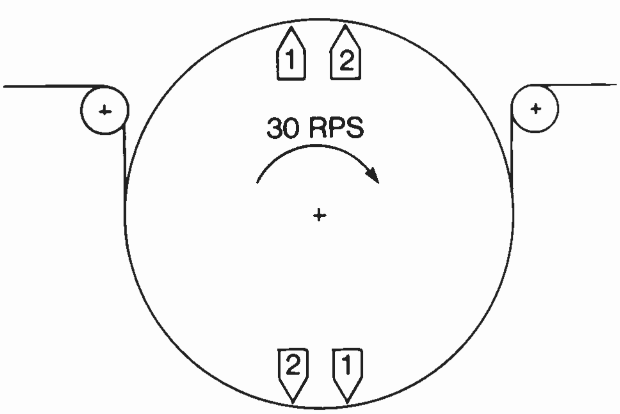


Figure 13. Helical scanner—½ inch tape formats.

Luminance and Chrominance Processing

In all four systems, the luminance (Y) and chrominance (C) channels are premphasized by two methods before modulation. In all four systems the luminance signal is converted to an FM signal and is recorded on the Y track. The four systems differ in the method of encoding chrominance and the FM deviation used as shown in the table above and in Fig. 15.

The M system carries the color information encoded in the I and Q format defined for the NTSC system. These are converted to a 5.0 to 6.0 MHz deviation FM signal for I and a 0.75 to 1.25 MHz deviation FM signal for Q. The FM signals are then summed (frequency multiplexed) and recorded on the color track as shown in Fig. 16. The Q signal is linearly recorded with the I level increased to provide the high frequency bias needed for linear recording.

The Betacam, BetacamSP and M-II systems carry the color information encoded as R-Y and B-Y. These are time compressed 2:1 so that each of the R-Y and B-Y components is contained in a period equivalent to one-half the active horizontal picture interval. An additional delay of approximately one-half H (horizontal scan frequency) is applied to the B-Y component to place the compressed B-Y information in the last half of the line. The information is then FM modulated as described in Figs. 16, 17 and 18.

During playback the decompressed R-Y and B-Y information is demultiplexed, and any system processing delay differences between the three channels (Y, R-Y, B-Y) are corrected.

Applicable Standards Documents

Note: Type L = Betacam or BetacamSP
Type M = M-2 (or M-II)

1. SMPTE/ANSI V98.35M—American National Standard for Video Recording—½ inch Type G—Cassette And Tape.
2. SMPTE 229M—Standard for Video Recording—½ inch Type L—Records.
3. SMPTE 230M—Standard for Video Recording—½ inch Type L—Electrical Parameters—Control Code, and Tracking Control.
4. SMPTE 238M—Standard for Television Analog Recording – ½ inch Type L—Tape and Cassettes.
5. SMPTE 249M—Standard for Television Analog Recording – ½ inch Type M-2—Records.

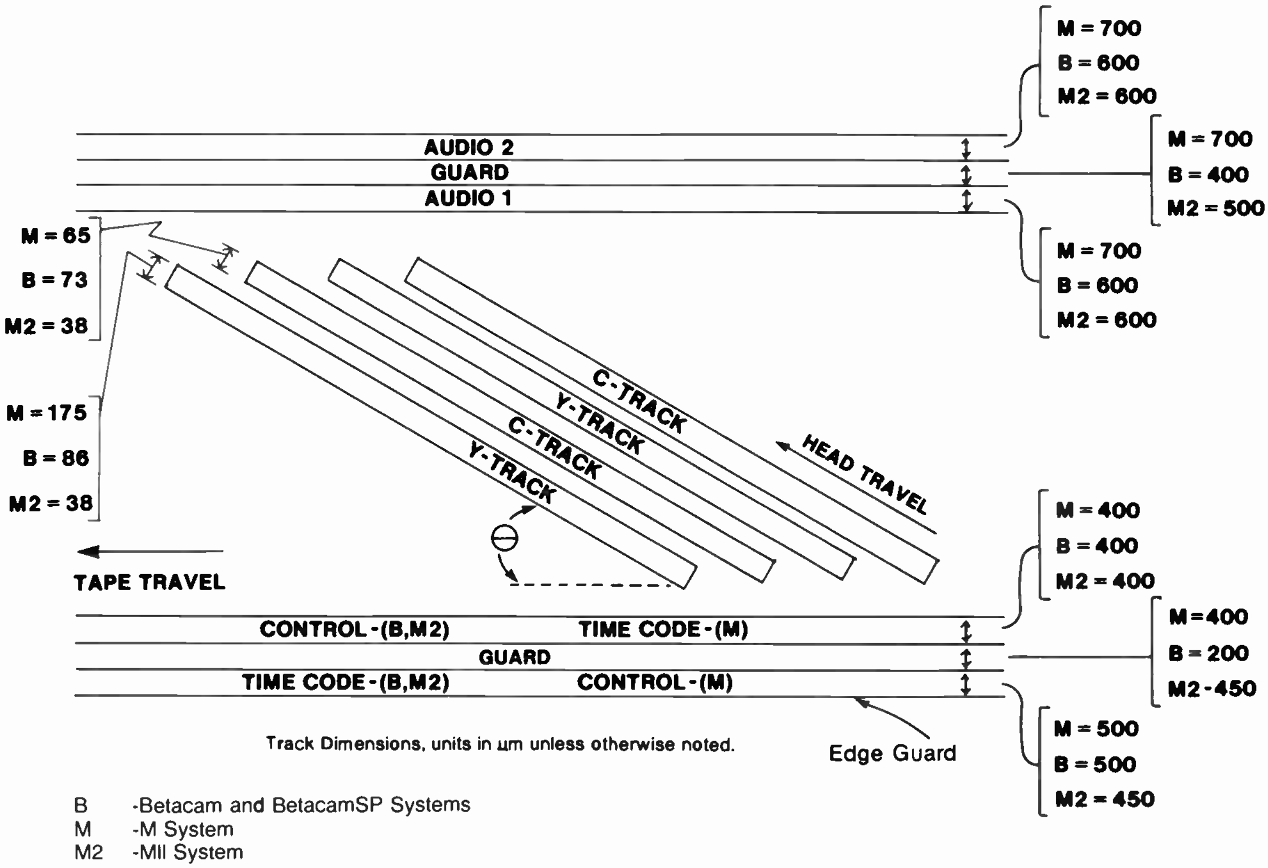


Figure 14. Video, audio and control track locations—½ inch component format systems.

6. SMPTE 250M—Standard for Television Analog Recording – ½ inch Type M-2—Videotapes and Cassettes.
7. SMPTE 251M—Standard for Television Analog Recording – ½ inch Type M-2—Electrical Parameters of Video, Audio, Time and Control Code and Tracking Control.
8. SMPTE 252M—Standard for Television Analog Recording – ½ inch Type M-2—Pulse Code Modulation Audio.
9. SMPTE RP142—Recommended Practice: Stereo Audio Track Allocations and Identification of Noise Reduction for Videotape Recording.
10. SMPTE RP144—Recommended Practice: Basic System and Transport Geometry Parameters for ½ inch Type L Cassette.
11. SMPTE RP148—Recommended Practice: Relative Polarity of Stereo Signals.
12. SMPTE RP155—Recommended Practice: Audio Levels and Indicators for Digital Audio Records on Digital Television Tape Recorders.
13. SMPTE RP158—Recommended Practice: Basic System and Transport Geometry for ½ inch Type M-2 Format.

½ INCH FORMAT ANALOG COMPOSITE S-VHS SYSTEM

Introduction

The S-VHS system is a helical-scan videotape recording system based on the standard consumer VHS cassette. It was introduced for professional applications in 1987. The system uses a cobalt doped oxide tape and a higher frequency carrier with wider deviation than the standard VHS systems for frequency modulation on the tape to achieve 400 TV line per picture height horizontal resolution.

The ease of use, equipment size, (a lightweight, single-unit camcorder), and low cost have won the S-VHS format system an advocacy in some areas for electronic news gathering applications.

Principles and Characteristics

The same helical scan architecture is employed as is used in consumer VHS systems in that the tape unspools from the cassette supply reel and is presented to the drum in such a manner that the heads slant across the tape to form part of a helix.

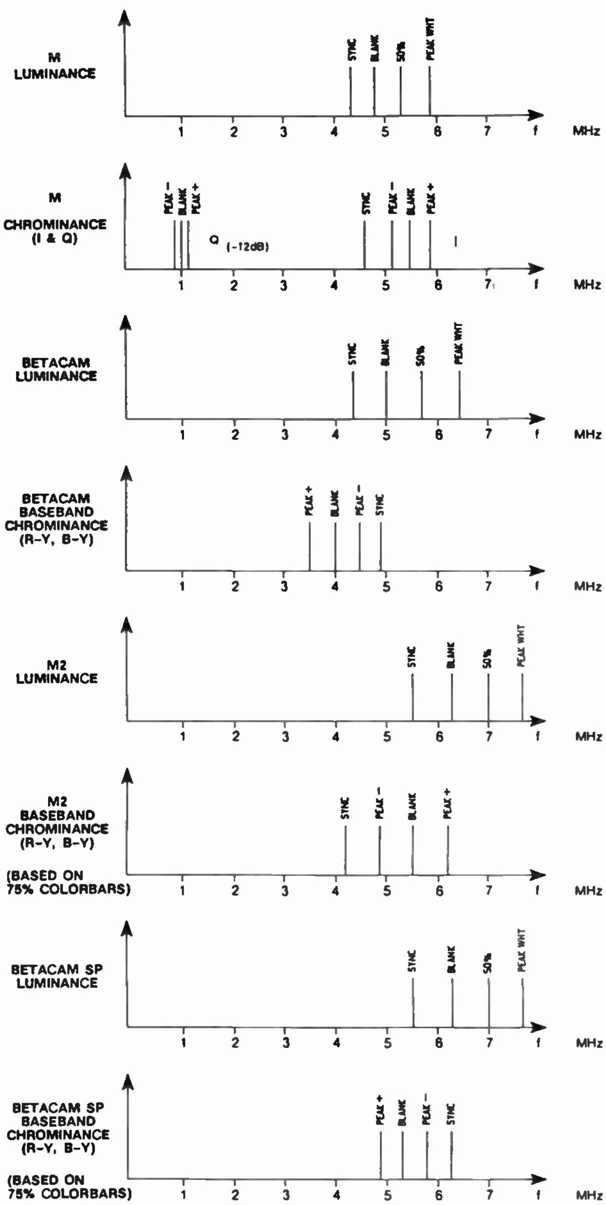


Figure 15. Frequency spectrum—½ inch component format systems.

The S-VHS system provides, in addition to the video channels, three longitudinal tracks, two for audio and one for the control track (Fig. 19). The video tracks do not interfere with the longitudinal analog tracks and can be independently edited.

The S-VHS system records luminance (Y) and chrominance (C) separately. The Y signal is recorded using FM carrier frequencies corresponding to the following reference levels:[28]

Peak White: 7.0 ± 0.1 MHz
Sync: 5.4 ± 0.1 MHz

The chrominance signal is down converted and direct-recorded using a carrier frequency of 629 kHz.

Although implementations may vary according to manufacturer and model, for professional equipment, incoming composite NTSC signals are separated into their Y and C components using adaptive comb filters, and during post processing, chroma enhancement restores some of the chroma bandwidth lost in the recording process.

Audio signal recording capability includes the use of both ac-biased linear recording and FM multiplex (Hi-Fi) techniques. Since each method provides for recording of two audio channels, the total capability is four channels or two stereo pairs.[29]

Advantage can be taken of the separate luminance and chrominance recording system if interfaces and supporting subsystems such as timebase correctors (TBCs)[30] are provided with separate access to Y and C. The ability to record and recover wideband video signals and process them provides the possibility of producing a final encoded NTSC composite signal relatively free from the cross-luminance and cross-chrominance artifacts of systems which are composite in their entirety.

Applicable Standards Documents

1. SMPTE/ANSI V98.32M—American National Standard—½-inch type H cassette—records.

D-1 COMPONENT DIGITAL VTR

Introduction

The D-1 format component digital videotape system was developed in the early 1980s to provide a format for the recording of digital video data conforming to the 4:2:2 level of CCIR Recommendation 601.[32] The D-1 system is capable of multigeneration recording of more than 30 generations using the digital interface, with little signal degradation of the audio and video.

Principles and Characteristics

The D-1 system is based on a 19 mm wide cassette based helical scan architecture employing oxide tape. The D-1 system provides for three different size cassettes: the L-size (maximum of 76 minutes of operation), the M-size (maximum of 34 minutes of operation), and the S-size (maximum of 11 minutes of operation).

The D-1 format system requirements can be met with one of three different scanner configurations incorporating either four or six video heads. One configuration uses four heads placed 90 degrees apart on the scanner with the tape wrapped approximately 261 degrees around the scanner. A second configuration uses four heads in two pairs of heads: each pair placed 180 degrees apart on the scanner and the heads within a pair separated by approximately 7.5 degrees, with the tape wrapped approximately 261 degrees around the scanner. The third configuration uses six heads in three pairs of heads: each pair placed 120 degrees apart on the scanner and the heads within a

TABLE 3
Typical performance specifications.

Format Implementation Comparison

	M	Betacam	M-II	BetacamSP
Drum Diameter	62.0 mm	74.49 mm	76.0 mm	74.49 mm
R.P.M. (nom.)	1800	1800	1800	1800
Writing Speed	5.63 m/s	6.89 m/s	7.09 m/s	6.89 m/s
Linear Tape Speed	204.5 mm/s	118.52 mm/s	67.69 mm/s	118.52 mm/s
Cassette Length (standard)	246 m	150 m	389 m	150 m
Play Length (standard)	20.0 min	21.1 min	95.0 min	21.1 min

Manufacturer's Specifications

Parameter	M[19]	Beta[20]	M-II[21]	BetaSP[22]	
Luminance Bandwidth (−3dB)	4.1	4.1 (-6dB)	4.5	4.5	MHz
Luminance SNR	50	48	50	51	dB
K-factor (2T)	3	3	<2	<2	%
Chroma Coding	I,Q	R-Y,B-Y	R-Y,B-Y	R-Y,B-Y	
Chroma Bamdwidth	1.3	1.5	1.5	1.5	MHz
Chroma SNR	52	50	50	53	dB
Diff Gain[23]	<3	<3	<2	<2	%
Diff Phase[24]	<3	<3	<2	<2	%
Y-C Delay	<30	<20	<20	<20	ns
Audio Bandwidth (±3dB)	40-15k	50-15k	50-15k	50-15k	Hz
Response	±2	±3	±1.5, −3	±3.0	dB
Audio SNR (3% distortion)[25]	53	50	74 (Ref. 26)	72 (Ref. 27)	dB
Audio Distortion	1.0	2.0	1.0	1.0	%
Wow & Flutter	0.15	0.15	0.10	0.10	% rms
Track Width	Y-175,C-65	Y-73,C-73 Y-86,C-73	Y-44,C-36	Y-86,C-73	um

pair separated by approximately 6 degrees, with the tape wrapped approximately 210 degrees around the scanner (Fig. 20).[33]

In the D-1 format, the input component video signals are sampled at 13.5 MHz for the luminance (Y) channel and 6.75 MHz for each of the two color difference signals (C_B and C_R). The video data are combined with pulse code modulation (PCM) encoded audio data for recording. Channel encoding employs nonreturn to zero (NRZ) modulation.

The most obvious and simplest method of recording digital data on magnetic media is to record a pulse for each data "1" and no pulse for each data "0". This method is termed "return-to-bias" or RB. NRZ is a

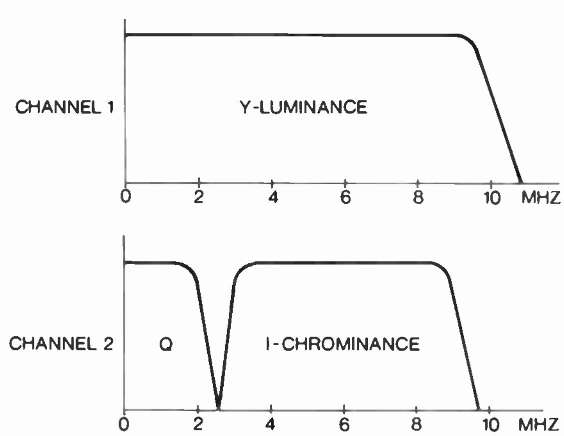


Figure 16. M System two-channel FM multiplexed spectrum allocation.

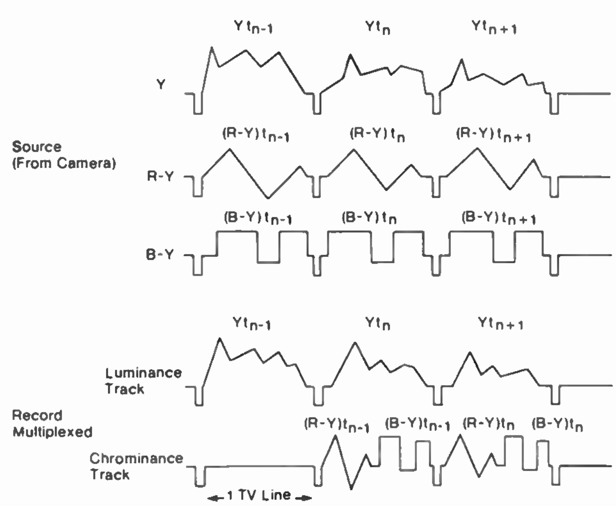


Figure 17. Betacam, BetacamSp, and M-2 system two-channel time domain multiplexed allocation.

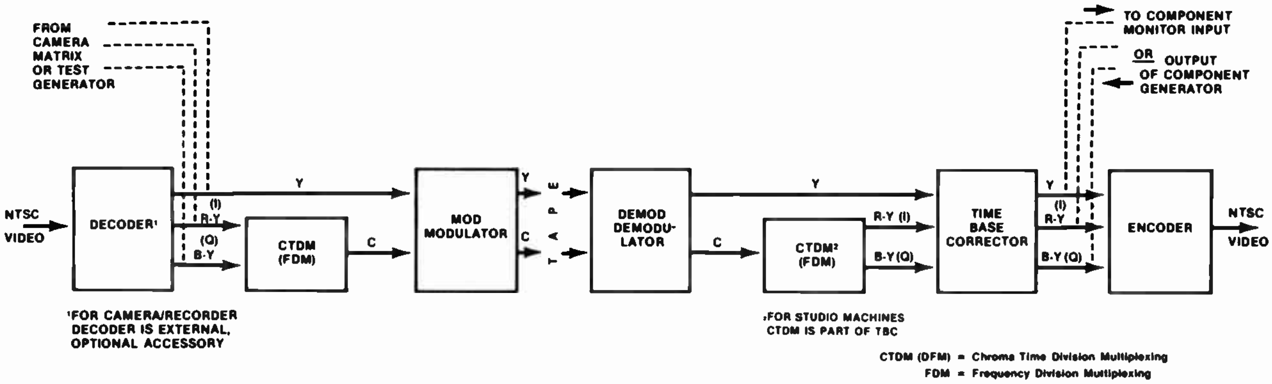


Figure 18. Component video recorder—block diagram.

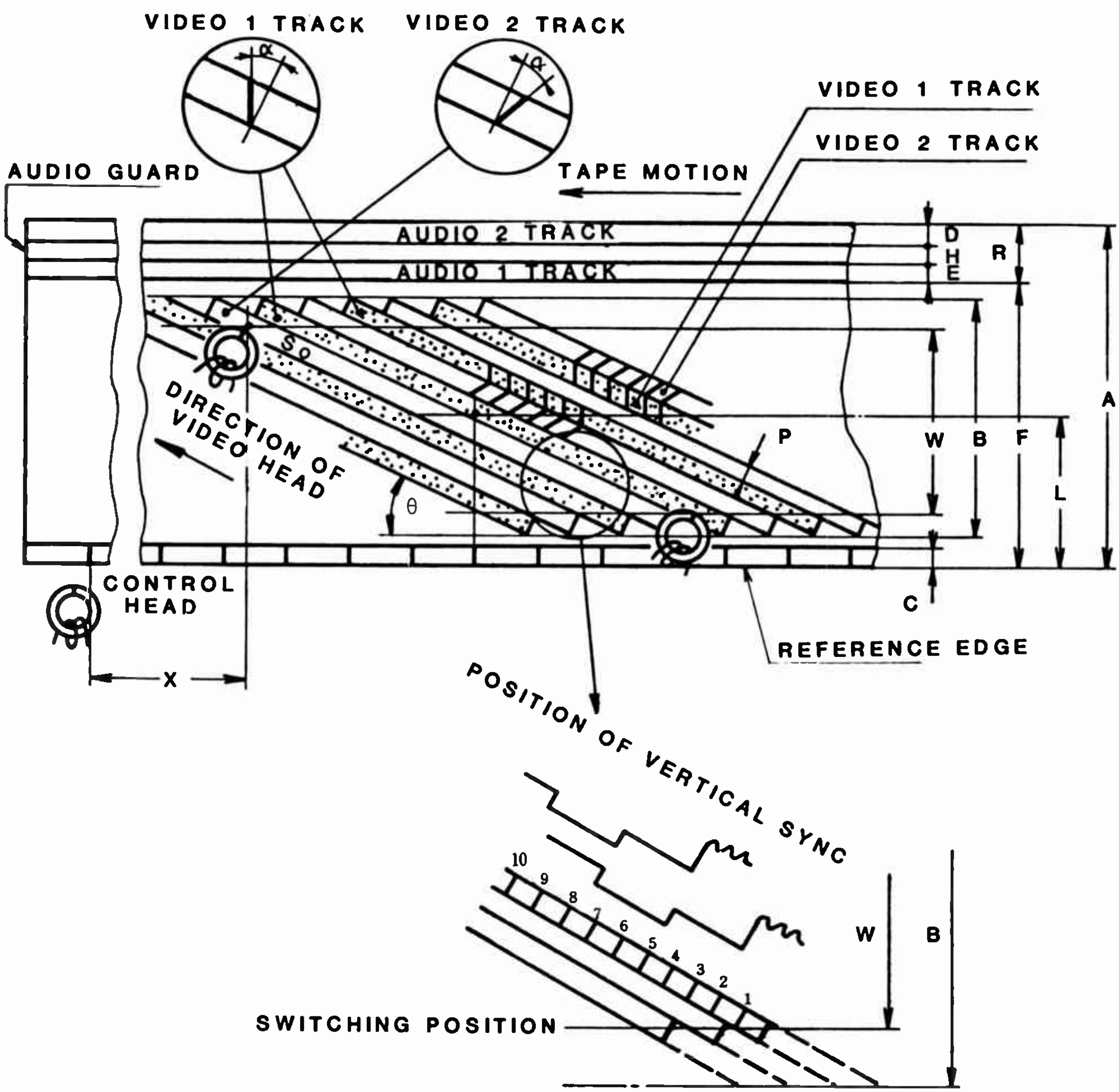


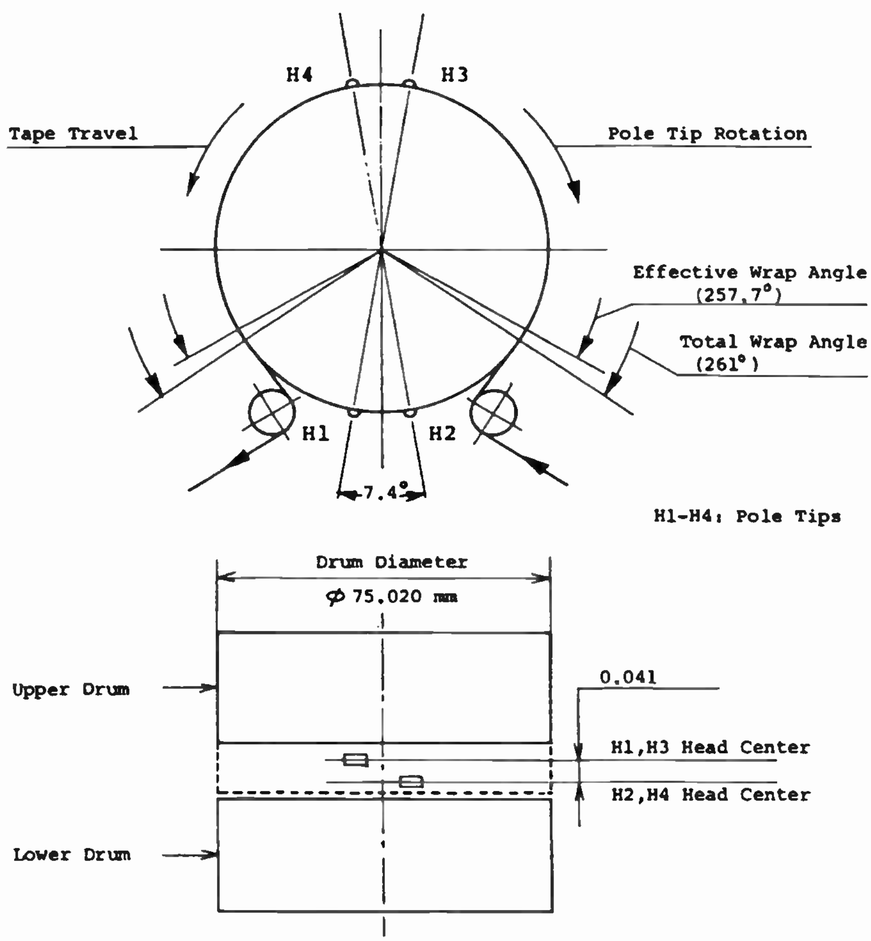
Figure 19. Video, audio and control track locations—½ inch VHS format system.

TABLE 4
Typical performance specifications.[31]
VIDEO

Parameter	Value
Horizontal Resolution	400 tv lines (app. 5 MHz)
SNR	>46 dB
Standard Recording Time	120 Mins.

AUDIO

Parameter	Mode	Value
Input Level		+4/0/−6 dB, 600 ohms
Frequency Response	Longitudinal Tracks	50 Hz—12 kHz
	Hi-Fi	20 Hz—20 kHz
Signal-to-Noise ratio	Longitudinal Tracks (Dolby)	48 dB
(Dynamic Range-Hi-Fi)	Hi-Fi	90 dB



Scanner configuration for design I

Figure 20. Helical scanner configurations—D-1 component tape format.

modification of the RB method which reduces the number of flux changes required by assigning a single transition to each bit, a flux change indicating that the current data symbol has changed from the prior signal.[34]

The digital video signals, in the form of eight-bit words, are remapped according to a specified distribution. Information received during horizontal blanking is not recorded. During vertical blanking, lines one through 13 (in field one) and 264 through 275 (in field two) are also not recorded. In the active portion of each horizontal line, 1,440 bytes of data are recorded with 720 bytes representing luminance values and 720 bytes representing 360 bytes each of the two color difference signal values. As protection against errors, the video data words are also subjected to block coding. The 1,440 samples per line times 250 lines or rows are considered as an array. The arrays are broken up into sectors of segments with 180 luminance samples and 90 pairs of chrominance samples being assigned to each sector of a segment. Two data shuffling sequences are applied. The first is an

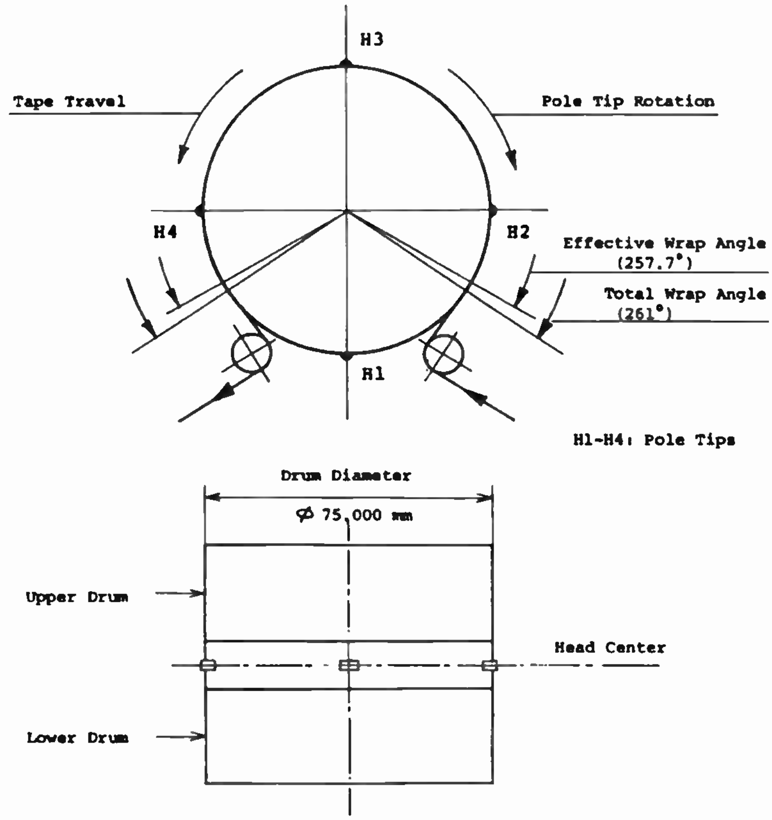
Pole tip relationships for design I, II and III

Parameters	Design I	Design II	Design III
Relevant figures	Figure 1	Figure 2	Figure 3
Minimum number of pole tips	4	4	6
Angular relationship	H1 – H2: 7.4° H3 – H4: 7.4° H1 – H3: 180° H2 – H4: 180°	H1, H2, H3, and H4 equispaced 90° ± .00833°	H1 – H2: 6.0° H3 – H4: 6.0° H5 – H6: 6.0° H1 – H3: 120.0° H3 – H5: 120.0° H5 – H1: 120.0°
Vertical displacement (mm)	H1 – H2: 0.041 H3 – H4: 0.041	± 0.002 max	H1 – H2: 0.0405 H3 – H4: 0.0405 H5 – H6: 0.0405
Maximum tip projection (μm)	45	45	60

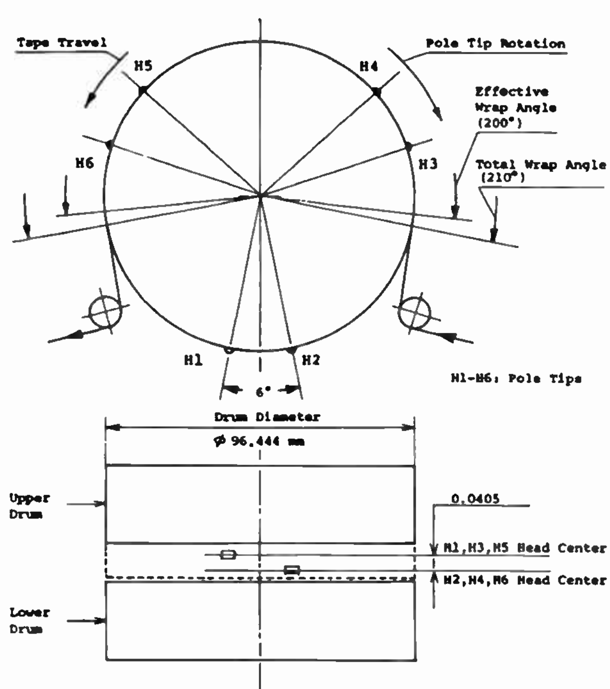
Scanner design parameters

Parameters		Design I	Design II	Design III
Scanner rotation speed (r.p.s.)		150/1.001	150/1.001	100/1.001
Number of tracks per rotation		4	4	6
Drum diameter Upper (mm) Lower (mm)		 75.020 ± 0.005 75.000 ± 0.005	 75.000 ± 0.005 75.000 ± 0.005	 96.444 96.400
Tape tension	IN	NA	0.6 ± 0.005	NA
	OUT	NA	1.0 ± 0.1	NA
Center span tension	(N)		0.8 ± 0.2	
Helix angle		5.4444° ± .0028°	5.4441° ± .0002°	5.4517°
Effective wrap angle		257.7°	257.7°	200.0°
Total wrap angle		261.0°	261.0°	210.0°
Scanner circumferential speed (m/sec)		35.3	35.3	30.3
NA = Not available at this time.				

Figure 20. Helical scanner configurations (cont.)

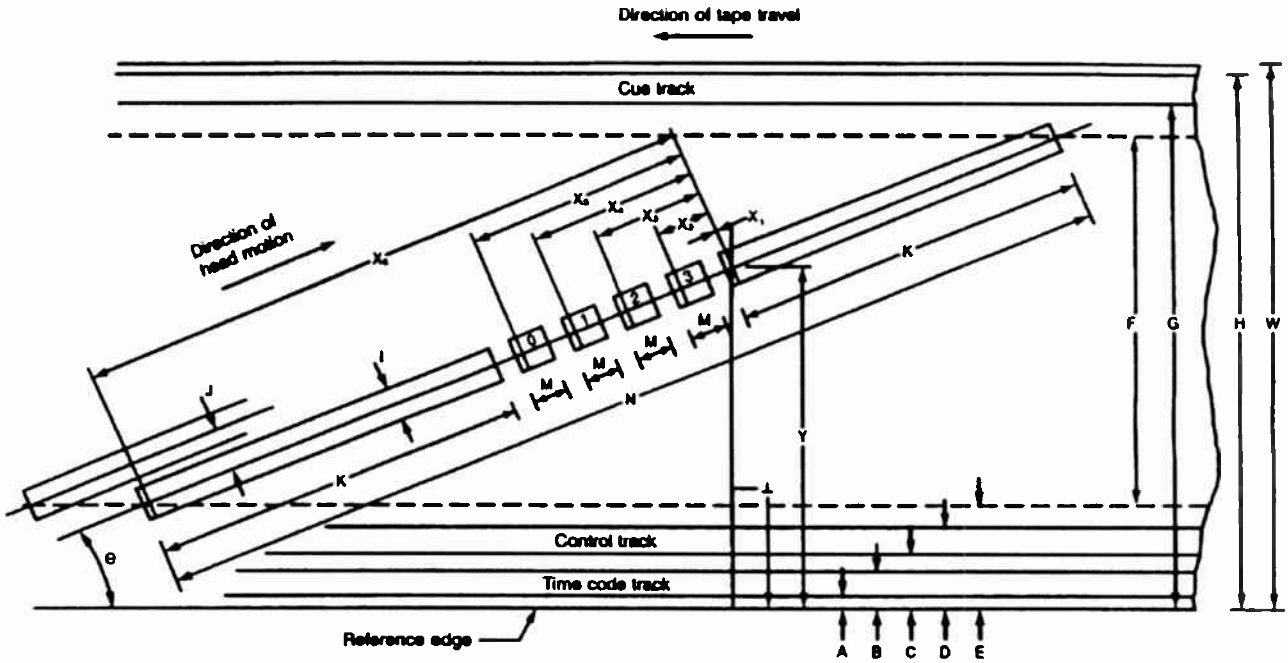


Scanner configuration for design II



Scanner configuration for design III

Figure 20. Helical scanner configurations (cont.)



NOTE – X dimensions (X_1–X_6) are determined by the end of preamble location, as defined in SMPTE 227M.

Figure 21. Video, audio, time code and control track locations—D-1 format system.

intrasector shuffle. This redistributes video and ancillary words within a single line prior to outer block error coding and also shuffles data and error correction codes within the sector prior to being written on the tape. The inner block and outer block protection codes are both forms of Reed-Solomon coding.[35]

Each helical track is recorded with data from the video channel and the four audio channels. The data are arranged so that there are six sectors per track as shown in Fig. 21. Two of the sectors are used for video data, and four sectors, each containing data from one of the four audio channels, are assigned to audio data.

The D-1 system, in addition to the video and audio channels, has three longitudinal tracks, one each for control, cue, and time code (Fig. 21).[36]

Applicable Standards Documents

1. SMPTE 224M—Standard for Television Digital Component Recording—19 mm Type D-1 Format—Tape Record.
2. SMPTE 225M—Standard for Television Digital Component Recording—19 mm Type D-1 Format—Magnetic Tape.
3. SMPTE 226M—Standard for Television Digital Component and Composite Recording—19 mm Type Format—Tape Cassette.
4. SMPTE 227M—Standard for Television Digital Component Recording—19 mm Type D-1 Format—Helical Data and Control Records.
5. SMPTE 228M—Standard for Television Digital Component Recording—19 mm Type D-1 Format—Time and Control Code and Cue Records.
6. SMPTE RP155—Recommended Practice: Audio Levels and Indicators for Digital Audio Records on Digital Television Tape Recorders.
7. SMPTE RP156—Recommended Practice: Bar Code Labeling for Type D-1 Component and Type D-2 Composite Cassette Identification.
8. SMPTE RP148—Recommended Practice: Relative

TABLE 5

Typical performance specifications.

VIDEO

Parameter	Format	Value
Input Video	NTSC 1 V p-p or CCIR Rec.601, 4:2:2	
Horizontal Bandwidth		5.75 MHz[37]
SNR		>54 dB

ANALOG AUDIO[38]

Parameter	Value
Frequency Response	50 Hz to 14 kHz
Dynamic Range	>44 dB
Distortion	<3% (200 nWb/m)
Wow & Flutter	0.1%

Polarity of Stereo Audio Signals. Identification of Noise Reduction for Videotape Recording.

9. SMPTE EG10—Engineering Guideline: Tape Transport Geometry Parameters for 19 mm Type D-1 Cassette for Component Digital Recording.
10. SMPTE EG21—Engineering Guideline: Nomenclature for Television Digital Recording, 19 mm Type D-1 Component and Type D-2 Composite Format.

D-2 COMPOSITE DIGITAL VTR

Introduction

The D-2 format composite NTSC digital videotape system was developed in the late 1980s to provide a digital format replacement for the C format machines. The D-2 system is capable of multigeneration recording of more than 20 generations using the digital interface with little signal degradation of the audio and video.

Principles and Characteristics

The D-2 system is based on a 19 mm wide cassette based helical scan architecture. The D-2 system uses the same basic cassettes as the D-1 but employs metal particle tape instead of the oxide tape used in D-1. The D-2 system provides for three different size cassettes, the L-size (maximum of 208 minutes of operation), the M-size (maximum of 94 minutes of operation), and the S-size (maximum of 32 minutes of operation).

A helical scan architecture is used in the D-2 format machines, in that the tape unspools from the supply reel and is presented to the drum in such a manner that the heads slant across the tape to form part of a helix.

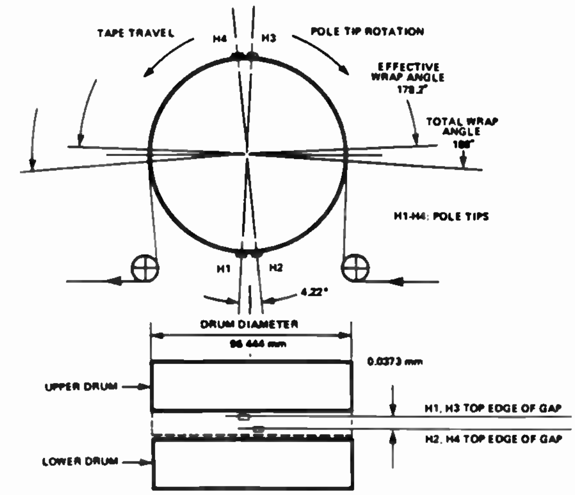
The D-2 format system incorporates eight video heads paired into four sets of video heads placed 90 degrees apart on the scanner, and the tape is wrapped around the scanner for slightly less than 180 degrees (Fig. 22). Two sets of video heads are used for recording and are placed 180 degrees apart, and the other two sets are used for playback and are also placed 180 degrees apart. Each pair of heads serves as either record or playback for two adjacent channels. The drum rotates at approximately 5,400 revolutions per minute.

In the D-2 format, the input video signal is sampled at four times the color subcarrier ($4f_{sc}$) or 14.318 MHz with the sampling phase on the I,Q axes. The video samples are quantized at eight bits and the channel coding is Miller-squared. In Miller-squared coding, the data stream is divided into three types of sequences:

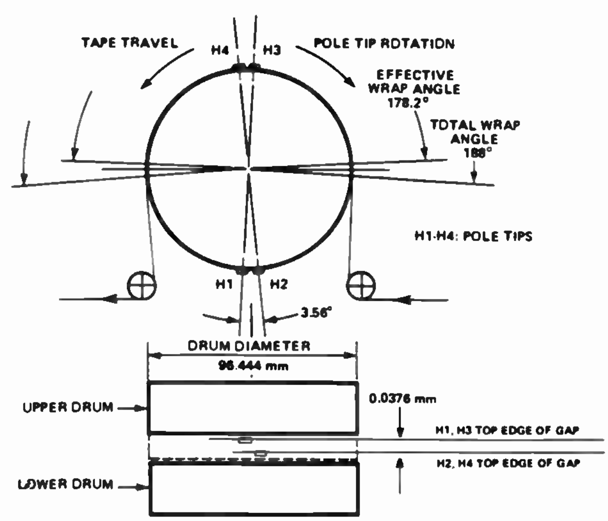
1. Any number of consecutive data "1's"
2. Two data "0's" separated by either no "1's" or any odd number of "1's"
3. One "0" followed by any even number of "1's"

Sequence types a and b are encoded such that data "1's" are assigned transitions in the middle of the bit cell, isolated "0's" are ignored, and transitions are inserted at the boundary of a bit cell between adjacent data "0's". Type (3) sequences are similarly encoded except that the transition associated with the last bit of the sequence is suppressed. Miller-squared coding has an advantage in that it is self-clocking.[39]

The video data is combined with PCM encoded audio data for recording. The four channels of audio are sampled at 48 kHz and quantized at 16 bits. Each helical track is composed of one video and four audio sectors with three segments of two tracks each and six tracks per field. The D-2 system, in addition to the video and audio channels, provides three longitudinal tracks, one each for control, cue, and time code (Fig. 23).[40,41]

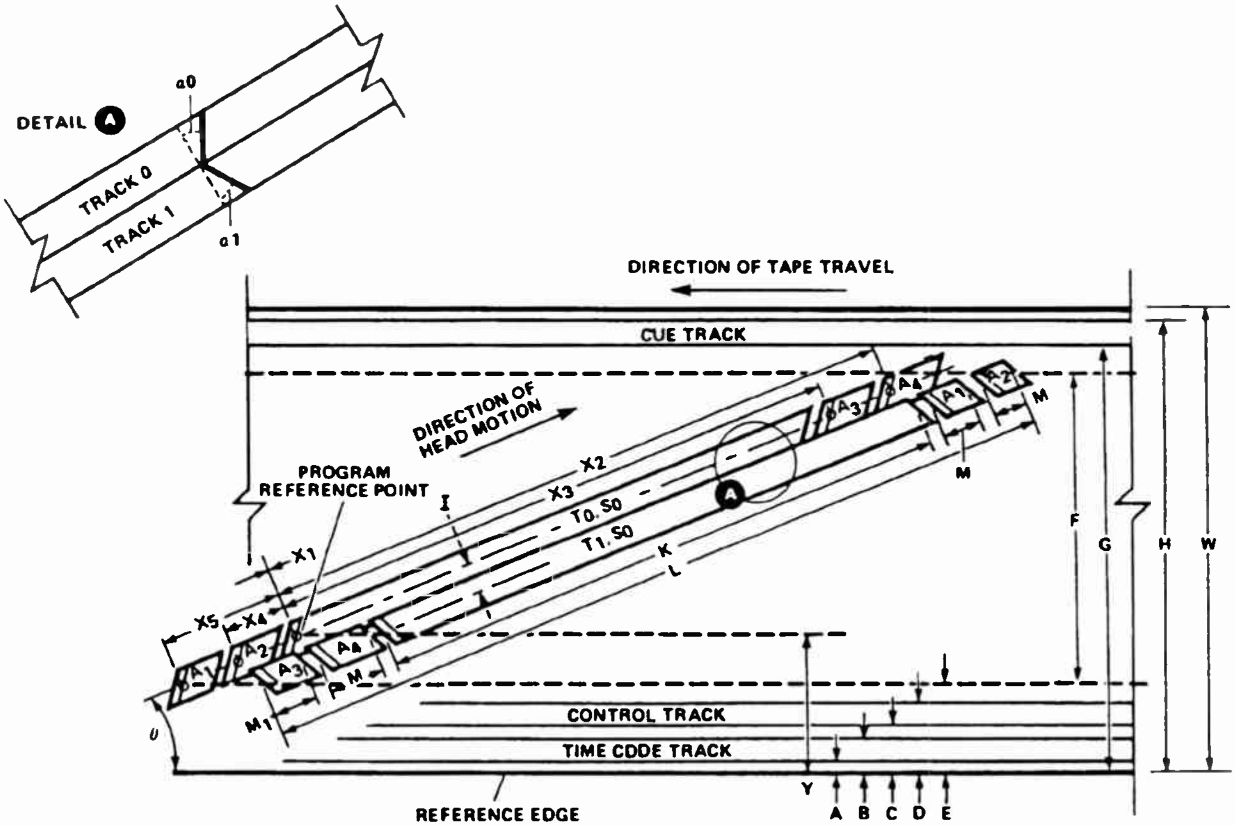


Scanner Configuration for Design 1



Scanner Configuration for Design 2

Figure 22. Helical scanner—D-2 composite tape format.



NOTES
1 A_1,. A_2, A_3, and A_4 are audio setors.
2 T_0 and T_1 are track numbers. S_0 is segment number (typical).
3 Dimensions X_1–X_5 are determined by the program reference point as defined in figure 2.

Figure 23. Video, audio, cue, time code and control track locations—D-2 format system.

TABLE 6
Typical performance specifications.
VIDEO[42]

Parameter	Format	Value
Input Signal	NTSC Composite	1.0 V p-p
Horizontal Bandwidth		6.0 MHz
Diff. Gain		≤2%
Diff. Phase		≤1°
Y/C delay		≤10 ns
K factor		≤1.0
SNR		>54 dB
Record/Play Time	Standard (M) Cassette	94 min.
Multigenerations	Digital	>20

AUDIO[43]

Parameter	Value
Frequency Response	20 Hz to 20 kHz (±0.5 dB)
Dynamic Range	>90 dB
Distortion	<0.05%
Wow & Flutter	Not Applicable

Applicable Standards Documents

1. SMPTE 226M—Standard for Television Digital Component and Composite Recording—19 mm Type Format—Tape Cassette.
2. SMPTE 245M—Standard for Television Digital Recording—19 mm Type D-2 Composite Format—Tape Record.
3. SMPTE 246M—Standard for Television Digital Recording—19 mm Type D-2 Composite Format—Magnetic Tape.
4. SMPTE 247M—Standard for Television Digital Recording—19 mm Type D-2 Composite Format—Helical Data and Control Records.
5. SMPTE 248M—Standard for Television Digital Recording—19 mm Type D-2 Composite Format—Time and Control Code and Cue Records.
6. SMPTE RP148—Recommended Practice: Relative Polarity of Stereo Audio Signals. Identification of Noise Reduction for Videotape Recording.
7. SMPTE RP155—Recommended Practice: Audio Levels and Indicators for Digital Audio Records on Digital Television Tape Recorders.

8. SMPTE RP156—Recommended Practice: Bar Code Labeling for Type D-1 Component and Type D-2 Composite Cassette Identification.
9. SMPTE EG20—Engineering Guideline: Tape Transport Geometry Parameters for 19 mm Type D-2 Composite Format for Television Digital Recording.
10. SMPTE EG21—Engineering Guideline: Nomenclature for Television Digital Recording, 19 mm Type D-1 Component and Type D-2 Composite Format.
11. SMPTE EG22—Engineering Guideline: Description and Index of Documents for 19 mm Type D-2 Composite Television Digital Recording.

½-INCH DIGITAL VTR (D-3)

Introduction

The introduction of the ½-inch component analog (M, Betacam, M-II, and BetacamSP) videotape formats proved that a small size videotape format with professional video and audio performance levels was achievable. The small size and light weight inherent in this system therefore led to the potential for a universal format used from studio to ENG/EFP. A possible next step would be to combine the size and weight advantages of the ½-inch format with the benefits and advantages of digital VTR technology. A composite digital VTR using ½-inch tape was developed by NHK in 1989 and introduced commercially by Panasonic in 1990 as the Dx system. SMPTE has begun the standardization effort, and the system carries the SMPTE D-3 designation.

Principles and Characteristics

The D-3 digital VTR system is based on a ½-inch wide cassette based helical-scan architecture employing metal particle tape. The system accommodates both a small (S) size (maximum of 125 minutes of operation) cassette and a large (L) size (maximum of 4 hours of operation) cassette. The small cassette is designed for use in camcorders.

A helical-scan architecture is used in the ½-inch digital format machines, in that the tape unspools from the supply reel and is presented to the drum in such a manner that the heads slant across the tape to form part of a helix. The drum rotates at approximately 5,400 revolutions per minute.

As in the D-2 format, the input video signal in the D-3, ½-inch digital format is sampled at four times the color subcarrier ($4f_{sc}$) or 14.318 MHz with the sampling phase on the I,Q axes. The video samples are quantized as eight bits with the channel coding employing an 8–14 modulation system. In the 8–14 system, the eight bit video samples are remapped into 14-bit words in such a manner that bit values (logic "1" or "0") are held for a minimum of two clock time intervals, and for a maximum of seven clock time intervals. This system limits both the low frequency and high frequency content of the data stream. The video data are combined with PCM encoded audio data for recording. The four channels of audio are sampled at 48 kHz and quantized at 20 bits. Each helical track is composed of one video and four audio sectors with three segments of two tracks each and six tracks per field. The ½-inch digital system, in addition to the video and audio channels, provides three longitudinal tracks, one each for control, cue, and time code (Fig. 24).[44]

The number of codes used in the particular implementation of the 8–14 coding modulation is only 4.7% of the possible combination of 14 bits. Consideration was given in selecting the particular codes to error detection as well as bandwidth limiting. Reed-Solomon error coding and one-field shuffling are also applied to improve the robustness to errors.

MAINTENANCE

Introduction

The key to maintaining any piece of equipment is a systematic approach for both attacking an unknown problem and preventing recurring problems. A systematic preventative approach maintains performance and reduces out-of-service time.

One can break down the maintenance of any tape format machine into two areas. One area is the routine procedures for prevention and upkeep. The other concerns itself with troubleshooting of failures and program analysis.

The items discussed in this section pertain for the most part to all equipment. The approaches to maintenance of analog and digital based tape machines differ in some aspects, and where applicable, those differences are discussed.

Operational Considerations

No system can perform well if not properly cared for. In helical scan systems the tape handling mechanisms must be properly aligned to provide maximum performance. The 1-inch type-C format system is a reel-to-reel system in which the tape handling mechanisms and the tape are more exposed to the environment than cassette based systems, and therefore, requires additional care.

Care should be taken to be sure that the tension on the supply and take up mechanism is proper, and that entrance and exit guides are properly aligned so that the tape is properly guided around the scanner. Guiding the tape around the scanner is critical to optimize track straightness.

Scanners wear, and after a certain number of hours of use (depending on the system implementation and tape formulation used), it may be necessary to replace the scanner or at least that portion which carries the heads. For each system, the operating manual for the equipment should be followed and the transport should be kept clean and free of oxide and contamination build-up. The environment has a great effect on the wear on the system and its subsequent performance. In less than optimum environments, consideration should be given to providing controlled, clean air of proper humidity and temperature range. This is particularly true in the case of automated tape systems.

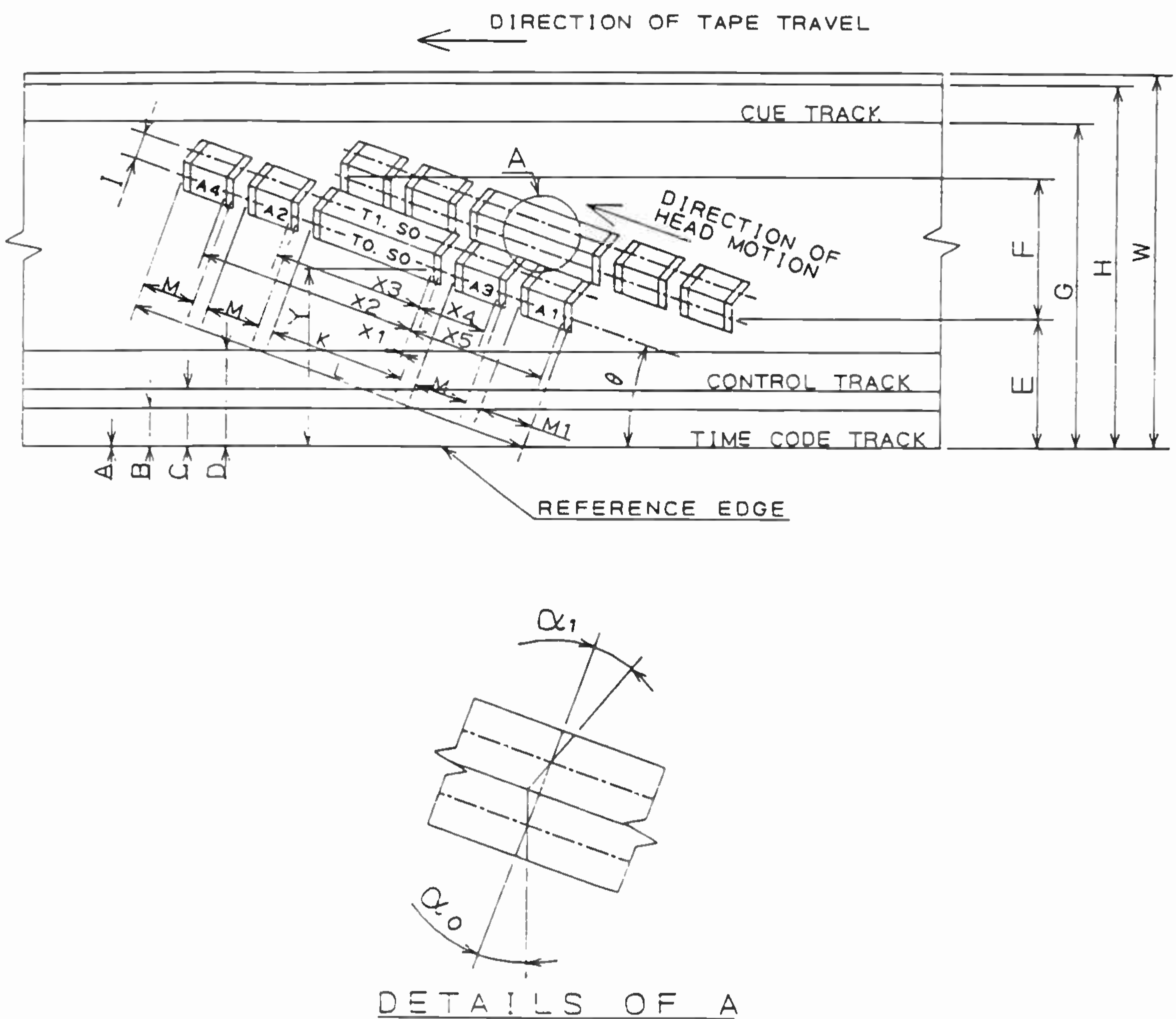
Routine Maintenance

The following should be performed on a regular basis, starting with what should be done most frequently.

1. *Operational Check:* An operational check should be performed before each extended operating session. This is merely a quick verification of all functions of the machine, from tape threading to editing.

All button lamps should be checked, where used, as they tend to burn out. Any lamps used to illuminate the cassette or tape threading area should also be checked. Tape machines which are incorporated into automated playback and record systems should be checked on a regular basis, particularly where the automation system does not provide a complete system checkout on a cyclic basis as part of its operational routine.

2. *Cleaning of Video Heads:* Cleaning of the video heads should be performed with a nonresidue producing evaporative solvent and a nonlint producing



NOTES

1 A1,A2,A3 and A4 are audio sectors.

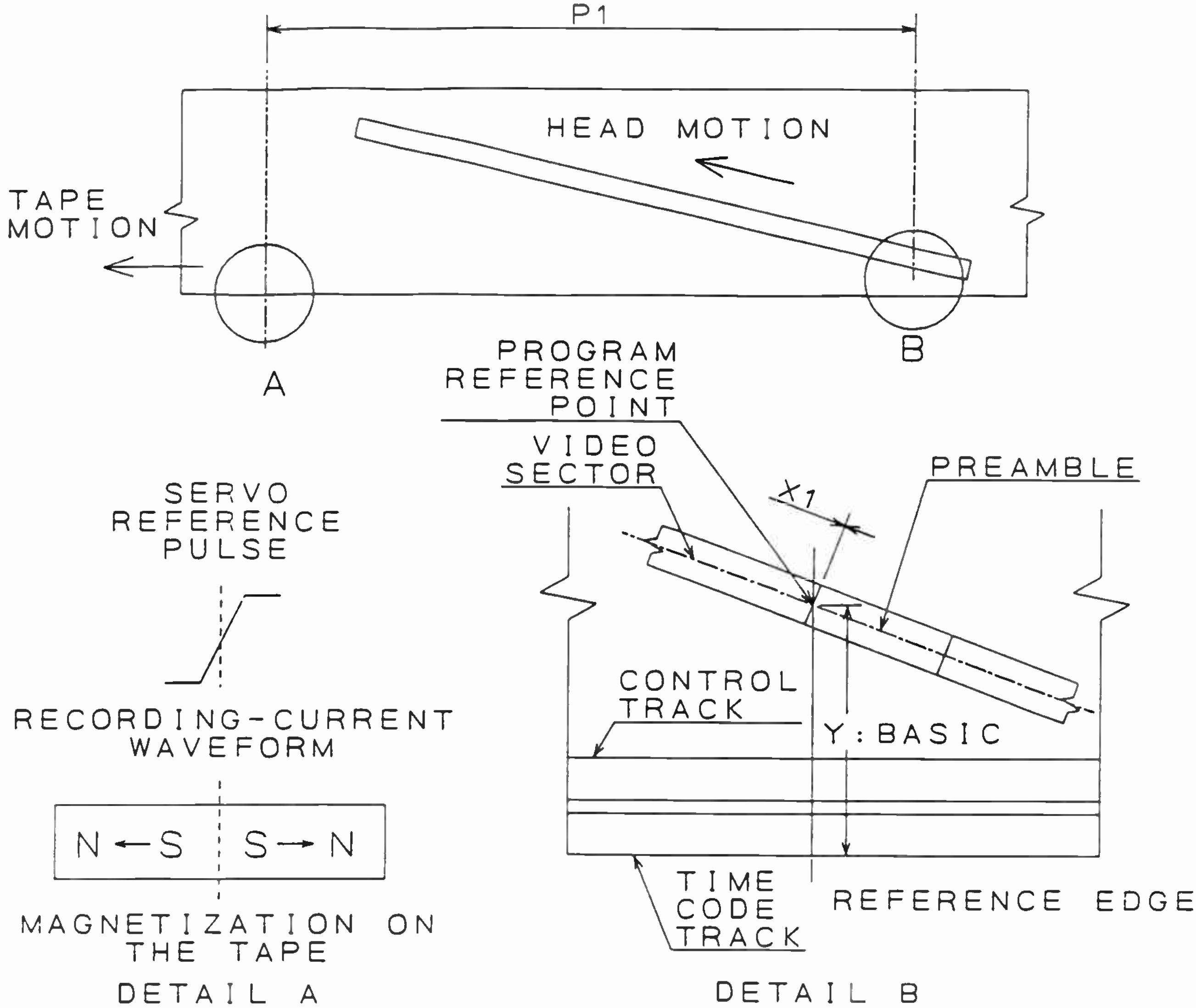
2 T0 and T1 are track numbers, S0 is segment number (typical).
Track number is identified by the difference of the azimuth angle.

3 Tape viewed from magnetic coating side.

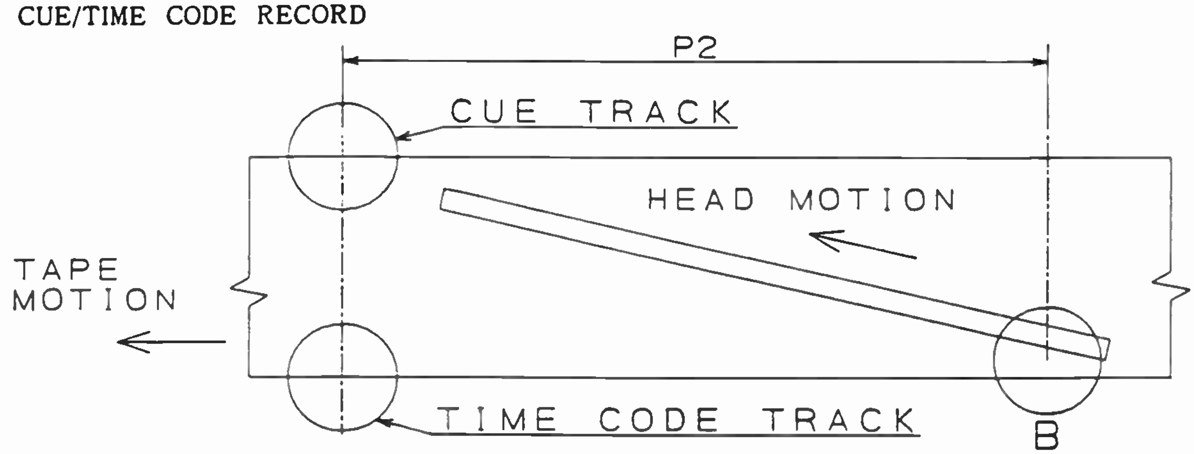
4 Dimension X1-X5 are determined by the program reference point.

Figure 24. Video, audio, cue, time code and control track locations—D-3, ½ inch digital format system.

1. CONTROL TRACK RECORD



2. CUE/TIME CODE RECORD



Location of cue, time code and control track record

Figure 24. Video, audio, cue, time code and control track locations—D-3, ½ inch digital format system (cont.).

Record location and dimensions

Dimensions in millimeters

	Dimensions	Nominal	Tolerance
A	Time code track lower edge	0	Basic
B	Time code track upper edge	0.450	±0.050
C	Control track lower edge	0.900	±0.050
D	Control track upper edge	1.300	±0.050
E	Program area lower edge	1.567	Derived
F	Program area width	10.090	"
G	Cue audio track lower edge	11.950	±0.050
H	Cue audio track upper edge	12.550	±0.050
I	Helical track pitch	0.0200	Ref
K	Video sector length	107.095	Derived
L	Helical track total length	117.667	"
M1	Audio sector 1 length*	2.266	"
M	Audio sector 2-4 length**	2.185	"
P1	Control track	176.000	±0.050
P2	Cue/time code track	177.430	±0.100
X1	Location of video sector	0	±0.050
X2	Location of start of audio sector A4***	110.155	±0.050
X3	" A2***	107.533	±0.050
X4	" A3***	2.623	±0.050
X5	" A1***	5.245	±0.050
Y	Program area reference	2.030	Basic
W	Tape width	12.650	±0.008

Dimensions in degrees

	Dimensions	Nominal	Tolerance
θ	Track angle	4.9192	Basic
α_0	Azimuth angle (track 0)	-20.019	±0.150
α_1	" (track 1)	19.981	±0.150

* : Audio sectors located at the start of helical tracks.
** : All other audio sectors.
*** : Audio channel numbers vary.

Figure 24. Video, audio, cue, time code and control track locations—D-3, ½ inch digital format system (cont.)

cloth wipe. (There are commercial products available for this application.) When the drum is cleaned, one must take care in gently cleaning heads in the direction of the head travel only, otherwise the heads can be damaged. The best technique is to turn the drum slowly while holding the cloth against it.

3. *Cleaning of the tape path:* All surfaces that are in contact with the tape should be cleaned, including the fixed heads (time code, audio/control track, erase, etc.) and all tape guides.

4. *Cleaning and inspection of pinch roller(s):* It is very important for the pinch rollers to stay clean. They easily collect oxide residue and lose their ability to properly engage the tape. The rollers should be inspected for aging or deformation and changed if necessary. Rollers can dry out and not work effectively.

5. *Head Degaussing:* Both rotary heads and fixed heads may, on occasion, require degaussing. The degausser is placed as close to the head as possible,

TABLE 7
Typical performance specifications.
VIDEO[45]

Parameter	Format	Value
Input Signal	NTSC Composite	1.0 V p-p
Horizontal Bandwidth		6.0 MHz
Diff. Gain		≤2%
Diff. Phase		≤1°
Y/C delay		≤15 ns
K factor		≤1.0
SNR		>54 dB
Record/Play Time	Standard (L) Cassette	125 min.
Multigenerations	Digital	>20

AUDIO[46]

Frequency Response	20 Hz to 20 kHz (±0.5 dB)
Dynamic Range	>100 dB
Distortion	<0.01%
Wow & Flutter	Not Applicable

turned on, and drawn away before it is turned off. This removes residual magnetization from the heads. This procedure need not be done as often as cleaning, but for many machines, particularly the one inch and U formats, it is easily done at the same time.

6. *Cleaning of head drum slip rings:* This need only be done when necessary but is included for completeness. The slip rings should be cleaned with a soft brush to remove dirt build-up. If necessary, an appropriate solvent can be used.

7. *Head replacement and optimization:* In most machines, the rotary video heads are not replaced individually; the whole upper drum assembly is replaced when any one head fails. The upper drum is easily changed, but it must be carefully aligned using a manufacturer supplied eccentricity gauge. The gauge, which must be attached to the chassis, is used to verify that the drum is centered on the flange. The drum is rotated slowly and the gauge is observed. Adjustments are made and the process repeated until the drum is centered to the manufacturer's recommended tolerance.

After replacement of the head assembly, adjustments should be made to optimize performance for the characteristics of the new head. The following adjustments should be made (reference the manufacturer's manual):

- Equalization
- Playback RF levels and Balance
- Record Currents (Y, Chroma, Erase)

These adjustments should be made several times during the lifetime of a set of heads as the characteristics change with head wear. On U format machines for example, manufacturer's recommendations are for replacement after 1,000 hours and optimization at 200, 500, and 750 hours. On most systems, adjustment at 25%, 50%, and 75% of the rated head life is recommended.

8. *Tension Detector Adjustment:* Most systems employ tension detectors which control tape tension by regulating power supplies to the take-up and supply reels. On the BVU-800 for instance, the detector is an electromechanical device consisting of an LED which shines through a slit onto a photosensitive cell. The slit is on a plate which is in contact with the tape and moves as a function of the tape tension. A varying control voltage is generated according to the slit position and is used to vary the reel motor torque (Fig. 25).

The tension detectors are replaceable and should be checked and adjusted as needed. The manufacturer recommends checking every 500 hours. Both physical alignment and calibration should be checked.

Troubleshooting

A thorough functional understanding of the machine is essential to quick problem isolation and repair. A minimum set of documentation should include a section on the theory of operation, block diagrams and schematics of all assemblies showing critical test points, assembly drawings, and parts lists, and a description of alignment and maintenance routines. Some of the areas of specific concern in videotape recorders are as follows:

1. *Threading:* The cassette loading process involves sensors, solenoids, and motors which detect and control the loading process. If a sensor becomes misaligned, the process will halt. If a solenoid or motor cycle is out of specification, a mechanical jam could

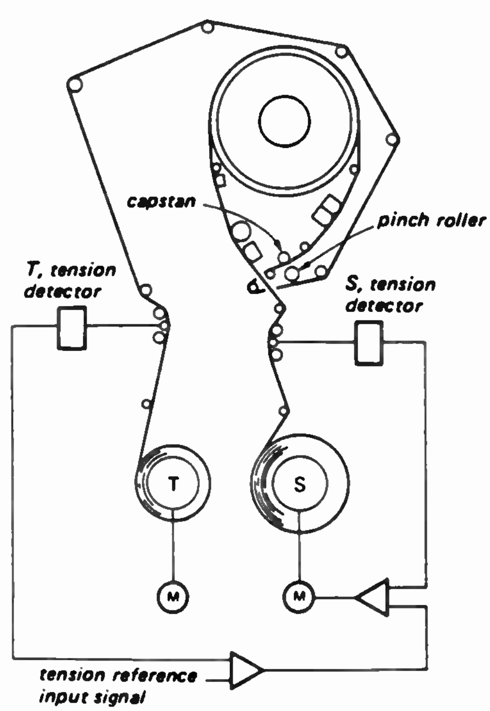


Figure 25. Videotape path & servo system interface—U format system.

occur. The appropriate manufacturer's manual should be consulted for the mechanical adjustments.

A good understanding of the control logic associated with the loading cycle will ease troubleshooting. Often the manufacturer outlines a logic timing diagram which displays when each event is supposed to occur. Also, the control logic is often designed to place the machine in a safe state (one that will not result in tape or machine damage) when it detects that certain cycles have not occurred. Understanding the possible conditions that result in the machine entering this state will be of help.

2. *Control System:* The control system interprets all input from the buttons, sensors, and other control signals and for actuating events in the sequence required. Most problems with the control system "hanging-up" are the result of a cycle that was not completed. This is unlikely to be a control system problem but more probably the failure to receive a proper response. Therefore, it is essential that all sensors operate properly.

3. *Tension and Tape Run:* In the case of cassette based machines, if the cassette successfully loads, but jams still occur, or if portions of the picture appear consistently noisier than other portions (implying a nonlinear RF envelope), the tension and tape run should be checked.

Tension: As previously mentioned, the tension detectors feed back a control voltage which in turn regulates the take-up and supply reel tension (Fig. 25). Other adjustments that affect tension are calibration of the brake motor and motor sensitivity.

Tape: The tape guides should be checked after the tension adjustments are made. The manufacturer's manual should be consulted for this. The objective is to be able to observe a uniform RF envelope while playing an alignment tape. This implies that the video heads are successfully staying within the recorded tracks on the alignment tape. If the entrance or exit guides are not properly aligned, the beginning or the end of the RF envelope will have a lower amount of detected RF. The other tape guides are used for aligning the tape properly as it passes by the stationary heads and for maintaining a uniform edge to edge tension as the tape turns and changes height. These must all be checked and adjusted to the manufacturer's specification.

Notes on Use of Alignment Tapes: For both audio and video systems alignment, procedures regarding the use of alignment tapes should be understood. The alignment tape should be stored under good conditions and handled with great care. One should be sure the machine is running with correct tension and that the tape path is clean and degaussed before the alignment tape is placed on the machine. It is best to run the tape all the way through without stopping because starting and stopping may stretch the tape and destroy its accuracy.

4. *Servo System:* There are five elements to a servo system:

- Input reference
- Motor or other device to provide movement
- Detector of output position
- Comparator to determine the difference (error) between the reference (desired) and achieved output
- Feedback path for error to be applied to allow corrective action

Adjustments to the many servo systems in a machine (such as that in Fig. 25) include calibration of references and of open loop (in which the feedback path has been disabled) responses.

5. *Audio System:* In most analog recording systems, alignment of the longitudinal track audio system is similar to an audio recorder except that there is a combined record/play head. Physical alignment should be checked and then an alignment tape played to verify playback bias. Equalization and levels should be adjusted to result in comparable playback through the set up playback electronics. The erase current also should be checked.

6. *Video System:* Alignment of the video system is similar and should be performed in the same manner as that described for audio, in that the playback electronics are aligned first according to the alignment tape, followed by record adjustments to optimize playback through the previously adjusted playback electronics. The particulars for each system vary, but, in general, alignment is a three step process.

- Playback amplifier optimization (as previously described)
- Demodulator adjustments (both luminance and chrominance)
- Modulator adjustments

The manufacturer's manual should be consulted for these adjustments.

7. *Power Supply:* The power supplies can be checked and adjusted if they are suspected of causing a problem. Parameters to be checked are:

- Purity of dc
- Regulation under varying loads
- Regulation over the specified range of input voltages

SYSTEM FEATURES

Use of the various videotape systems since the introduction of videotape recording has caused evolutionary changes in system capability. Various features and functions are available although many may be options depending upon the basic system, the tape technology utilized, system application, and manufacturer and model. In each case, the manufacturer's manual should be referenced to determine which features are available. These features include:

Control Track: The control track provides a constant flux level, alternating in polarity, and completing one cycle per frame. The polarity of the track-control record is such that the south pole of the magnetic domains point in the direction of tape travel when field 1 is recording.

Shuttle and Jog Search Functions: The search functions enable quick location of edit points. In the shuttle mode, a recognizable picture is obtained in a range on the order of from 1/30 to 10 times normal speed in both forward and reverse direction, depending upon the format and the manufacturer's implementation. Some machines offer high speed shuttle capability between five times and 32 times, depending upon the manufacturer and model. In the jog mode, the picture moves in one field increments according to the motion search control dial.

Dynamic Tracking: Playback head positioning is servo-controlled to maintain optimum tracking in still frame and other nonstandard-speed modes.

Editing Features: May include selectable preroll time, assemble, audio channel editing, preview, auto edit in/out functions, and field-by-field, or frame-by-frame, forward or reverse trim capabilities.

Dubbing: Dub mode provides the capability of directly rerecording the separated luminance and chrominance channels without the need to up-convert and encode to NTSC and then decode and down-convert.

Backspace Edit: On the ½-inch analog format camera/recorders, provides sequential recording without picture break-up at transition points.

Audio Noise Reduction: Dolby A,B, or C™ capability is included in some machines.

Additional Audio Tracks: Some systems offer additional audio channels recorded on portions of the video tracks using FM or PCM techniques and exhibiting better quality than the linear analog tracks.

Time Code Generator: Internal automatic time-code generator and reading for SMPTE/EBU Time Code.

Time Base Correctors: Normal signal reformatting, such as the color-under processing in U format systems, corrects time base error sufficiently for viewing video playback directly on color monitors or receivers. However, to meet broadcast specifications, additional time base correction is required. Time base correction is also required to provide proper signal timing in the stop motion or enhanced motion modes.

Error Correction and Compensation/Dropout Compensation: Analog systems depend upon a playback carrier dropout detector and compensator which replaces the missing RF information with a delayed signal, such as from an adjacent line to provide dropout compensation. Digital systems employ complex shuffling and encoding schemes which permit error detection and correction. The digital scheme used is particular to each system.

SMPTE TIME CODE

Evolution of Editing Techniques

Having developed a means of electronically recording video signals, the next obvious step was to develop a means of editing those recordings. The first steps emulated the film industry technique and made use of razor blades and splicing tape to cut and rejoin the physical segments of tape.

The second step made use of electronic counters and the control track pulses to define a position on a specific reel of tape from an arbitrary starting point. Obviously, what was needed was some means of defining each frame of the electronic recording in a unique and consistent manner.

The method developed was based on time code to provide for system synchronization during recording and playback. The standard adopted is documented in SMPTE 12M, "American National Standard for Television—Time and Control Code—Video and Audio Tape for 525-Line/60-Field Systems," and is referred to as SMPTE Time Code.

Description of the SMPTE Time Code

Introduction

SMPTE Time Code provides a mechanism for recording and reading out hours, minutes, seconds, and frames. The time code was originally designed as a longitudinally-recorded signal and was recorded on the Audio 3 track of C Format machines. Later, the specification was modified to allow for recording as a vertical interval signal.

The code provides an 80-bit digital word for each video frame. Within the 80-bit digital word, certain bit groups have been assigned to various time, synchronization, and user defined functions.

Longitudinal Track Application

Each television frame is identified by a unique and complete address consisting of 80 bits of information, numbered 0 through 79. The frames are numbered successively 0 through 29 with an arrangement that accommodates the NTSC color system of 29.97 Hz (drop frame), as well as, 30 Hz systems (Fig. 26).

Frame addresses start at the beginning of line 5 (±1 line) in fields 1 and 3 of the NTSC color system. The bit assignments are as follows:

Bit#	*Application*
0–3	Units of Frames
4–7	First Binary Group
8–9	Tens of Frames
10	Drop Frame Flag
11	Color Frame Flag
12–15	Second Binary Group
16–19	Units of Seconds
20–23	Third Binary Group
24–26	Tens of Seconds
27	Biphase Mark Correction Bit
28–31	Fourth Binary Group
32–35	Units of Minutes
36–39	Fifth Binary Group
40–42	Tens of Minutes

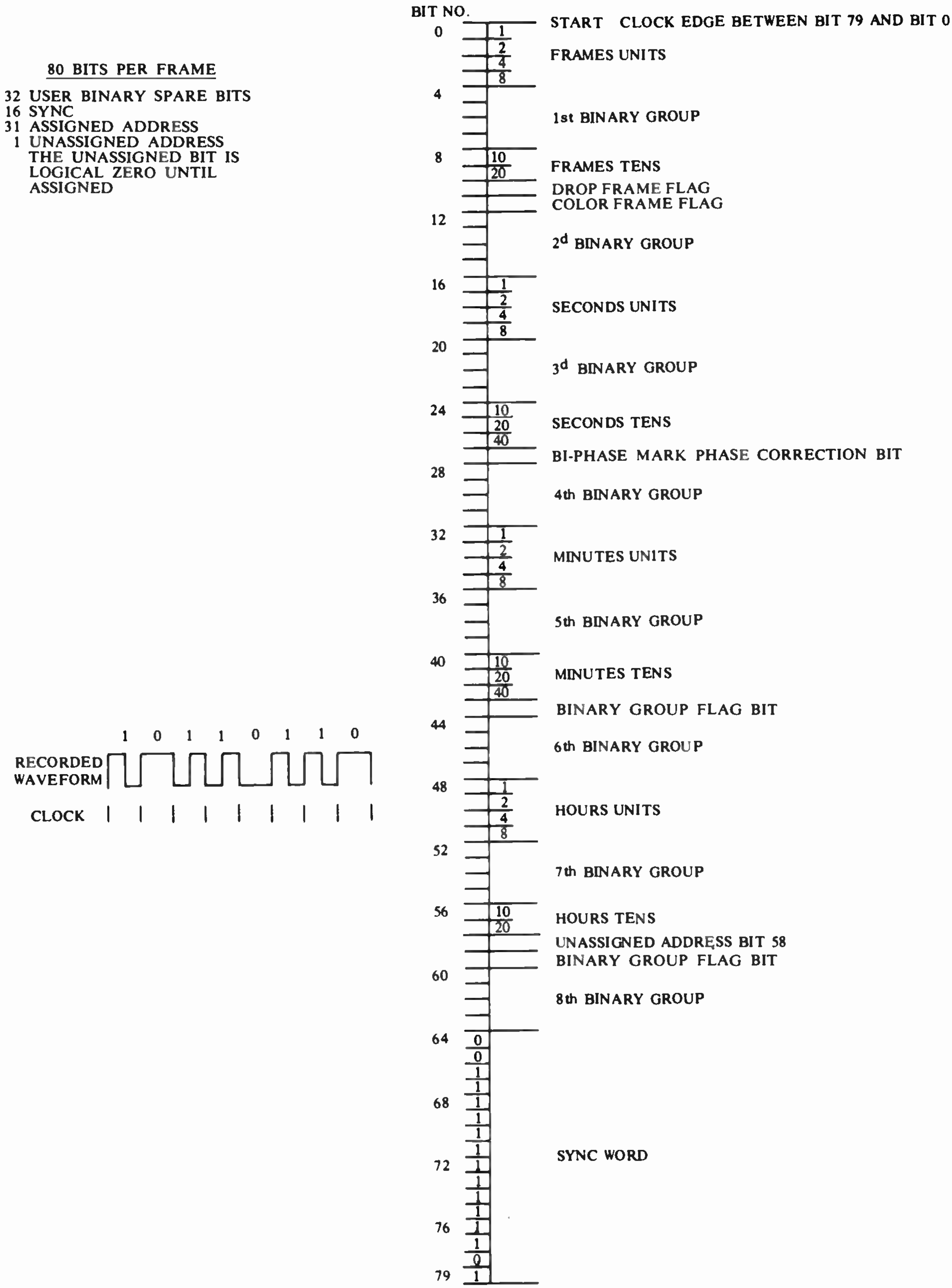


Figure 26. SMPTE time code—longitudinal bit assignment.

43	Binary Group Flag
44–47	Sixth Binary Group
48–51	Units of Hours
52–55	Seventh Binary Group
56–57	Tens of Hours
58	Binary "0" (Unassigned)
59	Binary Group Flag
60–63	Eighth Binary Group
64–79	Synchronizing Word

The eight Binary Groups are intended for storage of data by the users and the 32 bits are interpreted according to the settings of bits 43 and 59 as follows:

Bit 10: Drop Frame Flag: A "1" indicates that certain frame numbers are being dropped to resolve the difference between real time and NTSC color time.

Bit 11: Color Frame Flag: A "1" identifies NTSC color frame A.

Bit 27: Biphase Mark: This bit is set so that each 80 bit word will contain an even number of logical zeros in bits 0 through 63 (exclusive of bit 27).

TABLE 8

Assignment	Bit 43	Bit 59
Character Set Not Specified	0	0
Eight-bit Character Set	1	0
Unassigned	0	1
Unassigned	1	1

Bits 64–79: Synchronizing word in which bits 64–65 and 78 are fixed at "0", and bits 66–77 and 79 are fixed at "1".

Longitudinal Track Modulation

The modulation method used in longitudinal track recording is self clocking in that a transition occurs at the beginning of every bit period. A logic "1" is represented by a second transition one half bit period from the start of the bit. A logic "0" is represented by there being no transition within the bit period (see Fig. 27).

Vertical Interval Application

Each television frame is identified by a unique and complete address consisting of 90 bits of information,

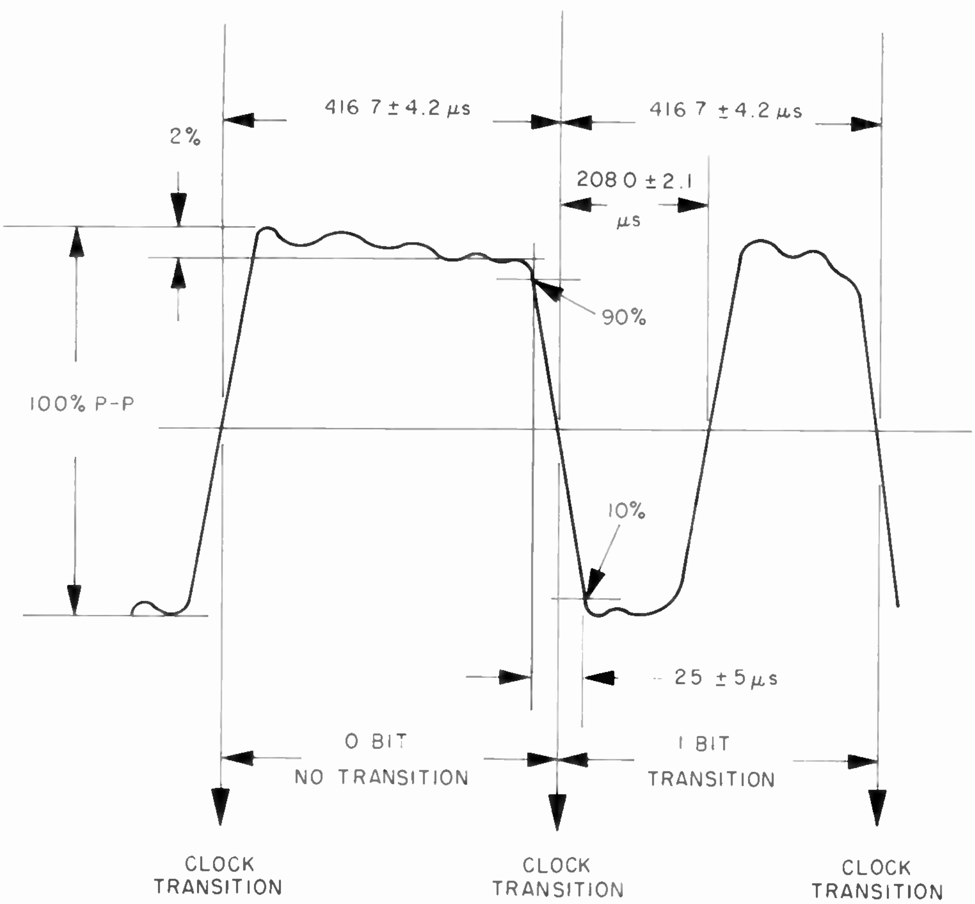
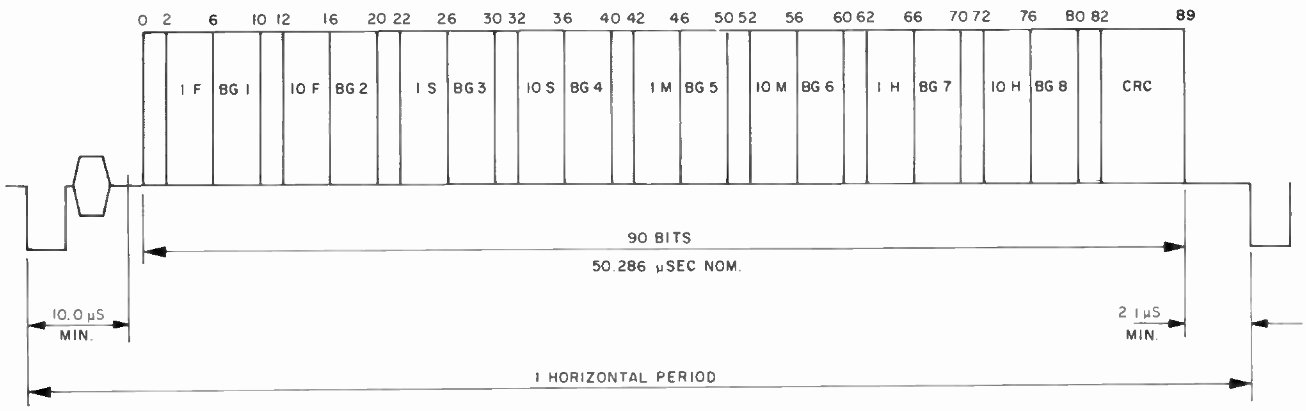


Figure 27. SMPTE time code—longitudinal record waveform.



0-1	Synchronizing bits	0 Fixed one 1 Fixed zero	30-31	Synchronizing bits	30 Fixed one 31 Fixed zero	60-61	Synchronizing bits	60 Fixed one 61 Fixed zero
2-5	Units of Frames		32-34	Tens of Seconds		62-65	Units of Hours	
6-9	First Binary Group	(BG 1)	35	Field Mark		66-69	Seventh Binary Group	(BG 7)
10-11	Synchronizing bits	10 Fixed one 11 Fixed zero	36-39	Fourth Binary Group	(BG 4)	70-71	Synchronizing bits	70 Fixed one 71 Fixed zero
12-13	Tens of Frames		40-41	Synchronizing bits	40 Fixed one 41 Fixed zero	72-73	Tens of Hours	
14	Drop Frame Flag		42-45	Units of Minutes		74	Unassigned bit	(Zero until specified)
15	Color Frame Flag		46-49	Fifth Binary Group	(BG 5)	75	Binary Group Flag	
16-19	Second Binary Group	(BG 2)	50-51	Synchronizing bits	50 Fixed one 51 Fixed zero	76-79	Eighth Binary Group	(BG 8)
20-21	Synchronizing bits	20 Fixed one 21 Fixed zero	52-54	Tens of minutes		80-81	Synchronizing bits	80 Fixed one 81 Fixed zero
22-25	Units of Seconds		55	Binary Group Flag		82-89	CRC Code (Cyclic Redundance Check Code. See 4.4.1)	
26-29	Third Binary Group	(BG 3)	56-59	Sixth Binary Group	(BG 6)			

Figure 28. SMPTE time code—vertical interval address bit assignment.

numbered 0 through 89. The frames are numbered successively 0 through 29 with an arrangement that accommodates the NTSC color system of 29.97 Hz (drop frame), as well as, 30 Hz systems (Figs. 28 and 29).

The bit assignments are as follows:

Bit#	*Application*
0–1	Sync Bits
2–5	Units of Frames
6–9	First Binary Group
10–11	Sync Bits
12–13	Tens of Frames
14	Drop Frame Flag
15	Color Frame Flag
16–19	Second Binary Group
20–21	Sync Bits
22–25	Units of Seconds
26–29	Third Binary Group
30–31	Sync Bits
32–34	Tens of Seconds
35	Phase Mark Correction Bit
36–39	Fourth Binary Group
40–41	Sync Bits
42–45	Units of Minutes
46–49	Fifth Binary Group
50–51	Sync Bits
52–54	Tens of Minutes
55	Binary Group Flag
56–59	Sixth Binary Group
60–61	Sync Bits
62–65	Units of Hours
66–69	Seventh Binary Group
70–71	Sync Bits
72–73	Tens of Hours
74	Binary "0" (Unassigned)
75	Binary Group Flag
76–79	Eighth Binary Group
80–81	Sync Bits
82–89	CRC Code

The eight Binary Groups are intended for storage of data by the users and the 32 bits are interpreted according to the settings of bits 55 and 75 as follows:

Bit 14: Drop Frame Flag: A "1" indicates that certain frame numbers are being dropped to resolve the difference between real time and NTSC color time.

Bit 15: Color Frame Flag: A "1" identifies NTSC color frame A.

Bit 35: Phase Mark: This bit is set as follows: A "0" represents the field in which the first equalizing pulses follows the preceding horizontal pulse by a whole line (field 1 or 3 in the color system).

Bits 82–89: Cyclic Redundance Check Code. The polynomial $G(X) = X^8 + 1$ is applied to all bits 0 through 81 inclusive.

Vertical Interval Modulation

The modulation method used in the vertical interval is a modified NRZ coding. A logic "1" is represented

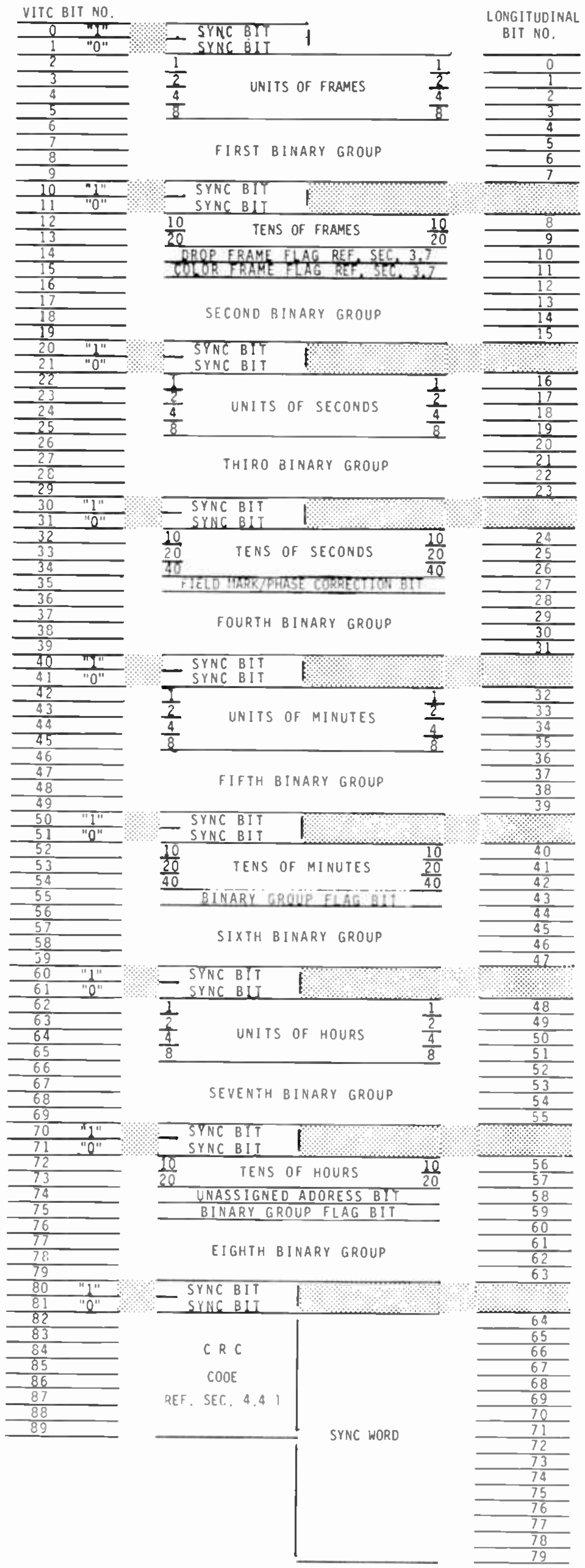


Figure 29. SMPTE time code—vertical interval record waveform.

TABLE 9

Assignment	Bit 55	Bit 75
Character Set Not Specified	0	0
Eight-bit Character Set	1	0
Unassigned	0	1
Unassigned	1	1

by a signal level of 80 ±10 IRE while a logic "0" is represented by a signal level of 0 ± 10 IRE. A transition occurs only when there is a change is data level (see Fig. 27). The bit rate of the information is F_e where,

$$F_e = F_h \times (455/4) \pm 200 \text{ Hz}$$

and F_h is the horizontal line rate.

Applicable Standards Documents

1. SMPTE/ANSI 12M—American National Standard for Television—Time and Control Code—Video and Audio Tape for 525-Line/60 Field Systems.

REFERENCES

1. Ginsberg, Charles; A New Magnetic Video Recording System," *SMPTE Journal*, Vol. 65, May 1956; pp. 302–304.
2. Frederick Remley, "An Overview of One-Inch Helical Video Recording," *One-Inch Helical Video Recording*, SMPTE, 1978, pp. 1–7.
3. SMPTE V16–861, Appendix C.
4. David K. Fibush, "SMPTE Type C Helical scan Recording Format," *SMPTE Journal*, Nov. 1978, p. 756.
5. SMPTE Engineering Guideline—EG 10—Tape Transport Geometry Parameters for 19mm type D-1 Television Digital Component Recording.
6. SMPTE Engineering Guideline—EG 21—Nomenclature for Television Digital Recording, 19mm type D-1 component and type D-2 composite formats.
7. James A. Howard and Lester Ferguson, "Magnetic Tape Reording Handbook," Hewlett Packard, 1967.
8. David K. Fibush, "SMPTE Type C. Helical scan Recording Format," *SMPTE Journal*, Nov. 1978, pp. 759–760.
9. For component systems reference SMPTE RP 125.
10. For composite systems reference SMPTE RP 244.
11. R. H. Stafford, "Digital Television," John Wiley and Sons, August 1980, pp. 286–288.
12. Dennis M. Ryan, "Mechanical Design Considerations for Helical scan Videotape Recorders," *SMPTE Journal*, Nov. 1978, pp. 767–771.
13. R. A. Hathaway and Ray Ravizza, "Development and Design of the Ampex Auto Scan Tracking (AST) System," *SMPTE Journal*, Dec. 1980, pp. 931–934.

14. J. Roizen, "Video Recording, Computer Graphics, Special Effects—An Overview" (1988).
15. As specified in SMPTE RP 86.
16. Reference Ampex VPR# Specification.
17. Reference SMPTE RP 87.
18. Reference SONY VBU-950 Specification.
19. Reference Panasonic AU300 Specification.
20. Reference SONY BVW 40 Specification.
21. Reference Panasonic AU660 Specification.
22. Reference SONY BVW 75 Specification with metal particle tape.
23. For composite signal output.
24. For composite signal output.
25. Linear channel standard, higher performance available with Dolby.
26. Dolby noise reduction is assumed.
27. Dolby noise reduction is assumed.
28. Reference IEC 60B—70, "Supplement to S-VHS Videotape cassette system."
29. Akira Hirota and Neil Neubert, "Recent Technical Developments in the S-VHS VCR for Broadcasting and Professional Applications," *SMPTE Journal*, May 1990, pp. 376–380.
30. Reference Nova Systems novasync-2 TBC.
31. Reference Panasonic AG-7500.
32. A description of the 525 line version of an interface meeting the requirements of CCIR Recommendation 601 is found in SMPTE Recommended Practice RP125.
33. SMPTE Engineering Guideline—EG 10, Tape transport geometry parameters for 19mm type D-1 Television Digital Component Recording.
34. S. W. Athey, "Magnetic Recording," NASA-SP-5038, Washington, DC, January 1966.
35. SMPTE 227M, Standard for Television Digital Component Recording—19mm type D-1—helical data and control records.
36. SMPTE 224M, Standard for digital component recording – 19mm type D-1—tape record.
37. Reference SMPTE RP125.
38. Reference DCT 100 specification.
39. SMPTE 247M.
40. N. Kaku, S. Ozaki, et al., "Mechnical Considerations in the Design of a Composite Digital VTR," *SMPTE Journal*, Aug. 1989, pp. 568–574.
41. Reference Ampex VPR-300 Manual.
42. Reference Ampex VPR-300 Specification.
43. Reference Ampex VPR-300 Specification.
44. Y. Oba and K. Murayama, "A New Small Format Digital VTR Using ½-Inch Tape," 1990 NAB Engineering Conference Proceedings, pp. 64–67.
45. Reference Panasonic AJ-D350 Specification.
46. Reference Panasonic AJ-D350 Specification.

5.7 Videotape Editing

Michael T. Keller
Keller Systems Technology, Southborough, Massachusetts

INTRODUCTION

The subject of videotape editing encompasses a collection of continuously evolving technologies. The purpose of this chapter is to review each of these technologies, examine some of the options, and examine how these options may be integrated to meet the needs of broadcast facilities.

Both facilities and demand for videotape editing were once limited at the average broadcast facility. Simple editing was done by using a pair of regular VTRs during network programming or after hours.

This was a relatively slow but effective technique that was practiced by videotape editors and engineers. This method prevailed in many broadcast facilities for production and news editing until the conversion to electronic news gathering (ENG). ENG produced a demand for high volume, fast turnaround editing. To accommodate this tempo, news editing moved out of the VTR room and into dedicated suites. Soon after the ENG conversion, the video magazine program format drove the demand for in-house videotape editing further. A hybrid editing suite was developed, using ENG recording formats. It provided more sophisticated effects and eventually developed into a mini A/B roll editing suite, employing both ENG and full-scale post-production technology.

Since the initial years of the ENG revolution, the news edit suite has evolved into a complete turn-key system, requiring minimal engineering. Similarly, evolving technology allows small and medium scale turn-key A/B roll editing systems to be built for magazine format editing.

It may appear that the technical design challenge has been diminished. Rather, it has been refocused: the number of choices to be made today are substantially greater than before. In previous versions of this chapter, a discussion of the difficulties of interfacing equipment with various machine control systems consumed the bulk of the material. The video signal system was always assumed to be a composite analog design. Today, the choices seem to have been reversed. The question is which signal system format (component or composite, digital or analog) to use with the now assumed "standard": serially interfaced equipment.

HISTORICAL PERSPECTIVE

The initial approach to editing quadruplex videotape directly emulated film techniques: cut and splice. The tape was played, edit points were marked on the tape with a grease pencil, and the undesired material was simply cut away. To provide as smooth a splice as possible to the VTR's servo systems, the frame pulses and matching tracks were visually identified under a microscope after "developing" the tape with a liquid suspension of ferrous particles. This form of "assemble editing" required great craftsmanship and edit decisions had to be right the first time.

Such mechanical editing techniques were replaced by the first generation of electronic editors. VTRs equipped with electronic editors contained systems for correct record timing and servo switching to switch smoothly from playback to record mode while the tape continuously rolled.

To edit, the operator, using two VTRs, would mark each tape at the preroll point, backspace each VTR by hand to allow time for lock-up, and manually start them as simultaneously as possible. If both VTRs locked-up before they reached the edit point, the operator would manually initiate the edit. If the editing operator's reactions were off slightly, causing the desired edit point to be missed, there would not be a second chance. There was a clear need for a method that would reduce the dependence on operator reaction time and allow the testing of an edit before actually performing it.

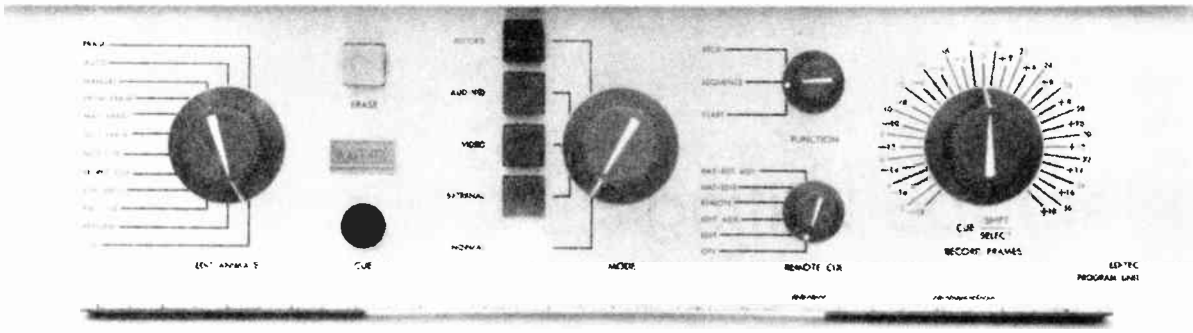


Figure 1. Ampex Editec. (Courtesy Ampex)

The Ampex Editec (Fig. 1) addressed these problems by allowing the operator to record an audio cuetone to mark the edit point. This tone was used to trigger the VTR's internal electronic editor.

An edit could be previewed by using the cuetone to switch the record VTR's input to the play VTR without turning on record and erase systems. The play/source VTR's output was then routed through the electronics to electronics (E to E) path of the record VTR. This became the commonplace method, but it still relied on manual prerolling of the transports.

Figure 2. Ampex HS-200 operating console. (Courtesy Ampex)

Lost in the transition from film to quadruplex videotape editing was the ability to examine edit points on a still frame. This problem was addressed by magnetic disc technology, as exemplified by the Ampex HS-200 (Fig. 2), which integrated the slow motion playback capabilities and programmable control of the HS-100 video disc recorder with cuetone control of quad VTRs. This allowed the slow motion and still-frame capability of the disc system to provide still-frame identification of edit points. The ability to program switcher transitions and effects between the HS-100 and VTRs under control of the edit marks was also contained in the HS-200.

The problems of synchronizing multiple VTRs for A/B roll effects were addressed by the using a predecessor of time code in the Ampex RA-4000 (Fig. 3). A unique address for each frame was laid down during recording of the cue track. This number stream provided references for transport control servo systems, allowing a desired frame to be located automatically. Using the code as a frame number reference, the multiple play VTRs and the record VTR could be automatically synchronized by the RA-4000.

One of the first systems to provide random-access editing was the CMX 600 system (Fig. 4), which used a light pen to select source edit points from a total of 30 minutes of video recorded on six disks. This capacity was attained by recording single monochrome fields with audio tracks encoded into the front and back porch of the video signal.

Although editing with the CMX 600 did not produce a master tape, the system generated a list of time-code numbers defining edit decisions. This list could be printed or punched onto paper tape. The CMX 300 on-line system would be loaded with the papertape and automatically assemble the master. Today the wide-spread use of analog and digital helical VTRs provides a reasonable emulation of the still picture editing ability of film.

SYSTEM TYPES

Editing systems are often identified by the level of output they produce, the method of accessing that

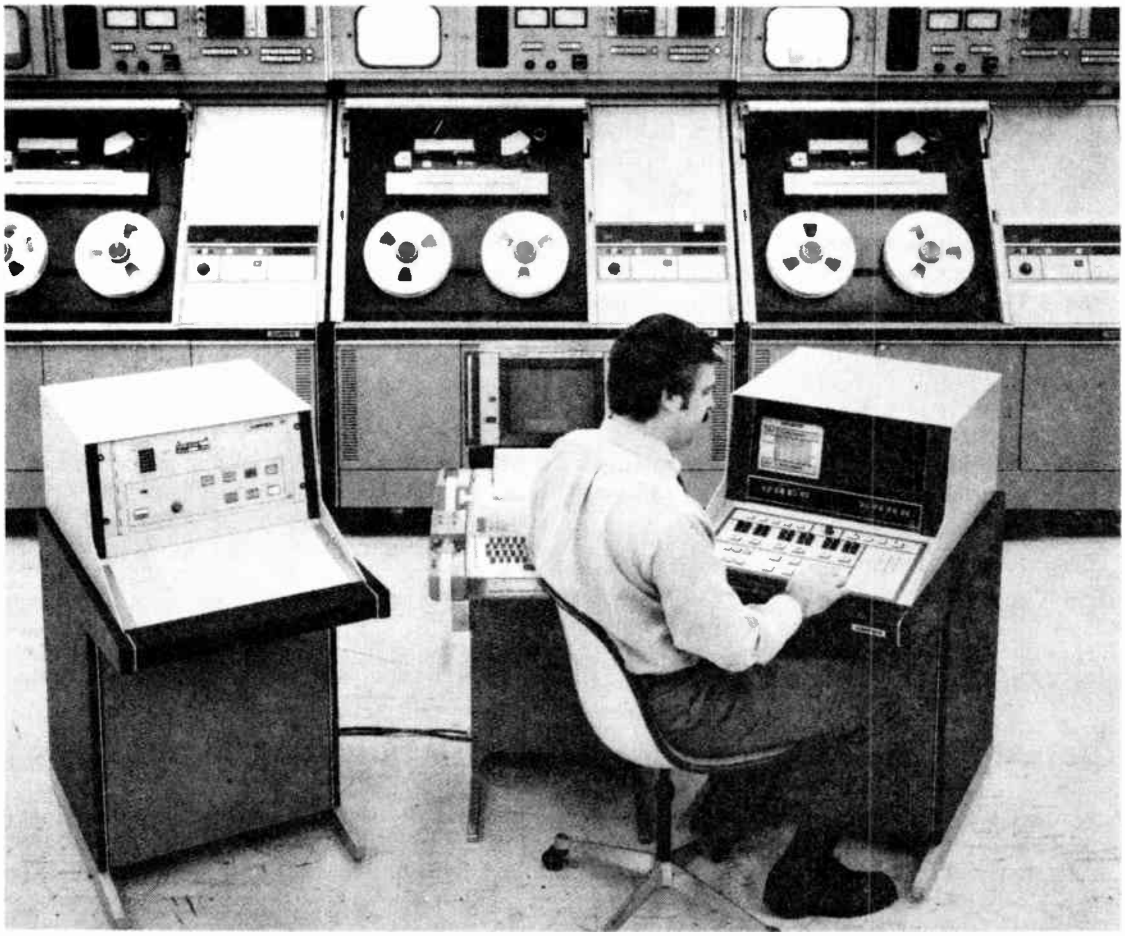


Figure 3. Ampex RA-4000 controlling AVR-1 VTR's. (Courtesy Ampex.)

material, the videotape formats used, and the type of signal system employed. An edit suite that produces a final product or master tape is referred to as an *on-line* suite. The format may be anything from Type C to VHS, but, as long as the resulting tape is the final product, that system is acting as an on-line suite. The on-line suite uses the facility's highest quality and most sophisticated technology, and therefore requires highly-skilled operators. This combination means a high hourly operating cost.

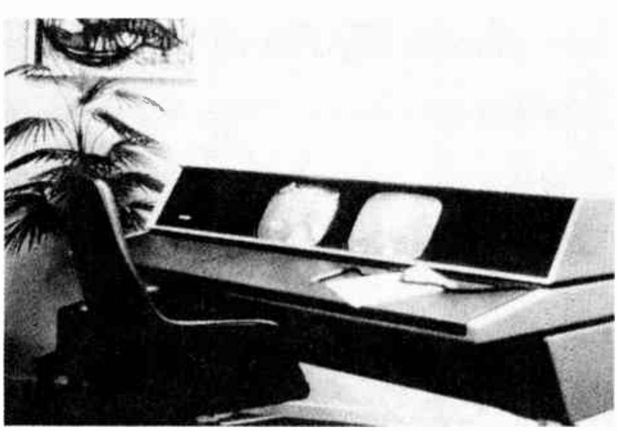


Figure 4. CMX-600 operating console. (Courtesy CMX)

Using an *off-line* edit suite is a method of holding down the number of hours used in on-line editing. Although the equipment used in some off-line facilities may technically surpass that of an on-line suite, the defining factor is that the resulting tape is not a complete ready to air program. An off-line room's product is most often a list of the desired timecode edit locations, called an edit decision list (EDL) and a cuts-only recording.

The off-line suite allows the creative person to edit and re-edit using dub copies of the source tapes until the continuity of the show is firm. The time spent identifying these basic cuts is better spent using a system designed specifically for that purpose. Once the off-line decisions have been made, the source material is taken into the slower and more costly on-line facility to duplicate the cuts made in off-line by using a computerized editor to auto-assemble the source tapes onto a record master.

The final on-line master is produced by adding the audio and video special effects and graphics to the material from the auto-assembled tape.

Another cost-saving technique is nonedit list generation. In its most elementary form, this consists of compiling a paper list of timecode numbers for the desired takes from raw footage. These numbers can be read from a timecode display that is integrated in or attached to the playback VTR. The list is then manually entered into the on-line editing controller.

A more efficient method uses a timecode reader and logging software with a laptop computer. This combination allows time code numbers to be entered into an on-screen list directly from the tape's timecode track with the press of a button. Notes about the selected material can then be entered next to a time code number in word processor fashion. After the logging session, the list can be down oaded to the online system via a serial port, transferred to floppy disk or printed.

OPERATIONAL TYPES

Editing systems can be divided into two operational types: linear and random-access. Linear editing is the most prevalent technique. Each source tape is individually loaded and edited into the master reel. Operating in a linear system imposes a number of constraints. The speed of searching for source material is limited by the tape length and the time and tedium of changing reels. A transition cannot be compared with another version. This tends to limit the choices readily available to producers.

Random access editors improve creative flexibility with a combination of techniques, such as user-friendly pictorial interfaces and instant access to source material. To provide high-speed retrieval, source material is transferred from tape to a random access medium, such as optical disk, magnetic disk, or tandem synchronized VTRs. To provide a pictorial index when searching the stored material, multiple compressed stills are displayed, with time code numbers superimposed. This technique is similar to the browse mode used in many still-store systems. Edit points may be chosen by selecting the desired material from the compressed stills with a cursor or keypad entry.

Editing and re-editing with word processor flexibility by using images to cut and paste transition points provides an intuitive feel to which artistically-inclined users quickly adapt. To appreciate the power and flexibility of random-access editing, a hands-on demonstration is a must. Examples of random-access editors are the CMX600 and 6000, Montage Picture Processor, and Avid Media Composer (Fig. 5). Unfortunately, storage technology has not produced a practical, economical, on-line grade, random-access recording medium. This has limited the previously mentioned systems to operating as off-line front ends for slower but higher quality linear on-line editing systems.

SYSTEM FUNCTIONAL BLOCKS

Fig. 6 lists the functional blocks of equipment often found in an on-line A/B roll edit suite. Each category, with the exception of the editing controller, could be found in a broadcast air-production environment. The distinguishing features that make them suitable for

Figure 5. Avid Media composer. (Courtesy Avid)

EDITING CONTROLLER	AUDIO TAPE RECORDERS
VTRs	AUDIO TAPE SYNCHRONIZERS
VIDEO EFFECTS SWITCHER	AUDIO MIXER
DIGITAL EFFECTS UNIT	AUDIO EFFECTS PROCESSORS
COLOR CORRECTOR	TEST, MONITORING AND TERMINAL EQUIPMENT
CHARACTER GENERATOR	
GRAPHICS CAMERA	

Figure 6. Edit suite equipment categories.

editing applications are typically the addition of remote control interfaces, time code capabilities, and signal matching.

CONTROL SYSTEMS AND INTERFACES

Editing Controllers

A common architecture in many editing control systems, (Fig. 7) divides the hardware into two categories. The first is the central computer running the high-level editing software and often supporting the screen display, keyboard, and disk storage facilities. The second is the collection of interfaces linking the host central processing unit (CPU) and the control ports of the video and audio equipment. In the case of the VTR interfaces, each has a dedicated microcomputer slave system running a software routine specifically written for that type of VTR. This division of tasks has a number of advantages over a single centralized processor. Software in the central computer is not affected by changes in the user's video and audio hardware; only the interface program is changed to match the new unit.

Edit controllers using this architecture have traditionally been built as dedicated hardware systems. The system is delivered complete by one manufacturer. A custom keyboard, an industrial buss computer and card cage for the host, and separate frames containing dedicated microcomputer interfaces for the video and audio equipment are standard. Other versions of this basic concept place all interfaces on the same chassis as the host CPU and change the keyboard layout.

Recently, systems using personal computers as the host CPU have appeared in the marketplace (Fig. 8). The personal computer (PC) runs the high level editing software while providing an inexpensive platform for disk storage and printing capabilities. The original keyboard is customized to match the editing application simply by changing the keycaps. To communicate, a network control card links the PC to the dedicated microcomputer interfaces of the VTRs and the effects switcher.

In addition to the obvious hardware economy of using PCs, there is the further advantage of using custom software to expand system capabilities. Nu-

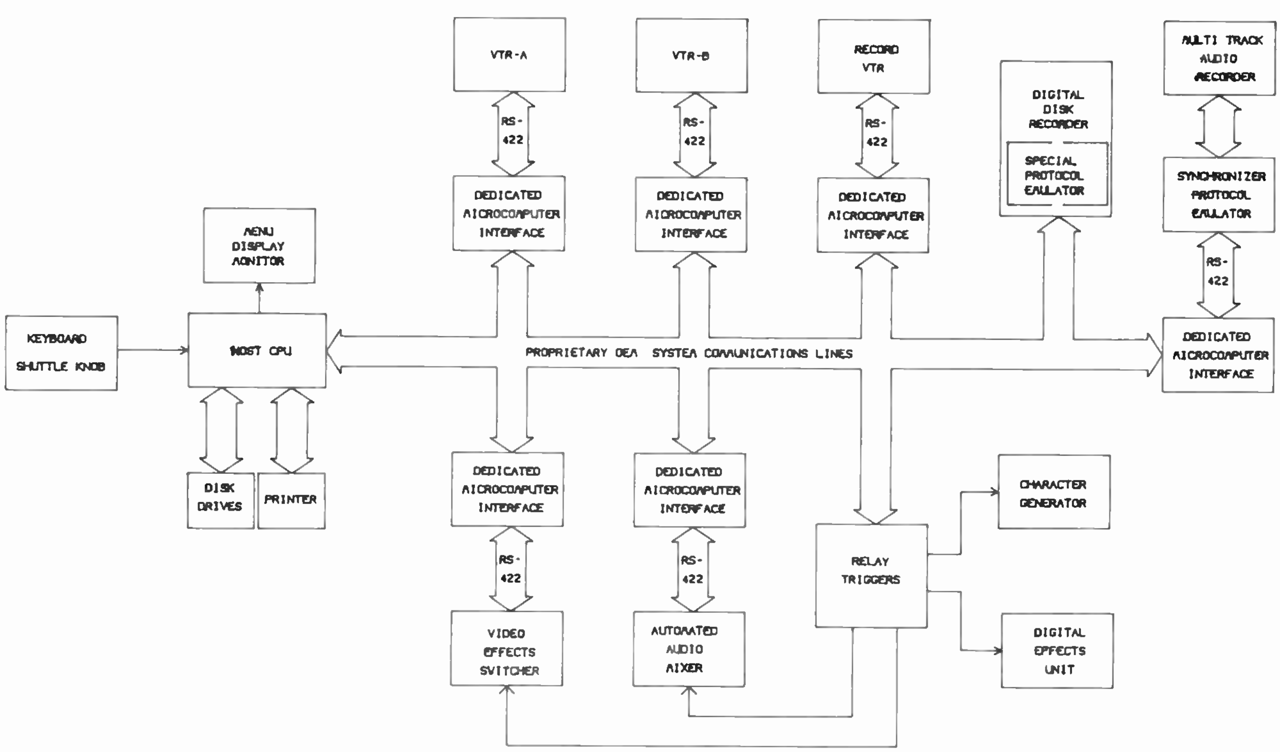


Figure 7. Generic editor control system architecture.

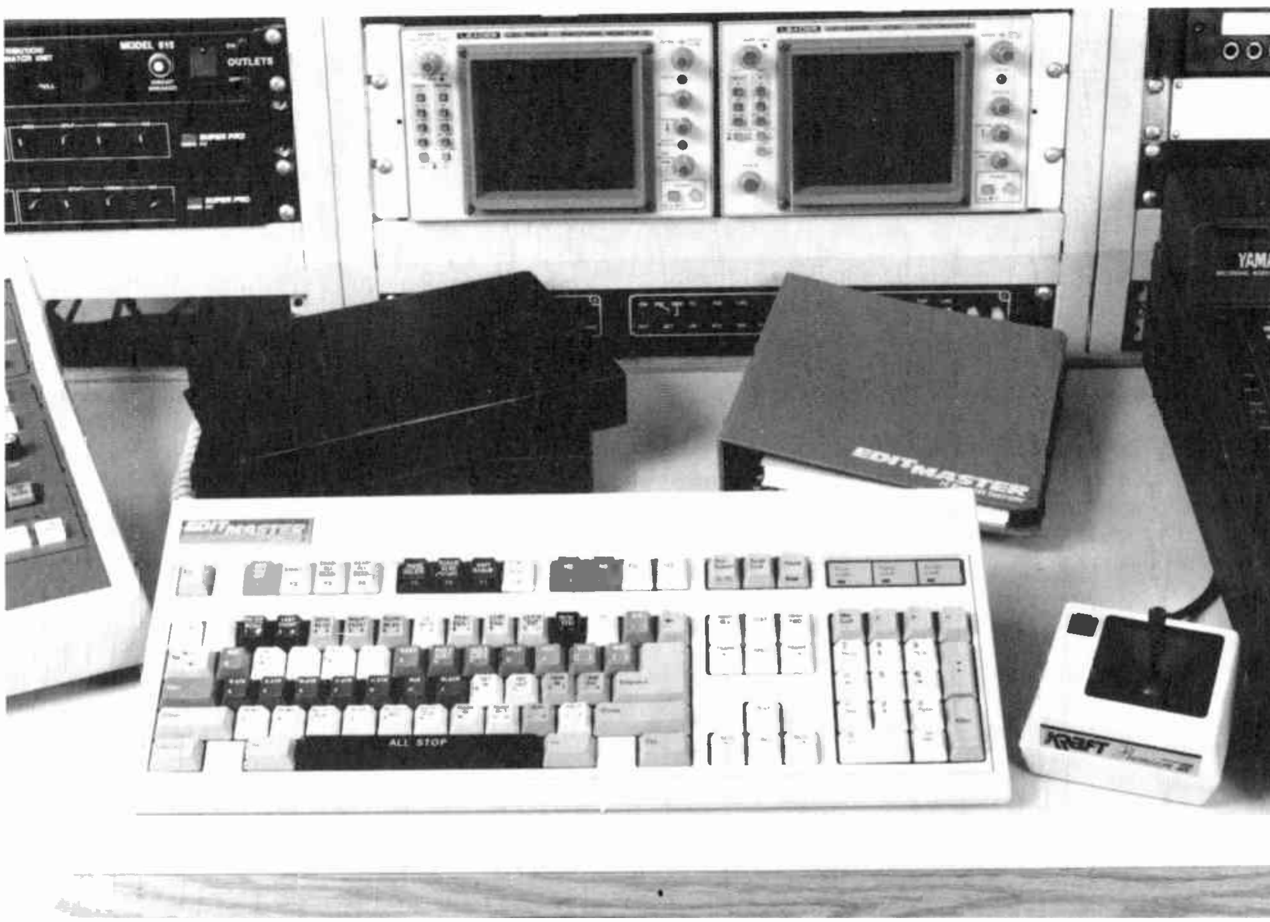


Figure 8. Edit master PC based editing system. (Courtesy CV Technology)

merous possibilities (such as networking facility resources into multiple edit suites) exist. This could offer options ranging from exchanging lists to sharing control of special format VTRs, much as word processing networks now share laser printers. As of this writing these are not yet available features but they are being developed.

Multiple sizes and formats of edit decision list (EDL) disks present an interchange problem for many users. EDL disks may be produced in different disk sizes, with various writing densities and for different operating systems. (See Fig. 9.)

Other differences exist in the way the edit information is described: the EDL format. In pursuit of interchangability, some manufacturers, in addition to their own format, provide an alternate "CMX compatible" edit list format. In facilities where incompatibility between different EDL disk formats is a recurring problem, a practical approach may be to centralize the interchange task by using a dedicated interchange station.

Editor	Size/Type	Operating System
CMX 3400	8" SSDD	DEC RT-11
CV-TECH.	5" DSDD	MS-DOS
GVG	D.5" DSDD	DEC RT-11

Figure 9. EDL floppy disk formats.

An interchange station consists of a PC equipped with floppy disk drives for the different formats, suitable drive controller cardS, and reformatting software. REFORMATTER by MicroTech Exports is a package of programs that addresses the interchange between DEC RT-11 format (used by CMX) and MS DOS (used by many personal computer systems). REFORMATTER provides the ability to read, write, and format RT-11 disks using a PC and to transfer their contents between drives of different sizes. In order to use this type of software, a disk controller card designed to use an eight-inch floppy disk must be used.

Due to the space limitations of the typical PC cabinet, it is often necessary to connect the eight-inch drive and power supply externally. Such a work station provides additional benefits, such as the ability to print lengthy edit decision lists without tying up an edit suite and to test troubleshoot faulty drives.

Interfacing VTRs

Interfacing the numerous models, types, and formats of VTRs to an editing controller has always been a challenge to the manufacturer and the installing engineer. Fortunately, advancing technology has reduced the number of unique hardware interfaces for broadcast grade equipment.

Adding a different VTR to an editing system presents several issues to the engineer. The most obvious one is connector compatibility problems between the remote control port on the VTR and the editor. The current generation of equipment incorporates a serial RS-422 remote control system using a protocol conforming to SMPTE RP 113–1983 and using nine-pin D-type connectors (Fig. 10).

If a serial port is not provided by the manufacturer, the addition of a parallel-to-serial converter and protocol emulator may be necessary. These interfaces contain a microcomputer-based system that allows switch-selectable conversion between several models of parallel-port-only VTRs and a common serial VTR protocol.

Once the connectors and communication standards are resolved, the engineer turns to matching the editor's interface software to that of the VTR and optimizing it. The edit controller is equipped with software drivers for various VTRs, either loadable from disk or loaded into a programmable read only memory (PROM). Unfortunately, since the VTR may have been manufactured by some other company or at a later date, the correct driver may not be available. This situation is similar to that often encountered by personal computer users when adding new hardware to an existing system.

The best course is usually to ask the editor's manufacturer for a copy of the current driver, and guidance in using diagnostic software. The major editing equipment manufacturers routinely arrange with manufacturers of video recording equipment to borrow production models of new VTRs to develop interfaces. While they valiantly try to keep in step with VTR software upgrades, it is a running battle. During the operational life of one popular family of VTRs, more than twenty versions of the driver were released. Some of these were the inevitable "bug fixes," but most were upgrades and speed improvements.

A few manufacturers allow the user to optimize their interface software manually by providing a "debug" mode. This allows the user to alter the default control table values and store the modified values in nonvolatile random access memory (RAM). Some manufacturers add adaptive software routines that learn the necessary offsets. The first attempt at synchronizing a VTR with such software may abort or require excessive preroll time. The next attempt is likely to go better as the system uses correction values which it learned on the previous try.

If the editor does not have a driver written specifically for a particular VTR model, limited control may be established by emulating an older but backwardly-compatible model. This often means that features found only on the newer product cannot be accessed and that overall performance is downgraded to that of the older model.

Once the link to the VTR is established, the interface provides the host computer with full remote control of the transport and editing functions, the tape position and time-code information, and synchronization of the transport to the house reference sync before it arrives at the desired frame.

External editing control systems rely on the record VTR to implement the electronic edit sequence. An operator, using the edit controller keyboard, determines the place to perform the edit, the tracks to be edited, the duration of the edit, and the editing mode. The editing record VTR provides the internal systems necessary to switch servo references and sequence the erase heads and record/play electronics. These timing sequences are determined by the physical constraints of the tape format, involving the video and audio head positions and tape speed.

The following example is based on the operation of the internal edit timing system of a Type C one-inch VTR: The reception of an edit record command starts a timer which controls the flying erase head. This head is located in the video scanning assembly (Fig. 11).

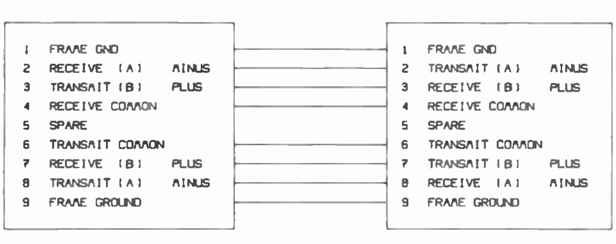


Figure 10. RS-422 9 PIN wiring.

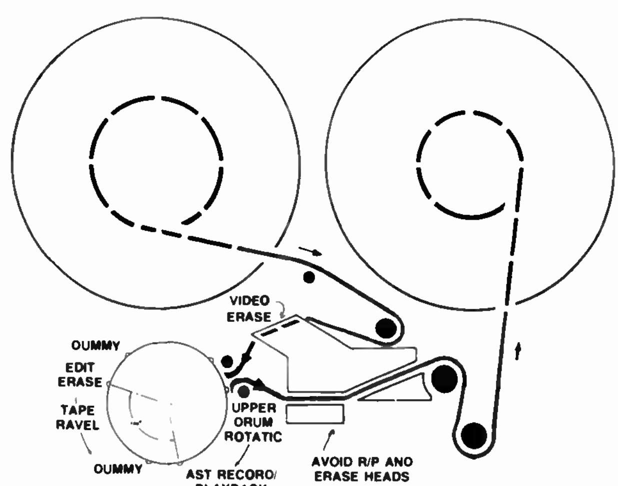


Figure 11. Head position diagram Ampex BPR-80. (Courtesy Ampex)

and erases video tracks which have travelled past the stationary video erase head prior to the start of the edit. The flying erase head traces the same path as the helical record/play (R/P) head and therefore can erase only the video track.

The record/play relay switches to the record position after the second timer in the sequence expires. This is followed by another delay timer edit interval when recording begins. At the end of the video edit interval, the sequence is reversed. The flying erase head shuts down first, then the record/play relay switches back to play, and the record amplifier is turned off.

To perform an audio edit, the edit timing systems control the audio erase current and the record switching for any or all channels. The audio erase current is applied to the erase head in advance of the edit point. This compensates for the distance between the audio erase head and the record/play head stack. When the erased segment arrives at the record/play stack, the record amplifier is ramped-up by the completion of its timer cycle, and the audio is edited in. Conversely, the erase command is removed several frames before the edit-out point reaches the R/P head and the audio record amplifier is ramped down upon reaching the edit-out point.

The internal VTR editor can perform edits in two modes: *assemble* and *insert*. *Assemble* editing is similar to nonedit recording in that the original control track is erased and a new one recorded. The new material is added onto the previous material upon arrival at the edit-in points in a building block fashion. The new material, having been recorded, then becomes the "previous" material to which another segment may be added. An edited tape is built by adding segment to segment over previously blank tape. This assembly of new material onto previously recorded blank material typically occurs for all tracks.

In the *insert* editing mode, the newly recorded material replaces previously recorded material for an edit interval shorter then the original recording length. The new material is recorded over previously-recorded material between the edit-in and edit-out points. The full width erase head is disabled and erasures occur on specific tracks. The original control track pulses are maintained, reducing servo instability caused by differences between the original and new recordings.

Time Code

An additional area of concern when interfacing a VTR to an editor is the handling of off-tape time code. Time code (specifically SMPTE Longitudinal Time Code (LTC) as described in ANSI Standard V98–12) is usually recorded on an address track specifically designed to handle the fast rise-time of its square wave envelope. LTC can be decoded at play or higher speeds, but as tape speed drops to 1⁄20 of play speed, signals reproduced by the stationary head become unreadable.

The popularity of helical scan VTRs with slow speed and still frame capabilities has made accurate reading of timecode at these speeds imperative. Some editing controllers interpolate a slow speed "tach" from the last-known LTC number or utilize control track pulses until LTC is valid. Fortunately numerous broadcast-grade VTRs now include both vertical-interval and longitudinal time code reader-generators. Vertical-interval time code (VITC) contains the same time and frame-numbering information found in LTC along with some additional improvements such as an error-checking technique called cyclic-redundancy check and the addition of a field-identification bit. VITC data is recorded on two separate lines in the vertical interval. This redundancy reduces code loss due to tape drop-outs. In order to accommodate these additional features, VITC requires ten more bits than the 80 bits in LTC (Fig. 12).

If a VTR is equipped with an internal VITC/LTC reader, provision must be made to determine which time code will be used. Many VTRs offer a choice of manual or automatic off-tape timecode selection. The automatic mode provides selection of LTC during all readable speeds, switching over to VITC during slow-motion. This would appear to be an ideal mode. Unfortunately, an ever-present operational trap when using VITC is attempting to perform a video insert edit with the new video containing VITC different from LTC. The edit controller may become confused when time code jumps at slow speeds, not knowing which code represents the current tape location.

The edit controller uses time code as the only basis to determine if a pair of edit points will produce a matched color frame edit. In a properly operating timecode generator, the even frame numbers identify NTSC fields 1 and 2 or Frame A and odd numbers identify fields 3 and 4 or Frame B. Bit 11 of the code signals that this code identifies a color frame sequence. If code is placed on tape without being color frame locked to the recorded video, the editor interface will supply invalid data to the host CPU.

Another source of possibly erroneous information is the phase relationship between the position of time code relative to video on tape. The boundary between bit 79 and 0, where one code number ends and the other begins, should be aligned with the first serration in vertical sync. Substantial mismatch can be caused by mechanical misalignment of a LTC record/playback head. This may cause the editor to abort the edit due to synchronization error. The VTR's capstan servo attempts to maintain lock to house sync by using control track written in proper phase with video. Simultaneously, the editor interface bumps the capstan servo via the tape speed override control line using the out-of-phase time code as a reference. The ensuing battle causes the VTR to remain unlocked.

The list of problems that are blamed on VTR interfaces but originate in time code is a long one. Fortunately, diagnostic equipment is available to help resolve time code problems (Fig. 13). The ADX time code analyzer graphically displays the status of all 80 LTC bits, as well as indicating code-to-video relationships.

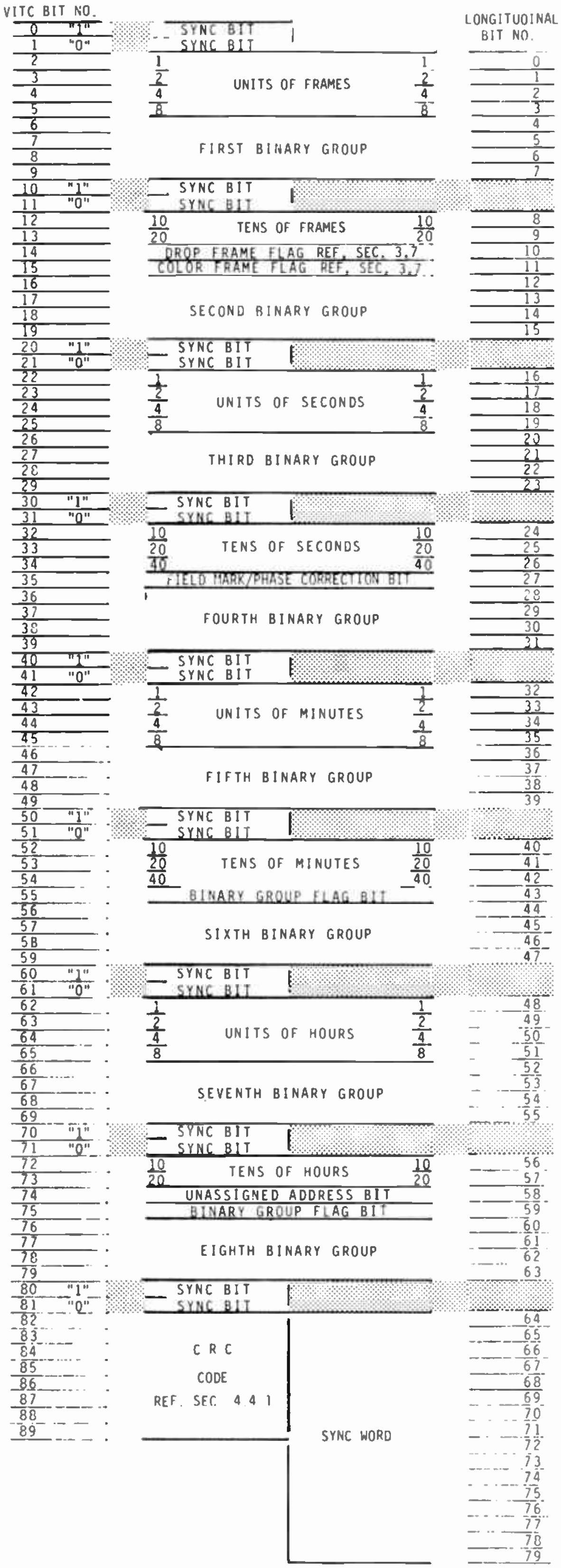


Figure 12. Relationship of vertical interval timecode to longitudinal code. (Reprinted with the permission of SMPTE)

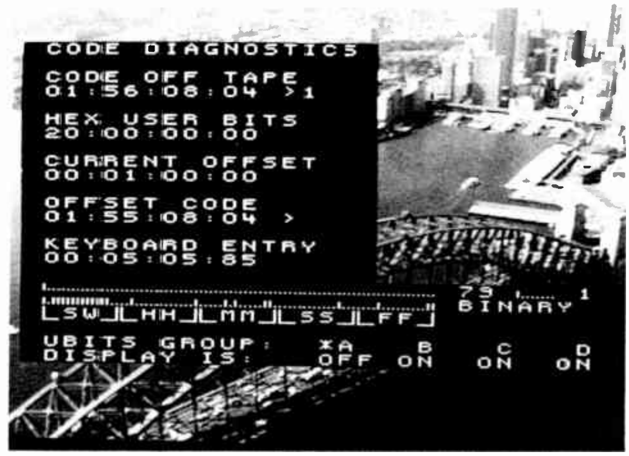


Figure 13. ADX-02 timecode analyzer diagnostic. (Courtesy ADX)

Audio Tape Interfaces

With only a few, very expensive exceptions, audio tape recorders lack serial interfaces designed for editor control. The addition of an external synchronizer and protocol emulator (Fig. 14) to a suitable audio tape recorder (ATR) provides a practical solution. This combination allows an editor to use the ATR as if it were an additional videotape transport.

The synchronizer (Fig. 15) of the emulator system provides the servo loop to lock the audio tape transport to system video sync. The emulation section converts the incoming commands for rewind, fast-forward, play, and other operations (along with capstan servo slewing voltages from the VTR format) to the ATR format.

This application places additional technical requirements on the ATR. A suitable unit must make available for remote control all transport-mode commands, capstan speed override and record/playback functions. Additionally, a broadband track for recording time code information (one which will function over a wide speed range without cross-talking into adjacent tracks) is a necessity.

Some users may not need the ability to command an ATR through emulation. Perhaps the ATR is only providing additional or substitute tracks for those of a VTR. With only a synchronizer, an ATR can follow the time code of an editor-controlled VTR and act as a slave transport.

Figure 14. Evertz emulator audio transport interface. (Courtesy Evertz)

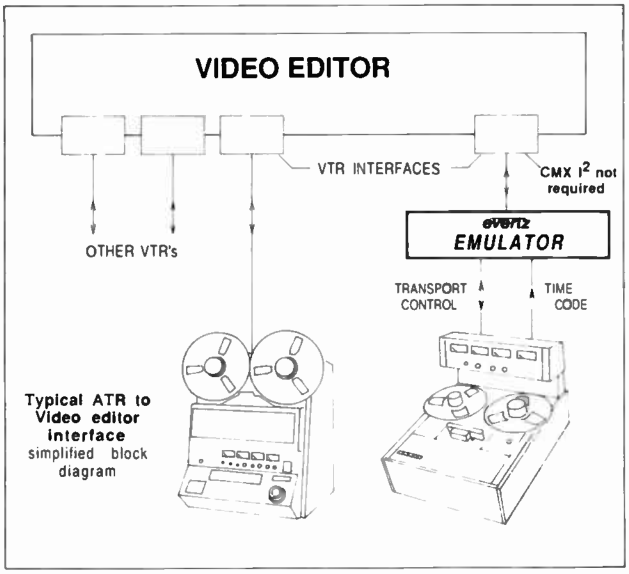


Figure 15. ATR to editor interface using emulator. (Courtesy Evertz)

Video Switcher Interfaces

An ability to control video switcher functions from the editor and store information in the edit list provides a number of operational advantages. Transitions such as dissolves, wipes, or keys may be performed repeatedly, reducing the risk chance of error found in manual operation.

The use of microprocessor systems to control most switchers has brought them conveniently and effectively under the control of the editors. Beyond allowing for remote control, many switchers (Fig. 16) extend the microprocessor system to provide an internal event memory system that stores control panel settings. Once an effect is stored in the switcher it can be recalled and triggered by the editor.

A switcher is often interfaced to the editor utilizing a combination of techniques (Fig. 17). Contact controlled by the editor are wired to opto-isolator inputs on the switcher along with an RS-422 type serial communications line. The contacts are used to trigger switcher memories with a minimum of operator keystrokes. It is sometimes more convenient to store an effect in switcher memory and trigger it than to build an edit list containing numerous registers describing it. The serial line provides the ability to control remotely a multitude of panel functions such as crosspoints and transition types. The serial data line can also select memory registers by address and trigger them or read them into the editor for storage in the EDL.

Even though the popular nine-pin, D connector, RS-422 serial communications configuration is commonly used for the hardware interface, a unique software driver for each manufacturer and model of switcher is usually required. The common exception is the Grass Valley Group Model 100 protocol which is often emulated by manufacturers of similar size switchers.

Digital Video Effects and Character Generator Interfaces

The interfacing of digital video effects systems or character generator systems is the least standardized of all video equipment that should be under direct and flexible control of an editor. A few serial interfaces exist, but their availability must be determined by contacting the edit controller manufacturer.

If a serial interface is available for a character generator, the control functions provided are usually limited to basic commands such as page changes, roll, crawl, pause, and speed. The selection of functions provided is often unique to the combination of editor

Figure 16. Video effects switcher with editor control. (Courtesy Ampex)

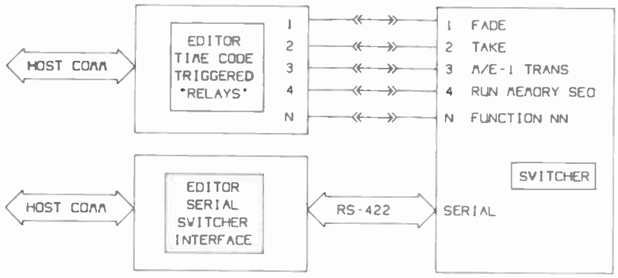


Figure 17. Video effects switcher interfaced with serial and GPI control.

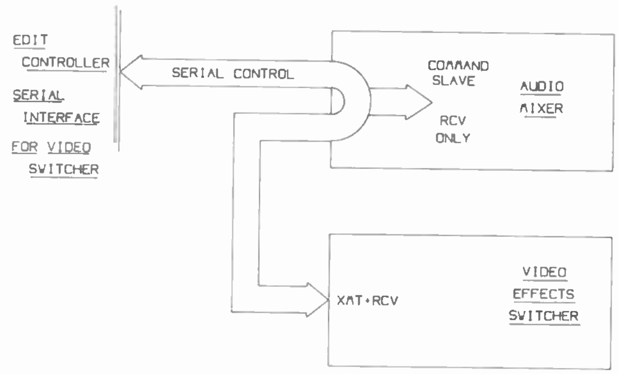


Figure 19. Audio mixer controlled as slave of video switcher.

and character generator (CG), making EDL interchange between edit suites limited.

If a serial system is not available, a set of external triggers for basic functions is provided for contact closure control from the editor. This topic is expanded under the section on "General Purpose Interfaces," which includes a diagram of a contact closure interface to a character generator via a keyboard emulator.

The limitations encountered when attempting to provide editor control of digital effects systems are similar to those encountered with character generators. An alternate method of controlling a digital video effects unit (DVE) may be possible via an indirect path. Some switchers are designed to control a DVE as an extension of the switcher's overall effects system. This configuration (Fig. 18) includes a serial control line that allows a digital effect to be selected and run through the switcher's control panel as part of a switcher transition. The DVE control selection is stored in switcher effects memory as part of the overall switcher effect. If the switcher memory is then stored in the EDL, the DVE move selection is, therefore, stored with it. The stored information may only specify the effect number and its duration. Design information for that effect may have to be stored on the DVE's own disk system.

Audio Mixer Interfaces

The ability to interface an audio mixer to the editing controller provides advantages similar to those achieved when the editor controls a video switcher and is accomplished with similar techniques. The level of editor control ranges from simple audio-follow-video switching to complete, full-scale automation.

In many editing installations, full scale automation may be too expensive. A method of limited control that still provides a labor savings can be implemented by slaving the audio mixer control system to commands sent to the video switcher (Fig. 19). In this system, the editor's serial control line for the video switcher is looped through the audio mixer. The loop allows commands issued to the video switcher to be monitored and followed by the control system of the audio mixer. This method will only provide an audio-follow-video mode, but creative operators can use a dummy video source and edits to extend the system's flexibility.

At the other extreme, an audio board designed for automated edit suite mixing provides the editor with remote control of channel faders, source selection, and preview monitoring. Functions accessible by the editing system provide the ability to select crosspoints that follow video or provide split audio using multiple channels. Level trimming data may be available to the edit controller to allow listings of trim values in the edit decision lists. Programmed transition modes available are "V" fades, cross fades, cut/fade and fade/cut. An "over" buss is selectable, allowing common voice-over mixes to be added into the master output.

These functions are made accessible for editor control through an interface to the internal microprocessor systems used by the mixer to operate the programmable channel amplifiers, solid state crosspoints, and control memory system. An audio mixer's control panel memory system allows settings and effects to be stored and recalled. Therefore, external control of this system, by either a serial interface or triggering by external contact closure, provides a powerful control link to the editor. The serial control protocol used by an audio console may be proprietary, particularly if the manufacturer is also a supplier of editing controllers. A few manufacturers who concentrate on building post-production consoles for the general market have made their protocols available to editing manufacturers. These protocols have become common enough that many editors will come equipped with drivers for them.

Figure 18. Editor control of DVE through video switcher.

General Purpose Interfaces

The general purpose interface (GPI) system is the simplest and most versatile of interfaces. The GPI is a contact closure or logic level that can be asserted at

GPI #	Controlled Device	Function
1	Digital Effects	Forward
2	" " " "	Reverse
3	" " " "	Stop
4	Video Switcher	Autotrans Trigger M/E-1
5	" " "	" " " " M/E-2
6	" " "	" " " " DSK
7	" " "	Fade to Black
8	Countdown Generator	Start
9	Audio Tape Recorder	Start
10	Character Generator	Start
11	" " " "	Stop
12	" " " "	Macro Trigger

Figure 20. GPI controlled equipment table.

a desired trigger time to activate a wide variety of external devices (Fig. 20). The trigger output may be either a latched or momentary signal.

Often the GPI is the only method available to control a device to which multiple users desire access. A patching system or GPI router allows flexibility when there is not a device to serve each user.

EDIT SUITE SIGNAL CONFIGURATIONS

The video signal system of an editing suite can be designed in a multitude of signal formats. The choice is based on the desired final product and budget. The analog choices are traditional NTSC composite or component signals in either RGB or a color difference form (such as Y,R-Y,B-Y or Betacam). Digital video choices are also divided into composite and component formats, with the CCIR 601 component format used by D-1 VTRs and the composite format used by D-2 and D-3 VTRs. The digital format signal can be carried in parallel or serial form.

Although the video signal formats may vary, the functional signal system flow of an edit suite remains essentially consistent (Fig. 22). The functional blocks shown consist of (1) sources (VTRs, still store, character generator, title camera, safe title, and test generators); (2) processing equipment (color correctors, time base correctors, framestores, and processing amplifiers); and (3) switching equipment (the central video effects switcher, the editor's preview switcher, and the monitoring strip switcher).

Previewing an edit may require the use of a preview switcher. An edit preview consists of viewing the record VTR's material up to the edit-in point, switching to the new source material from the effects switcher for the duration of the edit, and then switching back to the record VTR's source at the edit-out point. All of the equipment operates in exactly the same way as when an edit is being actually performed except that the record machine does not enter the record mode.

Depending on the editing equipment manufacturer,

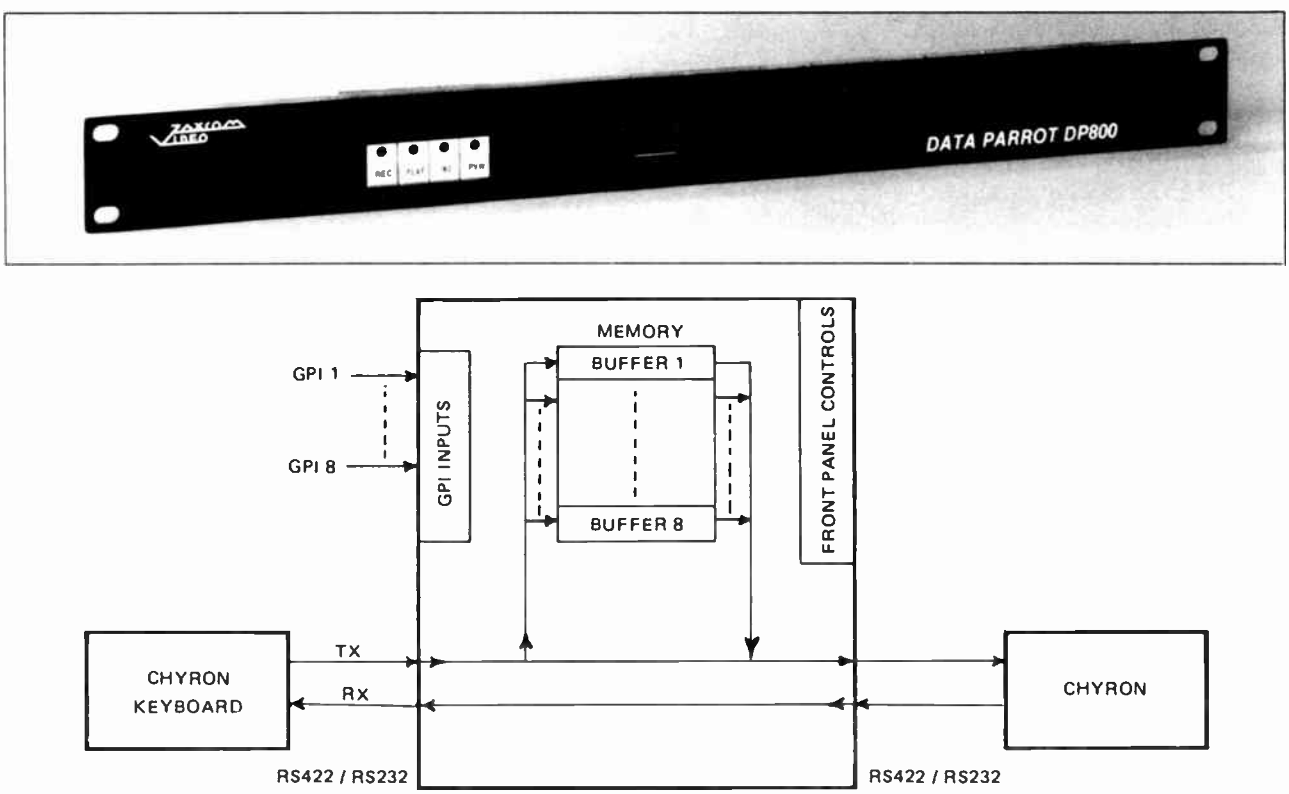


Figure 21. Keyboard emulator Data Parrot 800. (Courtesy Zaxcom)

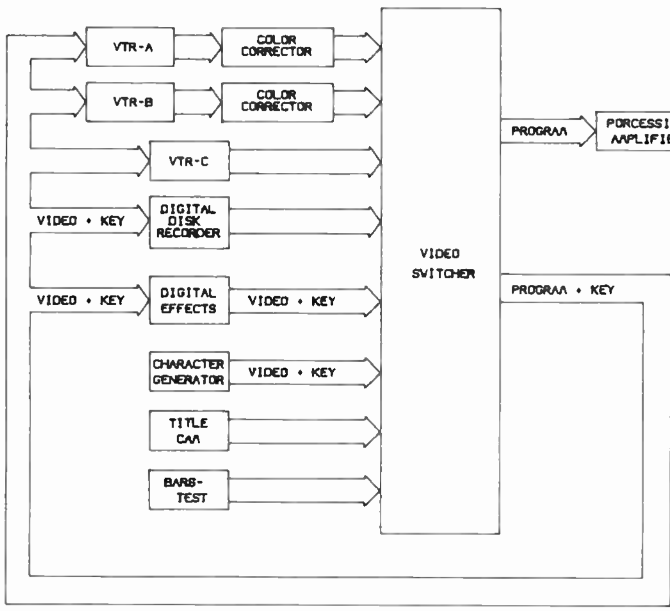
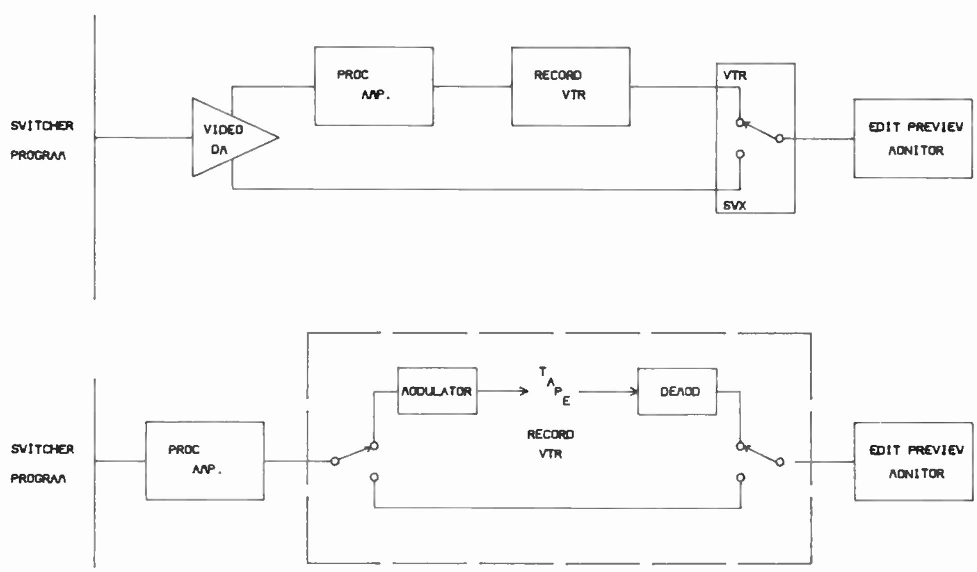


Figure 22. Edit suite general video functional.

one of two approaches will be used to accomplish this switching (Fig. 23). The upper drawing illustrates the use of a dedicated preview switcher, controlled by the editor, to switch between switcher program and record VTR at the edit points. The lower figure shows the alternate approach, using the record VTR's internal switcher to provide the preview switching.

Whichever method is chosen, it is imperative to ensure that the signal quality of the edit preview path is equal to that of the record path. In the past, preview system components often produced a lower quality video path that was tolerated since it did not affect the actual edit. The confusion caused to clients by the level and timing shifts of older preview systems was a



SWITCH POSTION	UP	DOWN	UP
MONITORED SOURCE	R-VTR PLAYBACK	— IN NEW EDIT MATERIAL OUT —	R-VTR PLAYBACK

Figure 23. Edit preview switching techniques.

constant thorn in the side of both the editorial and engineering departments.

A processing amplifier located downstream of the effects switcher reinserts blanking and provides clipping of excessive chroma and luminance levels. This reduces picture shifts by presenting the record VTR with consistent subcarrier-to-horizontal phase (SC/H) and blanking width.

A color corrector that can be placed in the output path of a source VTR provides a powerful tool for saving field tapes with poor color balance. Provisions must be made to advance easily the timing of the source VTR to compensate for its fixed delay.

Selection of an effects switcher for an edit suite is often based on counting inputs, key levels, wipes and other glitzy features without considering how well they can integrate into the video system. Beyond the fundamental task of mixing input sources, a flexible switcher can also resolve the problems of routing video and key signals to the inputs and from the outputs of a digital effects system (Fig. 24). Auxiliary and utility busses are used to provide technical monitor switching and device input selection.

Component Analog Video

A decision to build a component video editing system adds greatly to the complexity of the video and sync systems. This additional investment in time and money is usually prompted by a desire to reduce the generational losses and artifacts encountered in composite analog formats.

The difficulty of working in the component world with three times the number of video paths per source is readily apparent. Additional complexity is added by the variety of analog component formats. A common and upsetting discovery when building a component system is that rarely will all the functional blocks be of exactly the same format.

In RGB component formats, sync can be found on the green channel or as a separate signal. This can pose quite a problem for system distribution. In order to build a compatible distribution network the designer to build a compatible distribution network the designer must standardize on one sync format. RGB with separate sync (RGBS) has the inherent disadvantage of requiring four-cable wiring and patchbays, but is a more common format. The RGB with sync on green

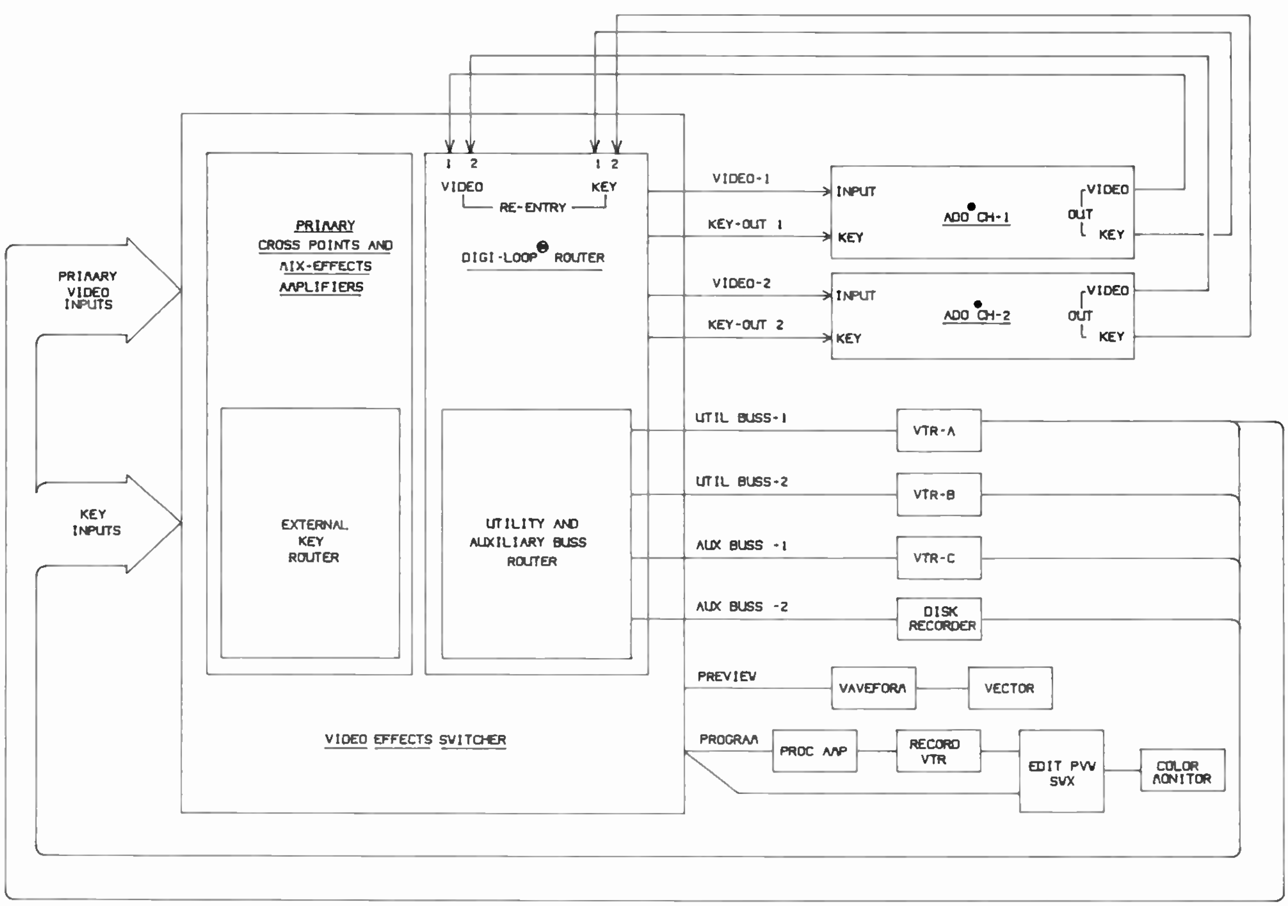


Figure 24. Edit suite video routing provided by switcher busses.

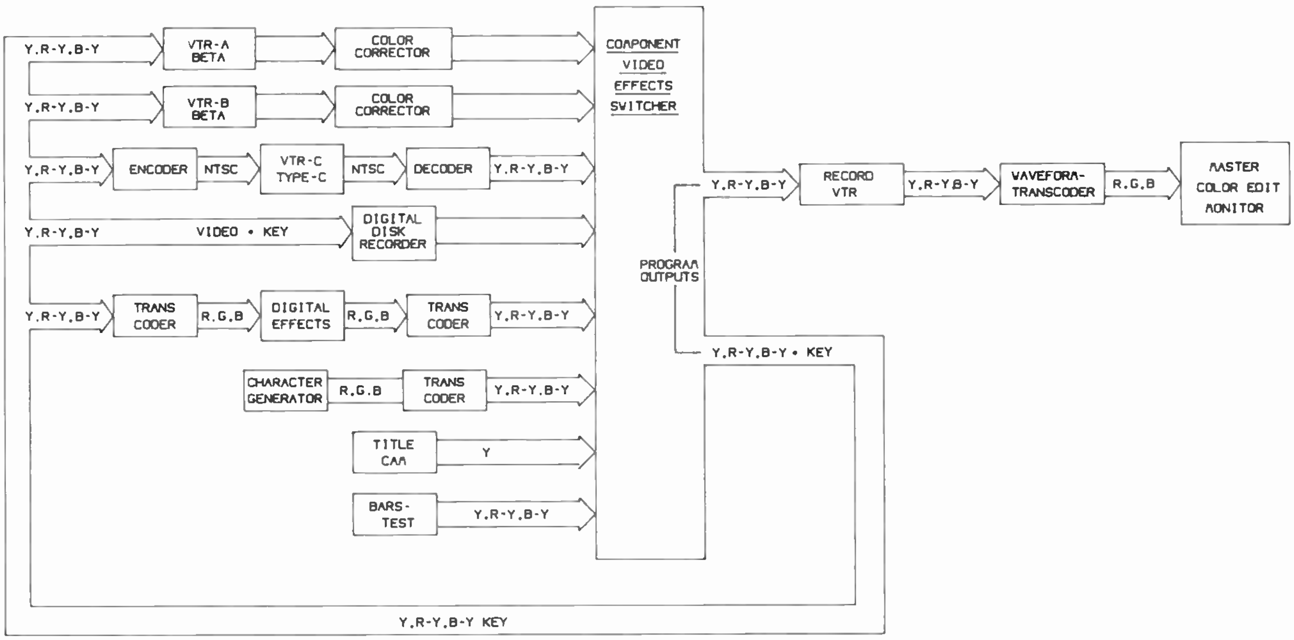


Figure 25. General component edit suite video functional.

format reduces the cable and patchbay jack count to sets of three, but may require more conversions than an RGBS system.

In color-difference systems, sync may be on luminance or separate, and setup is added to all three components. Fortunately, there are format transcoders and effects switchers that include multi-format inputs with jumpers for various sync conditions.

The functional drawing of an A/B roll component edit suite is shown in Fig. 25. The component format here is color-difference with setup and sync on green (as used in Betacam VTRs). The source and record VTRs are Betacam format but the character generator produces RGB outputs and therefore requires transcoding. A frame synchronizer with component outputs is included to provide a flexible method of importing composite sources. Other system devices that must be procured in their component version are: the color bar generator, monitoring-tech switcher and waveform, and color program monitor. In this example, the waveform monitor provides transcoded outputs to a standard RGB monitor. In a larger system, the digital video effects unit, distribution amplifiers, patchbays, routing switcher and edit preview switcher would also have to be three-channel devices; consequently, the complexity of the overall system grows very rapidly.

Digital Video

The all-digital edit suite can provide the multi-generation capabilities of which producers and graphic artists dream, but the price tag for this level of performance can be difficult to justify. The two most common forms of digital tape (19 mm) formats are composite digital (D2) and component digital (D1). The D2 format has entered the broadcaster's domain through its application in on-air cart machines. Component digital has been adopted readily by video graphics facilities, which need the multi-generation capabilities of digital recording to build complicated multi-layer graphics while avoiding the degradation typical of multi-generation analog video. The half-inch D3 format uses a video interface compatible with D2.

Once the source material is recorded on digital tape, the aim should be to stay in the digital domain throughout the editing processes. This means that the major signal processing components (effects switchers, digital video effects units, character generators, and still store) should be of the same digital format.

Both manufacturers and their customers give great consideration to the compatibility of major signal processing components. Unfortunately, compatible small-system components (such as edit preview switchers, routing switchers, test generators, color correctors, and color monitors) do not receive the same attention. The variety of models and configurations found in the analog-composite world has not been produced yet for the digital environment. In some situations, the only practical option may be to use analog equipment with the appropriate digital-to-analog and analog-to-digital convertors.

An interim step in the move to all-digital editing can be a mixed analog and digital composite configuration. One approach uses a digital composite switcher that offers internal transcoding to and from analog signals (Fig. 26). This approach provides flexibility by integrating analog and digital VTRs into one system, allowing the user to reap the advantages and economies

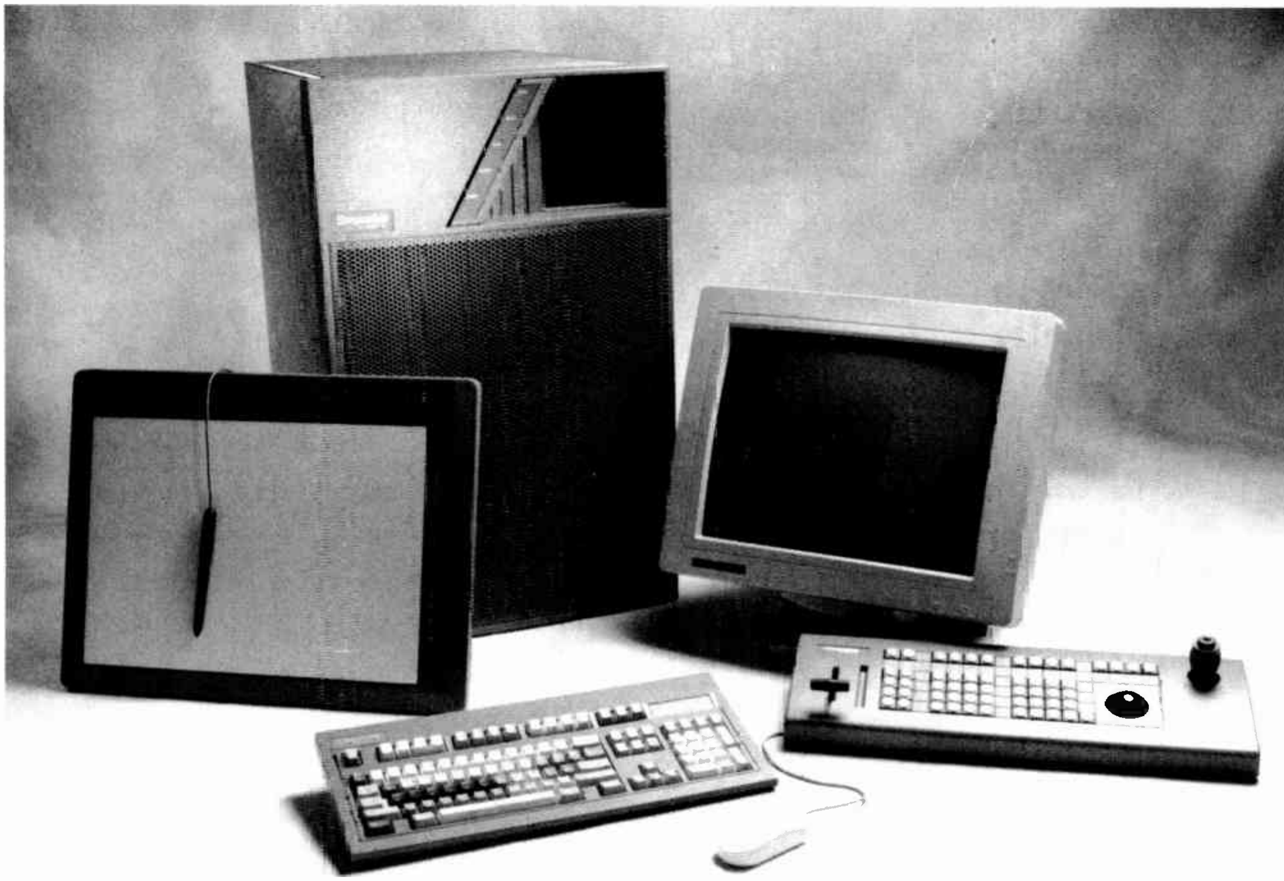


Figure 26. Integrated digital component suite. (Courtesy Digital F/X)

of each format. The drawback is degradation from successive analog-digital conversions.

An alternative to using discrete equipment for signal processing and editing in the digital domain is an integrated system such as the DFX-Composium (Fig. 27). The functions of an effects switcher, character generator, digital effects device, still store, paint system, and edit controller are implemented in a single unit.

In this system, all video and key processing occurs in a component-digital domain. Analog signals sources in either composite, RGB component, or Y/R-Y/B-Y component and 4:2:2 digital are converted to the internal 4:4:4:4 signal system, manipulated, and sent to the external world in any of the original formats. Using Microsoft Windows as the PC's user interface, the editor can control external digital disk recorders, multiple VTRs, and the internal effects switcher (Fig. 28).

Workstation Editing

The difficulties encountered in building and operating an edit system using discrete equipment have increased the demand for integrated systems at all application levels. A number of manufacturers have produced systems combining switcher, editor, and sometimes a character generator and digital video effects unit into one package. The user interface is often based on a modified version of the dedicated keyboard or control panel used on the discrete products.

The integration of these functions with a personal computer, with the user interface in one display, has given rise to the expression "desktop video." The Video F/X (Fig. 29) by Digital F/X takes an Apple MacIntosh II platform equipped with a frame buffer board and adds an external card cage containing a video switcher, audio mixer, and frame-accurate VTR controller. The user can import and export composite video using a proprietary scan conversion board. Graphics generated in Encapsulated PostScript format can be converted to fully anti-aliased NTSC composite and keyed or mixed over other sources.

The ramifications of this type of integrated system in the broadcast environment are difficult to define. Could a magazine format or news story be edited by nontechnical personnel? Possibly, but the future of desktop video systems will probably be similar to that of their hardcopy predecessors: desktop publishing

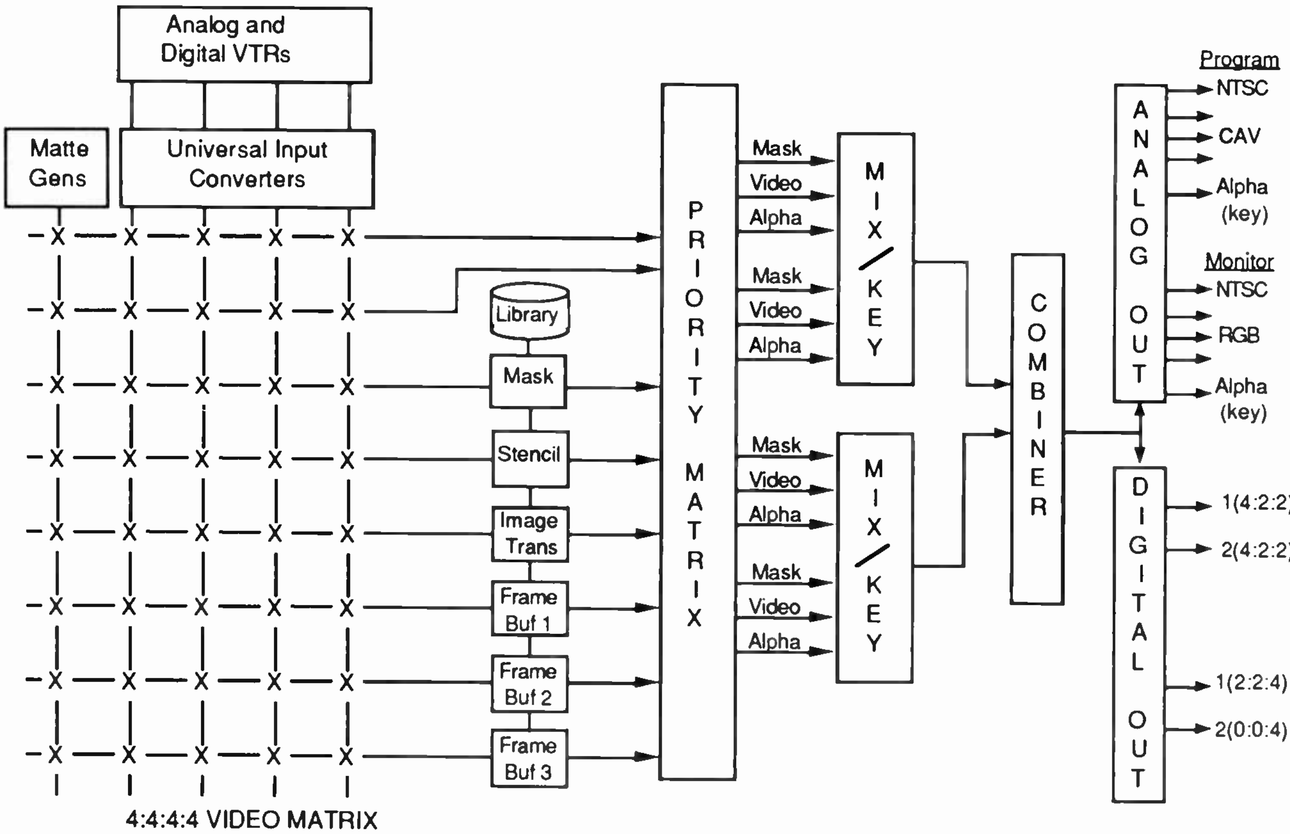


Figure 27. Composium video functional. (Courtesy Digital F/X)

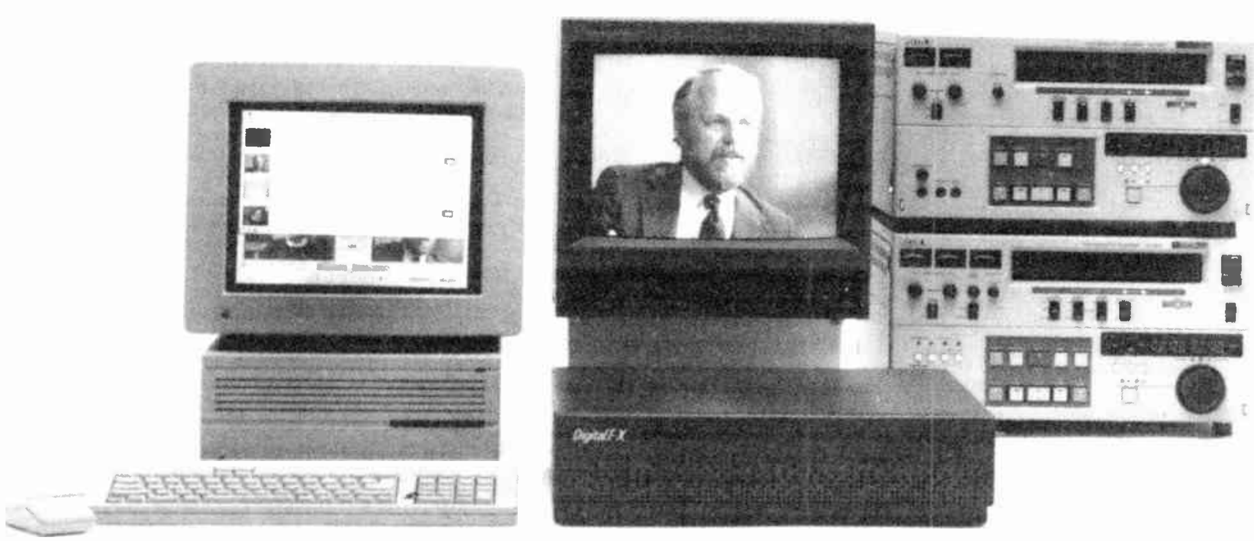


Figure 28. Desktop video production system. (Courtesy Digital F/X)

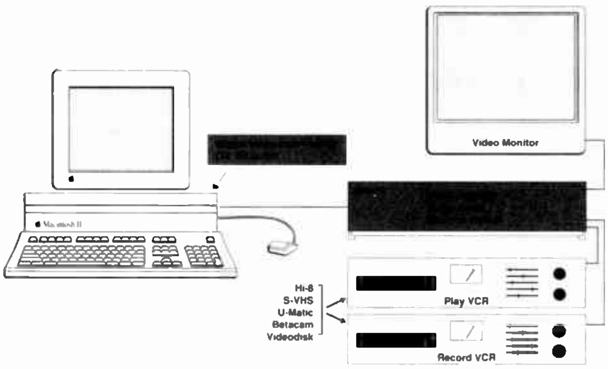


Figure 29. Video F/X system functional components. (Courtesy Digital F/X)

Figure 30. Advanced multi-channel audio mixer. (Courtesy Graham-Patten Systems)

systems. In the hands of motivated personnel, desktop publishing has become the off-line system of the printing industry.

Multichannel VTR Audio

As the newer generations of VTRs (Beta SP, D1, D2, and D3) move into the edit suite, the traditional two-channel-per-source mixdown configuration becomes too restrictive. The newer formats provide four channels of audio in addition to the address channel.

The "Hi-Fi" audio tracks are most often used only as sources because they are multiplexed into the video information and are therefore unavailable for audio-only editing. A second problem arising from multiplexing is the inability to monitor these tracks in shuttle or jog. Some users resort to dubbing the desired audio segments to the longitudinal tracks.

The extent of system changes necessary to accommodate the increase in available tracks depends on the frequency of use and type of application. An initial approach would be to make the additional tracks patchable and later, as demand increases, expand the number of board input channels.

Simultaneously previewing a four-channel audio edit may not be an everyday requirement. Buying a new preview switcher may be delayed if allowances are made to preselect the preview switcher's input. This configuration allows the operator to preview an edit using any two channels out of the available four. Providing this level of edit preview capability is often sufficient for many installations.

There are two ways to provide an existing system with simultaneous four-channel edit previews. If the existing audio mixer can be expanded to provide four-channel output, the addition of an external "add-on" four-channel preview switcher may be sufficient.

If the existing audio mixer and preview switcher must be replaced, consideration should be given to a system that integrates both functions (such as the Graham-Patten D-SAM in Fig. 30). This system addresses the multichannel expansion with digital-audio building blocks providing a multichannel preview switcher and monitoring system.

EDIT SUITE PHYSICAL CONSIDERATIONS

The physical design of an edit suite is a matter of adapting the best ideas available from other facilities and adding your own. The variables are so numerous that each facility (above the ENG application level) is a custom design fitted to the specific combination of people, product, and equipment.

The following collection of comments, concerns, and advice collected from editors, engineers, and clients may be helpful.

Client Access—The client can enter and exit the suite with minimal disruption to the technical operation (Fig. 31). Long hours of editing can be tolerated easier by everyone if the client can step out during a tedious technical problem without climbing over technical personnel.

Operator Access—The editor and his assistant can enter the machine room without disturbing the clients.

Maintenance Access—In case of failure, the maintenance technician can patch in auxiliary equipment without disturbing the client. Allowance should be made for routine maintenance of unassigned equipment without the necessity of disturbing the operators or clients.

Client Table—The placement of the client table should provide a comfortable workspace. It should strike a balance between the client being oppressively close to the operator and being a spectator in a stadium observing the action down on the field. The table itself should provide sufficient space for writing and for the mandatory piles of tape and have lighting that does not interfere with program monitors.

Telephones—Attention must be given to the need for privacy when locating phones for client and editor. It may not be desirable for the client to overhear the editor calling tech support.

Master Color Monitor—Life in the edit suite is greatly simplified with only one color monitor viewable from

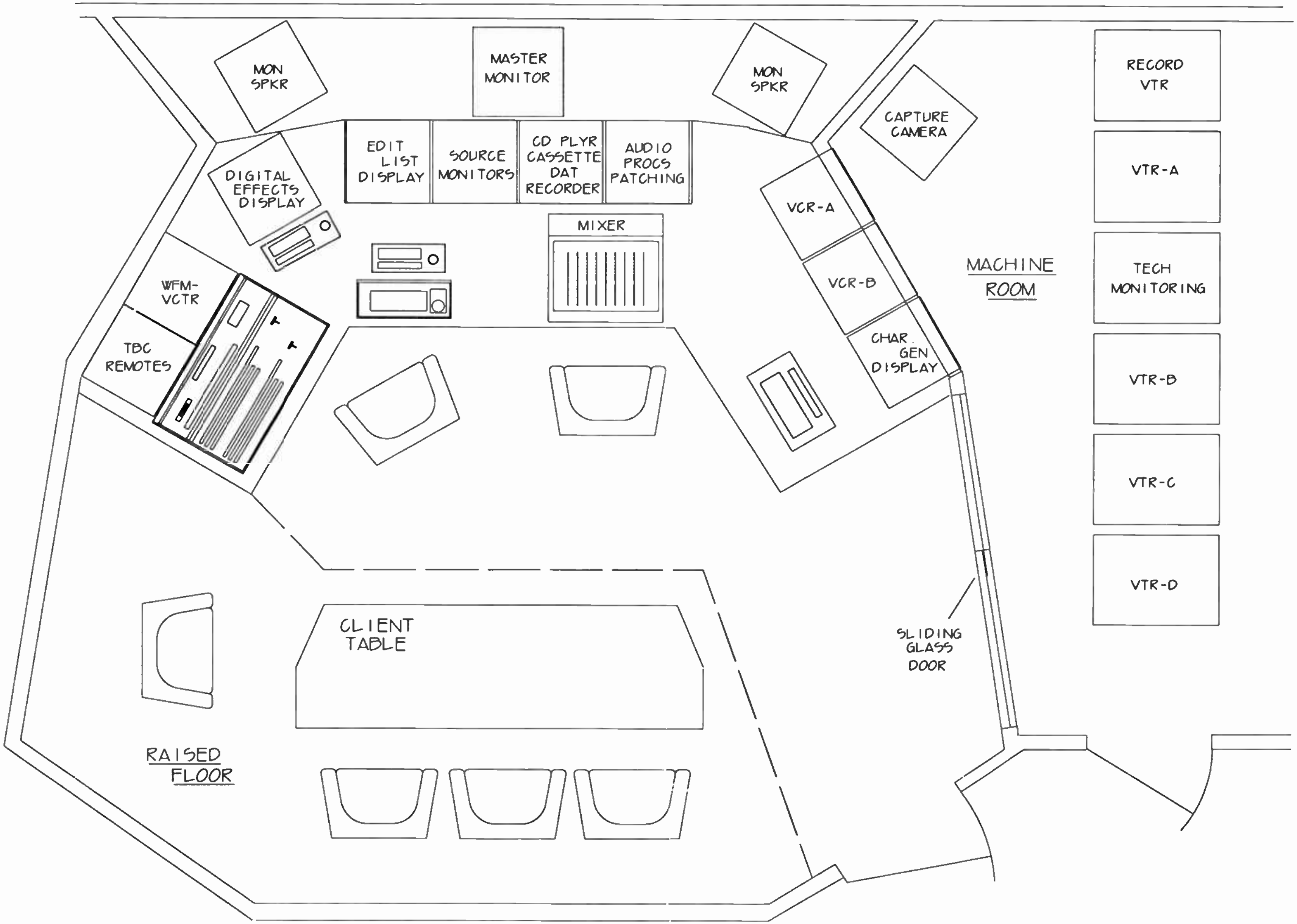


Figure 31. Floorplan of an on-line edit suite.

the client's position. In order to avoid confusion about which monitor is correct, any additional color monitors for color CG preview, etc. should be off-axis from the master monitor and, if possible blocked by the operator's body position.

Keyboard Placement—Switcher control panels, audio consoles, remote controls, etc., should be located so that typical operation doesn't place two simultaneously-needed controls out of reach. Height, width, and depth of cabinetry must be designed to allow comfortable, prolonged operation.

View Path—Placement of equipment and console heights should allow a parallax-free view of the program monitors (to the extent possible) for both operators and clients by minimizing the angle away from in-line view that will occur as the operator moves between positions. This is accomplished by placing primary controls directly in front of or adjacent to an operator without using a horseshoe approach which requires excessive neck twisting. There is never enough space directly in front of an operator.

Acoustic Path—Ensure that the client can hear the audio well enough to make a judgement. It is very easy to optimize the audio monitoring for the editor's position and unintentionally forget the acoustical environment at client's location. At the same time this balance is being established, it is necessary to ensure that the client can communicate with the editor without having to compete with the program material.

Lighting—Proper lighting conditions during an editing session will provide minimal glare on monitors and console surfaces, clear readability of controls, work light for scripts, and walkway lighting for safety. Attention to color temperature produced by lighting fixtures and monitors will minimize contamination. Room color should be neutral and nonreflective.

Service Lighting—Provide auxiliary lighting fixtures, or a higher level for the existing room lights, to give sufficient illumination for equipment or general maintenance. The method chosen should not prevent a quick return to normal lighting conditions.

Frequency Response—Minimize control room conditions that produce frequency response nonlinearities. Many edit points are identified by a specific aural cue. Loss or exaggeration of these cues may cause mistakes in editing. These conditions may also interfere with quality control.

Stereo Imaging—Placement of monitor speakers, room obstructions, and personnel all affect the stereo image. Frequency response and reverberation time must be controlled in order to avoid distorting the image. It is best to draw upon the experience of experts who have successfully dealt with these issues in dedicated mixdown facilities.

Technical Environment

The physical reliability of an edit suite is clearly tied to a number of factors not readily observable in a floor plan, photograph, or rate card, but which are every bit as important. These factors comprise the technical environment.

Temperature Control—Provide sufficient and quiet cooling of equipment without freezing clients. Operating consoles often allow too little ventilation for the modern generation of high-density electronics.

Humidity Control—Include a humidifier/dehumidifier system to suppress static electricity in winter and to reduce excessive videotape transport adhesion in summer.

Power Quality—Editing suite power sources should receive the same general consideration as those in a computer facility. The likelihood of losing an edit list due to a computer crash caused by a power line spike can be reduced by using any of the many line filters now commonly available for the larger personal computer systems.

Static Discharge Control—Editing systems, like computer systems in general, are prone to crashes induced by static discharge. Provide a safe discharge path for the static charge that builds up on personnel. Humidification, conductive floor treatments, and anti-static sprays are all good steps in reducing the likelihood of system crash due to static discharge. Combining these measures with operator static-awareness training can significantly improve system reliability.

Contamination Control—The contamination problems in an edit suite are the same as those in a studio control room. The difficulty is in diplomatically convincing some clients that, if dropouts are considered unacceptable, so too must be their behavioral causes, most notably smoking.

Equipment Centralization

As is often the case in broadcast facilities, many commercial post-production suites are built with a separate equipment room for VTR, digital effects, and terminal equipment. This technique offers the advantage of keeping the noise and heat generated by such equipment away from the operating environment. In case of equipment failure, patching over to alternate equipment service occurs with minimal alarm to the client.

A compromise between full centralization and dedicated equipment is usually the most practical approach. The need to have Betacam and U-matic VTRs in the suite is due to the short record time of field recorders. The typical 20 minute tape load triples the number of tapes handled compared to Type C. Fortunately, the reduction in size and power provided by interformat editing allows this type of equipment to be placed in the suite.

This problem can rarely be permanently resolved due to the continual increase in the power of technology that can be purchased per dollar. A typical situation is that the falling price of technology may allow an acceptable character generator to be purchased for dedicated use by each edit suite. Conversely, similar advances in technology may advance the state of the art in character generator capabilities. The new features provided by these advances will soon classify the dedicated CG as obsolete.

5.8 Film for Television

Richard W. Bauer
Eastman Kodak, Rochester, New York

INTRODUCTION

"Film for TV" holds many meanings for broadcasters, video engineers, and others in the motion picture and television industries. Most kinescopes hold fond but also horrendous memories. In the early days, a "kine" was flown from one coast to the other so that full U.S. broadcast coverage of the program material could be achieved. The photography of a cathode ray tube image (CRT) on black and white film, rushed through a process and immediately shipped to the far destination, with time always of the essence, made for a challenging use of film in television. The quality level of all components (CRT, film processing, handling, and personnel) left much to be desired in those early days.

Until the development of the quadraplex videotape recorder in the late 1950's, film was the only recording medium for television productions. The use of kinescope recorders to record television images on film is now only of historical interest.

Film was the primary distribution medium for the delivery of syndicated television programs to the local broadcaster for many years. Camera-tube (photoconductive) telecines were a fixture in local broadcasting studios. Today, most syndicated television programs are delivered by videotape or satellite link in the United States. Some local broadcasters maintain a 16mm film library along with their vintage camera tube telecines.

Motion picture film, however, is used for more than the production of theatrical feature films (movies). It is used for the production of television programs and commercials. It is also used for the production of industrial and educational programs. Nearly all of this film production is transferred to videotape using a telecine.

The development of *production telecines* capable of transferring camera negative film with programmable color correction has supported the continued growth of this industry. Flying spot scanner (FSS) telecines made by Rank Cintel and charge coupled device (CCD) telecines made by Bosch Fernseh are used in this production capacity. There are other suppliers of these types of telecines but these are the most prevalently used telecines in the United States.

Internationally, film remains the preferred medium for the origination and distribution of syndicated television programming. Motion picture film in 35mm width, four perforation pull down, and 24 frames per second projection is the standard worldwide. It can be projected or transferred via telecine to any videotape format, or line scan. Broadcasters can transfer the film to videotape in the local television standard before broadcast.

For the past three decades, film has maintained its role as the primary production medium for prime time television programming. In fact, the percentage of prime time programs originated on film has remained fairly constant (around 80%). Today, many of these programs are transferred directly from the camera negative film and post-produced on videotape. Most national television commercials are also shot on film and post-produced by independent television post-production facilities.

Another significant interaction and use of theatrical feature films for video is a large and growing market for videocassettes. Films are transferred to videotape and then duplicated for sale and rental of videocassettes and videodiscs. This secondary market often accounts for larger revenues than the initial theatrical film releases.

Summarizing this introduction of film for television: film is very important to television and television is very important to film. This relationship is more than 40 years old and continues to grow and prosper. Future developments such as high definition television (HDTV) will further enhance the value of this use of film for television.

FILM CHARACTERISTICS

Color motion picture film has three photosensitive layers. Exposure and processing produce dye images in these layers which, when projected on a large screen or scanned in a telecine, produce the color pictures. Although film in the 16mm width was the most common format for distribution of television programs for many years, it has been supplanted by the 35mm wide film.

The 35mm width is the main format for large screen projection in the theaters *and* is also used for the production of prime-time television programs. This format is used extensively in the post-production houses and for the production of television commercials. Film, although an old media, continues to be improved for speed, grain, and sharpness by the manufacturers, and Eastman Kodak Company estimates another 10x improvement is still possible with silver halide technology. The quality of film images, versatility of production techniques, and worldwide standardization keep 35mm film an important and valuable tool for the production of images for television broadcast. New versions of both flying-spot type and CCD film scanners continue to be developed for direct broadcast or transfer of film to video. Television film scanners have interchangeable optical blocks enabling either 16mm or 35mm format to be reproduced in telecine equipment.

Professional motion pictures, most prime-time television programs, and television commercials are originated on color negative film. From these originals, prints can be made for broadcast and distribution on either 35mm or 16mm formats. Duplicate negatives can be made and large numbers of prints prepared for theatrical release. Telecines are capable of reproducing from either color positives or color negatives. A program originated on color negative can be transferred directly to tape, edited electronically, and then broadcast. The same negative can be edited and either broadcast from film or a transferred tape. There are many choices available when the program or commercial is originated on 35mm film.

To display motion pictures on large screens in the theater, the projector advances the film one frame at a time and a rotating shutter cuts off the light while the film is being moved. A claw or Geneva mechanism advances the film by engaging perforations along the edges. Film 35mm wide has perforations on both edges, but the 16mm material used to make prints has perforations on one edge only; the space on the other edge is taken up by the sound track.

In camera-tube telecines, the film advances intermittently one frame at a time by means of the perforations just as in theater projections. But in the film scanners, film movements are continuous and the perforations are used only to drive a sprocket that generates framing pulses.

The film sound track may be either optical or magnetic. Telecine projectors and film scanners usually have interchangeable optical/magnetic sound playback heads. In the playback of an optical sound track, the lens imaging the exciter lamp filament at the film plane must be sharply focused on the emulsion side. Usually an adjustment is provided for this purpose, as the emulsion side in 16mm prints may be either towards the lens (preferred position) or towards the light source.

The standard motion picture frame rate around the world is 24 frames per second for both 35mm and 16mm formats. The 35mm format has 16 frames per foot; so at 24 frames per second, the rate of film movement in a projector or scanner is 90 feet per minute (fpm). With 40 frames per foot in 16mm format, the film passes through a projector or scanner at 36 fpm. The standard size of a 35mm picture frame is 0.825 inch wide by 0.600 inch high, and standard 16mm frame size is 0.380 inch wide by 0.284 inch high.

There are other film formats in use, but 35mm and 16mm wide films are the most common. The "super" formats, "super 16" and "super 35," provide larger picture areas.

When motion picture prints made for theatrical projections were a primary source of program materials for broadcasting, significant quality losses occurred. This was due to the inability of the television system to transmit the range of gray scale value needed to create acceptable pictures on motion picture screens. Committees of the Society of Motion Picture and Television Engineers (SMPTE) addressed this problem and published a report recommending the light density range in films for television be limited by adjustments in staging and lighting, rather than in the making of color prints. In this way, prints could be produced with a range of density values that the television system at that time could reproduce.

For creative and practical reasons, it was not always possible to adhere to the recommended lighting ratio of 2:1; nor to the recommended 60% reflectivity of the lightest scene element. Market requirements that made necessary the production of motion pictures for both purposes (theater projection and television display) have since changed, but at the time were extremely important.

The problem was addressed by the Eastman Kodak Company with the development of a low contrast print film. This film was designed with reduced upper scale contrast compared with the print film for projection on theater screens. The transfer characteristic of the low contrast print film was, of necessity, a compromise among several considerations. The contrast had to be as low as possible for good telecine performance, yet high enough to provide acceptable screen images in direct projection in review rooms. Also, the color saturation had to be high enough to maintain a chroma gain at a fairly low level. When a film with a contrast range of 160:1 is being squeezed into the television system, the shadow areas in the pictures are compressed and much of the shadow detail is lost. Some improvement was achieved electronically through the use of black stretch circuitry in the telecine, and further improvement was made by using low contrast print film for the television prints.

There were significant cooperative research and development programs among the telecine manufacturers and Eastman Kodak Company. Continued improvements in all elements of the film system, from the negative through the intermediates to the final print, contributed to major quality improvements in the broadcast of film on television. The design of circuitry to allow the use of negative working films for either transfer or broadcast also led to major quality improvements as well as economic gains for the use of film in television.

FILM FOR TELEVISION

Many programs are still originating on film; though most broadcasters today do not see the film itself, but rather a videotape transfer. Nearly 80% of the prime time programming on U.S. television originates on film whether the program is a dramatic or comedy series, made-for-television movie, or theatrical motion picture. A large percentage of commercials also originate on film but are transferred to tape for distribution and broadcast.

There are many reasons why film remains the medium of choice for origination. One of the main reasons is an undefined phenomenon called the "film look." This characteristic has defied quantification by performance parameters but continues to be a major consideration. Another advantage of film origination is the standard format worldwide. Programs originated on film today can be readily syndicated for distribution in today's 525 or 625 line standards or any of tomorrow's HDTV systems.

With today's family of color negative films, the cinematographer has the flexibility to choose high speed emulsions or slower, finer grain emulsions simply by changing the camera magazine. A comparison of film characteristics and video camera characteristics indicates that film still has the advantage for sensitivity and speed, tone scale reproduction, and resolution.

Color negative films with exposure index (EI) ratings of 500 are one to two stops faster than high-quality studio video cameras. In evaluating the gray scale transfer function of film and video, it is found that while film has an exposure latitude of about 10 stops with gamma correction, a video camera has a practical limitation of about six stops.

Although motion picture film offers significant advantages as a medium for the production of TV programming, videotape offers advantages for the handling and broadcasting in the studio. Films transferred to videotape can be broadcast and handled many times without losses due to dirt and scratches that may occur with repeated handling of film.

Many feature films are transferred to videotape and duplicated for distribution to the large and growing markets for the sale and rental of videocassettes and videodiscs. This secondary market often accounts for larger revenues than the initial theatrical film releases.

The Broadcast Telecine

Traditionally, telecines were used for the transfer and direct broadcast of syndicated television programs and feature films. Only print films were transferred, generally without supervision and no scene-to-scene color correction other than that provided by automatic gain or automatic black level circuitry.

The original broadcast telecines evolved from television camera technology combined with motion picture projectors. During the 1970's, flying spot and CCD telecines were developed for broadcast use. In the United States, telecines are rarely used for on-air broadcast, but some broadcast facilities maintain a telecine to transfer print film from their libraries to videotape for local broadcast.

The Transfer Telecine

The transfer telecine is used to transfer feature films to videotape masters for subsequent duplication to videocassettes and videodisc formats as well as for broadcast. Flying spot and CCD telecines are used by film-to-tape transfer houses in this market. A skilled colorist provides scene-to-scene color correction, using programmable color correction controls. The transfer telecine is designed to accommodate negative film, print film, low contrast print film, duplicate negative, or interpositive film.

Some television programs are edited on film and transferred to videotape in a similar manner with scene-to-scene color correction.

The Production Telecine

The production telecine is used in the production of television programs and commercials. Post production facilities use flying spot and CCD production telecines. Selected camera shots ("circle takes") are transferred from the camera negative film to videotape with a colorist providing scene-to-scene color correction.

Commercials may be edited directly (on-line), but television programs are typically edited off-line. Dubs are made to videodisc or videocassette formats for off-line editing. The edit decision list (EDL) generated in the off-line edit session is then used to assemble a finished videotape master in an on-line editing session.

For internationally syndicated television programs, a PAL master is made. Video standards conversion from the NTSC master has proven inadequate, due to compromised spatial resolution and motion judder. The preferred approach is to cut the negative film to a cut list generated from the edit decision list that was produced in the off-line edit session. The cut negative film is then transferred directly to a PAL master with scene-to-scene color correction. The recently introduced machine-readable edge numbers (Keykode™) will facilitate the process of negative cutting.

TELECINE DESIGN

Basic Principles

The basic function of a telecine is to convert an optical image on motion picture film to a television signal. This conversion involves an opto-electronic transducer and a scanning operation. The resulting video signal must be color matrixed, gamma corrected, and frame rate converted.

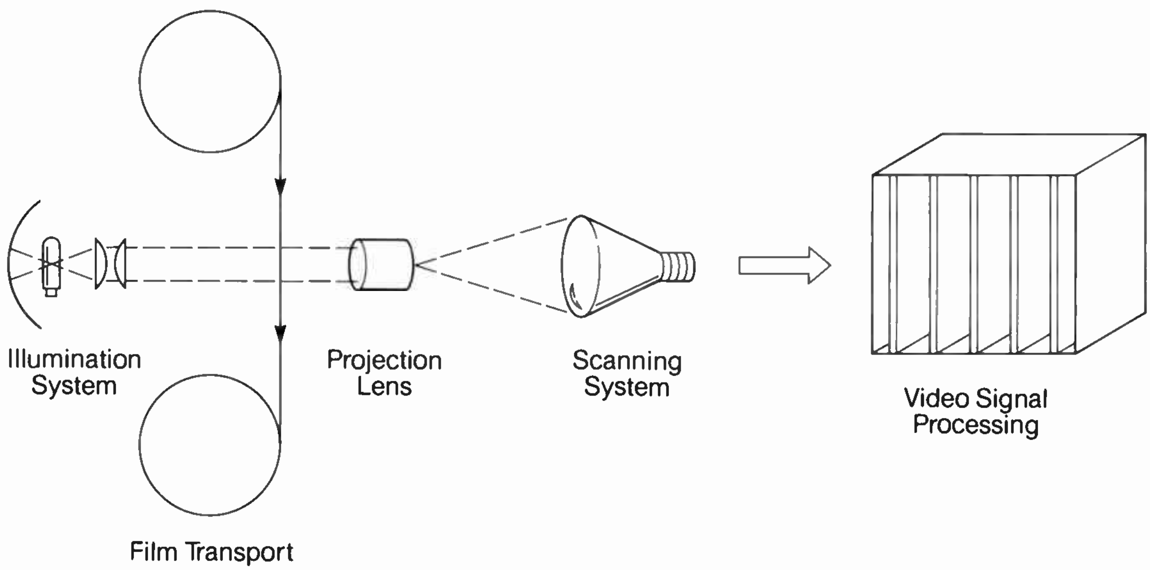


Figure 1. Generic telecine components.

There are three basic telecine designs that are commercially available: (1) photoconductive (camera-tube), (2) cathode ray tube flying spot scanner (FSS), and (3) charge coupled device (CCD) line array. An experimental laser telecine for HDTV has also been demonstrated by NAC.

Fig. 1 illustrates the major components of a generic telecine design. These components include: (1) film transport, (2) illumination system, (3) projection lens, (4) scanning system, and (5) video signal processing. Implementation of these basic components depends on the telecine technology and design.

Photoconductive Telecine

The photoconductive (or camera-tube) telecine represents the original telecine design. This design involves the combination of a synchronized motion picture projector with a television camera. As the design evolved, specialized video signal processing circuitry was developed to improve the image quality of the transferred video images. Photoconductive telecines were once sold by RCA, General Electric, Marconi, and others. Ikegami currently manufactures a photoconductive telecine for broadcast applications.

A simplified block diagram for a photoconductive telecine is illustrated in Fig. 2. The film projector head includes an intermittent transport, lamphouse, and relay lens. The intermittent film transport is designed to rapidly advance the film to the next frame while the lamp is shuttered and to project the film frame while the film is stationary. In the telecine application, the film advance and shutting is synchronized with the vertical blanking interval of the house sync generator.

The intermittent movement is either a pull-down claw (16mm) or Geneva sprocket movement (35mm). Frame rate conversion (from 24 frames per second film to 30 frame per second video) is accomplished by using a "3:2 pulldown" sequence. Each successive frame is pulled down and held for three video fields for the first frame and two video fields for the next frame, then three video fields, and so on. In this process four frames of 24 frames per second film are converted to five frames of 30 frames per second video.

The camera head is essentially a modified studio camera including beam splitter and three vidicon camera tubes. The film image is projected onto the photosensitive oxide of the beam-scanned photoconductive pickup tube, hence the name. The output signal is buffered by a preamplifier and processed by the video signal processing circuitry.

The first step is to *clamp* the signal and provide black level adjustment and color balance. The reference voltage for the clamp is varied, permitting dark current noise to be clamped out or "buried."

Flare correction compensates for the stray illumination (flare) due to reflections in the optical path or from the camera tube surfaces. The signal is integrated over the entire field and fed back to the black level control.

Shading correction compensates for uneven illumination, optical losses, and uneven sensitivity of the pickup tubes. Shading correction is implemented by adding or multiplying a combination of sawtooth and parabolic signals at both horizontal and vertical rates.

Since the *multiplicative gain control* is implemented by a four-quandrant multiplier. It can be controlled remotely, providing a gain control for setting white level and color balance. For live broadcast applications, a servo circuit for automatic white balance was developed. Typically, the green channel signal is held at a fixed level and the red and blue channel gains are varied to achieve matched signal levels.

The *gamma correction* circuitry corrects the video signal for the gamma characteristics of the pickup tube (1.0 for lead oxide, 0.7 for vidicon), the film (1.5 for

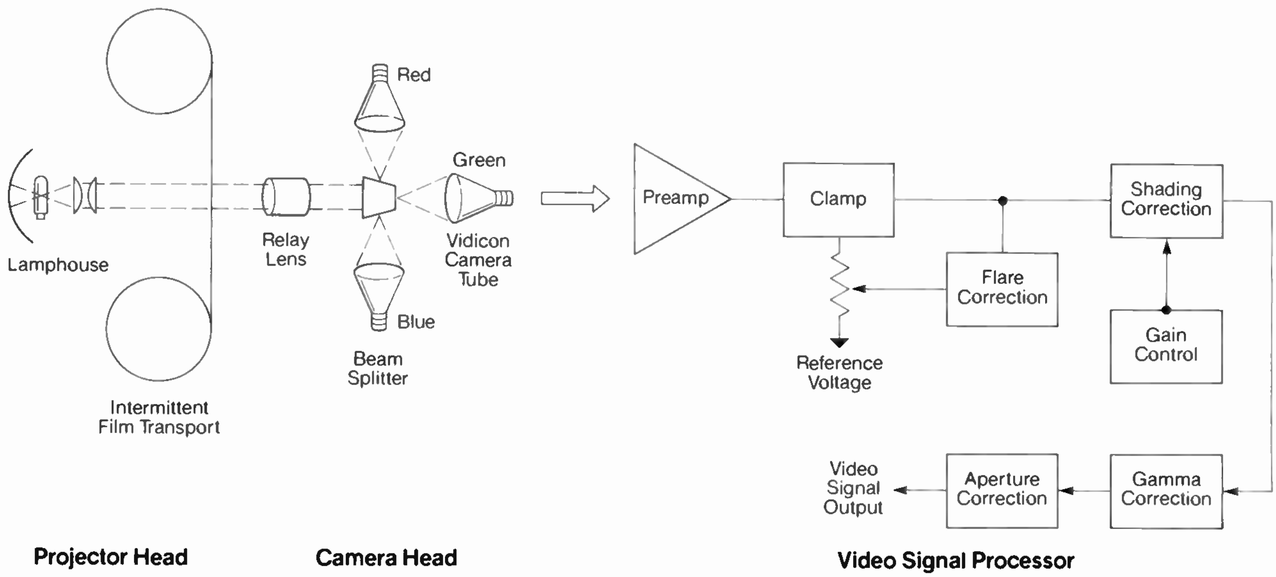


Figure 2. Block diagram for photoconductive telecine.

print film), and the display monitor (2.2). It also permits independent correction of each color channel to match the gammas of the film dye records. Studies have shown that the preferred gamma of a film transfer from a photoconductive telecine is 1.2 to 1.5. Practically, gamma correction circuitry must accommodate a range from about 0.2 to 0.6 for unity gamma pickup tubes and 0.3 to 0.8 for vidicon tubes.

Aperture correction compensates for the optical high frequency roll-off characteristics of the lens and scanning aperture of the pickup tube. Both *out-of-band* and *in-band* (or contouring) correction can be applied. Out-of-band correction is so named because the high-pass filter has a maximum response above the video passband.

The advantages of photoconductive telecines include simplicity of design and automatic operation for live broadcast of feature films. The primary disadvantage is image quality limitations due to flare, lag, and color misregistration artifacts. Also, the intermittent film advance, while adequate for print film, was never considered an acceptable transport system for color negative film.

CRT Flying Spot Scanner

The CRT FSS produces the video signal by scanning the film images with a very small spot of light and collecting the light transmitted through the film with a photomultiplier tube (PMT). A high intensity CRT is scanned by an unblanked electron beam.

The CRT flying spot scanner was developed by Rank Cintel and originally designed for 25 frames per second film transfer to 25 frames per second European television standards (PAL and SECAM). Early attempts at 30 frames per second NTSC designs involved a complicated "jump-scan" approach, controlling the scan to implement both interlace and 3:2 frame rate conversion. The development of the "Digiscan" frame store, which permitted the film frame to be progressively scanned and then interlaced, and frame-rate converted by controlling the output (read) rate, made the flying spot scanner design practical for NTSC. Today, the Rank Cintel FSS is the most commonly used telecine for production and transfer applications.

One of the fundamental innovations of the flying spot scanner design was the development of a continuous motion, capstan-driven film transport. The velocity of the film is monitored by a shaft encoder on a free-running timing sprocket that tracks the film perforations, and the timing pulses are used to control the capstan velocity via a servo loop. This transport has proven to be gentle enough to handle negative and intermediate film stocks in addition to print film.

One of the fundamental challenges in the CRT FSS design is the selection of a high intensity, short persistence phosphor with a broad spectral energy distribution for scanning color film dyes. The typical phosphor has little blue output, limiting the signal-to-noise ratio of the blue record of transfers from negative film, as negative film contains yellow masking couplers (0.90 D).

The basic block diagram for a CRT FSS is shown in Fig. 3. The scanning spot is divided into three color channels by a dichroic beam splitter where each signal is picked up by a PMT. The signal from each PMT is buffered by the head amplifier and applied to the afterglow correction circuitry. Afterglow correction is a high-pass filtering operation that compensates for the persistence (afterglow) of the phosphor.

The next step is shading correction, which serves the same function as that described in the photoconductive

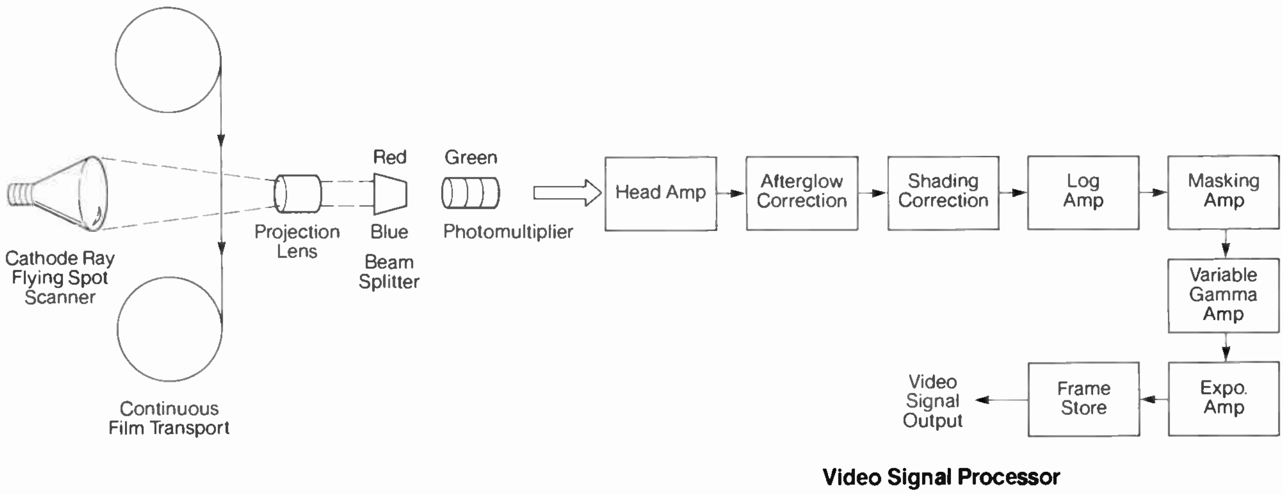


Figure 3. Block diagram of cathode ray tube flying spot scanner.

telecine design. Here, shading correction compensates for nonuniformity of the CRT and lens system.

The video signal processing is implemented in four steps: (1) log amp, (2) color masking, (3) variable gamma, and (4) expo amp. Color masking and gamma correction are implemented on logarithmic signals, and the resulting signal is exponentiated for display. The color masking operation is implemented by a resistor bridge matrix, with selectable options for different film stocks. Electronic color masking compensates for the "unwanted absorption," or cross-talk between the film dye records and the spectral response of the telecine.

The variable gamma function is implemented as a gain adjustment on the log signal. Black level (or lift) correction is applied by controlling the reference pulse to the log amp, and white level (gain) correction is applied at the expo amp.

A digital framestore is used to provide both interlace conversion and frame rate conversion (24 frames per second to 30 frames per second). This is accomplished by independent control of the input (write) clocks and the output (read) clocks. Aperture correction is also implemented digitally.

Recent advances in CRT flying spot scanners include an all digital video signal processing channel in the Rank Cintel URSA telecine. Also of note is the development of a pin-registered gate for image compositing (Steadi-Film) and the development of an electronic pin registration (EPR) system for real-time steadiness correction (Encore Video).

The primary advantage of the CRT flying spot scanner is the scan flexibility: zoom, pan, and anamorphic expansion are easily implemented by changing the scanning raster. The continuous motion transport handles film gently, making the transfer of camera negative film viable. Also, the lag and color misregistration artifacts of photoconductive telecines are eliminated.

The limitation of the CRT design is primarily the short life (2,000 to 5,000 hours) of the tube before it must be replaced due to phosphor burn or spot size deterioration. While not a real problem with NTSC or PAL transfer, prototype CRT HDTV telecines exhibit limited sharpness and signal-to-noise performance (particularly when scanning negative film).

CCD Line Array

The charge coupled device (CCD) line array telecine was first introduced in the early 1980's. As its name implies, the heart of the system is a CCD line array imager which converts the optical image to a video signal by transferring the charge accumulated in each photosite of the line array through a charge coupled output register. CCD telecines are manufactured by Rank Cintel, Bosch, and Marconi and are used in production and broadcast applications. The CCD telecine design also utilizes the digital frame store and continuous motion transport first implemented in the CRT flying spot scanner.

The illumination system is a high energy tungsten halogen lamp. While more than adequate for print film, the limited blue output of this lamp limits the blue channel signal-to-noise ratio obtained from negative film.

Sensor clocks are generated to control the integration time of each line and the pixel read-out rate. The commonly used Fairchild 1024 element photosensor has dual channel read-out with alternate samples interleaved.

Pattern correction removes any stripe patterns resulting from photosite sensitivity variations and output shift register mismatches.

The video signal processing in a CCD telecine is much the same as that of the CRT FSS telecine described earlier. A basic block diagram is illustrated in Fig. 4.

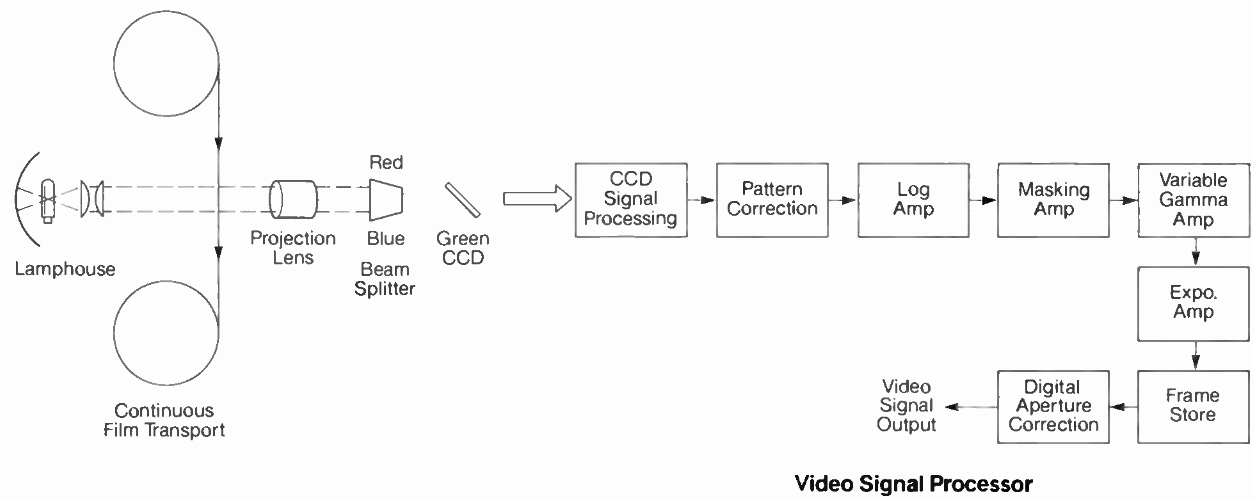


Figure 4. Block diagram of CCD line array telecine.

Recent advances include the implementation of an infrared-sensitive fourth channel in the Rank Cintel ADS-1 telecine to detect and conceal film scratches. Also, Bosch has developed an innovative Electronic Steadiness Optimization (ESO) system for the FDL-60 that scans the perforation and electronically corrects for vertical unsteadiness or "hop."

In 1989, Eastman Kodak demonstrated a prototype HDTV telecine using custom CCD imagers, digital signal processing, and a Xenon arc lamp and diffuse illumination system. The CCD imagers include a trilinear color sensor and a four quandrant detail sensor. Transfers from color negative and print film can be made without the limitations inherent in other HDTV telecine designs.

The principle advantage of the CCD telecine design is reliable operation and no need to replace camera tubes or CRTs. Automatic features have been implemented for broadcast applications. Fundamental disadvantages include scan inflexibility (zoom or pan must be implemented digitally) and limited still frame color-grading capability (a "freeze" frame must be used from the framestore).

FILM-TO-TAPE TRANSFER

Nearly all film shot for television display is eventually transferred, both images and sound tracks, to videotape. This is sometimes done just for post-production, but very often for broadcast and distribution. There are many reasons why film is still a widely used medium for origination for TV; including economics and logistics. But, there is also an important intangible which people call the "film look," which is a composite of the many physical characteristics of film. Dynamic range appears to be a very large contributor to the "film look" and is defined as the ability to record a long brightness range. That means there is tolerance for over- and under-exposure.

The introduction of the flying-spot telecine and the gentler film handling of the continuous motion transport contributed to the improved quality of film-to-tape transfers. Besides the transport improvements, the newer telecines incorporate better optics, significantly improved electronics, and the capability to transfer negative films. All of the new telecines, whether flying spot or CCD, use capstan drives, which are safer transports for film and incorporate control circuits for color and gamma correction. These improvements have opened new avenues for the creative producer to shoot on film and transfer to tape directly from the original negative.

Optically, the negative has all of the information that was in the original scene and it is a first generation image. There is some preference for printing the original negative onto an interpositive film stock and transferring to videotape from the interpositive (IP). An IP will generally retain the wide dynamic range of information recorded on the negative and it will also handle dirt and scratches better simply because it is a positive. Dirt and scratches on negative film result in white defects while those on a positive film transfer as dark defects. White artifacts or defects are more annoying and noticeable than the darker ones.

With interpositive film, however, a slight build-up of noise or grain occurs, but noise can be reduced during the transfer function. To the casual observer, the negative and IP transfers look very similar but a discerning viewer may have a preference for the subtle differences of one or the other.

Most of the film seen on television is transferred from prints. One reason is that is costs less to make a print than an interpositive. There is also the matter of familiarity. Most people have had a lifetime of experiences looking at movies and television in print form and frequently a print is the only material available.

Prints are one generation removed from the original. Any time you add a generation, whether film or videotape, there is an increase in grain or noise. Transfers made from prints may also show some loss of details in the highlights or shadows because of the inability of video systems to accommodate the wide density range of a print. These transfers tend to have a harder look with more saturated colors and higher contrast. Some of these factors can be corrected during the transfer process. For example, color reproduction can be controlled with masking and black level, and gamma controls can help to soften the look.

Two types of print film are used to make video transfers. There is the conventional print film, which is designed for optical projection, and there is a low contrast, or telecine-optimized, print film. Eastman Color LC print film 7/5380 is designed specifically for use with telecines. It is similar to regular print film except that it has reduced upper scale contrast which helps to make the blackest shadows light enough for reproduction by the telecine's limited sensitivity. The result is that instead of blocking-in, more of the details in the shadow areas will more faithfully reproduce on the video screen. This low contrast print film is fully process compatible with the normal contrast film so no extra costs are incurred for video. Many of the prime time network programs produced on film are now copied onto LC print film before they are transferred to videotape.

Many variables can influence the choice of film to use for the transfer to videotape. Transferring from the original negative provides the truest reproduction of the original scene as well as excellent tone-scale, color reproduction, sharpness and grain. It is also the most economical, and it saves time as no further printing is required.

Preparation of an interpositive is more expensive and time consuming, but most of the reproduction of the original is retained and excellent transfers can be obtained. This may be the choice also if a conformed IP is required for overseas theatrical release of the program.

If a slightly higher contrast is desired, with more color saturation but still good low light level detail, then a low contrast print might be chosen. A producer who wants snappy detail, higher contrast, and color saturation has the choice of the regular contrast print film as the transfer master. Some details may be lost in the dark shadows but this may fit the artistic preferences. By originating on film, the producers have many choices available to create just the look desired for any type of program.

TELECINE SET UP

A telecine is composed of several subsystems which work harmoniously to accomplish the task of scanning motion picture film and producing a video signal compatible with the television environment. Several film test targets have been developed which are designed to assist the set up and maintenance of telecine equipment. Through the use of film test targets, a telecine operator can ensure optimum operation and often diagnose system difficulties.

The telecine is a transducer which converts film to video format image records. Film test targets allow the operator a means to measure telecine resolution and signal processing characteristics. The telecine is also a creative tool, providing a television producer the means to alter video image characteristics such as color. Telecine control systems often offer the operator a myriad of controls and indicators with which to work when transferring motion picture film-to-video. Often it is useful to have the means to return telecine controls to a reference position to verify correct telecine performance and offer operators nominal starting points for a film to video transfer session. Typically a telecine transfer facility will maintain a selection of test films which facilitate the alignment of the telecine.

Telecine analysis film (TAF) is a target that is offered by the Eastman Kodak Company. TAF is an objective tool for initial setup and centering of the color grade controls on a telecine before the transfer of images from motion picture film to video. Kodak provides TAF as a benchmark and a tool for achieving optimum color timing control. After using TAF for telecine setup, the operator will be able to transfer a greater variety of scenes without major equipment adjustment. Kodak provides TAF so that film-to-video transfer facilities will achieve optimum image quality and satisfaction using Eastman (motion picture) film products.

TAF, which originates on Eastman color negative film, consists of an eight color bar test pattern, an eight step "neutral" gray scale, and a "neutral" gray surround. The color signals represent typical saturated colors that are encountered in motion picture production and can be used for telecine masking module adjustments should they be needed. Color reproduction of the test bars can be monitored by observing a color vectorscope connected to the telecine output video. Scene exposure information is encoded onto film in the form of film dye. The amount of dye formed on a minute area of film will be proportional to the amount of light energy which fell on this area at the time of scene exposure. Amounts of dye are typically measured by using a densitometer, which reports density values. TAF film gray scale includes typical film maximum and minimum density patches. These are useful in setting video white and black point controls on a telecine. Color channel imbalance can also be equalized easily by using TAF gray scale as an alignment tool. TAF is offered as original camera negative, print, and master positive materials in both 35mm and 16mm formats. TAF negative is particularly useful for setting up telecines for negative film transfers. Often, little telecine adjustment is needed to achieve good results after setting up a telecine on TAF.

The Society of Motion Picture and Television Engineers (SMPTE) offers several film test targets for telecine use. One of the more popular targets is the Television Alignment and Resolution Test Pattern. This target offers indicators to assist with scan sizing and position and resolution patterns of the calibrated wedge design which offer a range from 250 to 500 television lines. This target is widely used for critical focus adjustments for both optical and electronic subsystems of the telecine. A complete test target catalog can be obtained from SMPTE.

The European Independent Broadcasting Authority (IBA) also offers test films for television and recommended procedures for their use. The IBA provides targets which quantify telecine resolution, streaking, flare, geometry, and image steadiness.

MONITOR SETUP

The color picture monitor is the subjective reference for all film-to-video transfers. Although a video waveform monitor and vectorscope provide some objective signal levels, the color picture monitor is the basis for the artistic judgements of the colorist and the acceptance of the finished product by his client. Program interchange between facilities and consistent quality can only be assured by high standards for color picture monitor performance, alignment, and viewing conditions.

The SMPTE Working Group on Professional/Studio Monitors (T14.28) has worked to develop recommended practices and standards for color picture monitors. These include:

- Proposed Standard for Television—Professional 525-line Television/Studio Type A Color Picture Monitors—Performance (T14.280)
- Draft Recommended Practice for Critical Viewing Conditions of Color Television Pictures (T14.281)
- Draft Recommended Practice for Alignment of Professional Television/Studio Color Picture Monitors (T14.282)

Performance

Performance criteria for color picture monitors identified in the SMPTE standard include the following specifications:

- White point = 35 ft.L
- Color temperature = D65
- Colorimetry = SMPTE "C" phosphors
- Contrast ratio ≥ 50:1
- Gamma = 2.2 +0.4/-0.0
- Luminance nonuniformity ≤ 25%
- Color purity ≤ 6Δ E* CIELUV units
- Resolution ≥ 500 TV lines per picture height (without aperture correction)

Aperture correction should be switchable. Underscan capability is a must in order to see the full active video raster. Long- and short-term stability requirements are also specified in the SMPTE standard.

Alignment

A color picture monitor can be aligned with a test signal generator and a color analyzer (tristimulus device). The required test signals will be indicated in the following discussion of alignment procedures.

Scan size is adjusted in underscan mode so that all four corners of the raster are visible. Overscan mode should involve no more than 5%.

Geometry, *linearity*, and *aspect ratio* are adjusted using the cross-hatch test signal, turning on only the green beam, and adjusting pin-cushion and scan-linerity controls for visual alignment with a linearity overlay (ball-chart) over the CRT face.

Convergence is adjusted with the cross-hatch signal with all three beams on. The manufacturer's recommended adjustment sequence should be followed.

If *aperture correction* is used, the amount should be set to make the apparent brightness of the 2T $\sin^2$ pulse match that of the bar in the "pulse-and-bar" signal. Alternatively, the multiburst signal can be used to match the apparent contrast of the 4.2 MHz burst to that of the 3 MHz burst.

SMPTE color bars are used to set *chrominance amplitude* and *phase* for the decoder, viewing the blue channel only. The left blue bar is the reference bar and its brightness is not affected by the chroma or phase controls. The phase control affects the brightness of the inner two bars. The recommended adjustment sequence is to first adjust the chroma control and then the phase control so that all bars are of equal brightness (matching).

A 100 IRE window signal is used to set *color temperature* and *reference white*. First a visual comparator is used to adjust the RGB gain controls for a visual match to D65 color temperature. A photometer is used to set the contrast control for a reference white level of 35 ±3 ft.L.

Gray scale tracking is set using an unmodulated stairstep signal and adjusting the RGB screen controls to produce visual neutrals in the darker steps. As the screen controls and gain controls interact, this requires an iterative approach.

The last step is to use the PLUGE (Picture Line Up Generator) signal to set the *brightness* control so the darker patch just merges with the reference black level, but the brighter patch is clearly distinguishable at the normal viewing distance. The PLUGE signal of the SMPTE color bars can be used.

Viewing Conditions

The recommended viewing distance for the telecine colorist is four to six picture heights.

The monitor surround should be a neutral gray illuminated by D65 lighting in order to permit unbiased color judgements. A graded illumination is recommended to minimize fatigue, with a peak luminance of less than 3.5 ft.L (10% of reference white). A surround area of at least eight times the area of the color monitor is required. This can be obtained either by lighting a rear wall behind the monitor or by placing the monitor in a rear-illuminated panel.

All room lighting should be filtered to D65 color temperature and directed so that there are no spurious reflections from the monitor face.

PROGRAMMABLE COLOR CORRECTION

The introduction of programmable color correctors for telecines in the early 1980's was a significant factor in the development of telecine transfer houses providing negative transfer and video post-production. These techniques are now widely used for both commercial and program production.

At the heart of the programmable color corrector is a microprocessor with memory. The settings of the primary colorgrading controls of the telecine are encoded and stored. These controls include gain, lift, and gamma, and color balance in each. Colorists use a remote colorgrade control panel with joysticks, sliders, or track-ball controls. They roll the film to the start and stop frame of each scene and enter the corresponding frame counts. An optical shaft encoder driven by a timing sprocket monitors film frame count. The colorgrade settings for each scene are stored along with the framecount over which they apply. The film is transferred to videotape in one direct pass after each colorgrade setting has been programmed.

The programmable color corrector also stores control settings for zoom and pan. Aperture correction and noise reduction can also be programmed on a scene-to-scene basis. Secondary color correction is also built into many programmable color correctors. Secondary correction provides selectable hue alteration, utilizing from six to as many as 24 different color vectors. The term "secondary" refers to the processing of the telecine output signals.

Programmable color correctors are available from each telecine manufacturer. Other widely used color correctors include the Corporate Communications Sunburst and the Da Vinci units.

SOUND

In the film projectors used in camera-type telecines, the sound reproducer is located just below the picture gate. An optical system focuses a slit of light from an exciter lamp onto the film. Usually this lens is adjustable to accommodate films with the emulsion side either towards the light source of the projector or towards the lens. This is necessary only for 16mm format sound; 35mm is always in the same position with the emulsion towards the light.

A system of damping rollers smoothes out the movement of the film over the sound scanning drum following the intermittent frame-by-frame movement of the film in the picture gate. The sound in the film track is located 26 frames ahead (in the direction of film travel) of the corresponding picture frame for 16mm film, and 21 frames ahead for 35mm.

Variations in the light-transmitting properties of the sound track modulate the light beam as it passes through the film. The light falls on a photoreceptor, generating an output which, after suitable amplification, becomes the audio signal. The optical system in the sound reproducer must be sharply focused on the emulsion side of the film, where the sound track is located, for best high-frequency response. This is a critical adjustment because the focal length of the lens is quite short; a change of only 1/1,000 of an inch can cause noticeable high-frequency loss. For 16mm, the lens usually has only two positions: the back or the front of the film. On 35mm reproducers, the lens is fixed after factory set-up.

If the intensity of the scanning beam across the width of the sound track is not uniform, the reproduced sound may be distorted or low in level. Lack of uniformity is usually caused by dirt accumulating in the slit of the optical system. Incorrect positioning of the exciter lamp filament will also adversely affect sound quality. Films that are scratch-free and clean will also produce better sound quality.

Continuous film motion in flying-spot and CCD type scanners makes sound reproduction much easier, since there is no need for a system of damping rollers at the sound drum as in intermittent projectors used in camera-type telecines.

The Bosch FDL 60 CCD telecine and Rank flying-spot telecine scan optical and magnetic sound tracks directly at the capstan. A roller is mounted between the capstan and the picture gate to obtain the necessary picture/sound separation for 16mm and 35mm films.

Many 16mm films in circulation have magnetic sound tracks consisting of a stripe of iron oxide coated on the side of the film facing the projector lamp. The playback head is located 28 frames ahead of the corresponding picture frame. Projectors for television service usually have selectable optical and magnetic sound playback devices.

The azimuth of the magnetic playback head must be properly aligned. Ideally, the gap in the head should be positioned at exactly 90° to the direction of film movement. An incorrect azimuth setting results in loss of high frequency output, as does separation in the head-to-film contact. Periodic cleaning is essential to remove buildup of oxide that prevents good contact.

Today, most film is transferred to videotape for broadcast and it is easier to control the cleanliness, alignment, and pristine condition of the telecines' optical systems for both picture and sound in the transfer house.

In post-production and high quality film-to-video transfer operations, 35mm color negative film is the preferred format, instead of making prints from the negatives. Here the sound would be supplied on a separate high quality magnetic film which can be synchronized with the pictures. This way, the best television picture and sound can be obtained for broadcast or for mastering videocassettes for the home VCR market.

Significant quality improvements have been made in recent years in the sound amplifiers and speaker systems of home television sets. Better frequency response, signal-to-noise ratio, and even stereo sound now requires the transmission of clean, quality sound signals. This is possible whether the sound is transferred from film or tape, and if the usual precautions for dirt-free, scratch-free operations are followed.

FILM HANDLING, CLEANING, STORAGE

In the normal operation of a TV station, it is not easy to ensure that all films are spotlessly clean and free from physical defects. The methods of film handling adopted by station personnel and the degree of tolerance for image imperfections that broadcasters allow sometimes aggravate these conditions. Motion picture film will stand a considerable amount of abuse, but scratches and abrasions can cause a film to be rendered useless.

Films can be scratched in any piece of equipment where it is drawn over metal or plastic surfaces. A tiny nick in the gate of a telecine projector can put a severe scratch on the film from one end to the other. Constant attention is needed to make sure that imperfections are not being caused by faulty or dirty equipment. In this respect, film and tape handling requirements on their respective reproducers are very similar for cleanliness and meticulous maintenance of the mechanical parts.

Abrasions are usually caused by careless handling, such as excessive or erratic rewinding speed. Another very common cause of abrasions is the tightening of a loose-wound roll by pulling on the end of the film causing the convolutions to rub on each other. A particularly bad practice, often run, is to start winding at high speed and then allowing the film reels to coast.

With a little practice, the probable cause of a film scratch can be identified by noting the nature of the damage. A straight, uninterrupted scratch parallel to the film edges is most likely caused in a machine where the film moves continuously such as an FSS or CCD telecine. Camera-tube telecine projector scratches may show some discontinuities caused by the intermittent action of the claw and gate mechanism. Wavy scratches are usually caused by careless cleaning practices.

Torn or damaged perforations are usually caused by faulty projectors or damaged sprockets. Damaged perforations likely will give rise to unsteadiness in the television pictures. (Kodak Publication H-23, "The Book of Film Care" is an authoritative guide on all aspects of film handling. See especially pages 51–83).

Film Cleaning

When film and slides are reproduced in the television system, any particles of dust and dirt as well as scratches and abrasions on the film surfaces produce most unpleasant effects in the transmitted images. Electronic enhancement of the images invariably enhance the picture defects as well.

The large network centers rarely broadcast a scratched or dirty film, but local stations many times have no choice as they may receive prints that have been used many times. Few stations seem to have the time to clean every film prior to telecast, and replacements for defective prints are seldom available.

The preparation of film programs for on-air release or transfer to videotape usually involves much handling and rewinding, often at a high rate of speed. If the film's reels are bent or damaged, the film base and emulsion may be scraped off the edges of the film and trapped in the convolutions of the wound roll.

Static charges on the film surface, generated by high speed handling, can cause the particles of film support, film emulsion, or other airborne dust particles to cling securely to the film surface. Smoking in film handling areas should be forbidden as ashes and smoke residues can be deposited on floors, work surfaces and eventually on the films. Personnel traffic should be restricted in film cleaning and handling areas as dust and dirt particles emanate from clothing, shoes, and even hair. It is not necessary to create an antiseptically clean area to handle film, but clean room practices and common sense can prevent problems in film transfers or transmission.

Cleaning films to remove dust and dirt particles is neither difficult nor time consuming. There are several methods that have been in use for many years and recently Eastman Kodak Company introduced an effective and safe dry cleaning technique. The oldest, simplest, and most economical method is the use of plush pads and a cleaning fluid. In this method the film is placed on a rewinder and a plush pad is moistened with a cleaning solution (Kodak movie film cleaner with lubricant or one of several other commercial products). The pad is folded over the film strand and the film is slowly pulled through the moist pad and wound on a take-up reel. A fresh section of pad should be used for every few hundred feet of film, depending on the amount of dirt being removed. Care must be taken to keep the pad moist but not so wet that the film strand is wet when wound. If the film surfaces are wet as the roll is being wound up, drying marks may appear on the film surfaces. Cleaning pads can be made from pieces of plush velvet, 12 x 16 inches in size, folded over once and sewn around the sides with the edge turned inward. This is to prevent any plush strands from getting onto the film. These pads may be obtained from photographic supply stores.

Ultrasonic film cleaning machines are available to provide excellent cleaning, proper film handling, and high volume. These machines are available from several manufacturers and can be sized for various capacities. Both the ultrasonic cleaners and the plush pad cleaning technique utilize organic solvents. These solvents were especially selected to remove fingerprint grease, wax pencil, and other markings found on film. When used according to the manufacturer's recommendations, properly stored, and used with adequate ventilation, they are reasonably safe. New controls for use, storage and disposal of solvents and solvent vapors have made these methods of cleaning more expensive and less desirable.

Recently, Eastman Kodak Company introduced the particle transfer roller (PTR) as a proposed method for cleaning film. This is a polyurethane covered roller that is specially surface-finished to attain a tacky nature. By careful selection of a polyurethane for hardness and addition of additives to optimize electrostatic properties, and then carefully finishing the surface, a very efficient cleaning roller is obtained. Unlike other types of tacky-roller cleaners, the PTR is easily cleaned and rejuvenated merely by wiping the surface with a moist sponge to remove the dirt particles.

This new technology allows for installation of one or two PTR rollers on a telecine, projector, or rewinder. PTRs can be installed on either side of the telecine head and thus the film can be effectively cleaned when transported in either direction. (Further PTR availability and full descriptions of effectiveness may be obtained from Eastman Kodak Company.)

Film Splicing

Many different methods can be used to prepare films for on-air release or transfer to videotape. For example, a film program and all the short commercials, promos, etc., to fill a full half-hour broadcast period can be spliced together into a single, complete roll ready to run in a telecine projector or scanner.

Another method is to splice together all the commercials and promos on one reel, with short sections of leader between these items, and run this reel on one telecine projector or scanner while the main program is running on another. A complete on-air program or a videotape transfer can be produced by switching from one projector or scanner to the other.

Still another method is to transfer the film program and all the commercials, promos, etc., to videotape and package a complete program by electronic editing.

Whatever method is used in a station, some film splicing will always be needed. Each of the splices should be considered as a potential hazard in the telecasting of a program or in its transfer to tape. If a splice comes apart in the telecine, operating schedules will be disrupted, with accompanying losses of time and money.

Splicing film is not a difficult or complicated procedure. Good splices that will stand up to normal handling and repeated projections, running backwards or forwards if necessary, can be made in two different ways: either by joining with cement or by applying transparent adhesive tape to the two sections of film at the joint. (Kodak Publication H-23, "The Book of Film Care" deals with film splicing, pages 60–67.)

Cement Splices

When a cement splice is being made, the ends of the two pieces of film to be joined are overlapped in the splice area and the emulsion is removed from the lower section of film.

Then a little cement is applied to the scraped area, and the base side of the other piece of film is placed over it and pressure is applied until the cement dries.

Film cement is a solvent capable of dissolving the film base. Usually films are spliced in a small hand-operated machine, which has two hinged platens, registering pins to accurately position the two film ends by the perforations, a scraper for removing the emulsion in a narrow strip in the splice area, a knife to cut off the two ends neatly, and a hinged pressure block to hold the two ends in contact after the cement has been applied. The film cement is stored in a small bottle, with a tiny brush to apply the cement sparingly in a narrow area of the film at the splice. A good splice has sufficient strength after about 20 seconds to be removed from the splicer and wound up on a reel.

Tape Splices

Tape splices are easier to make, are less likely to come apart or break during projection than cement splices, and can be made on any film base. The tape used for making splices is clear, with an adhesive backing. Many different types of film splicing tape are available.

Tape splicing machines, readily available, make application of the tape and finishing of the splice very easy. An important advantage of tape splices in a television station is that no frames are lost when two films are being joined together. The film ends are simply butted in the space between frames and the splicing tape applied. Another big advantage is that tape splices can be taken apart easily later on, with no damage to the film frames, by simply peeling off the tape.

FILM AND THE FUTURE

High definition television has been under development for nearly two decades, and although worldwide standards are not established there is general agreement that HDTV will occur in the near future. HDTV has been used for limited program production in a few studios, and Japan began broadcasting an HDTV signal, called MUSE, in 1989. However, there is still significant disagreement over which signal formats are the best choices for production, distribution, and broadcasting.

Film has been, and continues to be, the worldwide standard for the origination and display of theatrical entertainment programs. For over thirty years it has been the preferred medium for the origination of prime time television programs.

Programs originated on film can be transferred to videotape in either of today's 525 or 625 line standards for distribution and sale worldwide. These same programs on film can be readily syndicated in tomorrow's high definition television formats, whatever they are. Video originated programs will be limited by the performance of today's line standards, or if produced with one of the proposed HDTV systems, will need to be converted to be used on the others.

There have been several studies made of the film and HDTV mediums for video program productions. The performance parameters compared were sensitivity and speed, gray scale transfer functions, contrast range, MTF and limiting resolution, noise and granularity, color reproduction, and artifacts. The results of one of the more recent studies was presented in a SMPTE paper by G. Kennel, L. DeMarsh, and J. Norris entitled "A Comparison of Color Negative Films and HDTV Cameras for Television Program Production." The conclusion reached was that 35mm motion picture film offers significant advantages as a medium for the production of HDTV programming, even if a single HDTV standard is chosen. Sensitivity, scene contrast range, practical and artistic-flexibility, speed, and worldwide standardization are all advantages of film origination for HDTV.

Film-Video Interface Definitions

The following definitions are not intended as literal television or motion picture definitions, but are meant to afford a mutual understanding of terms.

Motion Picture	Television
Color	Chroma
Overall printer color balance	Chroma Phase
Negative or positive scratches	Dropouts
Scratch removal process	Dropout compensator
Release print	Dub
Edited negative	Master
Dupe negative, internegative	Submaster
Poor printer contact, printer slippage	Banding
Fine-grain negative	High band
Projector and screen	Picture monitor
Film printer	Video recording head
Sound recorder	Audio recording head
Printer & process control system (for proper color balance)	Vectroscope
Method used to establish gamma and density	Waveform monitor
Motion picture projector	Film chain
Projector changeover	Multiplexer
A&B printing	A&B mix
Measurement of Density	Video level
Sets D-max	Pedestal
Frameline	Blanking
MP Camera	Video camera
MP film	Video magnetic tape

These are just some of the terms that can be related to the film/video interface and are presented only as examples. They have been excerpted from a proposed listing of motion picture nomenclature being proposed by the Society of Motion Picture and Television Engineers.

Many thanks to those major contributors listed below and to the Eastman Kodak Company for the funding that made this chapter possible.

Karolyn Alves	John Pytlak
Paul Kanerva	Richard Sehlin
Glenn Kennel	Walter Snyder

APPENDIX A: RELEVANT SMPTE/ANSI STANDARDS

Recommended Practices (RP) and Engineering Guidelines (EG)

Alignment Color Bar Signal	EGI-1990
Density, Films and Slides	RP46–1990
Image Area: 16mm Film	PH22.96–1982
35mm Film	PH22.95–1984
Review Rooms	SMPTE148–1984
Review Room Screens	RP41–1983
Slides and Opaques	SMPTE94–1985
2x2 Slide Mounts	RP9–1986
Monitors: Color Temperature	RP37–1969, R1982
Colorimetry	RP145–1987

Proposed SMPTE-RP

T14.280 Professional 525-line TV/Studio, Type A Color Picture Monitors

T14.281 Critical Viewing Conditions of Color TV Pictures

T14.282 Alignment of Professional, TV/Studio Color Picture Monitors

APPENDIX B: RELEVANT SMPTE AND OTHER PUBLICATIONS

SMPTE

1. Hunt, "Color Bars on Film for Setting Up Telecines," Feb. 1987, 78.
2. DeMarsh, "Evaluation of Color Reporduction in Film and Television," Sep. 1986, 624.
3. Staes, "Role of Film in Film-Plus-Telecine System," Sep. 1987, 565.
4. Spies, "Direct Assembly of Motion Pictures on Video Tape," Jul. 1987, 451.
5. Wood, "Interchangeability Between Film and Tape for Worldwide Distribution," 90:203, Mar. 1981.
6. Kriss & Liang, "Today's Photographic Imaging Technology for Tomorrow's HDTV System," 72:804, Aug. 1983.
7. Staes and Hayen, "Image Quality Transfer Through Film and Television," 90:196, Mar. 1981.
8. Markle, "The Development and Application of Colorization," 93:632, Jul. 1984.
9. Barrett, Callette, Lisk, Sager, and Suppala, "An Experimental Trinoscope for Improved Video to Film Recording," 93:746, Aug. 1984.
10. Poetsch, "FDL60—An Advanced Film Scanning System," 93:216, Mar. 1984.
11. Millward, "Flying Spot Scanner on 525 Line NTSC Standards," 90:786, Sep. 1981.
12. Lisk, Pytlak, and Barrett, "New Tools for Improved Telecine Quality," 93:6, Jan. 1984.
13. Powell and Kennel, "Noise in Film to Video Transfers," 96:16, Jan. 1987.
14. DeMarsh, "Optimum Telecine Transfer Characteristics," 81: Oct. 1972.
15. Staes and Markie, "The Interface of Color Negative Film and Telecine," 92:203, Mar. 1983.
16. Kolb, "Telecine Reproduction of Motion Picture Audio," 88:835, Dec. 1979.
17. Marsden, "Improved Automatic Color Correction for Telecine," 87:73, Feb. 1978.
18. Lees, et al., "High Performance CCD Telecine for HDTV," 99:837, Oct. 1990.

Other Articles and Publications

1. Reinking, "Film to Tape Mysteries Unravelled," *American Cinematographer*, Sep. 1988.
2. Edited by Fred Detmers, *American Cinematographer Manual*, 6th Edition, ASC.
3. Eastman Kodak Company, "TAF users Guide," Publication H-9, 1989.

4. McMurray, "Telecine Techniques to Retain Film Subtleties," BM/E, Nov. 1985.
5. Eastman Kodak Company, "TV Questions and Answers," Publication H-8, 1989.
6. Eastman Kodak Company, "Cinematographers Field Guide—Motion Picture Camera Films," Publication H-2.
7. Kennel, et al., "A Comparison of Color Negative Films and HDTV Cameras for TV Program Production," SMPTE 132nd Technical Conference, Oct. 1990 (unpublished).

APPENDIX C: FILMS FOR TELEVISION

Color Camera Films

- Eastman Ektachrome Film (Daylight) 5239 (35mm), 7239 (16mm)
- Eastman Ektachrome Film (Tungsten) 5240 (35mm), 7240 (16mm)
- Eastman Ektachrome High Speed Daylight Film 7251
- Eastman Ektachrome High Speed Tungsten Film 7250

NOTE: The above films are reversal camera films that have been used extensively for news gathering and documentary program production.

- Eastman EXR Color Negative Film 5245 (35mm) 7245 (16mm)
- Eastman EXR Color Negative Film 5296 (35mm)
- Eastman EXR Color Negative Film 5248 (35mm) 7248 (16mm)
- Eastman Color High Speed Daylight Negative Film 5297 (35mm) 7297 (16mm)
- Eastman Color High Speed Negative Film 97 7292 (16mm)

NOTE: The above listing covers some of the many color negative films used for TV program production and theatrical motion picture production used on television. For a more complete listing and description of chracteristics, see Eastman Kodak Company Publication H-5, 1989.

Color Release Film

- *Eastman Color LC Print Film 5380 (35mm), 7380 (16mm)
- Eastman Color Print Film 5384 (35mm), 7384 (16mm)
- Eastman Ektachrome Print Film 5399 (35mm), 7399 (16mm)

* This film was specifically designed with lower contrast (LC) for use as telecine.

5.9 Studio Audio

Walt Lowery
Broadcast Supply West, Tacoma, Washington

INTRODUCTION

This chapter describes the equipment and layout of audio facilities used in radio and television studios. First is an overview of typical studio layout and planning, and then of major components used in an audio facility. The purpose of this chapter is to provide guidance in designing or rebuilding a studio.

The quality of a station's sound (a radio station's only distributed product) is determined by what happens in the studio. The best source material and the best air talent will produce only marginal results when burdened by inadequate equipment. Even in television, audio can no longer be treated as a secondary technology. State of the art equipment, properly used in the studio, translates to ratings and rate card figures.

TYPICAL RADIO STUDIO LAYOUTS

Studio design is based on factors such as programming needs, available space, and creativity of the current and former engineers. Many stations in active markets have gone through changes in formats, managers, and engineers leaving the studios as a crazy quilt of equipment ripe for redesign and rebuilding.

It is possible to catalog most radio station studios under a few general categories. The main design parameter should be the current station format. The operational concepts will be much different for a music format than for a news/talk operation.

Music Formats

Studios built for music formats are the most common, particularly in smaller markets. The basic configuration is an audio console, two turntables or CD players, and multiple cart players arranged on a "U" or "L" shaped desk. Other associated equipment might include audio routing, transmitter remote control, and telephone interface equipment. This configuration is designed to handle rather simple programming needs. The announcer on duty mixes and switches his own audio, reads announcements, plays music from records or CDs, plays commercials from carts, and switches in the news and network programming from satellite.

At some stations, music is transferred from vinyl or CD to cart, making the studio an all-cart operation. In this situation the studio will contain four to six cart-playback decks and no turntables or CD players. Having all program material on cart makes the operation more "idiot proof." The operator needs only to talk, shove plastic boxes into slots, and push buttons. Going one step further, some stations are installing digital audio storage systems. Commercials and music are converted to data, stored in a computer on a hard drive, retrieved when required, converted back to analog audio, and fed to the console. This eliminates the physical handling of any audio storage medium.

Most music formats require positions for one combo operator and possibly an announce booth for news. With the popularity of the "Morning Zoo" format, many large market stations have studios designed for two or three on-air personalities. One member of the team will operate the console, one will assist with phone calls, pulling music and spots, and the third will handle news. The studio layout will vary with the duties handled by each member of the team.

News/Talk Formats

In this format, information going on the air is either live in-studio sound or short feeds from a large number of sources. Here the board operator may be more of an engineer than air talent. The control room will generally be surrounded by a number of small studios used for news and talk shows. The console will be used more for switching than audio mixing. A digital audio storage system is also helpful in this format as the operator is busy enough without handling carts.

Figure 1. Typical radio studio. (Courtesy Harris-Allied & Arrakis Systems.)

In some operations the audio console has been replaced with a computerized audio-switching system which is accessed by a touch screen. Some vendors build custom systems for their customers using PCs and touch screens linked by LANs and RS-232 connections. The operator sees the log displayed on the screen in sequence. He or she can shift events around, control audio source equipment, adjust levels, and even read copy and tags directly from the screen.

The Production Studio

A separate studio is generally used for production and transfer of music to cart or digital audio storage. The production studio will have turntables, CD players, reel-to-reel recorders, cart recorders, equalization, patch panels, and possibly a digital audio workstation. The production studio should be more flexible and handle a wider variety of audio mediums than the on-air studio. Special effects and equalization not needed in the main studios are frequently required in production. Four, eight, and sixteen-track consoles and recorders are found in major market production facilities. These consoles will have submaster mixing busses which allow the production operator to mix down two or three tracks at a time in producing complex spots or promos.

Digital workstations are becoming more popular because they allow editing without razor blades and splicing tape. Editing of individual tracks is also possible, providing a big advantage over reel-to-reel editing. Digital workstations are bringing to audio production the speed and versatility that word processing brought to typing.

Since all material produced in this studio will eventually be played on the air the quality of the equipment should be equal to, or better than, that used in the main studios. In smaller markets, the tendency to scrimp on equipment in the production room should be avoided.

TYPICAL TELEVISION STUDIO LAYOUT

Audio for television stations can be challenging because of the need for several types of audio mixes. Within the studio, a stereo program feed is needed along with a monitor mix for on-camera talent, a mono mix for a studio audience, a mix-minus for telephone hybrids, and possibly a mix in a different language for

Figure 2. AKG DSE-7000 digital audio workstation. (Courtesy AKG & Harris-Allied.)

a second audio program (SAP) channel. Monitoring of the preview channel is possible by using the mixer solo function. If a remote truck is involved, an interruptable foldback (IFB) system is required to feed program audio and cues to the talent at the truck. The console to handle all this is designed specifically for the task. The operator is assigned the job of mixing audio and nothing else.

Some of the work of handling the audio can be done by an audio routing switcher if it is tied to the video switcher logic. In this situation, the audio would be switched with the video source and any level adjustments would be made of by the station's audio processing chain.

Studio Planning

Planning a new studio or rebuilding an old one begins with a layout on paper or computer of all of the required audio sources and feeds. A good place to start is with console inputs. The engineer makes a list of all possible sources and assigns them priorities according to how often and how quickly they are needed by the operator. This will help determine the number of mixing channels and switched inputs needed for each mixer.

All frequently-used audio sources should be assigned to individual console mixers. Input switching and patching by the operator should be kept to a minimum to avoid errors and "dead air." It is common practice to run all line-level inputs through patch panels on their way to the console inputs. This allows for rerouting for special programming and "patching around" any problems which may develop. Microphone inputs are usually wired directly to the console as their positions are seldom changed and because signal-to-noise ratios are a factor when dealing with levels of −60 dBm. A possible exception is in a television studio, where the program may move from set to set.

Part of planning for a new console is ensuring that the levels from all audio sources will be compatible with the input levels required by the console. If this is not done, the operating positions of the mixers will be different, making it difficult for the operator to run the board properly. In extreme cases the operator may not be able to open the pot more than a fraction of a turn before driving the meters to the pin or may not be able to get enough gain even with the control fully open. Neither situation is acceptable and in most cases will result in distortion or poor signal-to-noise figures. Attenuators should normally operate at the two o'clock position (if rotary) or at about 60% (if the slider type).

All studio sources should be set to provide their normal output levels, and, where necessary, H-pads to the console inputs. The pads can be bought preassembled and mounted inside the equipment or on punch-blocks. Pads of 10 dB or greater are also useful in correcting impedance mismatches. Pads should also be used when two audio transformers are connected directly together. If an output transformer directly feeds an input transformer, the output transformer sees an inductive load and frequency response of the system will be affected. With an H-pad between them, the output transformer will see a linear resistive load. The accompanying chart gives resistor values for 600 ohm H-pads.

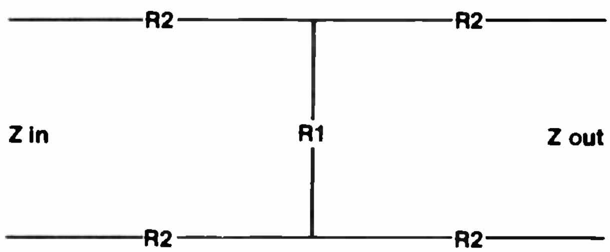
The worksheet shown in Fig. 4 is a suggested starting point for designing a studio. Engineers should prepare customized worksheets for each studio. The studio outlined in Fig. 4 is one of three in a larger station. The console has twenty mixer modules, each with two inputs. All audio is routed through patch panels and the outputs feed distribution amplifiers. Outputs from other studios appear on the patch panel in the main studio as well as on mixers. If technical problems take the control room off-line, the station then can originate programs from Studios B or C.

The reader should refer to Chapter 5.12 for guidance in the physical layout of the studio. Television engineers will also want to read Chapter 5.13 on studio lighting.

Buying Equipment

After the studio is designed, the next step is to prepare a shopping list and cost estimate for new

Loss (In DB)	R1 (Ohms)	R2 (Ohms)
3	1.8K	51
6	820	100
10	430	160
15	220	200
20	120	240
30	39	270
40	12	300



H-PAD

Z In = Z out = 600 Ohms

Figure 3. Typical H-Pad values.

equipment. The major components (console, cart players, professional CD players, and furniture) should be chosen early in the planning stage so that their quality will not be compromised if budget cuts are required later on. Allow for price increases, and do not forget sales tax and shipping costs.

Where to buy? Each engineer should develop a working relationship with a reputable broadcast equipment dealer. Absolute bottom-dollar may not be the best deal. Watch out for hidden "handling" or drop-shipment charges. Consider that there are no real savings in paying $15.00 less for a CD player that is not delivered in time to make the on-air date for the new studio. Experienced broadcast equipment salesmen give their best deals and service to customers to whom they sell on a continuing basis.

Fax the shopping list to the dealer a week or so before the cost estimate is due in order to allow time to research and work up a quotation. Competition in the broadcast supply business generally ensures that pricing between reputable dealers will vary by only about a few percentage points.

If management insists on competitive bids, limit this exercise to two bids. Time is more valuable than chasing nickels and dimes. If a regular dealer does not carry an item required for the project, ask his recommendation for a source. A salesman will most

Audio Source	Patch Panel #1 Jacks #	Jacks #	Console Miser #	Console Outputs	Patch Panel #3 Jacks #	Jacks #	DA Input #	DA Output #	Patch Panel #4 Jacks #	Jacks #	Termination
Microphone #1			Mixer 1-A	Prog. L	Jacks 1 & 2	Jacks 25 & 26	Input #1	Output #1	Jacks 1 & 2	25 & 26	Audio Processing
Microphone #2			Mixer 2-A					2	3 & 4	27 & 28	Studio B
Microphone #3			Mixer 3-A					3	5 & 6	29 & 30	Studio C
Spare			Mixer 3-B					4	7 & 8	31 & 32	Spare
CD #1	Jacks 1 & 2	Jacks 25 & 26	Mixer 4-A	Prog. R	3 & 4	27 & 28	2	1	9 & 10	33 & 34	Audio Processing
CD #2	3 & 4	27 & 28	Mixer 5-A					2	11 & 12	35 & 36	Studio B
CD #3	5 & 6	29 & 30	Mixer 6-A					3	13 & 14	37 & 38	Studio C
Spare	7 & 8	31 & 32	Mixer 6-B					4	15 & 16	39 & 40	Spare
Cart #1	9 & 10	33 & 34	Mixer 7-A	Audition L	5 & 6	29 & 30	3	1	17 & 18	41 & 42	Reel to Reel L.
Cart #2	11 & 12	35 & 36	Mixer 8-A					2	19 & 20	43 & 44	Studio B
Cart #3	13 & 14	37 & 38	Mixer 9-A					3			
Cart #4	15 & 16	39 & 40	Mixer 10-A					4			
Turntable #1	17 & 18	41 & 42	Mixer 11-A	Audition R	7 & 8	31 & 32	4	1	21 & 22	45 & 46	Reel to Reel R.
Reel to Reel #1	19 & 20	43 & 44	Mixer 11-B					2	23 & 24	47 & 48	Studio B
Cassette #1	21 & 22	45 & 46	Mixer 12-A					3			
Spare	23 & 24	47 & 48	Mixer 12-B					4			
									Patch Panel #3		
	Patch Panel #2			Mono	Jack #9	33	DA Input #5	1	Jack 11	35	Office Monitor
Studio B	Jacks 1 & 2	Jacks 25 & 26	Mixer 13A	Output				2	Jack 13	37	Music on hold
Studio C	3 & 4	27 & 28	Mixer 14A					3			
Spare	5 & 6	29 & 30	Mixer 14B					4			
News Network	7 & 8	31 & 32	Mixer 15A	Mix-Minus	#15	35					Hybrid #1
Satellite	9 & 10	33 & 34	Mixer 16A	#1							
Phone Hybrid #1	11 & 12	35 & 36	Mixer 17A								
Phone Hybrid #2	13 & 14	37 & 38	Mixer 18A								
RPV Receiver	15 & 16	39 & 40	Mixer 19A	Mix-Minus	#17	37					Hybrid #2
EBS Receiver	17 & 18	41 & 42	Mixer 20A	#2							
Spare	19 & 20	43 & 44	Mixer 19B								
Spare	21 & 22	45 & 46	Mixer 20B								

Figure 4. Example of a studio wiring worksheet.

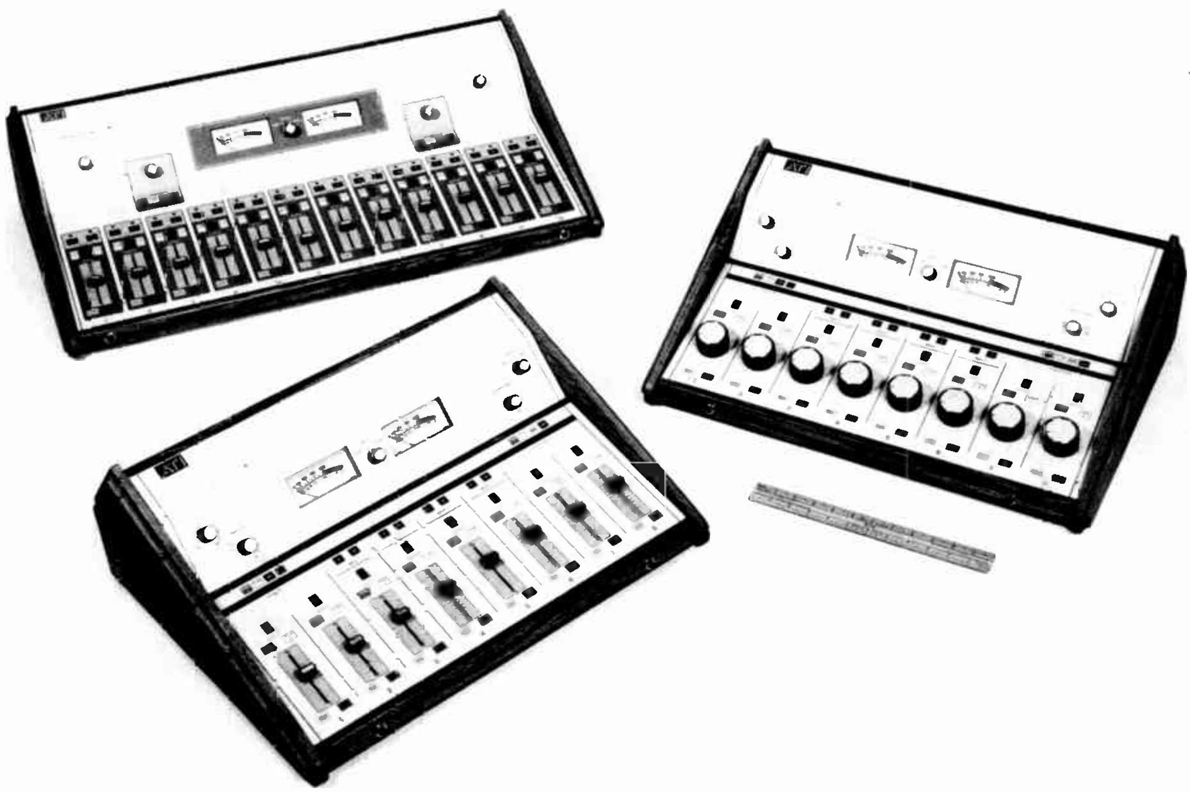


Figure 5. The Vanguard Audio consoles from ATI. (Courtesy ATI.)

likely know a good supplier and sometimes will be able to get equipment not in his normal line for his best customers. His extra service can be worth a lot more than a few dollars in a pinch.

AUDIO CONSOLES

Radio Consoles

Centered in the radio studio, in front of the disc jockey and the corresponding worn spot in the carpet, will be the main audio console. Smaller mixer consoles are usually found in news rooms and smaller production studios. The number of mixing channels limits the number of audio events that can occur simultaneously or in rapid succession. The station format will dictate the flexibility and ease of operation required. Although an operator-assisted, easy-listening format may be able to use a four or five channel console, it would be out of the question for a "Morning Zoo" type of fast-paced program with multiple events. These smaller consoles may not have an audition bus or switchable inputs on the mixers. Without multiple switched inputs, either one mixer per audio source or external switchers will be required.

The console found in most control rooms will have eight or more mixers. In smaller markets these will generally have rotary attenuators. For heavy on-air use, step attenuators are preferred because they require only occasional cleaning in order to maintain silent operation. These use make-before-break switches to move through a series of resistive pads. The step between contacts results in a two or three dB difference and is constant throughout the entire range of twenty or more steps.

Their rugged construction results in smooth, quiet, dependable operation, but with a large size and a rotary design. Program audio normally passes directly through the attenuator, but a switch will route the

Figure 6. The Arrokis 500 Series Console. (Courtesy Arrokis Systems, Inc.)

COMPLETE VCA SCHEMATIC

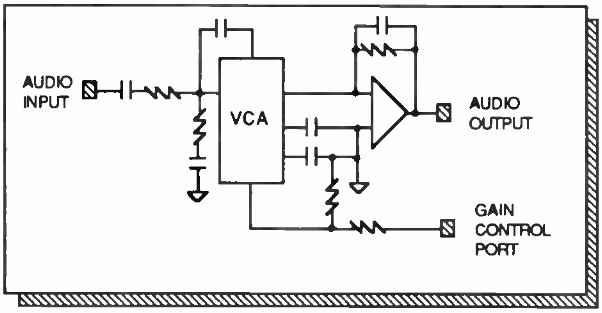


Figure 7. VCA schematic. (Courtesy Arrakis Systems, Inc.)

audio into the console cue bus when the attenuator is turned fully counter-clockwise.

Attenuators which depend on sliding a contact over a resistive element are subject to wear and build-up of dirt. The resistive element will be either carbon or conductive plastic. Normal wear changes the element's resistance and build-up of worn-off carbon may cause erratic resistive changes in the contact between slider and element. The results are noise and uneven operation between channels.

Noise is a major problem if the program audio is routed directly through the potentiometer. This design does allow the attenuator to be either rotary or a straight-line slider. In the case of a voltage controlled amplifier (VCA) design, audio does not pass through the control. While attenuator control noise will not be heard in program audio, the action of the attenuator will become erratic and nonlinear as the sliding contact attenuator is moved through its range. With the VCA design, only a sample voltage is passed through the potentiometer. This voltage controls the gain of an amplifier through which the program audio passes.

This console design allows the use of cheaper rotary or straight-line attenuators. As with any amplifier, the VCA will introduce some thermal noise and distortion. The amount is generally a direct function of price. A good console should have an overall distortion figure of 0.05% or less. The noise floor should be less than −80.0 dBm.

Modular design consoles offer the engineer several major advantages. Modules can be removed and swapped for troubleshooting. The design of the console can be changed if the station format changes. Mixers can separated by blank panels in order to group sources together. Custom panels can be created or purchased for special functions such as reel-to-reel recorder control or telephone line selection. If extra space is left in the mainframe, the console may later be expanded with additional mixers as the station's needs grow. These benefits often justify the extra investment in a modular console.

For on-air use, a cue channel is vital to allow the operator to cue records, receive cue signals from remotes and networks, and preview program material. Even if all program material is on carts, the cue channel is a must.

In the early days of radio, when all program material was live, the audition channel was the cue channel. Levels were preset, and program material was previewed using the audition channel. Some of these early consoles did not have a cue circuit. Today the audition channel is used by the operator to record network feeds for later use while the station audio is being mixed on the program channels. Some consoles allow audio from the mixers to be fed simultaneously into the program and audition channels, allowing on-air programming to be recorded using the audition channel. The audition channel can be monitored by the studio monitor system to preset levels and check audio quality of remotes. In talk formats, the audition channel

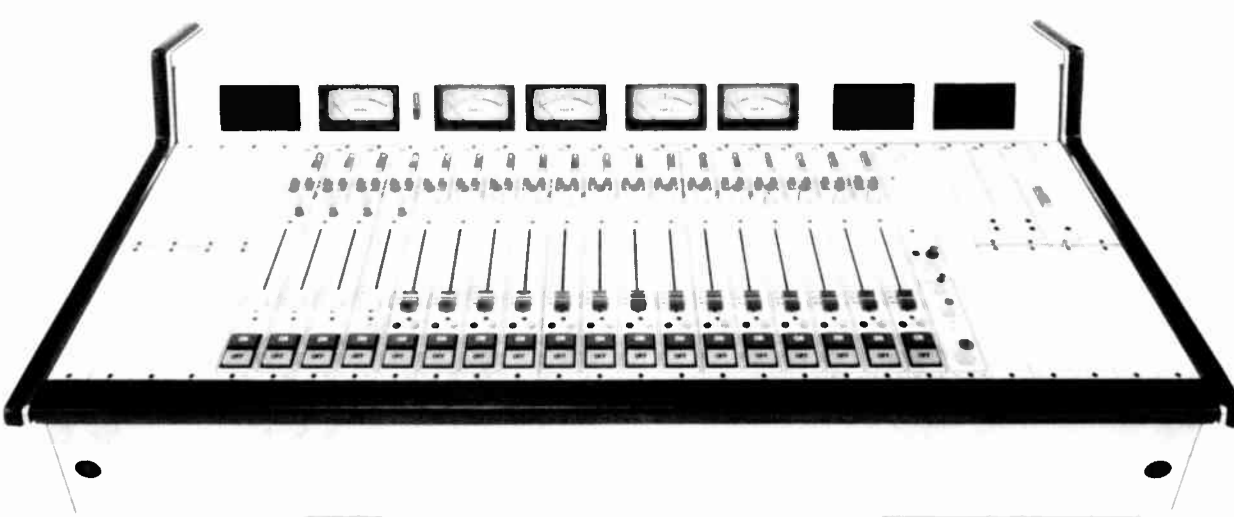


Figure 8. Auditronics 200 Series Console. (Courtesy Auditronics, Inc.)

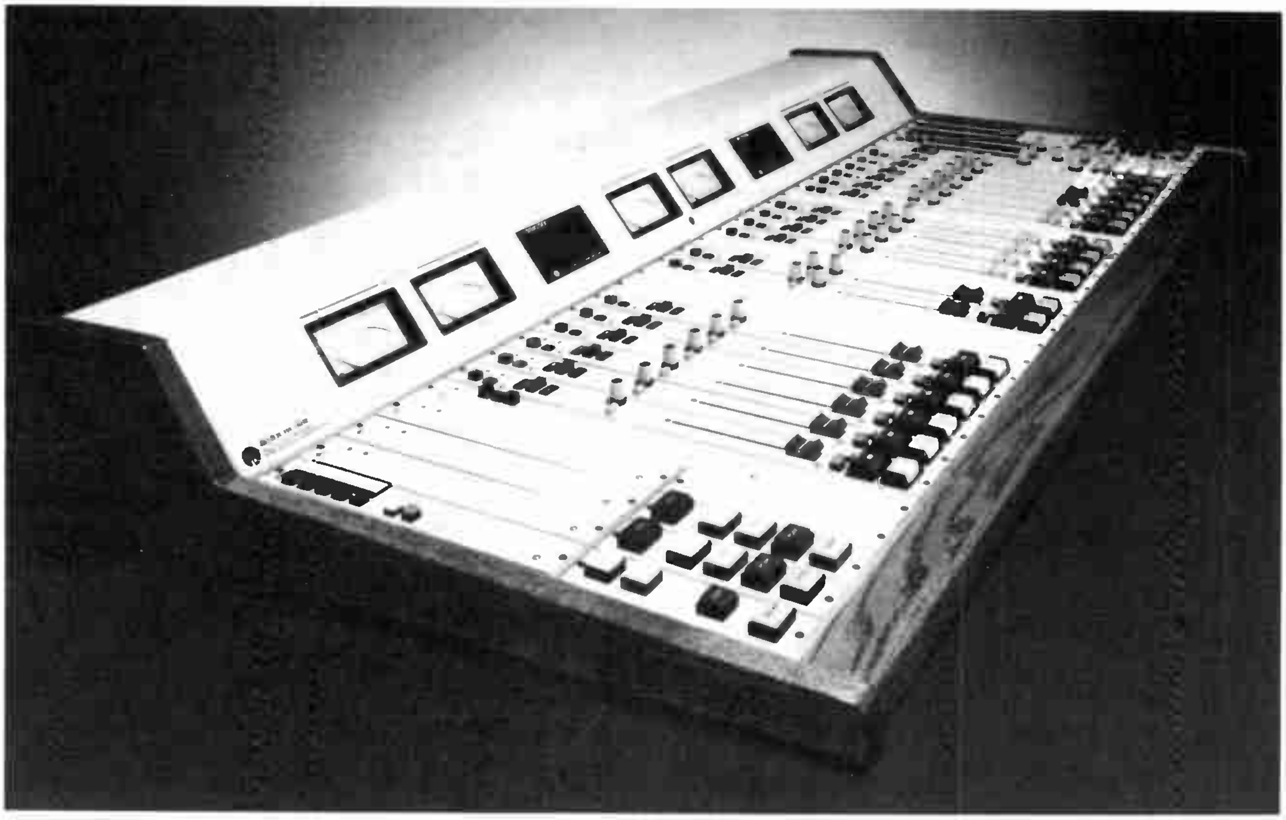


Figure 9. Pacific Recorders BMX Console. (Courtesy Pacific Recorders.)

is often used to mix program audio which is fed to a delay unit. The output of the delay system is fed into the program channels and to the transmitter.

Television Audio Consoles

The requirements of TV audio are complex because more of the audio sources are "live" (as opposed to radio, where most program material is recorded). In television installations, more mixes are required because of monitoring requirements.

Television consoles will include mono input modules for microphones, telephone hybrids and other mono-only sources. These modules will include a pan pot to allow left-right positioning of the apparent audio source. Often, equalization will be required, and equalization or filtering on each module is a popular option. The input will have a gain control or switch to allow stepping between microphone and line-level input. There will be a cue channel feed and solo button which allows stereo monitoring of an individual audio source in the control room monitors.

In the consoles, stereo input modules will contain all of the above features except that the pan pot will be replaced by a mode switch which enables the operator to select normal stereo, left-channel only, right-channel only, a mono mix, or reversed channels.

A major departure from consoles used in radio will be the ability to create multiple audio feeds by the use of submaster mixing busses. Mixer outputs are assigned to submaster busses, and these submasters can be added to the master mix and monitored separately. This allows the creation of several mix-minus feeds for special monitoring requirements. In video production, audio can be mixed manually or can be handled with a mixer controlled by the switcher.

Console Features and Options

A growing list of extras is now available on most larger audio consoles used in both radio and television. These make them more user-friendly and "idiot-proof."

Clocks and timers are available. The timer will reset to zero anytime that a new channel is selected so the operator will know how long a CD or cart has been running.

Mix-minus busses are becoming standard equipment. Mix-minus is necessary when putting telephone calls on the air to provide audio down the line to the caller. This feed is console audio minus the caller's voice. If program audio, which contains the caller's voice, were fed to the phone hybrid, a feedback path would exist. If multiple phone hybrids are used in a studio, a mix-minus feed for each hybrid is required so that callers will be able to hear each other's comments.

For television, a mix-minus feed is used to provide air monitoring for the talent on set. This is a feed of

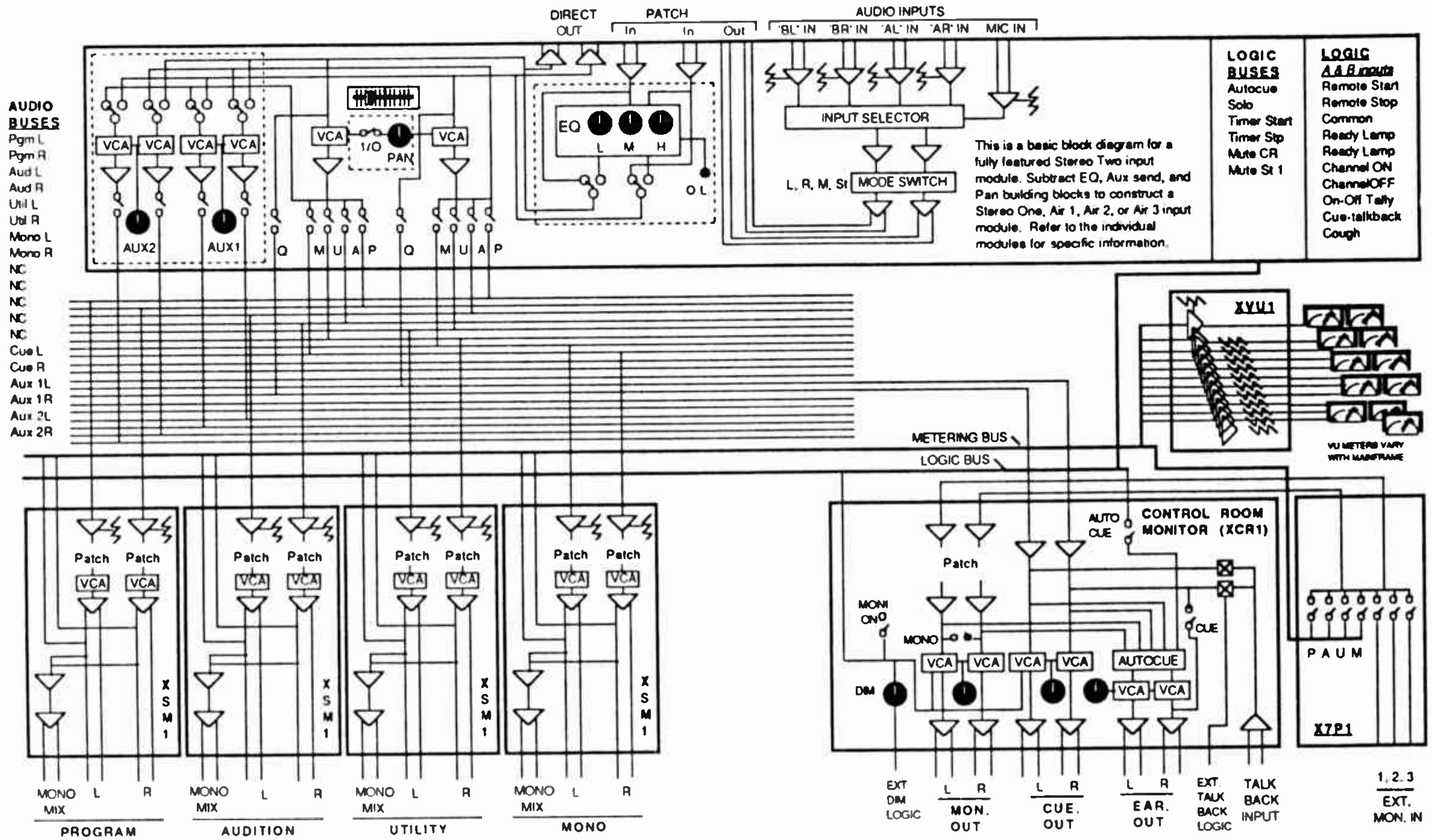


Figure 10. Audio console block diagram, the Arrakis 10000 Series. (Courtesy Arrakis Systems, Inc.)

Figure 11. Pacific Recorders STX Console. (Courtesy Pacific Recorders.)

program audio minus the talent's microphone, thereby preventing acoustic feedback while allowing the talent to hear program material and cues.

Consoles designed for radio production and for television will have prefader processing patchpoints. These allow the audio source to be routed to compressors, equalizers, or other processing devices before it is routed to the mixing bus. This is used primarily for microphone processing.

Monitor amplifiers may be external or built into the console. Built-in monitor amps are usually limited to five to eight watts because of power and space limitations. Many engineers prefer to use external, higher-power amplifiers to drive the studio monitors. The stereo monitor system should contain a stereo/mono switch or button. This allows the operator to check for out-of-phase program material and misaligned tapes. This switch can easily be added by wiring it between channels at the monitor gain control. At some stations, the control room monitors are operated in the mono mode. This insures that station personnel will be aware of out-of-phase conditions at least as soon as listeners. A well-designed console will have interchangeable input cards so that input sources can be moved to different mixers to allow for future changes in the studio. An alternate system would have

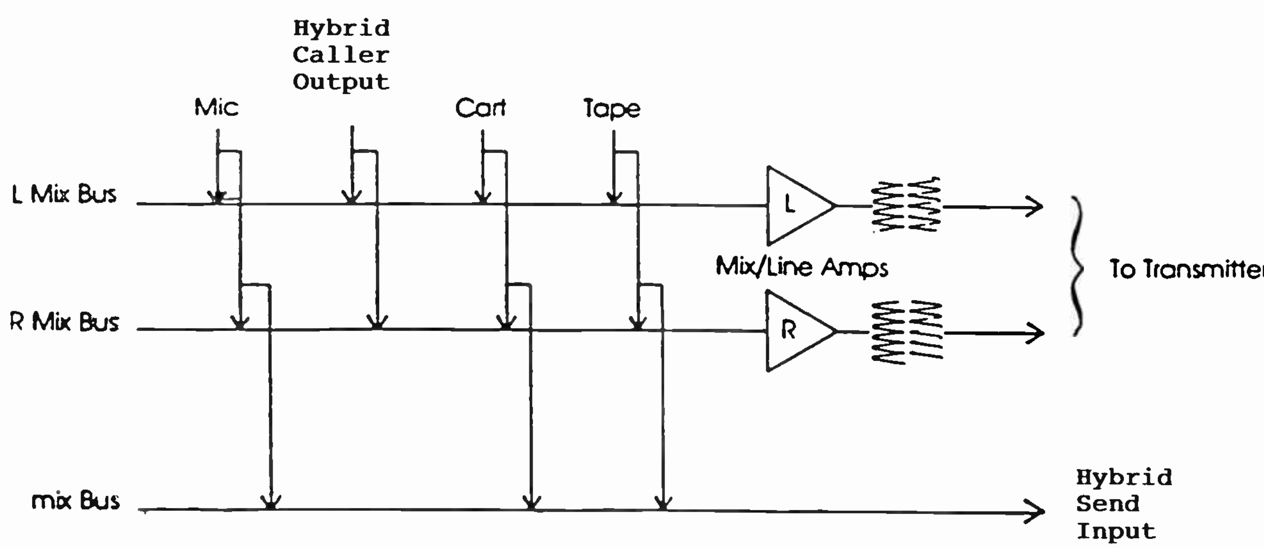


Figure 12. Mix-Minus for telephone hybrid.

switchable input levels and impedances. An input card could then be used at either microphone or line level. More consoles are being built with balanced bridging inputs. These make for a wider variety of input sources than were possible when everything was designed for 600 ohm, +4 dBm inputs. If a 600 ohm termination is needed a 600 ohm resistor is tied across the input. Other termination requirements are similarly met.

Going one step further, some consoles have programmable presets for console input configuration. The console inputs are set up for specific uses and the source selections are stored in data. The operator tells the console which use or program is needed, and all of the sources are automatically switched to the proper mixer.

Remote start contacts for cart machines and other program sources are standard on better consoles. They allow the operator to start the equipment by simply turning on the appropriate channel. Some console manufacturers are using logic circuits which allow the mixer to be turned on by pushing the start button on the cart player. When the cart machine recues, the mixer is automatically turned off. One important feature to look for in these systems is automatic disabling of this logic function when the input selector is switched to some other source. This eliminates the annoyance of having a cart machine start when the mixer is used for an auxiliary function.

Some manufacturers offer consoles with a choice of conventional analog VU meters or LED bar-graph metering. LED metering is available in three colors and can give indications of root mean square (RMS) and peak values. One model shows left, right, and peaks on the same display. Some believe that operators using LED displays are less likely to run a board with the meters "buried in the red."

With multitrack recorders and digital workstations in the broadcast production environment, there is a need for consoles with more than a single pair of left and right outputs. Four and eight-channel consoles are common in major market stations. Channel assignment switches route the mixer's audio to the proper bus, pan pots shift audio between left and right, and equalizers are available on each mixing module. At this point, the console is moving close to what was once found only in recording studios.

Engineers should be aware that digital technology is coming rapidly to the console industry. One major console manufacturer has developed a console in which all of the switching and level control is done digitally. The fluid situation with regard to standards makes it impractical to include such equipment in this edition of the *Handbook*.

AUDIO DISTRIBUTION AND ROUTING

Patch Panels

Patch panels (or jack fields) come in three basic types. In the ¼-inch jack size there is the tip/sleeve type which consists of one conductor and one shield.

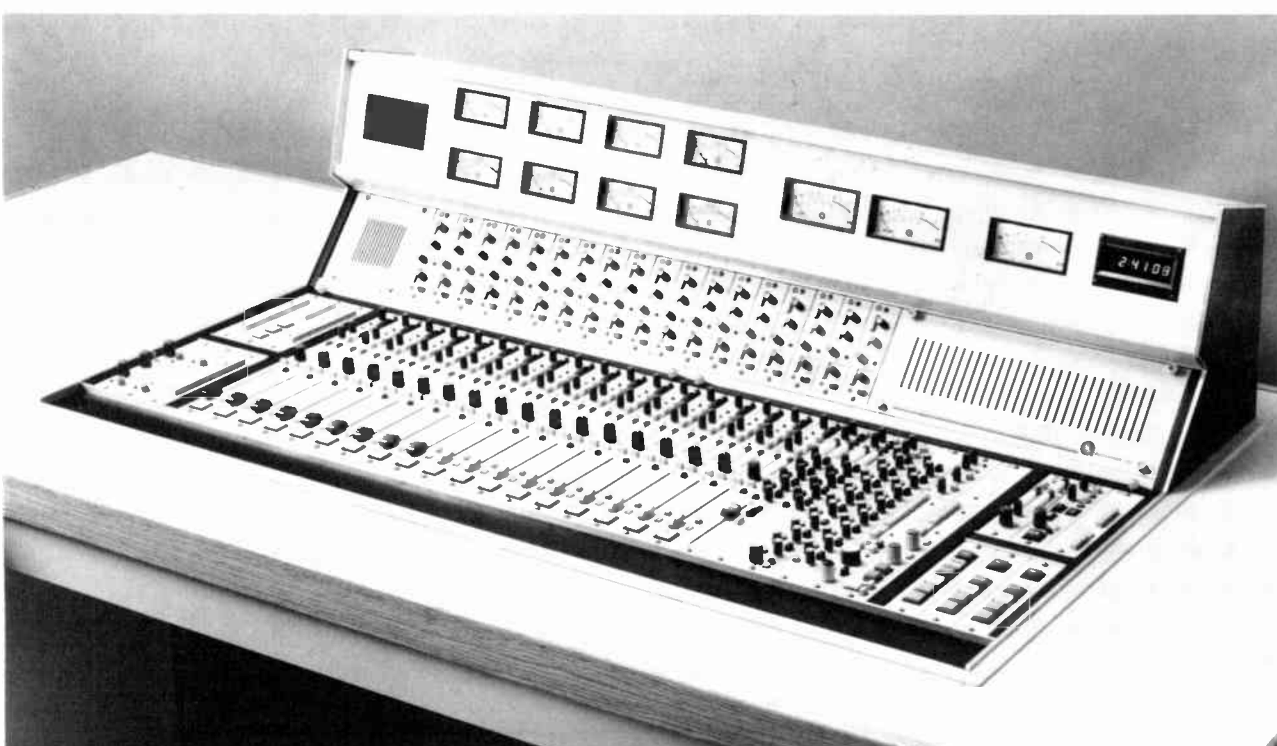
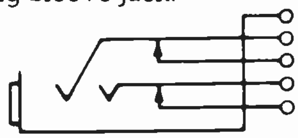


Figure 13. Auditronics 400 Series Console. (Courtesy Auditronics, Inc.)

T-R-S Tip-Ring-Sleeve Jack.



T-R Tip-Ring Jack (Sometimes referred to as Tip-Sleeve Jack).

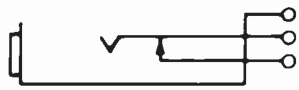


Figure 14. Patch panel Jack types.

This type dates back to the early days of radio and is obsolete. Four single plug cords would be necessary to patch a balanced stereo connection. More common is the tip/ring/sleeve ¼-inch panel. This configuration offers two conductors, each shielded. Two single-plug cords will complete a stereo circuit. Dual-plug cable assemblies are also available, allowing the operator the convenience of patching a stereo source with one cable, and, with obvious workings, avoiding channel reversals.

Quarter-inch patch panels are available in single rows of 24 jacks or dual rows of 48 jacks. With this configuration, stereo spacing is possible. Jacks are grouped in pairs with wider spacing between pairs than between individual paired jacks. When using a dual plug patch cord, it is impossible to cross patch or insert the cord in half of one audio feed and half of the circuit next to it on the patch panel. The dual plug will fit in paired jacks only.

The ¼-inch panels are also available with single rows of 26 jacks or dual rows of 52. These panels have standard spacing between all jacks and allow an additional stereo circuit on the 26 jack version and two additional stereo circuits on the 52 jack, dual row model.

Custom versions of the ¼-inch jack field are also available on special order. One version has three rows of 26 for a total of 78 jacks. The two lower rows are wired conventionally and the top row is bridged across the middle row and allows for monitoring without interrupting the circuit. Another custom item is a patch panel with special jacks in which not only are the conductors normalled but the shield is also switched if a patch cord is inserted. This configuration is used to patch microphone circuits.

One popular arrangement of patch panels comes built into a 19-inch rack mount chassis. The entire assembly may be mounted into an equipment rack. The jacks appear at the front of the rack and the termination is at the rear for easy access to equipment wiring. There is a misconception that this design provides protection from RF. Don't count on it: the bay fronts are phenolic material and some manufacturers use unshielded wiring inside the box.

Rapidly making its way out of the recording studios and into broadcasting is the bantam or tiny-telephone jack field. These plugs and jacks are 0.175 inch in diameter and are of the tip/ring/sleeve configuration. This bantam patch panel will take up about half of the space in an equipment rack that a similar ¼-inch jack panel would require. Ninety-six jacks can be accommodated in a 1¾-inch by 19-inch rack space.

Patch Panel Wiring and Termination

The jacks used in an audio patch panel have a set of contacts on each circuit which make contact when

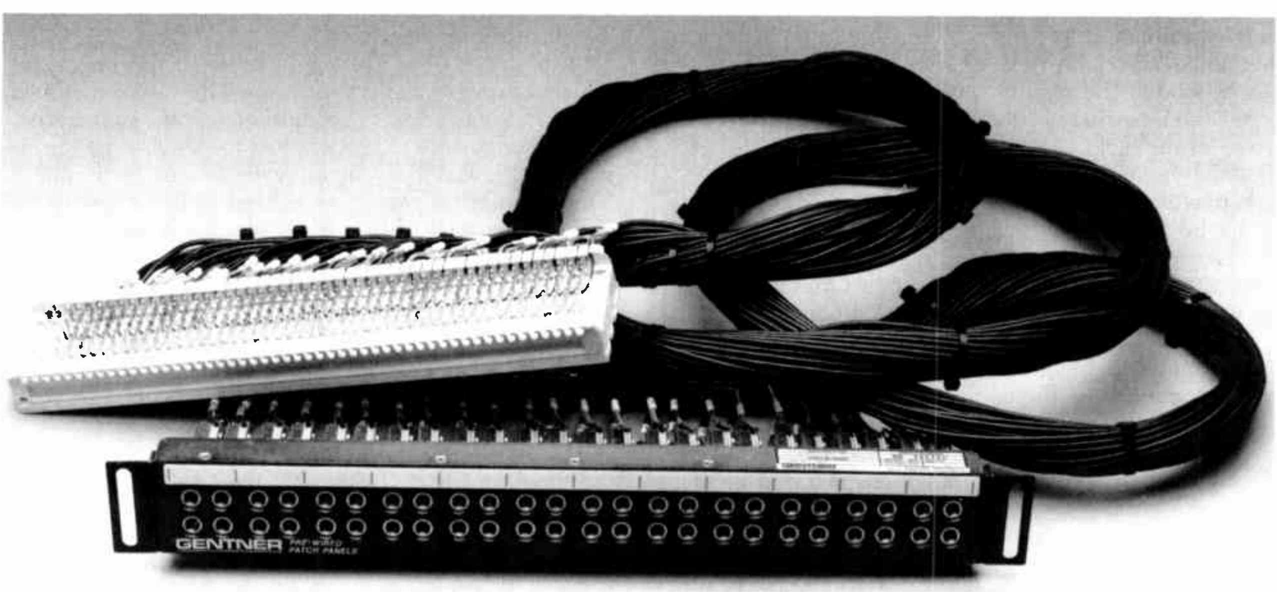


Figure 15. Pre-wired patch panel. (Courtesy Gentner Electronics.)

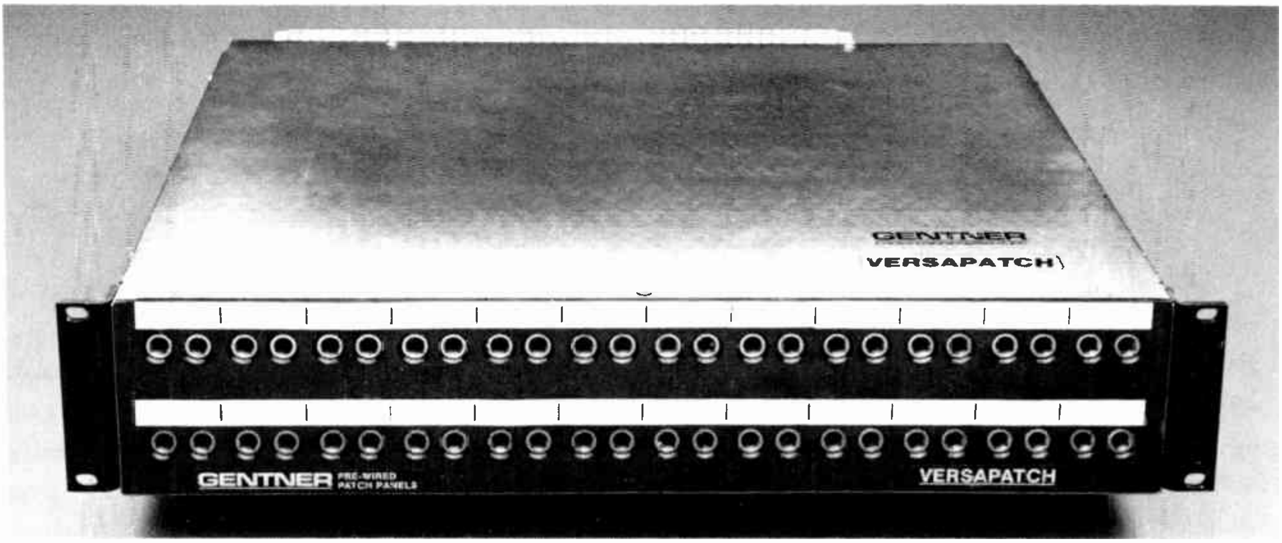


Figure 16. The Gentner Versapatch. (Courtesy Gentner Electronics.)

no plug is in the jack. When a plug is inserted, the contact is broken. This allows for *normaling*: an audio output wired to a pair of such jacks is transferred automatically to the pair of jacks associated with the corresponding input which it normally feeds, provided no plug is inserted. The proper procedure is to wire all outputs to the top row of jacks in a dual row patch panel and the normal inputs to the bottom row directly below them. When the circuit is not interrupted by the insertion of a patch cord, outputs would feed to the proper inputs, their normal connections.

The connection for the normals can be made by jumpers at the jacks, or this function can be brought out to the termination. If the normals are brought out to the termination, the engineer can determine whether he wants the circuit to be normalled or not. Changes can be made in normalling without removing the patch panel and unwiring the jumpers. Most patch panels are wired at the jacks.

If the engineer wants to reroute the output of the cart player normalled to Console Input 3 from Input 6, he/she would insert a pair of cords in the top row of jacks associated with the cart machine's output. This interrupts the audio path going to Mixer Input 3 by breaking the normal circuit. The other end of the patch cords is inserted in the jacks on the lower row associated with inputs for Mixer 6. This breaks the normal circuit from the audio device normally feeding Mixer Number 6 and puts the output of the cart player on that mixer. With conventional normal wiring, a circuit is broken whenever plugs are inserted in the jacks.

If the engineer wants the ability to monitor equipment outputs at the patch panel, half-normalling or top-row bridging can be used. Here, the jumpers for the top row of jacks are wired to the jack arms or wiper contacts with the audio connection. The normal contacts on these top-row jacks are not used. The circuit between the top-row jacks and the bottom-row jacks can only be broken by inserting a plug into the lower-row jacks. This allows for high impedance monitoring or meters to be bridged across the circuits without interrupting the audio feed.

Patch-panel jacks are usually wired to equipment outputs and inputs via termination blocks. The blocks may be mounted together in the studio or in a central point in the engineering area. Wiring changes can be made at this convenient point without pulling new wire through the station. The termination can be solder-type "christmas trees," wire wrap, Type 66 telephone punch blocks, or the newer types of punch blocks designed for audio wiring.

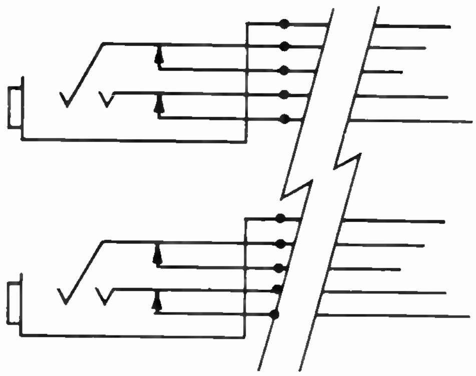
"Christmas trees" are popular with the engineers who trust only solder connections. They are sometimes found in remote trucks because of vibration concerns.

Most engineers have switched to punch-type terminations in which insulated wire is punched into a slotted connector with a tool designed for the purpose. The insulation is stripped away or displaced as the wire seats in the connection. Soldering is not required if the manufacturer's instructions are followed. The "66" blocks are built for single-conductor wire and were designed for telephone service. Stranded wire can be used if the connection is tested. Sometimes insulation may not be fully removed as the stranded wire is punched down.

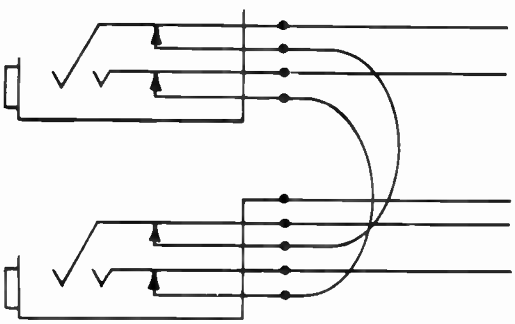
The safest approach would be to use punch blocks designed for the stranded wire used by the broadcast industry.

Buying prewired patch panels is almost always more cost-effective than buying the components and wiring them in the station. Allowing for normal interruptions, wiring a 48-jack patch panel can easily take up to 16 hours of a technician's time.

BO Normals brought out to termination.



AB Normals wired at the bay using a short jumper.



TRB Normals are bridged so that top row may be used for non-interruptive monitoring.

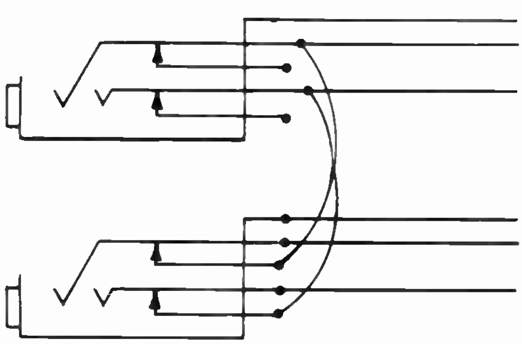


Figure 17. Wiring of normals.

Audio Routing Switchers

Replacing the patch panels in some applications is the audio routing switcher. This unit accomplishes the same function by switching the audio with relays or solid-state switches rather than by plugs and jacks.

The system has advantages in that it eliminates patch cords, can often be operated by remote control, and can often route audio to more than one feed at a time. Some of the more elaborate systems can be computer controlled. The advantage of using a routing switcher is greatly increased flexibility. Routing switchers are the only practical solution when many audio sources must be switched frequently, such as in a busy TV control room.

Routing switches are active devices which are subject to component failure and which add a small amount of noise and distortion to the system. The engineer should compare the specifications of different models before ordering one. An important consideration is what happens if there is a power interruption. Powering the routing switcher through an uninterruptable power supply (UPS) may be best.

The cost of a routing switcher is determined not only by quality but by its size. A stereo switcher with twelve inputs and twelve outputs will have 288 cross-points or switch-points (12 inputs × 12 outputs × 2 for stereo = 288). This can be pictured as two side-by-side matrices of 12 horizontal lines (inputs) intersected by 12 vertical lines (outputs). Each intersection is a cross-point where a connection could be made. One matrix would be for left-channel audio; the other, for right-channel audio.

In a television station, engineers must decide whether individual switchers are used for left, right, mono, SAP and time code or whether one large system may be used for all signals. With the larger-system approach, mono sources could be directed into left and right channels, channel reversals could be corrected and a mono mix could be created very easily. A switcher with sufficient cross points to handle all switching tasks will be more expensive than several smaller ones assigned to individual channels.

Distribution Amplifiers

When audio must be distributed to a number of locations on a continuous basis (without switching), a distribution amplifier (DA) is used. A typical application is sending a console output to several recorders, a patch panel and other studios, or distributing a satellite feed to all studios. A distribution amplifier (DA) eliminates the need for constant patching and switching as various pieces of equipment are needed. A DA is sometimes the only practical solution, as when audio must be fed in two directions on a constant basis (such as dual in audio processing chains in an AM/FM simulcast operation).

The typical DA has two stereo inputs and six to eight outputs for each channel. While there may be no input-level adjustment, it should have individual trim pots for each output. Some larger distribution systems will be modular, offering more versatility and avoiding unused outputs. One model offers four stereo inputs which can be assigned to any of its 14 stereo outputs by the use of jumpers.

Options on some larger DA systems include metering, input-level adjustments, audio compression, loss-of-signal alarms and redundant power supplies.

STUDIO MONITORS

The control room audio monitoring system is the first line of defense in spotting equipment failures and

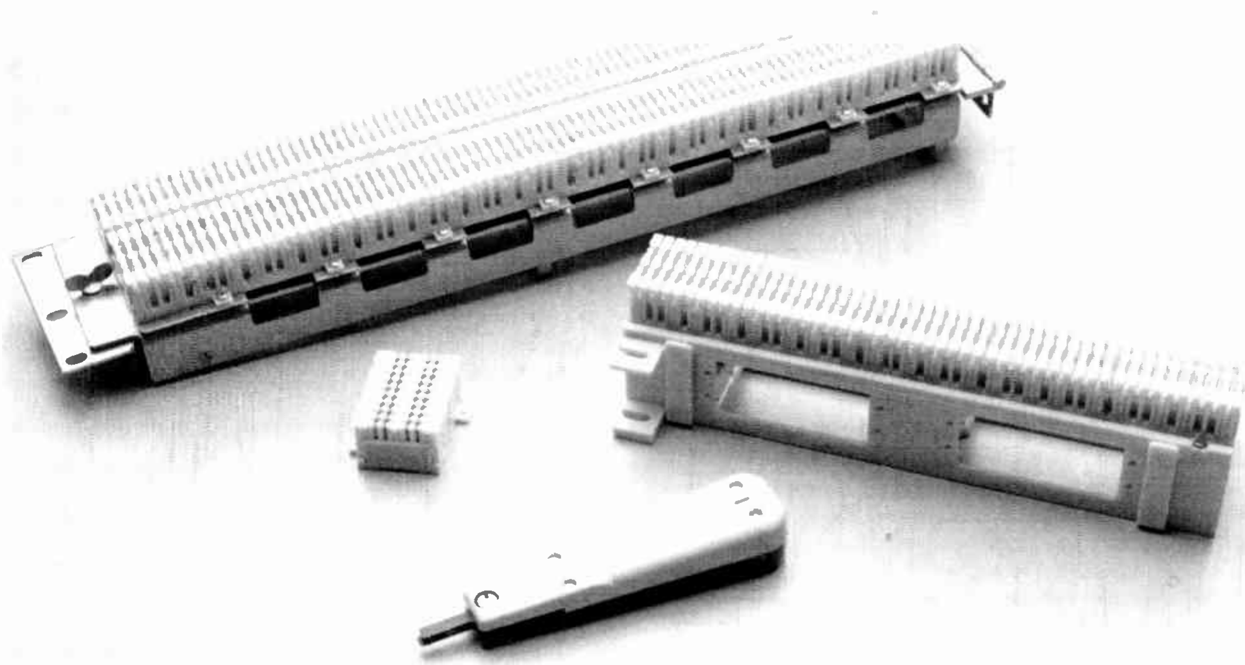


Figure 18. The Gentner FB-1 Punch Block. (Courtesy Gentner Electronics.)

Figure 19. ATI DA-10000 Distribution Amplifier. (Courtesy ATI, Inc.)

Figure 20. JBL 4412 Control Monitors. (Courtesy JBL.)

problems. For that reason, monitor speakers should be selected carefully.

In choosing monitors, room size will largely dictate the cabinet size. If the studio is large, use a monitor with 12-inch woofers, five-inch mid-range cones and horn or dome tweeters. Buy the best that the budget will allow. Small studios will be limited to a five or six inch woofer/mid-range and tweeter. Look for low distortion and flat response.

If at all possible, the studio designer should consider background noise sources, reverberation time of the room, interaction from walls and ceiling, and room equalization. To do this properly, involve testing the control room with a real-time analyzer and positioning the monitors for best results. This is seldom practical. When mounting the monitors on walls, use suspension mounts, preferably with vibration-isolating components. Position each monitor an equal distance from the normal operator's position, with the monitors directly facing that position. Use a sound-proofing material on as much of the flat wall surface in the room as possible.

In small studios, close-field monitoring may be the best choice. That is, simply place the monitors as close to the operator as practical. The theory is that the monitors will be closer to the listener than elements of the room environment which might affect the audio. The monitors could be set on a shelf above the console or suspended from the ceiling so that they are positioned directly in front of the operator at ear level or slightly above.

Another consideration is the power amplifier. Because of space and power requirements, audio console internal amplifiers are generally limited to eight to ten watts. The noise and distortion specifications may not be as good as those of commercial power amplifiers.

The power amplifier should be matched with the requirements of the monitors chosen. An underpowered amplifier will be pushed to clipping and distortion on peaks before a normal listening level is reached. This could damage the speakers. A better choice would be to run a more powerful amplifier conservatively. To prevent damaging the monitors with too much power, put fast-blowing fuses in the lines. Experiment with fuse values and listening levels to find the proper combination.

The wiring size between amplifier and monitors is important. Use at least #16 AWG for small amplifier-speaker combinations and use heavier wire for higher power or longer runs.

AUDIO SOURCES

Compact Disc (CD) Players

The CD player is now the audio source of choice in radio. CD technology encodes audio as digital bits which are recorded as etched holes on the surface of the disc. A transparent plastic coating protects the surface so that only a severe accumulation of dirt or scratches affect the playback quality. The bits are read by a laser beam focused on the spinning disc. Since nothing but the laser beam touches the disc, there is no wear.

The engineer should select the best equipment that the station can afford. Several manufacturers build CD players for broadcast and professional use. If the station must use nonprofessional consumer equipment, remember that consumer-grade equipment is designed for use in homes for a few hours a week. These machines will not last long when used almost constantly in a studio. Keep an extra machine in the station to replace the CD players when they fail. The output impedance and level of consumer machines are not the same as those of broadcast versions. If the console inputs are 600 ohms, use a matching interface. This

Figure 21. JBL Control 1 Monitors. (Courtesy JBL.)

will convert the −20 dB, high-impedance, unbalanced output of the consumer-grade CD player to 600 ohms, balanced at 0 dB for the console's input. Consumer equipment can also be difficult to cue and slow to start.

Turntables

Turntables are still found in radio stations, but their importance has greatly diminished. Some source material is only available on LP and some programs are provided to radio stations on vinyl.

Turntables are of two types, idler-wheel drive and direct drive. The once-common broadcast turntable of the idler-wheel drive type used a motor which turned at 1,800 RPM. The idler wheel was driven by the motor and, in turn, drove a hub near the center of the platter. This design minimized wow and flutter caused by fluctuations in motor speed. Further, speed stability was achieved by the use of a heavy platter. This system required a heavy, powerful motor to achieve the rapid starts necessary for tight cueing.

The direct-drive turntable is more popular because of its reduced noise, wow and flutter. Their speed control circuits constantly monitor and adjust speed, keeping it more accurate than if left to line voltage and

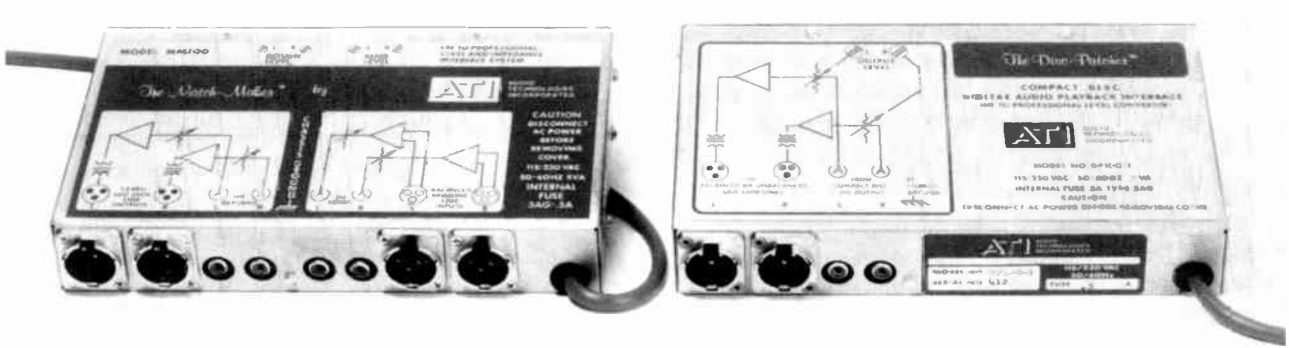


Figure 22. ATI MM-100 & DP-100 Audio Interfaces. (Courtesy ATI.)

frequency. This makes "speed enhancement" an exact science, enabling the operator to select an exact speed. Since this is an electronic system, there is no idler wheel to replace or bearings to lubricate. Circuit repairs may be a problem because the documentation provided with most direct drive turntables is sparse at best.

There are relatively few true broadcast tone arms available; most are designed for consumer use. These arms track quite well and are easily adjusted but are not rugged and are difficult to cue. The ideal broadcast arm is adjusted once for its cartridge and then left alone except for occasional testing.

When installing a tone arm, use the template from the turntable manufacturer and follow the instructions supplied with the arm. The tracking weight should be set as specified by the cartridge manufacturer. The choice of phono cartridge depends on the quality of its output. The less-expensive models are generally rugged and give the greatest life in on-air use. Better separation and high frequency response may be gained by moving up to more expensive, but less rugged, models. Avoid the consumer products whenever possible.

The stylus should be inspected daily and cleaned regularly. It should be replaced if it shows wear or if a defect is suspected. Keep a supply on hand.

The turntable's preamplifier is easy to neglect because it is mounted out of sight and never seen after installation. The important specifications of a pre-amp are noise and distortion and manufacturers express their figures in many ways. Noise will be referenced to different output levels and stated in decibels or RMS voltage. The more expensive models offer filtering, high-frequency cut or boost, and adjustable cartridge loading.

The most pressing concern for the engineer is the preamp's resistance to RF interference. These high-gain amplifiers are one of the first places that RF will make itself known if the studio is located at the transmitter site. The general rule that you get what you pay for applies here also.

Microphones

The choice of a studio microphone, unfortunately, is quite often left to chance. Dynamic microphones are the most popular for studio use. They are rugged, dependable, do not require power and are affordable. Ribbon microphones used in the past were heavy and delicate; but, with new technology, they are making a comeback. Condenser microphones are best known as recording studio equipment, but they are often found in radio production studios as well. Wireless microphones are handy for remotes and for television work. The wireless systems use either a miniature lavaliere mike and beltpack transmitter, or a hand-held design with the transmitter built into the case. Shotgun microphones are popular with TV news crews and on TV studio sets.

Chapter 5.1, "Microphones," deals with current microphone technology, and the engineer should refer to this chapter for advice on proper use and selection of a studio microphone.

Audio Cart Machines and Reel-to-Reel Recorders

For the past thirty years the cart machine has been the mainstay for playing commercials, jingles, and music. Stereo versions use three audio tracks on an endless loop of tape. Two of the tracks are used for left and right audio, and the third carries cue tones. A mono machine has two tracks, one for program and the lower track for cueing. Recorded carts cannot be traded between mono and stereo players, because the tracks will not line up with the different head formats.

In the NAB cart format, all tracks are equal in width. Another format has been developed in which the audio tracks are twice as wide as the cue track, providing better signal-to-noise specifications. The cue track width is minimal because its audio quality is not critical. This format is popular at stations which put their music library on cart.

A 1000 Hz tone is used as a cue signal to stop the tape after it has completed its loop in the cartridge and returned to the beginning of the audio. This stop signal is generated when the start button is pushed with the cart recorder in the record mode. Machines with optional features have secondary and tertiary tones for cueing and starting the next tape. The secondary or "aux" tone is at 150 Hz and is customarily used to trigger the next event in the program sequence automation systems. The tertiary tone is at 8 kHz and is typically used to warn air talent that the program material on the cart is coming to an end. These tones are inserted manually by the operator when the cart is recorded.

The reel-to-reel recorder is still a workhorse in radio studios. In the control room, it is used to record news feeds from networks and reporters. FM stations record music requests and contest winners on reel-to-reel for delayed playback. Some program material is still supplied on reel-to-reel tape. The control room machine is typically two-track stereo deck with speeds of 7½ inches per second (ips) and 15 ips.

The prime use for reel-to-reel machines is found in the production studio. Tape is edited by marking the edit points with a grease pencil and cutting the tape in a splicing block with a razor blade. The unwanted tape is discarded and the two edit points are spliced together. Words can be cut out and edit points tightened using this method. This method works well with mono or two-track stereo formats. If a multi-track machine is used, the operator should remember that he is cutting all tracks on the tape.

When using multi-track machines, the typical procedure is to record elements of the production on different tracks, and then mix all tracks down to a stereo mix as the production is transferred to cart. This makes adding tags or reading copy into "doughnut" tapes much easier. The music bed or agency tape is laid down on a stereo track and the copy is recorded on a third track. If a mistake is made, the voice track may be rerecorded. To allow this, multi-track recorders have a selective synchronization feature which switches the record heads of tracks not in the "Record" mode to the playback amplifier in order to synchronize all tracks.

Tape speeds of 15 ips and 30 ips are used in production work to give the best possible audio quality. Faster speeds make cut-and-splice editing easier because the audio is spread over a longer distance on the tape.

The reader should refer to Chapter 5.5, "Audio Magnetic Recording," on audio recording for more information on cart machines and reel-to-reel recorders.

DIGITAL AUDIO SYSTEMS

Digital Audio Storage

The most rapidly developing technology in the audio field is digital audio storage and editing. New products are being developed rapidly, and the engineer should study the products available before making an investment in digital equipment.

The most conventional form of digital audio storage uses ¼-inch or ½-inch audio tape recorded on an open-reel transport with stationary heads. The machine looks very much like an analog reel-to-reel recorder. Tape speeds vary from 7.5 ips to 30 ips and formats allow from two to 48 tracks. Sampling rates generally available are 44.1 kHz for CD mastering and 48.0 kHz for professional audio work. One model even features 88.2 kHz and 96.0 kHz sampling rates. Audio Engineering Society/European Broadcasting Union (AES/EBU) interface is standard equipment. Society of Motion Picture and Television Engineers (SMPTE) time-code reader/generators are available. Sixteen to 20-bit quantization is found on recorders now in production, and a dynamic range of 90 dB to 120 dB is possible.

This type of digital recorder offers the advantage of familiarity. Tape handling is identical to the analog reel-to-reel machines used for the past four decades. Since the tape format is open-reel, editing can be done with razor blade and splicing tape.

More common in the digital tape world is the rotating-head digital audio tape (R-DAT) machine. This format uses a rotating head much like video cassette recorder (VCR), and the tape is stored in a cartridge. Recording times are 60, 90, and 120 minutes. Razor blade editing is not possible, but time-code editing can be done with professional models.

Professional models feature fast cueing, instant starts, remote control, and time-code compatibility. DAT machines offer broadcasters the ability to record live events, concerts, and network feeds in the digital domain without sacrificing dynamic range and signal-to-noise ratio.

Of major interest in digital recording are systems which store digitized audio on hard drives, magneto-optical (MO) discs, or compact discs. By using bit-rate reduction techniques, system designers can greatly increase the storage capacity of hard disks. This technique discards bits which are not needed because they could not be heard by the human ear when reproduced. When decoded, the reduced-rate information reproduces an audio signal which is indistinguishable from the original even to the critical ear. Without bit-rate reduction, a 40-megabyte hard drive could store only four minutes of stereo audio. With bit-rate reduction the same drive can store 30 minutes of audio.

Digital Editing and Work Stations

Digital audio work station systems store multiple audio tracks on disk or in random-access memory (RAM). They allow the operator to edit tracks individually and produce a finished product by using keyboard or mouse rather than grease pencil and razor blade. The audio waveform for each track is displayed on a screen, allowing visual as well as audible cueing and editing.

Inputs can be two or more analog channels (through A/D converters) and, if the source material is digital, the operator can use optional digital inputs. Outputs may also be analog, digital, or both. The most prevalent standard is AES/EBU. It is the common denominator found on all equipment. Of course the digital input standard must match the digital output of the equipment feeding it. Until an industry-wide standard is in place, it is essential to be sure that any two pieces of equipment can work together.

To produce a spot with a digital audio work station, the operator puts the audio tracks and music bed in the system memory, and uses the work station to move component sounds, adjust timing, edit tracks and finally complete a stereo mix. Because this is all done with software, the original material is never destroyed as it would be with razor blade editing. Mistakes are easily corrected and editing experiments can be safely and quickly done. The benefit to the station is faster and more creative production, using a process quite similar to the word processing with which written text is edited.

AUDIO ENHANCEMENTS

Equalization

Generally, engineers want audio-frequency response to be as "flat" or linear as possible. In the audio processing chain, however, an equalizer can give character to the station's sound. Low frequencies can be boosted to produce a heavy "thumping" bass and the upper mid-range can be boosted to add "brightness." In production, an equalizer can be used to clean up commercial tapes and to remove noise from program material as it is transferred to cart.

Equalizers are available in two basic types. The best-known is the graphic equalizer, in which the audio spectrum is divided into a series of bands represented by a group of rotary or slider controls on the front panel. The operator adjusts the controls to affect gain in each particular band. The graphic equalizer with slider controls is easiest for nontechnical people to use because the controls give a visual picture of the frequency response curve which is being produced. A graphic equalizer is generally used in the station's processing chain to tailor the air sound.

The parametric equalizer is more versatile in the hands of a trained operator. With a parametric equalizer, a specific frequency can be selected for elimination or boosting. The parametric equalizer also allows the adjustment of filter "Q" or selectivity. The operator can "notch out" a hum or buzz with a sharp "Q," or a broad curve can be created to boost or cut a band of frequencies. The typical equalizer will have three or four sections to cover the entire audio spectrum. A parametric equalizer is a valuable tool in the production studio.

Noise Reduction Systems

Noise reduction techniques are used to reduce or eliminate source noise of audio tape, records, and radio links. The audio is encoded, recorded on tape, and then decoded to reproduce the original audio with a lower level of system noise. Audio companding (compression/expansion) is used to reduce the dynamic range of the program material, keeping it further above the noise floor of the recording medium. These dual-ended systems are used in audio tape recording and on radio subcarriers.

Single-ended noise reduction systems do not use the encode/decode process. These are frequency-sensitive gating devices which eliminate all audio in certain bands when the signal falls below a fixed level. The target range systems are centered around the frequencies associated with tape hiss. The theory of operation is that anything below the threshold must be noise and should be eliminated.

Effects Generators

Digital effects generators enable production directors to produce special effects for creative production. Some devices allow easy pitch changes of voice and music for effects and to compensate for speed adjustments made to time tapes perfectly to 30 or 60 seconds. If the tape speed is varied more than two or three percent, the audio pitch must be corrected.

Effects generators can produce "phasing," "flanging," and "echo" for production and on-air use. If reverberation is used in program material, the amount mixed in should be about 20 dB down from normal program level; otherwise, the effect will be distracting. If compression or limiting is used, it should be mixed into the audio chain after the limiter output. Otherwise the percentage of the effect mix will vary with the operation of the limiter.

SUMMARY

In this chapter I have tried to present the radio and television engineer with the state of the art in equipment and the standards of engineering practice. In reviewing the chapter on studio audio from the previous NAB Handbook, I was surprised by how many changes had taken place in the hardware realm in just five years. We now stand at the brink of digital signal processing and digital audio storage. The standards which we have worked with since the infancy of broadcasting are now falling to new technology. Even the "standard" 600 ohm, 0 dB connection between audio equipment is being superceded by the AES/EBU standard. The A3M and A3F connectors are making way for DB25s and fiber optic cable.

The broadcast engineer can look forward to source material from digital audio storage devices mixed on a totally digital console. A D/A converter will be required on the monitor amp. The console may be a touch-screen controlled computer. The digital processing chain will be linked by fiber optics or a digital RF link to the transmitter site. A digital stereo generator will connect to a digital FM exciter. The transmitter site may even be an uplink and the listener may receive a digial signal via satellite.

The challenge for the industry will be "retooling" by the equipment manufacturers to provide the new technology in hardware form. We will see new names emerge and some of the industry leaders fade from vogue as they fall behind the pace of changing technology. Pressure will be on broadcast engineers also to educate themselves in order to provide a competitive product to their listeners and viewers.

BIBIOGRAPHY

Blenkinsopp, Jeff: Broadcast Engineering, "Using Close-Field Monitors," Intertec Publishing, Overland Park, KS, July 1989.

Bytheway, David: Broadcast Engineering, "Planning For An Audio Routing Switcher," Intertec Publishing, Overland Park, KS, July 1989.

Dick, Brad: Broadcast Engineering, "Control-Room Monitoring Systems," Intertec Publishing, Overland Park, KS, July 1989.

Ennes, Harold: AM-FM Broadcasting, Howard W. Sams & Co., New York, 1974.

Runstein, Robert: Modern Recording Techniques, Howard W. Sams & Co., New York, 1974.

Various Authors: Radio World, Industrial Marketing Advisory Services, Falls Church, VA., Various issues.

5.10 Transmission Audio Processing

Robert Orban
Orban, A Division of AKG Acoustics Inc., San Leandro, California

INTRODUCTION

Transmission audio processing is both an engineering and artistic discipline. The engineering goal is to make most efficient use of the signal-to-noise ratio and audio bandwidth available from the transmission channel while preventing its overmodulation. The artistic goal is set by the organization using audio processing. It may be to avoid audibly modifying the original program material at all. Or it may be to create a distinct "sonic signature" for the broadcast by radically changing the sound of the original. Most broadcasters operate somewhere in between these two extremes.

Provided that the transmitted signal meets regulatory requirements for modulation control and RF bandwidth, there is no well defined "right" or "wrong" way to process audio. Like most areas requiring subjective, artistic judgement, processing is highly controversial and likely to provoke highly opinionated arguments between its practitioners. Ultimately, the success of a broadcast's audio processing must be judged by its results—if the broadcast gets the desired audience, then the processing must be deemed satisfactory regardless of the opinions of audiophiles, purists, or others who consider processing an unnecessary evil.

One mark of the professionalism of a broadcast engineer is his or her mastery of the techniques of audio processing. The canny practitioner has a bag of tricks that can be used to achieve the processing goal specified by the station's management, whether it is purist or "squashed against the wall."

FUNDAMENTALS OF AUDIO PROCESSING

Compression

Compression reduces dynamic range of program material by reducing the gain of material whose average or root mean square (RMS) level exceeds the *threshold of compression*. The amount by which the gain is reduced is called the *gain reduction* (G/R).

Above threshold, the slope of the input/output curve is the *compression ratio*. Low ratios provide loose control over levels, but generally sound more natural than high ratios, which provide tight control.

The *knee* of the input/output level graph can show an abrupt transition (*hard-knee*) into compression, or a gradual transition (*soft-knee*), in which the ratio becomes progressively larger as the amount of gain reduction increases.

The *attack time* is, generally, the time that it takes the compressor to settle to a new gain following a step increase in level. There is no generally agreed upon precise definition on how to measure attack time. Some measure it as the "time constant"; the time necessary for the gain to achieve 67% of its new value. Others measure it as the time for the gain to reach 90% of its new value for a given amplitude step (often 10 dB).

The *release time* is the time necessary for the gain to recover to within a certain percent of its final value after the level of the input signal to the compressor has been reduced below the compression threshold. It is sometimes convenient to specify the release time in dB per second if the shape of the release time is a straight line on a dB vs. time graph. However, this shape often is not linear. *Multiple time constant* (sometimes called automatic) release time circuits change the release rate (in dB/second) according to the history of the program, and according to how much gain reduction is in use. For example, the release time will temporarily speed up after an abrupt transient, to prevent a hole from being punched in the program by the gain reduction. The release time may slow down as 0 dB gain reduction is approached to make compression of wide-dynamic-range program material less obvious to the ear.

Delayed release holds the gain constant for a short time (typically less than 20 milliseconds) after gain

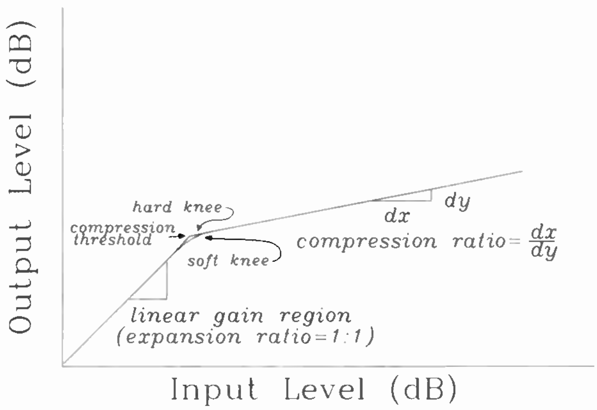


Figure 1. Input vs. output levels for compressors.

reduction has occurred. This prevents fast release times from causing modulation of individual cycles in the program waveform, thus reducing the tendency of the compressor to introduce harmonic or intermodulation distortion when operated with fast attack and release.

Expansion

Expansion increases the dynamic range of program material by reducing gain when the program level is lower than the *threshold of expansion*. The primary purpose of expansion is to reduce noise, either electronic or acoustic. Expanders are often coupled to compressors so that low-level program material is not amplified, thus reducing the noise that would otherwise be exaggerated by the compression. Expanders have attack times, release times, and expansion ratios that are analogous to those for compressors.

Peak Limiting and Clipping

Peak limiting is an extreme form of compression characterized by a very high compression ratio, fast attack time (typically less than 2 milliseconds), and fast release time (typically less than 200 milliseconds). In modern audio processing, a peak limiter, by itself, usually limits the peaks of the *envelope* of the waveform, as opposed to individual instantaneous peaks *in the waveform*. These are usually controlled by *clipping*. As a matter of good engineering practice, peak limiters are usually adjusted to produce no more than 6 dB of gain reduction to prevent offensive audible side-effects.

The main purpose of limiting is to protect a subsequent channel from overload, as opposed to compression, the main purpose of which is to reduce dynamic range of the program.

Peak clipping is a process that instantaneously chops off any part of the waveform that exceeds the *threshold of clipping*. This threshold can be either symmetrical or asymmetrical around 0 volts. While peak clipping can be very effective, it causes audible distortion when overused. It also increases the bandwidth of the signal by introducing both harmonic and intermodulation distortion into its output signal. Manufacturers of modern audio processors have therefore developed various forms of *overshoot compensation*, which is essentially peak clipping that does not introduce significant out-of-band spectral energy into its output.

Radio-frequency clipping (RF clipping) is peak clipping applied to a single-sideband RF signal. (A typical carrier frequency is 1 MHz). All clipping-induced harmonics fall around harmonics of the carrier (2 MHz,. . .). Upon demodulation, these harmonics remain at high frequencies and are removed by a low-pass filter. Thus RF clipping produces only intermodulation distortion, and no harmonic distortion. Ordinary audio-frequency (AF) clipping produces both. RF clipping is substantially more effective than AF clipping on voice because intermodulation distortion is considered less objectionable than harmonic distortion in this application. On the other hand, RF clipping is considered much more objectionable than AF clipping on music.

The *Hilbert transform clipper*[1] combines the features of RF and AF clippers. It acts as an RF clipper below 4 kHz (the region in which most voice energy is located), and acts as an AF clipper above 4 kHz to prevent excessive intermodulation distortion with music.

Unless a limiter has an attack time of less than about 10 μs, it will exhibit *overshoots* at its output. If the goal of the processing is to precisely constrain the instantaneous values of the waveform to a given threshold, it usually sounds best to control these overshoots by a limiter with 2 ms attack time followed by a clipper. Attempting to provide all peak control with the limiter does not sound as good, because the clipper affects only the offending overshoot and does not apply gain reduction to the surrounding signal.

When used in this way, clippers can cause audible distortion on certain program material. However, fast-attack limiters will cause audible clipping of the first half-cycle of certain program material, such as solo piano, harp, and nylon string acoustic guitar. Such distortion can be eliminated by a *delay line limiter*.[2] This device consists of two audio paths. The audio is applied to a pilot limiter, which has a very fast attack time. The gain control voltage generated by the pilot

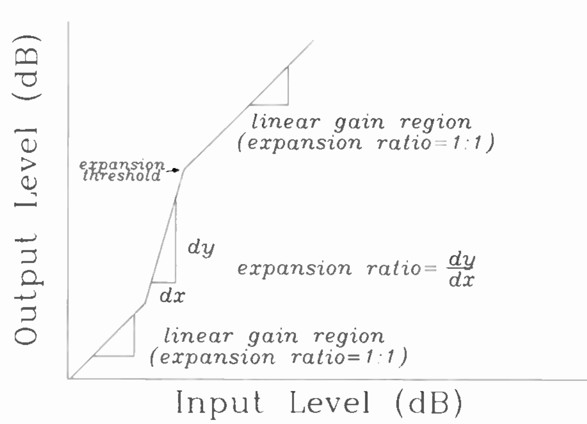


Figure 2. Input vs. output levels for expanders.

limiter is applied to a low-pass filter to smooth its sharp edges, and then to the gain control port of the variable gain amplifier that passes the actual program signal. To compensate for the group delay in the control voltage low-pass filter (which delays application of gain control to the audio), the audio is delayed equally by a delay line prior to the variable gain amplifier's input.

Gating

There are two fundamental types of gates, the *compressor gate* and the *noise gate*. The *compressor gate* prevents any change in background noise during pauses or low-level program material by freezing the compressor gain when the input level drops below the *threshold of gating*. Because it produces natural sound, it is very popular in broadcasting.

Many compressor gates will, instead of freezing, cause the gain to move very slowly to a nominal value (typically 10 dB of gain reduction) if the gating period is long enough. This prevents the compressor from getting stuck with an unusually high or low gain.

The *noise gate* is an expander with a high expansion ratio. Its purpose is to reduce noise. Because it causes gain reduction when the input level drops below a given threshold, the ear is likely to hear the accompanying gain reducing as a fluctuation in the noise level; sometimes called "breathing." This can sound unnatural. Therefore, the noise gate is most useful when applied to a single microphone in a multi-microphone recording. Usually, the other microphones will mask any "breathing," yet the noise reduction provided by the noise gate will still be appreciated during quiet program material.

Multiband Compression and Frequency-Selective Limiting

These techniques divide the audio spectrum into several frequency bands and compress or limit each band separately (although some interband coupling may be used to prevent excessive disparity between the gains of adjacent bands). This is the most powerful and popular contemporary audio processing technique, because, when done correctly, it eliminates spectral gain intermodulation. This occurs in a wideband compressor or limiter when a voice or instrument in one frequency range dominates the spectral energy, thus determining the amount of gain reduction. If other, weaker, elements are also present, their loudness may be audibly and disturbingly modulated by the dominant element. Particularly unpleasant effects may occur if the dominant energy is in the bass region, because the ear is relatively insensitive to bass energy, so the loudness of the midrange is pushed down by the dominant bass energy seemingly inexplicably.

Another type of frequency-selective limiting uses a program-controlled filter. The filter's cutoff frequency, its depth of shelving, or a combination of these parameters, is varied to dynamically change the frequency response of the transmission channel. Such program-controlled filters are most often used as *high-frequency limiters* to control potential overload due to pre-emphasis in pre-emphasized systems like FM (VHF), and television audio (NTSC and PAL), and in FM modulated transmission channels such as microwave links and satellite circuits.

Equalization

Equalization is changing the spectral balance of an audio signal, and is achieved by use of an *equalizer*. In broadest terms, an equalizer is any frequency-selective network (filter) placed in the signal path. In audio processing, an equalizer is usually a device that can apply a *shelving* or *peaking* curve to the audio.

A shelving curve starts off at a certain gain. As frequency changes, the gain increases (boost) or decreases (cut) asymptotically. Finally, the gain shelves off and does not change with further changes in frequency.

A peaking curve is bell-shaped on the frequency axis. As opposed to a shelving curve, it has a well-defined peak frequency. The shape of the curve can be uniquely defined by three *parameters: the amount of equalization* (in dB), *the frequency of maximum equalization* (in Hz), and the "*Q*," which is a dimensionless number that describes whether the curve is broad or sharp.

A *parametric equalizer* provides several peaking equalizers, in which the user has control of all three parameters. This type of equalizer is generally considered to be the most flexible and musical-sounding equalizer. Some parametric equalizers can also be used as notch filters.

A *graphic equalizer* provides a number of peaking equalizers (usually 8 to 31) distributed on octave or fractions of octave (¼ or ⅓) spaced frequency centers throughout the audible range. The controls for the amount of equalization are linear-throw faders, and are arranged on the panel in order of frequency. The positions of the controls, when considered together, thus provide a very rough graphic display of the amount of equalization provided by the entire equalizer. The advantage of a graphic equalizer is that it is easy to understand and quick to adjust. Its primary disadvantage is lack of flexibility. Usually, only the amount of equalization is adjustable, the "Q" and center frequency being fixed. However, a few manufacturers make parametric equalizers with graphic-style controls. These provide the advantages of both types.

Low-pass and high-pass filters remove spectrum at the top and bottom of the audible range, respectively. They are usually used to remove unwanted high- or low-frequency noise, and can also produce special effects (like telephone simulation).

These filters come with their rate of cutoff fixed in multiples of 6 dB/octave. 12 dB/octave and 18 dB/octave are popular. In addition, the shape of the region around the cutoff frequency has considerable effect on the listening quality of such filters. Bessel (constant-delay) filters have a gentle transition into cutoff, and sound pleasant and musically neutral. Butterworth (maximally-flat magnitude) filters have a sharper transition into cutoff. They are more effective at removing noise than Bessel filters, but have a more colored listening quality.

Equalizers are sometimes used on-line in transmission to create a certain sonic signature for a broadcast. Any of the types above may be used. Commercial audio processors may include equalizers for program coloration, or for correcting the frequency response of previous or subsequent transmission links. Sometimes the various bands of a multiband compressor or limiter are used as an equalizer by adjusting the gains of the various bands to achieve the desired equalized frequency response.

Loudness

One of the main uses of audio processing is to increase perceived loudness within the peak modulation constraints of a transmission channel. Assessing the effectiveness of audio processing thus requires a means of measuring loudness. *Loudness is subjective: it is the intensity of sound as perceived by the ear/brain system.* No simple meter, whether peak program meter (PPM) or VU, provides a reading that correlates well to perceived loudness. A meter that purports to measure loudness must agree with a panel of human listeners.

There are three important factors that correlate to subjective loudness:

1. The spectral distribution of the sound energy. (The ear's sensitivity vs. frequency—the ear is most sensitive from 2 to 8 kHz. Sensitivity falls off fastest below 200 Hz.)
2. Whether the sound energy is concentrated in a wide or narrow bandwidth. (Loudness summation: for a given total sound power, the sound becomes louder as the power is spread over a larger number of "critical bands" (about ⅓ octave).)
3. The duration of the sound. A given amount of sound power appears progressively louder until its duration exceeds about 200 milliseconds, at which point no further loudness increase will occur.

Torick and Jones have published a paper describing a meter for measuring the loudness of broadcast signals.[3] The FCC did an informal validation of the results of this meter, and concluded that it was effective in assessing whether commercials in television were noticeably louder than the surrounding entertainment programming.[4]

Additionally, the independently developed loudness measuring methods of Stevens and of Zwicker have both become international standards. Unfortunately, at the time of this writing there are no commercially available loudness meters using the Jones & Torick, Stevens, or Zwicker methods. Practically, the relative loudness of various program segments must thus be judged subjectively. It is wise to use several auditors when doing this, and to ascertain that they all have normal hearing by conventional audiometric tests.

GENERAL PERFORMANCE REQUIREMENTS FOR TRANSMISSION AUDIO PROCESSORS

Peak Modulation Control

The audio processor must control the peak modulation of the RF carrier to the standards required by the governing authority, such as the FCC in the United States. In AM, this usually means that negative carrier pinch-off must not occur at any time because this would cause splatter interference into adjacent channels. In FM and television (NTSC and PAL[5]), the peak deviation of the carrier must be controlled so that the modulation monitor specified by the governing authority does not indicate overmodulation. Because the rules often permit the modulation monitor to ignore very brief overshoots, the instantaneous peak deviation might exceed the peak modulation as indicated on the modulation monitor.

The requirements for peak control and spectrum control tend to conflict, which is why sophisticated nonlinear filters are required to achieve highest performance. Applying a peak controlled signal to a linear filter almost always causes the filter to overshoot and ring because of two mechanisms: *spectral truncation* and *time dispersion*. One can build a square wave by summing its Fourier components together with correct amplitude and phase. Analysis shows that the fundamental of the square wave is approximately 2.1 dB than the amplitude of the square wave itself. As each harmonic is added in turn to the fundamental, a given harmonic's phase is such that the peak amplitude of the resulting waveform *decreases* by the largest possible amount. Simultaneously, the RMS value *increases* because of the addition of the power in each harmonic. This is the fundamental theoretical reason why simple clipping is such a powerful tool for improving the peak-to-average ratio of broadcast audio: clipping adds to the audio waveform spectral components whose phase and amplitude are precisely correct to minimize the waveform's peak level while simultaneously increasing the power in the waveform.

If a square wave (or clipped waveform) is applied to a low-pass filter with constant time delay at all frequencies, the higher harmonics that reduce the peak level will be removed, increasing the peak level and with it the peak-to-average ratio. Thus even a perfectly phase-linear[6] low-pass filter will cause overshoot. *There is no sharp-cutoff linear low-pass filter that is overshoot-free:* overshoot-free spectral control to FCC or CCIR standards must be achieved with filters that are embedded within the processing, such that the nonlinear peak controlling elements in the processor can also control the overshoot.

If the sharp-cutoff filter is now allowed to be minimum-phase,[7] it will exhibit a sharp peak in group delay around its cutoff frequency. Because the filter is no longer phase-linear, it will not only remove the higher harmonics required to minimize peak levels, but will also change the time relationship between the lower harmonics and the fundamental. They become delayed by different amounts of time, causing the shape of the waveform to change. This *time dispersion* will therefore further increase the peak level.

When a square wave is applied to a linear-phase filter, overshoot and ringing will appear symmetrically on the leading and trailing edge of the waveform. If the filter is minimum phase, the overshoot will appear on the leading edge and will be about twice as large. In the first case, the "overshoot and ringing" are in fact caused by spectrum truncation which eliminates harmonics necessary to minimize the peak level of the wave at all times; in the second case, the overshoot and ringing are caused by spectrum truncation *and* by distortion of the time relationship between the remaining Fourier components in the wave.

Other Considerations

Except as required to achieve very specific artistic goals (most notably in some major-market high-energy hit-music formats), the processed audio should be free from unnatural subjective side-effects, such as *pumping* (a sense that the gain is constantly and unnaturally changing—a characteristic side-effect of wideband compressors and limiters when driven heavily), *breathing* (audible pulling up of background noise, cured by a compressor gate), and *hole-punching* (a sudden drop in loudness after a program transient, caused by the transient's inducing a large amount of gain reduction which then does not decay quickly, and cured by multiple time-constant release time circuitry).

The processor must be packaged so that it is easy to operate and maintain, and can work in high RF fields without compromise.

The processor should have setup controls with enough versatility to enable the subjective effect to be readily tuned to the requirements of the broadcasting authority operating it. For mixed-format applications the processor may have several presets, selectable by remote control, that permit the operator to set the amount of compression, limiting, clipping, and other parameters to complement the program material being transmitted.

Ordinarily, the processor should be equipped with sufficient remote control facilities to enable it to be interfaced efficiently with modern, automated plants. Most of the required facilities are specific to the application: for example, medium wave (MW), short-wave (HF), FM (VHF), or television.

The processor should have sufficient metering to permit it to be easily set up with tones or program material. The metering should also provide operations and diagnostic capabilities. Metering usually includes input level, output level, and "gain reduction" (the amount of limiting or compression) occurring in each variable-gain stage.

Processing for Stereo

Processing for stereophonic transmission is similar to processing for monophonic transmission, except that two audio processing chains are used. To preserve stereo imaging, the gains of the left and right automatic gain control and compression circuitry must be identical. Conversely, experience has shown that fast peak limiting and high-frequency limiting circuits sound best when operated independently (without stereo coupling), because the ear does not perceive channel-imbalance-induced spatial shifts with these fast time constants. However, the ear can perceive the loudness if one channel's being modulated unnaturally by a dominant element in the other channel when the channels are coupled.

The gain of the coupled elements is determined by the requirements of the transmission service. In FM, the channel requiring the greatest amount of limiting determines the gain of both channels. The processor operates by sensing the higher of the left and right channels and determing the gain of *both* channels such that the higher channel does not exceed a given level at the processor's output.

In AM, the gain of both channels is controlled by sensing and controlling the level of their sum (L + R), because the envelope modulation represents the sum of the channels.

SYSTEM CONSIDERATIONS

Subjective Audio Processing Concepts

Loudness is increased by reducing the peak-to-average ratio of the audio. If peaks are reduced, the average level can be increased within the permitted modulation limits. The effectiveness with which this can be accomplished without introducing objectionable side effects (like clipping distortion) is the single best measure of audio processing effectiveness.

Density is the extent to which the amplitudes of audio signal peaks are made uniform (at the expense of dynamic range). Programs with large amounts of short-term dynamic range have low density; highly compressed programs have high density.

Compression reduces the difference in level between the soft and loud sounds to make more efficient use of permitted peak level limits, resulting in a subjective increase in the loudness of soft sounds. It *cannot* make loud sounds seem louder. Compression reduces dynamic range relatively slowly in a manner similar to "riding the gain." Limiting and clipping, on the other hand, reduce the short-term peak-to-average ratio of the audio.

Limiting increases audio density. Increasing density can make loud sounds seem louder, but can also result in an unattractive, busier, flatter, and denser sound. It is important to be aware of the many negative subjective side effects of excessive density when setting controls which affect the density of the processed sound.

Clipping sharp peaks does not produce any audible side effects when done moderately. Excessive clipping will be perceived as audible distortion.

Building a System

Combining several audio processors into a good sounding system is tricky because of headroom and time constant considerations. The device driving a given processor must be able to drive that processor into full compression or limiting. If the driving device (for example, distribution amplifier) runs out of headroom before full limiting occurs in the driven device, then that device cannot achieve its full capability.

Beware of interactions between their attack times and release times when cascading several processors. It is wise to start the system with the slowest device. This is usually a compressor or automatic gain controller (AGC) with slow attack and release times, and with a compressor gate to prevent noise breathing. Such a processor does not significantly increase the *density* of the audio; it simply does gentle gain-riding to ensure that following stages are driven at the correct level.

The slow AGC is often followed by a multiband compressor with moderate attack and release time. Correctly-designed multiband processors have these time constants optimized for each frequency band. The low-frequency bands have slower time constants than the high-frequency bands. This multiband compressor usually does most of the work in increasing program density.

The amount of *gain reduction* determines how much the loudness of soft passages will be increased (and, therefore, how consistent overall loudness will be). Our hypothetical system reduces gain with the broadband AGC and the multiband compressor. The broadband AGC is designed to control average levels, and to compensate for a reasonable amount of operator error. It is *not* designed to substantially increase the short-term program density; the multiband compressor and peak limiters do that.

Modern audio processing systems usually add other elements to the basic system described above. For example, it is not unusual to incorporate an equalizer to color the audio for artistic effect. The equalizer may be any of the types described earlier and is usually found between the slow AGC and the multiband compressor. The multiband compressor itself can also be used as an equalizer by adjusting the gains of its various bands.

Various low-pass filters are often included in the system to limit the bandwidth of the output signal to 15 kHz (for FM), 10 kHz (AM in NRSC countries), 4.5 kHz (AM in EBU countries, and shortwave worldwide), or other bandwidths as required by the local regulatory authority. The final low-pass filter in the system is almost always overshoot-compensated to prevent introducing spurious modulation peaks into the output waveform.

High-pass filters may be incorporated to protect the transmitter. This is particularly important in high-power AM and shortwave installations exceeding 100 kW carrier power.

A transmitter equalizer that corrects the pulse response of the transmitter is found on high-end AM processors.

Location of System Components

The best location for the processing system is as close as possible to the transmitter, so that the processing system's output can be connected to the transmitter through a circuit path that introduces the least possible change in the shape of the carefully peak-limited waveform at the processing system's output. Sometimes, it is impractical to locate the processing system at the transmitter, and it must instead be located on the studio side of the link connecting the audio plant to the transmitter. (The studio/transmitter link (STL) might be telephone or post lines, analog microwave radio, or various types of digital paths.) This situation is not ideal because artifacts that cannot be controlled by the audio processor can be introduced in the link to the transmitter or by additional peak limiters placed at the transmitter. (Such additional peak limiters are common in countries where the transmitter is operated by a different authority than that providing the broadcast program.)

In this case, the audio output of the processing system should be fed directly to the transmitter through a link which is as flat and phase-linear as possible. Deviation from flatness and phase-linearity will cause spurious modulation peaks because the shape of the peak-limited waveform is changed. Such peaks add nothing to average modulation. Thus the average modulation must be lowered to accommodate those peaks within the carrier deviation limits dictated by government authorities.

This implies that if the transmitter has built-in high-pass or low-pass filters (as some do), these filters *must* be bypassed to achieve accurate waveform fidelity. Modern processing systems contain filters that are fully able to protect the transmitter, but which are located in the processing system where they do not degrade control of peak modulation.

Where Access To The Transmitter Is Available

The audio received at the transmitter site should be of as good quality as possible. Because the audio processor controls peaks, it is not important that the audio link (STL) feeding the processing system's input terminals be phase-linear. However, the link should have low noise, flattest possible frequency response from 30 to 15,000 Hz, and low nonlinear distortion.

If the audio link between the studio and the transmitter is noisy, the audibility of this noise can be minimized by performing the compression function at the studio site. Compression applied before the audio link improves the signal-to-noise ratio because the average level on the link will be greater. If the STL has limited dynamic range, it may be desirable to compress the signal at the *studio* end of the STL. To apply such compression, split the processing system, placing the AGC and multiband compressor sections at the studio, and the peak limiter at the transmitter.

Where Access To The Transmitter Plant Is Not Available

In some situations, the organization originating the program does not have access to the transmitter, which is operated by a separate entity. In this case, all audio processing must be done at the studio, and any damage that occurs later must be tolerated. A peak limiter would, however, be used at the transmitter to provide protection against overmodulation.

If it is possible to obtain a broadband phase-linear link to the transmitter, the processing system at the studio location can feed the STL. The output of the STL receiver is then fed directly into the transmitter with no intervening processing. A *composite* STL (ordinarily used for FM stereo baseband) has the requisite characteristics, and can be used to carry the output of the processing system to the transmitter. However, the output of a typical composite STL receiver is at the wrong level and impedance to directly drive a typical transmitter (most of which require + 10 dBm into 600 ohms). Therefore, the transmitter must almost certainly be modified to make it compatible with the composite STL. Use of a composite STL has many ramifications, and the installation of the processing system at the transmitter may be less complicated.

Where only an audio link is available, feed the audio output of the processing system directly into the link. If possible, transmitter protection limiters should be adjusted for minimum possible action as the processing system does most of that work. Transmitter protection limiters should respond only to signals caused by faults or by spurious peaks introduced by imperfections in the link.

Where maximum quality is desired, it is important that all equipment in the signal path after the studio be carefully aligned and qualified to meet the appropriate standards for bandwidth, distortion, group delay, and gain stability, and that such equipment is requalified at reasonable intervals.

Requirements for Studio-Transmitter Links (STLs)

If the STL is prior to the audio processor, the STL's signal-to-noise ratio must be sufficient to pass unprocessed audio. This means that the SNR of the link must be better than the sum of the desired SNR of the transmitted signal plus the maximum gain of the audio processor plus about 6 dB (a useful rule-of-thumb). If the STL follows the audio processor, its SNR must be 6 dB better than the desired SNR of the transmitter signal.

To ensure that the STL does not distort the shape of the audio waveform (preventing introduction of overshoot into peak-limited waveforms applied to the STL input), the frequency response must be flat (±0.1 dB) throughout the operating frequency range. The group delay must be essentially constant throughout this range (deviation from linear phase $< \pm 10°$). Phase correction can be applied to meet the requirement at high frequencies.

At low frequencies, by far the best way to achieve the specification is to extend the −3 dB frequency of the STL to 0.15 Hz or lower and to eliminate any peaking in the infrasonic frequency response prior to the rolloff. Poor AFC-loop design in STL transmitters is all too common, and this is the most likely cause of low-frequency response problems. Such problems can be corrected by applying equalization prior to the STL transmitter that is complementary to existing low-frequency rolloff, such that the overall system frequency response rolls off smoothly at 0.15 Hz or below. This solution is far better than clipping the tilt-induced overshoots after the STL receiver because the clipping will introduce nonlinear distortion, while the equalizer is distortion-free.

For highest quality, the nonlinear distortion of the STL system should be less than 0.1% total harmonic distortion (THD) throughout the operating frequency range.

Transmission Levels and Metering

Engineers at the transmitter and the studio consider transmission levels and their measurements differently. Transmission engineers need to know the peak level of a transmission commonly measured by an oscilloscope. Studio engineers need to know the line-up (or reference) level of a transmission commonly measured by a VU meter (as the approximate RMS level) or by a peak program meter (as the PPM level). For details, see the standard established by the Institute of Electrical and Electronic Engineers, *Recommended Practice for Audio Program Level Measurement* (Doc G-2.1.2/13, 1988).

Metering

The VU meter is an average-responding meter (measuring the approximate RMS level) with a 300 ms rise time and decay time; the VU indication usually lags the true peak level by 8 to 14 dB.

PPM indicates a level between RMS and the actual peak. The PPM reading has an attack time of 10 ms, slow enough to cause the meter to ignore narrow peaks and lag the true peak level by 5 dB or more.

Transmission Levels

The transmission engineer is primarily concerned with the peak overload level of a transmission to prevent overloading. This peak overload level is defined differently, system to system. In tape, it is defined as the level producing the amount of harmonic distortion considered "tolerable" — often 3% THD at 400 Hz. In FM, microwave, or satellite links, it is the maximum-permitted RF carrier deviation. In AM, it is negative carrier pinch-off. In analog telephone transmission, it is the level above which serious crosstalk into other channels occurs, or the level at which the amplifiers in the channel overload. In digital, it is the largest possible digital word.

Studio Levels

The studio engineer is primarily concerned with what is commonly called the reference level, operating level, or line-up level. This line-up level aids studio engineers in providing adequate headroom between line-up level and the overload level of equipment to allow for the peaks that the meter doesn't indicate. In facilities that use VU meters, line-up level is usually at 0 VU, which corresponds to the studio standard level, typically +4 or +8 dBm. In systems that use PPM, line-up level may be at PPM 4 (for the BBC standard) or at the studio standard level (often +6 dBm).

Transmission-Link Limiting

Transmission-link limiting devices are sometimes used ahead of the transmission link to protect it from overload. (These links might be STLs, satellite uplinks, inter-studio digital links, and the like.) These devices are usually used below-threshold (that is, with no gain reduction) as protection limiters to control peak levels. They only produce gain reduction when abnormally high levels are applied to their input due to operator error or unforeseen level variations at the source. This is useful to transmission engineers concerned with overload, and as useful to studio engineers concerned with headroom. For the needs of both engineers, such a limiter's output must be adjusted to be at or slightly below the peak overload level of the transmission channel.

To properly match the studio line-up level to the transmission protection limiter, the desired headroom must be known. For example, assume that the transmission protection limiter produces 0 dBm at its output at 100% modulation of the transmission link. Further assume that the line-up level in a production facility is designed to allow 8 dB of headroom. The input attenuator of the transmission protection limiter would then be adjusted so that studio line-up tone produces −8 dBm at the output of the transmission protection limiter.

This assumes that the amplifier or other link between the studio and the input of the transmission protection limiter has enough headroom to drive the transmission protection limiter into gain reduction without clipping this link. The transmission protection limiter only protects a link connected to its output. In the previous example, if the transmission protection limiter provides 15 dB of maximum protection, the system prior to the transmission protection limiter requires 8 + 15 = 23 dB of headroom above studio line-up level. If the link is simply an amplifier, this should be achievable without difficulty if the absolute level of the studio line-up tone is chosen carefully. In our example, if the amplifiers in the system clip at +21 dBm, the level of the studio line-up tone can be no greater than −2 dBm (i.e., 23 dB below +21 dBm).

AUDIO PROCESSING REQUIREMENTS FOR MW AND HF BROADCAST TRANSMISSION

In amplitude-modulated services, reception is usually compromised by noise and interference and may be further compromised by acoustic noise in the listening environment (like the automobile). Thus the processor must compensate for noise (electrical and acoustic) and interference by reducing dynamic range. This is most readily done by multiband compression and limiting to achieve lowest peak-to-average ratio without processing-induced side effects.

The processor must provide absolute negative peak control to prevent AM carrier pinch-off, which would otherwise cause out-of-band emissions. Additionally, the processor must incorporate overshoot-free filtering to control the audio input spectrum to the transmitter, thus preventing out-of-band emissions and interference. The permissible radiated spectrum is usually specified by national (FCC) or international broadcast authorities (most notably the CCIR) so as to make most efficient use of available radio frequency spectrum.

The processor may also be equipped with a receiver equalizer that compensates for the poor frequency response of the typical MW or SW radio due to narrowband RF and intermediate frequency (IF) stages.

Transmitter Equalization

The processor may provide a transmitter equalizer to eliminate tilt, overshoot, and ringing in the transmitter and antenna. Accurate reproduction of the shape of the processed waveform requires that the transfer function between the audio input and the modulated RF envelope represent a constant delay (which may be any positive number or 0) at all frequencies contained within the audio input signal. Failure to meet this criterion can result in tilt, overshoot, and *ringing* in the modulated RF envelope. The cause of overshoot and ringing as spectrum truncation and time dispersion at the high-frequency end of the system bandpass were discussed earlier. Tilt, on the other hand, is caused by problems at low frequencies.

Fig. 3 shows the response of a 10 kW plate-modulated transmitter to a 50 Hz square wave. The transmitter causes the waveform to tilt, which increases peak modulation in both positive and negative directions. The magnitude of the transmitter's frequency response is essentially flat to 50 Hz; the problem is caused by infrasonic rolloff. This rolloff is equivalent to that of a high-pass filter and is minimum-phase, which

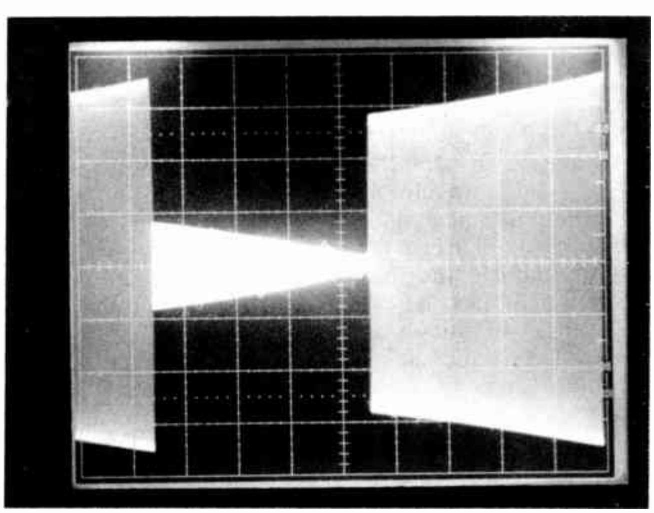


Figure 3. Tilt in plate-modulated transmitter.

introduces time dispersion, causing the shape of the waveform to change and further increasing the peak level.

Some transmitters contain high-pass filters at their audio inputs to protect high-power stages. *This location is absolutely inappropriate;* these filters can easily increase the peak-to-average ratio of the input audio by 3–4 dB. The correct location for a protection high-pass filter is in the audio processor where measures can be taken to prevent the high-pass filter from increasing the peak-to-average ratio at the audio processor's output.

Bounce

Linear errors can be equalized by predistorting the waveform in the audio processor. However, one major nonlinear error, commonly called *power supply bounce* caused by resonances in the LC filter elements of the transmitter's high voltage power supply. These resonances superimpose a sub-audible modulation onto the power supply voltage, resulting in a sort of very fast carrier shift that is too quick to be seen on a conventional carrier shift meter. The net result is to compromise the control of modulation peaks, particularly on strong bass transients which cause momentarily large current demands on the power supply, and which excite the resonance.

In some older transmitters, bounce has been known to compromise achievable modulation by up to 3 dB. Because bounce is not linearly related to the modulation, small-signal equalization cannot cure it. The most successful cure has been the use of a 12-phase power supply in the transmitter. The AC ripple from such a supply is down about 40 dB without filtering; a simple filter capacitor is all that is necessary to achieve adequate smoothing. Because there are no chokes in the power supply filter, resonance cannot occur. In all cases, bounce can be minimized by preventing excessive bass energy from being applied to the transmitter.

Slew Rate Limiting (Transient Intermodulation Distortion)

Transmitters using pulse-duration modulation (PDM) schemes are prone to problems with slew rate limiting. Because the PDM low-pass filter is located within the audio feedback loop of the transmitter, and because this filter is typically a multi-pole elliptic function filter with a cutoff frequency below 70 kHz, it will introduce very substantial delay into the feedback loop. This has two consequences: stability requires the amount of feedback applied around the transmitter to be limited, and it also requires that the open-loop gain of the modulator be rolled-off at a very low frequency. The first makes it difficult to design PDM transmitters with THD below 1%-2% at midrange frequencies, while the second renders "transient intermodulation distortion" (nonlinear behavior of the amplification stage prior to the frequency compensation stage) probable.[8] To minimize the probability that transient intermodulation distortion (TIM) will be bothersome, any amplification stage before the frequency compensation stage should be designed to be very linear to its clipping point, and to have sufficiently high headroom to accommodate the maximum rate of change to be expected at the transmitter's audio input.[9]

A transmitter can be qualified for TIM by one of the various difference-frequency intermodulation distortion tests. If the tests indicate that the transmitter has a low slew rate, it will not respond well to pre-emphasized audio and pre-emphasis will have to be reduced until the first derivative of the processed audio waveform seldom, if ever, exceeds the slew rate limit of the transmitter. Because of the benefits of pre-emphasis at the receiver, it is desirable to modify such transmitters to increase their slew rate, even if this means somewhat compromising harmonic distortion performance at low frequencies.

The NRSC-1 Audio Standard

As the North American AM band became more crowded, interference from first and second adjacent stations became more and more of a problem. Receiver manufacturers responded by producing receivers with decreased audio bandwidth, so that the encroachment of an adjacent station's modulation extremes would not be audible as interference.

This truncating of the bandwidth had the effect of diminishing the receiver's high-frequency response, but it was felt that lower fidelity would be less annoying than interference. To address these problems, the National Radio Systems Committee (NRSC) in 1987 formalized a standard for pre-emphasis and low-pass filtering for AM broadcast to provide brighter sound at the receiver while minimizing interference. See Chapter 3.1, "AM Transmitters," for more information on NRSC work.

AM Stereo Introduces a Pre-emphasis Dilemma

Certain AM receivers manufactured since 1984 for sale in North America, particularly those designed for domestic AM stereo reception, have a frequency response that is substantially wider than that of the typical mono AM receiver. The frequency response was widened largely to enhance the sales potential of AM stereo by presenting a dramatic, audible improvement in fidelity in the showroom. As these new receivers became more prevalent, broadcasters had to choose whether the station's pre-emphasis would be optimized for the new AM stereo receivers or for the existing conventional receivers that form the vast majority of the market.

If the choice was for conventional receivers (which implies a relatively extreme pre-emphasis), the newer receivers might sound strident or exceptionally bright. If the choice favored the newer receivers (less pre-emphasis and probably less processing), the majority of receivers would be deprived of much high-end energy and would sound "duller" and have less loudness.

NRSC Standard Pre-emphasis and Low-pass Filtering

In response to this dilemma, the NRSC undertook the difficult task of defining a voluntary recommended pre-emphasis curve for AM radio that would be acceptable to broadcasters (who want the highest quality sound on the majority of their listeners' radios) and to receiver manufacturers (who are primarily concerned with interference from first- and second-adjacent stations).

A "modified 75-microsecond" pre-emphasis/de-emphasis standard was approved (see Fig. 4). That provides a moderate amount of improvement for existing narrowband radios, while optimizing the sound of wideband radios. Most importantly, it generates substantially less first-adjacent interference than do steeper pre-emphasis curves.

The second part of the NRSC standard calls for a sharp upper limit of 10 kHz for the audio presented to the transmitter (see Fig. 5). This essentially eliminates interference to second and higher adjacencies. While some broadcasters believe that this is inadequate and that 15 kHz audio should be permitted, it is not likely that interference-free 15 kHz audio could only be achieved except by a reallocation of the AM band. The practical effect of widespread implementation of the 10 kHz standard is that 10 kHz radios will be feasible, and the bandwidth perceived by the average consumer (now typically limited by the receiver to 3 kHz) will be dramatically improved. On much mass-market consumer equipment, the difference between AM and FM reception will be less pronounced.

On April 27, 1989, The FCC released a Report and Order that amended section 73.44 of the FCC Rules by requiring all U.S. AM stations to comply with the occupied bandwidth specifications of the NRSC-2 standard by June 30, 1990. The NRSC-2 standard is an "RF mask" which was derived from the NRSC-1 audio standard by the NRSC. The purpose of the NRSC-2 RF mask is to provide a transmitted RF occupied bandwidth standard that any station with a properly-operating transmitter will meet, *provided that NRSC-1 audio processing is used prior to the transmitter*, and *provided that the station is not over-modulating*.

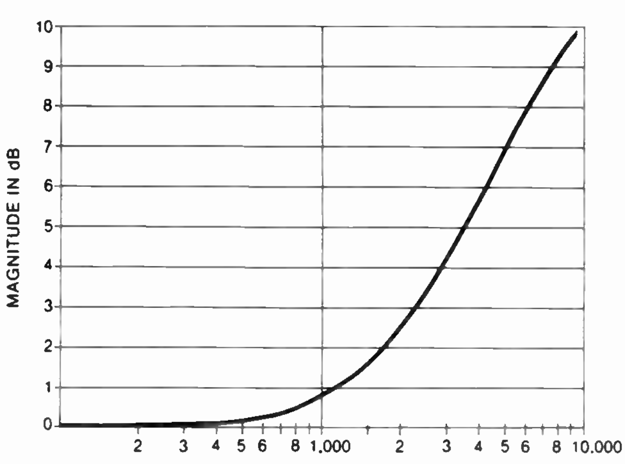


Figure 4. NRSC preemphasis curve.

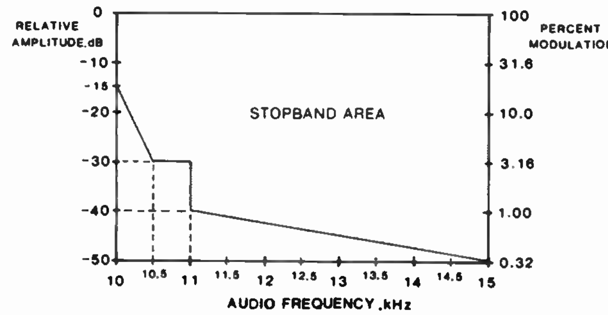


Figure 5. NRSC low-pass filter curve.

The Report and Order provides for "presumptive compliance" with the NRSC-2 occupied bandwidth standard: prior to June 30, 1994, any station whose audio complies with the NRSC-1 standard (an *audio* standard) is presumed to comply with the NRSC-2 standard (an *RF occupied bandwidth* standard) unless the station receives an Official Notice of Violation or a Notice of Apparent Liability from the Commission alleging noncompliance with the NRSC-2 occupied bandwidth standard.

The radio manufacturers participating in the NRSC stated emphatically that reduction in interference *must be demonstrated by broadcasters* before receiver manufacturers would be willing to release true wideband (10 kHz audio bandwidth) receivers to the mass market. This is rational—the receiver manufacturers can lose millions of dollars if they produce receivers that are rejected as noisy or interference-prone by consumers. In contrast, broadcasters can easily change pre-emphasis and filtering with relatively little expense.

Conformance with the NRSC standard, even if it were not required by the FCC, appears logical. It is likely that use of this more modest pre-emphasis and sharp 10 kHz filtering by AM broadcasters is the only factor that will eventually induce the receiver manufacturers to build and mass-market the high-fidelity, wideband radios that would allow AM stations to compete with FM in audio quality. The commitment to do so was strongly expressed by the receiver manufacturers involved in the NRSC's deliberations.

Audio Processing for AM Stereo

In all AM stereo systems, the envelope modulation is forced to a close approximation of the sum of the left and right channels to ensure compatibility with mono radios equipped with envelope detectors. To ensure minimum loudness loss compared to monophonic transmission, it is necessary to process stereo audio in the *sum and difference* format. This means that the left and right channel audio are passed through matrix circuits to create L + R (sum) and L − R (difference) signals. These signals are then passed separately through those parts of the processing that control modulation.

The C-Quam system and Kahn system (see Chapter 3.2 for descriptions of these systems) each have special individual requirements. To prevent clipping and distortion in the C-Quam decoder in the receiver, the negative-going modulation in the left and right channels must be no greater than −75% modulation (where 100% modulation is carrier pinch-off). Therefore, an audio processor for C-Quam must have both a sum and difference processor (which does the main processing) and a safety limiter to protect the left and right channels. This safety limiter is usually inactive, and typically only comes into play when the input program has sections which are momentarily single-channel, such as "ping-pong" stereo.

To generate the independent sideband wave, the Kahn system employs phase-shift networks ahead of its principal circuitry. These phase shift networks will severely distort the shape of peak-limited audio applied to their inputs. It is therefore necessary to split an audio processor for the Kahn system and to place the part of the processor that controls peak modulation after the phase-shift networks.

AUDIO PROCESSING REQUIREMENTS FOR FM (VHF) BROADCAST TRANSMISSION

The processor should provide a comfortably listenable dynamic range in domestic and automotive listening environments by applying subtle compression to the signal. Unless the program director requests otherwise for competitive reasons, such compression should be undetectable to the ear unless the original source is available for comparison.

The processor must provide high-frequency limiting to complement the pre-emphasis employed (50 μs or 75 μs, depending upon the Region in which the transmission occurs).

The processor must provide accurate peak control (as measured by a modulation monitor meeting the standards of the governing authority) in both the positive and negative directions. To ensure that absolute peak control will be retained at the system output, any system elements following the processor must have flat frequency response (±0.1 dB) and constant group delay (deviation from linear phase <±10°). Because the pre-emphasis networks and low-pass filters ordinarily found in stereo encoders do not meet these requirements, these should be bypassed. Thus, the processor should provide pre-emphasis and band-limiting for the transmission system. Its output must contain negligible energy above the bandwidth limit of the transmission system. In FM stereo broadcasting by the world-standard "pilot-tone" method, this bandwidth is limited to less than 19 kHz to prevent aliasing from the stereo subchannel into the main channel, and vice-versa. To protect the pilot tone itself (ensuring correct operation of the phase-locked loop subcarrier regeneration circuitry in the receiver's stereo decoder), the bandwidth must be further limited to no greater than 17 kHz. In practice, it is customary to begin the HF rolloff at slightly above 15 kHz to minimize group delay distortion in the low-pass filters used to effect the bandwidth limit. Nonlinear low-pass filters are usually used to prevent overshoot, enabling the processor to control peak deviation absolutely.

The processing system must be readily adjustable to achieve the subjective effect desired by the broadcasting authority operating it. To achieve a "competitive" sound in markets in which many stations compete for listeners, it may be necessary to add additional multiband limiting to the basic audio processing system (which usually consists of compressor, HF limiter, and peak limiter/clipper). Adding additional multiband limiting can create greater program density than the basic processing system alone without introducing spectral gain intermodulation.

AUDIO PROCESSING REQUIREMENTS FOR TELEVISION BROADCAST TRANSMISSION

The processor should provide a comfortably listenable dynamic range in domestic listening environments by applying subtle compression to the signal. Such compression should be undetectable to the ear unless the original source is available for comparison. Usually an available gain reduction range of 25 dB is adequate to handle the level variations encountered in typical operations.

The processor must provide high-frequency limiting to complement the pre-emphasis employed (50 μs or 75 μs, depending upon the Region in which the transmission occurs).

The processor should provide accurate peak control (as measured by a modulation monitor meeting the standards of the governing authority) in both the positive and negative directions. In general, the comments on FM (above) apply here as well.

The processor should control subjective loudness to prevent unpleasant inconsistencies when transitions occur between various program elements. This is most accurately achieved using technology similar to that developed for loudness measurement (discussed earlier). In essence, the processor uses a loudness meter in a servo loop to control loudness and ensure consistency of loudness between one program source and the next.

The processor should handle voice cleanly. The Hilbert-Transform clipper and delay-line limiter (see above) are effective for this, because neither creates audible clipping distortion on voice, even when the source is narrow-band (like optical film or telephone). Such narrow-band sources are extremely difficult for a conventional audio-frequency clipper to process without introducing some audible harmonic distortion on voice.

Audio Processing for Stereo Television

The general requirements for stereo television processing are not very different from the general requirements enumerated above. As discussed in the *Processing for Stereo* section earlier in this chapter, the processing elements with slow release time constants must be coupled to preserve the stability of the stereo image. In the North American BTSC system, the peak modulation criteria are complex. However, it can be shown that FM stereo-style processing will always prevent over-modulation in BTSC stereo, although it will not necessarily allow the most L + R modulation theoretically possible in this system. This style of processing is also appropriate for the other international stereo systems,[10] since it will always prevent over-modulation.

Because of the close proximity between the edge of the audio passband (approximately 15 kHz) and the stereo pilot tone (15.734 kHz), the BTSC system requires very sharp low-pass filters to prevent aliasing. It is impractical at the current state of the art to apply nonlinear overshoot compensation to these filters. These overshoots do not cause interference or problems in television receivers. Thus these overshoots must be accepted as inherent to the BTSC system, and must be ignored by modulation monitors designed as a reference for setting modulation levels. If these overshoots are not ignored, average modulation will be set too low and the viewer will experience annoying increases in loudness when switching from stereophonic to monophonic channels.

TECHNICAL EVALUATION OF AUDIO PROCESSING

Common swept frequency response, harmonic distortion, and intermodulation distortion tests are often used to evaluate audio processors. Therefore, it is useful to discuss why these tests may at times produce misleading results.

Definition of Linearity

A system can be tested for linearity as follows. Apply an input signal A to the system and measure its output. Let X be the output signal caused by input A. Then, remove A from the input and apply another signal B. Let Y be the output signal caused by the input B.

The system is linear if the following things happen: (1) If the input waveform is multiplied by a factor k to scale it, the output waveform also becomes scaled by a factor of k, but its shape is not distorted by the process of scaling. (2) If inputs A and B are applied to the system simultaneously, the system's output is X + Y (*superposition*).

It is clear that expanders, compressors and limiters are strongly nonlinear systems. The output of such a device is not scaled proportionally to its input; it is expanded or compressed. Similarly, when two signals are applied to such a device, its output is not the same as the sum of its response to either signal individually; superposition does not hold. Clippers are similarly nonlinear.

Sinewave Measurements and Nonlinearity

When predicting a system's response to program material by measuring its response to individual sinewaves, certain assumptions are made. The first assumption is that program material can be adequately represented as a sum of sinewaves (Fourier analysis). The second assumption is that superposition holds, so that the response of the system to single sinewaves also applies when several sinewaves are added together at the system's input. Thus, the sinewave results can be extrapolated to program material.

Because dynamic audio processing (compression, limiting, clipping, expansion, gating) is strongly nonlinear, the usual assumptions of superposition and scaling, which permit sinewave measurements to be extrapolated to complex program material through Fourier analysis, do not hold. Conventional harmonic and intermodulation distortion measurements, historically designed to measure slight departures from linearity in weakly nonlinear systems, are of very limited usefulness. Swept or spot frequency response measurements are not useful.

When making distortion measurements with tones, their relevance must be assessed psychoacoustically. Does the system output *sound* distorted when listening to the tones? For example, when measuring harmonic distortion using fundamentals in the 50–1000 Hz region, the higher harmonics are more significant than the lower harmonics because the higher harmonics are less readily masked by the desired fundamental. However, as the fundamental frequency is increased, the harmonics become less troublesome because the ear becomes less and less sensitive to them. Eventually, their frequency exceeds the passband of the system and they become irrelevant.

Similarly, SMPTE intermodulation distortion methods measures the level of 50 or 60 Hz sidebands around a high-frequency tone induced by system nonlinearity. Because these sidebands are within a single "critical band" (approximately ⅓ octave) of the high-frequency tone, they are maximally masked by it. Therefore rather high amounts of measured SMPTE IM distortion are not necessarily cause for concern. On the other hand, CCIF difference-frequency intermodulation distortion measurements measure the low-frequency difference tone caused by two high-frequency tones. Because the difference tone is far removed in frequency from the desired tones, it is not well-masked by them, and high amounts of CCIF IM are of some concern. See Chapter 7.1, "Broadcast Audio Measurement," for more information on distortion and measurements.

Subjective Listening Tests

There are few, if any, measurement techniques that can adequately predict whether the subjective effect of an audio processor will be satisfactory. The only effective way to evaluate nonlinear broadcast audio processing is by *subjective listening tests*. These must be done over a long time period, using many different types of program material, because a processor that sounds good on a certain type of program material may sound unsatisfactory on other program material having markedly dissimilar spectral balance or dynamics.

Usually, the subjective goal of broadcast processing is to have its action undetectable to the audience. In the case of processing in highly-competitive major market stations, some degradation of the program (as perceived on a high-quality monitor) is often accepted for the sake of maximizing "punch" and loudness. Moderate quality compromises are usually masked on smaller and lower cost radios and are noticeable only on higher-quality radios by critical listeners.

In all cases: *It is not appropriate to attempt to extrapolate the results of tone tests to program material, because superposition does not hold.*

ENDNOTES

1. R. Orban, "Increasing Coverage of International Shortwave Broadcast Through Improved Audio Processing Techniques," *Journal of the Audio Engineering Society*, June 1990.
2. British Broadcasting Corporation Engineering Division: "The dynamic characteristics of limiters for sound programme circuits," Research Report No. EL-5, 1967.
3. B.L. Jones and E. L. Torick, "A New Loudness Indicator for Use in Broadcasting," J. SMPTE, Sept. 1981, p. 772.
4. R.A. Haller, "An Update on the Technology of Loud Commercial Control," OST Technical Memorandum FCC/OST TM83-1, February 1983.
5. SECAM customarily uses AM sound, with the usual requirements for preventing carrier pinch-off.
6. A phase-linear filter has constant delay with frequency.
7. A minimum-phase filter has no zeros in the right half of the s-plane. As its name implies, there is no filter with the same magnitude response that can have less phase shift. Given the magnitude response of a minimum-phase filter, its phase shift can be computed (with the Hilbert Transform). This means that if a minimum-phase filter has constant group delay in its passband, this is associated with a certain type of magnitude response which rolls off gently around the filter's cutoff frequency: a minimum-phase filter with constant group delay in the passband cannot simultaneously have a highly selective magnitude response. Many textbooks provide the well-known mathematical details. See, for example, H.J. Blinchikoff & A.I. Zverev, *Filtering in the Time and Frequency Domains*, New York, Wiley, 1976, pp. 89–94.
8. Simply stated, almost all feedback systems contain a filter that forces the open loop characteristic to be either low-pass (all-pole; "lag compensation") or low-pass shelving, with poles *and* zeros ("lead-lag compensation"). Feedback forces the amplifier before this filter to present a pre-emphasized signal to the filter's input such that the total response of the system is flat. If the filter rolls off at 6 dB/octave starting at 15 Hz (a typical situation in an opamp like the TL072 or the LF353), this pre-emphasis *rises* at 6 dB/octave 15 Hz. High frequencies applied to this system will obviously challenge the headroom of the amplifier prior to the filter. For example, 20 kHz will be up 62.4 dB! If high frequencies drive the amplifier prior to the filter into clippling or substantially nonlinear operation, "transient intermodulation distortion" occurs. Because the clipping process is followed by a filter with a low-pass characteristic, harmonics generated by clipping will be de-emphasized, so difference-frequency IM tests are more sensitive than THD tests to this mechanism.
9. For a maximum audio bandwidth f, the required slew rate in percent modulation per microsecond is $0.0002\pi f\%/\mu s$. For 4.5 kHz, this is 2.827%/μs.
10. West German (dual-carrier); Japanese (FM subcarrier); English NICAM (block-companded digital).

5.11
Television Camera Robotics

Robert S. Murch
WPIX, New York, New York

Richard D. Slenker, Jr.
WTTG, Washington, District of Columbia

with contributions from
Bruce Levy
WTVH, Syracuse, New York

and edited by
Edmund A. Williams
Public Broadcasting Service, Alexandria, Virginia

INTRODUCTION

In recent years, there has been increased interest in the use of robotics for television cameras. In the past, remote control camera systems have been available utilizing simple pan and tilt control of camera position. With the application of high speed microprocessors and improvements in servomechanisms, robotic camera systems are now available at an affordable cost and level of sophistication which rival the performance level of a competent human camera operator in certain production situations. Performance and reliability have reached a point that a television station or production facility must seriously consider whether the use of camera robots can be an operational as well as financial asset to the facility.

Servomechanisms have been used in industrial settings for many years: from assembling automobiles to assisting in brain surgery. Their precision is well-known and the technology well-developed. It is only recently that robotics have been coupled to studio cameras in order to reduce operator costs, improve shot repeatability, and reduce errors. While camera robotics may not be a replacement for camera operators in fast-moving or dramatic productions where a high degree of operator flexibility and creativity is required, they are ideally suited for many less demanding productions.

News programs and productions requiring exact repetition of shots, such as commercials, are perfect applications for camera robotics. Moreover, most productions such as news and talk shows are scripted to a degree where use of camera robots can add to the quality of the presentation due to the smoothness and repeatability of shots.

CONSIDERATIONS FOR USING CAMERA ROBOTICS

An important reason for considering camera robotics is the potential for labor cost savings. Other operational considerations include the following:

- Repeatability of shots
- Scheduling
- Smooth execution of complex camera work
- Placing cameras in dangerous locations, or areas normally inaccessible to human operators ("point-of-view cameras")
- Reliability and durability

Labor Costs

The cost of camera robots vary for different manufacturers and different options. Consider the return on investment of a three camera robotic system that costs $200,000 to purchase and install, for illustrative purposes, and is depreciated over five years. Simply depreciated, the cost of the equipment is approximately $40,000 per year. If the system is operated by an existing video operator and if the station is employing three camera operators at a salary of $500.00 per week for each operator (or $26,000 per year, per operator) for a total of $78,000 per year, the equipment will pay for itself in about two and one-half years.

Looking at cost trade-offs another way, there is a savings of $78,000 per year in salaries (exclusive of benefits) less the $40,000 cost of capital for the equipment for a net savings of $38,000 per year. The return on investment is further improved with each additional shift of usage plus any savings of employee benefits.

Another pay-back calculation example: assume three camera robot systems are purchased at $100,000 each for a total of $300,000. At 10% simple interest and a three year payment plan with equal monthly payments, each payment is $10,833 or $130,000 per year. Further assume that a station or production facility has six full-time camera operators in two shifts at an average of $26,000 each plus 25% overhead for a total of $32,500 each or $195,000 per year. If four of the six operators are replaced (two from each shift) at $32,500 per year each, the savings is $130,000 per year or break-even for the first three years. One camera operator remains on each shift as the robot operator or to be on call for special production shots.

If salary or benefits is higher than those shown above or if the purchase or lease plan is longer than three years, actual cash savings can be realized immediately. A detailed analysis of the station production requirements versus the cost of staff and the costs associated with leasing or purchasing camera robots must be performed prior to initiating an investigation into replacing skilled studio personnel.

The increase in workload on the maintenance staff to care for the robots, including lubrication, adjustment, repair, testing, software checks and electronic service, must be factored into the decision to use robotics. However, the incremental increase can be expected to be relatively small.

The increase in additional production planning effort may add an extra few percent to the production costs. This is due to the time needed to set-up, rehearse, and store the positional information for each shot. However, once done for a given production, less time will be needed for future editions of the same program.

Scheduling

Camera robotics provide a high degree of scheduling flexibility as they are available 24 hours a day and are unaffected by holidays, vacation, illness, or overtime.

Repeatability of Shots

Camera robot systems have a degree of precision that provides highly accurate repeatability of all programmed camera shots and camera moves. Once a difficult shot is programmed, the system will execute the shot on-air exactly the same way it did in rehearsal. This factor is important when repeating scenes or to avoid accidentally shooting off-set, thus saving countless retakes or on-air mistakes. Camera robot systems also are designed to accommodate last moment changes and execute them with the same degree of accuracy as preprogrammed shots. From the program director's perspective, the repeatability of shots and the ability to quickly accommodate last minute changes provide an additional level of confidence that the production will occur as planned in rehearsal.

The servos that perform the movements of camera robots are able to position the pan-tilt head within 0.1°. The pedestal (dolly and trucking movements) will return to the same place in the studio to within a fraction of an inch and maintain the aim to within 0.1° using the tape guidance system or other guidance techniques. Slightly less accuracy can be obtained when the pedestal is operating in a "free" mode, that is off the tape or unable to obtain a reference. In any event, the accuracy is equal to or greater than what most human operators could produce.

Movement rates of the pan-tilt heads and pedestals are generally slower than what human operators can perform. This is due to, (1) the mechanical effort of the servo needed to move the heavy camera, prompter and pedestal, and (2) the desire to keep movements slow enough to prevent injury to someone in the way of the movement.

Studio productions may require dozens of different camera positions and up to 100 different shots by multiple cameras with varying angles, zoom positions, and framing differences. Camera robot systems have provisions to store hundreds of shots in memory (for longer, more complex productions or multiple shows) and can interface with zoom lenses, iris controls, camera control units, lighting systems, and virtually any other computer controlled studio facilities with external data ports (see section on interfacing below).

Smooth Execution of Complex Camera Work

Full-featured camera robots will execute virtually all the standard moves of similarly-equipped manual cameras including:

- Pan
- Tilt
- Zoom
- Focus
- Dolly (in and out)
- Trucking (left, right)
- Arc (curve left and right)
- Pedestal (up, down)

Each movement can be made at different and varying speeds and have gentle starts and stops.

The manner in which the transition from one shot to another occurs may be critical in some production operations. Some robotics systems allow shots to be linked together along with varying speeds of transition.

Camera robot systems allow an operator to control several cameras at the same time. For more artistic, creative productions, a director can plan a number of different, complex, simultaneous camera moves and program them into the system to occur in a particular sequence. When these moves are recalled, they will be replayed exactly as rehearsed, every time they are recalled. For productions which require repetitive shots, such as news programs, this ability relieves the boredom human operators suffer and the subsequent potential for error.

Some robot control systems will also link the zoom lens focal length and focus controls to the camera movements so that as desired objects within the field of view change distance, the focus will track as well.

much as a human operator would do. However, it should be noted that, if the mechanical system of the zoom lens is operating unevenly (sticky), the robotics system cannot correct such a defect.

Another caution is that robots will only perform the tasks that have been given to them, exactly as given. That is, if a scene that is changed after rehearsal, a human operator can make an instant and creative decision on adjusting shots while the robot cannot automatically make the on-air adjustment.

Compared with productions with human operators, more advanced planning may be required on the part of production personnel in setting the shots and entering them in the robot's memory. For example, the robot will not automatically focus the lens on a scene but will follow orders to change focus only if the operator programmed it to do so in advance.

Operation of Cameras Located in Inaccessible Places

Advances in robot technology and the use of the smaller, high quality color television cameras make it possible to obtain shots from places and angles never previously thought achievable. Camera robots allow great flexibility for those special "point-of-view" shots that give the audience fascinating and informative views. Examples of these locations include studio ceilings, on race cars, roofs of sports stadiums, tight spots along parade routes, and masts of microwave trucks.

Reliability And Durability

The reliability of any system is dependent upon the reliability of the individual components and subcomponents that make up the system. In the case of camera robotics, the principal components are the servomechanisms and the computer control systems. Servos, by their very nature, are very reliable and, if operated at less than their rated loads, will last indefinitely. However, in order to achieve a high level of reliability, a regular program of inspection, cleaning, lubrication, and preventative maintenance must be instituted for the mechanical parts of the robot system.

Servos are designed to withstand a certain amount of physical abuse. Some servos have built-in clutches on motors or gear trains to prevent stalling or stripping of the gears. Others can be stalled without burnout. Most will provide an indication of inability to react to commands, such as when an obstruction prevents a pan-tilt head or pedestal from moving to the desired position.

Servomechanisms are designed to operate in relatively hostile environments such as the high heat and humidity of summer, and in cold winter conditions. They are, therefore, suitable for use in locations that human operators would find uncomfortable. The range of conditions in which the servos can safely operate is likely to exceed the conditions acceptable for the camera itself.

The computer systems of most camera robots use PC components. The normal lifetime of most of these parts is on the order of several years. Little maintenance is required other than the usual routine inspection and cleaning. However, if mechnical memory devices are used (hard and floppy disk drives), an occasional check should be made of the disks using standard software available for this task.

Because most camera robots are likely to be operated 16 to 20 hours a day, and may in fact be left on 24 hours a day, the disk drives are the most likely parts of the computer to fail. A back-up drive or comoputer is recommended to help bridge the time when such a failure occurs.

INSTALLING CAMERA ROBOTICS

The major installation issues to be addressed are: (1) the impact installing robots will have on the staff, (2) potential locations for camera robots, (3) the smoothness of the studio floor, and (4) the location of the robot controls.

Personnel Considerations

The installation of camera robots can result in more effective utilization of production personnel. How management deals with the staff, who may believe that robots will replace human operators, may determine the level of success of the installation more than any other factor. Depending upon individual situations it may be possible for reassignment or attrition to help adjust staffing arrangements.

In addition to the changes in the technical and production staff due to the installation of robots, other departments in the station must also be directly involved in the project.

The use of camera robots will affect how a production is designed and executed. All production personnel, including producers, directors, and other production managers throughout the facility, must be aware of the advantages and the limitations of using robots. Initially, this awareness could begin with the engineering director providing information about robots and how they can improve production and reduce production costs. Later, after the staff has become used to the idea of using robots, production staff can be trained on how best to use the camera robots. After the initial introduction of the robotic system, it is essential for engineering and production departments to work jointly to develop daily standard operating procedures for the new robots.

Potential Locations for Camera Robots

While the studio is the obvious first choice for using robotic cameras, it is by no means the only location. Newsrooms are popular sites for robots where the camera is not required to be as full-featured as in the studio. Newsroom cameras may not need to be mobile and simpler robots (equipped only with pan, tilt, zoom and limited side-to-side travel, for example) may be all that is required. In any event, when considering installation of camera robots, carefully study how the camera will used. Consult with other stations that have installed robots and review the performance limitations of the system under consideration to insure that it meets the needs of the facility. While it is relatively easy to have an operator move a camera to a more suitable location, moving the robot may not be possible without considerable time and expense.

Camera robots also can serve to relieve operator boredom in production situations where a camera must be left in place for long periods and used only occasionally. Examples include parade routes, city councils or other public forums, race tracks, elections, and telethons. All that is needed for control is a data circuit between the control point and the robot. The operator at the control position can observe the response of the robot by the action shown on the monitor for that camera.

Robots come in a wide variety of configurations and cost ranges. Some of the most simple devices are the older, yet still desirable, pan-tilt mounts. At the next level are those with memory for different shots and variable speed, followed by those with capability for movement along one or more axes. The full-featured robot adds elevation to the other movements and some have scissor-type mounts (or cranes) to extend the elevation range at both upper and lower extremes.

Studio Floors

Camera robots are heavy (several hundred pounds) and may be required to execute shots while in motion. Thus, the surface of the studio floor is extremely important to the on-air look the cameras will produce. One of the best studio floors is built of concrete with a smooth epoxy paint covering, that together provide a degree of levelness of less than 1/8 inch height variation for every 12 linear feet. Many studio floors have a concrete base but have been covered with tiles that may not be entirely suitable for the weight of robots or may be chipped and require replacement.

Another possible and popular location for camera robots is the station newsroom. Many newsrooms have computer-style raised deck flooring that must be evaluated for suitability. The manufacturer should be consulted to insure that the floor will be suitable for the robots.

Control Point

An important part of installing camera robots is deciding where to locate the control point of the system. Possibilities include the studio, the production control room, the camera control position or a dedicated control room. An advantage of having the robot controls at the camera video control position is that the video operator can combine this function with that of the robot controls.

One approach is to have the camera control or video operator also assume the duties to control the robots, either from a studio control room or a dedicated video control room. Some installations locate the robot

Figure 1. A custom-made console is equipped with monitors for the cameras, program line output, and teleprompter.

operator on the studio floor. In some instances the control position is elevated. This has the advantage of giving the operator direct view of all camera movements.

Some systems have cable length limits between the camera control heads, the robot control panels, and the rack mounted control processing and memory storage units. Such restrictions may limit the operation for placement of control panels or rack systems.

Installation of the robot controls can be kept simple and straightforward. In Fig. 1, a custom-made console is equipped with monitors for the cameras, program line output, and teleprompter. There is one simple control panel for each camera and a common joystick panel used by the operator to, (1) preset the shots for each camera, and (2) to trim the shots as needed.

CAMERA ROBOT NAVIGATION SCHEMES

Camera robots utilize sophisticated computer techniques and servomechanism technology to provide the quickness, smoothness and repeatability required to be effective. In addition to robot systems where the pedestals are manually moved to a location by a person on the studio floor, camera robot pedestals that can automatically move about the studio are also available. A variety of schemes are used by the different manufacturers for the system to "know" where each pedestal is located. The sophistication of these navigation systems is such that most systems will not allow cameras to run into each other. As discussed below under "Safety," there are mechanisms on the camera pedestals, such as proximity and sonar detectors, to protect the cameras if they do come in contact with an object.

Studio camera robots must have some means for determining their position in the studio and with respect to other robots. There are several different approaches to determining how the robots move about the studio which include mechanical tracks, tape tracks, "targets," and wall mounted bar-coded signs. Of paramount concern is safety to both studio personnel and other robots and facilities (monitors, set pieces, microphone, and lighting stands) within the range of the robots.

New studio procedures must be developed and established when employing camera robots. Camera robots depend upon reference marks which cannot be covered or damaged. The robots also may move without apparent warning during set-up, rehearsal and on-air production. *Camera robots do not have peripheral vision!* Therefore, a new awareness of studio activities is required of all production personnel. Moreover, the damage a robot can cause to monitors, lenses and other studio equipment may cost more than the operator that was replaced.

There are several methods and associated technologies that are employed for controlling the position and keeping track of camera robots.

Mechanical Track System

The oldest and most limited method is the X-Y rail system which requires the construction of wooden or metal tracks or rails on the studio floor on which the robots travel. This system is much like, and was adopted from, the camera dolly systems used in motion picture studios. While the technique provides extremely smooth and repeatable shots, substantial work is required to move or re-configure the tracks if the studio set is changed or if a different perspective is desired. A simple track system may be ideal for use in the newsroom or dedicated news studio where limited movement is all that is needed.

While movement of the robots is limited by the rail system, it is important that a sufficient number of proximity and contact sensors be included to prevent cameras from colliding, cables becoming cut or damaged, or personnel being injured by the moving equipment.

Tape Track System

This basic navigation system uses a metal foil tape placed on the studio floor that functions as a track for the pedestal to follow. Some robot control systems allow the pedestal to move away from the tape for short distances. The robot then relies on counting wheel revolutions (tachometer pulses) to compute its position. Disadvantages to this system are the need to change the tape whenever there is a change in set placement or damage to the tape caused by moving heavy objects over it.

The Target System

The tape track technique can be expanded and simplified by having a home "target" on the studio floor. Each camera robot uses its target to establish a reference just prior to a production, and is then free to roam the studio floor. Of course, the further the robot moves from the target the greater the accumulated error will be, especially after changing positions several times during a production. Just as the tape track must be protected from damage, the target must also be protected yet made available to the robot whenever it is necessary to re-establish its reference.

Bar-Coded Signs

Another robot navigation approach is to attach bar code signs on one wall of the studio. Each camera robot pedestal then uses a laser in the pedestal base to scan the bar codes on the wall. By knowing the location of the bar code cards, the system can calculate the location of the pedestal. While this technique employs substantial computer programming power, accidentally blocking the view of the signs can render the robot blind to finding the reference. The operator will be signalled when this happens, but the carefully designed production may be adversely affected as a result.

Each of these navigation schemes has its advantages and disadvantages. When selecting a camera robotic system, the user should evaluate each scheme to decide which one is best for the facility in which the robots will be working. An evolutionary approach may be employed in which initially a system would be installed with manually positioned pedestals that would later be upgraded to full mobility, perhaps one camera at a time, for economic reasons.

SAFETY

Safety is an extremely important matter with robotics cameras. *The cameras must not run into people, the set, or each other!* Several methods are employed by different manufacturers to provide various levels of safety.

Limits: Virtually all robotics servos are current or torque limited. This means that if, for example, the camera pans left and the lens pushes into a floor monitor, the servo will sense the obstruction and stop movement in that direction. In some systems the servos will back up a little. This limiting action can apply to any of the several directions of movement of the robot. Moving pedestals may have additional safety features. For example, infra-red sensors or microswitches on the base of the pedestal will sense any obstruction and stop the pedestal movement.

Observation: One way of observing the movements of the camera robots is to install a ceiling mounted camera that takes a wide shot of the studio. The output of this camera can be displayed on a monitor in the camera robotics control area.

Signs: The installation of large and bright signs on the sides of the camera robots and the studio doors and walls will serve to remind production personnel that the cameras move by themselves and without warning.

"WARNING!
KEEP CLEAR!
CAMERA ROBOT MAY MOVE
WITHOUT WARNING!"

Cables: Making sure that the cameras do not get tangled in their own cables is both a safety as well as an operational consideration. Problems of this nature can be alleviated by careful planning of camera movement. The camera cable and the robotics control cable can be sleeved in a nylon webbing. This webbing combined with low skirts on the camera, will minimize cable management problems.

OPERATING CONTROL SYSTEMS

Three of the most common types of controls for robotics cameras include: joysticks, data tablets, and touch screens. Most control systems can operate several cameras (four to eight) at once. The software in most control systems is designed to prevent the operator from requesting the cameras to do something that (1) cannot be done (move to the other side of the studio in 5 seconds), (2) is dangerous (move to the same location as another camera), (3) change position while on-the-air (unless the operator over-rides the warning), or (4) exceed a built-in limit or range.

Joysticks

This popular and relatively simple method provides simple direct control over the camera movement. Typically one camera is selected at a time for direct control in the system. Using the joystick, the operator sets-up the shot to the director's satisfaction. The rate of movement is also established. The shot is then stored as a number or an alphanumeric designator in the computer memory. The shot can later be recalled during rehearsal or broadcast using that number or alphanumeric.

Data Tablets

These devices are similar to those used on some video graphics generators. The operator can draw a studio layout on the tablet, indicate the positions of the camera and then teach the system the desired camera shots. Shots can then be recalled by touching a specific area on the tablet. For example, touching the tablet at a location representing the front of the Sports desk could move a camera to a wide shot of the Sports desk. Touching the tablet at a point behind the desk could cause the camera to go to an effects shot. In addition to positional information the system also stores zoom lens settings and pan-tilt positions.

Touch Screens

Similar to other touch screen controlled devices, such as graphics generators, computers and test equipment, this control system allows the operator to set and recall camera cues by the touch of a finger on a screen. Typically these systems use menu driven color touch screens. On some systems on-air tally indication can be provided on-screen, driven by the tally signals from the switcher. This feature reduces the chances for error by the operator accidentally moving an on-air camera to its next shot.

Options

When the primary usage of a camera robot system is the presentation of news, two other newly developed operations may be of interest: auto-focus and auto-framing. Auto-focus works on the same principle as an auto-focus consumer film camera. The time delay between the transmission and reflected reception of a high frequency sound is used to determine distance which sets the zoom lens focal length. An override control is needed with this option because auto-focus may be "fooled" into focussing on a large foreground object rather than the intended subject.

Auto-framing is an option that uses sophisticated comparison technology to adjust the framing of a shot when a moving subject is involved. A predetermined shot is programmed into a video frame store. The system continuously compares the stored shot to the live shot and automatically pans or tilts the camera as required to keep the object and stored frame coincident. A news talent who is prone to body movement on the air can be kept properly framed automatically.

Limits

Most systems, regardless of the type of control, can control multiple camera moves simultaneously. Some systems can also warn of impossible shots. For example, giving a command to move a camera from one end of the studio to the other in three seconds will cause some systems to display a warning message. Using camera robotic systems simultaneously to move multiple cameras, however, can result in added finesse to productions since more complicated moves can be smoothly executed on a consistent basis.

Backup Systems

Because the entire studio operation may be contingent upon the proper operation of the camera robots, it is essential that backup power and computer arrangements be made. In event of an electrical circuit outage that could take the robot controls down, some kind of uninterruptible power supply (UPS) on the computer control system (not the robot mechanisms), to bridge the gap between outage and restoration of line power, should be considered.

During installation of the robot system some of the power supplies for zoom and other functions on the camera may be replaced with power from the robot system. To avoid a complete failure of the camera if the robot fails, the original power supplies could be left in place for manual reconnection if needed. This arrangement may also aid maintenance of the system.

The computer control itself is also a source of single-point failure that can take out the entire system. A backup computer would reduce the time to restore operation if the main computer had to be repaired. Of course, a means for switching to a manual control panel could also help solve this problem.

Backup systems become more important if the camera robot controls are interfaced with other studio facilities. More than just a camera position is at stake if a control system failure should occur.

INTERFACE WITH OTHER SYSTEMS

Camera robot systems have varying levels of interface capability. Some provide for extra extension panels (pushbuttons, keyboards, joysticks) to be used for addressing external devices or systems but tied to the robot control system. The cues and control signal can be initiated as part of the shot list for the cameras. Some of the possible options are described below.

Some caution is needed when considering the interconnection of multiple devices employing microprocessors. The interfacing software must be developed, rigorously rehearsed, *and carefully documented* to avoid one system causing problems with another. The knowledge of how the systems interact must be shared with others so that when problems occur, personnel are available to solve them quickly. Few situations are more exasperating than when it is impossible to accomplish a function because the computer is down or a bug has been found in the software.

Camera Control Unit: Many robot control systems provide an interface with the camera control unit (CCU) (such as iris, pedestal or blanking or black level, color balance, and other controls) if the CCU has a serial data port that can be accessed.

Switcher Automation Systems: The robot controls can be interfaced with some production video switcher and effects systems (if an external serial data port is available) that can, (1) give instructions to the switcher upon selection of a camera move, or (2) activate a camera movement upon selection of the camera by the switcher. The result is simultaneous selection and start of movement by the camera.

Lighting Systems: The interface of studio lighting control systems with the robot control system provides the ability to cue lighting sequences along with the activation of camera moves. The camera control could be used to provide a trigger to the lighting system to initiate a lighting change sequence or merely to turn on (or off) the lights for a set about to be used on camera or a set that no longer is needed. Again, a serial data port is needed on the lighting control system for the interface.

Audio Control Systems: Just as audio-follow-video (AFV) interconnection systems are common in many television stations to switch microphones along with camera takes, similar interconnect arrangements can be made with robot control systems. However, the use of a more intelligent controller than that in the switcher provides the ability to control the audio with more finesse. For example, in situations where a complex video switcher and effects system does not provide reliable tally or AFV signals, the robot controller may provide the ability to turn on a microphone slightly in advance of a camera take to that on-air talent and turn it off some time after the camera is off the talent (to avoid clipping dialogue).

Newsroom Facilities: An interface with electronic newsroom facilities and the robot controller provides the ability to cue prompters, turn on microphones and lights, and activate tallies and other signals. Cues in the prompting text can be used to switch cameras or select preprogrammed shots.

Data Modems: Because the robot control is a computer, virtually any device that can be connected to a modem can be controlled in some fashion. Thus, it becomes possible to control a camera or some other device at some remote location via telephone lines or other data link. For example, cameras and other devices can be controlled via modems at satellite news bureaus or other remote sites.

OPERATIONAL CONSIDERATIONS

Manual Control

Frequently camera robots are used daily for scripted or fairly predictable shows such as newscasts or talk shows. Occasionally it may be desired to use camera robots for telethons or other more complex or spontaneous productions that would be difficult to do with robotics control. On these occasions it may be desirable to have camera operators. Some camera robot systems have switches that allow the robotics to release control of the camera pan head and the pedestal and thus allow for almost instant manual operation. This feature (also called "backdrive") can also be used as a manual backup in the event of failure of the robotics system. Another advantage of having this feature is to mix manned and unmanned cameras to suit the needs of the production.

Some robots will allow a switch to manual control, memorize the new position, and, after returning to remote control, adapt to the new position and continue with the shot sequence. In other words, on some systems, remote and local control may be mixed at any time.

Operator Training

The operation of robotics camera systems is fairly straightforward. Therefore, training can usually be accomplished in a few sessions with a good instructor. After the operator is familiar with the system, it is important that directors and others involved in studio production agree on a uniform set of commands for robotics camera operation. Rehearsals before the first live session can also be helpful. As with other operating positions in a station, one primary and several backup camera robot operators should be trained for this essential position.

Maintenance

Camera robot systems are generally considered to be reliable and require little routine maintenance other than the general cleaning, lubrication, routine testing, and periodic software and electrical checks that any electro-mechanical device in daily use would require. Support from manufacturers varies, however, and the availability of parts and service must be resolved at time of purchase. Good documentation on mechanical components, electrical systems, and computer software will help insure fast repairs and fullest use of the system's capabilities.

Some robot controllers use standard personal computer (PC) hardware plus specialized keyboards, tablets, or touch-screens to help keep maintenance simple and straightforward.

SUMMARY

Camera robotic systems can be a cost-effective investment for a broadcast television station or production facility. The level of sophistication in computer control and servomechanisms has reached the point where camera robots can provide respectable performance in certain production situations. In addition to the cost savings in operators there is the potential for linking the robots to other automation systems to integrate programming activities. However, consideration of the use of robots must be made on the basis of evaluating the capabilities of the technology that can improve operational efficiencies without causing any degradation of production values.

Careful planning, both of people and equipment, selection of a manufacturer, proper installation, a good maintenance program, and a control system with the degree of sophistication that meets the needs of the station can result in an effective, efficient, and cost-saving installation.

5.12 Planning Broadcast Facilities

Malcolm M. Burleson
Alexandria, Virginia[1]

INTRODUCTION

This chapter contains information which may be used as a guide for general planning and cost estimating of facilities. It is beyond the scope of this chapter to describe all of the many details which go into the design of a facility. Rather, it is intended here to provide hints and reminders of some of the more important considerations that should be addressed.

There are many factors that go into determining the initial planning budget, as well as accounting for the final cost figures when the project is completed. Advance planning combined with close project coordination and continuous oversight by the owner's representative will result in a well-built, cost-effective facility, constructed on time.

Give the construction or refurbishment of the facility the same study that would be given to a major equipment purchase. Research the facility needs, available architects and contractors. Review local codes, construction specifications criteria, and cost accounting methods. Hire additional personnel to cover for staff that will be occupied overseeing design and construction.

Construction of broadcast and production facilities is not the same as simple office construction. They require special attention to elements such as control rooms, special acoustic and air conditioning treatment, and special technical services not normally found in routine construction. Years of experience have taught the lesson that numerous large and small broadcast facility projects all have one major thing in common: the need to plan ahead and to watch every detail, or else the barrage of changes will undermine the budget and completion date.

[1] Additional input and editing was provided by Edmund A. Williams.

GENERAL PLANNING CONSIDERATIONS

The early planning of a broadcast facility involves factors such as: consideration of the market to be served, site selection, studio requirements, nature of programming, expected hours of operation, and available capital.

First and foremost decide whether the studio and transmitter are to be combined under one roof or established in separate locations. Wherever practical, it is more economical to combine the studio and transmitter. While the initial equipment requirements may be less, more importantly, the day-to-day operating expenses are substantially lower. With the plant "all under one roof" significant savings in heating, air conditioning, building maintenance, travel time, and personnel can be realized.

When a combined operation is not practical, the second best approach is to place the transmitter and other facilities in a separate building and operate the equipment by remote control from the studio. With a good remote control system, a transmitter site can be selected that provides the best coverage and the studio can then be placed in the most convenient location for business and production. The building requirements at the transmitter can be kept to a minimum, with space only for the transmitting and support equipment, a small work area, and the heating and air conditioning system.

Building Codes

Virtually any construction or alteration of a building is likely to be covered by one or more codes. In addition to general building codes, there are electrical, mechanical, plumbing, and other specialized codes. The building at a separate transmitter site is likely to be subject to fewer building code restrictions, because it is away from heavily populated areas and fewer personnel are involved in its operation. While various national associations have proposed model codes, individual state and local jurisdictions adopt whatever codes, with variations, they find appropriate. Reputable architects, contractors and subcontractors are familiar with local codes affecting their area as they apply to a specific construction job.

Zoning Regulations

Zoning is another consideration which is dependent on the local jurisdiction and can have a significant impact on selecting a studio location, appearance of construction, access to the facility, and placement and size of towers and antennas (microwave and satellite). While it is sometimes possible to have zoning requirements changed to suit a particular construction requirement, this should not be counted upon when considering a new site. However, if a site is particularly desirable, an advance team can be employed to (1) investigate the zoning regulations, (2) determine whether they are enforced, (3) consider the attitude of potential neighbors to having the proposed facility constructed nearby, and (4) propose solutions to the problems the zoning regulations are likely to cause.

If the planned construction is for an expansion of an existing facility, a careful review of local zoning regulations is essential. It may be necessary to increase parking if there is a staff increase. Construction that was "grandfathered" in the past may be required to meet new requirements if a modification is made.

PLANNING FOR NEW AND RENOVATED FACILITIES

The first step in planning a new facility or renovating an existing one is to develop an outline of requirements that will form the basis upon which future decisions are made. Such an outline for broadcast facility planning might be developed as follows:

Site Selection

1. Select a site that provides adequate space, not only for immediate building needs but any expansion anticipated.
2. Provide adequate space, open areas, and parking for employees, guests, and studio audience (if under consideration), and a reasonable surrounding landscape environment, if possible.
3. Access for delivery trucks and a generous loading area with a loading dock is essential if shipments over 100 lbs are expected on a regular basis. The additional cost for deliveries and the occasional rental of a fork-lift will be paid for in the first year of dock use.
4. Accessibility to major (and preferably less congested) roads for mobile units, audience, and the general convenience of employees will save time and reduce frustration.
5. Space for a tower for line-of-sight and ENG microwave and communications antennas will be needed either on the site or on the building itself.
6. Zoning regulations will affect use of towers, antennas, appearance of construction, and facility identification or advertising signs.
7. Consider a space for a helicopter landing pad on the roof or adjacent to the building for possible use for news and traffic reporting.
8. Consider various environmental elements such as noise from nearby airports and highways, weather (is the site subject to high winds or snow accumulations?) and, of course, susceptibility to flooding or high water.
9. Take into account the special technical requirements that broadcast and production facilities have, such as installation of satellite antennas, access for remote vans, microwave paths, telco connections, and alternate or standby power.

Building Planning

It is essential to pay particular attention to efficient personnel traffic flow and the relationship of studios, administrative offices, technical areas, craft shops, storage, and building utilities, such as power, heating, air-conditioning.

The design, layout and space arrangements of a new facility should include means for easy coordination between such operational functions as programming, operations, and engineering; between on-air talent, dressing rooms and studios; and between newsrooms and their mobile units. Examples include:

- Entrance arrangements for employees, guests, VIPs, and audience
- Relationship of studios to loading areas
- Relationship of parking areas to employee and audience entrances
- Storage spaces—area, rooms, closets, nooks and corners
- Physical security inside to discourage unwanted entry and passage throughout the facility
- Controlled access areas for staff
- Outside security to discourage theft and vandalism to the building and employee and visitors cars and outside technical facilities (satellite antennas)

Budget Development Checklist

The following checklist is designed for estimating the scope and cost of station construction.

Preliminary

- Land
- Land tests
- Site clearing and preparation
- Architect and engineering fees
- Permits (zoning & building)
- Special consulting fees (decoration, landscaping, communications)
- Local zoning and building codes

PRELIMINARY PROJECT BUDGET

Station ______________________ Planned Start Date ______________

Planned Finish Date ______________

Brief description of construction: ______________________

Item #	Description	Estimate
1.	Land purchase or lease	______
2.	Property survey	______
3.	Title search, insurance	______
4.	Real estate commission	______
5.	Architects/engineers—fees	______
6.	Permits and licenses (if separate)	______
7.	Consultants: acoustic, structural, etc.	______
8.	Site preparation and demolition	______
9.	General construction and finish	______
10.	Optional construction items and finish	______
11.	Special construction and finish	______
12.	Furniture and fixtures	______
13.	Decoration—interior	______
14.	Landscape—exterior	______
15.	Special equipment—electronic (and other)	______
16.	Special equipment—installation of above	______
17.	Contingencies (including price increases)	______
18.	______________	______
19.	______________	______
20.	______________	______
	Totals	______

Prepared by ______________________

Date ______________

General Construction

- Architectural
- Electrical
- Mechanical
- Heating and plumbing
- Special in-house cabling (TV, telephone, computer, PA)
- Communications (microwave, two-way radios, satellite)

Interior Finish

- Wall covering (fabric or paint)
- Floor covering
- Special studio treatment

Furniture and Fixtures

- Decorative office furniture
- Standard office furniture
- Office area built-ins
- Working area counters, cabinets and built-ins
- Studio fixtures
- Draperies
- Art work

Telephone System

- Broadcast line facilities
- TWX and facsimile (FAX)
- System configuration
- Office intercom system
- Number of private lines
- System: owned or leased
- Backup arrangements (very important!)

Computer Systems

- Administrative
- Technical
- Production
- Backup arrangements (very important!)

Security Arrangements

- Control of public (entrances and exits)
- Control of all points of entry (including employees)

- Closed circuit TV systems, security's tours
- Controlled access doors and gates
- Separation of 9:00–5:00 areas from 24-hr. areas
- Night lighting and emergency power
- Surveillance of ENG and SNG units and parking areas

Miscellaneous Items (Often Overlooked)

- Copier or printer outlet (may require special power hookup and ventilation)
- Special ventilation for odor producing areas (art work, craft shops)
- Building maintenance equipment closets
- Drinking fountains and vending machine areas (power hookups)
- Space for air-conditioning subdistribution boxes, fans
- Special waste water treatment for film darkroom processing (local codes)
- Cable access to roof for various communications systems
- Provisions for microwave and satellite antenna mounts
- Power, communications, signal lines to SNG, ENG unit areas
- Use of walls for built-in filing cabinets

Building Construction: General

Here are some of the major elements regarding building construction that should be considered:

1. *Official Contact*—Restrict contact with the contractor to only those staff or employees given the authority to speak for the company and authorize changes. Route all official communications to and from the contractor through one individual, if possible.
2. *Change Orders*—Any changes in the design must be kept to a minimum if the budget is to be maintained. Changes in design are expensive and time consuming. Virtually any suggestion offered to the contractor by the facility staff will be considered a "change" by the contractor if "change" authority is not specifically designated as discussed above. The extra time spent in careful planning in advance of construction will more than pay for itself with fewer change orders. Coordinate any changes with all departments for a possible overlooked error. Make sure the contractor obtains written approval on all change orders and that the station understands the amount of additional cost involved. Several dozen relatively "small" change orders can result in many thousands of dollars of budget overruns with very little to show for the changes.
3. *Document the Construction*—Take pictures of all new construction areas at various stages in construction. Alert the architect that "as-built" drawings are to be provided. In some cases there is added cost, but advance notice will keep the cost down and the accuracy of the drawings will make future changes easier and less costly.
4. *Structural*—The selection of the main structural system will take into account local conditions, codes, availability of materials, flexibility for future changes, special loading conditions and spans, degree of fire resistance and effect on fire insurance rates. A careful study of local fire code requirements will result in designing a facility that meets the codes and avoids costly retrofit later (doors, fire control systems, ventilation).
5. *Partitions*—Plan the layout of the interior of the building to provide sound isolation in the studio and control room areas, yet enough flexibility for ease of removal or relocation or some of the walls and partitions in office and general purpose areas. In general, the walls needed for good sound isolation will be more permanent than those for offices and other areas. It is generally better to make rooms larger in the initial planning and construction stages and subdivide them later (after construction is completed) if necessary.
6. *Ceilings*—Ceiling spaces contain virtually all utility services: sprinkler system, heating ducts, water, sewer, power for lights (and more recently outlet power), communications (computer, signalling, security, production video and audio), conduits, and wire trays. A high degree of accessibility is required to the ceiling spaces as well as good acoustical quality and ease of cleaning. Attention paid to the ceilings during the design phase will result in easier and less costly access to the space later.
7. *Wall Materials*—Greater durability is needed in technical and production areas, due to the mounting of various equipment to the walls as well as frequent moving of equipment and supplies. Install wall guards or rails along all corridors and rooms in which equipment will be transported.
8. *Floors and Floor Materials*—Most office buildings have concrete floors. Avoid placing any conduit in concrete floors in order to reduce installation costs and increase flexibility later. Different floor types and coverings are needed in different areas. Get the best available that is durable yet will reduce traffic sounds and vibration. It will be costly and disruptive to replace it in the future. Use carpeting in areas where footfalls and dolly wheels are likely to transmit sound to a nearby studio, control room or edit booth. Consider the pros and cons of using raised computer floors in control rooms, video tape, and other electronic areas. Computer floors are expensive, require constant maintenance, and are difficult and dangerous to work with when installing cables. However, they are an ideal duct for ventilating cool air into the bottoms of racks and consoles. They may be more convenient to use in an area devoted exclusively to computer facilities and may be selectively employed in technical areas. Computer floors look better than overhead cable trays but trays are easier to wire and rewire.

9. *Studio Floors*—Floors for studios have altogether different requirements. Concrete slab floors are typically covered with either a rubberized surface (higher maintenance) or close fitting tile (lower maintenance, but subject to "pot-holes"). Slab floors for TV studios must be extremely level if smooth camera movement is expected.

   Studio floors are often isolated from the building structural members in order to reduce sound transfer. Consider this approach, which is expensive, only if the studio will be relatively close to interior (air conditioning, craft shops) or exterior (aircraft or highway) sound sources, or if a higher degree of sound isolation is required. Floors are more difficult than walls or ceilings to acoustically isolate from building vibration and noise transfer.

   Footfalls in corridors will be heard in adjacent studios and control rooms if the studio floor is not isolated or the corridor floor is not carpeted or acoustically treated.
10. *Studio Ceilings*—If the ceiling of the studio is directly under a roof, consider what acoustic and thermal isolation may be needed if the site is (1) near an airport, (2) subject to heavy rains, snow or hail, or (3) located where there are hot sunny days.
11. *Tape Vaults*—Give special consideration to this area which contains the station's major programming resources. Use fireproof construction and limit ventilation to no more than that necessary for proper temperature and humidity control. Install instruments to measure the temperature and humidity. Control access to the vault and keep logs of tape movement into and out of the vault.
12. *Sound Locks*—Sound locks are small vestibules to provide sound isolation between noisy places (such as corridors, newsrooms, offices, and control rooms) and studios. The sound lock, when properly used, effectively allows entry into or exit from a studio during a broadcast or recording session. Sound locks can have several doors and serve more than one control room or studio. The space used by a sound lock is not entirely wasted. The space above the ceiling can be used for the ballasts for fluorescent lights used in the studio or control rooms.
13. *Sound Lock Doors*—These special doors must be carefully designed and properly installed. In addition to their superior acoustic performance, sound lock doors must have:

    A. A double glass window with rubber mounting to allow a view of who might be using the sound lock to avoid opening both doors at the same time.
    B. All sound lock doors must open outward from the sound lock so that when a slight vacuum is created it more firmly seals the other doors.
    C. The outer door of the sound lock must have a latch and optional lock.
    D. The inner studio door must have no latch, which would cause an audible click in the studio, but handles must be provided.
14. *Windows*—The main product of a broadcast or recording facility is the sound material generated in its studios. Windows can be strategically placed in studios and control rooms to allow clients to view the activities, for visual communications between the areas, for tours to see the operation without disruption, and for employees to see who is working in the studios without having to enter or use the sound lock. The additional expense for properly designed and installed windows will be well worth the effort in the long run.

    Adding windows after facility construction or renovation is completed is even more expensive because; (1) the operation will be disrupted during construction, (2) the studio or control room may have to be rearranged to accommodate the new window, and (3) conduits and support members in the wall of interest may have to be moved, modified, or removed altogether.

    If after construction the window is not needed, it is a simple matter to cover it or add drapes to limit the times when it is appropriate for use.

    Where the location of the facility allows it, windows to the outside should be considered. This is particularly applicable for radio stations. In some instances studios are built in the interior of a building and exterior windows are not possible. However, when a radio studio or combo facility is built on an outside wall, a small exterior window will allow occupants to know a little more about the weather and have a connection with the outside world. Of course, if the facility is located next to an airport or major highway, the window invites more background noise than might be acceptable.

Engineering: General

1. *Main Power Service*—A new physical facility will require a new main power service and metering to be constructed by the power company, which may pass the cost on to the customer. An addition to an existing facility may require some construction for the additional power capacity. Research the costs with the power company well in advance of construction in order to reach a cost compromise. After all, studios are good electrical customers. Consider an alternate service entrance to (1) reduce the construction on the main service, and (2) provide a partial backup in event of a main service failure.

   Examine the path the main service takes from the substation. Consult neighbors for indications and frequency of power problems.

   It is not usual for a second service to be installed if convenient. That is, if one is within a block or so of the studio. However, it may be expensive to have it brought in when it is to be rarely used. It may be better to spend the money upgrading or hardening the main service.

2. *Power Substations*—Consider separate unit substations for: (1) air-conditioning and electric heating and other major motor loads such as elevators, (2) general illumination, (3) studio production lighting, and (4) technical facilities. However, arrange for a common metering system to keep the per unit usage costs low.
3. *Lighting in General*—Use fluorescent lights providing about 100-footcandles in all personnel working areas. Use dimmer controlled incandescent lights in control rooms and other technical areas during production activity and to provide ambiance. Include fluorescent lights in the same areas for installation, maintenance, and cleaning.
4. *Air Conditioning*—Studios require high volume, low velocity air ducts for sound control. Consider a separate system or stand-by components or cross-connecting arrangements for control rooms and editing suites. Install a separate air purging exhaust system for TV studios.
5. *Energy Savings Systems*—Novel energy savings techniques should be considered but only with a thorough understanding of the consequences of a system outtage.
6. *Miscellaneous Systems*—Telephone, public address, security's tour, door security, closed circuit TV, and fire alarm.

Electrical

General

The following sample specifications for the electrical system of a television station are based on the National Electrical Code, and are intended to show how such specifications may be tailored. Elements may be adapted to other applications, such as facility upgrading as well as radio station construction.

All state and local electrical, fire, and safety codes are to be observed.

Area Classification

All electrical installations, materials, and equipment shall comply with the classification "General Purpose" except for hazardous areas which shall be designed for Class I, Group D, explosion-proof conditions. (There are normally no hazardous areas in a broadcast facility.)

Incoming Power Service and Metering

Incoming power service is be high voltage (4.16 kV or 13.8 kV) due to high load requirements.

Standby partial or full incoming service shall be provided (if practical) with automatic transfer when normal service fails. One primary metering point shall be provided to obtain best possible utility rates.

Primary Distribution

Distribution within complex may be high voltage (4.16 kV or 13.2 kV) from a primary switchgear to unit substations located as close as possible to the center of the loads served.

Separate unit substations shall be provided for different type loads, as follows:

- A secondary 120/208 volt system to handle equipment and motor loads and a 277 volt system for all fluorescent lighting
- A secondary 120/208 volt system to handle receptacle, incandescent lighting and small equipment loads
- A secondary 120/208 volt system to handle studio production lighting only
- A secondary 120/208 volt system to handle technical TV loads only

Each unit substation shall include components as follows:

- Primary compartment with high voltage (HV) fused load break switch
- Open dry-type transformer with a delta HV primary and 120/208 volt or 277/480 volt, three-phase, four-wire secondary: with best possible sound rating of transformer
- Voltmeters and ammeters provided for each phase
- A main secondary air circuit breaker
- Molded case feeder circuit breakers and spares

Secondary Distribution

Power shall be extended from unit substations with cable and conduit to automatic circuit breaker panels and motor control centers.

Motor control centers shall be Class I, Type B, with combination magnetic, full-voltage starting, circuit-breaker-type motor starters or circuit breakers only for 480 volt, three-phase operation. Each starter shall have three thermal overloads.

Power panels shall be designed for 480 volt, three-phase, three-wire service. Panels shall be of the dead-front type with automatic circuit breakers of ampere rating as required.

Panels for receptacle and incandescent lighting loads shall be designed for 120/208 volt, three-phase, four-wire service. Panels shall be of the dead-front type with automatic circuit breakers of ampere rating as required. Ground fault detector breakers shall be installed as required by codes.

Lighting panels shall be designed for 277/480 volt, three-phase, four-wire service. Panels shall be of the dead-front type with 20 ampere automatic branch circuit breakers. Panels shall be similar to Westinghouse Type NH1B-4.

Conduit

Conduit shall be rigid steel, asphaltum painted when installed in concrete slabs, below grade and outdoors above grade.

Rigid aluminum conduit shall be used for exposed installation in mechanical equipment rooms, damp locations, and locations where exposed to mechanical damage.

Rigid aluminum or steel conduit shall be used for all feeder and subfeeder runs.

Steel EMT conduit with compression weathertight fittings shall be used for all other branch circuit wiring indoors and above grade.

Wire

HV cable shall be single conductor crosslinked polyethylene insulated and shielded.

Building wire shall be Type THW rated at 600 volt, 75°C. No. 12 AWG and smaller shall be solid copper. No. 8 AWG and larger shall be stranded.

Fixture wire shall be Type AF, 300 volt insulation.

Minimum wire size shall be No. 12 AWG, except No. 14 AWG for control wires. Maximum wire size shall be 500 MCM.

Grounding

Electrical grounding shall be provided in accordance with the National Electric Code. Equipment enclosures, electrical service, transformer neutrals, outdoor lighting standards, and cable shielding shall be grounded.

Insulated bushings and double lock nuts shall be provided at all panel boards and pull boxes in feeder runs and pull boxes shall be bonded through with bare copper wire.

A separate technical equipment ground system shall be provided as required.

Switches, Wiring Devices, Wall Plates, and Special Enclosures

Single pole switches shall be 20 amperes, 120/277 volt, AC, quiet type.

Duplex receptacles shall be 20 amperes, 125 volt, 2 pole plus U-slot ground.

Special outlets to be provided as required.

All wall plates for switch, receptacle, telephone and computer outlets shall be 0.06" stainless steel.

Telephone System

Two incoming underground services are required, one for technical use and one for business and office use.

The equipment room for the technical service shall be located close to master control and there shall be a cable-tray tie between the telco equipment room and master control.

The business and office system shall be complete, consisting of conduits from equipment room outlying telephone closets and interconnecting panels and thence to the various outlets as required. All installations shall be in accordance with the requirements of the local telephone company.

Interconnecting panels shall be steel with plywood backboard with full opening door, latch, cylinder lock, and trim.

Each telephone closet shall be furnished with plywood backboard for the installation of distribution equipment.

Conduits from the equipment rooms to each destination shall be ¾" minimum.

Telephone and computer outlets shall be four inches square with bushed aluminum cover plates.

At least two separately fused AC power circuits from the emergency power distribution system shall be provided in each telephone terminal room.

Note: Modern telephone systems are digital or multiplexed systems often requiring no more than two pair between instrument and telephone panel. To save substantial time and money in the future it is recommended that three or four pair be pulled into each office as a home-run to the panel. This will help to serve the needs of future communications services.

Public Address System

A complete PA system consisting of amplifiers, loudspeakers, and microphone shall be provided. In some cases this function can be accomplished with the telephone system.

Loudspeakers shall be located in corridors and other strategic locations.

System shall be zoned as required.

Fire Alarm System

Note: New codes have made fire alarm and control systems extremely complicated. The system must operate reliably. While it is expensive to evacuate a television facility, stations have been known to burn to the ground. False alarms can be kept to a minimum with knowledge of the system and operational procedures. Fire drills must be held annually and may be planned in advance to reduce the disruption to the operation.

The fire alarm system shall be closed circuit zoned, consisting of control cabinet, gongs, manual stations, and automatic fire and smoke detectors.

Manual stations shall be provided at each stairway on each floor and at all ground level exterior doors.

Automatic thermal or smoke detectors shall be provided in all areas except where sprinkler heads are installed.

Each sprinkler alarm valve shall indicate on the fire alarm panel zone annunciator as a separate zone when activated.

A detailed instruction tour of the system shall be provided by the vendor.

A Halon Fire Suppressant System is to be installed in the master control and video tape areas. One practice "dump" shall be provided under supervision of the station and witnessed by the fire marshall.

Note: Provide hand fire extinguishers at exit doors, in corridors and as required in shop and craft areas, and as required by local codes.

Cable Trays and Signal Conduits

A system of cable trays and signal conduits originating from master control shall be provided to studio control rooms, studios, microwave rooms, electronic maintenance shop, and computer centers.

In addition, a separate cable tray for microwave waveguide shall be provided from the antenna site to areas where the microwave receivers are located (news room or master control).

Special insulation (Teflon) must be used for all signal cables not enclosed in conduit. Because of the substantial extra cost for this cable, serious consideration should be given to running extra cable ducts between technical and production areas, and areas that might be used in the future as well as conduits to most other locations in the facility.

Studio Production Lighting System

Unit substation and dimmer board shall be located as close as possible to studio served.

Unit substation shall include the following:

- Electrically operated main circuit breaker to permit remote control from studio floor
- Transformer with six 2.5% taps, three above and three below rated primary voltage, to compensate for secondary voltage variations

Other work shall be as follows:

- Wireway with wiring from load side of dimmers to studio floor patch panel
- Studio grid wireways with load wiring to studio patch panel
- Control wiring from studio control console to dimmer board

In sizing unit, substations serving dimmer boards, a 50% demand factor may be applied to connect dimmer load.

An "on-air" studio warning light system shall be provided as required. Install "on-air" or "in-use" lights beside (not above) doors to specified control rooms, edit booths, and studios. Install "on-air" lights above windows to studios. Install "on-air" beacons in noisy scene construction, storage, and staging areas.

Security

The following security systems shall be provided:

- Supervision of all exterior doors on ground level, with the control cabinet in the guard's room
- Closed circuit TV cameras at key positions, with monitors in the guard's room
- Manual nonwired security's tour stations located throughout complex
- Electrically operated gates to control automobile traffic
- Controlled access doors and gates to main office and technical facilities

Emergency Systems

Power for the emergency system shall be provided with a water-cooled diesel generator set with generator output configuration to be 277/480 volt, three-phase, four-wire. The installation shall include (but not be limited to) accessories such as automatic transfer switch, output switchboard, battery starting set, oil storage tank, fuel pump, mufflers, and vibration isolators.

Generator set shall automatically sense power failure or 80% under-voltage, start engine, attain and maintain speed, and transfer designated emergency load. A manual override of start and transfer of load controls shall be provided.

Provide local transformer with primary delta 480 volt, three-phase and secondary 120/208 volt loads on emergency supply.

Install rechargeable battery operated lights in control rooms, studios, corridors, rest rooms, and other areas without windows but which are used by the public or extensively used by employees.

Loads on emergency supply shall include auxiliary lights in the generator room, stairway lights, exit signs, selected corridor lights, telephone and PA systems, selected technical lighting and heating, ventilating, and air conditioning loads required for transmission of limited live news programs, network and taped programs. *Do not connect the rechargeable battery operated emergency lights to the emergency lighting circuit as they will go out when the emergency power comes on.*

Note: Plan for more than mere evacuation procedures. Plan also what must be taken with evacuating staff if the alarm is more than just precautionary. That is, determine what steps should be to insure the continuation of programming? What documents, logs, tapes, data files and other information should be taken out by staff or immediately stored in fire-proof files (for later retrieval) to protect the business and programming operations of the facility.

Lighting—277 V Fluorescent, 120 V Incandescent

Lighting fixtures shall be completely installed with all required outlet boxes and accessories.

Lighting levels shall be in accordance with IES recommendations, with minimum 100 footcandles in working areas.

Fluorescent fixtures shall be used for general illumination. Fixtures shall be with 40 watt RS lamps and HP factor ballast, with best sound rating. Fluorescent fixture types shall be as follows:

- Recessed with acrylic lens diffuser to be used in areas with hung ceiling
- Surface or pendant mounted with wrap-around acrylic lens diffuser to be used in stairs and other selected areas with exposed ceiling
- Industrial RLM with porcelain reflectors to be used in mechanical equipment rooms, storage rooms, etc.
- Executive offices and conference rooms shall be provided with dimmer-controlled incandescent lighting using recessed fixtures in addition to fluorescent fixtures for maintenance and cleaning

Selected walls and art work shall be illuminated with recessed ceiling-mounted incandescent wall-washing fixtures.

Make-up room mirrors shall be illuminated with special bracket wall-mounted fluorescent fixtures. Dressing room mirrors shall be illuminated with special wall-mounted strips with incandescent bare lamps. (Consider the use of special color temperature lights for the dressing room and make-up areas so that the colors applied in make-up will appear the same under the studio lights.)

Outside Lighting

Outside lighting shall include illumination of audience concourses, entrances, parking lots, signs, building exteriors, planters, and the like, and shall be installed in accordance with IES recommendations.

Miscellaneous

Outlets for wall-mounted clocks operating on 120 v shall be provided in designated areas.

A clock system for use with the master clock in master control, studios, and news areas shall be provided as required.

Local office intercommunication systems shall be provided as required.

An audio-visual system shall be provided for all conference rooms.

Mechanical

General

Television studios require special consideration in solving the many problems entailed in the mechanical design due to the high-lighting capacity, noise criteria, air distribution, and entrances. Radio broadcasting facilities require many of the same considerations, but some of the problems may not be as severe.

Grounding Systems

Years ago when an AC ground on individual pieces of equipment was not required, a technical ground made sense. Today, the AC ground and technical ground may actually interfere with each other. In some instances it may be necessary to disconnect the AC ground from various pieces of equipment which may violate the codes and use only the technical ground. A better approach may be to harden the AC grounding system by insuring that it is tied to the building structure often. Install separate ground cables from rack areas to AC distribution ground points, and a separate heavy ground cable from the main AC service to a ground system beneath the building if new construction is involved.

Air Conditioning Design Criteria—General

The optimum summer and winter design conditions to be maintained by the air-conditioning system is 73° F dry bulb and 50% relative humidity for office and general use areas. However, these conditions may need to be varied depending upon (1) the side of building, (2) whether there are windows, (3) geographic location in the country and, of course, (4) in craft and technical areas where temperatures may need to be lower and humidity more closely controlled.

Air-Conditioning Loads

Studio and production lighting for television constitute the major portion of the heat gain and can exceed 75% of the total cooling load requirement. The unit lighting load requirement in the production area of the studio can equal 50 to 60 watts per square foot of floor area and in many cases this load can occur in any part of the studio since the production area and audience accommodations usually are flexible.

The studio air-conditioning system should be separate from the rest of the building so that it can be turned off or down during nonproduction times.

It is desirable to have a separate system or systems for the technical areas with ability to cross-connect to the studio or office system in case of failure of one system.

Transmission and solar heat gains are minimal since the exterior walls of television studios are well insulated and windowless for lighting and acoustic reasons.

Another contribution to the cooling load results from occupancy heat gain and the fresh air load. The fresh air requirement should be based on either 15 CFM per person, or the equivalent of one air change of fresh air per minute, whichever is greater. Fresh air quantities must conform with all code requirements.

Methods of Air Distribution and Noise Control

Proper air distribution and air movement in broadcasting studios are of critical importance. Systems should be designed so that within a zone of up to 12 ft above the floor, an air movement of 25 fpm is not exceeded. Air velocities exceeding the 25 fpm causes drafts and movement of performer's hair and clothing, and stage props which are often built of light materials such as thin canvas and plywood.

The air supply should be introduced at a level above the movable lighting grid system to prevent interference with closely spaced lighting system batten strips. Low-level return grilles located at the perimeter of the studio, in principle, would be desirable. However, due to the nature of studio operation, the grille uses valuable wall space and could be blocked off by the cyclorama curtains or by studio props, which would result in an ineffective return air system. It could also be a possible source of noise generation. Locating the return air outlets at a level above the air supply will tend to relieve the neutral zone before it can heat the ceiling and radiate heat downward. Proper location of return air grilles and maintaining low velocities will reduce the problem of air system "short circuiting."

Sound power levels (SPL) and noise criteria (NC) ratings for studios are of utmost importance and unless proper consideration is given to this problem, will result in an acoustically unpleasant studio. Noise level should be within a range of NC 20 to NC 25, so as not to interfere with studio performance, particularly during scenes where there is no conversation and no background sound effects. Duct velocities should be designed for approximately 400 feet per minute (fpm) within 10 ft of diffuser or register opening, 525 fpm within 10 ft to 30 ft from opening, 700 fpm within 30 ft to 50 ft of opening, and 800 fpm within 70 ft to 90 ft of opening.

All air ducts (supply and return) should be acoustically lined for sound attenuation and the sound power level of all outlets should be carefully checked to insure that it does not exceed the decibel rating at the end of the duct run, otherwise it will become additive (logarithmic) and negate a portion of the acoustically treated duct.

All exposed pipes should be insulated and all air ducts should be externally insulated to eliminate reflected sound in studios. Where ducts and pipes pass through walls or floor, the openings should be sealed with acoustical sound-deadening material. Ducts and piping should be suspended from vibration isolators. Where ducts and pipes pass through studio walls, flexible pipe and flexible duct connections of appropriate size should be provided to reduce the transfer of acoustic energy into the studio.

Any hole, pipe, or conduit that is not properly treated will allow the transmission of sound into the studio and effectively defeat the other expensive sound isolation measures that have been incorporated. Remember this when making modifications to studio facilities after construction is completed.

Mechanical equipment should be located remotely from the studio to eliminate transmission of sound and vibration. All equipment should be properly supported from vibration isolators. Sound traps for sound attenuation and flexible duct connections to prevent transmission of vibration should be provided for all air-handling apparatus.

Hire an acoustic consultant to make measurements of the installed studio air-conditioning system in order to insure the system meets design requirements.

Type of System and Control

Each studio should be served by its own air-handling apparatus and should consist of supply and return fans, filters, heating and cooling coils, sound traps, and a purge exhaust system.

The arrangement and selection of the component parts of the air conditioning system are highly dependent upon the economics, space conditions, and geographical location of the project. Special air flow schemes should also be considered. For example, an "economizer" cycle utilizing 100% fresh air during moderate seasons, or a fixed percentage fresh air system can be used. Preheating coils and reheaters may be required depending upon the percentage of fresh air used, the geographical location, and the outside humidity. Heating coils can be of either the steam or hot water type and cooling coils should be of the chilled water type.

In areas where freezing outside temperatures are experienced, special arrangements must be incorporated to prevent possible "freeze-up" of preheating coils and chilled water coils. Reheaters are often used in individual areas that have less heat buildup than a larger nearby area.

A separate purge exhaust system should be provided to permit the studio to be evacuated during periods when it is not in use. During the purge operation, the system should provide 100% outside air without attempting to maintain studio design conditions. Where an economizer-type cycle is provided, purging of the area can be accomplished by resetting controls to 100% outside air.

The control system should be arranged to control studio temperature, and where facilities are provided for audience participation, additional humidity control should be provided.

The installation of a supervisory data center provides operational supervision of the system and normally includes some remote controls for resetting of space temperatures and humidity, starting and stopping of air handling system, "read-out" of other pertinent air and water temperatures, and alarm indication.

A pneumatic temperature control system should be provided with standby air compressor with automatic cut-in features.

Stand-By Operation

Master control rooms, videotape rooms, editing rooms and the newsroom generally operate on a 24-hour basis, and an air-conditioning failure cannot be tolerated for more than a few minutes. A separate air-handling system should be employed to serve various combinations or these areas with provisions for standby equipment or cross-connection in the event primary equipment fails. This can be accomplished by interconnecting the ducts and appropriate dampers, with another air-handling system serving a noncritical area in the building such as office areas, thereby permitting the technical areas and news room to be satisfied during an emergency period. Another desirable feature to be incorporated in the system is provision for handling 100% fresh air in the event refrigeration equipment becomes inoperative.

Multiple refrigeration units and boilers should be provided so that in the event a single unit becomes inoperative, partial operation can maintain conditions in critical areas.

If used, the compressed air system that serves the videotape machines should be provided with standby compressors to insure continuous operation. Compressors that are water cooled can operate off the chilled water system and arranged so that in the event there is a loss in chilled water pressure, the cooling system will automatically switch over to city water.

In the event of an electrical power failure, an emergency generator should start automatically to maintain operation of the boilers, heating pump, and air-handling system serving the master control room and videotape room, pneumatic temperature control air compressor, and air compressor serving videotape machines.

STUDIO ACOUSTICS[2]

Introduction

When a microphone is opened in a studio or combo control room, it is expected to pick up only the voice of the on-air talent. But the microphone also picks up building rumble, office noise, air conditioning noise, and unwanted reflections from the room walls. The resulting broadcast signal can sound hollow, dull, or "boomy" instead of bright and clean. Poor acoustics can diminish intelligibility, and invites listener fatigue. A studio with good acoustic design and treatment is easy to "mic" and sounds good without the need for extensive audio processing. The sound from a studio with poor acoustics cannot be fixed with electronic processing.

Creating the ideal studio acoustic environment is often an expensive undertaking. It requires the expertise of an acoustical consultant working in cooperation with the architect, both of whom must be familiar with broadcast and recording studio construction. Studio designers often find it necessary to recommend types of construction which are several times more costly to execute than standard office construction. However, the additional expense will result in studios and control rooms that will be well isolated from intrusive noise and vibrations, and will project a clean and crisp "on-air" signal from live talent, free of unwanted resonances and echoes.

Estimating Costs

Many of the factors that affect the costs of acoustic construction are actually under the station's control. A Top 40 radio station whose talent works only a few inches from a single microphone, relies less on room acoustics for its live "sound" than does an All-Talk format station requiring several open microphones to pick up guests at a variety of distances. Because broadcasting is considered high fidelity, studio acoustics must be closely controlled.

Note: The more audio signal compression and limiting a station employs, the more evident will be the annoyance of poor studio acoustic characteristics. Conversely, the compression on stations with better studio acoustics will be less noticeable during live broadcasts and recordings.

When planning new studios, significant cost reductions can be realized by selecting a site which is removed from street, rail and air traffic, and manufacturing equipment. Within the building, keep the studios away from noise generating sources such as pumps and air conditioning equipment. Be sure to consider ceiling height, as large HVAC ducts are required to keep air flow noise at a minimum. If insufficient space exists for the installation of the large insulated ducts, expensive alternatives to solve the acoustic problems will be required.

Note: When initially sketching a studio floor plan, remember that studio walls occupy space and should not be drawn as single lines. All partitions have some thickness which will take up floor space. It is not uncommon for some acoustic partitions employing internal air spaces to exceed a 16" thickness. If space is not a premium, such constructions are generally more economical than thinner partitions with similar transmission characteristics. Absorptive surface treatments can easily add an additional 4" to 6" to each side of the wall profile.

The control of the transmission and reverberation of sound discussed below are two acoustics disciplines which apply to broadcast studio and control room environments. They are independently designed and require completely different methods and materials in creating studios suitable for broadcast and recording. It is important that each be understood, since confusion between the two can result in aggravating an existing problem.

Controlling Sound Transmission

Sound transmission occurs when a source of sound is carried from one area to another. The vehicle of transmission may be a direct path, a flanking path around an object, or a path carrying vibration. The most economical and practical means of limiting the transmission of noise usually requires the control or reduction of it at or near the source. Disturbances with low frequency components are the most difficult and expensive to contain. Completely enclosing the noise source in a highly absorptive room, and isolating its vibrations from the building structure, will limit the ability of noise to intrude upon other areas.

It is often cheaper to isolate the sound at the source than to isolate several rooms from the building structure. For example, if footfalls on a cement floor adjacent to or above the studio are causing problems, try carpeting those areas before attempting to increase the transmission loss of the studio construction.

Where treatment at the source is not practical, the most effective way of isolating an area against noise intrusion is to employ floating construction. With floating construction as shown in Fig. 1, the noise sensitive area is actually a room built within a room. The important aspect of this type of construction is that mechanical contact between the inner room and outer room is kept to a minimum, with the construction of both rooms creating as complete and uninterrupted an envelope as possible. This is most commonly achieved by resilient floor supports and suspending the ceiling of the inner room on springs or neoprene isolators, hence the term "floating".

Floating studios in radio stations generally are simpler because the inner and outer walls have fewer connections between them. Television studios, however, have ceilings that are significantly higher and resilient sway brace connections between the inner and outer walls may be required to stiffen the construction, and prevent buckling. Television and recording studios often have a variety of utilities provided in the

[2] This section was contributed by Alfred W. D'Alessio, Northeastern Communication Concepts.

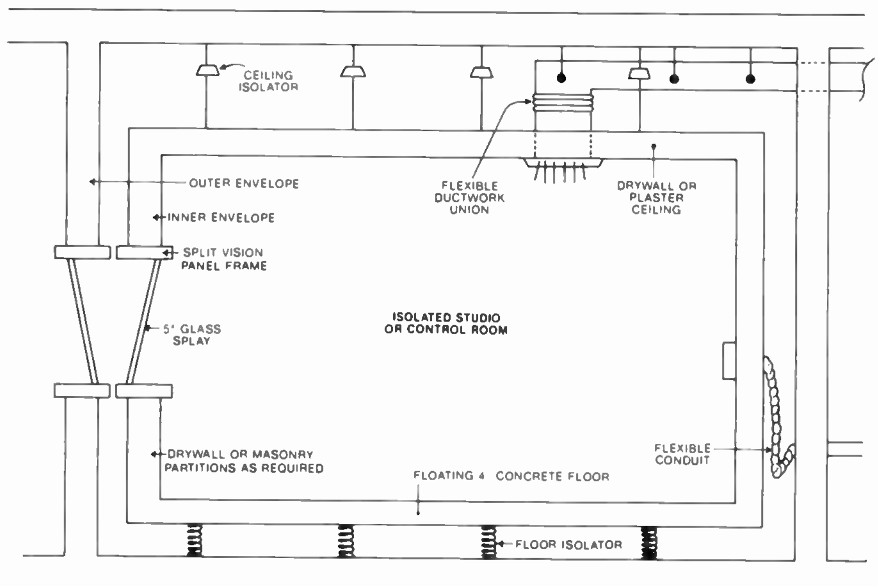


Figure 1. Floating construction schematic showing mechanical isolation between inner and outer rooms.

studio, such as electrical lighting and water connections, in addition to the audio circuits that must be isolated. In addition, there are more windows and doors in television studios that must receive special attention.

Depending on the intensity and proximity of the noise source, and the degree of residual noise which can be tolerated, some or all of the aspects of floating construction can be employed in designing a broadcast studio or control room. The resiliently isolated floor and ceiling is necessary to prevent structure-borne sound and vibration from entering the interior space. The double walls with the resulting air cavity between them are actually more effective than a single partition of the same mass. *In order to preserve this advantage, no service such as air conditioning ducts, plumbing, or electrical conduit should be allowed to make a solid mechanical contact between the outer and inner rooms. The use of flexible connections within the cavities is extremely important.*

When spring or neoprene isolators are employed for vibration isolation, the effectiveness of the isolator is measured by its nominal deflection, the relation of its natural period to that of the disturbing vibration, and the stiffness of the structure supporting it. Therefore, isolators with the proper characteristics must be specified. Isolators which are too lightly loaded will provide little isolation, while isolators which are too compliant may cause instability, or bottom out. If the isolation system exhibits a natural resonance equal to the disturbing frequency, the transmission of the disturbing frequency can actually be amplified.

One often overlooked source of vibration which may induce unwanted disturbances into adjacent areas is the control room monitor speaker system. Suspending the loudspeaker enclosure from the wall or ceiling on a resilient mounting will significantly reduce the structure-borne transmission of sound from the monitor.

Acoustic Materials

The materials used for the reduction of sound transmission in walls, ceilings, and floors are typically heavy, dense, and thick. Factors to be considered when designing a barrier for maximum transmission loss include the stiffness, resonance, mass, isolation, cost, and construction details of a partition. Typical acoustic partitions may be composed of brick or masonry block, poured concrete, lead, multiple layers of gypsum board supported by metal studs, or a combination of these. Regardless of the materials used, the most efficient barriers employ an integral airspace or cavities filled with a damping material such as fiberglass bats.

Isolation at the lower frequencies is more difficult and expensive and usually requires stiffer, thicker, and heavier construction than that needed for higher frequencies.

The single most important aspect in the construction of acoustical barriers is sealing. *Regardless of the materials used, the fit of the individual components with each other and the existing structure must be tight.* All joints should be filled and sealed with a resilient, nonhardening caulk. The smallest openings can have serious consequences, and render otherwise expensive construction no better than a far cheaper counterpart.

The table below shows how the sound insulation value (R value) of a hypothetical partition at some mid-frequency, is reduced due to acoustic leakage through various size openings.

Example: Partition with R value = 60 dB

TABLE 1
Reduction in the sound insulation (R) of a partition due to acoustic leakage through various size openings.

Size of Opening	Resultant R
0.0%	60 dB
0.1%	30 dB
1.0%	20 dB
10.0%	10 dB
50.0%	3 dB

Controlling Reverberation Time

Separate and distinct from sound and vibration isolation, is that part of acoustics which governs how a room "sounds" when a sound originates in it. This is almost entirely a function of the room size, proportion, and the ability of its contents to reflect, diffuse or absorb sound of differing frequencies. To complicate matters, all these parameters interact with each other in determining the reverberation characteristic of a room.

The major measurable acoustic characteristics are a combination of:

- The time that it takes a sound to decay 60 dB within a room, once the source of that sound is terminated (T60)
- How T60 differs with frequency
- How uniform the decay rate is
- The ratio of early and late reflections
- The natural room resonances (modes)

Room tuning is both an art and a science, and research continues to more fully understand the process. While divergent philosophies are often involved, most consultants agree on the following guidelines for broadcast and recording studios:

1. The smaller and more symmetrical a room is, the more noticeable will be its undesirable resonances. This is why many television announce booths sound more like stuffy little phone booths.
2. Avoid exceptionally long and narrow proportions, square rooms, rooms with concave walls, or rooms with a ceiling height equal to the height or width. Splayed walls and ceilings are dramatic, but necessary only as opposing surfaces which cannot be covered with mid and high frequency absorptive material, such as large vision panels and glass doors.

Controlling the T60 of a room yields the most dramatic results. A medium size radio studio with a T60 of approximately 0.3 to 0.4 seconds from 100 Hz to 6 kHz will yield a pleasant acoustic environment.

Unlike the massive solids that are used for acoustic isolation, the most common absorptive materials are light and porous. The most common materials that are commercially available for absorbing sound are carpeting, acoustic tile ceilings, and fiberglass or polyurethane foam wall panels. The difficulty in using these materials in broadcast studios is that they provide only mid- and high-frequency absorption.

Note: The exclusive and excessive use of these materials can cause a studio to become "boomy," by having a long T60 at low frequencies in proportion to the short T60 at the higher frequencies which these materials absorb.

An acoustic consultant can specify the design of resonant slot, hole, and panel absorbers, as well as extra thick mineral fiber materials to absorb low frequency sound. Boominess can also be decreased by adding a thick fiberglass blanket above a lay-in tile ceiling, and by using commercial absorptive materials in their thickest available form. Applying three or four foot widths of these materials with two or three foot spacings between them may also help balance the T60 of low and high frequencies. However, to avoid reflective echoes, no hard untreated surface should ever oppose another, either parallel or at an acute angle to it.

In combination control rooms ("combo" studios), often used in radio stations, mechanical equipment such as cartridge tape machines and reel-to-reel tape decks should be surrounded with as much absorptive material as practical. This will help absorb some of their mechanical sounds which might otherwise be reflected toward the host's microphone. Also, avoid placing the console microphone position too close to a vision panel. Whether omnidirectional or cardioid, any conventional microphone requires a free field behind it. A reflective surface behind the microphone such as a window or script panel will "color" the sound it picks up from the front.

Large television studios may derive as much of their acoustic characteristic from their sets and backdrops as from any materials purposely installed for acoustic purposes. However, because any portion of a studio wall may be exposed at one time or another, it's a good idea to cover the walls with absorptive mineral bats. The bats should be protected by a wire mesh, to keep them from disintegrating when props and sets are stored up against them.

The ceiling, above the lighting grid, should also be heavily absorptive, especially since large portions of the floor will remain reflective. It is important to keep the T60 of the television studio quite low to minimize the transmission of camera and crew noises, and to permit greater distances between talent and microphone without an "off-mike" quality. Hard, concave, acoustically reflective sets must be avoided, since any combination of these shapes tends to reflect unwanted sound toward the talent microphones in front of them.

The most reliable rule of thumb in acoustics is that treating low frequency problems is always more difficult and expensive than mid- and high-frequency work. If the budget is tight, always assign top priority to *sound transmission* considerations. Once the facility is built, little can be done to make up for economies made in the basic construction, while considerably more flexibility for improving interior room acoustics will remain.

Sound Locks

Once the design of the partitions has been decided, an entrance with the same sound insulation characteristic will be needed. There are few practical doors commercially available which meet or exceed the R values of the most common partitions used in broadcasting. Therefore, a sound lock entrance scheme is often required, which will attenuate sound through two doors and an intervening air space. (See earlier section on Sound Locks in this Chapter.)

The sound lock is a small vestibule between the studio and a hallway or other room that provides access to the studio as shown in Fig. 2. By entirely covering the walls of the sound lock between the studio and external doors with absorptive material, the efficiency of the sound lock will increase, decreasing the effect of opening one door. The use of a sound lock not only reduces the insulation requirement for each door, but also provides passage in and out of the studio or control room without exposing it to the full noise and disturbance of an adjoining area. Doors with good acoustic characteristics are available either in metal or wood, both of which feature a sound retardant core.

Like the partition requirements, the sound lock door must seal tightly within its frame. Lack of a good seal wastes whatever investment is made to procure a good sound rated door. Commercially available compression and/or magnetic type seals are recommended at the head and jambs of the door, while a mechanical drop seal and step saddle should be employed at its threshold. Some expensive prehung doors feature integral seals, and can lift hinges which actually lower the entire door into place tightly against the threshold saddle when it is closed. The greatest advantage of these doors results from the fact that they never have to be pushed shut to make a good seal all around the door. All other types require an oversize door check

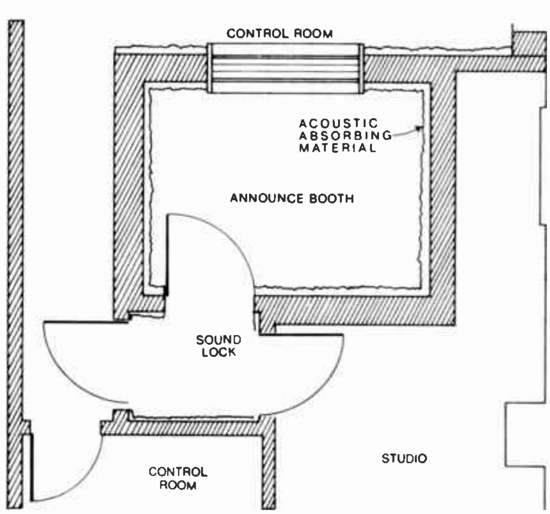


Figure 2. Soundlock entry to a small announce booth.

(pneumatic or hydraulic closer), and a latch to properly compress the seals when the door is closed.

Windows

Studio windows require special consideration. In order to match the transmission loss characteristic of the partition surrounding it, a studio window must be double glazed. The two panes of glass should be of different thicknesses to minimize the coincidence of each pane's deficiencies. The glass should be mounted in a resilient neoprene or felt lined channel, and caulked tight. Some provision should be made for removing either piece of glass for replacement or cleaning. The glass should be mounted with a 5° outward splay at the top to reduce acoustic and visual reflections.

Where a high degree of isolation is required, acoustic laminated glass should be employed. Each pane of this glass consists of alternating layers of plastic and glass; and it is lighter, less resonant, and a better acoustical barrier than standard plate glass.

Welded hollow metal frames are preferred over wood for both doors and windows for their stability. However, the void within them should be packed with either a cement-like or mineral fiber filler. It is important that the frames not connect the vertical partition sections in floating room or double wall construction.

Modular Rooms

In large cities where the cost of construction labor is high, small radio studios, announce booths, and control rooms may be economically built from modular components. The modular rooms are constructed from 4" thick, prefabricated hollow steel panels. A modular room usually includes a self-supporting ceiling and an integral floor. The entire room is assembled on rails which can be isolated with springs or neoprene from the building structure. Most companies offering modular rooms also include double and triple partitions, doors, windows, conduit, ventilation ducts, and a performance guarantee as available options. Because modular rooms can be disassembled and relocated, they are not usually considered leasehold improvements, and may offer certain financial benefits in addition to being somewhat portable.

Summary of Section on Acoustics

Regardless of the design, the installation contractor must be familiar with the stringent requirements of acoustic construction. The efforts of the general construction, electrical and mechanical trades must be coordinated by the architect both in the design and execution of the studio project to assure that no conflicts compromise the acoustic plan. The job should be inspected during construction by the acoustical consultant, to check for potential leakage paths brought about by existing field conditions, and conformance to acoustical details. If the station or studio is to be situated in an office building, or other shared location, don't forget that other tenants may be annoyed by high-level monitoring, tape rewind noises, and other sounds peculiar to broadcasters. Therefore, during the design phase, if the acoustic consultant recommends protection for adjacent building tenants, the extra effort will result in happier neighbors.

ESTIMATING FACILITY COSTS

General

Estimating the cost of a broadcast or recording facility is a complicated process. In addition to the technical equipment costs, which can be easily estimated, there is the cost of the physical facility construction or renovation which varies according to:

- The area or location in which the facility is to be built or renovated
- The quality of materials and level of attention to detail
- The time allowed for construction and time of year of construction
- The employment of energy saving techniques for long term savings
- Whether an architect, consultant, or station staff will be supervising the entire project

The costs of construction continue to increase. Materials and labor must be examined carefully by competent architects and construction engineers for the specific structure and area involved, if there is to be any accuracy in budgeting the project.

A major factor contributing to cost overrun on projects is inadequate initial plans and the resultant change orders during construction. Therefore, the owner should provide a supervisor directly involved in the project on a daily basis. While it is not necessary for the supervisor to be a construction expert, it behooves management to select someone who is skilled in project management, has a general understanding of the project, and the time to devote to it.

Standardized Construction Specifications

Most contractors employ standardized methods for estimating costs of construction and breakout costs in a standardized manner. An example of a standardized list of the different construction areas, disciplines and trades is the "Masterformat" published by Construction Specifications Institute. All construction is divided into numbered 16 super-categories and each super-category is subdivided into more specific construction type and third and fourth levels of specifications. For example, a contractor would refer to Section 16780 for Television Systems (Broadcast Video Systems) under the subhead of 16700 for Communications which is under Division 16 which is Electrical.

Another example of the use of the standardized codes would be for the major heading of 15 for Heating, Ventilating and Air Conditioning system under which are sub-headings (or products) including 15550—Heat Generation, 15680—Water Chillers, 15750 Heat Transfer, and 15850—Air Handling.

When working with competitive bidding on the project the standardized code numbers provide the means for comparing the costs for a given kind of work.

Because radio, television, and recording facilities do not fit neatly into any of the standardized construction practice categories, several organizations have attempted to provide rough estimates for facility construction.

1. The National Association of Broadcasters produced estimates in August 1990 for the construction of FM radio stations. Costs in the NAB brief for technical equipment only, buildings not included, showed $310,000 for a Class A FM station, $426,000 for a Class B FM station, and $770,000 for a Class C FM station.
2. The Public Telecommunications Facilities Program of the National Telecommunications and Information Administration in 1989 estimated the equipment costs for a public broadcasting station at $346,000 for a Class A FM station, $462,000 for a Class B and B1 FM station and $807,000 for a Class C, C1 and C2 FM station.

Construction Examples[3]

The two examples that follow illustrate construction costs for television facilities in large cities, with construction being completed between 1980 to 1983. The costs shown should be increased by 50% to reflect the increase in construction costs since then.

In each of the cases, a general contractor was retained and coordination was handled by station staff engineering personnel. Electrical costs include power to all equipment, but not wages paid to station staff technicians who installed the broadcast equipment. No attempt was made to estimate architectural and consulting fees.

Example A, New Studio Facility

For this addition to the studio facility there were three basic requirements which had to be met (see Figs. 3 and 4):

- Provide Corporate and Television Division offices
- Provide expansion for the News Department
- Provide a studio and ancillary facilities capable of producing syndicated shows as well as increased local station production

The area available to build in was limited because the property was divided into two types of zoning. This required the expansion to be planned for the area in which the zoning would allow such construction. The lack of public transportation presented a zoning requirement for additional parking.

The timetable for the construction of the addition was dictated by the desire to enclose the building before winter set in. Additional emphasis on coordination was provided by having both the corporate and Television Division personnel sharing space with the station. For all practical purposes the construction took approxi-

[3] Examples A and B were contributed by Richard J. Anderson, Metromedia Television.

mately nine months, not including engineering and construction drawings.

Example A. Cost for New Building—Studio with 25,000 Sq. Ft.

Activity	Cost	Per Sq. Ft.
Demolition & Site	$ 111,000	$ 4.44
Structural	255,000	10.20
General Construction	1,694,600	67.68
Decoration	93,700	3.75
Heating/Ventilating	315,000	12.60
Electrical	225,000	9.00
Plumbing	119,000	4.76
Architect, Elect. & Mech. Engineering	396,000	15.84
Special Woodwork & Built-Ins	55,000	2.20
Special Acoustical (including ceilings)	172,300	6.89
	Total	$137.46

Example B. Transmitter Building

In this example a transmitter site is to be relocated from downtown Houston to a site outside of town, and a new transmitter building (see Figs. 5 and 6) and towers are to be constructed. The following goals were set:

- Relocate from downtown site which was being surrounded by buildings taller than the existing transmitting antenna
- Locate in same general area as the existing Houston TV stations
- Permit an antenna height of 2,000 ft AMSL
- Provide a Principal City Coverage over the same area as the old transmitter site

A site was found that met FAA approval for a 2,000 ft AMSL antenna height. Two towers were to be constructed 100 ft apart. The foundation design included both towers and the transmitter building. The transmitter and antenna were selected to provide the best compromise between minimizing primary power consumption and providing five megawatts radiated power.

Example B Cost for New Transmitter Building—2,300 Sq. Ft.

Activity	Cost	Per Sq. Ft.
Structural	$ 29,000	$ 12.72
General Construction	77,000	33.77
Decoration	6,000	2.63
Heating & Ventilating	15,000	6.59
Electrical (including auxil iary generator, pad and roof)	67,898	29.78
Plumbing	10,000	4.39
Architect, Elect. & Mech. Engineering	14,730	6.46
Special Woodwork & Built-Ins	4,000	1.75
		$ 98.09

Typical Construction Cost Ranges

Activity	Cost Per Sq. Ft.
New Studios and Offices (new structure)	$82 to $138
Conversion for Studios & Offices (in existing structure)	$40 to $65
Transmitter Building (new structure)	$61 to $98

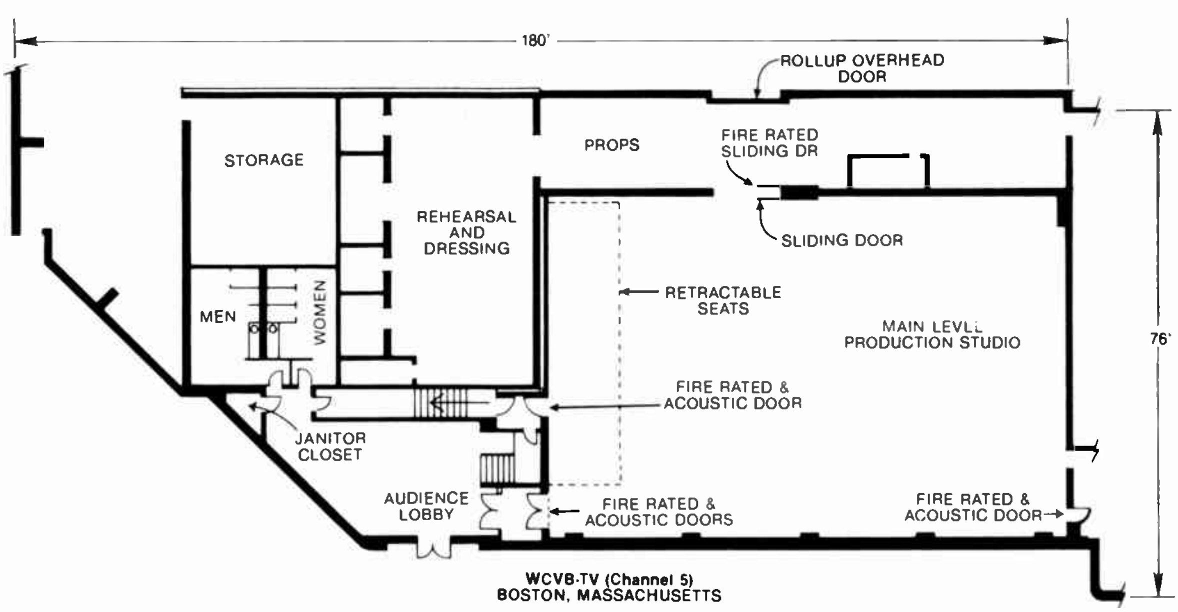


Figure 3. Production studio addition.

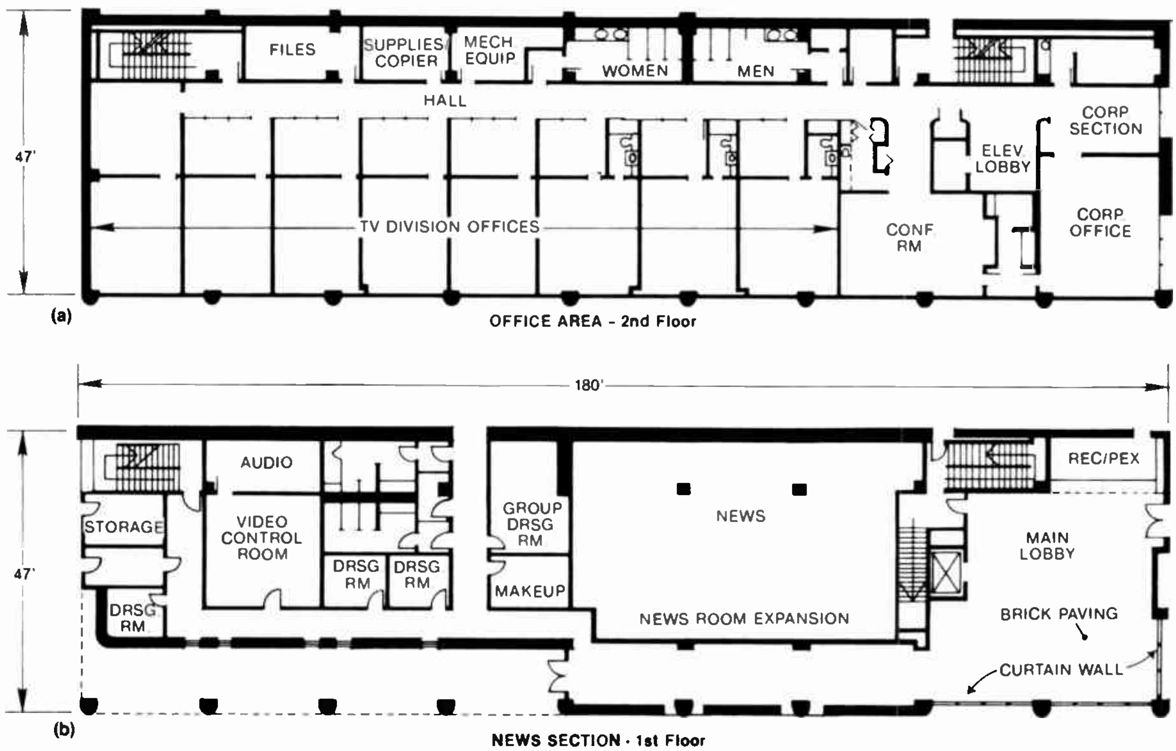


Figure 4. Facility layout.

When divided into separate elements the breakout is as follows:

Activity	Cost Per Sq. Ft.
Demolition and Site Work	$1.80 to $3.00
Structural	$9 to $14
General Construction	$18 to $50
Heating, Ventilating and Air Conditioning	$9 to $16
Electrical	$5.40 to $18.00
Plumbing	$1.35 to $4.00
Architect and Engineering Fees	$2.25 to $7.00
Special Acoustical	$2.70 to $5.00
Woodwork and Built-Ins	$1.08 to $3.00
Office Furniture	$2.70 to $7.00
Decoration	$3.60 to $7.00

Note: Electronic equipment installation and wiring is not covered by any of above estimates. A good general rule is to allow 15% to 25% for this in addition to basic equipment cost. Variations are due to location, personnel, working rules, etc.

AM AND FM STUDIO FACILITIES

Building Planning

One of the prime requisites for a successful broadcast station is the careful layout of studio, production, and administrative areas to achieve maximum effectiveness of space and personnel. The following four typical layouts depict a range of facilities from a small market minimum staff facility to an arrangement suitable for a large metropolitan operation employing a full complement of personnel. Each floor plan is handled differently according to the needs of different size stations.

In these examples, the control room, studio, and production facilities are in a centrally located core area although they need not necessarily be there. The suggested sizes of these areas should be considered as minimum from an operating standpoint with normal equipment complement. The layouts are presented as a guide for planning a modern, functional radio facility with considerations given to size of market, staff, and programming requirements.

Plan One. Small Radio Station

With approximately 1,800 sq. ft. (see Fig. 7) this floor plan provides adequate space for the small AM or FM station with a minimum staff. Since smaller staffs have several responsibilities, partitioned general office space is omitted in favor of a large news, recording storage, and general-use area at the rear of the building.

The transmitter and engineering workshop area is placed next to the control room, with a window recommended for a clear view of the transmitter meters. Alternate core area layouts are shown.

The building is of brick and plaster fascia, and includes a glass curtain wall. The building price will

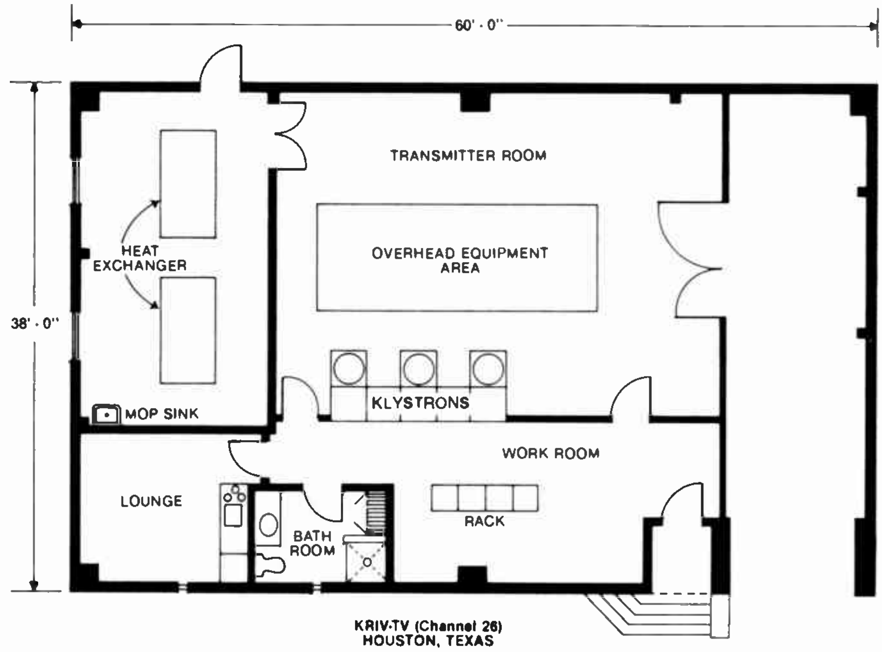


Figure 5. Transmitter building.

vary considerably depending on area construction costs; but a typical range is $60,000 to $75,000.

Plan Two. Medium Radio Station

In medium-sized stations, office space is increased for sales, promotion, and programming activities but without a substantially larger technical core area. This floor plan (see Fig. 8) takes the "small station" layout and expands it to approximately 2,500 sq. ft More room is provided for the sales staff and clerical help, and an impressive office for the general manager. Studio and control room space is slightly larger in anticipation of more equipment and activities in these areas. The news director, recording library, and chief engineer gain office space. Alternate core area layouts may be employed, and a few suggestions are indicated.

The building includes brick walls, weathering steel columns, and fascia, with a dark glass entrance and glazing strips. Cost is approximately $65,000 to $90,000, but may vary considerably, depending on construction costs in a particular area.

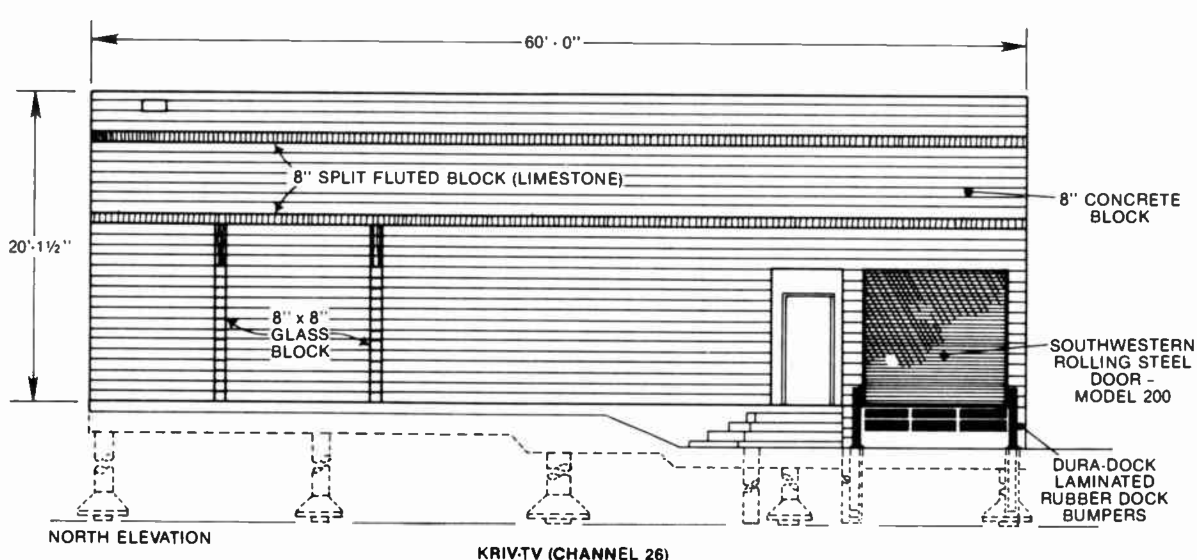


Figure 6. Transmitter building.

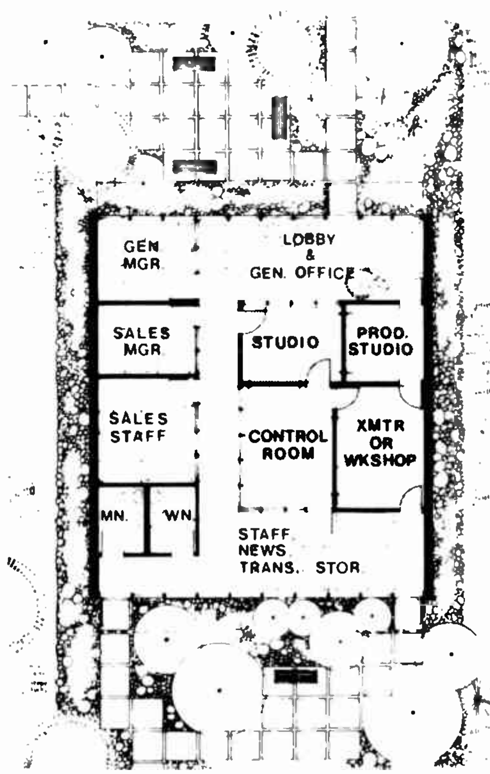


Figure 7. Plan 1—Small size AM station.

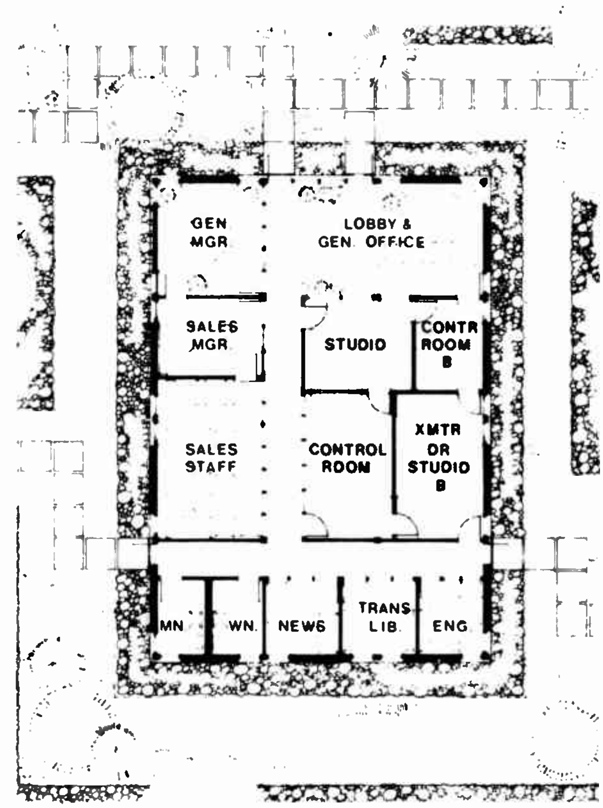


Figure 8. Plan 2—Medium size AM station.

Plan Three. Larger Radio Station

In this plan (see Fig. 9), the technical core area is adequate for two full-size control rooms, each with a large associated studio. Control rooms are separated by the transmitter, automation, or engineering area. This floor plan includes approximately 3,150 sq. ft and is suggested for stations planning both AM and FM operations. Additional office space is also allocated for the larger staff in this station. See alternate core area floor plans for additional layout ideas. This design features a mirror glass curtain wall building set in a reflecting pool. The building costs approximately $90,000 to $135,000.

Plan Four. Large Station

This 4,300 sq. ft studio/office complex (see Fig. 10) is an impressive broadcast center. Of primary importance is the location of all control room and studio space in the center of the building, eliminating the problem of outside traffic noise in a metropolitan area.

Operating personnel are assigned to the rear office areas, and the news room is strategically located near the control rooms and an outside exit to the news cars and trucks.

The building is made of exposed concrete and features a dark glass curtain wall. The building cost is approximately $100,000 to $150,000.

CONTROL ROOM/STUDIO DESIGN

Many present-day radio control rooms also double as studios and are called "combos" for combination control and studio. The combo room is where most of the station's announcing is carried on and may be the only studio area in a station other than the production studio (also a combo). Combos usually contain most of the audio equipment of the station. Frequently, a transmitter and record library will also be located in the control room.

In many instances the news room will be used with live mics for late breaking news events. Both situations make the problem of acoustics more complex but, because radio sells with sound, good acoustic design is essential when the best sound possible is required.

Design factors of combo control rooms are at best a compromise of the several design factors:

- Location of the control room within the studio building
- Isolation: elimination or substantial reduction of unwanted sound and noise, both internal and external
- Construction: special walls, ceilings and floors may be required
- Reverberation control: size the rooms and consider special wall and floor coverings to eliminate or

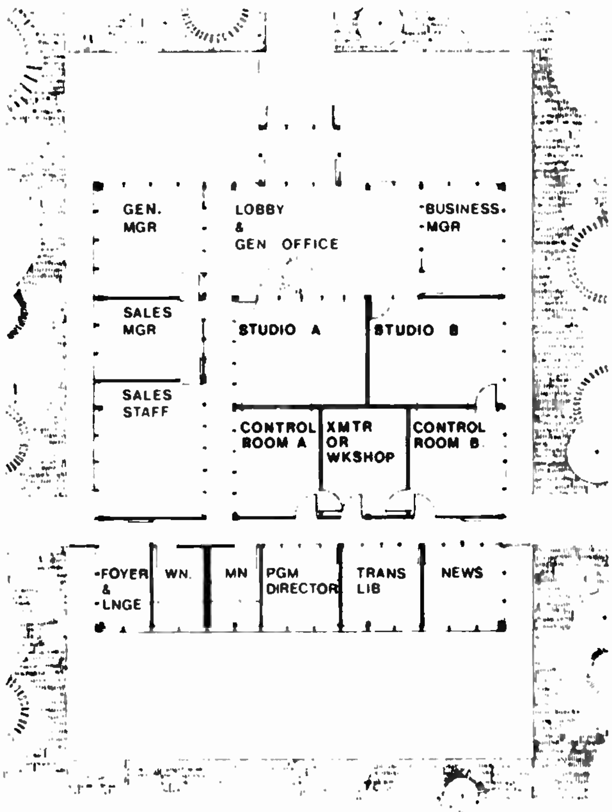


Figure 9. Plan 3—Large station. AM/FM dual control room.

reduce reflections and resonances that degrade the studio sound on the air

- Ventilation and air conditioning: technical equipment generates more heat than normal office equipment and studio guests add to the heat load which must be accommodated with the HVAC facilities
- Size and arrangement of equipment: There are as many arrangements as there are designers—plan

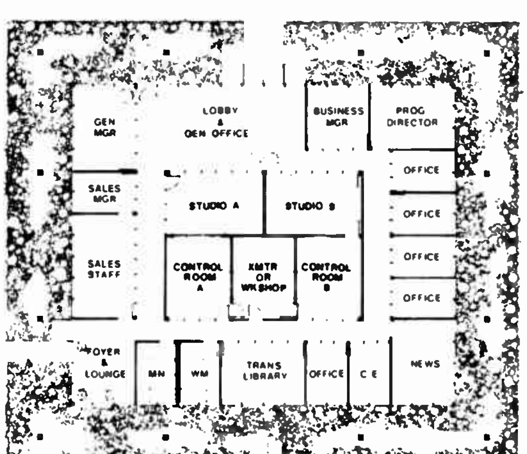


Figure 10. Plan 4—Large station. Metropolitan market full staff.

some flexibility for changes that will be needed after the facility is used for a while and for accommodating future needs

Refer to the "Studio Acoustics" section in this chapter for more detail in developing and designing a new or renovated studio facility.

RADIO STUDIOS ON A LOW BUDGET

Considerations Involved

Different sized stations require different construction criteria. A 50,000-watt AM clear channel or Class C FM in a major market and a 250-watt AM daytime or Class A FM in a rural area dictate that the small market station is going to be constructed quite different than its big city counterpart. In most cases, the selection of a studio site in a small market is dictated by what it costs to get the space. It is not unusual for space to be traded out in part or in full for advertising. Figs. 11 and 12 show examples of small radio broadcast studio and office facilities.

For the construction or renovation of a smaller facility, the chief engineer (or engineering consultant) may be presented with an existing suite of offices, a store front or even an old house that must be converted into studios and control rooms. After determining the nature of programming management intends to offer the community, the facility can be designed to meet the requirements. Most smaller stations use combo control rooms and studios, a simple production facility (which can be used as an emergency control room) and a newsroom with recording and editing facilities, and can be used for on-air use.

A larger studio can be fashioned from the lobby area, the general office area or by combining two or more management offices. With some advance planning, these areas can be acoustically treated and wired for use as acceptable studios for those occasional talent shows, church programs, talk shows or special group events that do not occur often enough to justify having a full-featured studio available. If the station is a new or growing operation and cannot afford to do all of the above, at least consider what functions may be required at a foreseeable future date. Plan now for what might be needed in two or three years.

Physical Layout

While the total amount of space may be limited, remember that combo control rooms are occupied for long periods by one or at most two people. Consider this as their office which should be large enough to house the equipment, move around, have visual access to other areas (including the outside) and be comfortable and well lit.

When laying out the studio location, draw a sketch of the available floor space and existing walls. Make a number of copies of this floor plan and start drawing in the studios, offices, newsroom, reception areas, and engineering. Make three or four versions of the plan. If the existing layout will allow the installation of new

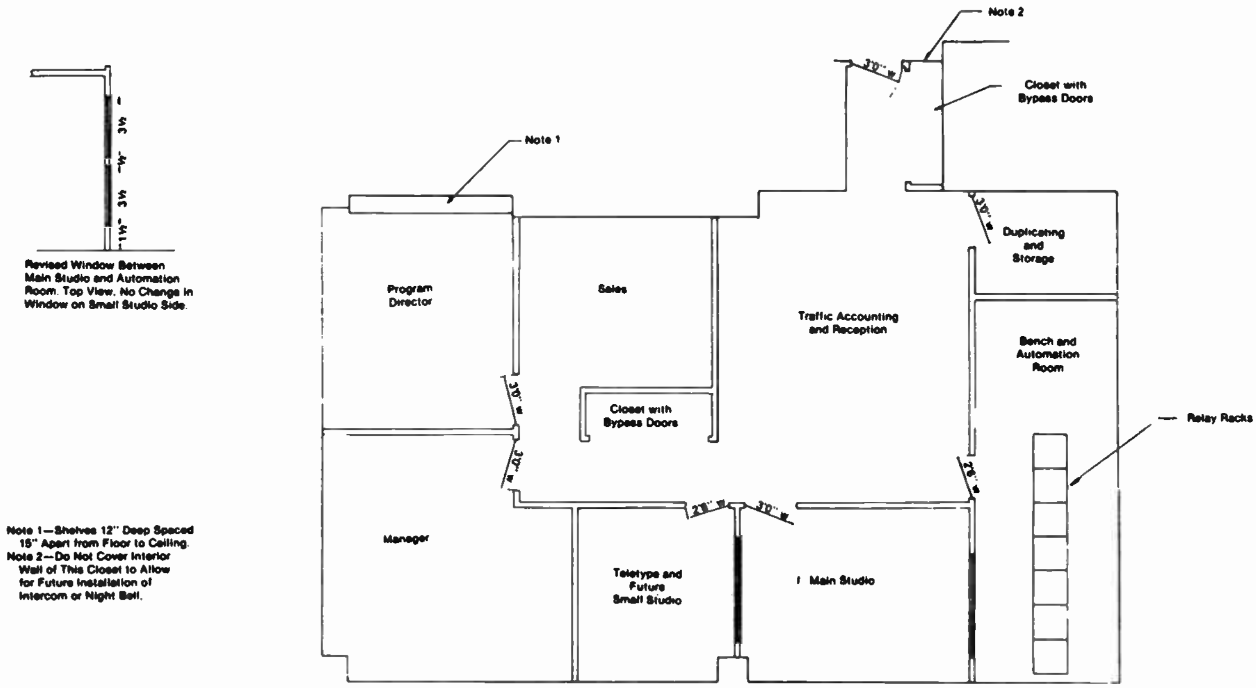


Figure 11. Studio layout for a small station.

walls, avoid perfectly square rooms. A room that is 8 ft by 8 ft by 8 ft is going to have a very definite and unpleasant resonance. Strive for 6 ft by 8 ft by 10 ft or dimensions in this proportion. Be sure to add extra wall thickness for the studios and control rooms.

Consider laying out the facility using scale cutouts of typical office furniture, studio equipment and consoles, plus storage cabinets and counter tops. Add some closets and undefined storage spaces for the future.

Keep the main studio and control room from beneath heavily traveled stairways and corridors, air conditioners, and front windows opening on busy streets. The control room floor should be very solid. The most desirable material is reinforced concrete. A wooden floor can be used if it is reinforced to prevent bouncing. Otherwise people walking on such a floor will cause problems with turntables and compact disk players. Sandbags or bricks in the base of the turntable cabinet will sometimes improve an otherwise bad situation.

Avoid placing the console up against a wall. In addition to being an acoustic problem, the operator will find looking at a wall to be uncomfortable. Locate the console in the middle of the room as a desk might be in an office. The wall space can be used for shelves, equipment, or other facilities.

Doors to studios and control rooms should be of solid construction with weatherstripping (felt, not rubber which will squeak) and a good quality door closer. Do not use a latch on the door which will click each time the door is opened and closed. Doors should not open out into busy hallways where people will walk into them as they pass by. The average door in most stations is 3 ft wide which can be reduced by 6". However, most furniture requires up to 30" to be moved through a doorway and more if it opens into a narrow corridor. The narrowest hallways should be no less than 3 ft 6". Note that hallways take up space. Every square foot of hallway that can be eliminated yields an equal amount of space that can be used for something productive.

Put windows in the studio and control rooms to the outside world. Sound locks, while desirable, are not essential, but can be combined with other purposes. Plan on plenty of lights but place the ballasts of fluorescent lights out of the production areas to eliminate acoustic noise and interference.

If the transmitter is at the studio site, locate it and the main control room so that meters on the transmitter and its associated monitoring equipment can easily be seen from the operating position.

Equipment

In a low-budget situation remember this rule of thumb, "Will the equipment to be bought pay for itself and make the station money?" The purchase of a \$200 directional microphone to do a one-time remote broadcast that pays only \$100 is poor business.

Use this logic when equipping the station. Create a list of those functions you will have to perform and spend only what is needed to get on the air and function. However, plan for the future! Design around a console with enough inputs to accommodate those

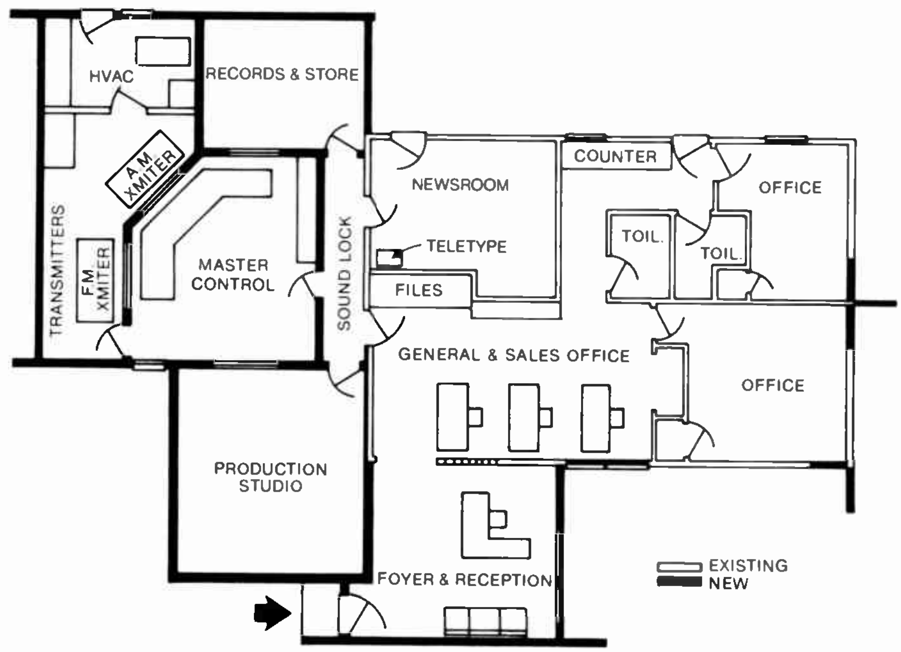


Figure 12. Renovated small station.

unanticipated needs. A four potentiometer board is not enough for most on-air operations except for simple production. Even the smallest stations should have provision for six or eight simultaneous inputs. Invest in a good quality telephone interface system. Virtually all stations now take phone calls for direct on-air use.

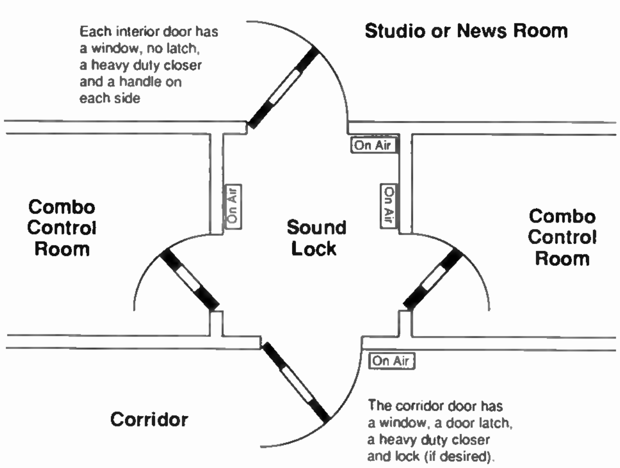


Figure 13. Sound lock design. When any door is opened, a slight vacuum is created in the lock, which seals the other doors. The first door is allowed to close which permits another door to be opened which, again, seals the other doors. The windows permit a look into the lock to determine if someone is in it before entering. The interior doors have no latch which would cause an audible click in the studio or control rooms. The heavy doors combined with the heavy duty closer prevent opening the corridor door too rapidly.

Consider used equipment with great care. While older equipment was built for long life, obtaining spare parts may be a problem and the general quality level may not meet today's needs.

Consider versatility and contingency in equipment configurations. The studio output can be operated through the automation system as a source. In the event of an emergency, the announcer can easily connect the output of the console to the STL, bypassing the automation system. In the event of an STL failure, the audio can be patched directly to the transmitter via the remote control system line or a telephone line.

Do not forget the patch panel in the control room. There is a growing tendency among engineers today to eliminate this important switching center. All high level inputs and outputs should appear here. Even console inputs should show up. This also applies to all recorders, tuners, and other audio sources. This practice makes for a very versatile operation and can save the embarrassment of dead air time if the console should fail. If this happens, a recorder can be patched directly to the transmitter, bypassing the defective elements in the system.

Consider every piece of equipment in the audio chain as a potential single point failure problem and address a plan on how to bypass it or do without while it is being repaired. Consider also a backup power generator. For a small station a modest sized generator can power the entire facility except for a central air conditioning system.

The construction of a new or renovated facility is a good time to implement a check list approach to routine testing of all equipment on an annual, quarterly, monthly, or weekly basis depending upon how often the equipment needs to be checked. The check list should also include such items as the air conditioning and heating system, backup generator, tower, antenna, lights, and the general condition of the facility and surrounding environment.

CONCLUSION

The construction or renovation of a broadcast or recording facility is a major undertaking. Only the briefest explanations of the various elements and concerns were provided in this chapter. It behooves station management to become intimately involved in the construction details or to have a knowledegable employee or consultant supervise the work. Advance planning and coordination with all departments and personnel will help design the best facility, and smooth the construction and the transition to the new facility.

REFERENCES

Master Format—Master List of Titles and Numbers for the Construction Industry, Alexandria, VA: The Construction Specifications Institute, 1988. A widely-used, uniform system of identifying materials, equipment, and systems used in construction.

5.13 Principles of Acoustics for Broadcast Applications

Lynn D. Claudy
National Association of Broadcasters, Washington, District of Columbia

INTRODUCTION

> Acoustics—"*The scientific study of sound, especially of its generation, propagation, perception and interaction with materials and other forms of radiation.*"—American Heritage Dictionary.

Reducing the above definition to its basic core, acoustics encompasses essentially *the total effect of sound*, including the characteristics of the sound source, propagation medium, and the reception of the transmitted sound. This chapter covers basic principles of acoustics relevant for broadcast applications. In addressing this subject, a review of the basics of sound wave propagation will first be presented. The behavior of sound in enclosed spaces will then be examined and design techniques for controlling sound within a room will be introduced. Sound isolation techniques, or controlling the intrusion of unwanted noises into enclosed spaces, will then be presented. The chapter offers a broad overview of acoustics, with a blend of theory and practice, with the goal of providing a balanced foundation on the subject. However, as the practice of acoustical design is still part science and part art form, the services of a professional consultant in this area will continue to be just as appropriate for the knowledgeable reader as for the layman.

THE NEED FOR ACOUSTICAL DESIGN IN BROADCASTING

The importance of the need for good acoustics in broadcasting applications cannot be overstated. Borrowing the garbage-in-garbage-out metaphor from the computing industry, it is easy to imagine that an excellent transmission system can be a wasted resource if the signal source itself is degraded. High ambient noise, improper acoustical treatment, or poor noise isolation can quickly alienate, fatigue or annoy a listener. Common sense might lead one to a different conclusion. Often one experiences poor acoustic environments in daily life and is not overly adversely affected: conversation does take place in noisy corridors, musical performances are well received in compromised settings, and so forth. However, a peculiar phenomenon, generically known as "the cocktail party effect", is that noise, excess reverberation or extraneous sounds are typically much less irritating to listeners when presented contextually in their original environments than when reproduced over radio or television service. The human binaural sense of hearing, visual and other cues allow concentration on a desired source and inhibition of unwanted sounds. A simple experiment can be set up to prove this point and provide insight as to the naming of this phenomenon. Carry a pocket tape recorder to a cocktail party and record the environment while moving from one conversation circle to another. While it is relatively easy to understand live conversations while at the party, an auditioning of the recording will in all likelihood be a cacophony of jumbled sounds, indistinguishable one from another. The lesson: acoustics, while being of significant importance in live settings, is ever more so important for radio and television broadcasting because any defects that are present at the source are magnified when the sounds are taken out of the originating environment and reproduced in another.

BASIC CHARACTERISTICS OF SOUND PROPAGATION

Sound is a wave phenomenon. The wave motion is of the longitudinal (as opposed to transverse) variety: i.e., the air molecules move to and fro in the same direction as the wave motion. Of course, there is no net motion of the air particles; sound doesn't create

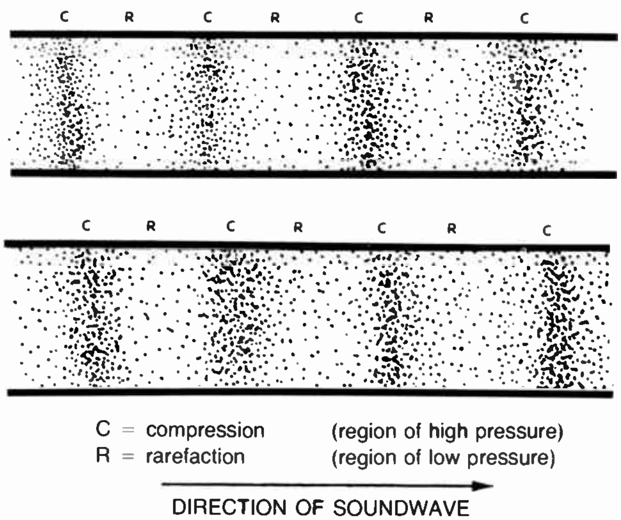


Figure 1. Illustration of sound wave propagation in a tube.

wind! Fig. 1 shows a sound wave traveling down a tube at two points in time. The distribution of air particles and regions of high and low pressure in the tube are also shown. As is evident, the air is alternately compressed and rarefied. In the second picture, at an instant of time later, it is apparent that the compression/rarefaction cycle is moving to the right.

The Speed of Sound

The speed of sound depends principally on the nature of the propagation medium in which it is traveling. Table 1 lists the speed of sound in various materials.

Looking at Table 1, it is easy to understand situations such as when banging on a pipe in the basement of a building results in the perception of a double bang in an upper floor room, due to the arrival of sound energy transmitted via the pipe itself reradiating in the room versus the slower airborne path through the room partitions and floor/ceiling structures. Also, it is worth keeping in mind that loudspeakers closely coupled to a wall structure can in some cases transmit sound via the structure to be reradiated from another point in the structure. Thus, since sound travels much faster in solid structures than in the air, sound energy can arrive at a listener or microphone before the direct sound path arrives, causing a weak but annoying pre-echo.

TABLE 1

Medium	Approximate Speed of Sound
Air (room temperature)	1,130 ft/sec
Soft Wood	11,000 ft/sec
Gypsum board	22,300 ft/sec
Steel	16,600 ft/sec
Water	4,900 ft/sec
Glass	17,000 ft/sec
Concrete	11,200 ft/sec
Hydrogen	4,100 ft/sec
Helium	3,000 ft/sec

The speed of sound at normal room temperature is approximately 1,130 ft/sec. However, the speed of sound is also dependent on temperature. At 32° F and 760 mm Hg the speed of sound has been experimentally verified to be approximately 1,087 ft/sec. A reasonably accurate simplified formula for the speed of sound is to assume a 1.1 ft/sec. increase or decrease for each degree F above or below 32°:

$$\textit{Speed of sound} = 1{,}087 + 1.1(T-32)\ \text{ft/sec}$$

where T is temperature in degrees Fahrenheit

It is worth noting that the ambient air pressure is not a factor in the speed of sound.

Relative humidity changes the density of air and has a small effect on the speed of sound. At typical room temperatures, the percentage change in velocity between 0 and 100% RH is less than 0.5%. This factor is usually ignored in practical situations.

Definitions and Sound Measurements

$$\textit{Intensity level} = IL = 10\log\left(\frac{I}{I_{ref}}\right)$$

where: I = Sound intensity (watts/m^2)
$I_{ref} = 10^{-12}$ watts/m^2

$$\textit{Acoustic power level} = PWL = 10\log\left(\frac{W}{W_{ref}}\right)$$

where: W = Acoustic power (watts)
$W_{ref} = 10^{-12}$ watts

$$\textit{Sound pressure level} = SPL = 20\log\left(\frac{P}{P_{ref}}\right)$$

where: P = Sound pressure (N/m^2)
$P_{ref} = 2 \times 10^{-5}$ N/m^2 or 0.0002 microbar

In a free field,

$$I = \frac{P^2}{\rho_o C}$$

where: ρ_o = Air density
C = Speed of sound

$SPL \approx IL$ in a free field

In a diffuse field,

$$I = \frac{P^2}{4\,\rho_o C}$$

$$SPL = PWL - 10\log(4\pi r^2)$$

when the source is omnidirectional (r in meters).

A decibel (dB) is a mathematically convenient way for expressing the ratio of two power-like quantities:

$$\text{dB} = 10\log\left(\frac{P_1}{P_2}\right)$$

where P_1 and P_2 are power quantities.

When adding decibel quantities, it must be remembered that power levels themselves can be added directly but decibels cannot. To add decibel quantities, the decibel notation must be rearranged so that the associated power levels are added and then converted back to decibel form:

$$dB_{total} = 10 \log \left[10^{\frac{dB_1}{10}} + 10^{\frac{dB_2}{10}} + \ldots 10^{\frac{dB_n}{10}} \right]$$

For an approximate rule of thumb in adding decibels, refer to the following table:

When adding two dB values that differ by:	Add to the higher value to to obtain the total:
0 or 1 dB	3 dB
2 or 3 dB	2 dB
4 to 8 dB	1 dB
9 or more dB	0 dB

SPL Attenuation as a Function of Distance

In a free field, sound intensity varies inversely with the square of the distance from the source. Hence, like light, sound follows so-called inverse square law. The difference in dB between the SPL at two different distances, d_1 and d_2, is then:

$$\textit{Difference in } \mathrm{dB} = 20 \log \left(\frac{d_2}{d_1}\right)$$

This is equivalent to a 6 dB loss per doubling of distance or a 20 dB loss per decade of distance.

Directionality of Sound Sources

In the practical case, sound sources are not omnidirectional and have an axis of main radiation. This can be quantified in the concept of directivity, or Q. The definition of Q is as follows: Q is the ratio of intensity along a given axis to the intensity which would be measured at the same distance if the same quantity of total acoustic power were being radiated omnidirectionally. The designated axis is usually taken to be the axis of maximum radiation, so $Q \geq 1$. This can be expressed as:

$$Q = \frac{(\textit{On-axis pressure})^2 \textit{ at some distance}}{(\textit{Mean sound pressure})^2 \textit{ at the same distance averaged over all directions}}$$

It can be shown that for a sound source having theoretical horizontal and vertical coverage angles of a and b,

$$Q = \frac{180}{\sin^{-1}\left(\sin\frac{a}{2}\sin\frac{b}{2}\right)}$$

Directivity is often expressed in decibel notation:

$$\textit{Directivity factor}\ (\mathrm{DF}) = 10 \log Q$$

The attenuation of SPL as a function of distance given previously can be modified to include the factor of source directionality:

$$SPL = PWL + DF - 10 \log (4\pi r^2)$$

where r is the distance from the source in meters.

Effect of Humidity on Sound Level

Humidity in the air will increase sound level loss in excess of that predicted by inverse square law alone. Fig. 2 shows a family of curves at different frequencies showing dB loss per 100 feet as a function of relative humidity (RH). In general, the attenuation increases with frequency and absorption is greatest between approximately 10% to 40% RH, decreasing above and below these levels for all frequencies.

The curves show why it is difficult to acoustically transmit high frequency energy long distances. For example, the SPL at a distance of 4 feet required to produce 80 dB SPL at 1,000 feet for 10 kHz at 20% RH is:

$$\underbrace{80 + 20 \log\left(\frac{1000}{4}\right)}_{\textit{inverse square law term}} + \underbrace{9\ \mathrm{dB/100\ ft} \times 10}_{\textit{humidity term}} = 218\ \mathrm{dB}\ \textit{SPL!}$$

SOUND IN ENCLOSED SPACES

Room Modes

For a closed tube of length L, a sound wave will travel the length of the tube and reflect, traveling back

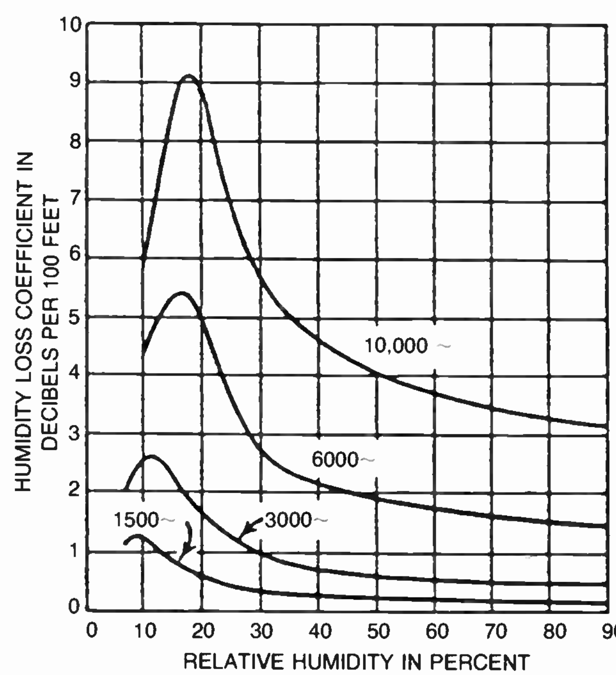


Figure 2. The effect of humidity on sound attenuation (dB loss per 100 feet).

and forth in the tube. The pressure distribution at any point in the tube at any given instant of time will be a result of the superposition of these reflecting waves as determined by the tube length, frequency, and speed of the wave. Standing waves will exist inside the tube at frequencies for which L is a multiple of half a wavelength as shown in Fig. 3.

The standing waves are characterized by strong reinforcement of sound at these frequencies. This effect is known as *resonance*. The ends of the tube will always represent displacement nodes and pressure antinodes.

Like closed tubes, rooms exhibit analogous characteristics of resonance, but the situation is now three-dimensional instead of one-dimensional. Three types of resonance or room modes can be described:

- Axial modes—resonant condition involving two parallel surfaces of the room
- Tangential modes—resonant condition involving four surfaces and parallel to two surfaces of room
- Oblique modes—resonant condition involving all six surfaces of room

The axial, tangential, and oblique room modes for a rectangular room can be calculated as follows:

$$F = \frac{C}{2}\left[\left(\frac{p}{L}\right)^2 + \left(\frac{q}{W}\right)^2 + \left(\frac{r}{H}\right)^2\right]^{1/2}$$

where: L, W, H = length, width, and height of room
C = the speed of sound
p, q, r = integers 0, 1, 2, 3, . . .

A set of p, q, and r represents one mode of vibration.

For the axial modes, only one of p, q, or r will be nonzero. For tangential modes, two of p, q, and r will be nonzero. Oblique modes have p, q, and r all nonzero.

The above formula can be simplified for the axial modes:

$$\textit{Axial mode frequencies} = \frac{Cp}{2L}, \frac{Cq}{2W}, \textit{ and } \frac{Cr}{2H}$$

$$\textit{for } p, q, r = 1, 2, 3, . . .$$

At lower frequencies, modal frequencies are spaced farther apart. As frequency is increased, the modes become spaced very close together. At low frequencies, the distribution of modes with respect to frequency determines the coloration of room response. In general, room dimensions that lead to an even and uniform spacing of room modes lead to the most natural sounding environment. Room dimensions that lead to common modal frequencies should be avoided since increased response at these coincident frequencies may cause irregular boominess. Similarly, a situation with large spacings (approximately 20 Hz or larger) between adjacent modes will also result in unnatural response.

Rooms where ratios of length, width, and height are related by small whole numbers should be avoided since this leads to coincident modes and consequent boominess at these frequencies. A room in the shape of a cube would be the worst case since axial modes associated with all three dimensions of the room would overlap. Beyond these simple guidelines, more concrete design criteria become complex. The search for optimum room ratios has been going on since the time of the 1940s. An example of the results of an optimized algorithm for achieving uniform mode spacing is shown in Table 2.

TABLE 2

Recommended room dimension ratios for optimum mode spacing.

	Set 1	Set 2	Set 3
Height	1.00	1.00	1.00
Width	1.14	1.28	1.60
Length	1.39	1.54	2.33

(Sepmeyer, L W.: "Computed Frequency and Angular Distribution of the Normal Modes of Vibration in Rectangular Rooms." JASA V. 37, N. 3 (March 1965).)

Room mode analysis and optimization is particularly important in small rooms such as studios, control rooms, announce booths and so forth. A look at the example in Fig. 4 shows why. Small rooms have few modes at the lowest audio frequencies and the spacing between them may be excessively large. The lowest mode in a room is equal to 565/L where L is the longest dimension of the room. Audio energy below this frequency will not be supported at all by resonance, and room response at these frequencies will be attenuated.

As stated previously, three types of room modes exist (in a rectangular room). While all three are significant, a reasonable first-order design and/or analysis can be made using only the axial modes. The

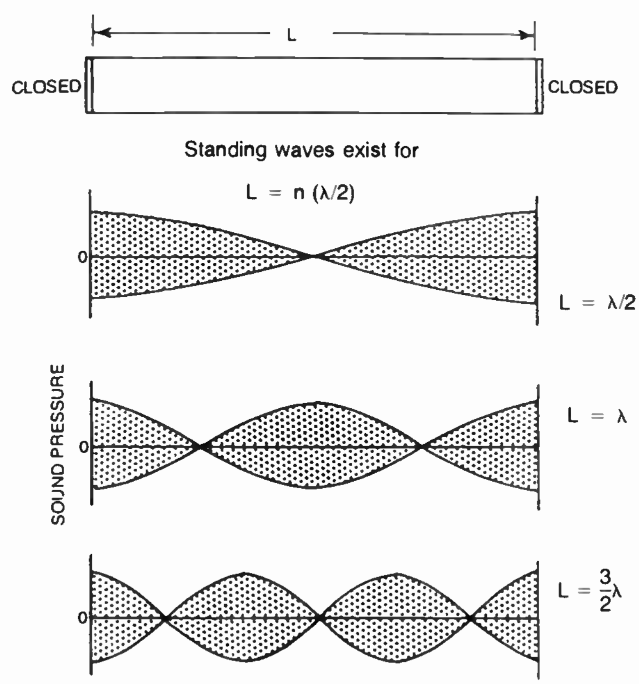


Figure 3. Standing waves in a closed tube.

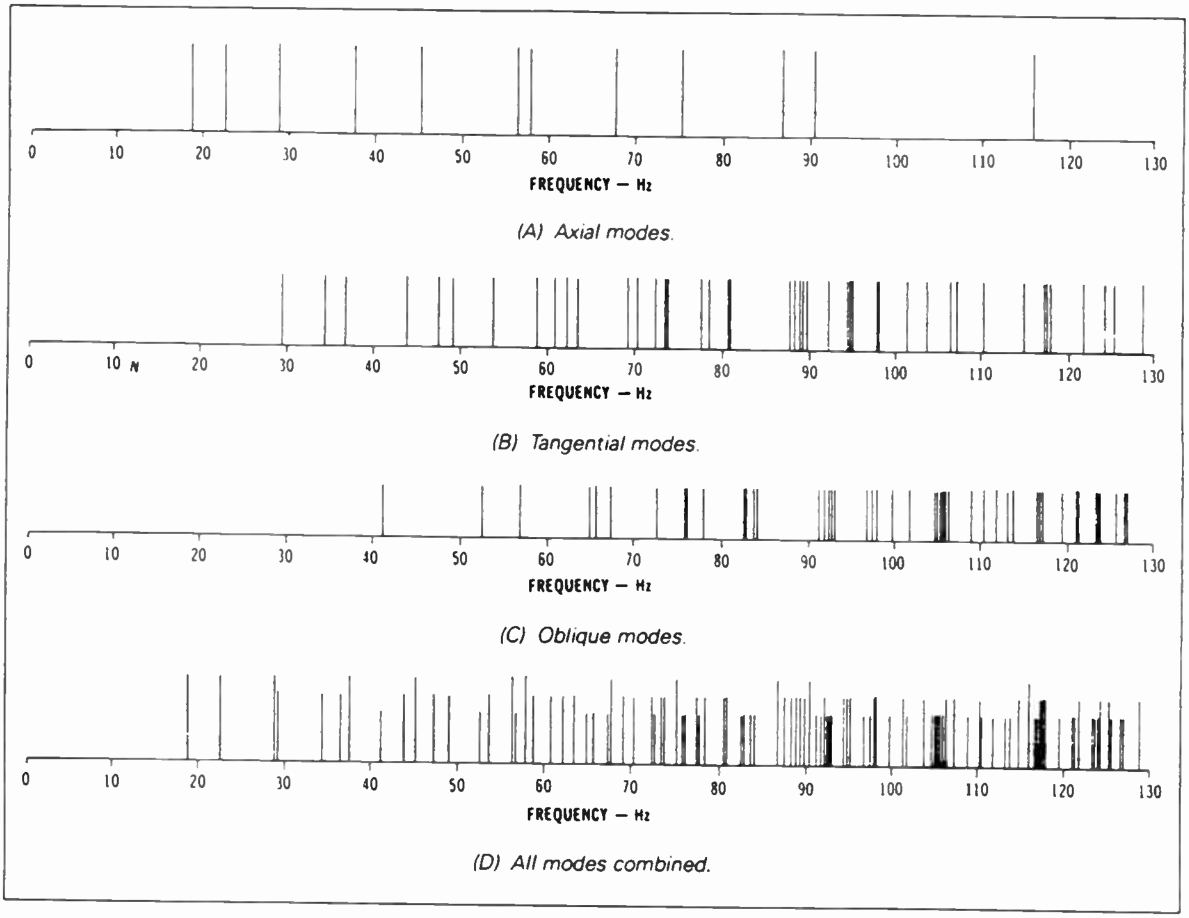


Figure 4. Modal frequencies for a room 30 ft x 25 ft x 19.5 ft. (Glen Ballou: *Handbook for Sound Engineers, the New Audio Cyclopedia*, H. Sams, 1987).

tangential and oblique modes involve four or six room surfaces respectively and are likely to suffer more attenuation due to the absorption of the surfaces and greater likelihood of meeting physical obstructions. In general, an optimized axial mode pattern will lead to an overall optimized modal pattern since the tangential and oblique modes will tend to fill in the gaps between the axial modes.

Because of the resonance problems of small rooms, minimum sizes have been investigated over the years. The European Broadcast Union recommends a preferred volume of 80 cubic meters for control rooms and 100 cubic meters for listening rooms. Less than 40 cubic meters is not recommended for use due to mode spacing problems.

Use of Nonparallel Walls

When the room is not rectangular, room mode problems do not vanish, they just become difficult to calculate. The modal distribution is certainly affected by room shape changes, but room resonances are not eliminated since the basic presence of room modes is more associated with volume as opposed to room shape. Nonparallel walls have been successfully used to promote good diffusion at higher frequencies and prevent flutter echoes. Flutter echoes are produced by repeated reflections from two flat parallel surfaces sufficiently distant from each other to produce a distinct echo. Usually an angle of 5° or a splay of one foot in 12 running feet is enough to destroy flutter echo problems. However, other more cost effective techniques exist to solve these problems.

Absorption of Sound

In general, when sound waves strike a surface:

A. A portion of the energy is transmitted through the surface;
B. A portion of the energy is reflected back into the room;
C. A portion of the energy is absorbed.

The figure of merit for the absorptive quality of a material is called the absorption coefficient, α.

$$\alpha = \frac{\textit{Energy absorbed by the surface}}{\textit{Energy incident on the surface}}$$

α ranges between 0 and 1, 0 being a perfect reflector and 1 being a perfect absorber. In practical terms, absorption coefficients of surfaces are used in room design to aid in controlling reflections. For room design, an open window is considered a perfect absorber since no energy that strikes the opening is reflected back into the room.

The total absorption of a surface is defined as the product of its absorption coefficient and the surface area. The unit of absorption is the *sabin*, named for the early 20th Century acoustician, Wallace Clement Sabine.

In many applications, it is convenient to define the average absorption coefficient of the room, $\bar{\alpha}$. $\bar{\alpha}$ is defined as the total absorption in the room (in sabins) divided by the total surface area:

$$\bar{\alpha} = \frac{\alpha_1 s_1 + \alpha_2 s_2 + \ldots \alpha_n s_n}{s_1 + s_2 + \ldots s_n} = \frac{\sum_{i=1}^{n} \alpha_i s_i}{s_T}$$

where α_n and s_n are the absorption coefficient and surface area of a portion of the room, and s_T is the total surface area of the room.

The average absorption coefficient is accurate only for the frequencies at which the absorption coefficient values are valid. Absorption coefficients vary as a function of frequency and are usually listed in tables at 125, 250, 500, 1000, 2000, and 4000 Hz.

Often materials are specified by their *noise reduction coefficient* (NRC) value. NRC is intended to be a convenient single number index of average sound absorbing efficiency. It is defined as the arithmetic average of absorption coefficients at 250, 500, 1000, and 2000 Hz:

$$NRC = \frac{\alpha_{250} + \alpha_{500} + \alpha_{1000} + \alpha_{2000}}{4}$$

While this has some merit, NRC has the considerable disadvantage that low (and high) frequency absorption is not considered, an important criteria to consider in critical broadcasting applications. Also, the variation of α with frequency is an important consideration as well in critical applications where optimizing the total absorption in a room at all frequencies is important. Since NRC is an unweighted average, two materials with the same NRC could have drastically different absorption versus frequency characteristics. Wherever possible, absorption in a room should be analyzed separately in the various frequency ranges of interest.

Effect of Mounting Method of Material on Absorption

Different methods of mounting a given material may give different absorption coefficient results. Tables of absorption coefficients must indicate the mounting method used when the absorption coefficient was measured to be useful.

Fig. 5 shows some of the various mounting designations used for standard measurements.

Table 3 shows the absorption coefficients of various commercial sound absorbers and general building materials.

Sound Absorbing Material Applications

Echo Control

Sound absorbing material can be used very effectively in specific trouble areas to stop echoes or flutter echoes (delayed sound reflections of sufficient intensity to be heard discretely above the general reverberant sound level).

Noise Reduction

Sound absorbing material can be used to control noise within a room by lessening the amount of reverberant (reflected) energy present.

$$\textit{Noise reduction (in dB)} = 10 \log \left(\frac{a_{after}}{a_{before}}\right)$$

where a is the total absorption in the room in sabins before and after room treatment.

Note that the total absorption must be doubled to lower the noise 3 dB and doubled again to achieve a total 6 dB reduction. A practical limit is quickly reached in attempting to achieve more than 6 dB to 10 dB of noise reduction using this technique.

Reverberation Control

The "liveness" or "deadness" of a room can be controlled by the introduction of sound absorbing material. This is perhaps the most common use of sound absorbing material.

Types of Absorbers

Porous Absorbers — Characterized by a material with deep pores and cavities. Sound energy entering the pores is dissipated by frictional and viscous resistance and/or vibrations of fibers of the material. Glass fiber is a typical porous absorber.

Panel Absorbers — Sound energy forces a panel into vibration. The vibrational activity converts the sound energy into heat.

Cavity (Helmholtz) Resonator — Analogous to blowing air across the mouth of a jug. At the resonant frequency, air in the jug neck vibrates back and forth as a single air mass. Sound energy is dissipated by frictional resistance in and around the neck.

Examples of Porous Absorbers

Acoustical blankets such as glass fibers are common porous absorbers. The absorption of a given material depends on its thickness, density, and relative porosity. In general,

1. Increasing thickness increases absorption (mainly at low frequencies).
2. Blankets of relatively low density have increased absorption as the density increases.
3. Blanket must be composed of interconnected open pores (closed cell foam is a poor absorber). As a guideline, if a blanket will pass smoke under moderate pressure, it will probably be a good absorber.

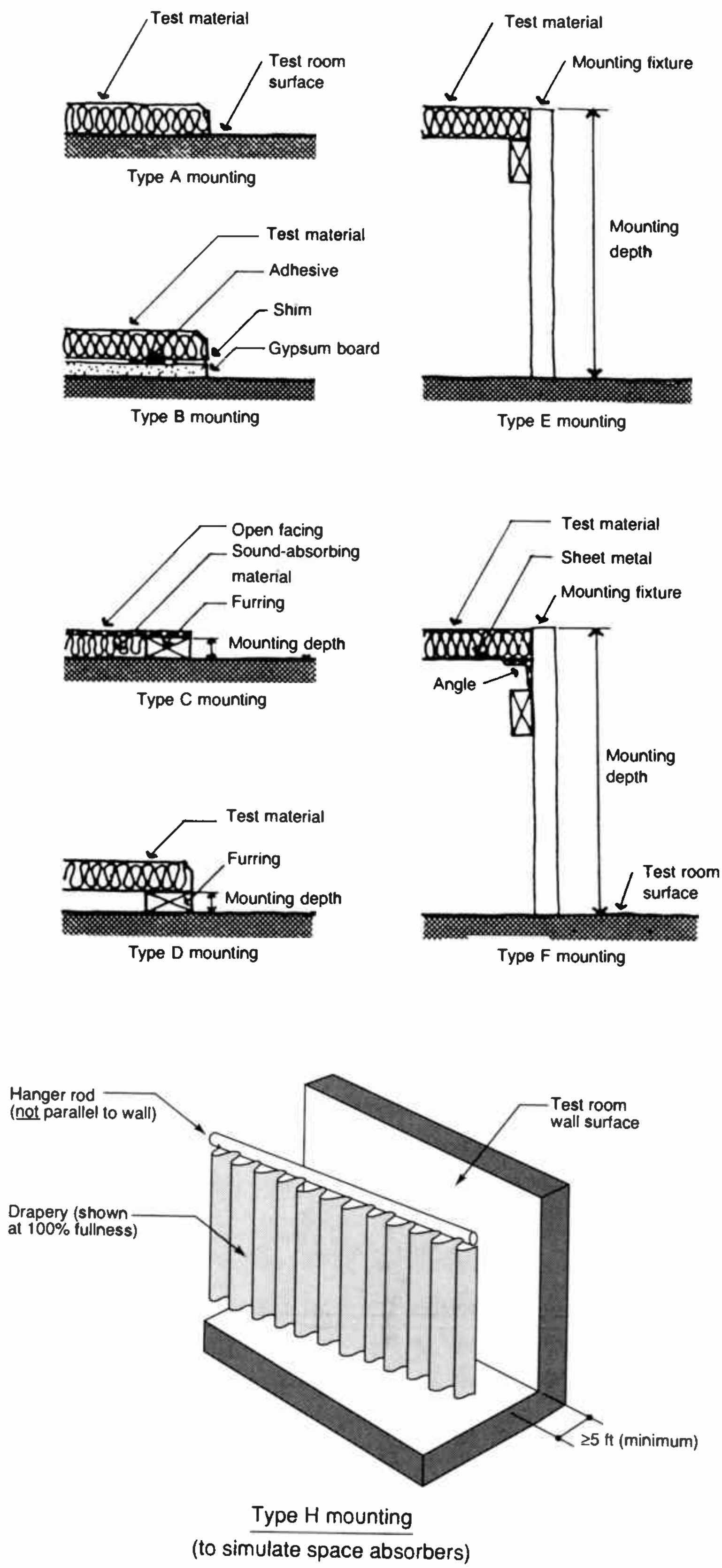


Figure 5. Standard mounting methods for measuring the absorption coefficient of a material. (ASTM C 423)

TABLE 3
Absorption coefficients of various materials. (Reprinted with permission from *Architectural Acoustics*, by M. D. Egan. Copyright 1988 by McGraw-Hill).

Material	Sound Absorption Coefficient 125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	NRC Number*
Walls[1–3, 9, 12]							
Sound-Reflecting:							
1. Brick, unglazed	0.02	0.02	0.03	0.04	0.05	0.07	0.05
2. Brick, unglazed and painted	0.01	0.01	0.02	0.02	0.02	0.03	0.00
3. Concrete, rough	0.01	0.02	0.04	0.06	0.08	0.10	0.05
4. Concrete block, painted	0.10	0.05	0.06	0.07	0.09	0.08	0.05
5. Glass, heavy (large panes)	0.18	0.06	0.04	0.03	0.02	0.02	0.05
6. Glass, ordinary window	0.35	0.25	0.18	0.12	0.07	0.04	0.15
7. Gypsum board, ½-in-thick (nailed to 2 × 4s, 16 in oc)	0.29	0.10	0.05	0.04	0.07	0.09	0.05
8. Gypsum board, 1 layer, ⅝-in-thick (screwed to 1 × 3s, 16 in oc with airspaces filled with fibrous insulation)	0.55	0.14	0.08	0.04	0.12	0.11	0.10
9. Construction no. 8 with 2 layers of ⅝-in-thick gypsum board	0.28	0.12	0.10	0.07	0.13	0.09	0.10
10. Marble or glazed tile	0.01	0.01	0.01	0.01	0.02	0.02	0.00
11. Plaster on brick	0.01	0.02	0.02	0.03	0.04	0.05	0.05
12. Plaster on concrete block (or 1 in thick on lath)	0.12	0.09	0.07	0.05	0.05	0.04	0.05
13. Plaster on lath	0.14	0.10	0.06	0.05	0.04	0.03	0.05
14. Plywood, ⅜-in paneling	0.28	0.22	0.17	0.09	0.10	0.11	0.15
15. Steel	0.05	0.10	0.10	0.10	0.07	0.02	0.10
16. Venetian blinds, metal	0.06	0.05	0.07	0.15	0.13	0.17	0.10
17. Wood, ¼-in paneling, with airspace behind	0.42	0.21	0.10	0.08	0.06	0.06	0.10
18. Wood, 1-in paneling with airspace behind	0.19	0.14	0.09	0.06	0.06	0.05	0.10
Sound-Absorbing:							
19. Concrete block, coarse	0.36	0.44	0.31	0.29	0.39	0.25	0.35
20. Lightweight drapery, 10 oz/yd², flat on wall (*Note:* Sound-reflecting at most frequencies.)	0.03	0.04	0.11	0.17	0.24	0.35	0.15
21. Mediumweight drapery, 14 oz/yd², draped to half area (i.e., 2 ft of drapery to 1 ft of wall)	0.07	0.31	0.49	0.75	0.70	0.60	0.55
22. Heavyweight drapery, 18 oz/yd², draped to half area	0.14	0.35	0.55	0.72	0.70	0.65	0.60
23. Fiberglass fabric curtain, 8½ oz/yd², draped to half area (*Note:* The deeper the airspace behind the drapery (up to 12 in), the greater the low-frequency absorption.)	0.09	0.32	0.68	0.83	0.39	0.76	0.55
24. Shredded-wood fiberboard, 2 in thick on concrete (mtg. A)	0.15	0.26	0.62	0.94	0.64	0.92	0.60
25. Thick, fibrous material behind open facing	0.60	0.75	0.82	0.80	0.60	0.38	0.75
26. Carpet, heavy, on ⅝-in perforated mineral fiberboard with airspace behind	0.37	0.41	0.63	0.85	0.96	0.92	0.70
27. Wood, ½-in paneling, perforated 3/16-in-diameter holes, 11% open area, with 2½-in glass fiber in airspace behind	0.40	0.90	0.80	0.50	0.40	0.30	0.65
Floors[9, 11]							
Sound-Reflecting:							
28. Concrete or terrazzo	0.01	0.01	0.02	0.02	0.02	0.02	0.00
29. Linoleum, rubber, or asphalt tile on concrete	0.02	0.03	0.03	0.03	0.03	0.02	0.05
30. Marble or glazed tile	0.01	0.01	0.01	0.01	0.02	0.02	0.00
31. Wood	0.15	0.11	0.10	0.07	0.06	0.07	0.10
32. Wood parquet on concrete	0.04	0.04	0.07	0.06	0.06	0.07	0.05
Sound Absorbing:							
33. Carpet, heavy, on concrete	0.02	0.06	0.14	0.37	0.60	0.65	0.30
34. Carpet, heavy, on foam rubber	0.08	0.24	0.57	0.69	0.71	0.73	0.55
35. Carpet, heavy, with impermeable latex backing on foam rubber	0.08	0.27	0.39	0.34	0.48	0.63	0.35
36. Indoor-outdoor carpet	0.01	0.05	0.10	0.20	0.45	0.65	0.20
Ceilings[6, 8–10]†							
Sound-Reflecting:							
37. Concrete	0.01	0.01	0.02	0.02	0.02	0.02	0.00
38. Gypsum board, ½-in-thick	0.29	0.10	0.05	0.04	0.07	0.09	0.05
39. Gypsum board, ½-in-thick, in suspension system	0.15	0.10	0.05	0.04	0.07	0.09	0.05
40. Plaster on lath	0.14	0.10	0.06	0.05	0.04	0.03	0.05
41. Plywood, ⅜-in thick	0.28	0.22	0.17	0.09	0.10	0.11	0.15

(continued)

TABLE 3 (continued)

Material	Sound Absorption Coefficient						NRC Number*
	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	
Ceilings[6, 8–10]† (continued)							
Sound-Absorbing:							
42. Acoustical board, ¾-in thick, in suspension system (mtg. E)	0.76	0.93	0.83	0.99	0.99	0.94	0.95
43. Shredded-wood fiberboard, 2 in thick on lay-in grid (mtg. E)	0.59	0.51	0.53	0.73	0.88	0.74	0.65
44. Thin, porous sound-absorbing material, ¾-in-thick (mtg. B)	0.10	0.60	0.80	0.82	0.78	0.60	0.75
45. Thick, porous sound-absorbing material, 2 in thick (mtg. B), or thin material with airspace behind (mtg. D)	0.38	0.60	0.78	0.80	0.78	0.70	0.75
46. Sprayed cellulose fibers, 1 in thick on concrete (mtg. A)	0.08	0.29	0.75	0.98	0.93	0.76	0.75
47. Glass-fiber roof fabric, 12 oz/yd^2	0.65	0.71	0.82	0.86	0.76	0.62	0.80
48. Glass-fiber roof fabric, 37½ oz/yd^2 (*Note:* Sound-reflecting at most frequencies.)	0.38	0.23	0.17	0.15	0.09	0.06	0.15
49. Polyurethane foam, 1 in thick, open cell, reticulated	0.07	0.11	0.20	0.32	0.60	0.85	0.30
50. Parallel glass-fiberboard panels, 1 in thick by 18 in deep, spaced 18 in apart, suspended 12 in below ceiling	0.07	0.20	0.40	0.52	0.60	0.67	0.45
51. Parallel glass-fiberboard panels, 1 in thick by 18 in deep, spaced 6½-in apart, suspended 12 in below ceiling	0.10	0.29	0.62	1.12	1.33	1.38	0.85
Seats and Audience[1, 5, 7, 9]‡							
52. Fabric well-upholstered seats, with perforated seat pans, unoccupied	0.19	0.37	0.56	0.67	0.61	0.59	
53. Leather-covered upholstered seats, unoccupied§	0.44	0.54	0.60	0.62	0.58	0.50	
54. Audience, seated in upholstered seats§	0.39	0.57	0.80	0.94	0.92	0.87	
55. Congregation, seated in wooden pews	0.57	0.61	0.75	0.86	0.91	0.86	
56. Chair, metal or wood seat, unoccupied	0.15	0.19	0.22	0.39	0.38	0.30	
57. Students, informally dressed, seated in tablet-arm chairs	0.30	0.41	0.49	0.84	0.87	0.84	
58. Person, adult (total number of sabins)[13]	2.5	3.5	4.2	4.6	5.0	5.0	
Miscellaneous[3, 9, 11]							
59. Gravel, loose and moist, 4 in thick	0.25	0.60	0.65	0.70	0.75	0.80	0.70
60. Grass, marion bluegrass, 2 in high	0.11	0.26	0.60	0.69	0.92	0.99	0.60
61. Snow, freshly fallen, 4 in high	0.45	0.75	0.90	0.95	0.95	0.95	0.90
62. Soil, rough	0.15	0.25	0.40	0.55	0.60	0.60	0.45
63. Trees, balsam firs, 20 ft^2 ground area per tree, 8 ft high	0.03	0.06	0.11	0.17	0.27	0.31	0.15
64. Water surface (swimming pool)	0.01	0.01	0.01	0.02	0.02	0.03	0.00

* NRC (noise reduction coefficient) is a single-number rating of the sound absorption coefficients of a material. It is an average that only includes the coefficients in the 250 to 2000 Hz frequency range and therefore should be used with caution.

†Refer to manufacturer's catalogs for absorption data which should be from up-to-date tests by independent acoustical laboratories according to current ASTM procedures.

‡Coefficients are per square foot of seating floor area or per unit. Where the audience is randomly spaced (e.g., courtroom, cafeteria), mid-frequency absorption can be estimated at about 5 sabins per person. To be precise, coefficients per person must be stated in relation to spacing pattern.

§The floor area occupid by the audience must be calculated to include an *edge effect* at aisles. For an aisle bounded on both sides by audience, include a strip 3 ft wide; for an aisle bounded on only one side by audience, including a strip 1½ ft wide. No edge effect is used when the seating abuts walls or balcony fronts (because the edge is shielded). The coefficients are also valid for orchestra and choral areas at 5 to 8 ft^2 per person. Orchestra areas include people, instruments, music racks, etc. No edge effects are used around musicians.

Test Reference

"Standard Test Method for Sound Absorption and Sound Absorption Coefficients by the Reverberation Room Method." ASTM C 423. Available from American Society for Testing and Materials (ASTM), 1916 Race Street, Philadelphia, PA 19103.

Sources

1. L. L. Beranek, "Audience and Chair Absorption in Large Halls," *Journal of the Acoustical Society of America*, January 1969.
2. A. N. Burd et al., "Data for the Acoustic Design of Studios," British Broadcasting Corporation, BBC Engineering Monograph no. 64, November 1966.
3. E. J. Evans and E. N. Bazley, "Sound Absorbing Materials," H. M. Stationery Office, London, 1964.
4. R. A. Hedeen, *Compendium of Materials for Noise Control*, National Institute for Occupational Safety and Health (NIOSH), Publication no. 80-116, Cincinnati, Ohio, May 1980. (Contains sound absorption data on hundreds of commercially available materials.)
5. H. F. Kingsbury and W. J. Wallace, "Acoustic Absorption Characteristics of People," *Sound and Vibration*, December 1968.
6. T. Mariner, "Control of Noise by Sound-Absorbent Materials," *Noise Control*, July 1957.
7. J. E. Moore and R. West, "In Search of an Instant Audience," *Journal of the Acoustical Society of America*, December 1970.
8. R. Moulder and J. Merrill, "Acoustical Properties of Glass Fiber Roof Fabrics," *Sound and Vibration*, October 1983.
9. "Performance Data, Architectural Acoustical Materials," Acoustical and Insulating Materials Association (AIMA). (This bulletin was published annually from 1941 to 1974.)
10. W. E. Purcell, "Materials for Noise and Vibration Control," *Sound and Vibration*, July 1982.
11. W. Siekman, "Outdoor Acoustical Treatment: Grass and Trees," *Journal of the Acoustical Society of America*, October 1969.
12. "Sound Conditioning with Carpet," The Carpet and Rug Institute, Dalton, Ga., 1970.
13. Knudsen, V. & Harris, C., "Acoustical Designing in Architecture," American Institute of Physics, 1978.

Glass fiber is certainly one of the workhorse materials for sound absorption applications. It is interesting to note the effect of thickness and mounting arrangement on the absorptive qualities of this material. Fig. 6 shows that increased thickness mainly affects the low-frequency performance. (High-frequency absorption is more a function of the surface texture, and is relatively independent of thickness). Fig. 7 shows the effect of increased low frequency absorption as a function of an airspace between the glass fiber and the wall. Fig. 8 shows that the effect of packing density on the absorption performance is rather small.

Drapes and curtains are also often used to attenuate reflected sound energy. However, drapes also tend to have high absorption at high frequencies and low absorption at low frequencies. Listed below are some guidelines to increase the low frequency absorption.

- Use heavy base material with a lining
- Use 100% to 200% gathering of the drape
- Hang at least 6" from the wall

Examples of the absorption of drapes are listed in Table 3.

An example of a commercially available cellular foam type of absorber is shown in Fig. 9.

Carpeting is often used as an absorber. However, absorption of carpeting at low frequencies is relatively poor. In general, a foam rubber or hair felt underlayment can improve low frequency absorption significantly. Typical absorption figures (with concrete as a reference) are listed in Table 3.

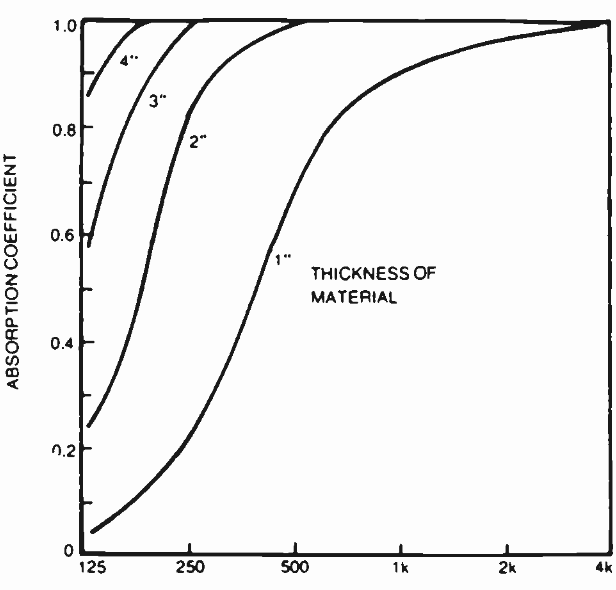


Figure 6. Effect of thickness on absorption characteristics of glass fiber (3 lb/cu. ft. density) mounted directly on a hard surface. (Reprinted, with permission, from book #3096 *The Master Handbook of Acoustics—2nd Edition*, by F. Alton Everest. Copyright 1981, 1989 by TAB Books, A Division of McGraw-Hill Inc., Blue Ridge Summit, PA 17294-0850. (1-800-233-1128).)

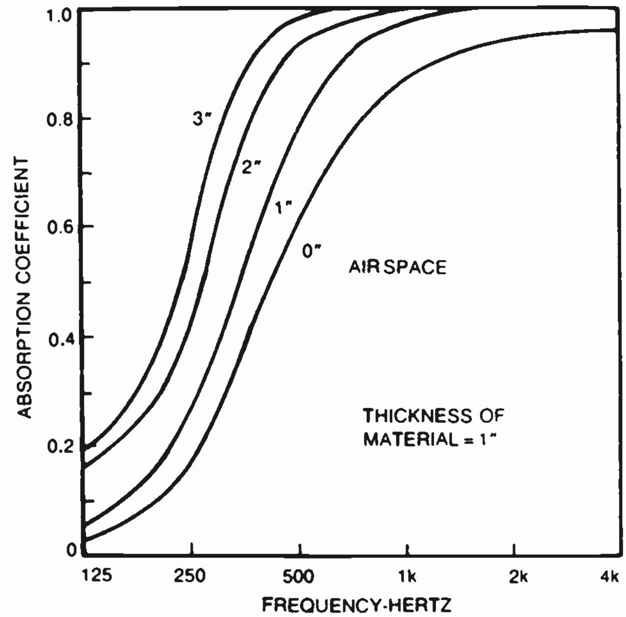


Figure 7. Effect of airspace on absorption characteristics of glass fiber (3 lb./cu. ft. density). (Reprinted, with permission, from book #3096 *The Master Handbook of Acoustics—2nd Edition*, by F. Alton Everest. Copyright 1981, 1989 by TAB Books, A Division of McGraw-Hill Inc., Blue Ridge Summit, PA 17294-0850. (1-800-233-1128).)

Acoustic Absorption Characteristics of People

Absorption by people and objects is specified in absorption tables as either the number of sabins per person or as an absorption coefficient provided by an audience based on normal seat and aisle spacings. Typical values are listed in Table 3.

Panel Absorbers

A panel with an enclosed air space behind it forms a resonant system with the air mass behaving as a spring and the panel as a mass. If the panel is thin, the resonant frequency and hence the frequency of maximum absorption coefficient can be shown to be:

$$f_{resonant} = \frac{170}{\sqrt{md}}$$

where: d = Air space depth (in)
m = Mass per unit area of panel (lb/sq. ft)

A broader absorption characteristic can be achieved by filling the air space with absorptive material, such as glass fiber.

Cavity Resonators

There are three basic types of cavity resonators:

1. Individual units
2. Perforated panels
3. Slit resonators

An example of an individual prefabricated cavity resonator is shown in Fig. 10, which is essentially a

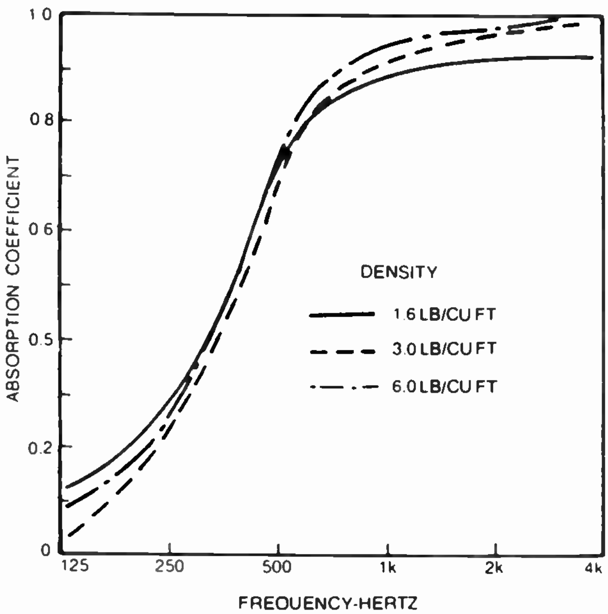


Figure 8. Effect of packing density on absorption characteristics of glass fiber (1" thickness). (Reprinted, with permission, from book #3096 *The Master Handbook of Acoustics—2nd Edition*, by F. Alton Everest. Copyright 1981, 1989 by TAB Books, A Division of McGraw-Hill Inc., Blue Ridge Summit, PA 17294-0850. (1-800-233-1128).)

slotted concrete block. Another type of commercially available individual unit is shown in Fig. 11.

A perforated panel of significant thickness (greater than approximately ⅛") spaced away from a rigid backing exhibits the absorptive behavior of a cavity resonator. The resonator has an approximate resonance frequency of:

$$f_{resonant} = 200\sqrt{\frac{p}{dt}}$$

where: p = Percentage of open area of panel
d = Air space depth (in.)
t = Panel thickness + 0.8 × hole diameter (in.)

A slit resonator consists of a number of slats spaced away from a rigid backing with air spaces (slots) in between the slats as shown in Fig. 12.

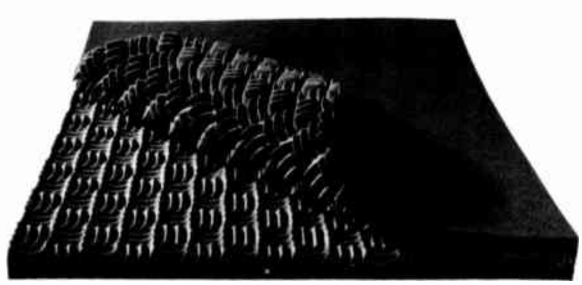


Figure 9. An open cell foam acoustical material. (SONEX photo courtesy of Illbruck Inc.; SONEX Acoustical Products Div.)

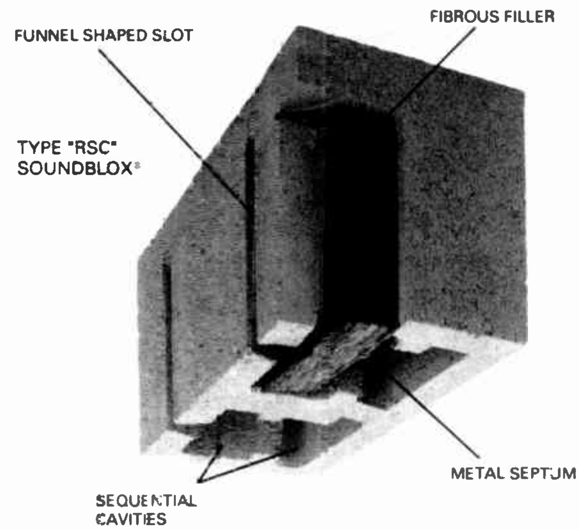


Figure 10. A cavity resonator used in construction. (Type RSC SOUNDBLOX® courtesy of the Proudfoot Co. Inc.)

The resonant frequency can be calculated as follows:

$$f_{resonant} = 2,160\sqrt{\frac{s}{dD(w+s)}}$$

where: s = Width of slot (in.)
d = Thickness of slat (in.)
D = Depth of air space (in.)
w = Width of slat (in.)

In the example in Fig. 12, a resonant frequency of about 250 Hz results. Broader absorption characteristics can be achieved by using slats and slots of varying widths or nonparallel air spaces.

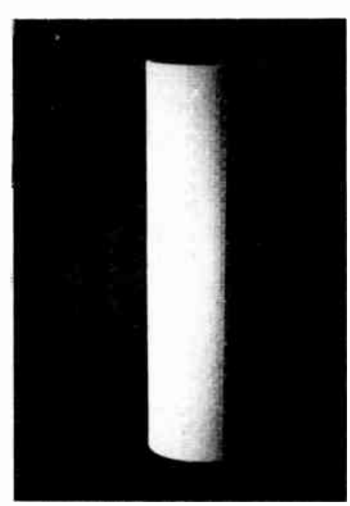


Figure 11. A portable cavity resonator. (TubeTrap™ Photo courtesy of Acoustic Sciences Corp.)

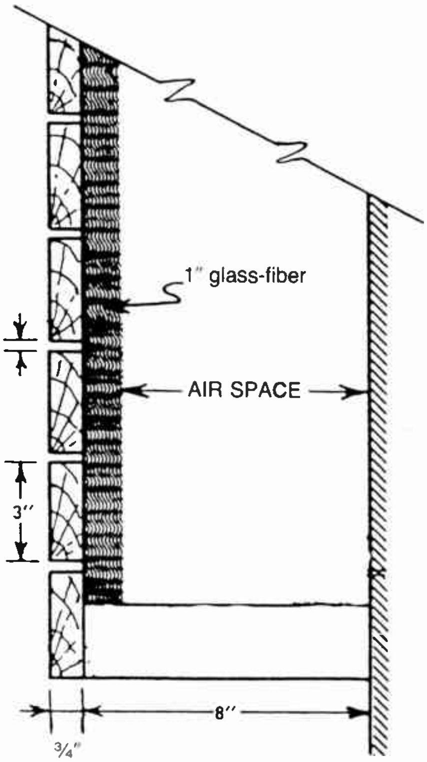


Figure 12. Example of a slit resonator.

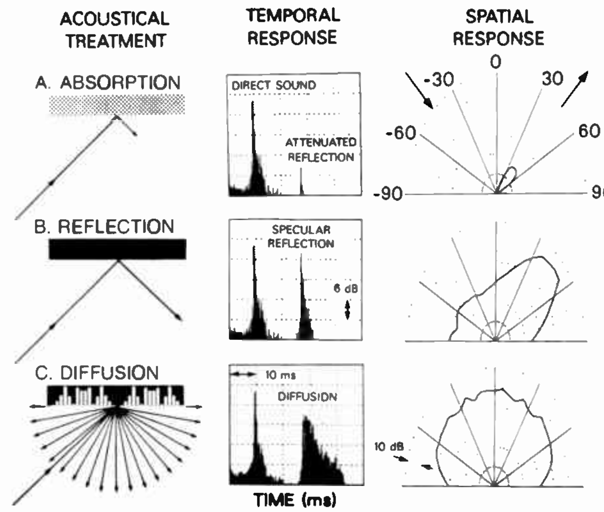


Figure 13. Reflection, absorption, and diffusion from acoustic surface treatments. (Courtesy of RPG Diffusor Systems)

Diffusion

Diffusion of sound in a room refers to the extent to which sound energy is uniformly distributed throughout a room. Making an analogy with baking a cake, diffusion would represent an indicated measure of how well the batter was mixed and individual ingredients of the recipe were dispersed throughout the mixture. In pursuing good room acoustics, maximum diffusion is desirable, all other things being equal. In a purely diffuse sound field:

1. at a given location, sound waves are equally likely to be traveling in any direction, and
2. the sound pressure will be equal at all locations throughout the room.

These criteria are determined by the pattern of reflections within the room. Diffusion in a room can be maximized in several ways. The introduction of oddly shaped protrusions aids in increasing diffusion. The success of many famous concert halls built in the 19th century can be largely attributed to the florid architectural features that offer diverse reflective and scattering properties for incident sound. Also, the intentionally irregular distribution of absorptive material in patches will increase diffusion. Increasing the randomness of the location of patches of absorptive material will also increase diffusion.

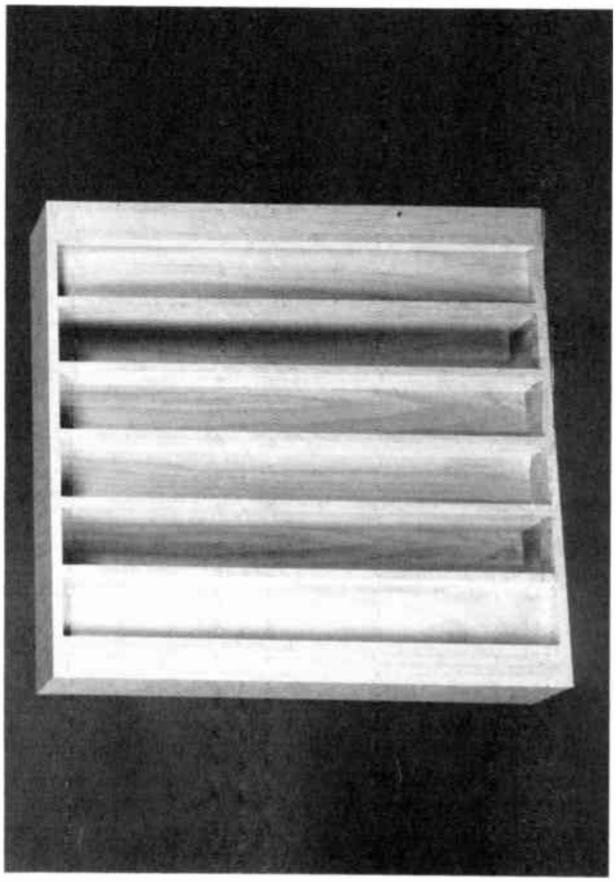
The range of possible effects of surface treatment on incident sound energy are shown in Fig. 13. As shown in (B), a hard flat surface will reflect sound in a specular manner just as light is reflected from a mirror, following the relationship that the angle of incidence (relative to a line perpendicular to the surface) equals the angle of reflection and is equal in magnitude. At a location where both direct and reflected waves arrive, the reflected wave is a delayed replica of the incident wave. Similarly, sound absorptive material applied to a surface yields an attenuated and delayed replica of the incident wave as shown in (A). Finally, a diffusive surface, as shown in (C), reflects the incident energy equally over a wide angular range and also corresponds to a widening of the energy received as a function of time.

Diffusing elements designed specifically for that purpose are now available and the design of these units is quite refined. While traditional surface relief ornamentation, as referenced above, is useful as a diffusion element, it generally does not provide broad bandwidth wide angle diffusion. An ideal diffusor would provide sound diffusion which is not a function of frequency, angle of incidence, or observation angle. Examples of several commercially available products designed with these goals in mind are shown in Fig. 14.

Based on the construction of a series of wells of different depths derived mathematically, the diffusors shown achieve a uniform angular distribution of reflected energy from a wide range of incident angles for mid and high frequencies.

Sound Level as a Function of Distance in an Enclosed Room

In an enclosed room, the situation is more complicated than the case of a free field; i.e., an environment with no reflected sound energy. The total sound pressure level (SPL) at any point in the room will be



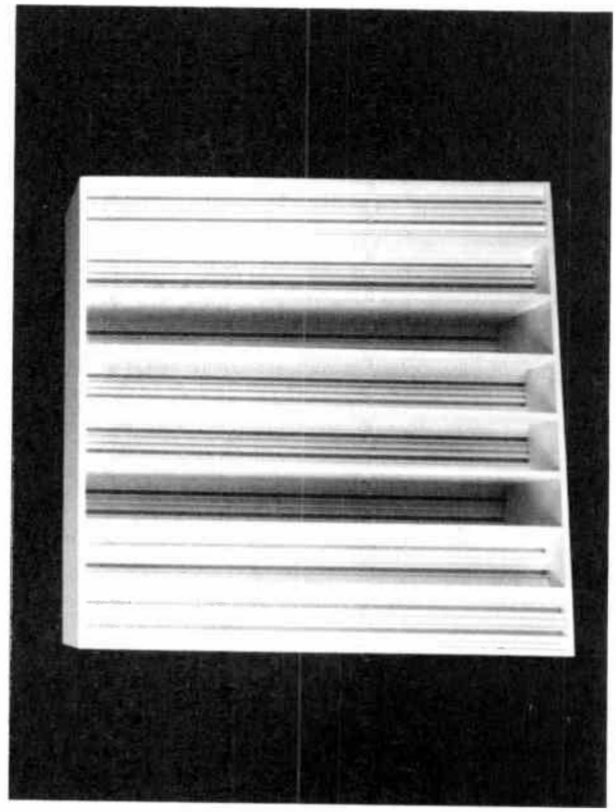


Figure 14. Examples of commercially available diffusors. (Photos courtesy of RPG Diffusor Systems)

the result of contributions directly from the source (referred to as the direct sound field) and sound energy associated with multiple reflections from the room surfaces (referred to as the diffuse or reverberant sound field). The direct sound field varies inversely with the square of the distance from the source; the diffuse sound field is, by definition of being diffuse, equal at all points in the room. It can be theoretically shown that, for an enclosed room, at a distance r (in feet) from a sound source:

$$SPL = PWL + 10 \log\left[\frac{Q}{4\pi r^2} + \frac{4}{R}\right] + 10.5$$

where: $R = \dfrac{S\overline{\alpha}}{1 - \overline{\alpha}}$

Q = Directivity of source
$\overline{\alpha}$ = Average absorption coefficient
S = Total room surface area

(R is sometimes called the Room Constant)

The graphical form of this equation is shown in Fig. 15.

Sound Decay in a Room

When a sound source stops in a room, the SPL at a given location will not decrease to zero instantaneously as in the free field case. Rather, the sound energy in the room will decay over a period of time due to reflected sound energy gradually dissipating as a result of the absorptive qualities of the room surfaces. This sound decay is called reverberation. The reverberation time of a room (RT_{60}) is defined to be the amount of time required for sound to decay 60 dB:

$$RT_{60} = \frac{0.049V}{S[-2.3 \log(1 - \overline{\alpha})]}$$

where: V = Room volume (ft^3)
S = Surface area of room (ft^2)
$\overline{\alpha}$ = Average absorption coefficient

This form of the RT_{60} calculation is known as the Norris-Eyring equation.

When the average absorption coefficient is small, the formula can be further simplified:

$$RT_{60} = \frac{0.049V}{S\overline{\alpha}} \ (when\ \overline{\alpha}<0.1)$$

This is the classic Sabine formula for RT_{60}.

One other form of RT_{60} concerns conditions where different surfaces have significantly different absorption characteristics. This is known as the Fitzroy equation:

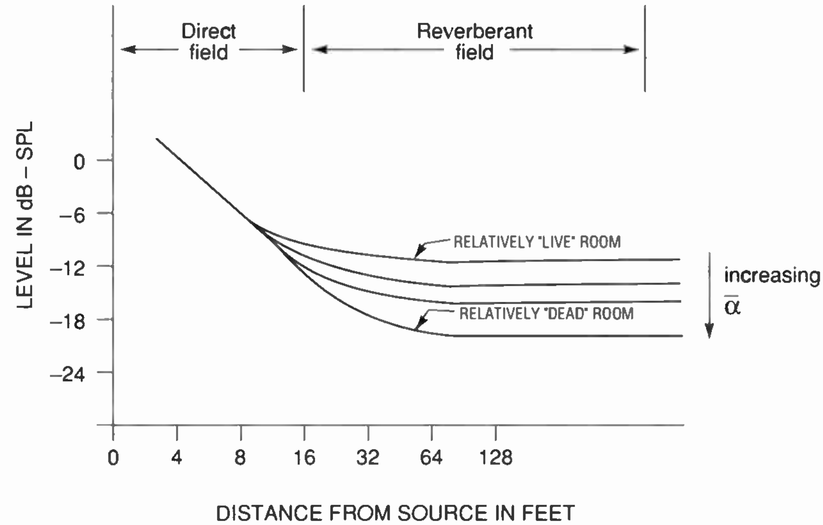


Figure 15. Sound attenuation as a function of distance and the amount of total room absorption.

$$RT_{60} = \frac{0.049V}{S^2}\left[\frac{2xy}{-2.3\log(1-\bar{\alpha}_{xy})} + \frac{2xz}{-2.3\log(1-\bar{\alpha}_{xz})} + \frac{2yz}{-2.3\log(1-\bar{\alpha}_{yz})}\right]$$

where surfaces xy, xz, and yz each represent parallel surfaces and $\bar{\alpha}_{xy}$, $\bar{\alpha}_{xz}$, and $\bar{\alpha}_{yz}$ are their associated average absorption coefficients.

In very large rooms or at high frequencies, the effect of excess absorption due to humidity and other effects must be taken into account:

$$RT_{60} = \frac{0.049V}{S[-2.3\log(1-\bar{\alpha})] + 4mV}$$

where: m = excess sound attenuation in dB/ft (see Fig. 2).

Limitations of RT_{60}

A basic premise of the RT_{60} formula is that the room exhibits a uniform rate of decay of sound. This in turn would require that the sound field is completely diffuse, an assumption generally more close to being true in large "live" rooms than small "dead" rooms. In small, very absorptive rooms, where all significant sound energy dies away in a few reflections, the validity of a statistically based tool like reverberation time becomes questionable. In these cases, the reflection profile itself must be considered on a more specific basis. In critical cases, in addition to addressing the statistical decay of the room versus frequency as represented by reverberation time, it is important to consider the strategic placement of absorptive, reflective, and/or diffusive materials to provide control of first or second-order reflections. For example, strategic reflection control may be important in a studio control room on the side walls and ceiling between the loudspeakers and listening position, as well as the front wall between the loudspeakers and the rear wall, which is often intentionally made diffusive according to several popular studio design philosophies.

While the absorptive qualities of common materials have been well documented, directional scattering properties have not been quantified adequately to permit accurate design use. Work is currently underway to develop a "directional scattering coefficient", which would describe directional reflective aspects of a material, and may someday be another useful tool in the acoustic palette.

Optimum reverberation time has received a lot of attention over the years. Being a subjective figure of merit, numbers that can be universally agreed upon for all circumstances will probably never exist. Certainly, the optimum RT_{60} varies with the size of the room and the intended application. For control rooms and listening rooms, the European Broadcasting Union recommends an RT_{60} of 0.3 seconds ± 0.1 seconds at mid frequencies (200 Hz to 2500 Hz). Fig. 16 shows another example of optimum RT_{60} curves for small listening rooms and studios.

RT_{60} varies as a function of frequency. When not specified, mid frequencies around 500 Hz are usually the assumed frequency range. For critical listening applications, the optimum reverberation characteristic is sometimes deemed to be flat as a function of frequency. However, for a natural sounding environment with 'warmth' it is often desirable to have higher RT_{60} at low frequencies compared to mid and high frequencies. Fig. 17 shows the recommendations of the EBU in this regard.

NOISE CONTROL TECHNIQUES

Common sense is often the most important ingredient in acoustical room design and also the first forgotten concern when practical matters arise and introduce

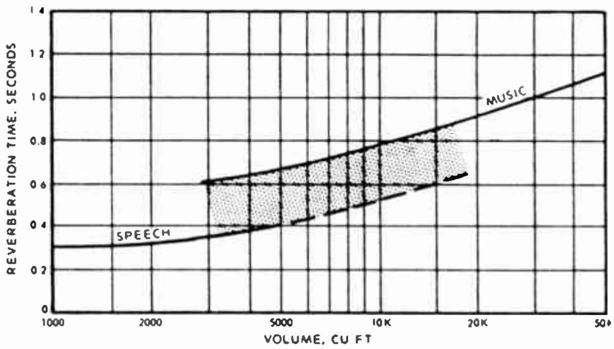


Figure 16. Optimum reverberation time (at mid frequencies). (Reprinted, with permission, from book #1696 *Acoustic Techniques for Home and Studio—2nd Ed.*, by F. Alton Everest. Copyright 1984 by TAB Books, A Division of McGraw-Hil, Inc., Blue Ridge Summit, PA 17294-0850. (1-800-233-1128).)

conflicts. In optimizing a room design for broadcast applications the following checklist identifies the highest priority items that should be considered, listed in approximate order of importance.

1. Select site for least noisy area.
2. Perform noise survey to determine required acoustical isolation.
3. Select position of room within building.
4. Select sound isolation technique.
5. Control noise within building including structure borne and airborne noise.
6. Design room geometry for good diffusion and room mode spacing.
7. Select and place sound absorbing material for optimum room acoustic response.

Numbers 1 through 5 refer to the control of noise and are intentionally a higher priority than criteria such as room shape and optimum reverberation time, especially for broadcast applications. A quiet room is both the most important design goal and also potentially the most expensive and difficult goal to achieve.

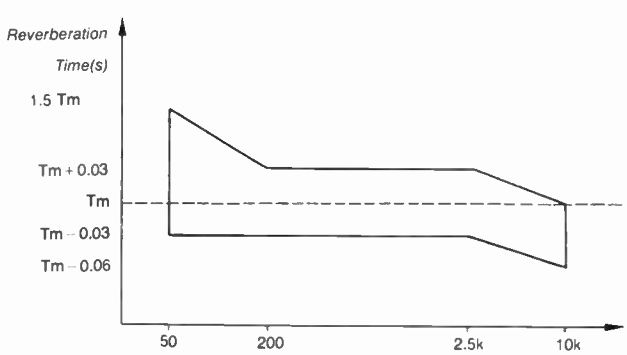


Figure 17. EBU recommendation on optimum reverberation time as a function of frequency. (Reprinted from EBU Technical Recommendation R22-1985: Acoustical Properties of control rooms and listening rooms for the assessment of broadcast programs).

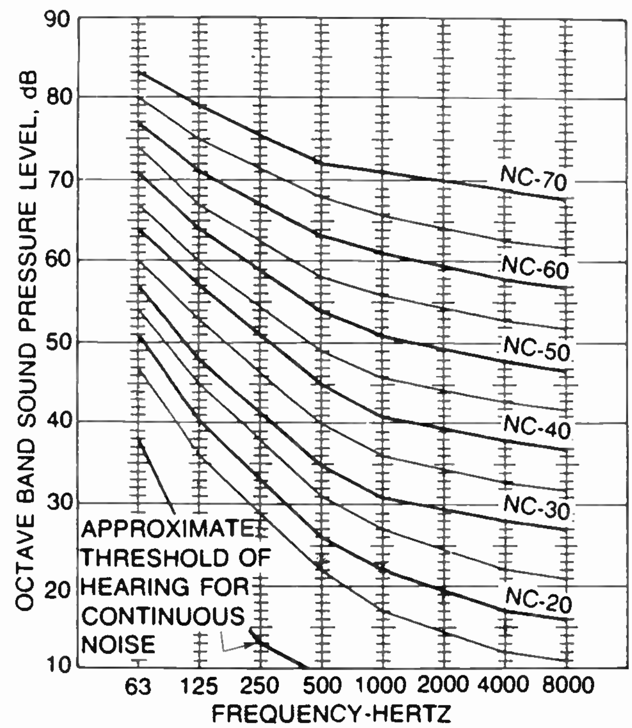


Figure 18. Noise Criteria (NC) curves. (L. Beranek, "Revised Criteria for Noise in Buildings." Noise Control, V. 3, N. 1, Jan. 1957, pp 19–27).

Below are four basic approaches to reducing noise within a room, listed in rough order of preference:

1. Select room location.
2. Reduce output of noise source.
3. Design barrier partitions between noise source and room.
4. Reduce the noise energy existing within the room.

Assessing Noise Levels

The noise criteria (NC) curves were developed through experimentation and knowledge of the sensitivities of the human hearing system to provide a single number figure of merit for maximum permissible noise level for a given activity. The NC number is given approximately by the value of the NC curve in the 1200–2400 Hz frequency band. Fig. 18 shows the NC family of curves.

For assigning an NC rating to an arbitrary room:

1. SPL readings are taken in the 8 octave bands from 63 to 8000 Hz.
2. The NC rating is defined as the lowest value NC curve which lies wholly above the measured data.

For broadcast studios, a rating of NC-20 or less is desirable.

The NC curves were developed in the late 1950s. Since that time, other noise assessment curves have been proposed (the "preferred noise criterion" (PNC) curves in 1971, the RC curves in 1981 and the "balanced noise criterion" (NCB) curves in 1989). While these newer curves may gain in importance in the future, the NC curves are by far the most widespread curves used today for specifying and measuring ambient noise levels.

Airborne Noise Reduction

For sound energy striking a partition, a transmission coefficient can be defined for the partition:

$$\tau = \frac{\textit{Energy transmitted through the partition}}{\textit{Energy incident on the partition}}$$

Transmission Loss (TL) is then defined as:

$$TL = 10\log\left(\frac{1}{\tau}\right)$$

and represents the decibel reduction of sound energy through the partition. The overall TL through a partition composed of several areas (for example, a wall with a door and window) can be calculated as follows:

$$Composite\ TL = 10\log\left[\frac{S_T}{\tau_1 s_1 + \tau_2 s_2 + \ldots \tau_n s_n}\right]$$

where: S_T = total surface area of partition

τ_n and s_n = the transmission coefficient and surface area of the nth element of the composite partition.

The disastrous effect of cracks and air gaps on achieving high TL partitions can easily be illustrated with this formula. For example, using $\tau = 1$ for a crack, consider the effect of a ⅛" crack under a door having a TL of 30 dB. The composite TL is then approximately 26 dB. Applying this same situation to a door with TL = 50 dB, yields a composite TL of approximately 28 dB, a 22 dB loss in isolation due to the air gap!

In partition design, the need for avoidance of cracks and gaps and the importance of gasketing and sealing cannot be overstated—high sound isolation simply cannot be achieved if these factors are ignored or compromised.

Noise Reduction Between Rooms

Noise reduction (NR) is the actual difference in SPL measured in a room containing an offending noise source and the room under test. NR is determined by the area of the dividing partition (S), the total absorption in the receiving room (a), and the TL of the partition:

$$NR = TL + 10\log\left(\frac{a}{S}\right)$$

Thus it can be seen that the actual isolation between rooms can vary both above and below the TL of the partition, although in practice, NR will typically be within 6 dB of the TL.

STC Rating

The Sound Transmission Class (STC) rating represents a single number figure of merit for overall acoustic isolation of a material. TL is plotted on ⅓ octave bands and compared with a standard contour curve to determine the STC rating. The higher the STC, the better the material for reducing sound transmission. While convenient as a figure of merit, in critical applications, such as broadcasting, it is always preferable to consider transmission loss at different frequencies to meet specific design objectives. There is no shortcut that can overcome the fact that transmission loss varies as a function of frequency.

Transmission Loss of Solid Materials

The TL of a solid wall tends to have three defined frequency regions as shown in Fig. 19: a mass controlled region at low frequencies, a plateau region (also known as a coincidence dip) and another mass controlled region above the plateau region. For rigid materials, especially at low frequencies, TL increases about 5 dB for each doubling of surface weight. In general, heavier materials have greater sound isolation capability.

Discontinuous Construction

Increased mass is one of two effective methods to reducing the transmission of sound energy. The other technique is the proper application of airspace, sometimes called the discontinuous construction technique. The TL of a wall can be significantly improved if two or more layers are used with an airspace in between. With a significant airspace, the composite wall behaves as two independent walls, and the composite TL can approach the sum of the individual TLs. Placing glass fiber between the walls can improve the STC 3 dB to 10 dB due to damping the resonant coupling of the wall panels. The STC performance of common wall constructions is shown in Fig. 20.

Windows

Achieving windows with high acoustic isolation requires a somewhat complex structure. A double pane with a 6" air gap between the panes is desirable. The glass must be resiliently mounted within the frame without cracks or air gaps. For best performance, the two panes should be different thicknesses and the

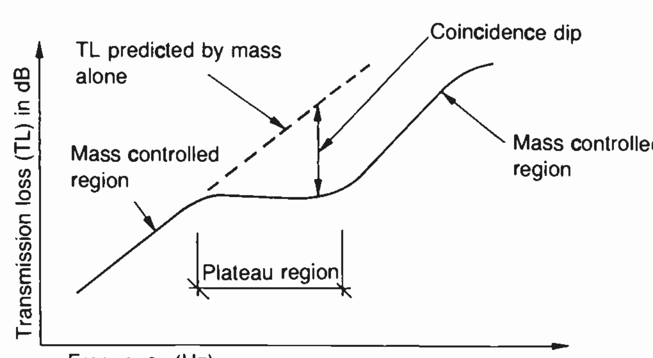
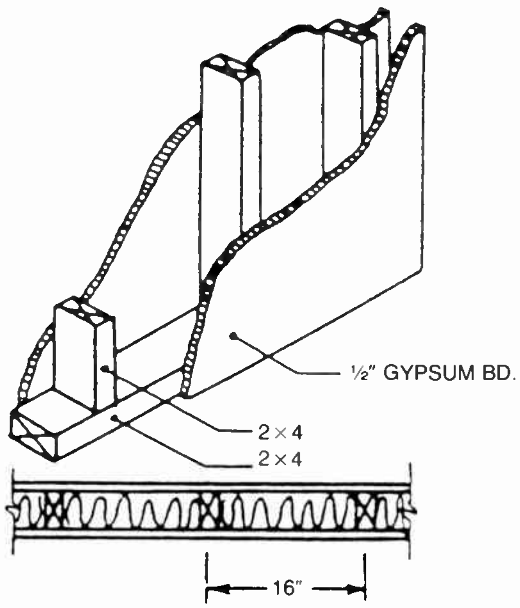


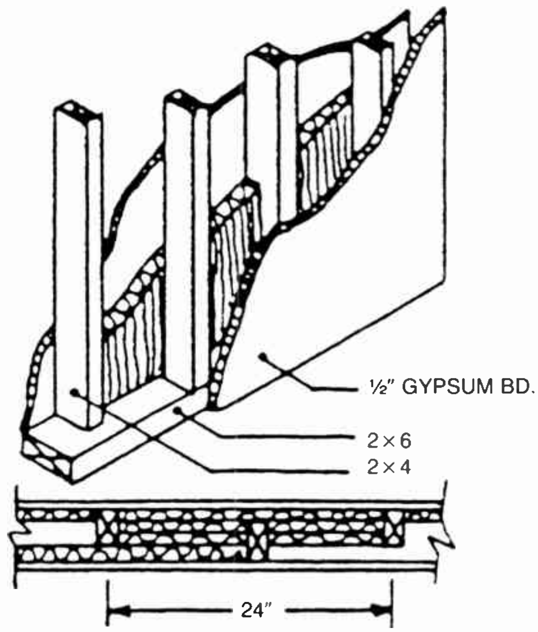
Figure 19. General transmission loss (TL) profile of solid materials.



Construction—standard stud partition

	Transmission Loss (dB)						STC
	125 Hz	250 Hz	500 Hz	1K Hz	2K Hz	4K Hz	
without glass fiber	15	27	36	42	47	40	35
with 3½" glass fiber	15	31	40	46	50	42	39

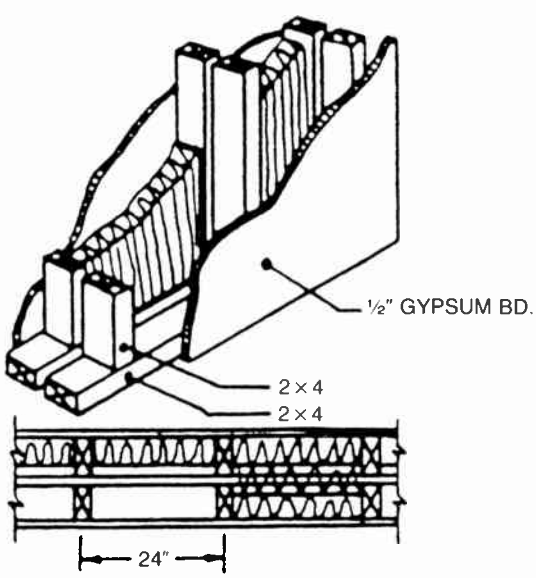
Figure 20A. Standard stud partition. (Reprinted with the permission of Owens-Corning Fiberglas Corporation)



Construction—staggered stud partition

	Transmission Loss (dB)						STC
	125 Hz	250 Hz	500 Hz	1K Hz	2K Hz	4K Hz	
without glass fiber	22	23	36	46	52	41	38
with 3½" glass fiber	31	37	47	52	56	50	49

Figure 20B. Staggered stud partition. (Reprinted with the permission of Owens-Corning Fiberglas Corporation)



Construction—double wall

	Transmission Loss (dB)						STC
	125 Hz	250 Hz	500 Hz	1K Hz	2K Hz	4K Hz	
without glass fiber	24	32	39	48	52	39	39
with 3½" glass fiber	29	43	54	63	66	52	52

Figure 20C. Double wall partition. (Reprinted with the permission of Owens-Corning Fiberglas Corporation)

perimeter of the air gap between the panes covered with sound absorbing material. Often one pane is angled with respect to the other, a characteristic which has more advantages in visual glare reduction than affecting the TL significantly.

Doors

High isolation doors are difficult to build and maintain since the gasketing and sealing of the door must be extremely precise and not degrade with time or wear and tear. Because of the difficulties of constructing and maintaining doors with high isolation, it is often more economical and reliable to create high isolation entry ways through use of a "sound lock", that is, an outer door leading to a vestibule with an inner door.

The overall TL from the outside hall to the inside of the room via the vestibule can be very high even with only moderately stringent construction techniques. Also, by sequentially entering or exiting a sequence of doors, the room is never exposed to the total loss of isolation inherent to a single door partition.

Floating Construction

Floating construction is a combination of discontinuous construction and resilient mounting techniques. Floating floors are solid slab floors which are completely isolated from the structural floor by a resilient underlayment or resilient isolators. Walls may be built attached to the floating floor and a ceiling may be resiliently hung from the structural ceiling, resulting in an actual room-within-a-room with very high isolation possibilities. This is shown in Fig. 21.

This type of construction can be extremely expensive but is sometimes the only avenue to achieving high levels of sound isolation, especially at low frequencies, in high noise environments, or when very low ambient noise levels are required.

Vibration Isolation

Vibrations and sound energy produced by mechanical equipment can be transmitted throughout a building via vibration of the structure, and reradiated as sound energy in a particular room. In general, vibrating equipment can be effectively isolated from the building structure by mounting the equipment on resilient mounts. The mass of the equipment and the compliance of the resilient mounting form a resonant system. Vibrations of the equipment at frequencies much higher than this resonant frequency will effectively be prevented from being transmitted into the building structure. As a rule of thumb, the resonant frequency should be ⅓ or less of the lowest desired frequency of effective isolation. The lower the resonant frequency, the lower the level of transmitted vibrations for a given vibration frequency.

Impact Noise Reduction

Impact noise, as the name implies, refers to such mechanisms as footsteps, objects dropped on floors, slamming doors, and so forth. Reduction of impact noise may in many cases be effectively treated by the obvious: a rug on the floor, castered chairs, and other measures can drastically reduce impact problems by "softening the blow." To obtain better improvements,

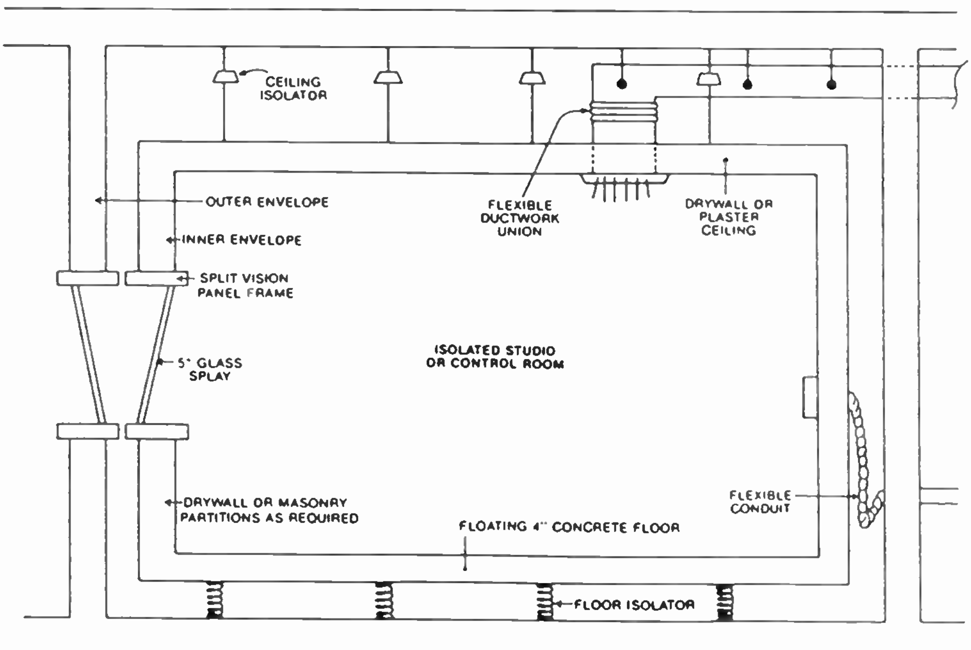


Figure 21. Floating construction technique.

the same techniques to improve TL such as discontinuous construction can be effective.

Similar to the concept of STC, Impact Isolation Class (IIC) is a single number rating system to assess a barrier's effectiveness at arresting transmission of impact noise. The IIC method is based on the use of a standard "tapping machine" which supplies a known impact noise profile. SPL readings are then taken in ⅓ octave bands in the receiving room and compared with a standard contour to determine the IIC rating.

HVAC Considerations

Often the HVAC system in a room is the primary source for noise introduced into the room and also determines the level of acoustic leakage between rooms. With proper design, these limitations can in large measure be avoided.

A major contributor to HVAC noise is the fan noise itself which propagates down the supply ducts and subsequently enters the room. Ventilation ducts can be lined with absorbing material such as 1" glass fiber to reduce the noise at the end of the duct. Prefabricated duct silencers are also available which are placed in line with the duct and offer significant sound attenuation characteristics. Structure borne transmission of fan noise can be stopped by coupling the fan motor to the duct system via a canvas or rubber coupling to break up the vibration path.

The amount of noise in a duct system is also strongly dependent on the air velocity in the duct. Low noise design requires low air velocities which, for the same amount of total airflow, leads to the use of either multiple ducts or larger duct cross sections, both of which imply higher costs. Diffusor designs to minimize turbulence induced noise at the entry point to the room are also important.

The layout of supply and return ducts can be a hidden source of poor acoustic isolation between rooms. As shown in Fig. 22, higher isolation is achieved when the supply or return duct paths connecting two rooms are longer. The optimum situation is to have completely separate ducts back to the fan source; however, this may not be practical in some situations.

CONCLUSIONS

This chapter has presented basic principles of acoustic design that are relevant to the design or renovation of broadcast facilities. Hopefully, this material will offer some perspective and insight to the reader. There are, however, many different techniques of using these principles to achieve specific acoustic design goals and a wealth of literature and experience exists in this area. The value of a professional acoustical consultant in a major design or renovation project should not be underestimated. Acoustical design techniques can be very expensive and mistakes can be even costlier. With a basic understanding of acoustic principles, users will be better able to explain their needs to a consultant, understand the ramifications of various design tradeoffs and develop the ability to make informed and competent decisions in the area of acoustics.

BIBLIOGRAPHY

L.L. Beranek: *Acoustics*, McGraw-Hill, 1954.

L.L. Beranek: *Acoustic Measurements*, John Wiley, 1949.

L.L. Beranek: *Music, Acoustics and Architecture*, John Wiley, 1962.

L.L. Beranek: *Noise and Vibration Control*, McGraw-Hill, 1971.

D. Davis: *Acoustical Tests and Measurements*, H. Sams, 1965.

D. & C. Davis: *Sound System Engineering*, 2nd Ed., H. Sams, 1987.

J. Backus: *The Acoustical Foundations of Music*, W. Norton, 1969.

V. Mankovsky: *Acoustics of Studios and Auditoria*, Hastings House, 1971.

V. Knudsen and C. Harris: *Acoustical Designing in Architecture*, Am. Inst. of Physics, 1978.

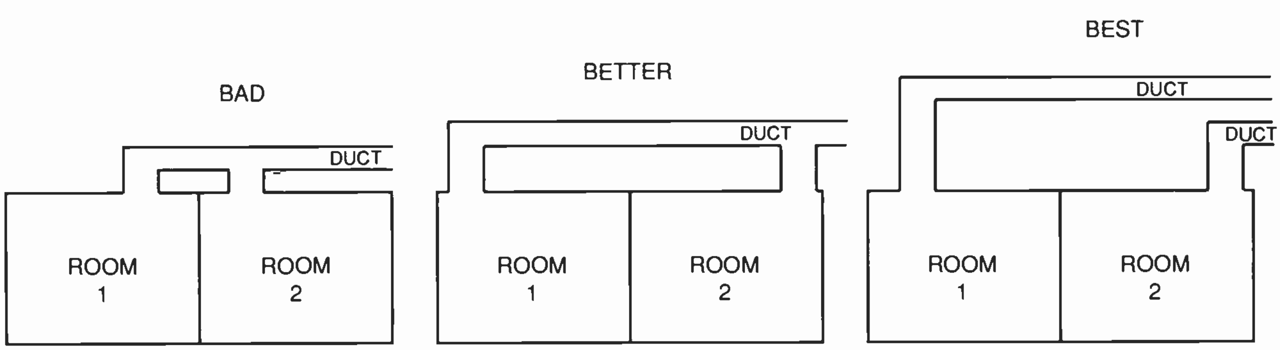


Figure 22. Effect of duct location on sound isolation.

L. Doelle: *Environmental Acoustics*, McGraw-Hill, 1972.

M.D. Egan: *Concepts in Architectural Acoustics*, McGraw-Hill, 1972.

M.D. Egan, *Architectural Acoustics*, McGraw-Hill, 1988.

F.A. Everest: *Acoustic Techniques for Home and Studio*, 2nd Ed., Tab, 1984.

F.A. Everest: *How to Build a Small Budget Recording Studio from Scratch . . . With 12 Tested Designs*, Tab, 1979.

F.A. Everest: *The Master Handbook of Acoustics*, 2nd Ed., Tab, 1989.

J. Cooper: *Building a Recording Studio*, Recording Institute of America, 1978.

A.P.G. Peterson and E.E. Gross: *Handbook of Noise Measurement*, 9th Ed., General Radio Co., 1980.

L.F. Yerges: *Sound, Noise and Vibration Control*, Van Nostrand Reinhold, 1969.

W.W. Seto: *Theory and Problems of Acoustics*, McGraw-Hill, 1971.

C.M. Harris: *Handbook of Noise Control*, McGraw Hill, 1981.

P.H. Parkin and H.R. Humphreys: *Acoustics, Noise and Buildings*, F. Praeger, 1957.

M. Rettinger: *Acoustical Design and Noise Control*, Vols. 1 and 2, Chemical Pub. Co., 1977.

Glen Ballou: *Handbook for Sound Engineers, the New Audio Cyclopedia*, H. Sams, 1987.

C. Gilford: *Principles and Applications of Room Acoustics*, Peter Peregrinus Ltd, 1972.
25.

D. Hall: *Basic Acoustics*, Harper and Row, 1987

5.14 Lighting for Television

Bill Marshall
Imero Fiorentino Associates, New York, New York

BASIC PRINCIPLES OF LIGHTING FOR TELEVISION

The continuing development of increasingly sensitive television cameras has allowed the television lighting director to reduce the effort required to provide high footcandle levels of illumination and devote more attention to the creative aspect of his craft. Despite the many advances in cameras and digital signal processing, the lighting director must remain aware that the home receiver with its limited range of contrast (30:1 approximate) continues to be the weakest element of the television system.

Since television is a two dimensional medium, lighting of the subject is critical to portrayal of three dimensional forms. Even the impressive resolution of recently developed high definition television can only replicate an image as it is revealed by light and shadow. Creative lighting, rather than simple illumination of the subject, will both model the form of the subject and maintain its relative balance of intensity to the rest of the image. Lighting design can be most accurately defined as *constructive use of controlled light for a predetermined objective*.

A major portion of the shots used in television production are closeups of one or a few faces; therefore, lighting of faces is the most important task. The standard techniques for modeling faces with light in still photography were simply a modern adaptation of the same approach used by portrait painters for centuries.

The Three-Point System

The basic arrangement of lights for television is commonly referred to as the two- or three-point system. The first point or element is the *key light* which is the principal illumination of the subject's face. It is placed near the lens axis or slightly to one side of the camera and generally at an elevation of about 30° above the horizontal line between the subject's face and the camera lens. The elevation should be adjusted relative to the particular structural elements of each subject's face, such as length of nose or depth of eyes relative to the brow. The location in practical terms will obviously be determined by where or how the fixture can be supported. In studio work, this means the key light is very often suspended much lower than the general grid height to gain precise positioning.

The second point or component of this lighting approach is *back light*. Located above and to the rear of the subject, the back light creates a glow on the hair and a highlight on shoulders separating the subject from appearing to be part of the background. The combination of a camera mounted key light and stand mounted back light is the most prevalent arrangement used in electronic news gathering (ENG) and electronic field production (EFP). Ironically, the use of the camera-mounted key light alone often creates a somewhat less than ideal picture but has come to have direct association with realism.

The third element of the classic three-point lighting system is *fill light*. A single key light can often provide modeling that is too severe in the absence of other ambient light sources. Fill light generally softens and blends the back and key light accents while maintaining their purpose to highlight and separate the form and can bring up overall illumination level.

While a set-up such as described above is relatively easy with one camera and one subject, the reality of television production is that there is generally more than one camera and multiple subjects. To light an actual production, the lighting director will expand the basics of the three-point system to cover all the subjects and for all the camera angles. This often requires clever fixture arrangements wherein the function of a fixture has multiple uses. For example, in a typical two person interview, the lighting design might be arranged such that one subject's key is simultaneously serving as the other subject's back light.

While carefully positioning light fixtures for modeling, the lighting director must also consider the control of the shadows that the light sources will create. For each camera angle, the key light and back light relationship must be maintained. Microphone boom shadow are the most difficult to control. Ultimate control of boom shadows requires careful coordination between the boom operator, the director, the lighting director, and the scenic designer. Other common shadow problems generally within the lighting director's control include elongated nose and chin shadows which occur from improperly positioned key or back light fixtures.

It should be noted that projected shadows, shapes, or patterns can be some of the most powerful techniques available to a lighting director in suggesting time, place, mood, or even pure background decoration. The character and quality of light is often more apparent from the shadows it casts. Shadows can be hard, soft, transparent, or even different color than the apparent source. Because shadows are so important, major lighting arrangements should be carefully planned and drawn to scale. Section views in scale will help to determine where the shadows will fall and if there will be any scenic conflicts. This preplanning will save a tremendous amount of time in the studio or on location.

Once the lighting of the faces is established, only then does the lighting director plan the background lighting. While a TV program can be set in almost any conceivable background, the lighting of faces continues to determine the relative brightness of the backgrounds. The reflective value of a face ranges between 28% and 41% while a white wall can be 96% reflective. A lighting director must therefore be very careful to use lighting to maintain a proper balance, since the viewers' attention will be naturally drawn to the brightest spot of a television picture. Lighting intensity along with scenic element selection are critical to controlling the viewers' focus.

A generic background found only in the studio is the *cyclorama* or "cyc," used to provide a broad, uniform, and neutral background. It may be fabric or a plastered wall and is usually arranged to merge with the floor so that there is no visible joint. A cyc can easily be overlit, and it will, in contrast, appear to darken the skin tones of talent in front of it. Projection of light patterns on the cyclorama is often an inexpensive way to create a setting with light; but again, the patterns which decorate a long shot should not appear undesirable or too bright relative to natural skin tones in the close-ups. Elaborate scenic treatment may require all types of built-in specialty lighting and a whole assortment of fixtures for each unique requirement.

LIGHTING LEVELS

There is no one proper level of light except as required to allow a particular camera to make a good picture under specific conditions. The level required will vary by camera type, lens, type of action, and any existing ambient lighting conditions, as well as other purely aesthetic considerations. The lighting director must control the relative balance of the lighting for every given situation whether it is in brilliant sunlight or in a dark dramatic stage setting. There are often situations where the natural conditions greatly exceed the limited 30:1 contrast range of the television broadcast receiver so it is also the lighting director's responsibility to lead the director away from unresolvable situations.

As with any camera lens system, the amount of light required is a direct relationship between the pickup medium, the lens aperture, and the depth of field desired. Long focal length lenses used in sport events have smaller apertures by design. Since the fast moving action demands a high depth of field, the lighting level must therefore be quite high to satisfy this combination of factors. Conversely, the exact same camera in a dimly lit studio close-up may require a fraction of the illumination level of the sporting event. Good picture quality and proper level can only be determined through the eye of the video camera.

There are many devices available to the lighting director to control the levels of illumination. First of all, most fixtures accept lamps of varying wattages within the same housing, and there is quite a wide variety of television studio or location lights available. How far the fixture is physically placed from the subject will also directly affect intensity. Intensity can also be controlled by a scrim, which is actually a wire screen inserted in front of the lamp. The greater the density of the screening is, the more the light level is reduced. Actually, any medium placed in front of a fixture will reduce its output to a predetermined degree. Spun glass or any of the many other types of sheet diffusion materials available will both reduce intensity as well as reshape the pattern and quality of the light. Spun glass softens or diffuses the light. Even colored filters reduce light output relative to the coefficient of transmission of the particular color. For example, a blue filter passes very little light while a yellow-green filter will reduce the output light level very little. The most precise and convenient level control of incandescent lighting is dimming. While seldom used outside a studio theatrical environment, it alone does not eliminate any of the above mechanical means of level control. While dimming does change the color temperature of the lamp, it can still be used effectively. There is more discussion of dimming and its control later in this chapter.

COLOR TEMPERATURE

For accurate color rendition, the television system must operate within a consistent range of color temperature. Color temperature is measured in degrees Kelvin (°K), and varies dramatically from type of artificial light to time of day and atmospheric conditions of daylight. A television camera must be white balanced for the same lighting conditions as the scene will be shot. Segments shot separately which are to be seamlessly edited together must be shot under consistent color temperature light for accurate color rendition between segments. In our daily routine, the eyes experience the effects of different color temperatures while the brain interpolates color rendition. Recalibration of the camera for varying lighting conditions has somewhat the same effect, yet it is often important to retain some of the original lighting color and character. The television system is actually quite flexible, permitting color temperature swings of ±300°K without major effect to visible color rendition.

Windows generally present the greatest color temperature problem because exterior daylight is a tremendously different color temperature and is extremely bright relative to the artificial interior lighting . Large sheet filters which correct both color temperature and intensity are available for this purpose; however, if the window area is extremely large and daylight dominates the scene, it is generally easier to work with the high color temperature artificial light sources rather than incandescent.

LIGHTING FIXTURES

Requirements on lighting equipment differ considerably between the studio and on location. On location, there is often little control, and conditions can vary from full sunlight to little or no light. Studios are usually designed and equipped to give the lighting director considerable control over all aspects of lighting: intensity, color, placement, and screening.

The following sections discuss considerations and typical equipment for studio and location lighting.

Location Lighting

Since lighting design is deliberate control of light for a desired effect, good lighting design may not even require the use of any fixtures at all. In many daylight situations, a lighting director may choose to simply use reflectors and scrims to achieve the desired effect.

Movement of the sun is predictable so the survey of the site must be made at the same time of day you will eventually shoot at in order to select workable camera positions and to determine the proper equipment. The sun as a key source is very harsh and requires intense fill light for the talent. Reflectors are very effective for this purpose. They require no electrical power, but must be attended by a person at all times to keep them oriented correctly to the moving sun. Lightweight folding reflectors are convenient to pack and set-up but do not present the necessary hard, stable surface needed to provide a smooth field of light. The heavy solid panels are less convenient but are stable. Another way the harsh light of the sun can be controlled in intensity is by using large silks or nets to soften or to shadow the talent. Nets and silks under 6 ft. × 6 ft. attached to frames are called butterflies; up to 20 ft. × 20 ft. are called overheads. Wind is an obvious factor in the decision to use these tools.

Figure 1. For remote lighting, compact kits which contain a large assortment of fixtures, stands, reflectors, and grip hardware are available. (Courtesy of Lowel-Light Manufacturing, Inc.)

For smaller remotes, the lighting director's selection of fixtures is often governed by weight and portability relative to lumen output. There is a tremendous variety of extremely ingenuous and lightweight fixtures available. These generally are packaged in kits with a full assortment of stands and grip hardware to accommodate any situation which may occur while on location. Many of these kits are designed to incorporate dichroic filters to match the 3,200°K incandescent lighting to the 5,600°K daylight. This filter reduces the light output quite noticeably.

For larger location work involving daylight conditions, HMI® lighting has become the standard of both the film and television industries. These highly efficient ballasted arc source lamps come in a wide variety of sizes from 200 watts to 20,000 watts. They provide about 80 lumens per watt at 5,500°K which is four times the output of incandescent light. HMI® lighting is often the best solution for interior situations with tremendous window exposure. As in any exterior situation, the lighting director must carefully monitor the changing color temperatures of daylight and remember that the light of a clear sky late afternoon is quite a different color and character than mid-morning when the sky was overcast.

Figure 2. HMI lamps are most often used in fresnel lens housings and are available in a wide variety of sizes. The associated lamp ballast is a separate unit from the lamp housing. (Courtesy Arriflex Corporation)

"HMI®" is a registered trademark of OSRAM, although it is often used (or misused) in referring to lamps of other manufacturers. The name refers to the basic elements which are combined in the lamp's quartz envelope to create the unique color temperature: "H" for mercury (Hg), "M" for the various metal halide rare earths, and "I" for the halogen iodine and bromine.

Often a lighting director is called upon to shoot in situations where the existing lighting is a low ceiling covered with fluorescent fixtures and they must be used because they characterize the space. Fluorescent light does not provide a full spectrum light and the most common type of "cool white" tube has very little red. While it is possible to color correct each fluorescent lamp using a minus green color correction gel, it is often impractical. By color balancing the camera under a representative mix of the fluorescent and the 3,200°K talent lights, acceptable skin tone can be produced, though the background may still appear greenish.

There are a wide variety of fluorescent lamps available and some offer quite a high color rendering index. A line of fluorescent studio fixtures is now available; however, control of the light is limited by the relatively large size of the source: a long fluorescent tube as compared with a small quartz bulb.

Exterior shooting at night for television requires great simplicity to look real. Large HMI®s are very useful because one large source can supply the basic illumination for a very wide area. Within an exterior scene there are always elements supposedly lit by artificial sources, either seen or imagined. With the camera balanced for 3,200°K, quartz lights can be used for the people areas while the overall HMI®s at 5,500°K will seem very blue. If the HMI® light is too blue, it can be corrected slightly until it seems to be the proper gray-blue of moonlight. Using smaller HMI®s as back and rim light will further enhance the moonlight effect. In attempting to provide realism, lighting must suggest the correct mood. A pretty picture is more important than realism. Too much realism can be as distracting to the audience generally as too little.

Studio Lighting

Working in a studio the television lighting director has the greatest control over lighting quality. A TV studio is by definition an idealized environment for production of television programming.

In a well equipped studio the television lighting director has quite a variety of lighting fixtures or luminaries available to accomplish the basic objectives of TV lighting: separation, modeling, accent, illumination, and directing viewer attention. This variety of fixtures and the various qualities of light which they produce have been developed to enhance the efficiency of studio operation. The luxury of utilizing a basic fixture and manipulating the quality of light by external diffusers, cutters, reflectors, etc. for each shot as in

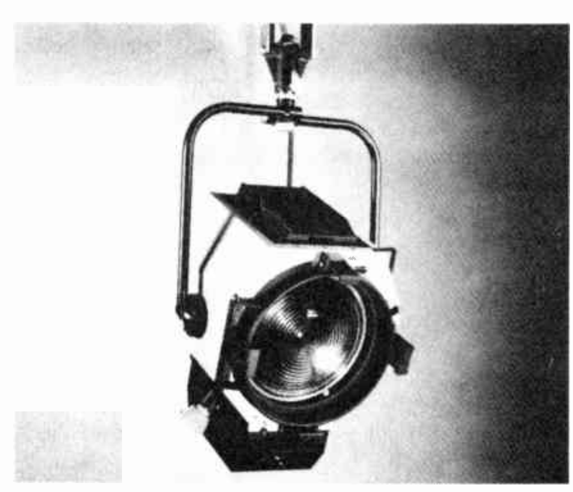


Figure 3. Studio quality fresnel spotlights are much more durable and have superior optical systems compared to more inexpensive theatrical units. (Courtesy of Colortran)

Figure 4. Studio quality fresnel spotlights must offer good quality long leaf barndoors as an essential accessory. (Courtesy of Strand-Century)

film is not normally possible in multi-camera television production. Therefore, a complement of fixtures with given characteristics provide a lighting director with a palette of choices to design the lighting.

Spot Lighting—Fresnels

In a studio, the most common and useful fixture is the *fresnel lens spotlight*. The lighting from a fresnel is very controllable and has a smooth even field. By a simple adjustment of the position of the lamp, the fixture will produce either a narrow spot or wide flood beam of light. Equipped with barndoors this beam can be further shaped to virtually any pattern. Slots in front of the lens allow color, diffusion media or screens for intensity control to be slipped in place. A complement of single and double thickness, and half and full frame screens should be provided for all fresnels even when dimmers are available.

For most 14 foot high grid studios, the six-inch, eight-inch, and ten-inch fixtures from 500 to 2,000 watts are the most common units. In smaller studios or tight applications, "baby size" fixtures are often employed. These units enclose the same wattage lamps in small housings. A number of people prefer these smaller size units for their easier handling and somewhat different optical characteristics. As noted previously, the higher the grid, the higher the wattage fixtures are required. Therefore, in higher studios, the standard complement of fixtures might be 5,000 and 2,000 watt units or even up to 10,000 watts. Conversely, fresnel lens units are also available in sizes down to a three-inch lens with a 150 watt lamp.

It is important in selecting a line of fresnels for the studio to weigh carefully both the optical and mechanical features:

1. *Stability*. Poorly made fixtures will not focus properly once they are hot.
2. *Balance*. Fixtures should also be well balanced, even with barndoors in place, to maintain both focus and direction.
3. *Accessories*. The manufacturer should also provide a complete line of suitable accessories including screen sets, diffusion frames, stand mounting hardware, and barndoors which are long enough to control spill adequately.

Broad Area Lighting

There are several fixtures commonly used to provide fill, or base, lighting to cover a wide area.

Scoops. *Scoop floodlights* are used for a variety of soft, diffuse requirements; most commonly grouped to provide fill, or base, light to a setting. These are very simple fixtures available in several sizes, where bigger reflectors are generally better. Scoops are used almost invariably with diffusion media, so it is essential that a diffusion frame is ordered with each scoop fixture.

The scoop is the only fixture in which a tungsten halogen lamp should *not be used*. The large old fashioned PS-shape incandescent lamp produces a much softer quality of light, because of its larger filament area. Some manufacturers make focusing scoops, which is generally a pointless additional cost when used with diffusion material.

Other Fixtures

While the majority of lighting can be done with fresnels and scoops, there is a variety of fixtures which produce a different quality of light. For a harder rectangular field of fill light a *broad* can be used. It is very similar to a scoop in concept, except its housing/reflector creates a somewhat less diffuse light. Like the scoop, the broad is almost always used with diffusion media.

The *softlights* are another type of diffuse fill, or base, light fixture, in which the light sources are totally concealed and the light is reflected onto the set in

Figure 5. A typical scoop unit. (Courtesy of Mole-Richardson)

Figure 6. Softlights produce a very diffuse source and have switches to select as many lamps as necessary. (Courtesy of Mole-Richardson)

an indirect manner. These units are not particularly efficient but are unequaled in providing a shadow-free light. They are available in a variety of sizes. Large 4,000 to 8,000 watt units are pretty unwieldy for many applications, but the smaller units in the 1,000 to 2,000 watt range are becoming very popular as fill lights because the indirect nature of the source is very easy on the eyes of on-camera talent.

Probably, the most efficient incandescent fixture used in television lighting is the *Par* unit. The lamp is a sealed unit containing both the reflector and the lens, and is available in five different beam spreads. While limited to very large studios and remotes, the Par is an inexpensive and powerful tool for the lighting director.

This type of lamp is manufactured with an internal *par*abolic reflector and lens. Par lamps do not offer very precise control or even a very large field of illumination, and are therefore limited to large studios or remotes, which rely on their powerful output. Par lamps and their corresponding fixtures are manufactured in many sizes and wattages but only the 1,000 watt Par 64 is used extensively in television lighting. This type is 3,200°K and is available in five beam widths.

The *ellipsoidal* reflector spotlight is a unique effect light in a studio. The optical system of this unit allows the beam of light to be hard or soft edged and the projected pattern of the beam can be precisely shaped

Figure 7. The Par 64 is one of the most widely used fixtures in large studios and theatrical type applications. (Courtesy of Altman)

Figure 8. Shutters on ellipsoidal spotlights allow the beam of light to be shaped into hard edge shapes or to project patterns. (Courtesy of Colortran)

Figure 9. The three level legs allow the Century stand to fold flat for storage. The head can grasp numerous flags, cutters, cookies, etc.

by internal shutters. It is most commonly used to project patterns on the cyclorama. Numerous photo-etched metal patterns are available in a variety of optical systems, but the wide-angle short-throw sizes are most useful in small and medium studios.

Accessories

There are a number of accessories which make the lighting system work. Despite the expenses and care expended on planning the grid system, it cannot satisfy every fixture mounting requirement; every studio must have some fixtures on *rolling floor stands*. In addition, many types of *clamps* and general grip equipment are available for special mounting situations.

One of the key accessories in the lighting system are *extension rods*. These devices allow fixtures to be hung below the level of the fixed grid. For example, a key light at a 30° angle (elevation) above the talent could be moved much closer to the talent and provide more light if it were hung lower. Counter-balanced devices such as pantographs and spring load telescoping hangers may be used, but these cannot be locked or clamped in place the way a simple rod can be. Another drawback of the flexible devices is that they must be precisely adjusted to the weight of a particular fixture and cannot be relied upon to stay in that position. Extension rods, which are nothing more than pipes that may be adjusted up and down, are the cheapest and best choice.

The *Century stand* (Mole Richardson), and its variations, features a versatile clutch grip which can grasp extension arms and all types of material used in light control, such as flags, cutters, cookies, and support reflectors at virtually any angle. Century stand use is limited only by one's imagination. No matter how many are available, they will probably all be put to good use.

Ultimately, the proper quantity of fixtures, accessories, and other components of a lighting system for a studio of a given size can vary widely according to the requirements of a specific situation. For example, a studio with a fixed grid and a tight production schedule will function more efficiently with a heavy saturation of fixtures. With a large quantity of fixtures, the need to relocate and adjust units is reduced. On the other hand, with sophisticated motorized grid systems, the ease of relocating fixtures will reduce the total quantity of fixtures needed for a given size of studio. The choice of either of these two approaches often depends on labor costs within a particular facility. Too often the expense of a few extra fixtures is not compared to the additional time and labor required to fully utilize a minimal complement of equipment or the cost of elaborate rigging. Final decisions on equipment purchases should be based on a careful analysis of specific production requirements, and such factors as the effective cost of downtime, to setup for the next scene or production.

It should be evident from the above general discussion of lighting problems that the close cooperation between the lighting director and the video shader is absolutely essential. The video shader is in control of a marvelously versatile picture machine whose limits are stretchable in the name of art. Great television lighting is frequently adequate lighting enhanced by creative video control.

STUDIO DESIGN

In designing a television studio, the lighting system is a major consideration, as it is closely related to the physical size and shape of the room. While the

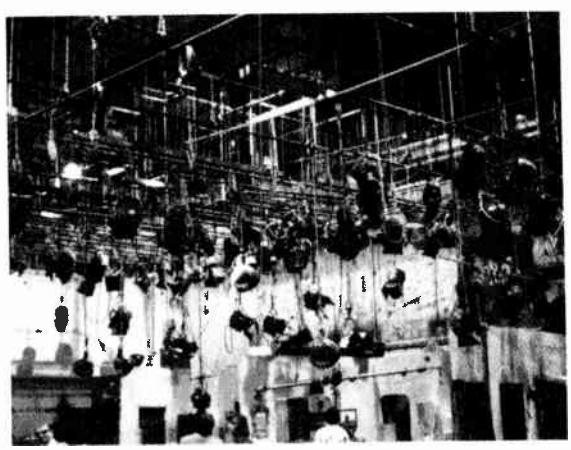


Figure 10. This large fixed grid studio illustrates the relationship of the various aspects of the studio lighting system. (Courtesy of WPBT)

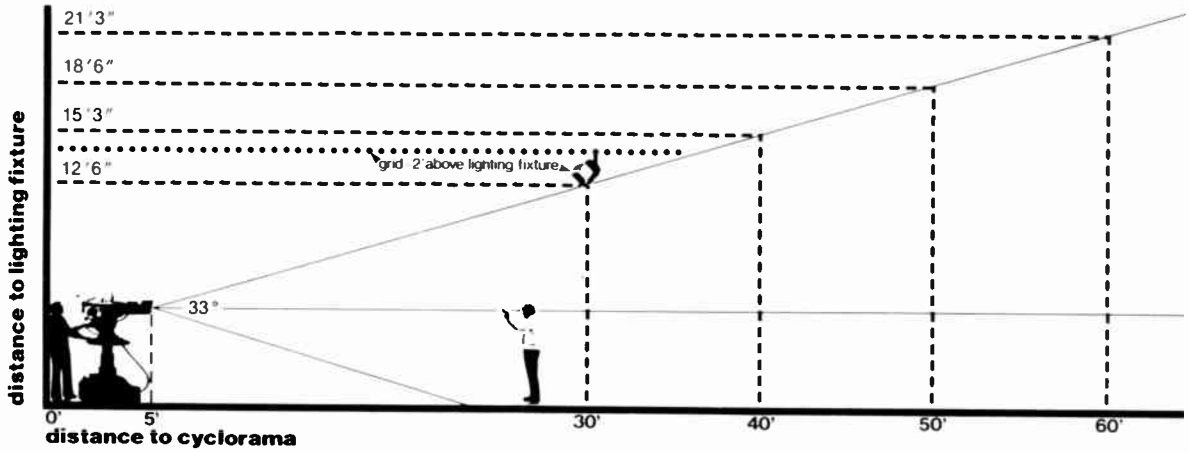


Figure 11. This drawing illustrates how the vertical aspect of a camera lens at eye level determines the height of a studio grid.

television industry has seen tremendous technological improvements throughout its evolution, the basic physics of light and its ability to describe three dimensional form cannot change. Consequently, many older studios have upgraded virtually every piece of broadcast equipment while the basic components of an older, well-designed lighting system still function extremely well.

Studio Size

Ideally, a studio provides the optimum environment for any type of production. In practice however, studios of various sizes tend to function best for particular types of production. The typical broadcast plant requires several sizes of studios most effectively to service its programming schedule. Typically, a small studio (1,200 sq. ft.) is dedicated to news, interviews, and public affairs, while 2,400 sq. ft. is the most common size of small production studio. For general production, a studio of 5,000 sq. ft. or more will impose few limitations. The size of a studio must be carefully determined by existing and future programming requirements, keeping in mind that the larger the studio, the fewer the limitations. These are important decisions, since the lighting system must be planned relative to the size and specific requirements of the studio.

Studio Height

When determining the height of a new studio, the lighting suspension system must be given careful consideration. The grid height for small and medium size studios is a function of the current 4:3 (width:height) TV aspect ratio and the typical wide angle zoom lens. Most zoom lenses can cover approximately a 45° field. By calculating the maximum width of a set or shot, the height of the grid can be determined by applying the aspect ratio. Keep in mind that the studio lighting fixtures will hang approximately two feet below the suspension structure. In addition, a normal studio pedestal and cameraperson will prevent the lens from getting closer than approximately five feet from a wall or any other obstruction, further reducing the maximum coverage of the lens. This method describes the theoretical maximum picture possible, and the height the fixtures need to be mounted to stay out of the picture. However, television is primarily a close-up medium, and limitations in grid height can be overcome by various camera angles and special fixture mounting systems. Current experiments in high definition television may bring about a new, wider aspect ratio (16:9 or 5.33:3) in the future, which will actually reduce the studio height requirement.

SUSPENSION SYSTEMS

The suspension system for the studio lighting is a critical factor in determining studio height. All types of lighting suspension systems are commonly referred to as the "grid." In its simplest form the system consists of a series of pipes suspended below the studio ceiling, in a regular pattern. A grid must allow lighting fixtures to be hung anywhere over the entire studio. Actually the grid can consist of several different types or a combination of suspension systems.

Fixed Grid

A fixed grid is the most common and least expensive system of suspending lighting fixtures. The pipes are generally laid out on a four-foot square spacing, which provides adequate flexibility in hanging positions. The most common fixed grid height is 14 feet. This height offers a good compromise between easy ladder reach and adequate clearance for wide shots. A fixed grid that is much higher than 14 feet greatly increases the amount of labor needed to install and adjust fixtures. For a large, multipurpose studio, a fixed grid can be a serious limitation. Bi-level fixed grids have been utilized to create a higher apparent background without

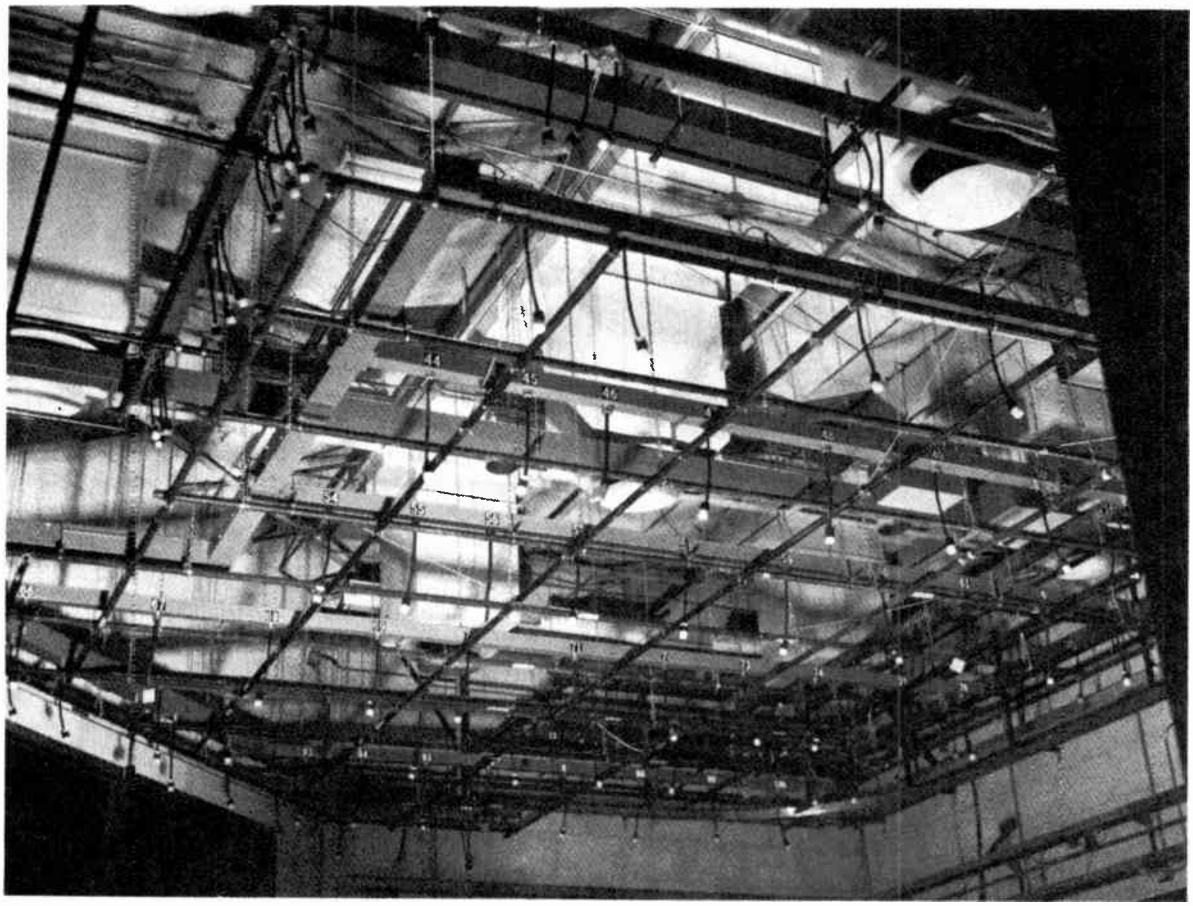


Figure 12. This fixed grid studio employs two levels to create a higher apparent background. Multiple pigtail outlets are positioned for cyc lighting units. (Courtesy of WFAA-Dallas)

additional cost, and to keep the major portion of the grid at a reasonable height. To solve this problem more efficiently, several other systems have been developed.

Catwalks

Catwalks are generally also hung at a fixed level. Catwalks provide a measure of increased efficiency over a fixed grid. Although they are somewhat more expensive, catwalks allow studio electricians to work on lighting at the same time carpenters are handling the scenery. This can be an important time savings on a tight production schedule. Catwalks are usually arranged to create a fixed grid utilizing their handrails in conjunction with an extension rod on each fixture. When hung too low, a catwalk creates the same height limitation as a fixed grid, but generally a catwalk is higher than a fixed grid, with extension rods lowering the fixtures to a more effective height. Unfortunately, catwalks do not eliminate actually getting on a ladder to adjust the fine focus of each fixture. Catwalks primarily save time in set up and strike and offer the crew a safe work platform above the studio.

Recently, most fixture manufacturers have developed pole-operated yokes for their units. Each fixture has a socket, which can be reached by a pole crank to adjust each of the basic functions, pan, tilt, and spot-flood. There is a reasonable limit to how long a pole can be easily manipulated, but this feature considerably enhances fixed, high-level grid systems. It is not possible to focus studio fixtures by this method as quickly or efficiently as by hand, but this system does provide a means of focusing otherwise unreachable units blocked by scenery and is seriously worth considering for any type of grid system.

Battens

In very large studios, a basic theatrical staging factor — "flyout clearance" — in addition to the TV aspect ratio, determines the grid height. Just as in a theatre stagehouse, larger-scale TV production requires that scenery be suspended by some arrangement so that it may be raised for storage and lowered into place for use. Flown scenery commonly requires a minimum grid height of 40 feet. To raise and lower the scenic units and the lighting fixtures, a regular pattern of long battens or pipes run across the width of the studio. These battens run on steel cables and are manually balanced by cast iron weights. These counter-

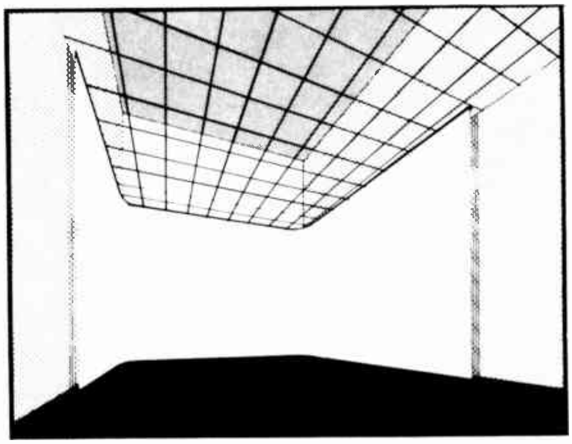


Figure 13A. Fixed grid plan.

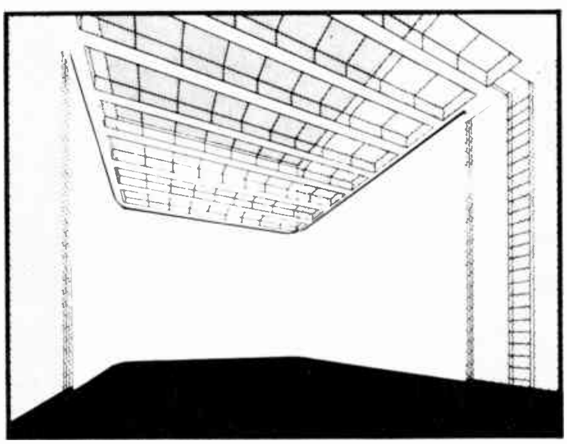


Figure 13B. Catwalk grid plan.

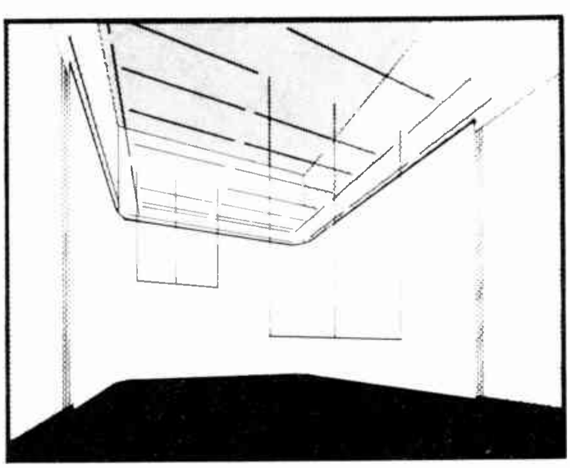


Figure 13C. Batten plan.

weighted battens solve the weight problems but add additional work in rebalancing the counterweight arbors every time fixtures are added or removed from the batten. Rigging for counterweight, typically mounted along the studio wall, encroaches on valuable floor space.

Winches

To overcome the disadvantages of counterweight systems, many installations use electric winches to raise and lower the battens. These winches can be operated by sophisticated control systems which allow the batten to be lowered for service or adjustment and then returned to an exact preset trim or level. The motor will lift any load within its designed capacity.

Many variations on the motorized batten system have been developed. Often the battens have been made shorter in order to increase flexibility . This increases the number of motors required and the cost.

Modular Grids—Self Hoisting

To take advantage of the convenience and power of electric winches without excessive cost, a number of modular grid systems have been installed. These modules are essentially small pipe grid sections, which can individually be raised and lowered. They offer a flexibility similar to short battens, but since they cover a larger area, fewer motors are required. Because these modules are designed to be self-hoisting, the structural requirements of the studio are greatly simplified, and the grid remains clear for motors and steel hoisting cables.

In studios which have either counterweighted or motorized battens, a full walk-over grid should cover the entire studio. This grid usually is made of steel grating or channels, as in a theatre stagehouse, and provides support for the adjustable rigging. This walk-over grid provides easy access for any overhead suspension task a production may require.

In developing a fully integrated grid system, it is essential to coordinate all the building's structural, electrical, and mechanical systems in relationship to each other and the grid. Whereas in normal construction, many of the mechanical elements in a ceiling are placed where convenient, it cannot be over-emphasized that improperly planned or installed ductwork and conduit runs can interfere with installation and operation of the grid system, and put an added strain on talent and crew by reducing the effectiveness of the cooling system.

CYCLORAMA

Designing the cyclorama is integral in planning the studio lighting and grid systems. The grid or suspension system must provide a mounting position for the eye lights located at the proper relationship to the cyc. Also, the total area of the cyclorama will affect the calculation of the studio power service.

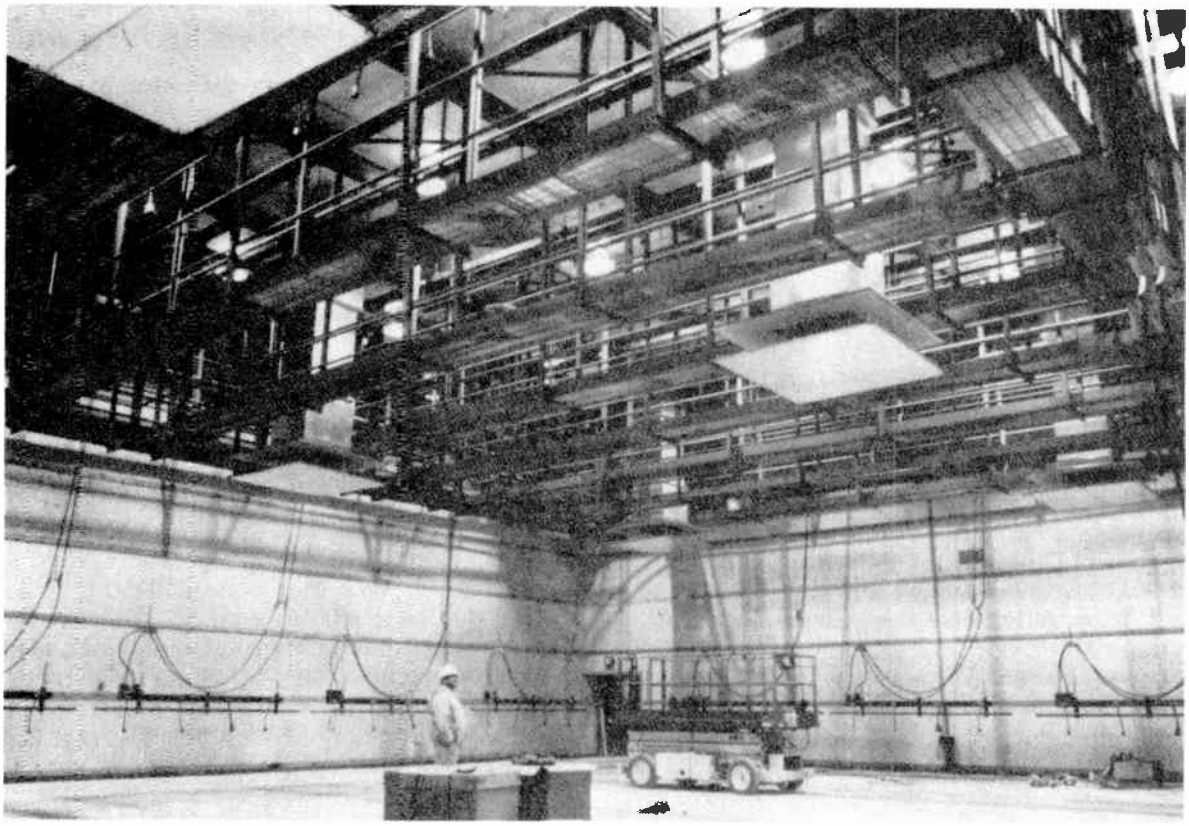


Figure 14. In a catwalk studio, lighting units are positioned in the open bays between catwalks. The motorized battens around the perimeter allow a higher backlight position where ceiling height is limited.

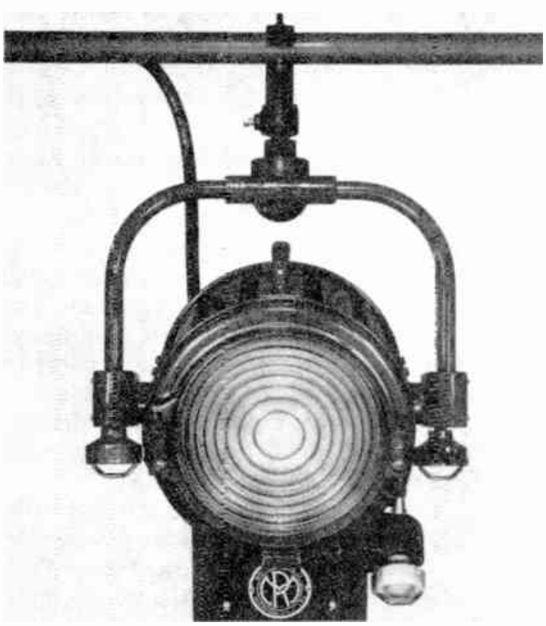


Figure 15. By inserting a long crank into the adjustment cups, this pole operated fixture can be panned, tilted, or spot/flood focused from the floor. (Courtesy of Mole-Richardson)

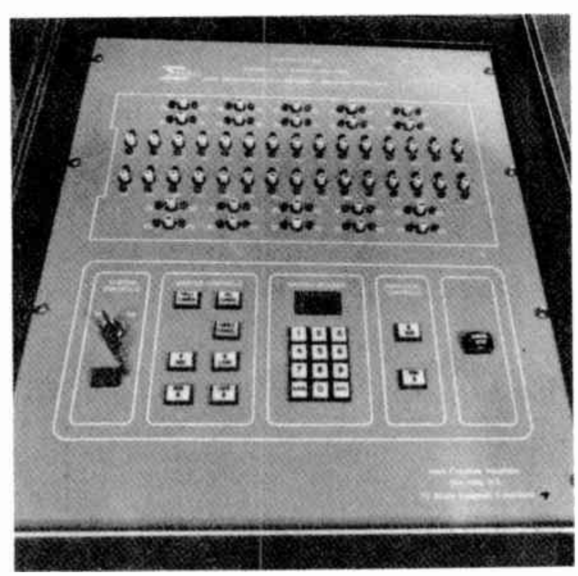


Figure 16. Computerized control panel for the motorized batten system at WNET. The panel includes an illuminated display which mimics the actual batten layout. (Courtesy of Peter Albrecht Corporation)

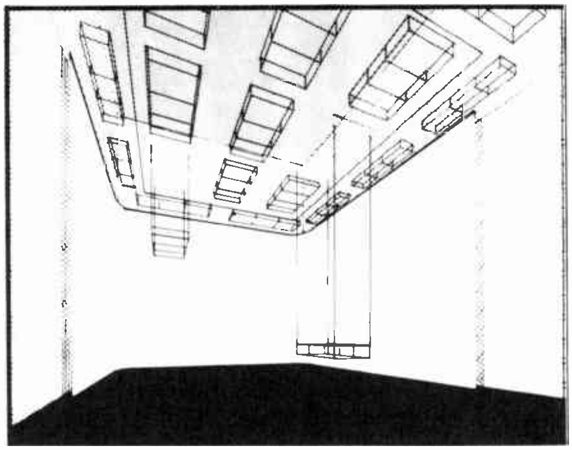


Figure 17. Modular grid plan.

Hard Cyc

Small studios often incorporate "hard" cycloramas: smooth plastered surfaces which actually blend flush into the floor. These provide the ideal infinity effect which draperies cannot equally simulate. Hard cycloramas can be easily painted any color as needed and are especially effective for certain chroma-key techniques that can be spoiled by seemingly-invisible seams. Hard cycloramas are generally limited to small studios since their hard surface area invariably creates acoustical problems.

Cyc Pit

In very large studios, a cyc pit is a compromise solution for creating this infinity effect. The pit contains and conceals the bottom cyclorama lighting, which is essential for a tall cyc, and the bottom of the drapery. When it is shot from the proper angle, the pit will simulate a background without a horizon.

For most studios, a drapery cyclorama is the most convenient solution. This seamless drapery is hung on carriers which roll along a track and allow it to be positioned anywhere around the perimeter of the studio. Two parallel tracks permit another type of background to be pulled in front of the stretched cyc. Switches on the track system allow the draperies to be easily transferred to the front or rear tracks. Leno, or filled scrim, is the most common drapery material, although seamless muslin is an inexpensive alternate material. Recently, a translucent plastic rear-projection screen type material has become a popular material because it can be lit from either the front or the rear.

In large studios where experienced lighting directors are available, true white cycloramas are used to achieve greater color intensity on the cyc. However, in smaller studios, where the talent occasionally must work close to the background or where limited control equipment

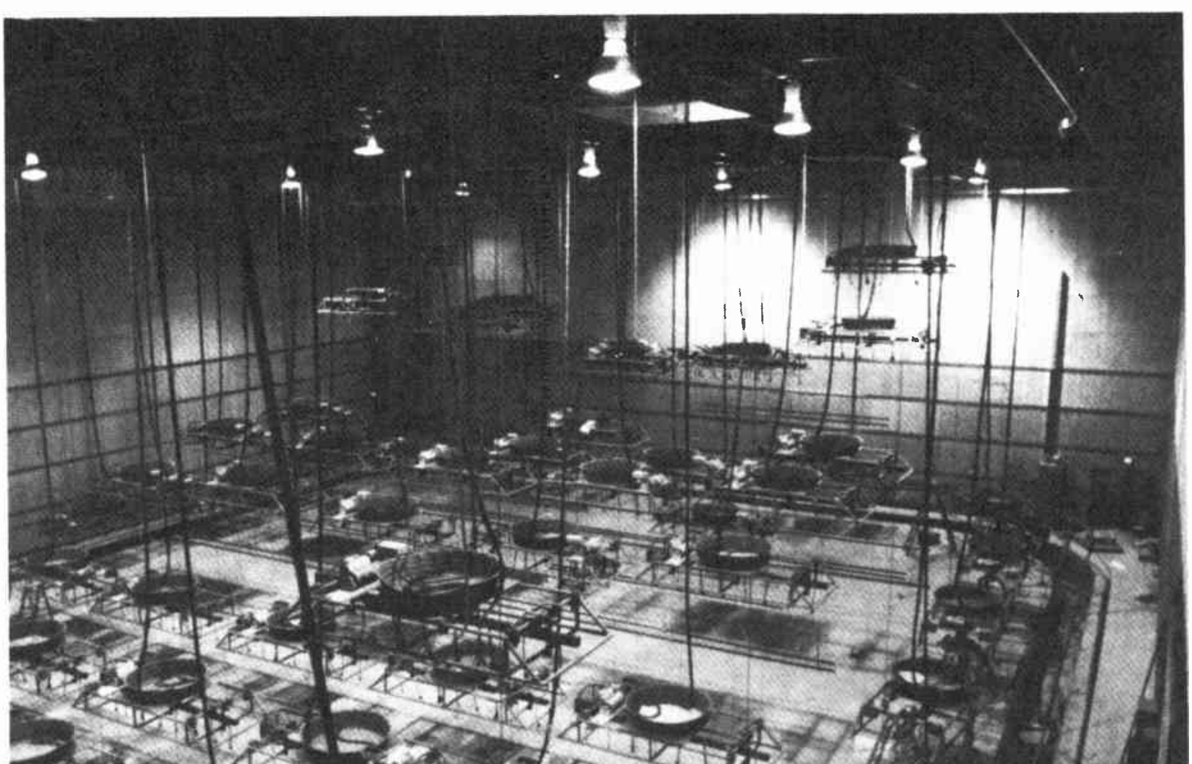


Figure 18. Modular self-hoisting grids allow variable height suspension of fixtures in a regular pattern through the studio. The round tubs serve as cable collectors for lighting and winch power. (Courtesy Dallas Communications Complex)

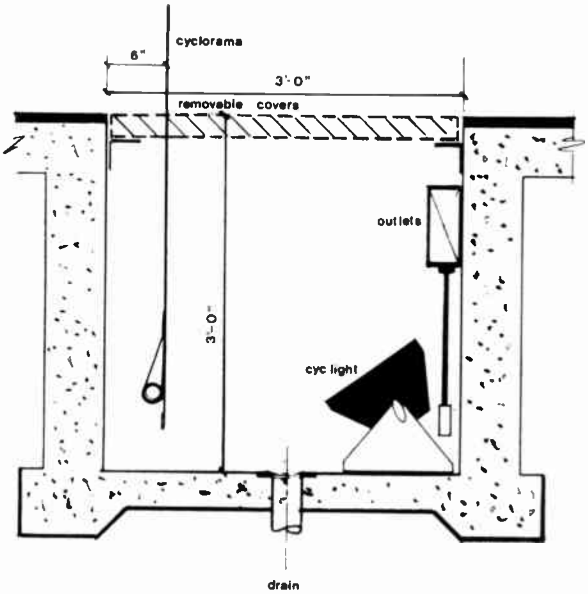


Figure 19. Cyc Pit: Section view through a typical cyclorama pit.

is available, a 60% TV white cyc should be used for better control of contrast.

Drapery cycloramas are generally furnished with jack chain weights. Removable pipe weights bent to match the shape of the track should also be provided to create a wrinkle-free background. One of the most common errors in cyclorama design is an insufficiently large radius of curvature at the corners of cyc. No matter what material the cyc is made of, a larger radius is easier to light evenly and accomplish the desired effect.

Cyclorama Lighting

Generally the arrangement of doors into the studio will be the primary determining factor for the location of the cyclorama within the studio. In the basic design of the grid or suspension system, the type of cyc lighting system should be predetermined, and the proper hanging system provided accordingly.

There are two basic types of cyclorama lighting fixtures. They are striplights and a type of widely-used fixtures commonly know as "far cycs." Striplights are continuous rows of quartz halogen or Par lamps, which for an 18 foot high cyclorama, should be mounted five to six feet from the cyc, while the far cyc units should be mounted seven to eight feet from the cyc. The entire suspension system should be designed around these dimensions. Far cycs generally will light the cyc as evenly as striplights, with less wattage. Because far cycs are mounted a greater distance from the cyc, they force the talent further away from the background. Generally, striplights should be limited to three colors; otherwise the separation between alternately colored lamps is too great to provide even coverage.

ARCHITECTURAL CRITERIA

Electrical

With the basic studio size determined and the net production area (NPA) defined by the cyclorama, it is possible to determine the power requirements for

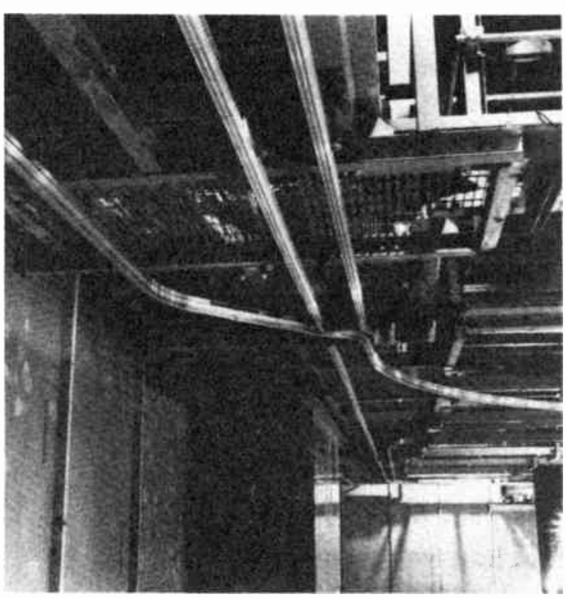


Figure 20. Cycloramas and draperies can be shifted to various track configurations by utilizing transfer switches. (CBS-NY) (Courtesy of Peter Albrecht Corporation)

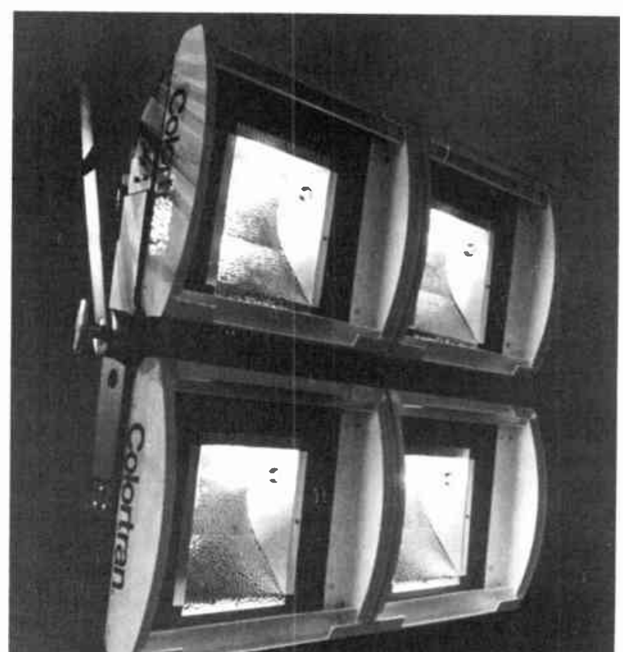


Figure 21. Four unit Far Cyc fixtures are very effective in lighting tall cycloramas. (Courtesy Colortran)

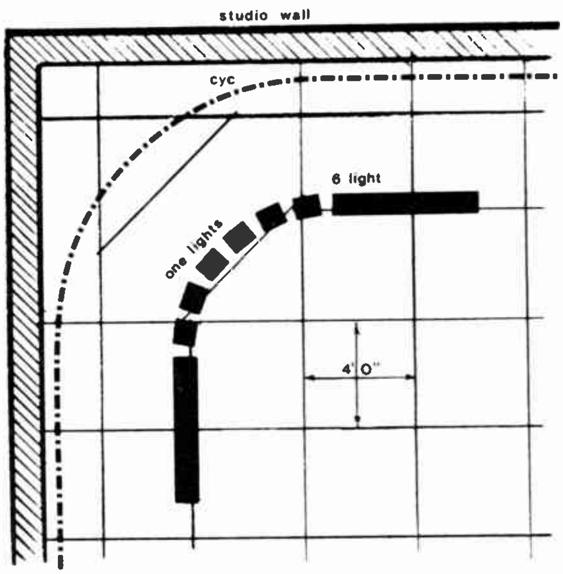


Figure 22. Cyc Light Layout: Typical arrangement of strip cyc lights relative to the cyc curve.

any given studio. The power requirements for studio lighting is a direct function of area and the required level of illumination. This power requirement remains consistent regardless of the grid height. For a lower grid, a greater quantity of smaller wattage fixtures are used, while a higher grid will require fewer fixtures of increased wattage. In either case, the total watts per square foot will remain roughly the same.

An average of 55 watts per square foot of NPA has been proven in production to provide sufficient power for any normal television lighting requirement. This method of calculating the studio load will provide sufficient power for virtually any situation. This is generally more power than is actually required for the average studio with today's state-of-the-art equipment, but it allows for over-lighting by novices and higher levels necessary for special situations. It also provides sufficient power for an average cyclorama. For very tall cycloramas which may require double rows of fixtures at the top, bottom, or both, cyc lights should be calculated as an additional power requirement based on the wattage per lineal foot of cyclorama according to the lighting system chosen.

This calculated power service describes a real maximum probable load, and the feeders must be able to supply this full amount of power. Only certain limited productions will ever require this full amount. Also, note that for dimmer-per-circuit systems, the dimming capacity will be far greater than this calculated power service. That is, the power service need not be equal to the full rated capacity for the dimmers, since the larger number of dimmers is a matter of convenience rather than power requirements.

Air Conditioning (HVAC)

Because the maximum lighting load seldom occurs, it is unnecessary to use that maximum load as the basis for the air conditioning capacity. Production practice has shown that a derating of 60% can be applied to the maximum load and still provide sufficient capacity for full-period shooting. Of course, any other heat generating devices and the population of the studio should be included in the air conditioning calculations. A properly designed HVAC system will require very large ducts to meet stringent acoustical requirements. These large ducts can often interfere with a grid system and should be closely coordinated.

Determining Studio Power and HVAC Requirements

1. Net Production Area (NPA):

NPA = Actual studio area in sq. ft. minus area behind cyc and other areas unusable for production

2. Approximate Number of Studio Lighting Outlets Required:

$$\frac{\text{NPA}}{15\text{ Sq. Ft.}} = \text{Number of outlets required (patch panel)}$$

NOTE: The actual number of outlets is determined by specific layout.

$$\frac{\text{NPA}}{18\text{ Sq. Ft.}} =$$

Number of outlets required (dimmer-per-circuit)

NOTE: Breakdown of 20A circuits and 50A circuits varies by grid height.

3. Studio Lighting Load:

NPA Sq. Ft. × 55 watts = Total watts

4. Studio Lighting Power Service:

$$\frac{\text{Total Watts}}{120\text{V}} = \text{Total watts}$$

$$\frac{\text{Total Amps}}{3} = \text{Power Service 3 phase/4 wire 120/208V}$$

NOTE: Round off to the next larger standard panel size.

5. Studio Lighting Heat Load for HVAC:

Studio lighting load (kW) × 60% (Diversity) = Heat Load for HVAC Design (kW)

6. Dimmer Room Heat Load for HVAC:

HVAC Heat Load (Item 5 above) × 5% = Dimmer Room Heat Load

ELECTRICAL DISTRIBUTION

For a studio to be flexible, lighting equipment power must be distributed uniformly throughout the studio. At the grid level, power is commonly distributed through prewired plugging strips. These strips are mounted directly to the grid and catwalks, or the strips fly in and out on battens or moveable grid sections.

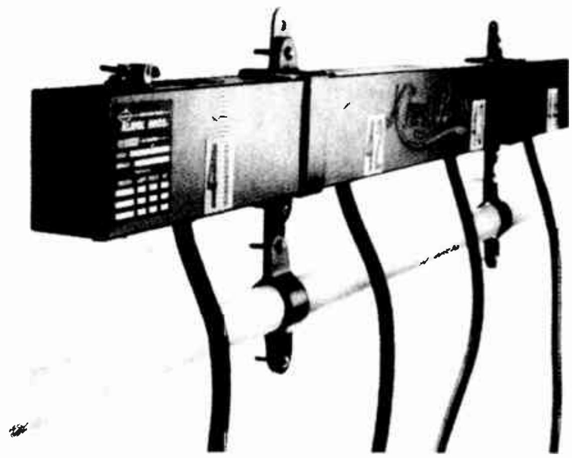


Figure 23. Grid mounted connector strips provide a convenient means of distributing large quantities of circuits throughout the studio. (Courtesy of Kliegl Brothers)

Each circuit terminates in the studio in a pigtail-outlet. There are two types of connectors in common use: stage pin connectors and twist lock. Stage pin connectors are less expensive and more common in rental equipment. If you envision renting additional fixtures on occasion, this may be an important consideration. In addition, the cost savings of stage pin connectors recur with fixture and cable purchase. Twist lock connectors have a positive locking feature.

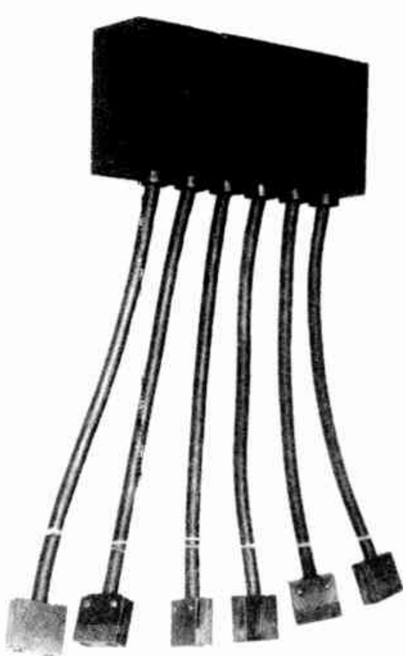


Figure 24. Wall mounted outlet boxes are generally located around the perimeter of the studio. (Courtesy of Kliegl Brothers)

The final choice should be based on specific requirements.

The number of circuits and their capacities is also related to studio size. The actual number of circuits is based initially on the square footage (approximately one outlet for every 18 sq. ft. of NPA) and then adjusted as necessary to conform to the particular grid system and the cyclorama layout. Dedicated circuits for the cyclorama are often overlooked. The cyc lights consume a large number of circuits, and once they are hung in place they will seldom, if ever, be moved. Because of the rather wide spacing of the cyc units, it is sometimes more efficient to feed these lights from individual grid-mounted junction boxes rather than a plugging strip. Also, when using far cyc lights, it is often convenient to double up the outlets to take full advantage of the 20A capacity of the dimmers.

Additional circuits should be located around the perimeter of the studio at grid level. This is a natural back light position for a set facing away from the wall. Properly located circuits save considerable time in running jumper cables.

Mounted outlet boxes should be provided at 30 inches above floor level, around the perimeter of the studio wall. Generally, the governing factor for their placement is the layout of floor mounted cyc strips. In addition, these outlets are used for miscellaneous lights on floor stands and practical fixtures on the set.

In every studio, the majority of circuits will be rated at 20 amperes, with a smaller quantity of 50 ampere circuits. As studio size and grid height increase, the ratio of 50 ampere circuits increases to accommodate the higher wattage fixtures needed for the longer throw distances. Fifty ampere circuits are distributed around the studio grid with a slightly higher density around the perimeter for back light. On adjustable-height grids, a full complement of both 20 and 50 ampere outlets should be provided to accommodate the proper fixtures

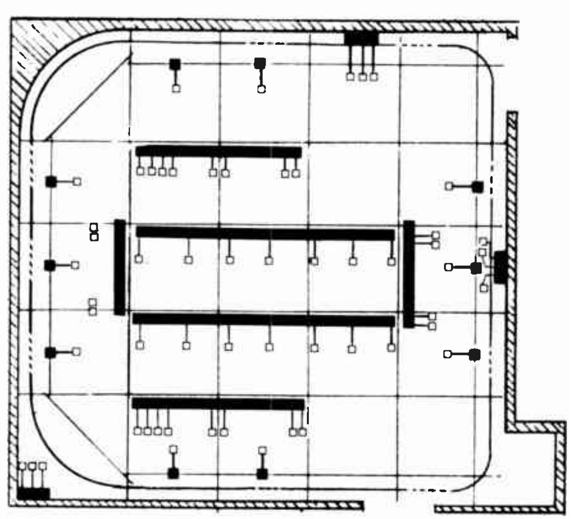


Figure 25. Typical studio distribution plan for a small industrial studio.

for the varying heights. For very high grids, over 25 feet, 100 ampere circuits are also recommended.

LIGHTING CONTROL

The importance of a dimming and control system cannot be underestimated. Dimmers allow easy control of numerous fixtures, balancing and recording of levels, and the blending of colors. A dimmer system frees the lighting director of the unnecessary burden of calculating and controlling the loads through more labor intensive mechanical methods. Electronic dimming and control allows the execution of complex lighting cues, which are a very effective production element.

Most studios today are outfitted on a basis of one dimmer-per-circuit. In this type of system every circuit terminates in its own dimmer, with the integral circuit breaker protecting both the dimmer and the circuit. The dimmer-per-circuit system gives the lighting director individual control of each individual fixture, or group of fixtures, plugged into that circuit. Normally, the studio is outfitted with 20A and 50A dimmers.

Dimmer Bank

The individual dimmer units generally plug into electronic equipment racks to form the dimmer bank. This modular system also allows quick plug-in substitution of faulty dimmer modules. Depending on the manufacturer, in excess of 96 individual 2.4 kW dimmer modules can fit into a single rack. All modern studio quality electronic dimmers utilize SCRs (silicon controlled rectifiers). These units are very reliable and are universally available in 2.4 kW, 6.0 kW, and 12.0 kW (20A, 50A, and 100A @ 120V) ratings. Only dimmers which have sufficient filtering to prevent unwanted RF interference and excessive filament vibration should be considered for use in television studios.

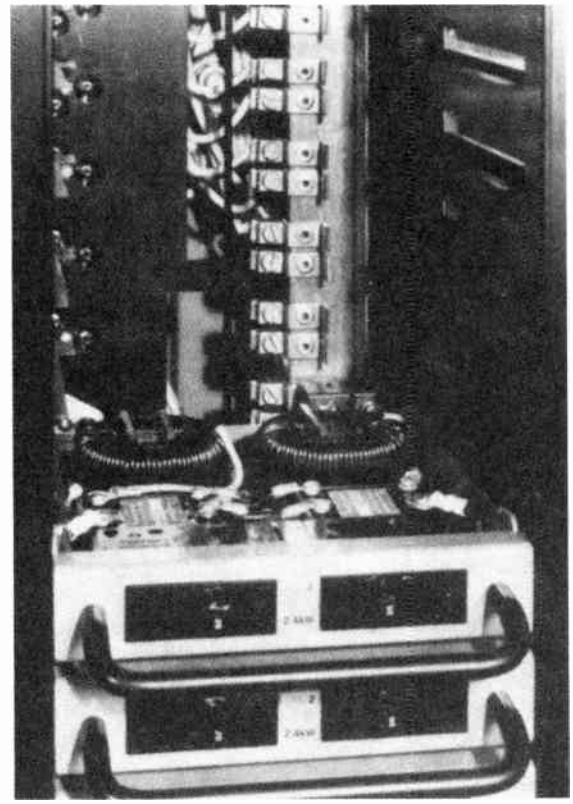


Figure 27. Tray mounted dual dimmer modules allow high density and quick replacement. (Courtesy of Colortran)

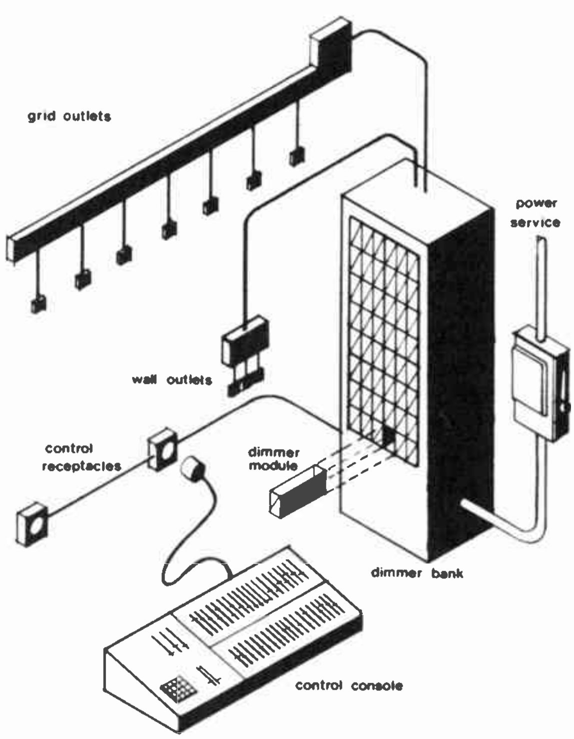


Figure 26. Dimmer system riser diagram: indicating the basic components of the dimmer system.

The location of the dimmer bank is an essential part of the initial studio space planning. The dimmer room should be centrally located to minimize the length of all the circuit homeruns which avoids voltage drop and excessive installation cost. This room should be sized to allow sufficient space for required conduit radii, access to the feeder lugs, and adequate front clearance as specified by the local code.

Virtually all SCR dimmers are approximately 5% inefficient. Therefore, they could create heat equal to 5% of the energized lighting load. Since the maximum lighting load is an infrequent occurrence, the dimmer room cooling should more reasonably be based on 5% of the diversified load on which the studio HVAC is based.

Control Consoles

In manual electronic dimming systems, an individual potentiometer or slider is required to operate each

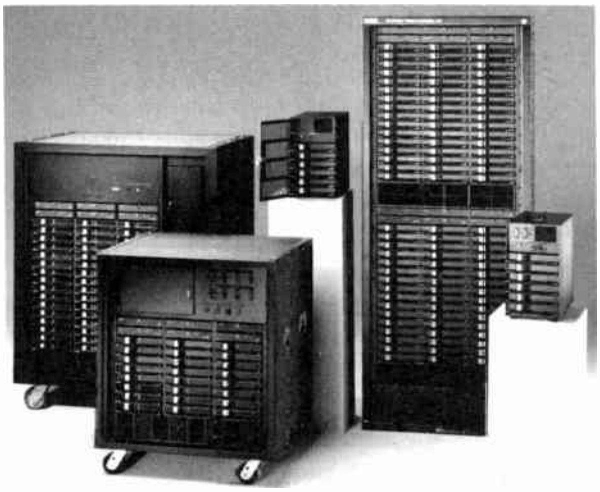


Figure 28. Dimmer systems are packaged in a variety of rack configurations to match the needs of every size studio. (Courtesy of ETC/LMI)

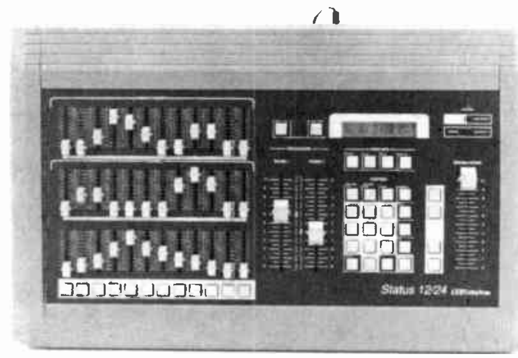


Figure 29. For small studios, very compact consoles offer memory presets and control patching via the keypad. (Courtesy of Colortran)

dimmer. To retain a particular dimmer setting, duplicate sets of sliders are necessary. Each preset group of these sliders is called a scene. While two scenes are sufficient for many productions, even a relatively simple production could require many more. The physical size of multi-scene control panels is cumbersome, and requires extensive paper work to record the setting for each preset.

Sophisticated control is possible through patching dimmers/circuits to a reasonable number of control channels. The simplest systems physically resemble the standard manual two scene preset system. On these systems, each slider represents a channel rather than a hardwired individual dimmer. The specific number of control channels is a matter of convenience. Obviously, more channels afford greater control within a single scene.

In a dimmer-per-circuit system it is not unusual for a medium size studio to have in excess of 200 dimmers. To control this large quantity with only 24 channels, it is necessary to patch the dimmers to channels utilizing only the dimmers required. The patching function allows any dimmer to be controlled by any channel. For example, the blue cyc lights load may require 12 separate dimmers. By patching them into the same channel, they will operate together in perfect unison. All this patching occurs within the console and does not require any cord, plugs, or diode pins. For small to medium size studio consoles with over 100 miniature potentiometers in two or three presets are now available to image channel patching and provide the operational convenience of manual systems.

After patching the dimmers to channels and setting the desired levels, the entire preset can be stored in a memory by assigning a memory number. This preset can then be recalled by keying the appropriate number or operating the slider to which the preset has been assigned. This type of system is very economical and is suitable for most small and medium studio situations.

For studios which encounter more complex production requirements, there are a number of systems which resemble a personal computer with specialized keypad and controls with much greater capabilities. One type of system which is common to several manufacturers utilizes a single CRT to display various functions. In operation, the screen displays the channel numbers and below each channel is a two digit number for the intensity. At the bottom of this field of numbers is the "cue sheet" which displays various operational functions such as cue numbers, fade times, etc.

The control panel on this type of computer control generally has a series of submaster sliders, X, Y faders, numerical keypad and a function keypad, and some tactile type of encoder for easily altering fade times and intensities.

The most complex systems employ dual CRT displays to actively display more information. These systems offer so many esoteric features that most studios would never begin to tax their full potential.

Computerized lighting systems store their memory on either floppy disk or data cassettes for reuse. This permits a complete copy of all the settings in memory. In addition, most larger systems are available with

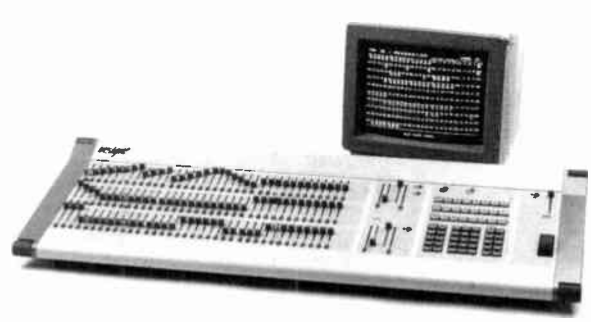


Figure 30. For medium size studios, memory consoles with up to 108 manual channels offer both manual and memory control. (Courtesy of ETC/LMI)

Figure 31. The most sophisticated consoles utilize dual CRT displays and can control very large quantities of dimmers with equal number of discrete channels.

varying degrees of internal backup memory systems. The operational software is permanently stored in ROM (read-only memory) within the machine.

Most computer lighting systems were originally designed for the legitimate theater, where the easy daily repetition of very complex multipart cues is the primary goal. Some manufacturers have modified their programming to be more sympathetic to television's somewhat unpredictable demands. In selecting a system with this caliber of sophistication, one must carefully determine which features are really necessary.

Computerized lighting systems have greatly simplified the installation of the control wiring. While manual systems require at least one wire for every dimmer, these computerized systems can be run over a single coax or a few twisted pairs for the entire system. An advantage to this simplified cabling is that the console can be easily relocated to any of the plug-in stations. In addition, when the main computer is located outside the studio, a small focus remote the size of a handheld calculator can be used in the studio to activate fixtures as necessary for focusing or other simple operations.

This allows the console to be located in the most convenient position for a particular phase of the production. It also allows several studios to share a more sophisticated system as required, since the largest systems can plug into the same control wiring as the smallest.

Most television production people have the opportunity to be involved in planning a new studio maybe once or twice in their entire careers. Even years of experience in studio production are not necessarily the best preparation for coordinating studio requirements into a construction process. Often, a new studio provides an opportunity to acquire a complement of equipment, which is more sophisticated than the existing staff's level of experience. In this situation, an experienced lighting director should be consulted in preparation of the equipment requirements and evaluating the quantities required. While manufacturers are sometimes helpful in this area, they are still primarily interested in selling you their product, and no single manufacturer offers a full line of suitable equipment in every area. By working with a consultant, you can share their experience in the planning and design of the lighting system for your studio.

6.1 Radio Electronic News Gathering and Field Production

Jerry Whitaker
Beaverton, Oregon
and
Skip Pizzi
Broadcast Engineering Magazine, Overland Park, Kansas

Radio stations have used the remote location broadcast for decades to bring the listener an added sense of realism and excitement. Although the concept of the "remote," as it is better known, has not changed substantially over the years, the means to accomplish the task has moved quantum leaps in terms of performance, ease of operation, and reliability. Radio electronic news gathering (RENG) systems of today can be configured to provide virtually any degree of sophistication required by the station. As with any other area of broadcasting, the key to a successful RENG system is thoughtful planning.

PLANNING THE RENG NETWORK

The importance of careful planning of a RENG system cannot be overemphasized. The network should be configured based on the precise needs of the station. All persons that will be involved in the use of the system should be consulted to determine just what type of arrangement will be needed. Whether a station's format is "all news" or "all hits," everyone in the news, production, and engineering departments should sit down and define the requirements of the network. At such gatherings, engineers should resist the urge to be negative when someone asks for a level of performance that is not practical. Even though the engineers may know that it is impossible to provide every reporter on the staff with a separate frequency that can be received at the studio from anywhere in town, at least listen to what the users would like the system to do. The realities of station economics and the laws of physics can be explained after the desires of the participants have been outlined.

Many, perhaps most, RENG systems were built on a piecemeal basis, as needs dictated and economics allowed. The lack of a unified plan has often led to RENG systems that are cumbersome to operate and, in the long run, more expensive to build than necessary for a given level of performance.

The size and layout of the station's market will have a substantial effect on how the RENG network is designed. A system intended to cover a sprawling urban area of 10,000 square miles will be configured much differently than a tightly clustered urban center covering 2,000 square miles. The number of stations in the market that are involved in RENG activity may also affect how a system is designed, and what types of equipment are used. Stations in major metropolitan areas may find that few, if any, frequencies are available for RENG activity.

Program material can be returned from the field to the studio through either of two common routes: wired telephone lines or wireless transmission systems of various types. The route back to the studio will depend upon a number of factors, including the location of the event, the availability of telephone lines, the amount of setup time provided and the duration of the broadcast.

WIRED VS. WIRELESS

Until the 1960s, the word "remote" was rarely spoken without reference in the same sentence to "the telephone company." Wired systems, either using the dial-up network or leased broadcast loops, provided the vast majority of interconnections from remote broadcast sites to a station's studio facilities. Since that time, however, RF systems have assumed an important role in remote activities because they inherently offer greater flexibility and generally provide the user with a higher-quality audio link. Given sufficiently frequent use over time, they may also be more cost-effective than leased program circuits from the telephone company.

Radio systems are ideally suited for broadcasts of relatively short duration and from several different locations during a short span of time. Meetings, speeches and sporting events, however, are probably best handled by a wired arrangement. The amount of frequency congestion in the origination area will also have an effect on which method a station will choose for the greatest reliability. Urban areas in which secure remote pickup unit (RPU) channels are difficult to find may be best suited for wired links.

The amount of lead-time provided before an event to be covered will also have a significant effect on the route chosen for getting the program audio from the remote site back to the studio. Broadcasts that are scheduled weeks in advance are obvious candidates for use of a telephone company loop. Spot news events, on the other hand, do not lend themselves to planning in advance. For such applications, then, a radio system is extremely useful.

The cost of telco facilities must also be considered. Unless the loop is to be left in place for a long period of time, installation charges can become prohibitive, especially if high-performance equalized lines are needed. Many stations are able to justify the cost of a RENG RF system based solely on the telco savings that can be anticipated.

The best approach, therefore, implements both wired links and a wireless system. Large-scale systems are often built using both interconnection methods, either as various links in the chain or for back-up protection, in the event of a partial system failure.

Digital telco services have recently become commonplace in most telco operations. Among these, the DS1 service (also referred to as "T-1") can be useful for these applications, especially when multiple and/or bidirectional circuits are required between remote site and station. The DS1 format is a bidirectional digital service with a data transmission rate of 1.544 megabits per second (Mb/s). It is primarily used to carry 24 voice-grade channels each with 3.5 kHz bandwidth, but can be easily reconfigured to virtually any standard telco line format for leased circuits by means of plug-in cards on terminal racks at each end, up to its total data capacity. For example, a DS1 link could be configured to carry several 15 kHz stereo program circuits bidirectionally for transmission from, and monitoring to, the remote site, bidirectional 5 kHz communication paths, bidirectional FAX lines, voice circuits, and data control lines. In some cases, a half-DS1 circuit or some other partial-carrier arrangement may be obtained, when capacity needs are not so great. Not available in all areas, this service is called *fractional T-1*. In most cases, for local loops, the DS1 digital service will be significantly more cost-effective than its multiple analog equivalents. Depending on the encoding equipment, 15 kHz channels may not provide "CD-quality" audio, although very good audio quality will result. For absolutely critical uses, such as in an STL or a music remote feed, several manufacturers provide black-box digital processors that offer superb audio quality on DS1 paths. Earlier devices of this ilk require the entire DS1 bandwidth for a single 15 kHz stereo signal and may not offer bidirectional capability. Later and more sophisticated codecs implement digital compression algorithms to allow multiple, high-quality 15 or 20 kHz stereo pairs, plus auxiliary circuits, in bidirectional fashion. (For more on these services, see Chapter 6.4, "Telco Audio Program Service.")

WIRED SYSTEMS

Wired communication systems for news, sports, or programming can take a number of different forms: from basic telco equalized loops to sophisticated single or multi-line frequency extension systems using the dial-up network, to newer digital telco services. An equalized line offers the user a simple, reliable link to the studio. The drawbacks include inflexibility, installation lead-time, and installation or rental costs. These lines also operate only in one direction. Leased circuits of a digital variety (DS1 or T-1) also require lead-time, and carry some moderate to high installation costs; although both of these parameters are now generally less than analog equalized loops. Cost of service is also typically less than equalized loops, especially since these lines are bidirectional. And DS1 reliability and audio quality are high. But the dial-up network gives today's broadcaster the greatest degree of flexibility in terms of when and where, including bidirectional ability; however, dial-up audio quality leaves a great deal to be desired. The most popular way around this problem is through the use of frequency extender systems, as illustrated in Fig. 1.

There are a variety of extension methods, each with a different way of accomplishing the task. Generally speaking, however, audio from the remote source is heterodyned up or down to frequencies that will pass through dial-up circuits (300 Hz to 3 kHz), then detuned back to normal at the receive end. The one-decade response of the phone line can thereby be made to pass not more, but a different, set of frequencies than it would normally carry. The intent is to make the limited response not sound that way to the human listener.

The simplest of these shifts the effective passband of the phone line down by 250 Hz, making it approximately 50 Hz to 2750 Hz. This has the effect of adding over two octaves to the low-end cutoff of the typical phone line, while only sacrificing a small portion (about ⅙) of an octave on the high end. This is a result of the heterodyning's frequency-linear action, versus the logarithmic (octave-based) spectral sensitivity of human hearing.

More sophisticated shifting systems add a second or even a third phone line to the process, to take higher frequency components of the remote audio and shift them DOWN to fit the phone line's passband, reversing the process on the receive end. For example, in a two-line unit, remote audio is split into two frequency bands by a filtering network. The higher frequency components are shifted lower by one conversion circuit and fed into one phone line, while the lower frequency components are shifted upward by a second conversion

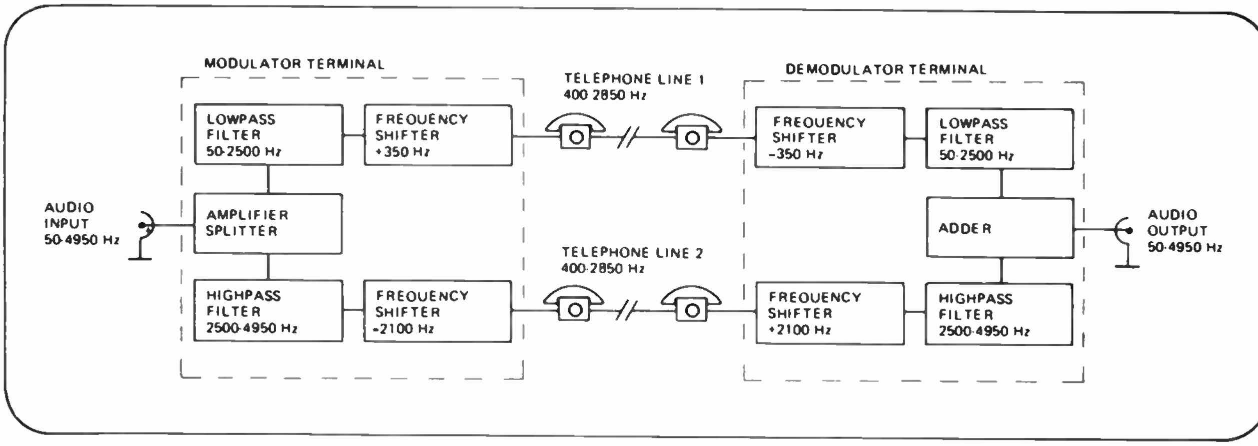


Figure 1. The basic two-line frequency extension process for dial-up telephone links. (Courtesy of C. N. Rood.)

circuit, and applied to a second telephone line. At the studio demodulator end, the two signals are frequency-shifted back to their original values, filtered and recombined to form the output of the system. These multiline systems usually include a multiband compandor circuit for reducing phone line noise.

An interesting variant using a single-line encoder/decoder, called half-speed transmission, is shown in Fig. 2. Audio from a news or sports story is recorded at the higher recording speed on a two-speed analog tape recorder. When the material is fed to the station, the low speed is used for playback and the recorder output is fed through a single-line bandwidth extension system. At the studio end, the feed is recorded on the low speed of another two-speed machine. When the program material is played back on high speed at the studio, an effective doubling of frequency response (a one octave increase) at the high end is achieved with a concurrent loss of one octave on the low end. The low end loss is of little consequence, usually, since the frequency extension adds over two octaves, which is more than enough in most cases. With this arrangement, it is thus possible to squeeze a signal bandwidth of 100 Hz to 5 kHz through a single dial-up telco line. Feed time takes twice as long, of course, and cannot be done live.

Setting Up a Wired Remote

If a wired remote is chosen for a particular broadcast, the station has a number of options in terms of equipment and configurations. The decision on how to originate the remote location programming will depend upon the requirements of the particular broadcast. Some generalizations can be made, however, that apply to most events.

A program transmitted back to the studio via a standard dial-up telephone company line without any bandwidth extension will usually be brief in duration, if for no other reason than the poor audio quality typical of such an arrangement. Spot news reports are common examples of this method of program return.

Small, battery-powered mic-to-line amplifiers are available to drive dial-up telephones through direct connection to the tip and ring wires of the phone company cable or through clip leads at the handset microphone pins. The direct connection method of coupling is preferred over the handset connection, because the former bypasses the telephone hybrid coil assembly with its associated level loss and possible distortion. But for this direct connection method, the device feeding the phone line (known as a "coupler") should have the ability to "seize" the line (meaning to hold it open, in an off-hook condition), which the

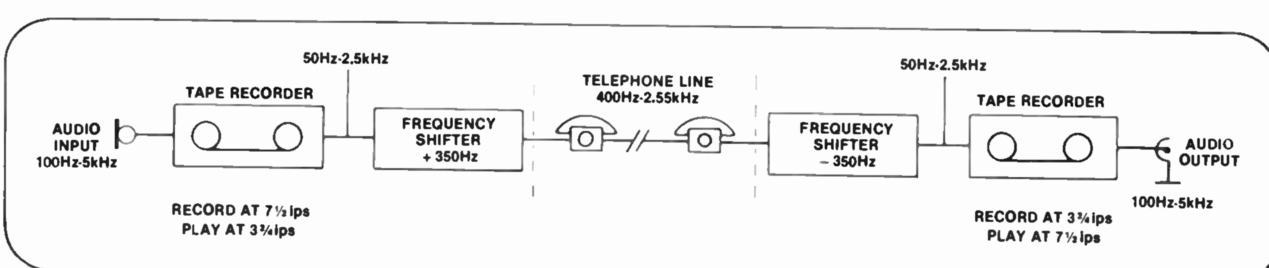


Figure 2. A frequency doubling technique using two-speed tape recorders and a single-line bandwidth extension system. (Courtesy of C. N. Rood.)

telephone instrument itself does when using the handset-type connection. With such a seizing coupler, a call can be established with the telephone instrument, then the coupler can be switched on, and the handset of the telephone hung up, with the coupler holding the line open by itself. Many couplers (based on the original Western Electric "QKT" style) will not hold the line without an instrument off-hook on the same line. The advantage of the seizing coupler is twofold. First, the handset microphone will not contribute its output to the phone feed, so the signal transmitted will be only what comes from the amplifier or mixer plugged into the coupler's input. Secondly, the transmitted level of the audio signal will be higher without the instrument off-hook. This will maximize the signal-to-noise ratio (SNR) of the feed.

Output level of the device feeding a phone line (whether by coupler or handset) should be carefully controlled, since the dynamic range of the dial-up network is somewhat limited. Transmit level should not exceed +8 dBm for program material, or 0 dBm for test tones. Be aware that the received-end noise floor may be only 25 dB to 30 dB below this on local calls, and considerably less on some international long-distance calls.

A program sent to the studio over dial-up telco lines using bandwidth extension equipment can provide impressive audio quality. Reasonably flat frequency response from 50 Hz to 5 kHz is possible in real-time using a two-line system, or at half-speed with a single-line system; the newer three-line systems can pass 7.5 kHz audio in real-time. The multiband compandor systems used by the latter are required to reduce the dial-up network noise, which exhibits a rising characteristic with frequency and also becomes more subjectively noticeable as the high-frequency cutoff of the system is raised.

Two- and three-line systems can also be operated at half-speed, with the expected effective doubling of their high-frequency passband (one octave added), and a commensurate loss of an octave on the low end. This is usually a worthwhile tradeoff in terms of resulting audio quality. Thus the dial-up network can actually pass 15 kHz audio, on three lines, at half-speed. For more information on frequency shifting systems, and the use of the dial-up network in general, see Chapter 6.3, "Telephone Network Interfacing."

The equalized broadcast loop is probably the most popular method of relaying lengthy remote programming to the studio. Some stations prefer to order unequalized lines and adjust the loop themselves for the required frequency response. This procedure can be effective on relatively short telco lines. Generally, however, any audio loop that goes through more than one exchange should be left to the phone company, which has had many years of experience in making miles of twisted pair cable perform properly.

If a station decides to equalize the line and not rely on telco, equalization should be applied only at the studio (receiving) end. Applying frequency-selective boost to a telephone line input can raise components of the signal to a point that will cause crosstalk into other lines or clipped audio due to the action of network protection devices.

The receiving equipment for a wired remote broadcast should be given careful consideration. Sophisticated telephone interface equipment is available today from a number of manufacturers for dial-up applications. This hardware can ensure that maximum audio fidelity is recovered from the line. Some new generation interface equipment includes automatic gain control circuits, equalization, and dynamic noise reduction systems. Alternatively, a station may add such hardware to the audio path downstream from the interface. (See also Chapters 6.3 and 6.4.)

If DS1 (T-1) circuits are employed, a relatively large rack of encoding hardware will be required at the remote site. This may either be leased from the telephone company or purchased from a number of manufacturers. The number and bandwidth of audio lines available is adjustable, but any configuration will require its own complement of encoding cards for the rack, and decoding cards in the station's receive rack. Service may be ordered by the day or month, and because the service is digital, little "tweaking" is required by telco personnel, making set-up time and installation costs lower than equalized program loops. When using these circuits (or analog loops), dial-up backup service in hot stand-by mode may be advisable, bidirectionally. For more on digital services, see "Emerging Technologies" below, and Chapter 6.4, "Telco Audio Program Service."

Remote Site Equipment

The audio equipment used at the site of the remote will vary as widely as the types of programs broadcast. Fig. 3 illustrates a typical application for either a wired or wireless relay system.

In this example, a four-channel audio mixer is used to mix the sources and drive the telco loop, phone, or RPU transmitter. Careful attention should be given to the connection of the mixer output to the telephone line. A phone coupler should be used between the mixer and the telephone, unless the mixer is specifically designed to work directly into a hot dial-up line (one with DC voltage across it). This caution applies to a connection made either to the phone line tip and ring wire, or to the telephone set through the handset terminals.

As shown in Fig. 3, two microphones are used: one for the announcer and another for interview guests. An output is taken from the local PA system to pick up audio from meetings, speeches, music or whatever. For more complex and high-budget music remotes, a separate mixer or audio truck is used to "premix" the stage microphones down to a stereo feed, which is then fed to the mixer in Fig. 3. But PA feeds of musical acts are not always inclusive of all the elements required, nor are they properly balanced for the radio mix. For this reason, a separate radio mix is often required if proper radiophonic balance and control of the stage event's mix is critical.[1] Fig. 4 shows a more

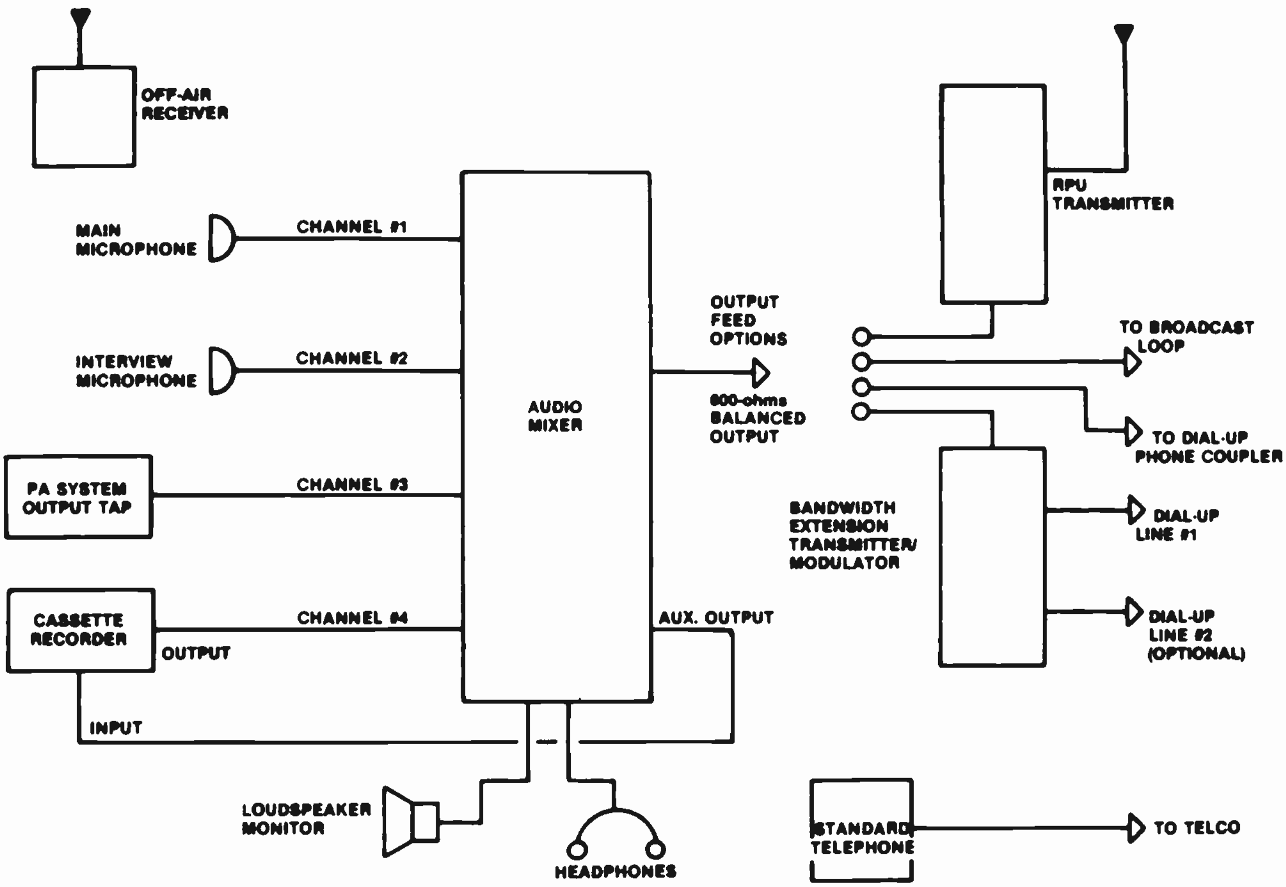


Figure 3. A typical equipment configuration for a medium scale RENG broadcast.

complex remote setup, in which two separate mix positions are used.

When multiple mixers (such as the PA and separate radio truck just mentioned) all require access to the same microphones, a number of options are available. The most primitive would be to put separate microphones up for all mixers, but this becomes extremely impractical and unsightly, especially for things like handheld vocal or announce microphones. In these cases, a shared arrangement is best, using proper microphone splitting techniques. Good transformer-isolated microphone splitters are recommended for this. One mixer remains directly connected to the microphones, while the others are fed by the secondary of a bridging input transformer. The microphones see only one load, and the consoles remain electrically isolated from one another. Typical microphone splitters also provide separate ground-lift switches for each channel (or in some cases, a single, ganged switch), by which ground loops between consoles can also be eliminated. If condenser microphones are used, their phantom power can only come from the console receiving the direct feed, since DC supply voltage from the transformer-fed consoles will not pass back across the transformers.

A cassette or DAT recorder is often useful at a remote broadcast because it gives added flexibility to the remote crew. The recorder input signal can be taken from an auxiliary output on the audio mixer, allowing interviews or material from the PA system to be mixed and recorded for later use on the air. This is always preferred over relying on a recorder at the station end, because the phone line degrades the audio quality, especially when the dial-up system is used, but also because the line may fail during the broadcast. In either case, an on-site deck provides a good-quality recording for later production and broadcast.

In cases where a separate stage mixer/truck is used for the stage premix, it is this "dry" premix that is usually recorded on site, so that these production recordings are of the actual event and do not contain any continuity or other local production elements added at the broadcast mixer. The full broadcast can be recorded for archival purposes either at the remote site or at the station.

If critical, multi-station monitoring and communication is necessary at the remote site, an interruptible foldback (IFB) system may be required. This device combines the function of intercom and monitor, such that a director, either on-site or at the station, can

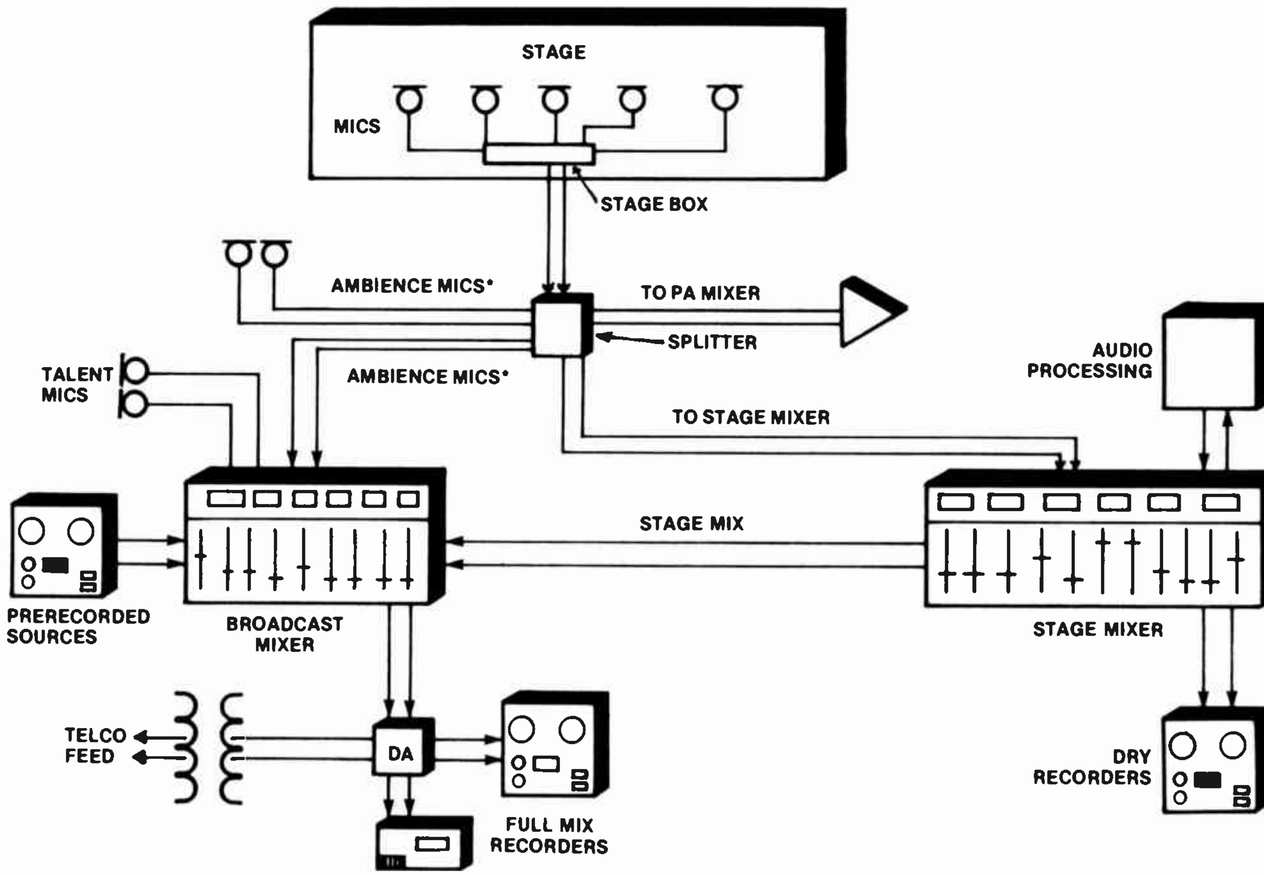


Figure 4. Typical setup for a large scale remote, using two separate mixing positions. This arrangement is ideal for events such as festivals or conventions, where several separate performances with set-changes in between will be broadcast in a long, continuous program from the site.

communicate with and cue various personnel at the remote (talent, stage managers, floor directors, engineers) via headset or loudspeaker, using a multi-station intercom. The director's control panel for the IFB system is typically equipped with push-to-talk, individual-station, and "all-call" (talk to all stations) buttons. What differentiates the IFB from a standard intercom is that when the director is not talking to a station, that station receives program audio. The director's cues temporarily interrupt the program audio, which returns after the director's message has ended, hence the term IFB. And unlike a traditional intercom, the IFB system is usually not bidirectional, but rather feeds one-way from the director out to the various receive stations. Fig. 5 shows a typical IFB layout in a complex remote situation.

Most IFB systems allow the use of two or more different program audio sources to be selectively routed to different stations, so that one station could be fed the dry stage mix, for example, while another station heard the whole remote transmission with continuity included. Both would hear the director when their channels were designated for communications, regardless of which program audio channel were selected. Off-site IFB (in which the director is not at the remote location but back at the studio) is often referred to as a "PL" (private line). Dial-up or leased telco lines are used for PLs, along with wireless return links or cellular phones, in many ENG cases.

Set-Up and Testing Procedures

A tone oscillator is useful in setting up the remote broadcast. Most mixers designed for remote applications include an oscillator that can be switched on to the program channel. This feature is especially useful when a telco line is employed to return program audio to the studio. If the oscillator has multiple frequencies, so much the better. Phone line tests for frequency response, SNR, headroom, and relative polarity for stereo broadcasts should be performed. Phase response and distortion tests are also useful, if time and test equipment permit. If any noise reduction or other

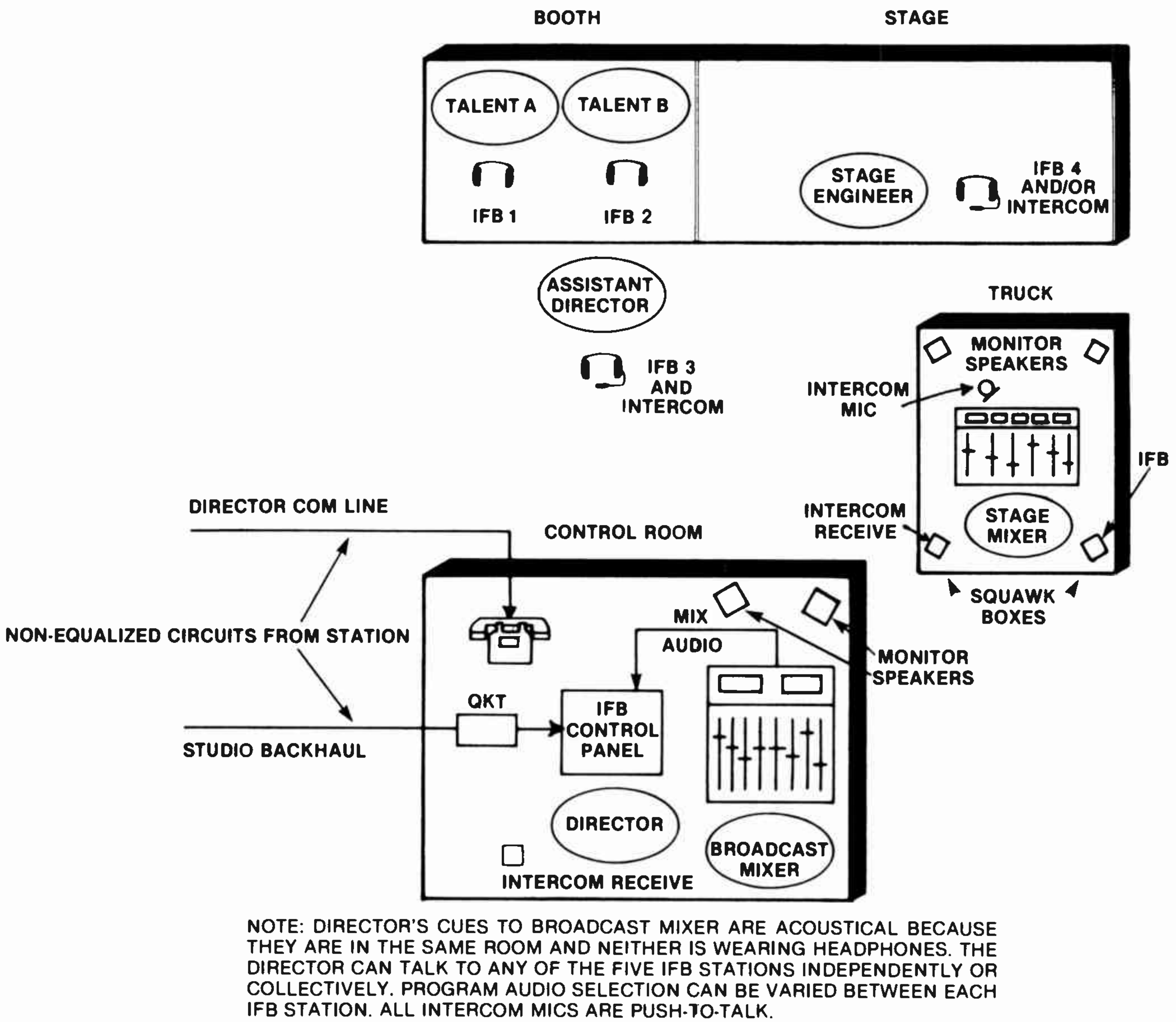


Figure 5. Communications plot for a complex live remote broadcast. Stage mixer can use simple intercom (with no program audio) for communications to stage and other personnel, while main control room uses interruptible foldback system (IFB) to coordinate entire broadcast.

enhancement devices are being used across the line, testing should be performed with them bypassed. After the line proves satisfactory, engage these devices and recheck.

Having the proper monitoring facilities is important to the success of any remote broadcast. A loudspeaker (or well-matched pair for stereo broadcasts) and set of headphones should be provided for the remote crew. It is often desirable to have several headphones available for use by personnel at the remote site. Not all portable mixers can support a loudspeaker and multiple headphone outputs, and so a separate power amplifier and headphone booster may be needed. Speakers and headphones selected should be reliable and familiar. Crew members should listen to familiar music over these while setting up, to attune themselves to the speaker(s) in this environment, however, since different rooms and speaker placements will produce drastic changes in the sound of any speaker.

After checking phone lines or wireless links for continuity back to the station, set up the monitoring system first, and check it for clean audio. Then add each other element or subsystem of the setup one at a time, checking the monitors for continued clean response after each. In this way, when any deleterious effects are heard (hum, buzz, hiss, etc.), it will be fairly simple and quick to track down the offending hardware or interface method.

An off-air receiver is a requirement for nearly all remote broadcasts. The receiver gives the remote crew a way of checking the total link and allows easy cuing of talent at the event.

A separate dedicated telephone set is suggested for complicated remote broadcasts. The phone provides an easy means of communicating with the studio. It can also serve as a backup line for program audio in case the RPU system or telco loop should fail. And in cases where the remote broadcast originates from outside the coverage area of the station, it is essential for monitoring as well as communicating. If digital telco service is employed, it is generally not much more expensive to have this return (or "backfeed") line be of the same fidelity as the transmit circuit.

WIRELESS SYSTEMS

RENG has come a long way since the narrow-band walkie-talkie days, when a 5 watt "portable" unit was the size of a briefcase and weighed 35 pounds. Today's gear is small, lightweight and can deliver excellent audio quality. Stations can now go into the field for news and special event programs and maintain studio-like sound. The audio quality of remote broadcasts is more important now than ever before because radio station transmission systems and consumer receivers are constantly improving. Further, the listening audience is becoming more discriminating and demanding of news operations. Today's competitive marketplace requires more than just sound from the field. It demands clean, quiet audio with good frequency response and low distortion. TV news has shown the public that this is possible and has conditioned consumers to expect it.

Radio channels typically used for RENG work are used on a shared basis, and so receipt of a license is no guarantee of unlimited interference-free operation. Indeed, an unused channel is the exception, not the rule, in most larger urban areas of the United States. The frequency coordination process is a complicated procedure that requires careful thought and planning, and generally a great deal of lead-time. Broadcasters in a given geographical area rarely have to decide whether they wish to become involved in frequency coordination efforts. Usually the need for coordination is painfully obvious to all persons involved in RENG activity in the region.

The main driving force behind coordination efforts has been the Society of Broadcast Engineers, which has set up a National Frequency Coordinating Committee to encourage and support local coordination efforts, and to provide whatever support might be needed in this regard. The subject is discussed in detail in Chapter 1.3, "Frequency Coordination." Spectrum congestion is a sad fact of life to many stations engaged in RENG today. Users must recognize that coordination is vital to the reliable operation of remote broadcast systems, since spectrum congestion will no doubt become worse in the future, rather than better.

Licensing Procedures

RENG work is done on two primary bands of frequencies set aside by the FCC for remote pickup unit (RPU) operation. A number of frequency groups are allocated near 150 MHz and 450 MHz. Some assignments are also made on frequencies in the 25 MHz region. A particular broadcast station is not restricted to a maximum number of RPU systems that it may put into operation. The needs of the station and the budget available for equipment purchase are, instead, the major controlling factors.

Most RENG activity is currently centered in the 150 MHz and 450 MHz bands. In these slices of spectrum, three major license classifications exist: automatic relay station (ARS), base station, and remote pickup mobile station.

ARS systems are designed to receive program material on one frequency and retransmit on another. In this way, the average area of the RENG system can be extended considerably.

Base stations are, as might be expected, fixed-position transmitters used for communication between the central point and one or more remote points. Base stations may, in the event of emergency conditions, be used as a program relay channel for Emergency Broadcast System information.

Remote pickup mobile stations consist of vehicle-mounted and portable (hand-carried) transmitters. They are usually licensed as a system in conjunction with a principal base station, or stations. Remote pickup mobile station licenses generally specify a minimum and maximum number of mobile transmitters allowed in the RPU system. Standard divisions include from one to four stations, from four to 12 stations, from ten to 20 stations and from 20 to 50 stations.

The Commission's Rules require that the transmitter power for a RPU station be limited to a level necessary for satisfactory coverage of the service area. In any event, not more than 100 watts of transmitter power output will be licensed. RPU transmitting equipment operating onboard an aircraft is normally limited to a maximum transmitter power of 15 W. A mobile station consisting of a hand-carried or pack-carried transmitter is restricted to not more than 2.5 W power output.

All RPU transmitting equipment must be type-accepted by the Commission and checked each year (for units with more than 3 W output) for frequency accuracy, deviation and RF power output. FCC Rules also require that RPU transmitters rated for 3 W or greater must be equipped with a circuit that will automatically prevent modulation in excess of the authorized limits.

There are virtually no operator requirements for the use of a unit in the RPU service. Any person designated by and under the control of the licensee of the station may operate the equipment. An operator's license, as detailed in Part 13 of the Commission's Rules, is not required.

Building an RPU System

In view of the serious spectrum congestion problems that exist today in many areas of the country, any RPU/RENG system should be designed to be as spectrum-efficient as possible and, equally important, to be as immune to undesired transmissions as possible. Even if the system will be operated in an area that currently does not have a spectrum congestion problem, there is no guarantee that such a problem will not surface in the near future. In any event, a well-engineered system is also a spectrum-efficient system.

The first rule of spectrum-efficiency is to use only the effective radiated power (ERP) necessary to do the job. There is no justification for putting 15 W into the air when 5 W will provide the desired (or acceptable) signal-to-noise figure from the receiver. Ideally, all transmitters in a RENG system would therefore be equipped with continuously-variable power output stages. The operator at the remote site would then run the transmitter with only enough power output to reach the required SNR figure at the receive (studio) point. With some types of units, this method of operation is possible, but in the majority of cases, continuously-variable power output transmitters are not available. User-modification of existing equipment is not an acceptable solution, since such work would most likely invalidate the transmitter's FCC type-acceptance.

A more practical solution, therefore, is to purchase RENG transmitters of several different power levels operating on the same frequency (or frequencies). All of the popular RENG broadcast equipment manufacturers offer units with different power output levels. With some equipment, a low-power transmitter is used and an optional power amplifier module is added between the transmitter and the antenna to give the needed RF output.

Directional receive and transmit antennas are a good idea from both an efficiency and coordination standpoint. The use of a pair of high gain antennas makes it possible to achieve a much greater ERP for the same transmitter power. Of equal benefit in a crowded urban area is the elimination of any nonessential radiation. Through the use of directional transmit and receive antennas, stations can establish more secure channels by placing the radiated energy where it will do the most good (from the transmit end), and rejecting unwanted signals from other directions (at the receive end).

A simple and sometimes effective coordination tool is cross-polarization. Two stations on adjacent frequencies may achieve as much as 25 dB RF isolation through the use of different polarizations of transmit antennas, matched by like polarization at their respective receive antennas. Cross-polarization results in varying degrees of success, depending upon the frequency of operation and the surrounding terrain. Line-of-sight paths usually will provide good results, but urban centers with their highly-reflective buildings, generally cause polarity shifts in the transmitted signal that may significantly reduce the benefits of cross-polarization.

Path Engineering for Fixed Stations

Careful path engineering should be performed prior to any licensing work to determine if the proposed locations of base station and ARS installations will be able to achieve the desired results without using excessive amounts of transmitter power. There is much more to path engineering for a RENG system than simply pointing the transmitting and receiving antennas at each other (when directional antennas are used) and turning the equipment on. Base station and ARS systems are fixed-position installations that cannot always be located in the best possible geographic locations because of space availability problems, excessive construction or site rental costs, or local or federal licensing difficulties. In such cases, the required path is not the ideal path, and the link will have to be engineered around these fixed points.

The site selection process for repeaters and receivers should also take into consideration the RF environment in which the equipment will be working. Multi-user locations, such as the World Trade Center in New York City or Mount Wilson near Los Angeles, are very good transmit sites, but terrible receive sites. For such situations, a remotely-located receiver (in the case of a repeater system) should be considered. The two sections of the ARS would then be tied together with telco facilities.

The use of a telco loop to feed a remotely located transmitter introduces several familiar problems, such as noise, crosstalk, distortion (if repeated several times), limited frequency response and installation and service delays. In some arrangements, the use of a telco loop to the transmitter (or receiver) is the only economical way to complete the link. Although such a hybrid system is not the ideal configuration, it will get the job done. Here again, the newer digital telco services can be helpful, because of their higher audio fidelity and freedom from crosstalk, and their ease in bidirectional, multiline applications.

Planning for any RENG system should begin with an accurate, detailed U.S. Geological Survey (USGS) map covering the proposed path. Note should be made of any natural obstructions or Fresnel clearance obstructions (such as mountains, hills, or vegetation) or manmade obstructions (such as buildings, water tanks or transmitting towers) in the proposed path. The transmitting and receiving antennas should be plotted so that a minimum of 0.6 Fresnel Zone clearance is obtained over 4/3 earth radius. Information on obtaining USGS maps may be found in the chapter on Information Sources in Section 1; Fresnel zone clearance is covered in more detail in the Chapter 4.2, "Microwave Engineering for the Broadcaster."

When planning an RENG path, a profile drawing of the transmitting and receiving antenna site, the terrain, and any obstructions in between, should be made on graph paper set to 4/3 earth radius. The use of such graph paper will compensate for the curvature of the earth and the normal refraction of VHF and UHF frequency signals when determining Fresnel Zone and obstruction clearance. Simple height above sea level is insufficient to determine whether a natural or manmade obstruction will interfere with the RENG signal on a long-distance path. Once a proposed path has been drawn, a visual inspection must be made of the area for any problems that could degrade the performance of the system. Particular attention should be paid to items not documented on the USGS maps, such as buildings and towers.

The terrain from the transmitting antenna to the receiving antenna must be examined not only for obstructions but for reflection possibilities as well. A large body of water will usually cause problems for a RENG system operating in the UHF frequencies. If the water is an even number of Fresnel Zones from the direct path, signal attenuation will likely occur at the receiver. Temperature changes and tidal conditions will also have an effect. Likewise, thick vegetation or forested areas can be reflective to RF signals when wet, creating a similar (but not so troublesome) problem. Generally, the solution to reflection problems is to change either the transmitting or receiving antenna height or to employ a diversity reception system, if a long-term solution is needed.

Determining the Fade Margin

A gain and loss balance sheet should be computed to determine the fade margin of the proposed system. An adequate fade margin is vital to reliable performance of the system, because a link that is operating on the edge of the minimum acceptable receiver quieting will encounter problems later down the road. Normal component aging in the receiver or transmitter can cause a loss in received signal level and thus degrade the system performance. Likewise, new construction near the transmitting or receiving site can degrade the path, resulting in poor performance. Atmospheric conditions (such as severe weather in the area or ice on the transmitting or receiving antennas) can also cause sharp fading, and even a complete loss of signal, if an adequate fade margin above minimum receiver quieting is not provided. The RENG system fade margin can be computed by using the following equations:

$$G_s = G_t + G_{ta} + G_{ra}$$

where: G_s = Total system gain (dB)
G_t = Transmitter power output (dBm)
G_{ta} = Transmit antenna gain (dBi)
G_{ra} = Receive antenna gain (dBi)

The values for G_{ta} and G_{ra} are gathered from the antenna manufacturer's literature.

(Note: dBi = dBd + 1.1 dB, approximately.) The value for G_t is given by the following formula:

$$G_t = 30 + 10 \log P_o$$

where: G_t = Transmitter power output in dBm
P_o = Transmitter power output in watts

Next, the system losses are computed:

$$L_s = L_p + L_l + L_c + L_m$$

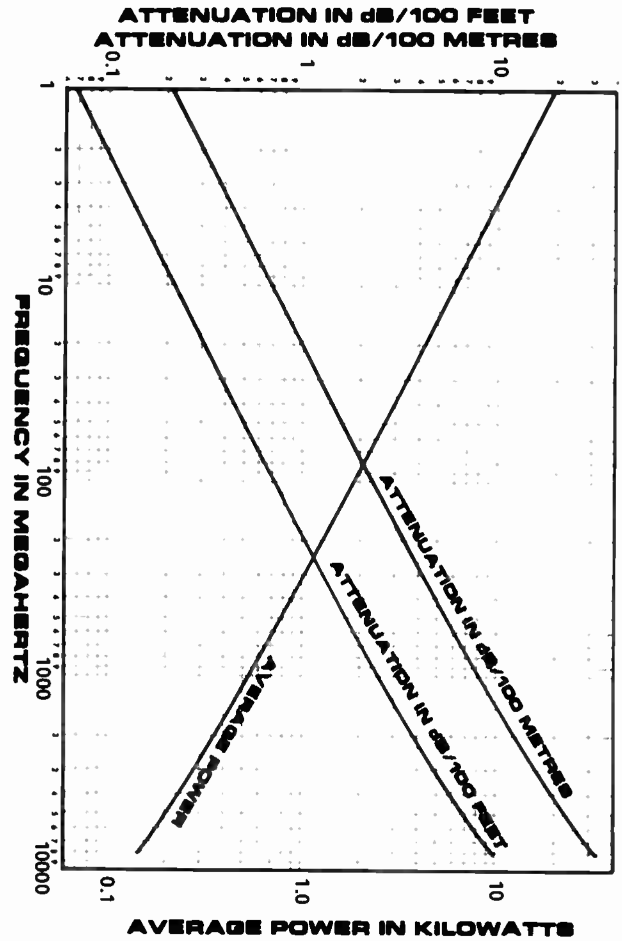


Figure 6. The attenuation and power handling ratings for ½-inch foam dielectric coax. (Courtesy of Andrew Corp.)

where: L_s = Total system losses (dB)
L_p = Path loss (dB)
L_l = Transmission line loss (dB)
L_c = Connector losses (dB)
L_m = Misc. losses (dB)

The values for L_l and L_c can be determined from the manufacturer's literature. Fig. 6 shows typical loss values for ½-inch foam-filled transmission line. A reasonable value for connector loss with components normally used in ½-inch transmission line installations is 0.5 dB. The value for L_p can be found by using the following formula:

$$L_p = 36.6 + 20 \log F + 20 \log D$$

where: L_p = Free space attenuation loss between two isotropic radiators (dB)
F = Frequency of operation in MHz
D = The distance between the antennas in statute miles

Now, the fade margin can be calculated:

$$\text{Fade Margin (dB)} = G_s - L_s - R_m$$

where: G_s = Total system gain (dB)
L_s = Total system losses (dB)
R_m = Minimum signal strength required for target SNR (dBm, a negative number)

G_s and L_s are determined by the equations previously shown. R_m (receiver sensitivity) is determined from the receiver manufacturer's specifications. If the manufacturer gives a receiver sensitivity figure in microvolts, the following formula can be used to convert to dBm:

$$R_m = 20 \log x (V_r \times 10^{-6}/.7746)$$

where: R_m = Minimum required signal strength (dBm)
V_r = Receiver sensitivity (microvolts)

R_m can also be found by using a communications receiver test set, available at most commercial and industrial communications radio service shops.

In order to predict accurately the performance of a RENG radio link, the value of R_m must be determined carefully. Many receiver manufacturers specify V_r for 20 dB of receiver quieting. This level is a convenient reference point; however, it should not be used for fade margin calculations. For maximum system performance and reliability, the fade margin determination should be made based upon the signal level required to provide the minimum acceptable receiver signal-to-noise performance.

The recommended fade margin for a 150 MHz band RPU system is at least 10 dB plus 2 dB for each 10 miles of line-of-sight path distance greater than 10 miles. At 450 MHz, the fade margin should be increased to a minimum of 15 dB plus 3 dB for each 10 miles of path distance greater than 10 miles. These fade margins are designed to limit periods of performance degradation of the radio link to 1% or less during worst-case environmental conditions. The fade margin assumes transmit and receive antenna clearance above the ground and all obstructions of 50 to 100 feet.

While it is important to provide an adequate fade margin, needlessly high fade margins should be avoided because of the spectrum congestion problems that may result.

Other Planning Considerations

Path engineering for remote-location broadcasts is seldom done for RENG activities because of the transient nature of such events. Rough estimates should be made, however, of the geographical areas of interest before attempting remote feeds. It is well worth the time spent to conduct a coverage survey of the primary areas of interest for RENG activity when planning an overall system to determine which locations provide good or marginal performance. A little planning and work ahead of time will save many problems (and probably dead air) once the system is put into operation.

For base or relay installations that, for one reason or another, cannot use frequencies in the RPU bands for program or control data interconnection, there is often an alternative to the telephone company audio loop back to the studio. Private common carriers in many markets are installing competition audio circuits, which can be leased. If the RENG installation is co-located with a television relay station, it may also be possible to back-haul the RENG audio on one of the TV microwave system's subcarrier channels.

The selection of receiving sites must be made with care, keeping in mind the area of coverage required of the receiver. The best location for a RENG system is not always the highest building in town. Placing a receive antenna at a high elevation in a metropolitan area can result in poor performance of the system in the downtown area, since the gain of many omnidirectional vertically polarized antennas decreases as the antenna is raised above the transmitting point. Tall buildings are excellent for point-to-point relay transmissions, but are generally unsatisfactory for wide-area coverage in a metropolitan region.

An inexpensive installation option is available to AM broadcast stations that do not want to erect a separate RENG transmitting tower at the main transmitting site. An isocoupler can be installed at the AM tower base that will pass the RPU transmitter frequency with good efficiency (90% is typical), while at the same time presenting a high impedance to the AM band energy. Isocouplers are available in various frequency and power ranges. Installation of these devices may change the base impedance of the AM tower slightly, thus an engineering consultant should be contacted before installation work begins.

Every effort should be made to locate the receiving antennas of a RENG system as far away from high-power transmitting antennas as possible. This should be attempted regardless of the frequency separation between the receive unit and the suspect high-power transmitting antenna. Failure to achieve adequate physical separation may require the installation of filters of various types on the receiver front-end.

In order to keep system losses to a minimum, a low-loss transmission line should be used, such as the ½-inch foam-filled coax shown in Fig. 6. The transmission line and connectors must be made watertight if exposed to the elements. Each connector should be sealed with a silicone dielectric compound and then wrapped with good quality tape. Unless this is done, rain may eventually work its way into the connector and cause signal loss or VSWR problems. The line should be grounded (using a recommended grounding kit) at the point where it leaves (or enters) the equipment building and where it starts its climb up the tower (unless the vertical distance to the antenna is less than 10 feet). This will prevent any high voltage transients caused by lightning from entering the equipment building, and thus the RENG equipment. The advantages of using a low-loss line are illustrated in Table 1. The two ends of the transmission line (at the receiver and transmitter) are probably the easiest parts of the hardwired system in which loss can be introduced, and so care should be taken to install the lines and connectors according to good engineering practice.

A short length of flexible coax is generally used on each end of the two transmission lines for connection

TABLE 1

Typical cable loss for popular types of transmission line. Note the poor performance figures for RG-58/U. LDF4–50 is ½-inch foam-filled line, LDF5–50 is ⅞-inch line, and LDF7–50 is 1⅝-inch line. (Courtesy of Scala Electronics.)

TRANSMISSION LINE COMPARISONS

L = Loss in dB per 100 feet.
E = Approximate power transmission efficiency of 100-foot length.

	150 MHz		450 MHz		950 MHz	
	L	E	L	E	L	E
RG-58/U	6.0 dB	25%	12.0 dB	6.3%	20.0 dB	1%
RG-8/U	2.5 dB	55%	5.0 dB	31%	9.0 dB	13%
LDF4–50A	0.85 dB	83%	1.7 dB	67%	2.5 dB	55%
LDF5–50A	0.48 dB	90%	0.9 dB	85%	1.55 dB	71%
LDF7–50	0.28 dB	94%	0.56 dB	88%	0.88 dB	84%

to the equipment and antennas (when ½-inch or larger coax is used). This "pigtail" is normally no more than 18 inches long. See Fig. 7.

Antenna Considerations

The selection of an antenna for use in a RENG system is an important decision because of the effect the antenna has on system performance and spectrum usage. The usual RENG antenna has, until recently, been the omnidirectional vertical whip with a small amount of gain. Many system planners, however, are now being forced by interference concerns to use directional antennas with moderate amounts of gain. The low power levels commonly used with RENG equipment and the RPU-band frequencies make it possible economically to achieve increased ERP through the use of high-gain transmit antennas. The use of high-gain antennas also concentrates the radiated signal where it will do the most good, and minimizes radiation in directions that may adversely affect the RENG activities of other stations in the same, or nearby, communities.

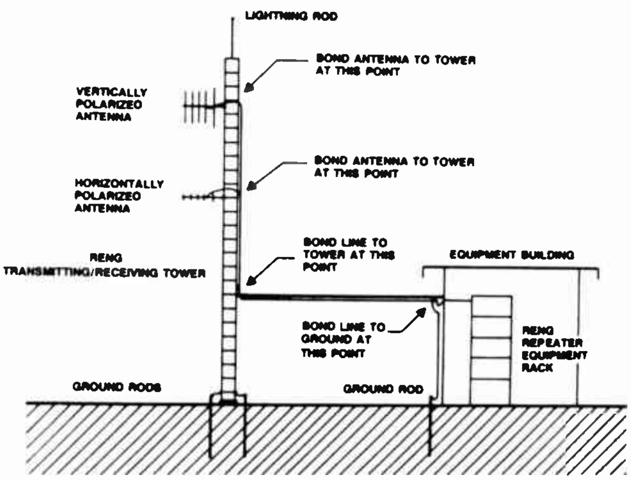


Figure 7. The recommended installation practices for RENG antennas and transmission lines.

The omnidirectional base station antennas commonly used in the 150 MHz and 450 MHz bands are vertically polarized units with 4 dB to 6 dB gain. Electrical beam tilt is sometimes available. Depending upon the manufacturer, up to 20 degrees downtilt can be provided on 150 MHz antennas, and up to 11 degrees is common for 450 MHz omnidirectional units. Large amounts of beam tilt are normally used when the antenna is to be mounted on a structure that is substantially above the surrounding terrain, thereby improving the antenna's close-in coverage.

The typical directional RENG antenna is a medium gain five-element Yagi. Such a unit provides about 9 dB to 10 dB gain over a reference dipole, with a front-to-back ratio of approximately 14 dB to 18 dB. Fig. 8 shows the radiation pattern for a commonly-used five-element 150 MHz Yagi. This particular antenna measures 40″ by 40″ by 4″ and weighs eight pounds. Thus it is small and light enough to be used on remote broadcasts. It is also suitable for permanent installations using either horizontal or vertical polarization. These antennas may be stacked in two and four bay arrays (with suitable phasing harnesses) for additional gain and directivity.

Most Yagi antennas are made to match the specific frequency requirements of the user. Multiple frequency operation using a single antenna is possible, however, with reasonable VSWR numbers, as long as the operating frequencies are not removed from the cut center frequency by more than one to two percent.

A more recent addition to the RENG user's bag of electronic tricks is the broadband log periodic antenna, which can be used on any channel within a wide band of frequencies. Such antennas provide a smooth pattern with minimal sidelobe radiation and a high front-to-back ratio (typically 25 dB in the 150 MHz band). Nominal gain for 150 MHz operation is 7 dB. Units can also be stacked to provide additional gain and directivity. Such antennas are usually larger and heavier than the familiar Yagi; however, they allow use of the antenna for virtually any frequency within the specified band at low VSWR levels (a maximum of 1.5 to 1 is typical). Fig. 9 shows the radiation pattern of a log periodic antenna designed for use in the 450 MHz

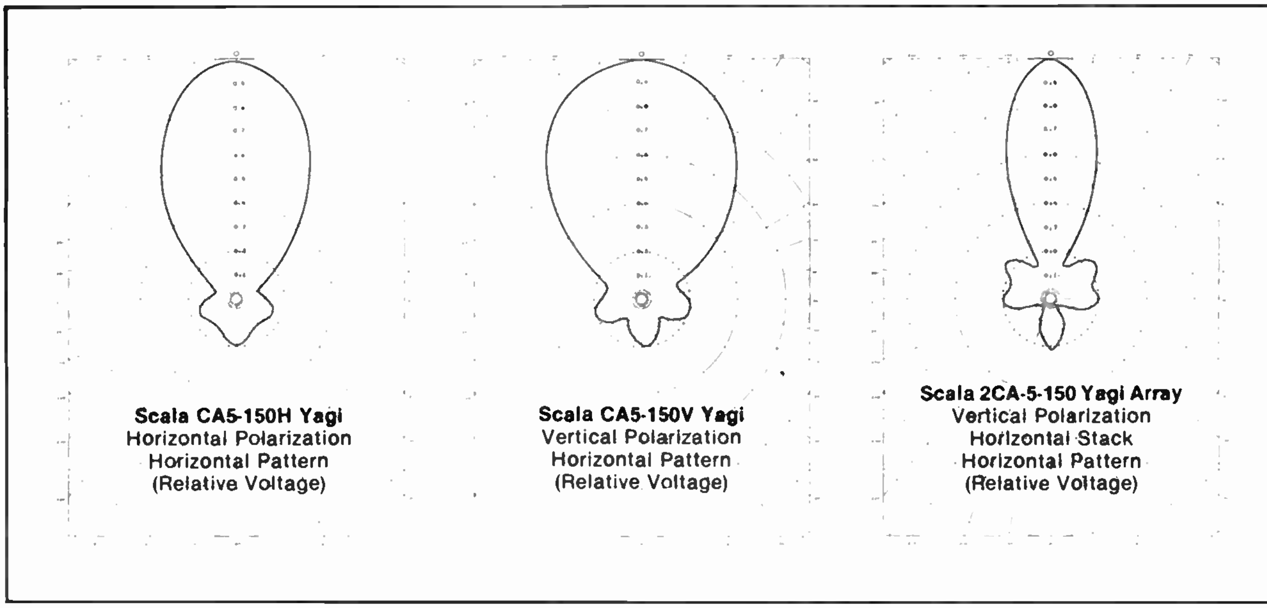


Figure 8. Radiation patterns for the CA5–150 five-element Yagi antenna made by Scala Electronics for use in the 150 MHz frequency band.

band. Horizontal or vertical polarization is available. The antenna shown in Fig. 9 has a gain of 8 dB and a front-to-back ratio of 35 dB. Such a unit is therefore well-suited for operation in areas with high spectrum congestion.

Just as a TV or FM broadcast antenna must be protected against icing problems, so should antennas used in RENG applications. Although antenna deicers are not used in RENG installations, a radome is often available for an antenna to protect it from damage or degradation in performance due to snow, ice, or salt spray.

Transmitter-Receiver Considerations

Whatever the configuration of the planned RENG system, there are several important points that should be considered. Most of these items apply to receiving equipment, which usually present the greatest problems to a system designer. Transmitting equipment must also be selected with care, but the receiving links in a RENG system are the ones most often subjected to conditions that may make good performance difficult.

A receiver should be selected that has sufficient dynamic range and headroom to allow the system to deal with strong adjacent-channel signals, as well as very weak and very strong co-channel signals from transmitters in the network. A receiver with inadequate headroom will clip and yield distortion. Wide-dynamic-range active devices should be used in the receiver front-end, such as gallium arsenide field effect transistors (GAsFETs).

The need for a preamplifier or cavity preselector network ahead of the first RF stage should also be considered. RF preamplifiers can add sensitivity, but they can also cause overload conditions in the presence of medium-level co-channel signals. Preselectors are often necessary at mountaintop or antenna farm locations because of the high-level RF signals present at such sites. It is not uncommon to have a 1 kW land mobile paging transmitter operating in the 454–455 MHz range located near an RPU band receiver that is working in the 455–456 MHz frequencies. High power FM or TV transmitters can also cause desensitizing of the receiver front end, unless adequate bandpass filtering has been included in the receiver design.

The locations commonly used for relay sites are seldom ideal from an environmental standpoint. They are often inaccessible during portions of the year, very hot in the summer and very cold in the winter. For these reasons, rugged equipment should be selected, in order to minimize downtime. Temperature extremes can also cause problems for frequency-determining elements, as well as accessories such as cavity filters, preselectors and preamplifiers. Since relay sites are often difficult to reach, equipment should be designed for easy maintenance, preferably through module replacement. A spare stock of modules should be kept at the site so that the system can be quickly returned to operation. The defective module can then be serviced at the studio, or returned to the factory for repair.

Regular performance tests should be made of the RENG system, just as an engineer would do with any other important chain of equipment at the station. Regular checks and measurement often allow the engineer to spot problems that could cause a total system failure if left unattended. If trouble is experienced with a piece of receiving equipment, the possibil-

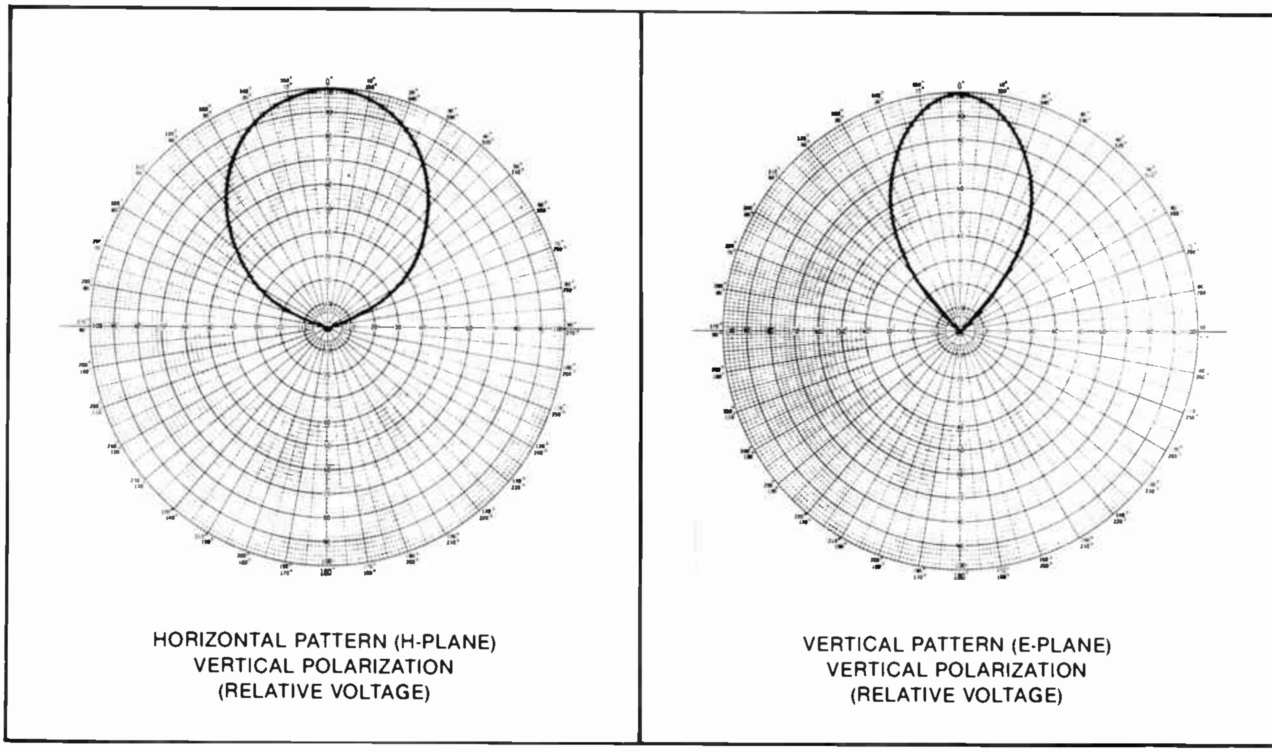


Figure 9. The radiation patters for the Scala CL-400 broadband log periodic antenna, designed for use in the 450 MHz RPU frequency band. (Courtesy of Scala Electronics.)

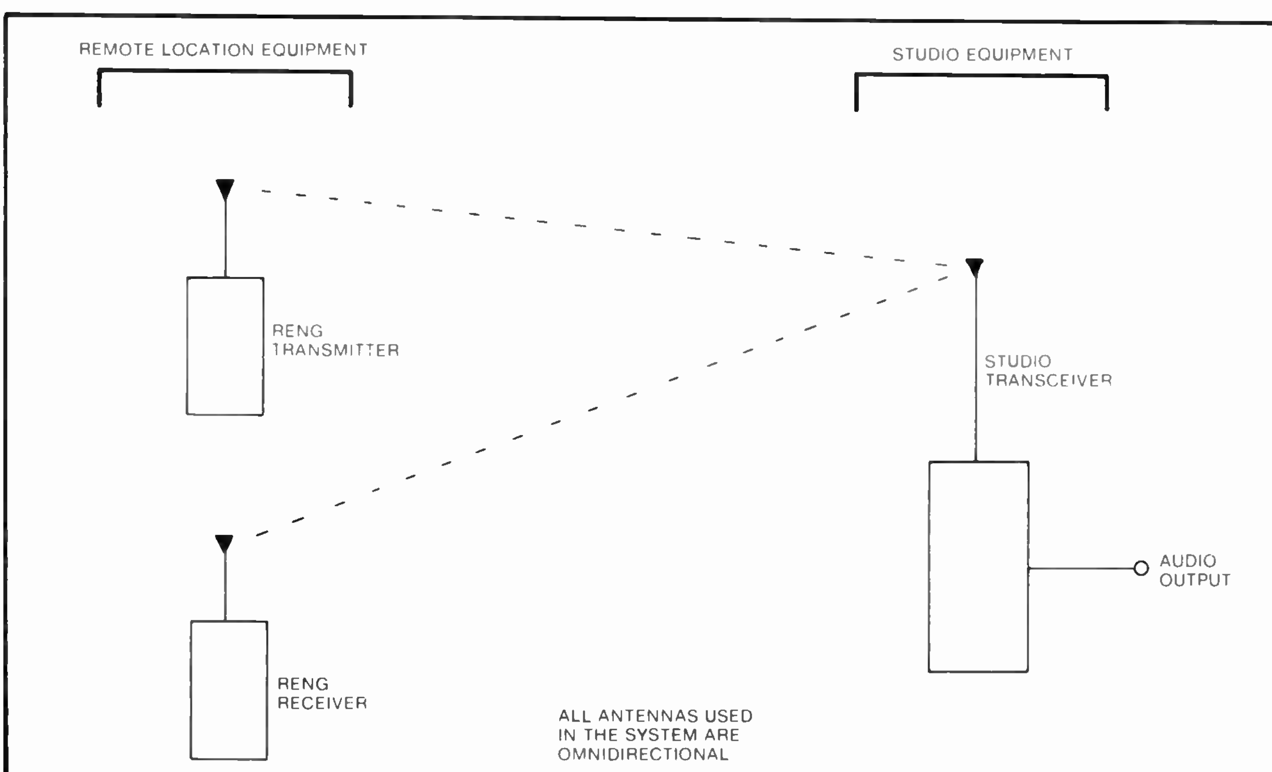


Figure 10. The basic RENG program relay system using a single hop from the remote location to the studio.

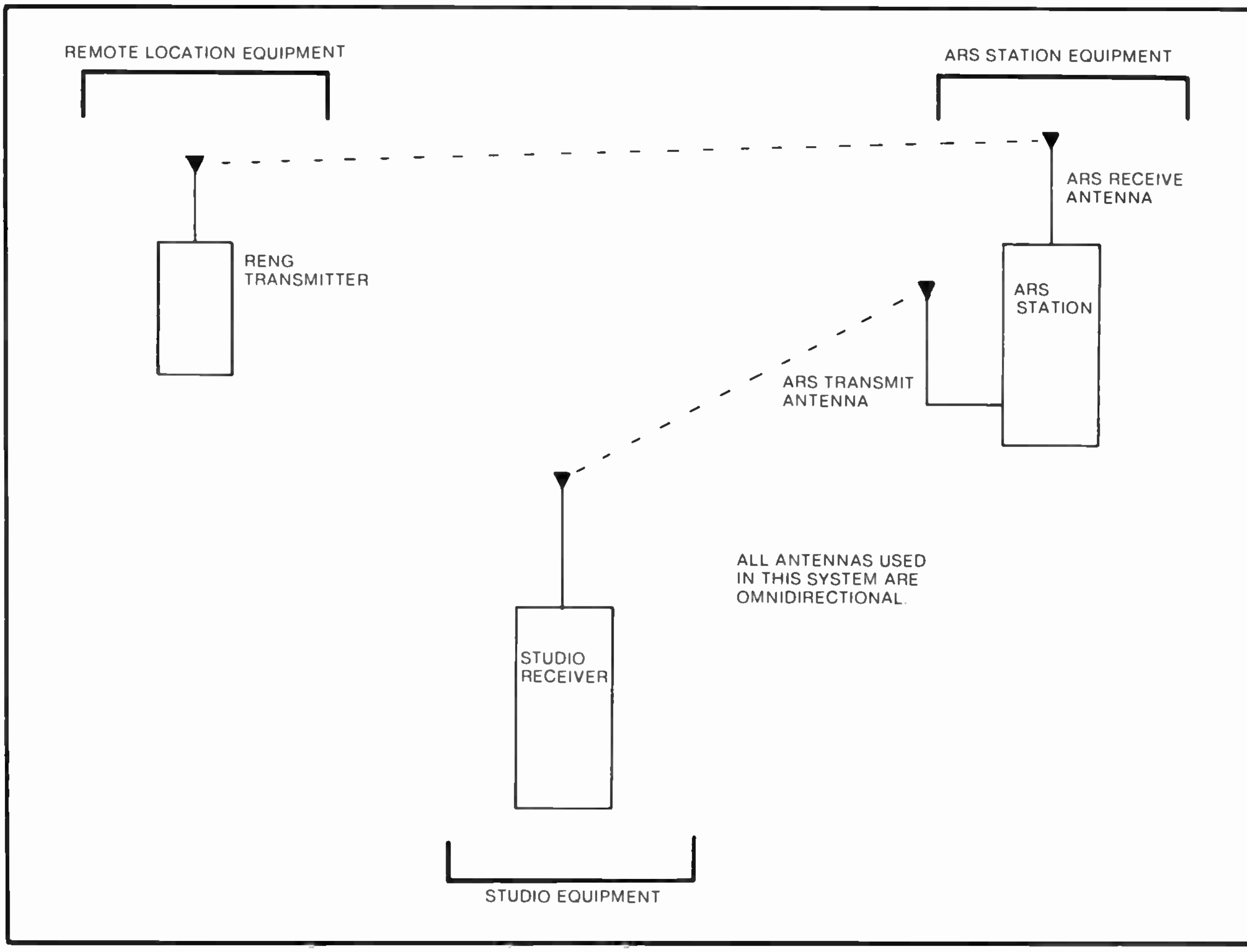


Figure 11. The basic RENG program relay configuration using an ARS station between the remote location and the studio.

ity of interference from other services should not be overlooked. A spectrum analyzer is invaluable for such work.

System Configuration

The requirements of users will vary greatly from one station to the next and from one market to the next. There are, however, several standard system configurations that can be modified to fit the requirements of most users. These range from the simple point-to-point program relay system common on many small-scale operations, to complicated multi-point relay installations with automatic signal-quality voting circuits.

Fig. 10 shows the basic RENG program relay system in which one or more transmitters on a particular frequency are used in the field, and a signal receiver is located at the studio. All antennas used in the system are omnidirectional. While there is much to be said for system simplicity, such an arrangement is not practical in an increasing number of urban areas because of spectrum congestion problems and the need to cover large geographical areas.

A simple alternative approach for a station with a good yet conveniently located transmitter site (not co-located with the studio) is as follows. The RPU antenna is side-mounted on the station's broadcast tower, and the RPU receiver system is housed in the shack or transmitter building with the station's broadcast transmission equipment. RPU audio is routed from the receiver output back to the studio via a high-quality transmitter-to-studio link (TSL). Simple control of the RPU receiver (power on/off, frequency selection) can be broken out on spare relays or closures from the broadcast transmitter's remote control system.

A more sophisticated approach may be called for in rougher terrains. The system configuration shown in Fig. 11 overcomes the geographical coverage area problem through the use of an automatic relay station. The range of a RENG system can be greatly extended through the use of an ARS. Such systems also make it possible to use lower power transmitters in the field, since the transmitter at the program origination point need only be powerful enough to reach the ARS site. This often allows the use of smaller and lighter remote transmitters, usually hand-carried or pack-carried units. The arrangement shown in Fig. 11 will satisfy the requirements for wide area coverage and is sufficient for radio markets where spectrum congestion is not a problem. Because all antennas in the system are omnidirectional, however, the configuration is not suitable for use in larger urban areas which are experiencing frequency allocation problems. For such applications, a more sophisticated approach is needed to RENG activity.

Fig. 12 shows a high-performance two-point RENG system designed for operation in spectrum-congested areas. At the remote site, two transmitters and two antennas are used. The communications transceiver is used for conveying setup information between technical personnel at either end, and to relay cues and coordinating information. The low-power transmitter and its associated directional transmit antenna are used to relay the program audio signal to the studio. At the studio site, a communications transceiver, feeding an omnidirectional antenna, is used for the setup information, cues and coordination work. The multi-antenna receive system is used for program audio pickup.

The "cues and orders" radio system shown in Fig. 12 is used for general purpose communications not requiring wide frequency response and a high signal-to-noise ratio. The lower power program relay transmitter and directional receive and transmit antennas provide a secure and quiet channel, without causing interference to other RPU band users in the area.

Between remote broadcasts and when beginning the initial set-up procedure for a remote, the omnidirectional antenna is patched into the broadcast-quality RPU band receiver at the studio through the coaxial switch K1. Once contact has been established with the remote crew, one of the directional antennas (which are mounted on a common mast driven by a remote-controlled antenna rotor) is switched into the studio receiver. The polarization of the transmission from the remote site is planned before the remote crew leaves the studio. Selection of either horizontal or vertical polarization is made during the frequency coordination process, or at the discretion of the user. Engineers may find that a particular polarization may yield better results from certain geographical areas, and in such cases, that polarization would be chosen. Once the proper antenna has been selected, the antenna rotor is

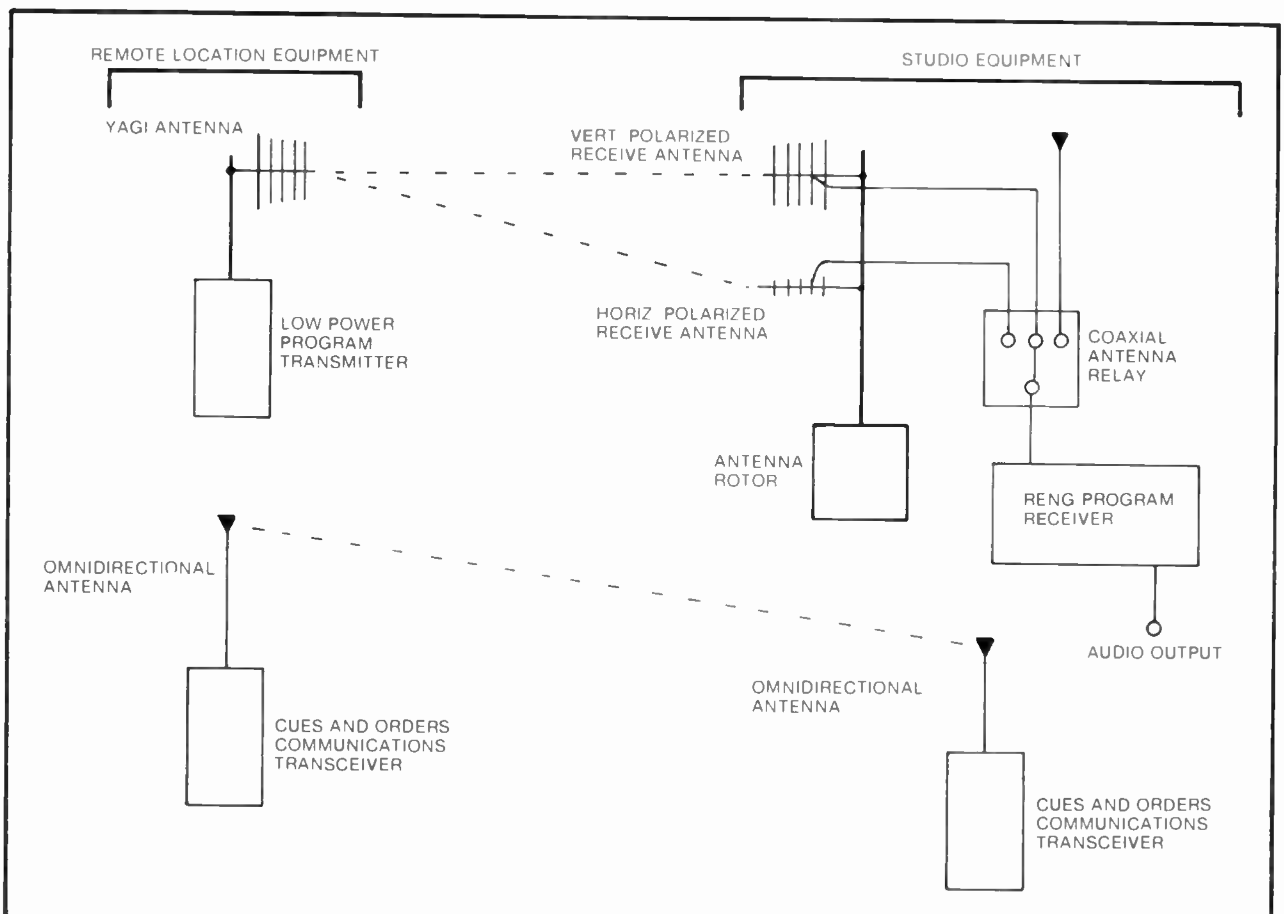


Figure 12. A high-performance two-point RPU system designed for operation in frequency-congested areas.

adjusted for maximum received signal strength. The studio operator then talks the remote crew into the best position for its Yagi transmit antenna. At this point, the antennas are locked down and the link is ready for the remote broadcast.

If a variable power transmitter is used at the remote site, or transmitters of various power levels are available, the transmitter power would next be adjusted to the point necessary to achieve the required SNR performance at the receiver. After power output adjustment, the antennas on both the receive and transmit ends should be checked again for correct positioning.

While this process may be time consuming and require the purchase of additional equipment, it will assure a high quality, secure RF link from the field to the studio. This system will also result in a minimum of unwanted radiation to other RPU band users.

Fig. 13 shows a high-performance, secure-channel automatic relay station (ARS). The same antenna selection and positioning procedure is used in the ARS installation as was used in the two-point system of Fig. 12, except that the antenna switching and positioning work is done by remote control. The link for this remote control system can be a subcarrier on the main station broadcast signal, a separate dedicated radio link, a dial-up telephone patch or a leased telco data or voice loop. A standard broadcast transmitter remote control system is used, with the common channel on-off/up-down functions performing the necessary switching and positioning work at the ARS site. For stations with multiple site capability on the main transmitter remote control system, the ARS remote points can be simply treated as other "transmitter sites" and controlled as such from the master unit.

A monitor receiver is included at each ARS installation to inhibit activation of the ARS transmitter if a transmission is already in progress on that frequency. As shown in Fig. 13, the control commands are received over a subcarrier receiver from the main station transmitter. The relay station logic interfaces the remote control unit with the receive antenna coaxial switch and the antenna rotor control box. During set-up, the telemetry section of the remote control unit provides an audio frequency shift-keying (FSK) signal

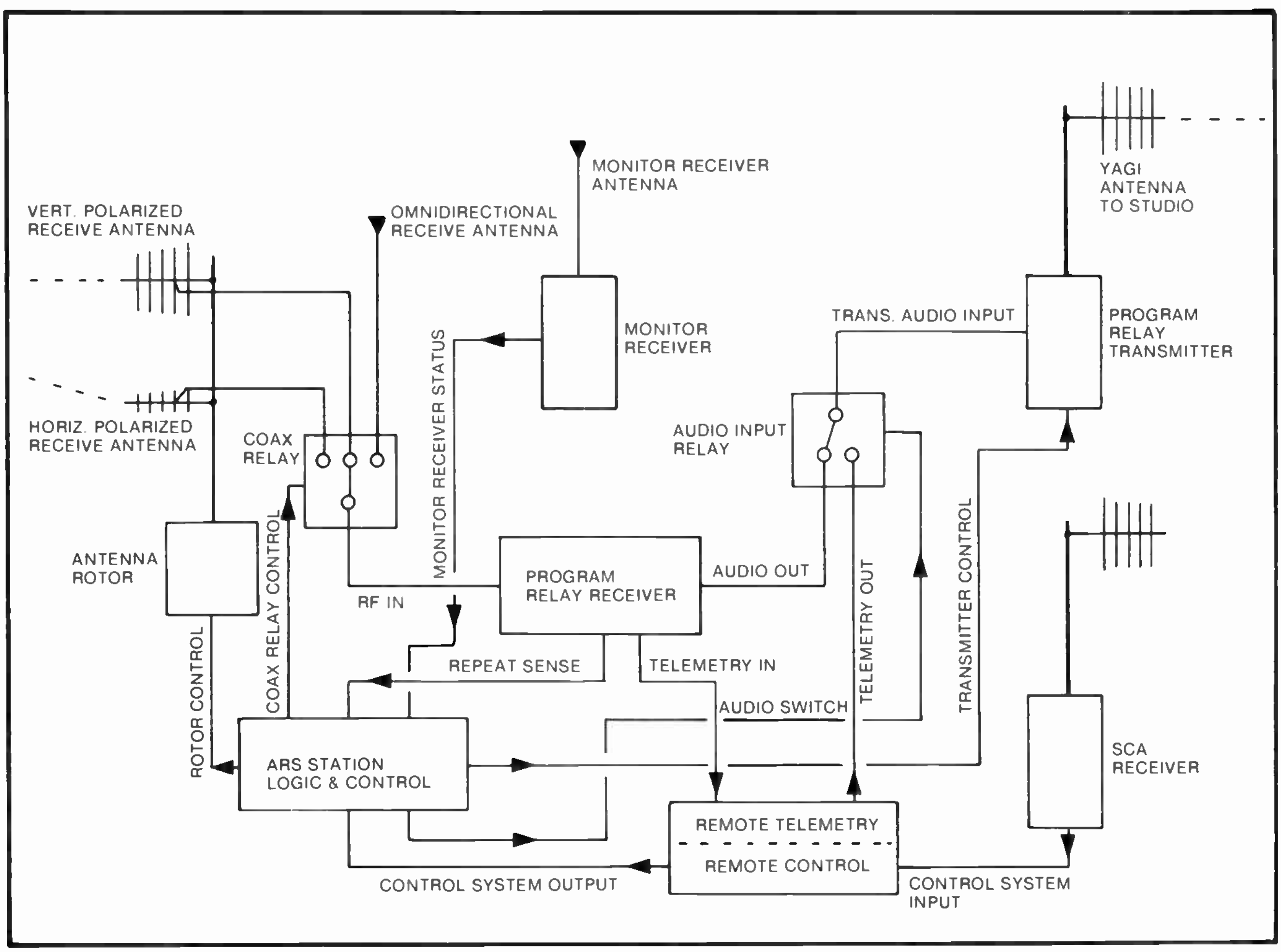


Figure 13. A high-performance secure-channel ARS station with remote control of system functions.

that is sent back to the studio control unit via a telco line or the relay (ARS) transmitter, as shown.

The secondary communications authorization (SCA), omnidirectional program, monitor and relay transmit antennas are all fixed in position. Only the directional receive antennas, one set for horizontal polarization and the other for vertical polarization, are movable. An arrangement such as that shown in Fig. 13, will provide maximum flexibility and minimum risk of program audio disruption.

In a system where two or more of the ARS stations shown in Fig. 13 are used, an arrangement such as that shown in Fig. 14 may be implemented. The studio remote control unit is used to determine which of the ARS stations is allowed to repeat the program traffic. Those stations which will not be used to repeat the program material would be instructed by the studio operator to remain inactive. For multiple site ARS operation, as shown in Fig. 14, individual directional receive antennas, or a single directional receive antenna mounted on an antenna rotor, may be used to receive the ARS traffic at the studio.

For protection against system failure, an equalized telco loop can be installed between each ARS point and the main studio. With this backup provision, an equipment failure in the relay gear would not interrupt a remote broadcast.

One of the problems sometimes experienced with ARS equipment is the possibility of a desired signal opening the system, and an undesired signal keeping it open after the desired traffic has ended. This can occur if the tone burst method of repeater keying is used. For example, a valid tone burst signal unlocks the ARS system and then undesired noise or traffic holds the channel open after the desired traffic has ended by prohibiting a loss-of-carrier indication from the receiver. The ARS will thus be stuck open until the level of the interfering signal drops to a point that allows the receiver to squelch and generate a loss-of-carrier command to the ARS system logic. In order to maintain positive control over the system, a means should be provided to override the ARS logic by remote control from the studio.

Fig. 15 shows some of the ways the remote location

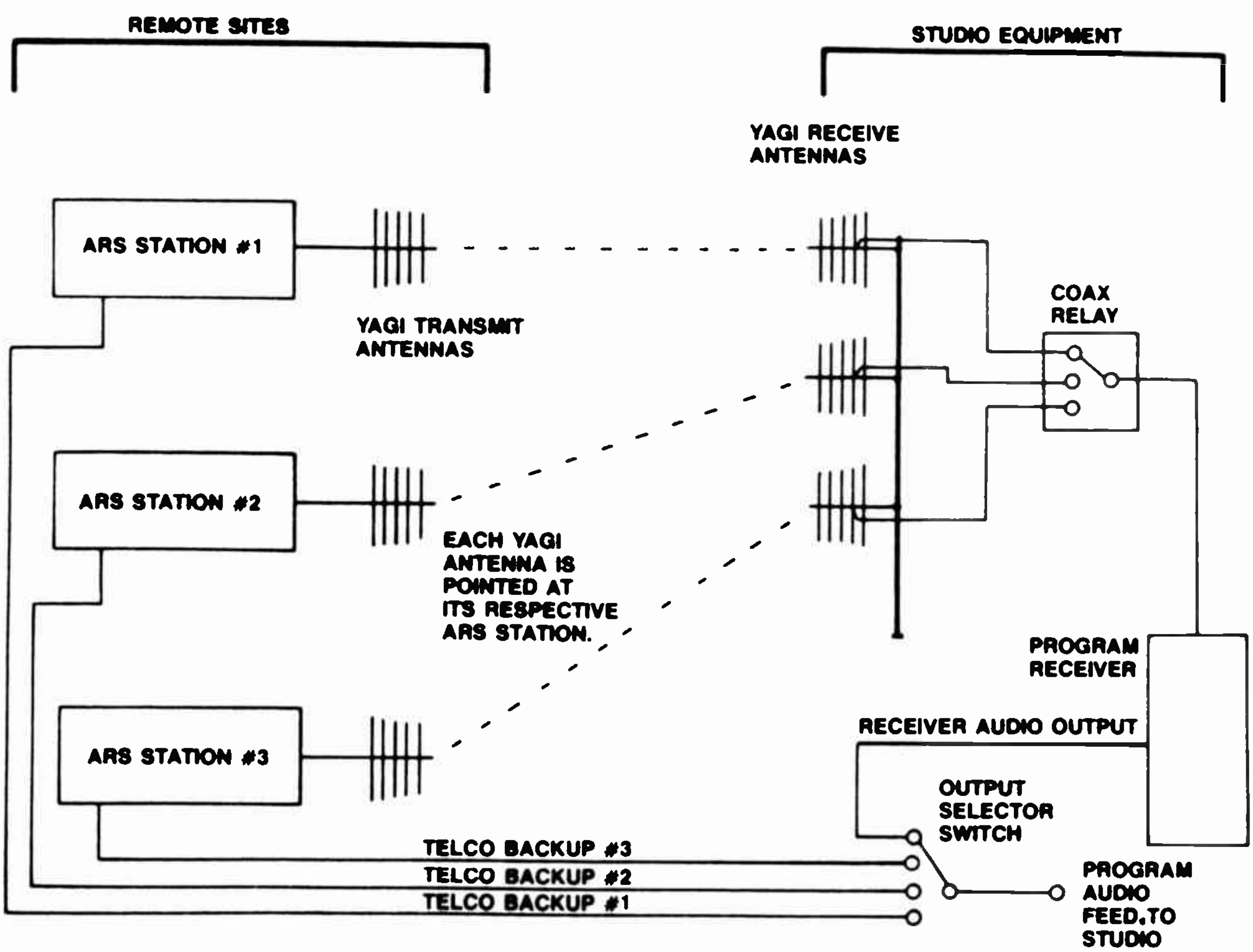


Figure 14. A multiple site ARS network feeding a central studio control point. Note each ARS station is also connected to the studio via a land line for backup protection.

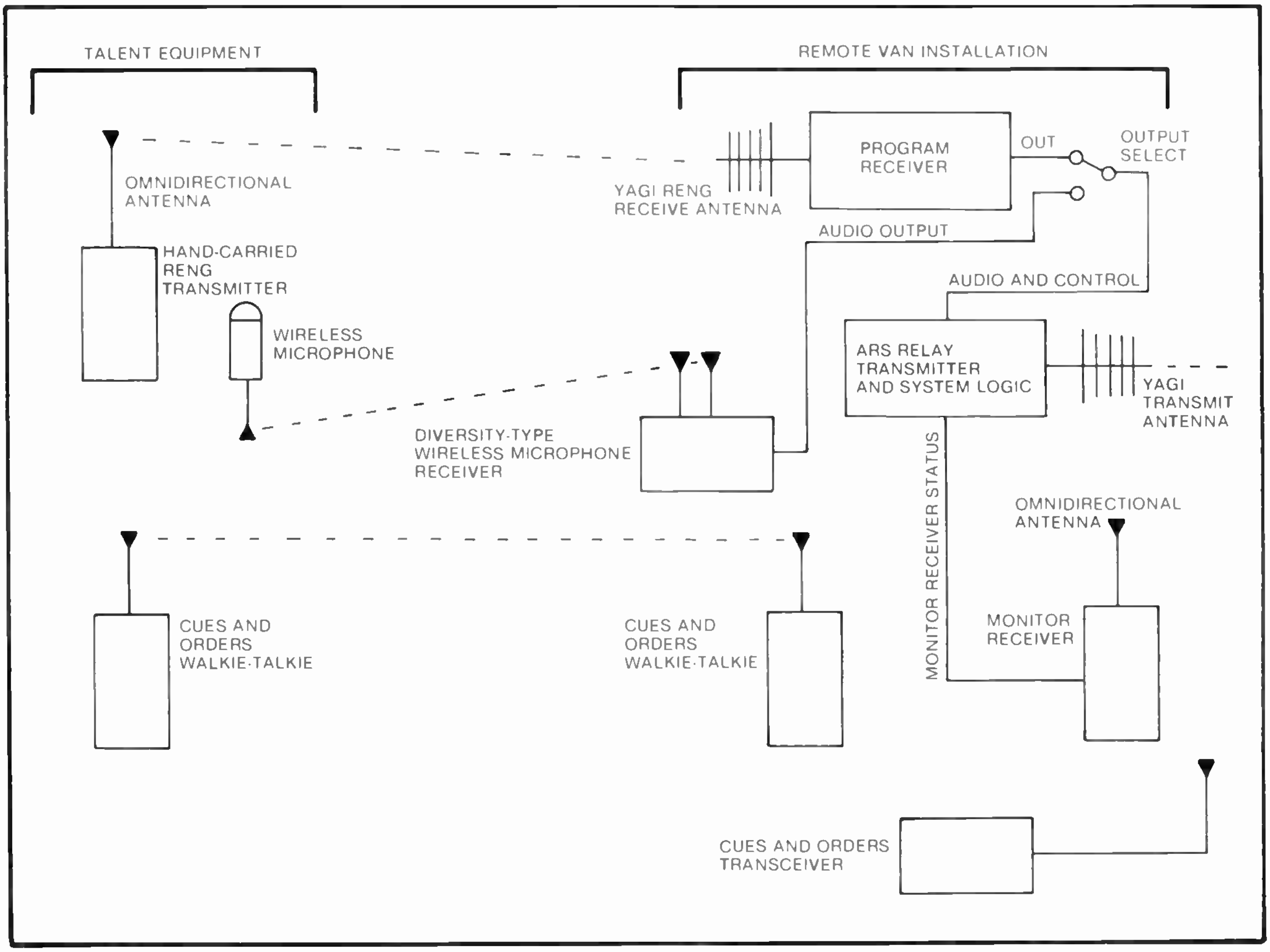


Figure 15. The use of an ARS system at the event site for added range and talent flexibility.

program audio can be transmitted. As mentioned previously, the communications transceiver is used for cues and orders from the studio location. The program channel signal can consist of a hand- or pack-carried transmitter, which directly feeds the studio receiver or one or more ARS systems. Fig. 15 also shows a repeater station configuration that can be used when a high-power transmitter is required to reach either the studio or the ARS relay point. The use of a repeater, configured as a standard ARS station, in a car or van outside the remote location also gives the talent at the event greater flexibility, because a small hand-carried transmitter can be used there, rather than a larger unit with antenna and power cables attached. This arrangement is also ideally suited for use with a wireless microphone, which gives the talent an even greater degree of flexibility. The receive antenna at the remote van can be either an omnidirectional unit or a Yagi. The system shown in Fig. 15 includes a monitor receiver to prevent ARS transmission over traffic already in progress.

There is a limit, of course, to the number of times a signal can be repeated and still maintain good audio specifications. Moreover, each added hop in the path between the remote site and the studio increases the chances of a spurious signal interrupting the remote feed. Each additional hop also increases the complexity of the system and the vulnerability of the link to equipment failure. The design goal for any RENG system should be to keep the arrangement as simple and direct as possible, while still providing talent flexibility, backup protection and high performance.

WIRELESS MICROPHONES

The use of wireless microphones to free-up the talent at a remote broadcast is gaining popularity with stations involved in RENG activity. The advantages to the talent are obvious: complete freedom of movement and nothing to carry around but a microphone and air monitor receiver. There are no controls or meters for talent to worry about. The range of a wireless mic is somewhat limited, but a properly designed system for remotes that are more-or-less stationary can provide simple set-up and coverage of an event.

The receiver used in conjunction with the wireless microphone may use either diversity or nondiversity reception techniques. A nondiversity receiver is used where multipath cancellation is not a problem, such as in open areas or when conducting fixed-position interviews. If, on the other hand, the wireless mic is to be used in several places and the possibility of multipath cancellation exists due to nearby reflective objects, a diversity receiver is recommended.

The diversity receiver uses two antennas, located in different areas of the event site. A minimum separation of 20 feet is usually recommended. The receiver automatically selects the stronger of the two signals for demodulation. The switching of RF sources occurs silently without any "squelch-type" noise bursts.

Many wireless microphone systems include audio companding circuits to extend the dynamic range and lower the apparent noise floor. A properly engineered wireless microphone system can be treated by engineering personnel as essentially a piece of wire between the microphone and the audio console input. Both VHF and UHF frequencies are used. An often-overlooked FCC regulation requires licensing of all wireless microphone systems.

REMOTE CUES AND ORDERS

Communications with a remote crew from the studio can be accomplished in one of several ways. The simplest method is an over-the-air cue in which the talent simply listens to the station's air signal and takes his (or her) cue from the studio announcer or a prerecorded introduction cart. Other methods include use of the station's subcarrier signal for cuing information; or a separate, dedicated, radio link specifically used for cuing instructions, either from the remote truck (as shown in Fig. 15) or from the main studio.

If a station needs a more sophisticated intercommunication system, a trunked 800 MHz radio system can be considered. A five- or ten-channel trunked repeater acts like a small telephone exchange in which the number of users (telephones) exceeds the number of channels (trunk lines). Telephone system theory is used to predict the busy level that can be expected during periods of heavy radio traffic. Three minute time-out timers are usually included in mobile transmitters to enforce time limits.

These trunked systems can tie into the regular telephone system at hilltop repeater sites or at trunked base stations. Broadcasters interested in 800 MHz trunked radio should contact their local area land mobile operator to see if such a system is available.

In certain situations, a station may be able to design and license a UHF business radio system for dispatch and coordination of RENG crews. These systems offer the user the luxury of not encountering a busy signal, as may occasionally happen in a trunked system. As with the trunked network, no programming is allowed on a UHF business radio system.

One of the problems often encountered when carrying remote broadcasts on an automated station is the need to have an operator stand-by during the broadcast to trip the automation system to the next event when the talent at the remote site gives the proper cue. The simplest way around this problem is through the use of a subaudible tone that is high enough in frequency to not interfere with the ARS subaudible tone that may be used for repeater equipment, and low enough in frequency so that it does not interfere with normal program audio. For example, in a system where the ARS access subaudible tone is 25 Hz, an "advance system" control tone of 45 Hz could be used. In order to prevent premature automation system trip commands, the program audio input to the remote location transmitter would be passed through a high-pass filter to remove any audio components below about 60 Hz. At the studio, the receiver audio output would run through another high-pass filter to remove any control tone signals from the automation system program channel feed.

A more recent solution available in most cases is the use of the cellular telephone system for this purpose. A portable or transportable cellular phone has become a common part of most RENG systems today. When more than just simple voice communication is required, such as interruptable foldback (IFB) or other audio signals to talent headphone or monitor speakers, an audio interface to and from the cellular phone may be used on both ends of the link. Unlike the standard dial-up system, the cellular phone system is "four-wire" end-to-end (meaning that it uses a *pair* of RF channels, one for transmit and one for receive), so that no hybrid or gating circuitry is required. The same type of hardware may be used for feeding RENG program audio from the field when no other method is possible. Audio performance will typically be somewhat to greatly degraded from what is expected from the standard dial-up phone feed.

EMERGING TECHNOLOGIES

At this writing, a new addition to the wired type of RENG transmission systems is becoming available. It is generically referred to as the switched digital network, and includes "Switched 56" and ISDN (integrated switched digital network) services. The former uses a 56 kbit/sec data rate, and the latter will offer a variety of 64 kbit/sec channel configurations. Both services terminate to the customer on twisted-pair telco cable, and operate bidirectionally. Although originally designed for convenient and versatile computer data communications, wideband program audio (7.5 kHz to 15 kHz at present) can be placed on these relatively low-data-rate channels through the use of a "black-box" adapter on each end of the line. The adapter implements a data compression algorithm, which reduces the data rate of the digitized audio signal to a rate that will correspond to the line's; but does so in a way that is audibly imperceptible: by using psychoacoustic principles of masking. In this way, real time, high-quality audio will be only a (digital) phone call away. ISDN is intended to replace the existing public switched telephone network, and widespread, if not universal, deployment is expected by the end of the 1990s. For more detailed information on these systems, see Chapter 6.4, "Telco Audio Program Service."

Another new system that may have application for radio remotes is the fiber optic snake. Here, a single, robust, yet lightweight cable carries several dozen separate microphone or line level channels from one part of the remote site (typically the stage area) to another (typically the mix position), replacing the traditional heavy copper multipair cable. Also unlike its predecessor, the fiber optic system is impervious to the pickup of EMI, nor does it generate any. Microphone-level audio is preamplified at the head- or stage-end, then digitally encoded and multiplexed, converted to the optical domain, and sent down the fiber. The reverse process takes place at the console end, terminating in line-level analog balanced outputs, to be plugged into a conventional mixer's line inputs. Splitting can also be accomplished without difficulty or degradation in the digital (electrical, not optical) domain with these systems. A helpful feature is the addition of some remote control ability for microphone preamp gain from the mix position, since the preamps are located at the head-end of the snake. Since this is an active rather than passive system, some provision must be built in for return signals running in reverse from the console back to the stage. A typical configuration is 56 send and eight return paths on a single cable, about the size of a standard mic cable.

CONCLUSION

A RENG network should be planned and constructed with long-term service and frequency coordination requirements in mind. Areas that currently do not experience spectrum congestion problems may encounter them in the near future. It pays, therefore, to design a system that is spectrum-efficient and relatively immune to interfering signals. New digital services, both telco-supplied and of an RF variety, should also be considered, and the progress in these areas carefully monitored for application on future links. It is always easier (and cheaper) to do the job right the first time.

REFERENCES AND BIBLIOGRAPHY

1. "Engineering Radio Remotes," by Skip Pizzi. *Broadcast Engineering*, January 1986.
2. "Remotes Revisited," by Skip Pizzi. *Broadcast Engineering*, January 1991.
3. "Dealing with Telco Lines (Parts 1 and 2)," by Skip Pizzi. *Mix*, April/May 1985.
4. "Telco Audio Program Service," *NAB Engineering Handbook, 8th Edition*, Chapter 6.4.

APPENDIX

Excerpt from FCC Rules and Regulations

§ 74.402 Frequency assignment.

(a) The following frequencies may be assigned for use by remote pickup broadcast stations:

(1) *Group A* (kHz) : 1606,[1] 1622, 1646.

(2) *Group D* (MHz) : 25.87,[2] 26.15, 26.25, 26.35.
Group E (MHz) : 25.91 ; [2] 26.17, 26.27, 26.37.
Group F (MHz) : 25.95 ;[2] 26.19 ; 26.29 ; 26.39.
Group G (MHz) : 25.99 ; [2] 26.21 ; 26.31 ; 26.41.
Group H (MHz) : 26.03 ; [2] 26.23 ; 26.33 ; 26.43.

(3) *Group I* (HMz) : 26.07 ; [3] 26.11 ; 26.45.
Group J (MHz) : 26.09 ; [3] 26.13 ; 26.47.

(4) *Group K_1*[8] (MHz) : 152.87[3], 152.93[3], 152.99[3], 153.05[3], 153.11[3], 153.17[3], 153.23[3], 153.29[3], 153.35[3].
Group K_2[8] (MHz) : 161.64[5] ; 161.67[5] ; 161.70[5] ; 161.73[5] ; 161.76[5].

(5) *Group L* (MHz) ; 166.25[4].
Group M (MHz) ; 170.15[4].

(6) *Group N_1* (MHz) : 450.050 ; 450.150 ; 450.250 ; 450.350 ; 450.450 ; 450.550 ; 455.050 ; 455.150 ; 455.250 ; 455.350 ; 455.450 ; 455.550.
Group N_2 (MHz) : 450.0875 ; 450.1125 ; 450.1875 ; 450.-2125 ; 450.2875 ; 450.3125 ; 450.3875 ; 450.4125 ; 450.4875 ; 450.5125 ; 450.5875 ; 450.6125 ; 455.0875 ; 455.1125 ; 455.-1875 ; 455.2125 ; 455.2875 ; 455.3125 ; 455.3875 ; 455.4125 ; 455.4875 ; 455.5125 ; 455.5875 ; 455.6125.

(7) *Group P* (MHz) : 450.01[6] ; 450.02[6] ; 450.98[6] ; 450.-99[6] ; 455.01[6] ; 455.02[6] ; 455.98[6] ; 455.99[6].

(8) *Group R* (MHz) : 450.650[7], 450.700[7]. 450.750[7], 450.800[7], 450.850[7], 455.650[7], 455.700[7], 455.750[7], 455.800[7], 455.850[7].
Group S (MHz) : 450.925,[7] 455.925.[7]

[1] Subject to the condition that no harmful interference is caused to the reception of standard broadcasting stations.

[2] Subject to the condition that no harmful interference is caused to stations in the broadcasting service.

[3] Subject to the condition that no harmful interference is caused to stations operating in accordance with the Table of Frequency Allocations set forth in Part 2 of the Commission's Rules and Regulations. Applications for licenses to use frequencies in this group must include statements showing what procedures will be taken to insure that interference will not be caused to stations in the industrial radio services.

[4] Operation on the frequencies 166.25 MHz and 170.15 MHz is not authorized (i) within the area bounded on the west by the Mississippi River, on the north by the parallel of latitude 37°30' N., and on the east and south by the arc of the circle with center at Springfield, Ill., and radius equal to the airline distance between Springfield, Ill., and Montgomery, Alabama, subtended between the foregoing west and north boundaries; (ii) within 150 miles of New York City; and (iii) in Alaska or outside the continental United States; and is subject to the condition that no harmful interference is caused to government radio stations in the band 162–174 MHz.

[5] These frequencies may not be used by remote pickup stations in Puerto Rico or the Virgin Islands. In other areas, certain existing stations in the Public Safety and Land Transportation Radio Services have been permitted to continue operation on these frequencies on condition that no harmful interference is caused to remote pickup broadcast stations.

[6] The use of these frequences is limited to operational communications, including tones for signalling and for remote control and automatic transmission system control and telemetry.

[7] The use of these frequencies is limited to the transmission of program material and cues and orders immediately necessary thereto.

[8] Frequencies in Group K_1 and K_2 will not be licensed to network entities. Frequencies in Group K_1 will not be authorized to new stations for use on board aircraft.

6.2 Television Field Production: An Overview

Carl Bentz
Broadcast Engineering Magazine, Overland Park, Kansas

Television brought the dimension of pictures by radio into our lives. Yet, over the years, technology has achieved technical successes almost as important as was the mere existence of early telecasts. Once bound to a studio by sheer size of equipment, TV production is now almost at home in the field. Going remote is more than just buying battery-powered equipment, however. This overview of field production will examine some of the aspects that impact the development of outside broadcast capabilities.

WHAT IS FIELD PRODUCTION?

Whether it is called *electronic news gathering* (ENG), *electronic field production* (EFP), *outside broadcasting* (OB) or *remote*, the concept of field production is the same. Field production is an escape from the artificial environment of studio walls into the reality of the outside.

Electronic News Gathering

ENG adds interest to newscasts even though it is the simplest approach to field production. Visually, the viewer is among marathon runners at the finish line, sitting on the bank with the winner of a fishing contest, or stroking on the last coat of paint in an assist-the-elderly neighborhood project. The viewer can experience the terror of a burning building, the fury inside a tornado and the hammering force of a military attack in some country halfway around the world, all coming to us live by microwave and satellite, or, perhaps, tape delay. ENG may be as simple as a one-operator camera-recorder unit, but it can be as complex as transoceanic satellite relay fed by a portable earth station from a production vehicle and a vast team of engineers, technicians, and equipment operators working different parts of a global network.

ENG puts heavy demands on equipment because of a need for immediacy and flexibility. Attributes of ruggedness and portability with ease of setup and operation have produced several generations of cameras, recorders, and assorted peripheral products, each answering some new requests of news production. Directors want usable pictures and sound now, never mind that gamma, color balance, and signal-to-noise are not quite optimized. Signal processors at the station will correct or hide those problems! Meanwhile, simplified operator controls and microprocessor-controlled setup have reduced the chances for operator technical error.

Electronic Field Production

EFP moves outside the news hour. From soap operas to documentaries, being on location via tape is more realistic and interesting than being live inside a studio. After all, studios require sets to recreate the outside, while just outside the studio door is a set already furnished, ready for use. The term production suggests video recording for eventual assembly into a final program, but EFP can also be live.

EFP equipment may not need to endure the rough handling that news crews give, but portability, ease in operation, and reliability remain criteria upon which many technical managers specify products. Because EFP projects are more planned in nature than ENG, the automatic sophistication of ENG equipment may be substituted with increased stability, more technical adjustment, and features for expanded signal cosmetics. Products suitable for ENG can serve EFP applications, but they may not provide the long-term quality of signal under every circumstance that is expected from EFP equipment.

Outside Broadcasting

A hazy boundary divides EFP from OB and the televising of program length remotes. Viewers expect Rose Bowl parades, political conventions, presidential inaugurations and tennis matches from Wimbledon.

Seeing an event as it happens brings us the experience in a way that tape does not seem to duplicate.

In outside broadcasting, flexibility and rugged design continue in importance, but performance quality is even more important than with ENG and EFP. Automatic over-rides may be disabled for manual adjustments to achieve optimum signals. The news can accept a noisy picture with some color imbalance and distorted sound, if the content is urgent and the exposure is brief. Still, for longer segments and full-length programs, particularly where considerable expense is involved in acquiring rights, setting up and post production, producers and sponsors expect near studio quality.

How does field production pay back the expense it brings? Viewer satisfaction is one way. The visibility of a TV crew at an event heightens the station's commitment to, and involvement in, the community in the perspective of the viewer. The station becomes more a part of the locale, which brings increased community acceptance of the station and its products (programming and commercial messages) and revenue.

PLANNING FOR FIELD PRODUCTION

Preparing for remote capability is a major step in station growth and a major investment. It is not a step that can be taken lightly, nor is it one that can be returned to the store if the final system does not suit someone or if a mistake was made. As a result, planning is critical. One of the primary concerns must be budgeting. Management will want the most for the least dollars, but cost effectiveness and reliability should be balanced against the dollars spent. Both personnel and equipment considerations will bring all departments of the station into every step of the process. Programming and production may know what they want to do, but engineering must make it work. The sales department must continue to bring in revenue so accounting can make the payments and the extra salaries.

The planning phase is the time to express the wish list, along with the realizations of harsh reality and limitations. Compromises are inevitable, and from them will come decisions on the proposed extent of field production activities. Answers to a number of questions must be found. For example:

1. Who will operate what equipment? Labor agreements may dictate that engineering staff handle all camera and recording operations. In another station, production people may do some of the technical work. No matter which approach is used, it must be determined if (a) there are enough people currently on the staff to handle the extra assignments, or (b) how many people, and with what skills, must be hired.
2. To what extent is remote operation expected? How many hours per week or month will be involved? Will this activity be limited to the local area or will it extend to a larger surrounding region?
3. Do plans warrant purchasing a production vehicle? What size of vehicle will meet station needs? How should it be outfitted to meet the expected production projects?

ASSESSING OPTIONS

Because of the vast number of questions to be answered, some method should be devised to organize your thoughts and evaluate each answer, relative to the entire project. One approach to consider is a type of spread sheet. The first column will list the questions. The answers to those questions will most likely fall into two categories: those affecting the project positively, and those with negative effects. Develop a scale to rate responses to each question. If every question is equally important, use a standard 1–10 range for ratings. On the other hand, questions may have different weights toward the final decision. Use column two to list the point value for that question. Column three will rate the positive effects, while column four will contain negative valued ratings. Setting up these questions with their numerical answers as a simple computer spread sheet is an easy but effective way to tally responses and develop an assessment.

Realize that this is not a valid statistical approach. It may be necessary to create several spread sheet analysis forms for different parts of the project. It will, however, provide a means to assess the answers in planning remote production equipment, because it permits playing a 'what if' game, getting immediate answers, and refining both the questions and the answers in successive iterations.

SPECIFYING A REMOTE PRODUCTION VEHICLE

The market for remote production vehicles (RPVs) is a limited one, which forces the cost of such a vehicle to be high. Errors in specifying the vehicle can prove very expensive. Avoiding buyer's remorse, endless criticism, inconvenience, and potential trouble, the best approach to shopping for a production vehicle means asking the right questions; because unlike other TV plant projects, RPVs can't be planned for expansion. Once a basic chassis is ordered, a number of aspects of the RPV become fixed.

Levels of Production

There are several levels of production requirements that will determine the overall scope and subsequent price tag of the project. On-site commercial videotaping requires the least from an RPV; live professional sports coverage can stretch the capabilities of the best equipped system. A good deal of the complexity of a vehicle is reduced if the vehicle is used only for remote videotaping. Such an operation leaves all editing, graphics insertion, effects system, audio sweetening, and other post production tasks in the studio. When the vehicle becomes the source of finished productions for live telecasting, the complexity and cost increase considerably because of the extra equipment and personnel requirements.

Single-unit Vehicles

Once the RPV use is identified, an initial budget can be developed and the search for a vehicle type started. The single-unit vehicle (with cab and box on a common chassis) is more maneuverable, easier to drive and park, and far less expensive than a tractor trailer system. For many stations, the step beyond the typical ENG vehicle is a cutaway van chassis with a cube-shaped box attached to the frame directly behind the driver. The height of the box offers adequate headroom and is commonly about 14 feet long with a seven-foot width. Approximately 100 square feet of useable area is available for work and storage space.

Single-unit vehicles have drawbacks. One is a 29-foot maximum box length and resulting storage space limitations. Because of the drive shaft connecting the engine to the rear driving wheels, the under-the-truck storage is reduced. Belly box storage depth is usually limited to about a three feet, which prohibits its use for long tripods, light kits, cables, and bulkier items.

It may be instructive to contact a local truck dealer for the cost of unmodified vehicles of various sizes. The addition of two or three cameras, one or two VTRs, control equipment, an on-board generator, and possibly a microwave mast and transmitter electronics quickly raise the price. A relatively simple two-camera industrial panel van will probably start upwards of $100,000. On the plus side, this unit will function well as a backup ENG vehicle when it is not involved in other field production assignments.

For somewhat more room, consider a stretch van, recreational, ambulance-type vehicle or straight truck, each of which may be outfitted as a very serviceable RPV for two- to four-camera productions with on-board graphics. A straight truck, with a 29-foot box provides approximately 200 square feet of deck space. Systems based on straight trucks are often used for multicamera sports events, with four cameras, slow motion VTRs, graphics, and a variety of sources. As a rule, they do not include a generator or microwave mast. If shore (on-site) power is not available, a generator can be rented locally.

The weight load and its distribution in the vehicle may require attention. A large concentration of weight over or behind the rear axle may create problems in steering. Ideally, the greatest concentration of weight should be fairly far forward. However, if the weight between the front and rear axles is too great, a tandem axle may be needed to relieve some of the strain. Such an axle must be ordered before the vehicle is constructed.

RPV builders may include a tag-along axle behind or a pusher axle in front of the original rear drive axle. The tag-along is particularly useful in stabilizing the vehicle, but does nothing in regard to weight distributions. A pusher axle improves load distribution and adds stability, while also reducing chassis flex. Modifications to the suspension may include heavier springs and shock absorbers, as well as electrical or hydraulic stabilizers.

The vehicle chassis size is directly related to the number of cameras the unit will be expected to handle. For a complex unit such as a complete 40-foot semi-network type vehicle including multiple cameras, switching and distribution equipment, recorders, titling and effects packages, expect a multimillion dollar price tag.

Interior Designs

The layout of an RPV interior does not provide for a great deal of flexibility. It would be wise to examine some existing vehicles to see what types of layouts might fit the expected requirements. It may also be a worthwhile investment in time to construct an actual-size cardboard mockup of an RPV interior, including control panels, rack layouts, and monitors. Such a mockup is an inexpensive engineering tool to get feedback from the staff and management before holes are drilled and wires are cut. In working with actual-size models, the staff will get a better feel of the proposed environment. Correcting a mistake in cardboard is much easier and less expensive than fixing an error in sheet metal and angle iron in the real vehicle. Combined with a logical, documented analysis of needs and wants, the mockup will help to prepare the purchase order.

While not as realistic as the life-size approach, miniature cardboard models can be used to check the fit of equipment into a floor plan. How well this works depends upon the accuracy with which the models are scaled. Computer assisted design (CAD) software offers another method by which various arrangements can be checked for proper fit. The basic floor plan is laid out in one layer of the drawing, a second drawing layer is used to locate racks and control equipment. If each element of the proposed design is created as a separate object, it is a relatively easy procedure to manipulate the positioning of the individual objects.

If an error in judgment or a lack of organized input to the builder results in alterations that must be made to the finished RPV, the minimum cost of the changes will be nearly threefold compared to the cost of doing it right the first time.

Equipment and people working in an enclosed area operate more reliably with a comfortable environment. When the truck is parked in an asphalt lot outside the stadium on a hot Sunday afternoon, heat buildup inside the vehicle can be considerable. Proper insulation and an air conditioning system designed for the heat load will keep the staff and the equipment at their best performance level. Formulas are available in texts and handbooks on heating and air conditioning for calculating the optimum capacity for an air conditioning system, given the wattage rating of the equipment, the number of people that will be in the area, insulation rating, outside temperature, and other factors.

MAKING IT MOBILE

In the case of the large production van, some arrangement must be considered to transport the vehicle from place to place. The cost of leasing a semi-tractor varies with the length of the lease period, with long-term contracts being more economical. The lease plan usually includes normal and emergency maintenance and repairs, licensing, and taxes required for the tractor.

To purchase a semi-tractor, plan on a price tag over $100,000 for a new unit. A serviceable, used tractor will probably carry a cost about half that of the new one, depending upon age and condition. Remember to add the weight of the trailer and all equipment and supplies when specifying the size and load capability of the tractor. If the vehicle will be used over a wide area, considerations may include a tractor with sleeping berth for long-haul trips. Keep in mind, however, that reliability is a key. If the production system is to be in position at a specified time on a given date, it must be there!

An additional expense in the operation of a semi-tractor is the driver. Reliability and experience are very important. Requirements for drivers of network vehicles include a spotless driving record and experience with all types of driving conditions in all four seasons and in all areas of the country. Because the driver is in charge of a multimillion dollar package of equipment, a very professional attitude toward the job is important.

The Department of Transportation also has requirements for drivers of vehicles with a gross vehicular weight of more than 10,000 pounds. DOT regulations involve a record of duty status report to be kept by the driver. (The records to be kept may be found in DOT regulation Title 49 CFR, Section 395.8.) Requests by DOT inspectors to inspect the record could result in substantial monetary fines and a forced out of service period for a driver if rules have been violated. In some areas, additional local requirements more strict than those of the Federal Government will also apply. The driver must know all such rules for areas where the truck will travel and must meet medical physical requirements specified by the DOT.

ALTERNATIVES: TO LEASE OR BUY

If there are uncertainties about the field production activities or the amount of expected use is limited, leasing presents a practical alternative. With a leased vehicle, production concepts can be checked out. Feasibility of remote activities for the station can be proven. Then, if the station later determines that the purchase of a vehicle is the proper step, valuable experience will have been gained toward specifying the purchase.

If the station owns its vehicle, it is responsible for continuing equipment maintenance, vehicle upkeep, and financing. Any time the system is out of service, it is a financial liability, so the need to keep all parts operating is essential. Under a leasing program, most major maintenance is handled by the leasing organization. The availability of maintenance facilities and the maintenance program offered should be important criteria in the selection of a leasing company.

The cost of leasing will vary with the equipment package desired, the size of the system, and with the area of the country. Many of the plans arrive at a total from a per-item equipment schedule. The agreement can be written to include a complete operating crew for the required number of hours per day.

The purchase of a turnkey production truck is expensive, even for the simple ENG super van configuration. If the purpose is strictly ENG, perhaps a four-wheeled vehicle would provide the solution. With some protective storage for the camera, recorder and other essentials, and a microwave system, there should still be enough room for a camera operator, microwave technician and the reporter. Such a project might be easily handled by members of the engineering staff.

As the size of the vehicle increases, so does the complexity of the installation and the significance of the decision whether to do it yourself or buy a turnkey system. Outside manufacturers usually build trucks following a general design that reflects their experience and user feedback. Input from experienced station personnel will be valuable for whomever builds the vehicle and helps to determine a configuration that maximizes the vehicle's benefit, usefulness, and practicality, while remaining satisfying to the buyer.

If the plans include live remotes, the production vehicle will probably require microwave equipment to feed signals back to the studio. In addition, some plan should be made for two-way communications for cueing purposes. Fortunately, there is a range of choices, including cellular telephones, subcarrier transmissions, and mobile radio. As the distance between the remote site and the station increases, the communications selections could include a channel on the microwave system or even a connection through a mobile satellite facility. Another approach to get the program back to the station could use an established common carrier network.

In most large metropolitan areas today, the use of RF communications channels, particularly in the microwave bands, also means dealing with problems of interference. An extensive frequency coordination program has been instituted by the Society of Broadcast Engineers. Chapter 1.3 in this *Handbook* provides an in-depth discussion of the problems and procedures of frequency coordination.

LEASE-OUT: AN AID TO FINANCING

The status of owning a vehicle and the convenience of not having to arrange for a production truck every time one is needed are plus factors. If the station's production schedule does not keep the system constantly in use, then it could be leased to others as a means to help in financing the purchase. Planning for this arrangement should include answering some additional questions.

1. Will the equipment complement make the truck desirable to others?
2. Given the station's planned uses of the vehicle, what is the likelihood that the station might need the vehicle when someone else is using it?
3. Will the vehicle need a staff when on lease-out assignments? Depending upon the experience of the lessee, it may be advisable for the owner-lessor to provide some experienced staff to help ensure safe and effective operation.
4. How much effort would be required of the station's staff to prepare the vehicle for use by others and how much effort would be required to service and test the systems when the vehicle is returned?

PLANNING THE EQUIPMENT PACKAGE

Shopping for remote production equipment is as critical as buying for the studio. The right purchase includes several specific criteria:

- Quality
- Reliability
- Operational simplicity
- Maintainability
- Compatibility with other equipment used by the station
- Price

Often equipment suppliers make special package offers, which provide significant cost savings over the purchase of individual items. While a program of getting bids on equipment packages takes time and may involve a good deal of paperwork, the process can result in a price tag that will allow an extra or two to be included.

The following sections include points to consider in various equipment areas. Most of the comments are directed toward portability. If a facility has already had experience with field production equipment, that experience will no doubt be a primary factor in future equipment purchases. Past experience could also be a reason to use an entirely different line of thinking. Applications of the decision spread sheet can prove quite useful in this stage of system planning.

Acquisition Equipment

If the station is already active in ENG, has a major investment in the equipment of a single format, and has experienced good results with that equipment, then the purchases for additional field production units should stay with the same format. Where several systems are in use, the greater the similarity among the cameras and recorders, the easier it will be for operators to switch from one system to another, as the need arises. The more familiar an operator is with the equipment, the greater the effort the operator can put toward visual concerns. The performance advantage of a new format may be outweighed by such practical operational concerns.

An initial evaluation of equipment, based on published specifications, is another opportunity to use a spread sheet approach to compare information about different models. List the models as column heads, let the rows be the more important specifications. Numbers can be filled-in from manufacturers' literature. The annual *Equipment Reference Manual* published by *Broadcast Engineering* uses this approach for a number of product areas and may prove useful. Table 1 is an example of items of interest for cameras.

TABLE 1
Spreadsheet to compare NTSC cameras.

Attribute	Model 1	Model 2	Model 3
CCD type & pixel array			
Sensitivity & max. sens.			
Resolution depth of modulation			
S/N Ratio			
Output format			
Adaptors: NTSC composite RGB<R>Y/R-Y/B-Y R/C (S-VHS)			
Cables: Multicore, Triaxial Coaxial, Fiber optic			
Docking VTR			
Control units: RCU, LCU, CCU Multicamera CU			
Camera Head Weight			

Field Cameras

Selection of a portable camera for ENG has been simplified by improvements in charge-coupled device (CCD) technology. For practical reasons, few cameras with pickup tubes are available today. The CCD camera is generally smaller, lighter weight, and more rugged, while the quality of the pictures generated by CCD devices nearly matches that of tubes. It is unlikely that viewers will be able to tell the difference. The image format has settled almost entirely on the half-inch size with a liberal array of automatic controls. Chapter 5.2, "The Broadcast Television Camera System," includes a detailed discussion of CCD types, noting their operational characteristics and capabilities in producing the sharpest picture with motion.

Two areas are left open for discussion: the type of output and the lens. Many of the current generation of cameras are designed to mate with a series of adapters. The adapter determines the output signal format: composite with multicore, coax, triax, or fiber optic cable; luminance and color components; Y/C components; and adaptations for docking recorders in Beta SP, M-II, S-VHS, or Hi8. A range of camcorder systems adds to the possibilities.

The trend in lens designs meets the special characteristic of CCD cameras. Because CCDs are cemented to the beam splitting prism and cannot be moved individually for fine focus, as was the case with tubes, the lens system must be optimized to compensate for chromatic aberration. While an uncompensated error may be visible as a slight fringing of color, lenses designed for CCDs bring all three color images into focus at the same plane.

Lens length may be partially based on a matter of personal choice and partially on the type of event to be covered. Lenses with a greater zoom ratio permit greater flexibility in camera location for coverage of some events, but one characteristic of long lenses should be kept in mind: the longer the lens focal length, the more pronounced are even minor movements. As lens focal length increases, the need for a more stable support than the operator's shoulder may be required.

Another factor in camera selection may be of importance particularly to camera operators who rove through crowds or shoot video at sports activities. Several special shapes of the camera package are designed for better operator vision. Standard camera designs create a blind spot to the right of the operator. A more ergonomic design reduces the height of the camera body in such a way that the operator can see toward the right while shooting. Although the appearance may seem odd at first, many operators come to appreciate the design feature after they are acquainted with it.

Operators should be aware of the relative levels of illumination in which they are shooting. Most CCD cameras can produce usable images to quite low levels. The use of a shutter with CCDs has become common place to avoid smearing in the pictures as a result of highlights or difficult contrast conditions. Integral neutral density filtering and an automatic iris are important in coping with widely differing or rapidly changing levels of light. For reference, the following Table notes several common lighting conditions and the relative light levels that may be encountered:

TABLE 2
Light levels of common lighting conditions.

Full sun	100,000 lux
Overcast at noon	10,000 lux
Interior by wind	1,000 lux
Interior work area	100 lux
Full moonlight	0.2 lux

Nearly all portable cameras include two or more levels of color correction filters to assist in color balancing for various light sources. These filters, along with color balance memories, can be used to produce realistic color in the output pictures.

Field Recording

Field recording formats include the half-inch analog component Beta SP, M-II, and S-VHS lines; the composite one-inch D-2 and half-inch D-3 digital formats; and high band 8mm (Hi8) products. Making a choice among these presents a real challenge, because all of them perform quite well. Beyond the problem of compatibility with existing equipment, consider how the product will be used. If the function will be only ENG, where limited generations of the material are required, analog component and Hi8 products may be the most cost effective. If the material acquired in the field will go through multiple generations in post production, the composite digital format is suggested. Additional guidance on recording equipment will be found in Chapter 5.6, "Video Tape Recording."

ENG cameras typically provide for microphones to be mounted directly on the camera body with a connector tied into a program audio circuit. The purpose of the video acquisition will decide the type of microphone pattern to be used. For shoots that involve general audio coverage, an omnidirectional or general cardioid pattern will serve nicely. If it is necessary to pickup only specific talent on the audio track, a shotgun unidirectional instrument will prove a better choice. This feature can avoid the problems of microphone cables at the remote site. Many instances will still require the talent to use a handheld microphone for best vocal pickups. With a miniaturized wireless microphone receiver attached to the camera, a wireless handheld or body pack microphone solves the cable problem. Wireless operation does have the drawback of possible interference from external sources.

When multiple microphones are used, a need for some audio mixing capability exists. If the portable recorder includes multiple audio channels, mixing may be deferred until post production. Much more freedom and creativity is available during post production work when separate audio tracks are recorded. The subject of microphones is discussed in greater detail in Chapter 5.1, "Microphones."

Cables and Links

More a concern for EFP and OB projects is the question of getting the signal back to the truck. For cameras that operate on the various types of cables, the choice must be made between multicore, coax, triax and fiber cables. Fiber optic options permit the greatest distance between the camera and the production vehicle of the four types. If distances are greater than those suggested for a single fiber optic cable run, relay units are available to double the length. During the construction of some stadium facilities, the needs for television are often included in the design. Cabling may be permanently installed to various locations in the stadium, reducing the setup time required for the TV crew.

Cableless cameras may be outfitted with a miniature microwave system to beam the output to the central production center. These systems may include a tracking antenna, which permits a return channel for communications between the truck and the operator. If the line-of-sight path is blocked, a relay becomes necessary. Such a relay in an open stadium setting might use an antenna (and electronics) held aloft by a balloon attached to the upper, outer wall of the stadium.

The relay unit remains relatively fixed-positioned with respect to the truck parked outside the stadium, while the tracking system in the camera unit keeps the primary transmitted energy aimed at the relay antenna.

GETTING BACK TO HEADQUARTERS

Live remotes bring together every complexity of the television broadcast industry. Not only must the remote truck operate perfectly, so must the program link from the remote originating site to the base station. Technology has produced a variety of methods for transporting the program signal back to the station. Options include wired connections (via telephone company facilities), microwave video radio (ENG), and satellite relays. The distance between the production site and the station can be a deciding factor in the transmission choice.

Terrestrial Microwave

If the site is local, for example, the most expedient link between the two locations is probably by microwave, especially if line-of-site operation is possible. If obstacles to a direct line path exists, perhaps the use of a reflector can redirect the signal around the obstacle. Microwave system design is detailed in Chapter 4.2, "Microwave Engineering for the Broadcaster."

Although microwave frequencies may already be assigned in a metropolitan area, a check on frequency coordination is advisable to make certain that interference can be avoided. Often cooperation between stations is needed to ensure interference-free operation for special events. A case in point is the renowned July Fourth celebration in Boston, during which a live performance of the Boston Pops Orchestra is telecast nationwide by PBS. In addition to the site-to-station channels, local site microwave between boats, helicopters, nearby buildings, and the production truck are in use. Because of a limited number of frequencies, the production is made possible only through cooperation between the various local stations in frequency coordination and equipment sharing. Coordination activities for such scheduled events typically begin months in advance.

Another approach to terrestrial microwave operation was the solution for an all-news television format in Boston. An extensive regional network of microwave transmission and reception facilities was constructed to capture live coverage from the region extending south to Connecticut and north to New Hampshire. On special occasions, the reach of the system was extended to New York City. Signal transmissions to WNEV-TV (now WHDH-TV) from four established news bureaus locations in Massachusetts came through five relay points, which also collected signals from mobile units and helicopter facilities. A mixture of 2 GHz, 7 GHz, and 23 GHz microwave was used in constructing the system.

As distances between the remote site and base facilities increase, microwave by common carrier services becomes practical. Demands for more and more communications channels has lead to the use of digital transmissions to accommodate multiplexing as well as bandwidth reduction through data compression. Several services are now offered using the technology of digital transmissions with DS0 through DS3 channels. Switched 56 service, which extends to some 400 metropolitan areas in the United States covered by the telephone company, is another convenient approach to signal transmission over long distances. The service provides a bidirectional 56 kilobits/second data path through dial-up terminals.

Satellite Backhaul

Program relays by satellite permits the remote site to be almost anywhere in the world. Many of the relay and backhaul services available by satellite also involve digital signal carriage with data compression to help deal with the amount of traffic that must be transmitted at all times. Depending upon the destination for the signal, the satellite system may be the more practical route to be used. In most larger urban areas, occasional-use satellite time can probably be obtained on short notice and is competitively in price with fiber.

For remote operation, the development of smaller Ku-band terminals and antennas offers several advantages. The size of a C-band antenna was too large to easily include an uplink antenna as a part of a remote production vehicle. As a result, such C-band equipment was often taken along to the remote location on a separate vehicle. Ku-band uplink equipment can actually be mounted on a production vehicle and immediately deployed upon arrival of the truck at the site. An option provides the convenience of almost completely automatic satellite antenna alignment through the LORAN navigation system. The engineer needs only to key into the control computer both the geographical coordinates of the site and the satellite through which the relay will pass. Within minutes the system is setup and operational.

Fly-away C-band and Ku-band equipment is available with folding dishes in the two-meter diameter range. Electronics, antenna, and mount can fit into a half dozen or so suitcases that can be checked as airline luggage. Equipment manufacturers and the operators of satellites and commercial teleports can provide technical and cost information on which to base planning decisions for the use of such facilities. With the growth of switched fiber networks, the available combinations of satellite, fiber, and terrestrial microwave links offer almost unlimited possibilities for live remote production anywhere in the world.

COMMUNICATIONS

Keeping in touch with the remote production crew is one of the secrets of a successful production session. Even for a simple ENG assignment, where only a camera operator is dispatched to gather local color for a special closing segment, communications between the operator and the station should not be forgotten. It may be necessary to direct the operator to another assignment before returning to the station or the operator may require assistance from the station before returning. As the production project becomes more involved, communications becomes even more important. When ENG goes live, cues to the technical staff and talent are as crucial as those to a remote team located halfway around the world.

Options now available for keeping in touch with mobilized members of the staff are widely varied. Cellular telephones and two-way radio provide convenient bidirectional links between the control room and the remote van over a wide area of operation. Both require licensing. Cellular telephones can also be used to transmit data between two points, including facsimile type data transmissions. Just as facsimile messages have become popular throughout modern business, the remote copier capability is particularly valuable in sending graphic type material that would require much more complex equipment for handling.

When the remote production crew leaves the area served by the TV station, two-way communication with the station can be carried along with program signals by microwave or satellite through an order wire subchannel of the existing connection. The roaming feature of cellular telephones can keep a remote crew in almost constant contact from most areas of the United States and Canada. Undoubtedly this will expand to a worldwide connection, as satellites play a greater role in linking terrestrial communications facilities.

In many cases, two one-way paths may already exist. The station that is transmitting multichannel television sound (MTS) stereo signals is also authorized to carry a second audio program (SAP) subcarrier, as well as a professional (PRO) subcarrier. Of the two, the SAP signal was originally intended to be related to the program being aired at the time. One suggested purpose was to provide audio in a second language. In many areas, that capability remains unused, and could be investigated as an outward bound signal path for production purposes. The PRO signal was designated as a means for technical communications from the station to remote crews. In most instances, the PRO facility can be used for technical communications as well as cueing through an interruptable foldback (IFB) signal. The return path can be added to ENG or EFP microwave as a subcarrier. In both cases, once the license for the primary service is established, adding the extra subcarriers is a matter of appropriate generators and receiving equipment without any special requirement for frequency coordination. Additional information regarding SAP and PRO subcarriers is covered in Chapter 3.6, "Multichannel Television Sound."

Until the spring of 1991, some concern was expressed by the lack of privacy provided by a SAP or PRO channel. Technically, anyone with equipment to receive the station's main carrier could add subcarrier receivers to access the secondary signals. In a metropolitan area where a good deal of competition between stations usually exists, the use of the PRO channel was slow to catch on, because other stations could eavesdrop. In early 1991, the Federal Communications Commission approved the use of digitally encoded transmissions on the subcarriers. Such encoding in essence scrambles the intelligence being carried on the subcarriers.

Communications among the staff at the remote site may mean another challenge for engineering. Operators of wired cameras remain in touch through the intercom circuit available with all cameras. Roving, nonwired camera operators are generally equipped with two-way radios or wireless intercom equipment. In a facility such as a stadium, where similar communications are required on a frequent basis, numerous wired systems may be permanently installed and staff members are required only to be plugged into the network.

These are only a few of the equipment areas that must be carefully considered before the purchase order for a remote production vehicle and equipment package can be finalized. You will also need to consider video switchers, audio mixers, on-board audio and video recording equipment, signal distribution and monitoring products. No less important to the comfort and efficiency of the operations staff are such concerns as the furnishings in the vehicle.

CONCLUSION

The purpose of this chapter was to provide an overview of the field production activity. A number of topics have been noted here, with the hopes that, as new or expanded remote capabilities are developed, these comments will serve as reminders of areas that should be considered to make remote productions successful. Readers will also find many useful subjects in the material presented in the preceding chapter, "Radio Field Production."

BIBLIOGRAPHY

From *Broadcast Engineering* (Intertec Publishing Corporation):

"Building an ENG Network," Karl Renwanz, Jan '86, pg. 26–42.

"Intercom System Design," Bob Tourkow, March 1986, pg. 74–96.

"Satellite Uplink Trucks," Brad Dick, Jan '86, pg. 44–62.

"Planning A Remote Production Vehicle," Ned Soseman, Jan '87, pg 68–74.

"Taking to the Skies," Steven D Hardy, Jan 1987, pg. 76–80.

"Mobile Mast Safety," Richard Wolf, Jan '88, pg. 26–32.

"Applying Cellular Technology," Brad Dick, Jan '88, pg. 36–44.

"Catching the News Crew," Ron Smith, Jan '88, pg. 46–56.

"Powering Portable Production," F. Kurt Shafer, Jan '98, pg. 58–78.
"Planning an STL system," Ronald Balonis, October 1988, pg. 116–138.
"Making the Most of Remotes," Rick Lehtinen, Jan '90, pg. 38–48, 99.
"Remotes Revisited," Skip Pizzi, Jan '91, pg. 26–46.
"Communicating with the Field," Rick Lehtinen, Jan '91, pg. 48–60.
Equipment Reference Manual, (See most recent annual issue).

6.3 Telephone Network Interfacing

Steve Church
Telos Systems, Cleveland, Ohio

INTRODUCTION

From earnest political talk shows to raucous morning zoos, listener involvement via telephone is an important programming element at many radio and television stations. When it is desired to create a two-way connection with listeners, a studio interface with the public switched telephone network is necessary.

As well, radio news departments rely extensively on "phoners" to get reporters and newsmakers on the air in a timely fashion. Why are the people who run local TV news so concerned with avoiding the dread "talking head" — that is, the anchor simply reading a story into the camera? Because they have discovered that *being there* is better. The same is true for radio.

And, the telephonĕ network, being accessible from nearly everywhere, is routinely used for remote program origination.

In this section, we'll explore all of these ways to integrate the ubiquitous telephone network into broadcast operations.

DIAL-UP NETWORK BASICS

While it is not necessary to know everything the telephone engineer knows about the telephone network, it is helpful to have a general understanding in order to make the best use of it in the broadcast environment. Let's begin with the standard *subscriber loop*.

Subscriber Loops

The telephone line pairs provided by the phone company are known officially as subscriber loops, *trunks*, or simply central office (CO) lines. (Actually, "trunks" used to refer only to lines destined for private branch exchanges (PBX), and may have included special signalling as well. Now, though, with even the smallest subsciber having a PBX, almost every business phone line is a trunk.)

Talk Battery and Ringing

A DC voltage called *talk battery* and the conversation audio appear together on each phone pair. The DC leaves the exchange at −48 volts and is limited to 20 to 100 mA by a series resistor. The resistor's value is selected to complement the resistance of the loop. The DC resistance of the loop itself varies from a few to 1,300 ohms depending on length. (A little known fact: a garden-variety standard "Model 500" telephone set incorporates a varistor which, by sensing loop current, automatically adjusts voice level to compensate for various loop lengths.)

Ringing causes an AC voltage of 90 VRMS at 20 Hz to be superimposed on the line. Talk battery is maintained during ringing so that the resulting potential has a sine shape shifted negative by 48 volts.

Parameter	Typical U.S. Values	Operating Limits
Talk Battery Voltage	-48 VDC	-47 to -105 VDC
Loop Current	20 to 80 mA	20 to 120 mA
Loop Resistance	0 to 1300 ohms	0 to 3600 ohms
Loop Loss	8 dB	17 dB
Distortion	- 50 dB	N.A.
Ringing Signal	20 Hz, 90 VRMS	16 to 60 Hz, 40 to 130 VRMS
Noise (objective)	- 69 dBm0 to 180 mi, - 50 dBm0 to 3000 mi (-16 dBm0 talk level) (C msg weight)	

Figure 1. Phone loop characteristics.

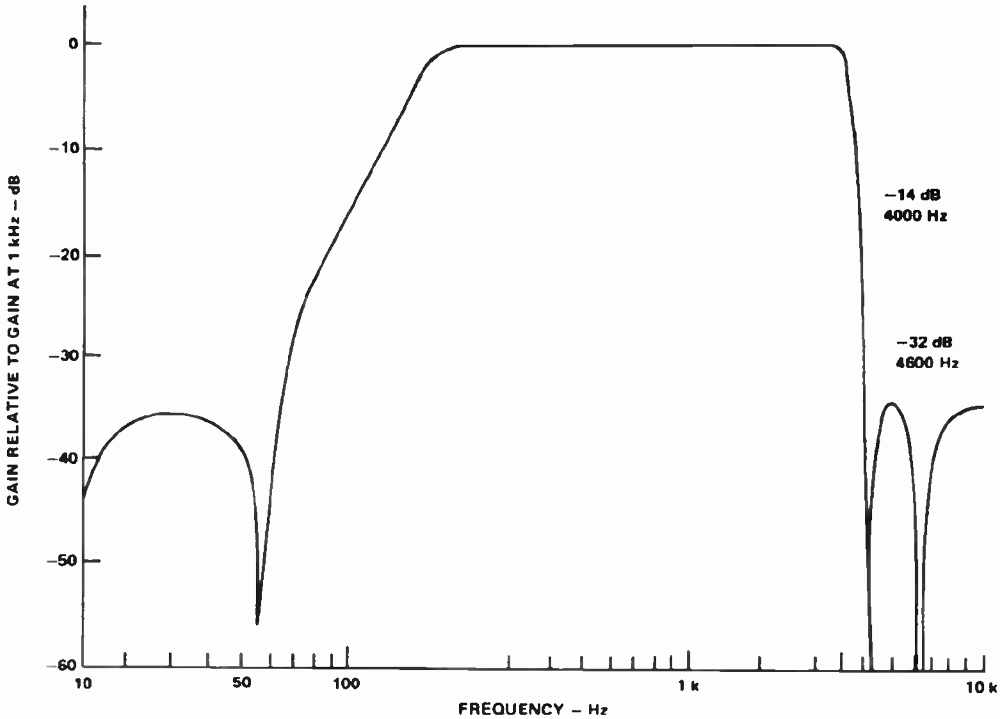


Figure 2. The low-pass filters required for digital transmission limit frequency response. This response curve is for a CODEC which is widely-used in the telephone network. (Note also the significant low frequency roll-off.)

Frequency Response

For ordinary subscriber loops, the phone company specifies a frequency response of 300 to 3400 Hz. In the not too distant past when all local calls were connected at the exchange by metallic contacts, better frequency response was likely to be had on many conversations. Long distance calls, however, have almost always gone by the way of band-limited microwave or satellite links so frequency response on them has never been better than the promised 300 to 3400 Hz.

Most telephone calls are digitized at some point in the transmission path. As described below, such calls are *strictly* limited to a 3.4 kHz bandwidth by sharp low-pass filters.

Noise and Level

A 1971 Bell System survey of the phone network nationwide determined that the average conversation had a level of −16 dBm. Of course, as anyone who has wrestled with broadcast-to-telco interfacing knows, incoming level varies tremendously, with a range of perhaps −40 to −10 dBm.

Send audio (that is, audio fed into the telephone line) must be limited to −9 dBm as specified in FCC Part 68. Audio loss on any given local loop is limited by tariff to 8 dB or less. This loss limit, however, applies only to the loop from the CO to the subscriber and does not include the rest of the signal path. Also, the 8 dB loss may occur at each end of a conversation path: once at the calling party end and again at the called party end for a total of 16 dB.

In telephone engineering, a different approach from that with which we are familiar is used for measuring and reporting noise. The phone engineering people measure noise "upside-down," defining a reference noise floor and then measuring *up* from there. The reference noise level is one picowatt, which corresponds to −90 dBm. Thus, a noise level of −60 dB relative to 0 dBm would be reported as "30 dBrn noise" (dBrn = dB above reference noise). Note that, according to this method, the higher this number, the poorer the noise.

Be aware also that when telephone engineers measure noise, they are measuring only *idle channel noise*. This is an important difference since in digital systems, idle channel noise is not the same as the traditional signal-to-noise measurement in analog systems. That is because noise in a digital system will generally increase when a signal is present. This effect is called *modulation* or *quantization* noise and is primarily dependent upon the number of bits used for quantization.

A *C-message weight* filter is employed when determining phone line signal-to-noise ratio. The C-message curve was developed years ago to simulate the frequency response of an old-style telephone earpiece and, accordingly, it has considerable low-frequency roll-off. This means that a line can have significant hum and other low frequency noise and can still meet

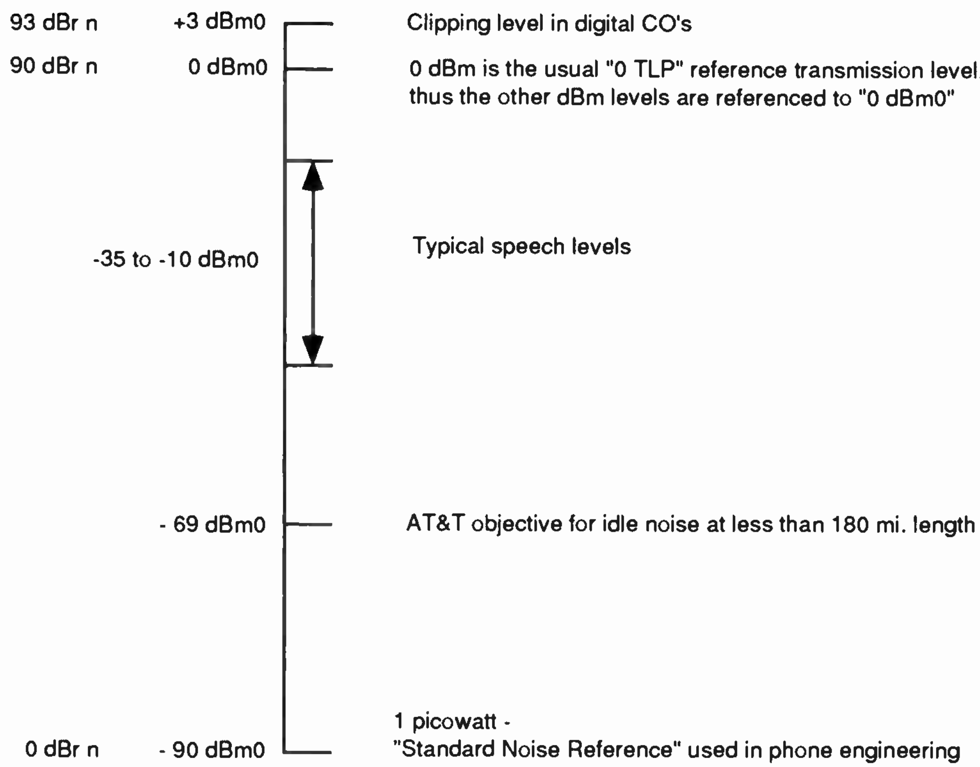


Figure 3. Chart showing signal and noise level references used in telephone engineering.

the officially mandated noise specs. While this makes life easier for the phone company technicians, it can be troublesome when a broadcaster is trying to use phone audio on the air.

If noise is a serious problem, try to get the technician to switch the noise meter to the flat position, without use of the C-message filter. The measuring set usually does have this option available.

DTMF Tone Dialing

DTMF (*dual-tone multiple frequency*) dialing uses two frequencies for each digit in order to avoid "talk-off": that is, the tone detector accidentally sensing voice as a dial command. As well, the frequencies were carefully chosen to avoid problems with harmonic distortion that might otherwise cause false detection. There are four "low group" frequencies, one for each button row, and four "high group" frequencies, with one assigned to each column. Tolerance is $\pm 1.5\%$ for the encoder and $\pm 2\%$ for the digit receiver. The time required to recognize any digit tone is 50 ms with an interdigit interval of another 50 ms. Low group tones are supposed to be sent at a level between -10 and -6 dBm; ideally, tones in the high group are transmitted with 2 dB greater level in order to compensate for high-frequency roll-off in the phone line.

Loop Start and Ground Start

Central office lines come in two basic configurations, loop start and ground start. Loop start is the kind that is most common. In this kind of circuit, the CO provides talk battery to the line at all times and detects that an off-hook condition is occurring when current begins to flow between the tip and ring. (Incidentally, the names "tip" and "ring" originated with the description of the circuits being on the tip and ring of the patch cords that were once used by telephone operators.) With ground-start circuits, the CO waits for a connection from the ring wire to ground before connecting talk battery, at which time the terminal equipment removes the ground connection to establish a balanced talk path. When the calling party hangs-up, a ground-start circuit removes talk battery. A loop-start circuit may or may not provide a momentary interruption or reversal of the talk battery when the calling party terminates.

Many PBXs are designed to work with the ground-start circuits because the possibility of collision is reduced, where the phone system tries to seize a line for an outgoing call just as that line is ringing-in.

Disconnection: Calling Party Control

Loop-current interruption occurs on most telco lines when the calling party hangs up. It is sometimes referred to as CPC, or *calling party control*, since the *calling party* controls *your* equipment when he hangs up. The CPC may turn off an answering machine, for example, or extinguish the winking light on a held line on a key phone. The CPC interruption was probably never intentional, having been a by-product of early

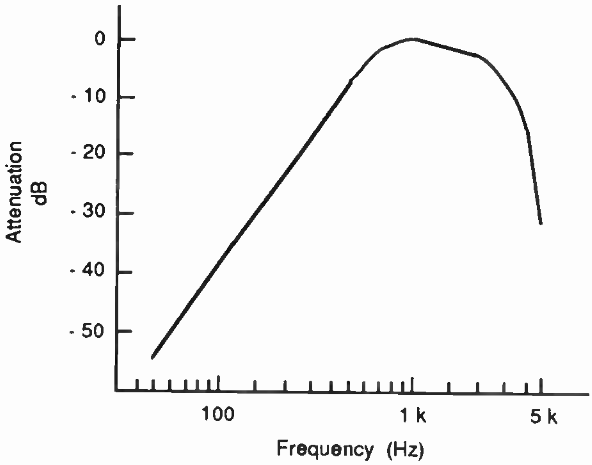


Figure 4. C-message weight frequency response curve.

mechanically-switched relay-controlled exchanges. Thus, some phone lines do not provide this function or they provide it unreliably. However, with the proliferation of answering machines which rely upon CPC, most central office equipment now has this capability designed-in. In some cases, it is necessary to specifically request this feature from the phone company on a per-line basis.

Loop-current *reversal*, on the other hand, has long been a phone company signalling method. First used between the telephone company's central offices, loop-reversal was later employed to communicate with some large organizations' PBX systems. Thus, lines which are set up for PBX use, or originate at central offices with large concentrations of business customers, sometimes use this method. (However, the preferred and more modern situation for PBX control is to use either ground-start lines or *E&M signalling*.)

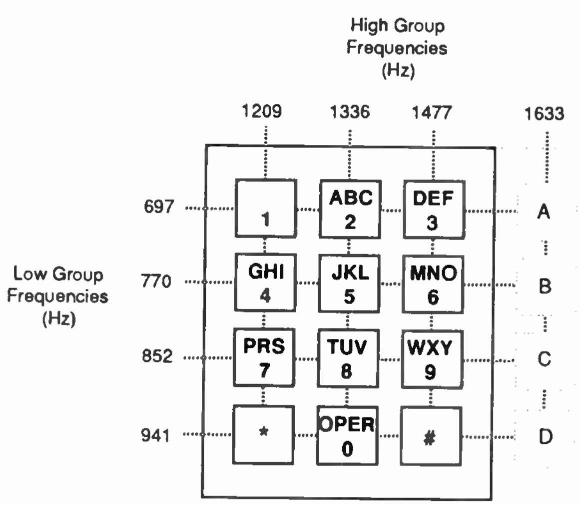


Figure 5. DTMF tone keypad frequency assignment. The four tone pairs in the last column are for special applications.

While most exchanges do provide CPC, there are some that do not reliably provide it, or provide it after a variable time delay; and most PBXs don't generate it. There is but one certainty: every (USA) telco central office eventually returns dial tone to its lines when the calling party hangs-up. Thus, we can use the presence of dial tone as a back-up to insure a disconnect will occur even when the loop-current detection methods may fail. An important consideration is to prevent false "talk-off" from noise, applause, or other spectrally-rich audio. One way to accomplish this is to use software-based statistical methods to ensure that the dial tone is *really* present before terminating the connection.

Loading Coils

A typical #24 gauge phone pair attenuates a 3 kHz signal 2.5 dB per mile due to losses of high frequency from capacitive effects. On an eight mile line, high frequency attenuation would thus be 20 dB, a significant amplitude distortion. Loading coils are toroidal inductors which counter the effects of the phone pair's natural capacitance. While the coils are effective at flattening out the response within the voice band, there remains very significant roll-off above 3.5 kHz.

Physically, load coil banks are long cylinders, with the individual donut-like coils stacked one on top of the other inside. They are typically placed at 3,000, 4,500, or 6,000 foot intervals along the phone cables. Generally, loading coils are found only on cables of greater than three miles length.

Loading coils also usually adversely affect the performance of hybrid interfaces since the coils can create resonant peaks and phase anomalies in the phone line's impedance curve which are difficult to compensate for. This is a reason why simpler hybrids are more likely to operate satisfactorily when your station is physically near the serving CO.

Two-Wire and Four-Wire

The typical telephone circuit is *two-wire*, so named because it arrives on two wires. A *four-wire* circuit uses a separate wire pair for each of the send and receive transmission directions: four wires altogether.

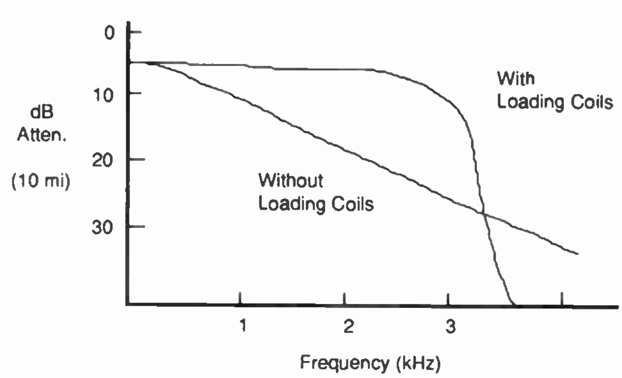


Figure 6. Frequency response with and without loading coils.

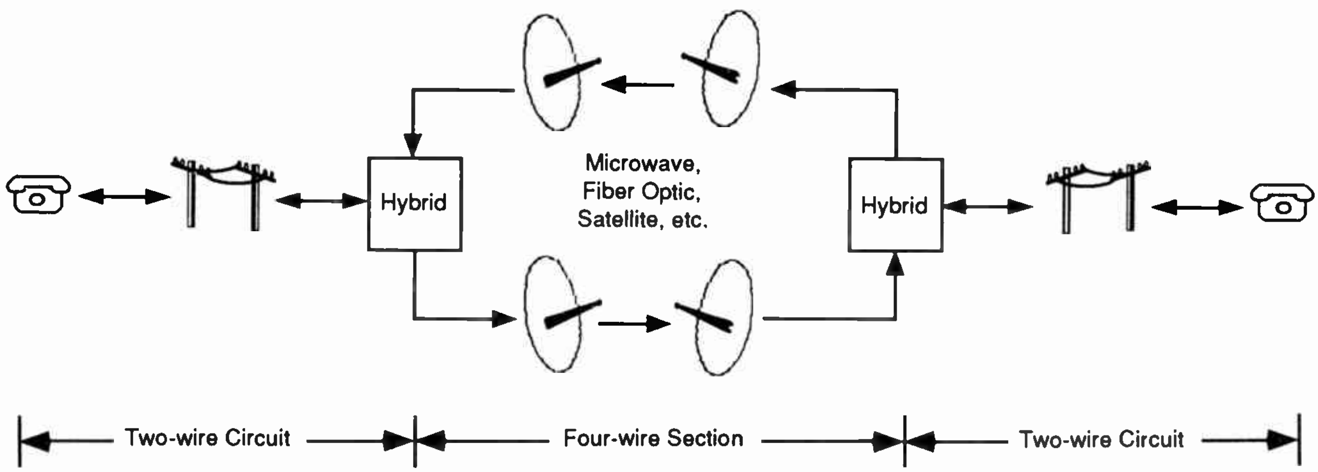


Figure 7. Two-wire circuits have both audio directions on a single pair of wire which are separated for switching and long-distance transmission into four-wire signals with hybrids.

Four-wire circuits are required for long-distance transmission because it's the only way to get sufficient gain without feedback *singing* on long paths. Also, noncopper transmission media such as microwave radio, satellite, and fiber optics are one-way only. So, the two-wire circuits familiar to us are converted before transmission to four-wire using hybrids (discussed in detail later). Almost all central office switching is done with four-wire.

Foreign Exchange (FX) Loops

FX provides local telephone service from a central office which is outside (foreign to) the subscriber's exchange area. If a station is located in the suburbs and the choke network central office (described below) is downtown, FX loops will be needed to connect your lines. When the phone is picked-up, you get the dial tone not from your local suburban CO, but from the downtown office. Foreign exchange service is also sometimes used to extend your coverage into another city, so that people can call the station without paying a toll charge and, as well, calls can be made within that city without incurring toll charges. For instance, if the studio is in Cleveland and the goal is to serve listeners in Akron as if they were local, FX service could be helpful to creating a local presence in Akron.

An FX loop is a four-wire circuit with hybrids at each end or terminating central office. Since FX loops add an extra layer of hardware to the phone audio, they are another source of problems for on-air interfacing. They usually are engineered to have a few dB loss (to prevent hybrid singing, as described later). And, they add to the impedance complexity of the line.

FX circuits are usually expensive, and pose certain technical challenges. FXs have separate four-wire send and receive paths with hybrids at each end. Since, as we will learn later in this chapter, hybrids are imperfect, a potential for a special kind of feedback called "singing" exists. This results from the inevitable leakage from the send to the receive ports at each hybrid. How do the phone people solve this problem? By inserting a pad: anywhere from 5 dB to 8 dB is common.

As optical fiber replaces copper for interexchange connections, FX service will likely pose fewer problems.

Choke Networks

Most stations need special high volume exchanges for contest and request lines. The choke network works by diverting calls beginning with the unique choke prefix around the local serving central office and sending them directly to the choke switching exchange, usually located downtown. The phone company dedicates very few talk paths (wire trunks or special carrier equipment) to the task of connecting the caller's serving CO choke ports to the choke exchange. The usual switching and routing process is by-passed. Unfortunately, only a very limited number of paths are generally provided. In the densely populated Los Angeles area, for instance, only three connections exist from most central offices. Often, as well, the poorest facilities are given over to the high volume service.

Generally, unless you are near the choke central office, the FX circuits described above are employed to connect the choke CO to your serving CO. This is one of the reasons why choke lines often have a lower level than standard lines.

Because of their higher complexity, choke lines usually have bumpier impedance curves as well, making good hybrid performance difficult to achieve due to the problem of finding appropriate balancing network values. This is especially a problem with simple analog hybrids.

In some areas, the FX circuits are being replaced by internal call-forwarding. This means that a published number is actually being software forwarded to a "real" number originating from your local serving CO. The

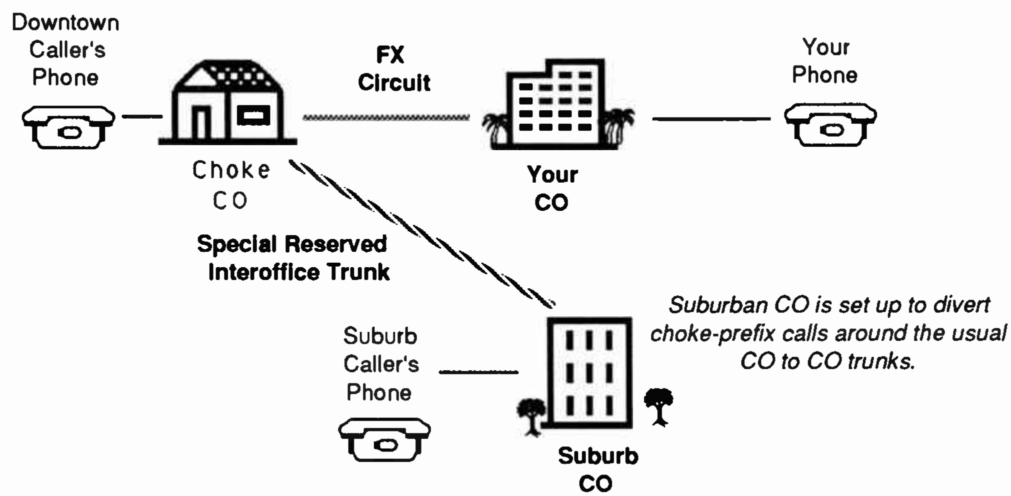


Figure 8. Typical "choke network" transmission path.

main advantage to this approach is lower cost, since you don't have to pay a premium for an FX circuit. However, there usually is a smaller call-forwarding charge.

Digital Telephone Systems

In 1962, Illinois Bell installed the first digital telephone transmission system and the era of modern digital telecommunications began. This, the first ever T-carrier system, was the first widespread commercial application of digital audio. Telephone engineers like digital technology for the same reason broadcasters do: the potential for reduced susceptibility to noise and other disturbances, as well as improved ability to switch, monitor, and maintain.

PCM and Speech Coding

Speech audio in the telephone network is encoded using the common *PCM (pulse code modulation)* approach. The speech is sampled at an 8 kHz rate and is represented in 8 bits. Audio is converted to and from the digital domain using specialized analog-to-digital (A/D) and digital-to-analog (D/A) devices called *CODECs* (CODer/DECoders).

Can eight bits represent the wide dynamic range of telephone audio signals? Since, in a perfect digital quantizing system, each bit adds 6 dB to the maximum dynamic range the system can handle, an 8 bit system could contain a dynamic range of only 48 dB. Moreover, at low signal levels, the distortion would rise to an intolerably high level. Indeed, telephone researchers long ago determined that at least 12 to 13 bits would be required for good "toll" telephone quality. The solution was to be found in a bit-reduction technique which utilizes instantaneous companding.

A companding system requires a compressor for encoding and an expandor for decoding. A special scheme called *uLaw companding* is used for telephony, which causes the step sizes to be arranged approximately logarithmically, with the smaller steps being used for lower signal levels and increasingly larger steps for the higher level signals. This means that the quantization noise (and distortion) is approximately a fixed percentage of the signal amplitude, regardless of its level.

Anti-aliasing filters limit frequency response to less than one-half of the sampling rate. Phone CODECs use a sampling rate of 8 kHz so that frequency response is absolutely limited to less than 4 kHz. In practice, filters with steep cutoffs above 3.4 kHz are employed. These are generally the limiting factor with regard to telephone system audio response.

Noise and Distortion on Digital Phone Circuits

The 8/13 bit quantization scheme results in less than high-fidelity. Often, the problem lies in the specific implementation rather than in any inherent shortcoming in the standard or the technology. One important quality limitation results from the anti-aliasing and reconstruction filters in CODECs. These filters usually have an ultimate roll-off of around 35 dB. Audio above the 4 kHz Nyquist frequency will alias and appear in the 300 Hz to 3.4 kHz band as distortion. Thus, typical

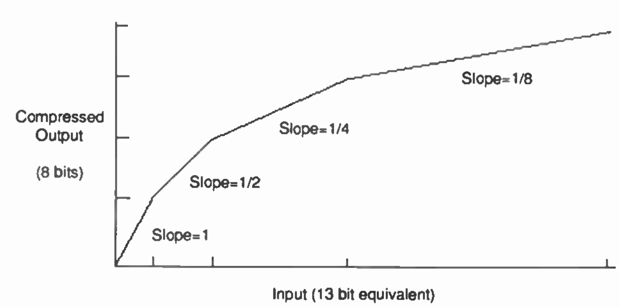


Figure 9. uLaw PCM coding in CODECs causes the quantization noise to be approximately a fixed percentage regardless of level.

CODECs have distortion of 2% to 3% from aliasing. The strange "raspy noise" that seems correlated with the speech sometimes heard on a telephone circuit is a result of the effects of this kind of distortion combined with audible quantization errors.

CODEC filters generally use switched-capacitor technology which tend to be fairly noisy. Some newer, better CODECs avoid the switched-capacitor problems by employing the same *delta-sigma* over-sampling and digital decimation concept used for high performance digital audio conversion.

Centrex

Centrex replaces customer-owned PBX or key equipment with telco central office switching capability. A connection from each phone set goes directly to the CO. The idea is to eliminate user up-front costs and transfer maintenance responsibility to the telephone company. Varying service requirements are accommodated without customer equipment changes. Centrex seems to be especially popular with universities.

One problem limiting the growth of Centrex is the present lack of "feature phones" with multiple feature buttons (LCD displays and the like) as found on almost any other modern business phone system. Features in Centrex rely upon the use of the star and pound tone pad keys in combination with other tone buttons, generally an awkward and confusing situation for inexperienced users. This problem is expected to be solved with the introduction of a new digital phone service called *ISDN* described later in this chapter. ISDN will permit very sophisticated phones to be used on Centrex lines. Indeed, initially, ISDN may be available only on Centrex lines.

Cellular Telephones

Cellular extends the dial-up network to many places where a wire connection would not be considered practical. Cellular transceivers operate in the 800 MHz range and automatically select the appropriate frequency from among the 666 FM channels assigned for this service. Low power is used so that the frequencies can be re-used in adjacent areas. The mobile phone varies its power according to the level of signal received at the base location. A very useful feature for on-air use of a cellular phone is the signal strength meter provided on some units. Some phones also allow you to see the send power value. Often, the antenna's pattern is quite directional due to its position on the vehicle, so moving around while observing the level indication can help to make remotes sound better. For fixed remotes, a Yagi antenna can be used. The user receives the benefit of higher gain and directionality; and at 800 MHz, Yagis are very compact. Calls from one cellular radio to another have better fidelity than those which are transmitted via telco landlines. Try calling from a mobile to another cellphone located at the studio to make this determination.

Most equipment designed for use with wired phone lines can be connected to cellular phones with an

Figure 10. Interface to cellular telephone for broadcast on-air use. (Courtesy Comrex.)

adapter provided by the phone manufacturer. These adapters, intended for lap-top computer modems and portable fax machines, provide an interface to any broadcast equipment, which can connect to a phone line as well. There are units especially designed for broadcast use which have provision for audio input and output for direct connection to mic mixers and the like.

FCC Regulations

FCC requirements for connection of equipment to phone lines are to be found in the Rules and Regulations of the Federal Communications Commission, Part 68: Connection of Terminal Equipment to the Telephone Network. Order these from the government printing office, or visit a library reference room which has the Code of Federal Regulations series.

PBX AND KEY SYSTEMS

Now that we know a bit about the nature of the phone network, we can explore what happens after the lines become "ours." We will want to use some of what the phone people refer to as CPE (*customer provided equipment*). That is, all of the equipment connected to the phone line after the official *demarcation point*. In this section, we will look at the various styles of PBX and key systems available both for general office and on-air use, followed by a look at systems designed specifically for studio application.

PBXs, or *private branch exchanges* are found where there is a need for a large number of extension phones. PBXs are miniature central office exchanges, allowing local phones to call each other as well as being able to access trunk lines for incoming and outgoing calls. PBX systems often have a number of specialized features for call routing and control. Traditionally, PBX systems have used only single-line phone sets as terminals, with special functions like transferring and conferencing being accessed by "flashing" the switch hook or by using the tone pad in a special way. Sometimes, these systems are called PABXs (*private automatic branch exchanges*).

Key systems are generally used where the number of extensions is smaller. Key systems are generally distinguished from PBX systems by the presence of each of the CO lines on an individual button on each extension phone.

Recently, the distinction between PBXs and key systems has lessened. Most PBXs have available feature phones, which can button-access individual lines as key phones do, as well as provide numerous other advanced functions. At the same time, most key systems now have PBX-like features and are expandable to large numbers of extensions.

1A2 Key Systems

While nearly all businesses have gone to the higher-tech "skinny-wire" electronic phones for the business office, many on-air installations continue to rely upon 1A2 key systems. Key systems offer the advantage of providing a direct metallic connection to the CO line. That means that no frequency response error, noise, distortion, or time delay is introduced. Often, these issues are not fully considered in the design of the more complex business phone systems. As well, cost is favorable and full schematics and other documentation are readily available.

Leading from the key service unit (KSU) to each phone is a thick cable with 50 conductors (25 pairs). The *tip/ring* pair carries the telephone audio. As mentioned, these are direct connections to the telco CO lines. The *A leads* tell the key system which lines are in use and also signal a hold condition. Selecting a line causes a connection to be made in the phone set from the A lead to another wire, the *A-common*. The A lead is normally at −24 VDC and A-common is at ground potential, so when a line is selected, the A lead goes from −24 VDC to ground. If the A lead is broken before the tip/ring is disconnected, the system puts the line on hold. The *lamp-leads* light the phone's line buttons with 10 VAC from the KSU's power supply and are returned via the *lamp grounds*.

Circuit schematics and wiring charts showing the wire assignments and other necessary installation information are available from suppliers and some broadcast telephone equipment vendors.

Electronic Telephone Systems

Electronic phones can be recognized by the much smaller two to eight-conductor wiring used to connect them. Because the cable is smaller and less expensive, and complex features are easier to implement, these systems are universally preferred for general office application.

While the systems are tremendously varied, most have several features in common. The cable from each phone set to the common equipment, must convey:

1. Power to operate the phone.
2. A two-way data path to signal user actions from the set to the switch, and operational and display status from the switch to the set.
3. The speech audio.

We'll now briefly explore each of the usual approaches for electronic phone wiring and communication.

Separate Pair per Function

The early electronic phones used a separate pair for each of the three functions, and thus required three (or more) pairs. The AT&T Merlin system, for instance, still uses this design. The center pair is the audio. Another pair is for the serially-transmitted control and display data, and another handles the phone's power requirements.

Two-Pair, Phantom Power

The most common approach used for electronic phones is the two-pair, four-wire scheme. The AT&T

Pin #	Wire Color	9 Line 1A2	5 Line 1A2
26	WHITE/BLUE	Line 1 tip	Line 1 tip
1	BLUE/WHITE	Line 1 ring	Line 1 ring
27	WHITE/ORANGE	Line 1 A	Line 1 A
2	ORANGE/WHITE	A circuit common(gnd)	A circuit common(gnd)
28	WHITE/GREEN	Line 1 lamp ground	Line 1 lamp ground
3	GREEN/WHITE	Line 1 lamp	Line 1 lamp
29	WHITE/BROWN	Line 2 tip	Line 2 tip
4	BROWN/WHITE	Line 2 ring	Line 2 ring
30	WHITE/SLATE	Line 2 A	Line 2 A
5	SLATE/WHITE	Line 9 A	A circuit common(gnd)
31	RED/BLUE	Line 2 lamp ground	Line 2 lamp ground
6	BLUE/RED	Line 2 lamp	Line 2 lamp
32	RED/ORANGE	Line 3 tip	Line 3 tip
7	ORANGE/RED	Line 3 ring	Line 3 ring
33	RED/GREEN	Line 3 A	Line 3 A
8	GREEN/RED	Line 8 A	A circuit common(gnd)
34	RED/BROWN	Line 3 lamp ground	Line 3 lamp ground
9	BROWN/RED	Line 3 lamp	Line 3 lamp
35	RED/SLATE	Line 4 tip	Line 4 tip
10	SLATE/RED	Line 4 ring	Line 4 ring
36	BLACK/BLUE	Line 4 A	Line 4 A
11	BLUE/BLACK	Line 7 A	A circuit common(gnd)
37	BLACK/ORANGE	Line 4 lamp ground	Line 4 lamp ground
12	ORANGE/BLACK	Line 4 lamp	Line 4 lamp
38	BLACK/GREEN	Line 5 tip	Line 5 tip
13	GREEN/BLACK	Line 5 ring	Line 5 ring
39	BLACK/BROWN	Line 5 A	Line 5 A
14	BROWN/BLACK	Line 6 A	A circuit common(gnd)
40	BLACK/SLATE	Line 5 lamp ground	Line 5 lamp ground
15	SLATE/BLACK	Line 5 lamp	Line 5 lamp
41	YELLOW/BLUE	Line 6 tip	
16	BLUE/YELLOW	Line 6 ring	
42	YELLOW/ORANGE	BL, AG, or spare	BL, AG, or spare
17	ORANGE/YELLOW	SG, LK, or spare	SG, LK, or spare
43	YELLOW/GREEN	Line 6 lamp ground	
18	GREEN/YELLOW	Line 6 lamp	
44	YELLOW/BROWN	Line 7 tip	
19	BROWN/YELLOW	Line 7 ring	
45	YELLOW/SLATE	B or B1	B or B1
20	SLATE/YELLOW	R or R1	R or R1
46	VIOLET/BLUE	Line 7 lamp ground	
21	BLUE/VIOLET	Line 7 lamp	
47	VIOLET/ORANGE	Line 8 tip	
22	ORANGE/VIOLET	Line 8 ring	
48	VIOLET/GREEN	Line 9 lamp ground	
23	GREEN/VIOLET	Line 9 lamp	
49	VIOLET/BROWN	Line 8 lamp ground	
24	BROWN/VIOLET	Line 8 lamp	
50	VIOLET/SLATE	Line 9 tip	
25	SLATE/VIOLET	Line 9 ring	

Figure 11. Telephone color code and 1A2 key system assignments. The pin numbers indicated are for the Amphenol "Blue Ribbon" connectors used to terminate 25-pair cables.

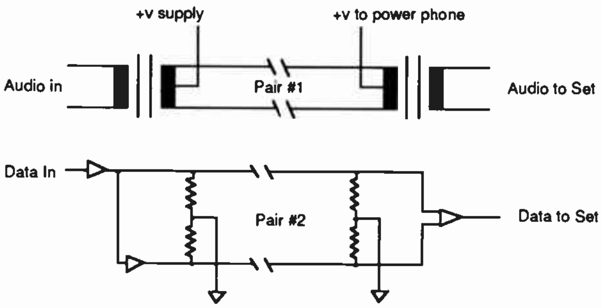


Figure 12. The most common "electronic phone" approach uses four wires, with one pair for audio and another for control and display data signalling. Power is phantom applied.

Spirit system and the popular NEC and TIE systems among many others use this approach. Talk and data each use one of the two pairs. The power is applied between the two pairs in a way similar to the method used for phantom powering condenser microphones in recording studios.

A transformer at each end of the audio pair permits the phantom power to be added. The data pair will probably use resistors to obtain a "center tap," rather than transformers since the data signal has a DC component which could not pass through a transformer.

Two-Pair, Power Not Phantom

Some two-pair systems put the data on one pair and the audio on the other. Power may be on the data pair or on the audio pair. In the latter case, the audio pair resembles a central office line so that the phone ports may be universal: either single-line sets or feature phones can be plugged-in without hardware changes in the PBX. At least one of the Panasonic systems uses this technique. The center pair, again, is generally the audio.

Data Over Voice

The Mitel Superset phones use a unique scheme that requires only one pair for all three functions. How do they do it? The data is amplitude shift modulated onto a 32 kHz carrier "over voice" and then the combined voice and data are AC coupled across the DC power voltage.

The most advanced systems use a pure digital bit stream for both voice and data. The phone set contains the CODEC for conversion to and from the analog and digital domains. The pure digital approach is used in the AT&T System 85, in the Northern Telecom Meridian family, in the newer Mitel systems with the Superset DN phones, and in the digital version of the NEC NEAX 2400, among many others.

Interfacing to Electronic Phone Equipment

With the proliferation of electronic phones in broadcast stations, consideration needs to be given to the possibilities and problems of interfacing with on-air use. A number of techniques for accomplishing broadcast interface are described below. However, these are best reserved for casual use of the phone such as for the occasional on-air request or contest winner call because the broadcast hybrid interface cannot determine when a new call is selected, so it cannot adjust its null to the new line before the conversation starts. However, since the hybrid can null on voice during conversation, in perhaps four seconds null will be achieved. This is okay if only a portion of the call is to be aired; as is common with on-air requests, contest winner calls, and the like. Another shortcoming of the direct-to-electronic phone approach is that the line switching "clunk" is not muted. This is no problem when calls are not to be aired directly and sequentially.

Another potential problem is audio quality. The primary impediment is usually noise, most often the result of the data signals cross-talking into the audio. Buzz from the power supply sometimes finds its way into the audio. Often, frequency response is limited by line coupling transformers which are too small, or other causes. Poorly designed digital systems may suffer from quantization and aliasing noise and distortion.

Few PBX manufacturers publish specs on audio performance. Since, clearly, this is of importance to those of us who need to get decent quality from phones for on-air use, we'll want to make sure that the audio is at least reasonable. Ask the phone system dealer for audio performance data. If possible, arrange to measure a system yourself before you purchase it.

For serious phone use, it is usually better to consider the complete on-air switching systems from broadcast interface manufacturers.

Interfacing to Electronic Phones

First, we'll consider ways to connect our broadcast interfacing equipment directly to the skinny-wire cable which feeds the phone. Since this might not be possible in some cases, such as with the pure digital systems, we'll then investigate other ways to get audio for on-air use.

Direct Connection to the "Skinny Wire"

When the phone system uses the separate-pair approach described above, the center two wires on the modular plug are usually the audio path. Since the phone's control functions stay active even when these connections are broken, it is possible to intercept the audio signal here for feed to the interface. Most broadcast interfaces provide a "loop-through" connection, which feeds the phone line back out when it's not active. Thus, the unit may be series connected with the audio pair. That way, you have normal telephone function preserved when the interface is not in use. When the interface is active, the phone serves merely as a controller, with no audio reaching the phone's network or handset. Wiring the hybrid's on/off functions to the console's switching logic accomplishes automatic operation.

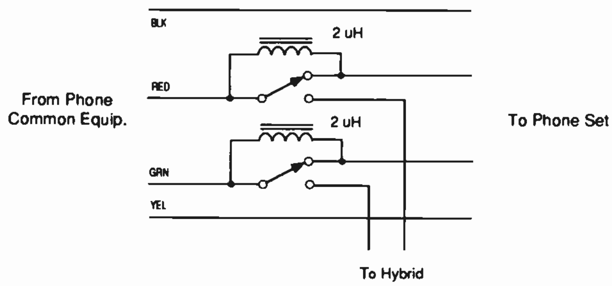


Figure 13. If the center two wires on many electronic phone systems convey audio, they may be used to feed broadcast equipment. The inductors bypass power to the phone set when the studio interface is active.

When the phone uses the two-pair phantom approach described above, again the audio is likely to be present on the inner pair and may be intercepted for interfacing use—if the DC connection is maintained. One way to do this is to provide a by-pass for DC with inductors. 2 uH has proven acceptable in experiments performed on some phone systems.

Special System Ports: Faux CO Lines

Since fax machines and modems need connections which look like central office lines, many systems provide ports for this use. They may be connected to broadcast interfacing equipment as if they were CO lines. Sophisticated PBX's have programming features which allow these ports to be configured in various and potentially useful ways. For example, they may be set up for "private line" ringing. That is, when a given incoming CO line rings, the call may be directly sent to the selected port.

Unfortunately, with most PBX systems, awkward operation may result, since the only way to move a call from a phone set to the port may be to transfer it using multiple button punches, rather than the usual simple place-on-hold-and-pick-up-elsewhere operation. Taking calls in sequence on-air may be extremely difficult—or impossible.

Speaker-Phone Tap-Off

One way to get low-cost interfacing is to take advantage of the switching-type interface that many phone set internal speaker-phones provide. The procedure is to tap off the speaker with a transformer and pad to the console's required input level. You may continue to use the phone's internal microphone or you can provide an external send audio source by substituting for the phone's internal microphone. Again, you will probably need a pad and transformer. The input feed is set so that appropriate switching action and proper send levels are obtained.

Handset Adapters

Adapters are available which plug into the phone set's handset modular jack and convert the microphone and earpiece signals into a signal which emulates a standard CO line. While useful in some applications, this approach is likely to offer a lower quality feed. This is because the phone set's network remains in the signal path causing impedance bumps and other problems.

Intercepting the Serial Data Stream

Why can't we just emulate an electronic phone set by generating and decoding the phone system's serial data? It does seem that this would be a good solution. However, phone system manufacturers insist upon maintaining their data protocols as a deep secret. That means that broadcast manufacturers are unable to design direct emulation equipment. Of course, even if we had the protocols, there is the problem of accommodating the dozens of communication methods employed by PBX designers.

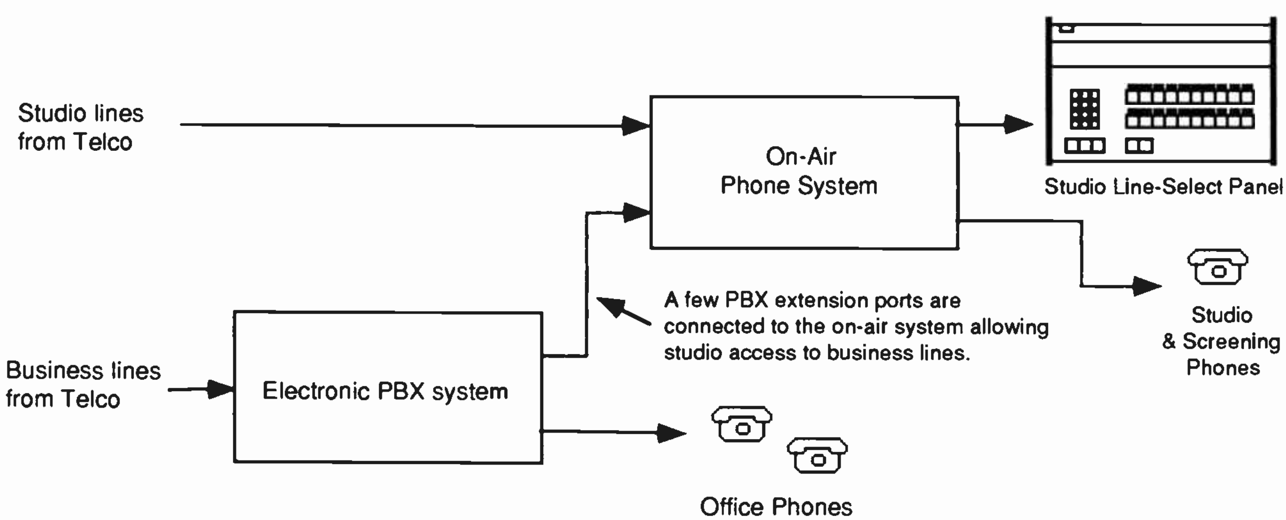


Figure 14. One way to integrate the on-air system with the station business phone system. Ports intended for single-line phone sets are used as input to the on-air system.

The Evolving Phone: Future Directions

The Move to Digital

As time goes on, probably all but the most inexpensive systems will use the pure digital approach. All of the integrated circuits (ICs) and technology being developed for the integrated services digital network (ISDN) (as described later) will be useful for PBX equipment application, as well. At first, these systems will be more difficult to interface. But, as the ISDN switching protocols become standardized, broadcast interface manufacturers may be able to provide equipment which could directly connect to the PBX in place of, or in series with, the studio phone set.

The Open PBX

With the open PBX system, the PBX manufacturer provides complete documentation on an interface which can provide control of all of the important aspects of phone switching including call set-up and routing functions. A standard data port is provided so that outside vendors may supply systems to work in concert with the phone equipment. These open PBXs may eventually offer a universal method for broadcast equipment to coordinate with the station's office phone system.

BROADCAST INTERFACING: AUDIO

One-way Interfacing

The QKT

Often, there is a need to take audio in only one direction at a time from a phone or broadcast. Newsroom phoners are a common application. If there is no requirement for a two-way conversation, a simple interface will do. A simple device that was once available from the phone company, the QKT, is required. This small box was permanently wired into a phone instrument or line, and provided a quarter-inch phone jack output for feeding a line-level signal to a console or recorder input.

A QKT is nothing more than a transformer, a capacitor, and a zener diode limiter. The capacitor provides DC blocking so that the transformer does not become saturated with the phone line's DC potential. In order for the coupler to hold the line by drawing loop current, eliminate the capacitor and use a transformer which can withstand the loop current without producing distortion. (Such a transformer is the SPT117 from Prem Magnetics.) When sending audio into the phone line, remember audio level should be limited to −9 dBm. The QKT has back-to-back zeners for this purpose; you may want to add them to your home brew interface if you expect audio levels to get out of hand.

When using a coupler, it is most convenient to have the telephone instrument on-line and equipped with a push-to-talk switch on its receiver. This is because the phone's receiver has to be off-hook while a feed is

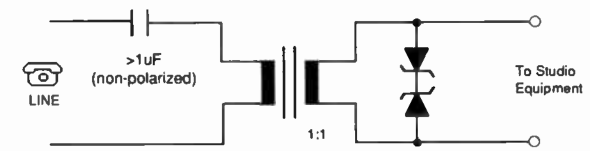


Figure 15. Simple one-way-at-a-time interface. The capacitor is for DC isolation and is not required when a transformer which can sink loop current is used. The zeners are chosen to properly limit transmission levels to the required −9 dBm.

coming in, and the switch turns off the receiver's mouthpiece microphone when it is not depressed, thus insuring that noise from the studio side will not be included in the recording. Since a QKT works in both directions, it can be used to send audio down the phone as well. This is useful in the production studio for letting clients hear those commercial masterpieces before they go into the control room.

When hooking up to a multi-line phone, connect to a point where the tip/ring is present after line selection. The most convenient place is usually right at the phone network. Use headphones to find the spot.

Two-way Interfacing

The QKT's limitations become apparent when it is necessary for the caller to hear the announcer and the audience to hear the caller simultaneously. A more complex method is needed because of the requirement to have isolated send and receive audio signals.

The Switching Technique

This technique uses gain switching to keep the send audio from appearing at the receive output. Two electronic switches are used in such a way as to ensure that either the send or the receive path is closed at any given time, but never both simultaneously. A decision circuit compares the send and receive levels, with the direction of transmission being determined by the relative signal strengths.

In practice, voltage controlled amplifiers are used to provide soft attenuation control rather than the absolute on/off that you would have with simple switches. The most common speakerphones work this way.

The disadvantage of the switching technique results from the unidirectional nature of these systems. The

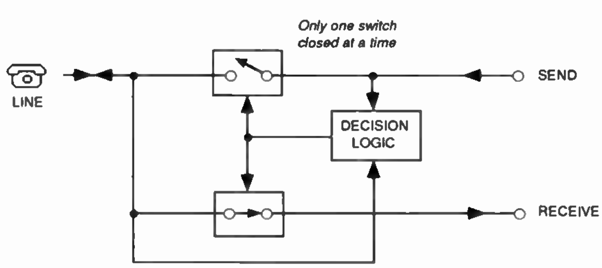


Figure 16. Switching interface allows two-way conversation, but only one way at a time.

primary problem is that the caller cannot be heard while the announcer is speaking. Also, noises in the studio can sometimes cause a caller to disappear momentarily, especially on weak calls.

The most sophisticated switching systems have circuits to compensate for varying levels and background noise conditions in order to prevent inappropriate switching.

The Hybrid

Hybrids were invented long ago in order to separate the send and receive signals from the common two-way phone pair. Early hybrids were made from transformers with multiple windings. Nowadays, most hybrids are made with active components and are known as *active hybrids*. Both circuit types use the same principle and achieve the same effect.

In Fig. 17, the first op-amp is simply a buffer. The second is used as a differential amplifier, and the two inputs are added out-of-phase (subtracted). If the phone lines and the *balancing network* have identical characteristics, then the send signals at the second differential amp will be identical, and no send audio will appear at the output.

The balancing network is a circuit consisting of capacitance and resistance, and sometimes inductance, forming an impedance network. Depending on the hybrid's application, this circuit can be very simple; or it can be comprised of a large number of components and have a very complex impedance characteristic.

R1 and the phone line form a voltage divider, as does R2 and the balancing network. If the phone line and balancing network are pure resistances, then, clearly, the phone line and the balancing network must have the same value in order for the signals at the differential amp to have the same amplitude and for complete cancellation to occur.

The phone line, however, is not purely resistive, but a complex impedance, causing both the amplitude and phase to vary as the send signal frequency varies. Loading coils, two-to-four wire converters, transformers, repeaters, T-carrier systems, and other telco systems are responsible for significant impedance bumps.

Only when the impedance of the balancing network is the same as the phone line, and the signals at the dif-amp are matched in both amplitude and phase, will full cancellation of the send signal be achieved. Otherwise, leakage results: the scourge of hybrids.

Because the phone company's requirements are not generally too stringent, they usually use a simple network with compromise values of resistance and capacitance. Their goal is to get an average of about 12 dB rejection, with 6 dB acceptable on difficult lines; just enough to prevent feedback in a system with back-to-back hybrids. When the situation calls for better performance, modules with a number of R and C elements which can be switched in or out are employed, the switches being set to match the network to a particular line.

A terminology note: Since the hybrid may be used to interconnect two and four-wire circuits by separating (and combining) the two signal directions, it is sometimes called a *two-to-four wire converter*.

Broadcast Hybrid Application

In broadcast application, the studio mixing console combines the output of the hybrid and the announcer's microphone audio. As discussed above, the hybrid output consists of both the desired caller audio and the undesired leakage: that is, the announcer audio; but *phase-shifted because of the phone line's reactance*. If the amount of leakage is too great and the phase shift too extreme, the announcer sound will suffer

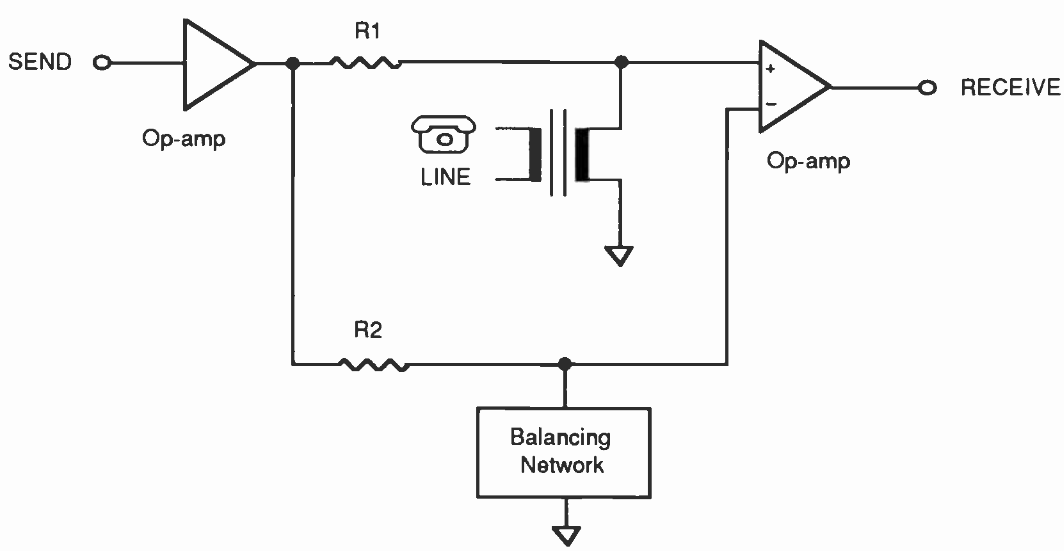


Figure 17. Op-amp hybrid. The second op-amp is used as a differential amplifier to perform the required subtraction for nulling.

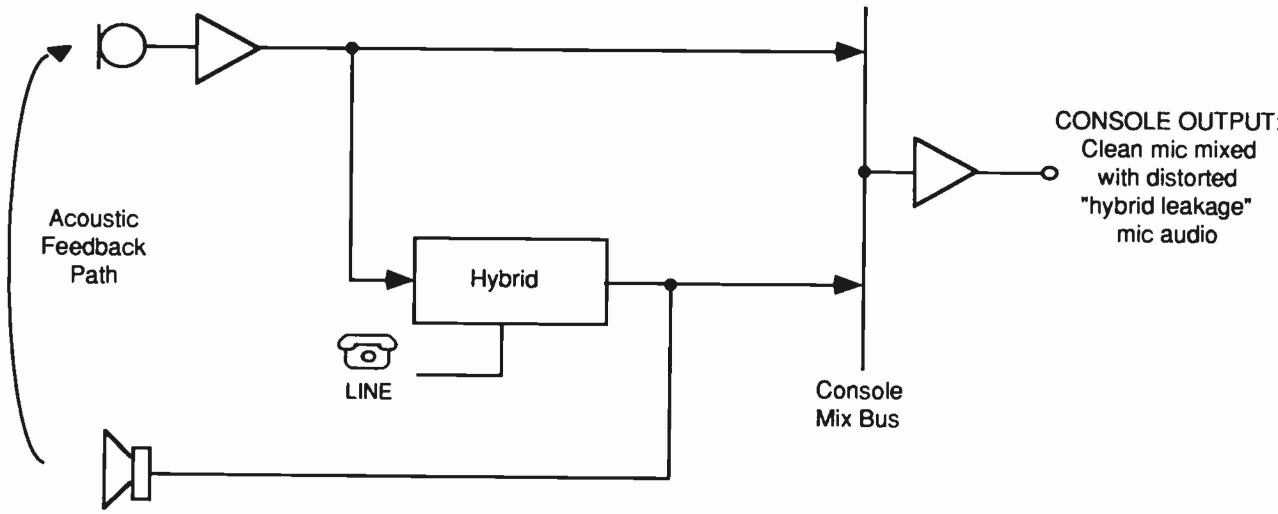


Figure 18. Block diagram of typical studio arrangement with telephone hybrid. Announcer audio is combined with hybrid output potentially causing problems with announcer voice distortion. The acoustic path is a possible source of audible feedback.

degradation as the original and leakage audio combine in and out of phase at the various affected frequencies. When this occurs, the announcer sounds either "hollow" or "tinny" as the phase cancellation affects some frequencies more than others. Another effect of too-little trans-hybrid loss is that feedback can result from the acoustic coupling created when callers must be heard on an open loudspeaker.

A hybrid, then, will be useful for broadcast only when leakage is kept acceptably and consistently low.

The plots of phone line impedance vs. frequency and phase shift shown in Fig. 19 are the result of measurements performed on phone lines at a radio station in the midwest. They indicate the wide variation seen on typical telco lines as provided to broadcasters.

The lines with smooth curves have impedance characteristics which could be emulated with a simple resistor-capacitor combination. These lines would work fairly well with a simple hybrid, since an RC balance network would match the impedance characteristic closely enough to make the cancellation of send audio at the hybrid output good enough to prevent coloration of the announcer audio.

Those other lines are quite another story! While it is theoretically possible to construct a balance network to match these difficult lines, practical considerations usually keep this approach from being used: the impedance characteristic required is too difficult to produce using resistors and capacitors. If the hybrid is to be switched among a number of lines, the line characteris-

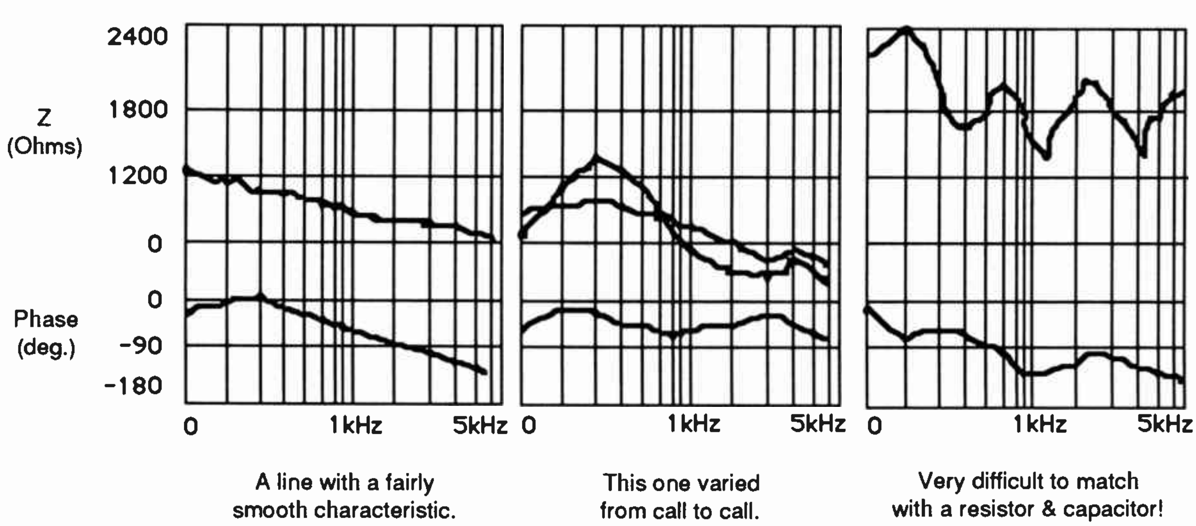


Figure 19. Impedance vs. frequency curves for some typical phone lines.

tic would have to be consistent from call-to-call with nearly the same impedance curve.

Hybrids with an automatically-adjusted resistor/capacitor balance network are a good start. They are capable of good results when the phone lines to be used have a smooth impedance curve, but poor performance will still result when the line impedance characteristic is not a simple first-order RC curve and cannot be adequately matched by a simple RC network.

One way to improve performance on widely varying or difficult lines is to trade single-frequency rejection for wide-band loss by making the phone line look less reactive to the hybrid with some series resistance. However, there is a limit to the amount of resistance which can be inserted, since the send and receive audio levels drop with increasing resistance.

Combining the Hybrid and Switching Techniques

In this case, the hybrid produces as much send-to-receive isolation as can be achieved. Then a ducking or override function causes the "dynamic" rejection to be greater than the hybrid alone can produce. With this approach, when send audio is present, the receive gain is reduced. Thus, leakage, also, is minimized. However, since the level from the phone is also reduced when the announcer is speaking, there is a sacrifice of full-duplex operation. A user adjustment in the control signal path permits variation of the amount of receive ducking, allowing full duplex operation when the hybrid alone produces sufficient rejection, or speakerphone-like operation whereby the caller is turned almost completely off when the announcer speaks. As a practical matter, this control is usually set to provide the minimum amount of ducking which provides adequate send-to-receive leakage suppression.

Digital Signal Processing Hybrids

DSP (*digital signal processing*) offers a very powerful and effective technology to improve hybrids. DSP is the process of operating on analog signals which have been converted into the digital domain. Since the signals are numbers, mathematical operations can be performed to manipulate them before being returned to analog. Complex processing functions either impractical or impossible to be done with analog circuit elements are achievable in DSP.

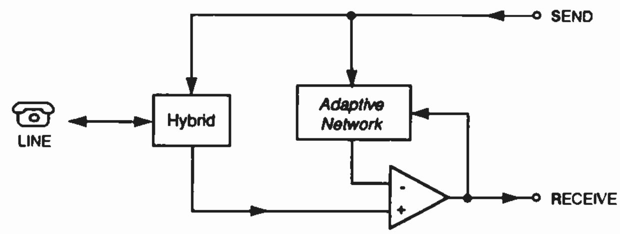


Figure 20. In the DSP hybrid, the digital balancing network continuously adjusts to the phone line impedance characteristic. When the adaptive network transfer function is identical to that of the phone line, perfect cancellation is achieved. Since the adaptive network is a digital filter which can create almost any required curve, performance is superior to the analog hybrid alone.

With the DSP hybrid, natural simultaneous conversation is possible without distortion of the announcer audio. To accomplish this, the announcer and caller audio signals are digitized and processed in a system which makes use of a specialized DSP microprocessor. The DSP processor is much like any other microprocessor. It has input/output, memory, a CPU, etc. The distinguishing difference is the very high speed and specialized instruction set optimized for signal processing. The digital hybrid incorporates software programmed to perform the hybrid function. The technique, *convolutional least mean square adaptive filtering*, is capable of very accurate synthesis of the required balancing transfer function for maximum cancellation. Unlike resistor/capacitor analog schemes, the adaptive filter can create the complex multiple break-point impedance vs. frequency curves required by difficult-to-match phone lines. The send and receive signals are constantly compared in a feedback loop with the leakage becoming an error control signal which drives adjustment of the *digital balancing network*.

The performance advantage of the digital hybrid technology is striking. On a typical phone line with a fairly smooth impedance curve, an analog hybrid might attain 15 to 20 dB trans-hybrid loss. A digital hybrid will likely produce 40 dB or better trans-hybrid loss. On lines with difficult impedance curves, the analog hybrid's performance will usually be so poor as to

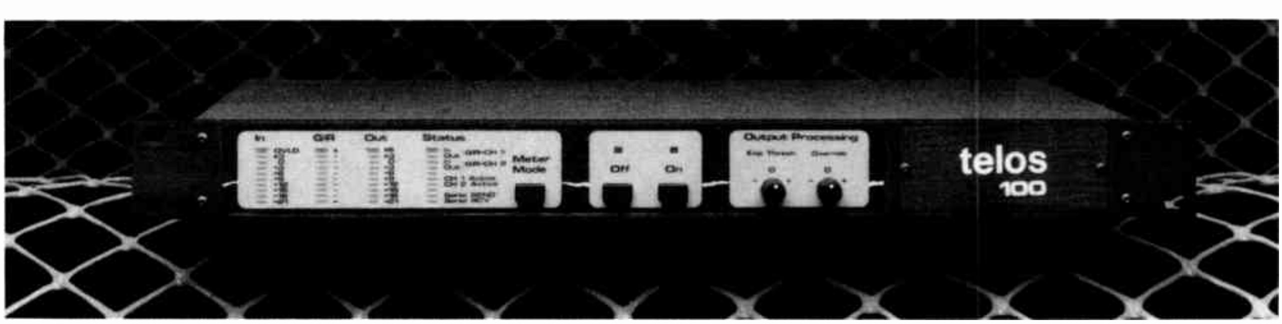


Figure 21. Telos 100 digital hybrid.

prevent its use, while a digital hybrid would perform acceptably.

When a call is initially established, a brief mute/adaption period provides an opportunity for the system to adjust to the phone line prior to the call going on air. The caller hears a noisy tone, but none of this tone is heard on the air since the output is muted. This has the incidental benefit of removing the line switching "clunk." Adaption continues as the conversation proceeds, using voice as the reference signal.

While in the digital domain, other operations in addition to the hybrid adaptive balancing can be performed. Automatic gain control (AGC) can take advantage of digital techniques to significantly improve upon the same functions implemented in analog. For instance, cross-coupling to the hybrid section is possible in order to avoid the output AGC confusing hybrid leakage with low-level caller audio and inappropriately increasing gain. AGC may be smartened in other ways as well. An adaptive floating expansion threshold, for example, improves noise-gating quality.

Evaluating Hybrid Performance

The amount of hybrid rejection, or *trans-hybrid loss*, directly affects the on-air audio and is the most critical measure of hybrid quality. *The true test of hybrid performance is determined by measuring the amount of rejection across the entire audio frequency range, preferably with pink noise as a test signal at the send input.* Any hybrid with an adjustable R and C balance network can produce high rejection at a single frequency, since both phase and amplitude at a single frequency can be adjusted for good cancellation. Voice is clearly not a single-frequency source.

Another thing to keep in mind is that the trans-hybrid loss is not the same as the *observed* difference between the caller level and the leakage at the hybrid's output, although the two are related. That is because the typical phone call is maybe −20 to −25 dBm (on choke lines, even lower) and the send level (to the caller) from the hybrid should be −10 dBm. That means that the hybrid has to use up 10 to 15 dB of its trans-hybrid loss *just to get even*. The remainder becomes the observed difference.

Other important performance characteristics include signal-to-noise ratio, distortion, and (for a digital unit) number of bits in the audio path. The operation of the dynamic functions (the AGC, noise gate, and override ducking) make a significant contribution to a hybrid's effective performance.

Improving Phone Audio Quality

Due to its limited frequency response and fairly high distortion, the audio from the phone has the poorest quality of our on-air sources. Thus, it generally pays to make telephone audio less of an "earsore", so that it does not stand out more than is necessary from other program material. Filtering and equalization are the primary tools.

Filtering

On a dial-up phone line, there is very little audio above 3.4 kHz, but there is noise. Thus, a filter with a very steep roll-off above the telephone pass band will reduce phone line noise significantly without affecting conversation audio.

The low-end can be improved as well. Low-frequency hum is often a problem; usually 60 Hz mixed with its second harmonic, 120 Hz. Thus, it is often helpful to have a sharp roll-off starting at 200 Hz or so. Again, there is very little audio present at these frequencies.

Additional Processing

An equalizer used to shape the frequency response of the phone line within its audio bandwidth can result in marked improvements in perceived quality. A typical phone line has an excess of energy at around 400 Hz and considerable roll-off at both the top and bottom ends of its pass band, so the idea is to compensate by adding gain at both. Boosts at 2.5 kHz and 250 Hz and a cut at 400 to 500 Hz with a parametric equalizer will help achieve better sound. Since every phone line is different, the ear is usually the best instrument to evaluate the results.

Another effective processing device is the expander, or noise gate. These devices may be used to reduce gain "between the words" of a conversation, thus making phone line noise less objectionable. On extremely noisy lines, however, the gating action can make noise more distracting by causing it to come and go with the words. In such cases, it might sound better to leave the gate off and let the noise remain present at a constant level. A unit with variable threshold and duck ratio can be adjusted so that the optimum compromise may be achieved between the benefit of reduced noise and audibility of the effect.

BROADCAST INTERFACING: ON-AIR SYSTEMS

With phones an important part of programming at many stations, systems to enable convenient, high-quality on-air integration of phone conversations are essential.

Design Characteristics of Importance to On-Air Systems

On-air phone systems are specifically designed for use in the broadcast studio environment. While many business phone systems offer similar functions (line selection and status indication, conferencing, etc.), they are generally awkward to operate and may have other limitations such as the audio quality flaws described above. For on-air use, we have somewhat different criterion from that of the typical business phone.

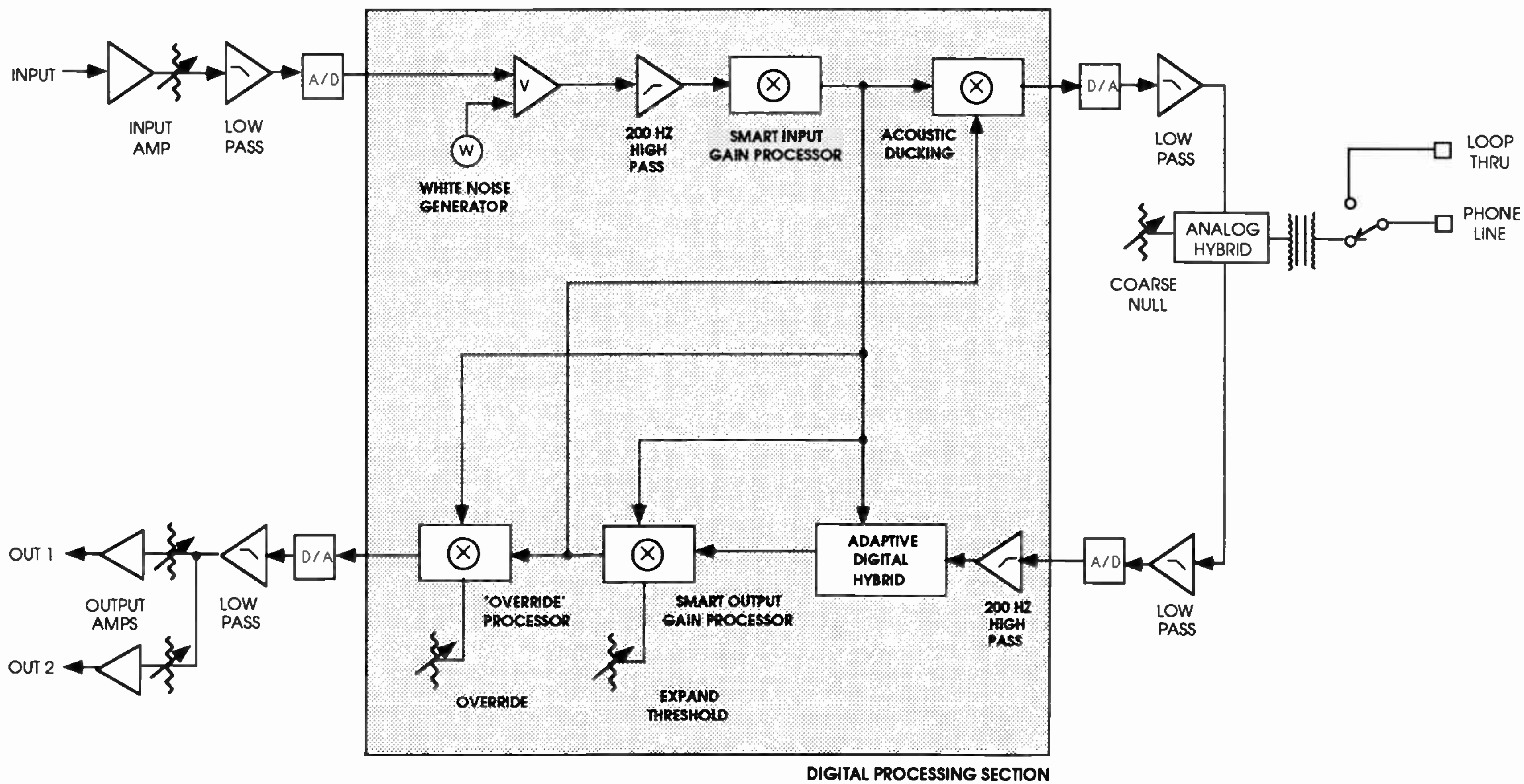


Figure 22. Signal flow diagram for a DSP-based hybrid (Telos 100). All processing, including the adaptive hybrid and smart AGC are performed in the digital domain.

Figure 23. Gentner "Teleswitch." Self-contained selector for up to five on-air lines. Used in combination with a hybrid for a complete on-air system.

Ergonomic Requirements

Line selection and other functions must be accomplished with a minimum of hassle and with an intuitive feel. Unlike with a telephone set, broadcast line selection panels have large illuminated buttons. Features are limited to those necessary for the on-air application in order to avoid operator confusion. Panels which drop into an open position in the studio mixing console (so that the line selection buttons are located near the channel on/off, fader, and audio switching functions) ease operational confusion.

Audio Quality Considerations

While the phone network would not be considered a high-fidelity source, it clearly does not help to degrade it further by adding additional noise, distortion, or frequency response impediments. For that reason, broadcast phone systems are designed with attention paid to these issues.

The phone system output should be free of inappropriate switching sounds and an air talent should be able to access and manipulate lines without any pops or "clunks."

Conferencing Capability

Most broadcast systems allow any number of lines to be switched to air, even if only a single hybrid is present. But, unless you are blessed with excellent phone lines, you will want additional hybrids with each connected to the other through a multiple mix-minus arrangement. That way, it will be possible to have amplification between callers. Without multiple hybrids, callers might have difficulty hearing each other, since you are at the mercy of the phone company's line level.

Special Features

Desirable features for an on-air phone switching system include:

Busy/unbusy. To prepare for a contest, all lines may be busied-out and then returned to readiness after the contest has been announced.

Automatic next line selection. Pressing the NEXT button picks up the line that has been holding the longest. If no line has been holding, the longest ringing-in line is selected.

Call length timer. So that callers be limited to a certain duration.

Held caller timer. To tell which line has been holding longest and for how long.

Automatic answer and message play. Provides an automated way to answer contest callers in sequence and for talk shows without someone to screen calls.

Mix-Minus: Getting the Send Audio Feed

The feed-to-caller signal has come to be referred to as *mix-minus*, so called because it is often the *mix* of all of the console's active inputs *minus* the phone

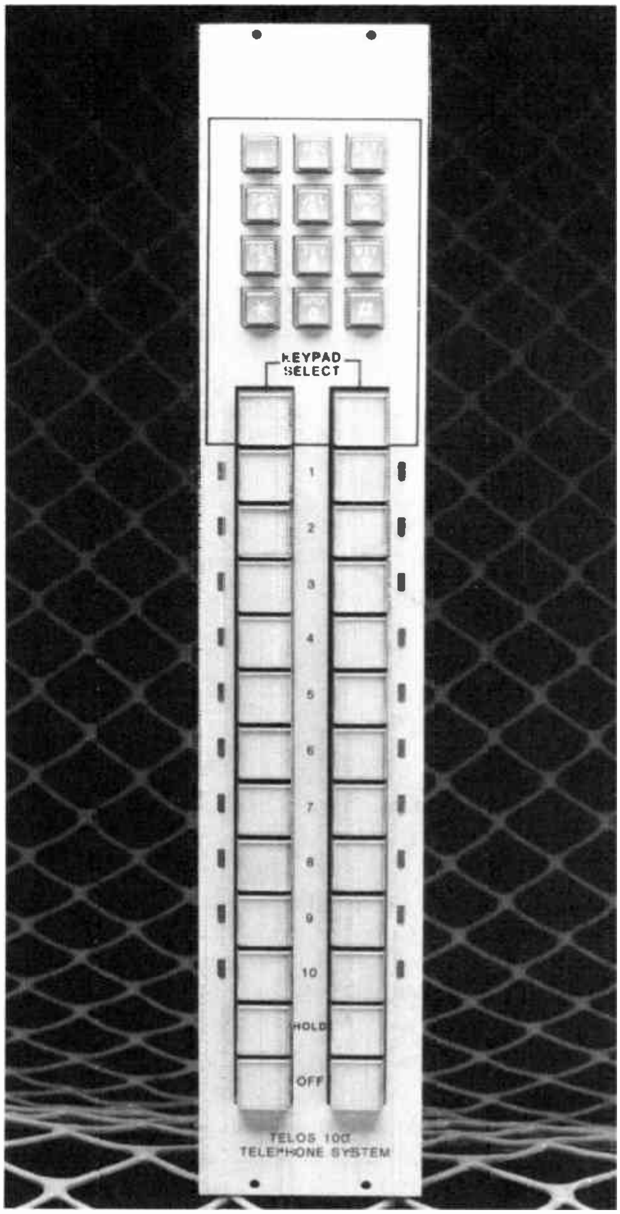


Figure 24. Panel used to allow control of line selection from within studio audio console. Connects to telco interface modules from Telos.

hybrid's output. A mix-minus feed is necessary because the hybrid will create a feedback path if it is forced to "chase its tail."

Usually, the mix-minus is a mix of only the studio microphones, but may sometimes include other audio which is to be sent to the phone such as contest sound effects from cart machines.

To create the required signal in simple installations, a feed taken from the main announce microphone may be all that is necessary. The *patch send* output available on many consoles is precisely what you need. In installations where multiple mics are to be used, a combiner of some sort is required. This may be a small outboard mixer or a homebrew op-amp summer or even a resistive combiner. Better consoles offer special purpose busses which may be used for mix-minus, often with provision for selective switching of sources into the phone feed. If you need to do something with an older console which does not have special busses, there is a device (made by Henry Engineering) which accomplishes the mix-minus by subtracting the hybrid audio from the console program output with a differential amp scheme. This unit generates a mix-minus signal true to its name: all sources except the phone itself will feed the phone.

Recording Phone Calls

Some stations may want to record calls for later playback. An interesting technique is as follows: The mix-minus goes to one track of a stereo tape machine, while the other channel gets the hybrid output with the caller audio. The result is a two-track tape with the announcer and caller separated. To play back, the console's input mode is set to mono. You can adjust the relative balance, if need be, upon playback. Your production department gets a tape which facilitates extraction of contest squeals, etc.

Integration of On-Air System's with PBXs

To interconnect the on-air system with the front office PBX, there are a number of possibilities.

1. *Segregate the studio and office phone lines.* Ports from the PBX configured to look like CO lines feed an input or two on the studio system so that calls taken by the receptionist can be put on the air.
2. *Route all lines through the PBX.* The studio lines are programmed in the PBX to be forwarded to the ports which feed the studio system. Some audio degradation may result.
3. *Simply parallel the two systems.* With no cross-coupling of line status information, there could be trouble if a line is inadvertently picked-up on one system while the other is using it.
4. *Route the on-air lines through the broadcast system.* This is possible if the broadcast system brings out a loop-through connection preventing PBX phones from picking-up active on-air lines.

Remotes With Phones

Many shows which make extensive use of phones have taken to the road. Phones need to be fed to the remote talent so that they can hear and responds to a caller, and the caller needs to hear the talent. In many cases, the remotes are sufficiently distant that the talent cannot hear callers from an off-air monitor. Even if he could, the profanity delay would be a problem, since he needs to hear the phone pre-delay.

In the usual situation, the talent hears callers via some kind of return link, usually a standard phone line. This line is fed with a different kind of mix-minus: a mix of everything on the program bus minus the

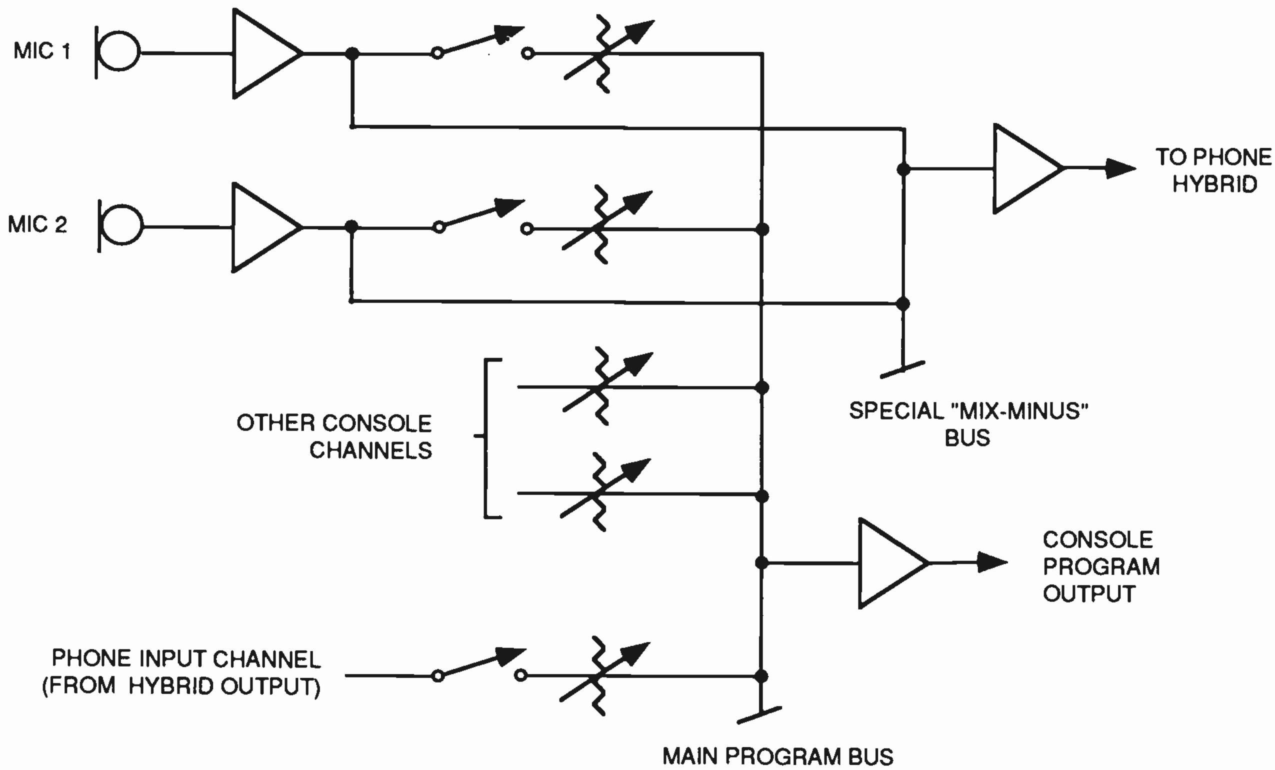


Figure 25. Simplified studio block diagram shows the mix-minus required for hybrid feed.

remote talent audio. The talent is mixed to his own headphones locally with a full-fidelity, nondelayed mic signal.

As for the second half of the equation, the caller hears the talent because the remote feed is added to the telephone mix-minus buss. No problem with a set-up which permits selective assignment to the phone mix-minus.

The most common problem with this arrangement is a result of insufficient trans-hybrid loss combined with satellite delay. The closest an earth station can be to a satellite in geo-stationary orbit is 22,300 miles; which means that, at 186,000 miles/sec, the two-way transit time is approximately one-half second.

If the hybrid isn't doing a good job of preventing the send audio from leaking to its output, the special remote send mix-minus is corrupted. If any of the announcer's audio from the remote site is returned via the monitor feed, it will be delayed by the satellite link. The answer is to have the best possible trans-hybrid loss. If the hybrid has variable caller ducking, it could be adjusted upward to enhance effective isolation.

Talk-Show Screening Software

In its simplest form this software, running on a personal computer, provides communication from a talk show screener/producer to the air talent, and signals who is on line waiting to talk. It replaces the "paper pieces on the window" system employed for years at many talk stations. The better packages offer a number of convenient features: display of weather data, liner messages, and other information; storage of caller data for demographic analysis; remote operation via modem.

A serial port on the broadcast system can enable the computer display to reflect current line status. Software could be developed to take advantage of lap-top computers to extend full control capability and status display to a remote site.

T-1 DIGITAL SERVICE

An ordinary copper phone pair can carry a much wider signal than the 3.4 kHz required for a single voice conversation. Indeed, a pure metallic path of reasonable length is easily capable of passing frequencies in excess of 100 kHz. Thus, multiplexing can be used to carry a number of voice channels over a single pair of wires.

T-1 Basics

A single telephone voice channel in digital form takes 64,000 bit/s (resulting from the 8 kHz sampling rate at 8 bits per sample). This is known to telco engineers as digital signal level zero, or DS-0. 1.5 Mbit/s is about the highest rate that can be supported reliably on copper pairs over the standard 1 mile distance

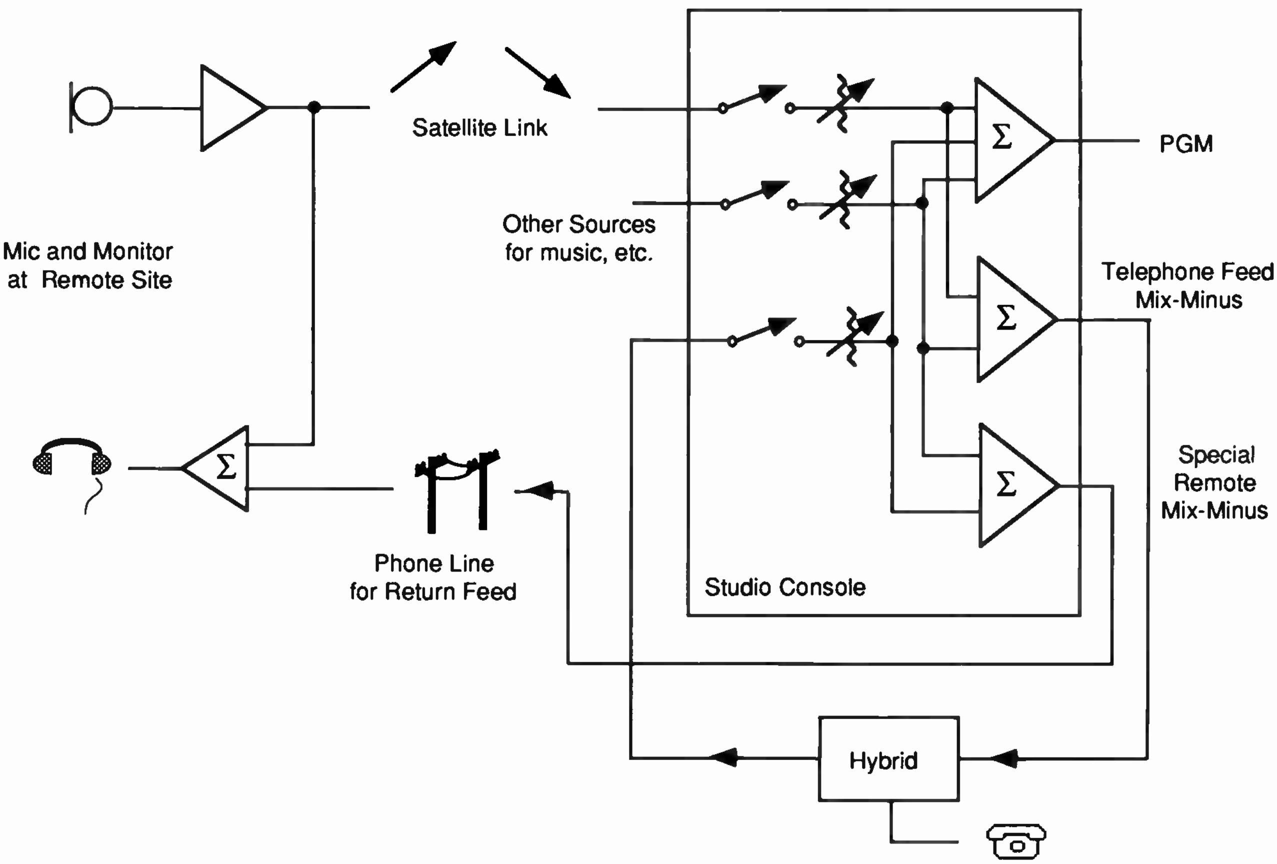


Figure 26. Set-up for remote broadcasts with phones.

between manholes; and thus repeater sites. 24 channels can be multiplexed using the new digital system (64 kbit/s × 24 = 1.536 Mbit/s). To create the T-1 bit stream, the 24 eight-bit channels are assembled end-to-end serially and the equivalent of another 8 kbit/s channel is added for synchronization. Thus the ultimate data rate becomes 1.544 Mbit/s. This is the number often seen associated with T-1. It is the rate also called DS-1. The signal is then converted into a digital bipolar bit stream in a special format called "binary 8-zeroes suppression," or B8ZS. The voltage varies between −3V and +3V.

In 1983 T-1 service began to be offered to the public. The service has grown tremendously in popularity,

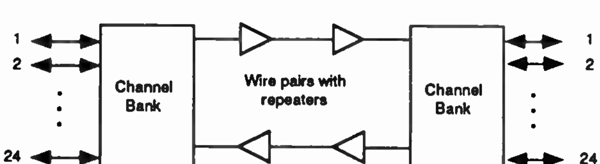


Figure 27. 24 channel carrier bank.

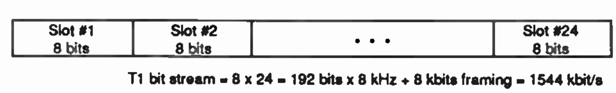


Figure 28. T-1 bit stream. 24 audio channels are transmitted sequentially.

particularly with large users of telephone service. These users like it because it is often less expensive than traditional analog service over the long haul and because it offers the advantages of a direct digital connection to the telephone network; advantages which become especially significant when large amounts of data as well as voice are to be transmitted.

Most long-distance carriers offer service on T-1. Long distance carriers are allowed by tariff only to deliver signals to the local *point of presence* (POP). The user is responsible for the so-called local loop, which may be obtained from your local telco or from any of the numerous "bypass" providers which are springing-up in many large cities. A user may provide its own bypass by installing a private link, such as a wideband microwave radio system.

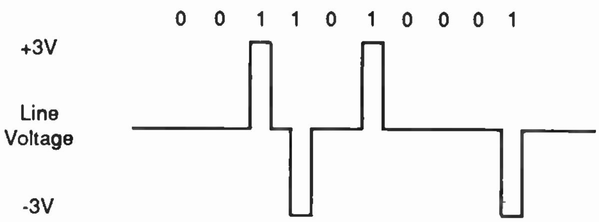


Figure 29. T-1 waveform. Bipolar voltage variation is encoded according to the "alternate mark inverted" technique.

Using T-1: The Customer Provided Equipment

Despite the difference in capacity and service, T-1 arrives at the end user site as two conventional telephone wire pairs, one for the data send and another for receive. The physical connector used to be a DB-15 type, but the current standard is the common RJ-48C, an 8-position modular plug.

The CSU

The T-1 line is first connected to a piece of equipment called the channel service unit (CSU). The CSU was considered part of the network (and still is outside the US), but is now almost always customer provided and may also be merely included as an adjunct section in a complete T-1 interface solution, which includes everything necessary to convert the digital feed into useful phone service. The CSU contains the last signal regenerator as well as a number of testing and maintenance features, such as provision for loopback testing by the central office. It may also include a system to collect and report error statistics.

The DSU and Multiplexer

The other components of a complete T-1 system are the digital service unit (DSU) and the multiplexer. In modern systems these functions are almost always combined into a single unit. The DSU handles the remaining digital "housekeeping" functions and data conversion from the bipolar T-1 format to standard serial data. The multiplexer, sometimes called a "channel bank," is where the multiple voice (or data) channels are combined into the single bit stream required for T-1 transmission. Each voice channel is converted to and from digital using the special CODEC A/D and D/A converters described earlier. In order to simulate typical telco lines, talk battery is added, ringing voltage is generated, and loop current is detected.

Channel Cards

Generally, multiplexers are constructed using a modular circuit card approach so that the available digital bandwidth may be used as desired. For example, multiple channels may be combined to create greater audio bandwidth for high-fidelity feeds. A traditional 15 kHz *program audio channel*, for instance, takes 6 of the 24 voice channels. Multiplexer manufacturers offer a number of options or *plugs* for their channel banks. Data in RS-232 or RS-422 format may be carried at almost any desired speed. Compressed video may be transported. Computer local area interfaces (LANs) may be connected so that separated computer networks operate as if they were together in the same building. Ports for special high-speed direct digital FAX machines are available. Newer systems allocate bandwidth dynamically between the competing users using *packet switching*, which allows both digitized voice and data streams to be divided into small packets and then efficiently combined.

A number of techniques are employed for bandwidth reduction of audio signals. The most common approach, adaptive differential pulse code modulation (ADPCM), allows 3 kHz telephone audio channels to be halved from the usual 64 kbit/s rate to 32 kbit/s so that two can be carried in a single T-1 64 kbit/s DS-0 channel. The same concept allows 7.5 kHz audio to be carried in a single DS-0. More exotic approaches rely on "perceptual coding" to further reduce bit rate and increase capacity.

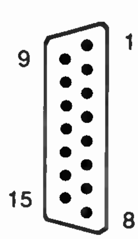
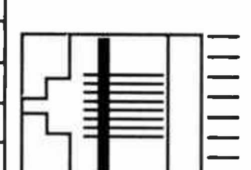
DB-15	PIN		PIN	RJ48-C
	11	R (Receive from Network)	1	
	3	T (Receive from Network)	2	
	9	R1 (Transmit to Network)	4	
	1	T1 (Transmit to Network)	5	
	N.A.	Optional Shield	7,8	
		(all others - no connection)		

Figure 30. T-1 connector pin-out. Either DB-15 or RJ48-C modular connectors may be used.

"Fractional" T-1

Telco providers are now able and willing to provide less than a full 24 channels of T-1 capacity. The low cost of fractional T-1, showing a cost breakeven (in 1990) with as few as three voice-grade lines is expected to speed the migration to digital. The provider delivers a standard T-1 line, but only turns on as many channels as you request. As your requirements increase, more channels can be activated. In time, it may be possible to expand and contract capacity in real time, paying only for what is needed and are using at any given moment.

Issues of Importance to Broadcasters

Many local telcos are aggressively pushing T-1 for simple telephone service to customers who have large numbers of phone lines. Long distance carriers offer significant financial incentives to convince you to switch to T-1 for long distance and 800 service. And with the installation of fiber optics reducing the cost and increasing the availability of T-1, it is becoming more commonly used for program transmission. Indeed, it is possible to allow a composite FM stereo signal to be relayed in the T-1 digital format. Since T-1 may be transported on copper, microwave, fiber, satellite, and even laser without degradation, it is possible to have an essentially perfect audio path from-and-to about anywhere. The NBC network and Voice of America (VOA), among others, make use of T-1 systems to move multiple channels of varying bandwidth from Washington to New York, and in VOA's case, ultimately to destinations in Europe.

What if a service provider offers T-1 for your standard phone service? Here are some considerations: If the serving central office is analog and the digital equipment is added to the signal path, noise and distortion may be increased. Also, the line will have more impedance "bumps," making the hybrid's job more difficult. On the other hand, if the CO is digital and the switching equipment allows the signal to pass through entirely in the digital domain, moving the audio conversion to your site may offer quality improvement over traditional analog delivery; especially if the local loop is lengthy. Another consideration: since all service will depend upon a single set of circuits, reliability could be reduced. Consider having back-up circuits in place.

Some T-1 terminal equipment has problems which cause the on-air hybrid interface to work very poorly, with bad cancellation the result.

Another possible trouble is that ADPCM encoders leave a bit to be desired with regard to fidelity. High-frequencies tend to sound a little muted and sometimes "buzzy." Thus, it is probably wise to steer clear of such encoding for a channel bank which serves lines intended for on-air use.

ISDN: DIGITAL PHONE SERVICE

Integrated services digital network (ISDN) is a refined and "consumerized" version of the T-1 system described above, allowing direct digital connection to the telephone network. In addition to the quality advantages end-to-end digital transmission offers for basic voice service, users may implement direct computer-to-computer links without using intervening inefficient, slow modems. Remember that the basic DS-0 digital channel operates at a bit rate of 64 kbit/s. This will also be ISDN's basic rate. Compare this to the common 2.4 kbit/s modem speed and it becomes evident why the promise of ISDN causes so much excitement among people who need to ferry lots of data. Imagine file transfers, computer graphic exchanges, and faxes at 10 to 32 times the present speed, with the dial-up convenience and low cost of the traditional telephone network.

Basic Rate ISDN

With this service, two 64 kbit/s voice or data channels, called "B" or *bearer* channels, and one 16 kbit/s "D" or *data* channel on a single telephone pair. The B channels may be used together for simultaneous voice and high-speed data transmission. The idea is to allow conversing with someone while sending (and receiving) a fax or computer data at the same time. The D channel is only for signalling between the terminal equipment and the telephone central office switching system. That is, for call set-up and the like.

One characteristic of ISDN important to broadcasters is that the B channels are true full-duplex. Unless intentionally introduced, there is absolutely no cross connection between the send and receive paths.

The single wire pair which brings ISDN to the user site is called the "U" interface. A device called the *network termination* or NT-1 converts the two-wire bidirectional feed into a four-wire connection called the "ST" interface (using, incidentally, a simplified version of the adaptive hybrid concept discussed earlier).

The terminal equipment may be as simple as existing telephone sets or may be some kind of integrated terminal, which would include computing power as well as voice capabilities. Low-grade compressed video is another possibility.

Primary Rate ISDN

Primary rate ISDN has a data rate equivalent to T-1, 1.5 44 Mbit/s. It should replace end-user T-1 eventually, since it will make immediate dial access to other primary rate ISDN subscribers possible. This high-capacity ISDN service will appeal to the same subscribers who now find T-1 appropriate. Further uses will, no doubt, evolve to take advantage of the unique combination of high capacity and dial access.

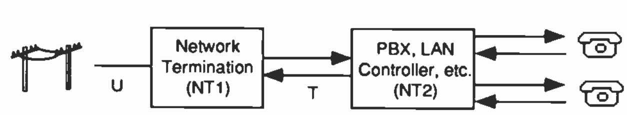


Figure 31. ISDN termination.

Video-conferencing might be an important application. Current plans call for primary rate ISDN to be "23B + D," for 23 64 kbit/s bearer channels and a single D channel, also at 64 kbit/s for setting up and taking down connections and other network control functions.

There is a proposed form of ISDN called *B-ISDN* which would take advantage of the growing availability of fiber optic facilities to achieve a data rate of 150 Mbit/s: a truly enormous data capacity. This B-ISDN service would be delivered to the end-user premises on a fiber cable. Full broadcast quality video, massive computer text/graphic/hyper-media networks, and other exotic services are anticipated.

ISDN and Broadcast: Hi-Fi Remotes on Dial-Up Lines

ISDN makes high-quality digital remotes possible with dial-up convenience. Studio-quality audio from anywhere without hassle and at reasonable cost will be possible. Talk radio will perhaps take on a different character when callers have the same near-studio quality. Will subscribers have hi-fi phones? Yes, ISDN proponents expect that even residential users will want 7.5 kHz phones, eventually.

Some broadcasters are using ADPCM digitizers to transmit 7.5 kHz audio over AT&T's 56 kbit/s Accu-Net lines. The International Telecommunication Union telephone standards committee, CCITT, has defined a standard for 7.5 kHz audio over a 56 or 64 kbit/s channel using sub-band ADPCM (described in CCITT recommendation G.722) and a small number of European networks and stations are presently taking advantage of the technology.

By using "perceptual" coding methods such as MUSICAM or ASPEC, it is possible to pass 15 kHz bandwidth audio over a telephone-grade 64 kbit/s channel.

DIAL-UP REMOTES: BANDWIDTH EXTENSION

The era of obtaining equalized broadcast loops at reasonable cost from the local telco has apparently come to an end. Thus, we are motivated to find ways to use the ubiquitous and low-cost dial-up network for program remotes. The problem is that the 300 to 3.4 kHz frequency response and 45 dB or so S/N that the dial network provides is not adequate for modern broadcast needs.

Until we can make use of the emerging digital telephone network, frequency shifting and amplitude companding methods which can be applied to the existing dial-up network may be acceptable for many remote origination projects.

Frequency Shifting Using One Line

Frequency shifting offers a way to squeeze more high frequencies into a line than it will normally pass. More accurately, the process allows different frequencies than the usual 300 to 3.4 kHz to be passed through a standard phone line.

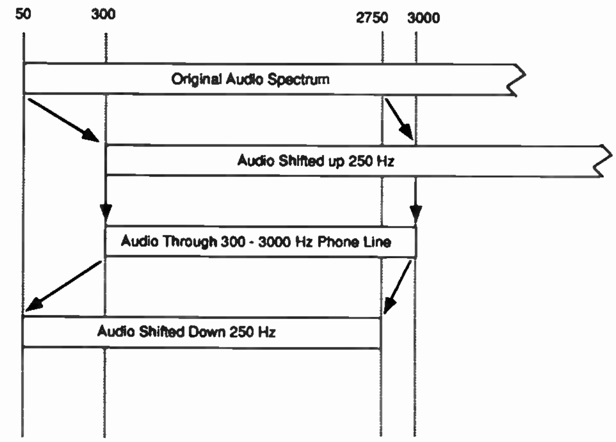


Figure 32. A frequency shifting bandwidth extender allows improved low end response at the expense of a small loss of high frequency audio.

The popular single-line frequency shifting units move all frequencies up by 250 Hz in the encode mode (going into the phone line) and down by 250 Hz in the decode mode. The result is a 250 Hz improvement at the low end at the expense of a 250 Hz loss at the high end. This means a typical phone line's response will be changed by the shifting process to be 50 to 2750 Hz. The 250 Hz loss at the top is not very significant due to the logarithmic nature of audio frequency response.

The shifting function is accomplished by heterodyning the input audio with a low-frequency carrier. The phasing single sideband (SBB) generation method is employed to allow only one sideband to emerge at the output (the carrier and other sideband having been cancelled in the SSB process). Encoding and decoding can easily be accomplished in the same unit, since only a simple signal path change is required for an encoder to decode, and vice-versa.

Subjectively, the resulting frequency-shifter processed audio sounds less "telephone-like." However, the result of improvement at the low end without high-end enhancement is often a somewhat muddy or flat sound.

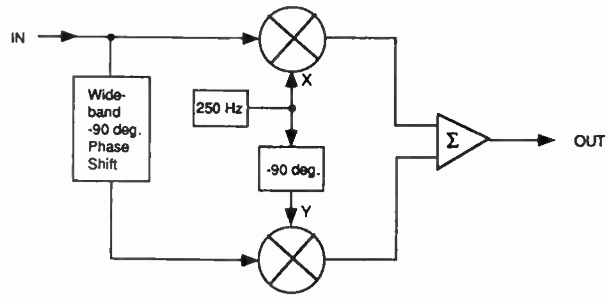


Figure 33. Single-line frequency extender uses SSB techniques to shift the audio 250 Hz at each end. A decode system is shown: signals at X and Y are reversed for encoding.

Comrex Two Line System

Model 2XP Encoder and Model 2XR Decoder

Figure 34. Comrex 2XP two-line frequency extender allows transmission of 50 to 5 kHz audio on two dial-up phone lines.

One approach to improving subjective quality is to boost the high end with a sharp EQ rise above 2 kHz. A parametric EQ or custom filter is preferred so that a high-Q curve can be obtained.

Decoding must be done from a direct input. That is, frequency-shifted audio should be decoded as it comes out of the phone line, before it goes to tape. If the shifted audio is recorded without decoding, with the intention of later playing the tape back through the decoder, interesting but unairable results may occur. These are due to the additional slight frequency shifts introduced by the recorder. The decoder is not expecting to see these, and is quite intolerant of them.

Frequency Shifting Using Two or More Lines

Despite the benefits derived from extending the frequency of a phone line to 50 Hz, the frequency response obtained from a single phone line is still limited to a very "low-fi" 3.4 kHz. Response to 5 kHz is possible with the use of two phone lines (the half-speed transmission technique described below extends response, but cannot be used when real-time transmission is required).

In a two-line extension system, frequency selective crossover networks split the program audio into two bands each of about 3 kHz or so in width. These bands are then heterodyned to occupy the available phone bandwidth using the same kind of process as in the single-line shifters. In the Comrex unit, input audio is separated into bands of 50 to 2400 Hz and 2500 to 5000 Hz. The first band is then shifted up 250 Hz, while the second band is shifted down 2 kHz. A five-band companding system is employed to improve signal-to-noise performance.

Digital signal processing (DSP) has come to bandwidth extension. There is now equipment which uses DSP to obtain a 7.5 kHz passband by using *three* phone lines. Fourier transform methods are employed rather than the SSB technique. The Fourier transform changes a signal from the time to the frequency domain, while the inverse Fourier transform does the opposite. In the DSP shifter, both are used. After a forward transform, the frequency bins are offset (up or down, as appropriate) and the signal changed back to the time domain.

A caveat: The Fourier transformations create enough time delay that off-air monitoring is generally not possible.

Half-Speed Transmission

Another more elaborate approach to high-end improvement is the half-speed transmission technique. With this approach, another encode process is used in addition to the frequency shifting already described; this time operating to improve high-end response rather than low-end. A phone line with fixed bandwidth can

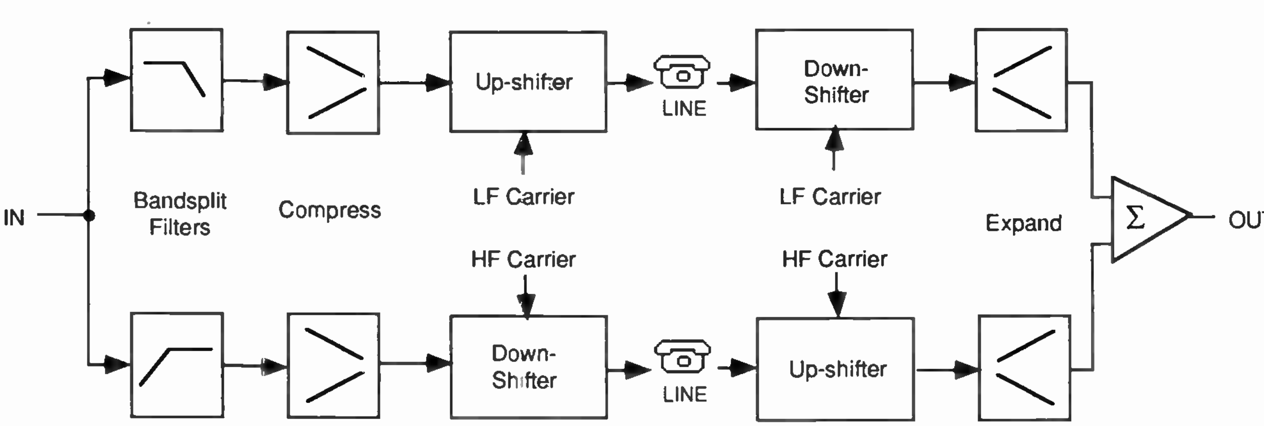


Figure 35. A two-line frequency extender allows a 5 kHz audio bandwidth with two phone lines.

Figure 36. Gentner EFT3000 three-line unit uses digital processing to extend high-frequency limit to 7.5 kHz. Requires three telephone lines.

be made to have more if we are willing to trade off time for better fidelity.

The technique is very simple: the recorder material is played at half-speed; after reception, the received material, having also been recorded, is played back at twice the record speed. What goes through the phone line is nearly unintelligible, since it is half speed audio. Of course, this method cannot be used when real-time transmission is desired, and toll calls cost twice as much: another example of the futility of expecting something for nothing!

The combined shifting and half-speed techniques result in an effective bandpass of 100 to 5500 Hz.

INTERFACING PRODUCTION INTERCOM SYSTEMS TO TELCO LINES

To aid communication with the field crew during television remote projects, it is often desirable to connect the production intercom system to dial-up telephone lines. Smooth integration of live news remote feeds, for instance, requires that production personnel at all locations can communicate with each other in a simple, troublefree fashion. This is especially true when multiple remote sites are involved; as with election coverage, major sporting events, and telethons. Ideally, crews at each location would use the intercom system without regard for the distances involved. Most often, access to the dial-up phone network is available by wire or cellular, so an interconnection of the intercom system to the telephone network may be the solution.

Four-Wire Intercom Systems

Four-wire systems are those in which the two speech directions are kept separated in the switching and distribution process. While it would be possible to use special four-wire telco circuits (or two standard loops)

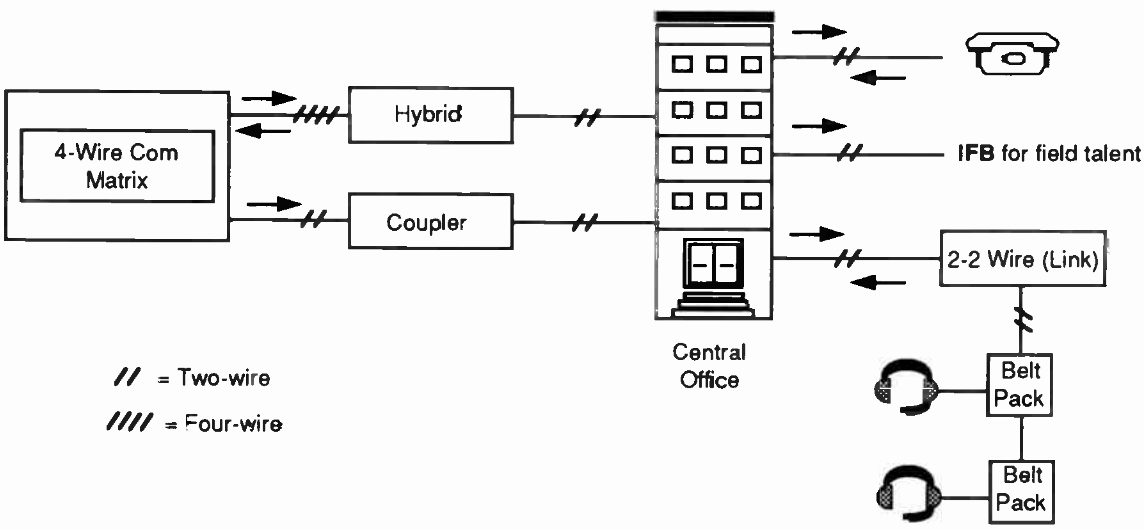


Figure 37. An arrangement which integrates a four-wire switching matrix with telco lines, an IFB feed, and a two-wire "party-line" intercom system for field production work.

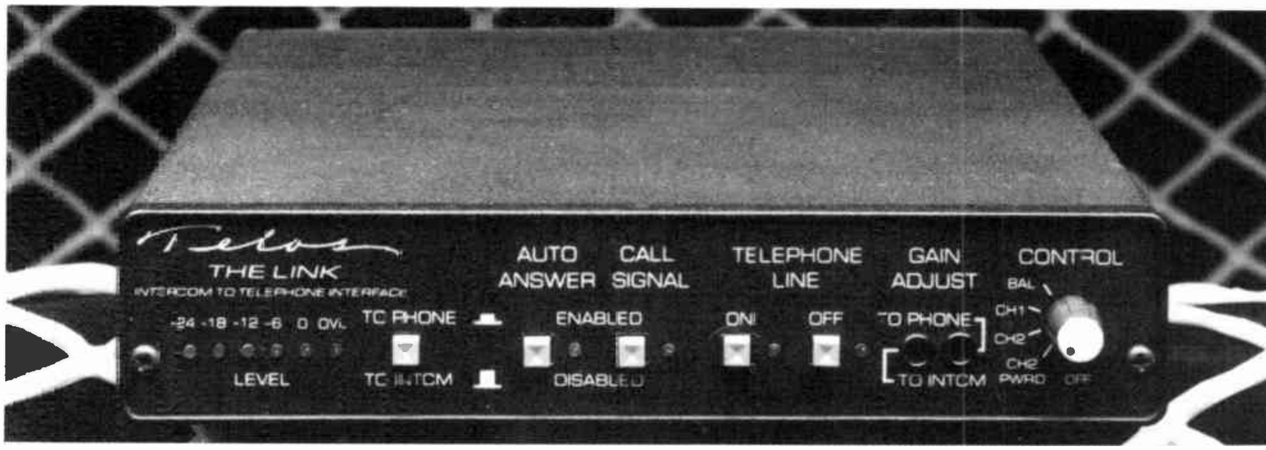


Figure 38. Telos "Link" interface. Single-box solution to interconnecting two-wire intercom systems to dial-up telco lines.

to maintain independent signal paths to remote sites, it is more economical and convenient to be able to use a single phone line. To accomplish this, our task will be to create effective conversion between the two-wire phone line and a port on the four-wire intercom matrix. It will be necessary to separate the intermingled send and receive speech signals on the phone line with a two-to-four-wire converter, or hybrid.

Trans-hybrid loss performance will be important when (1) intercom stations with open loudspeakers and mics are to be used, and (2) when conferencing of multiple telephone lines is desired. In the first case,

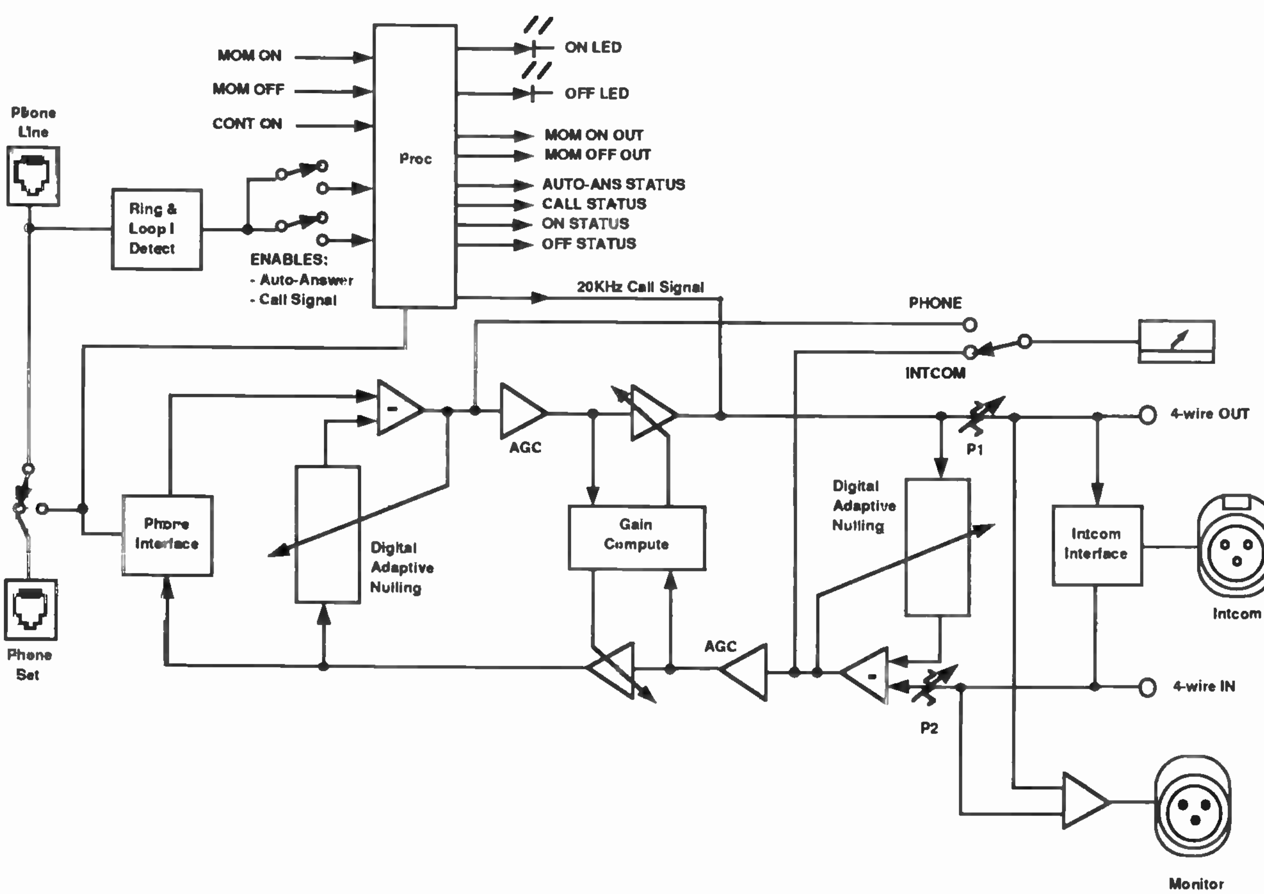


Figure 39. Two-wire intercom-to-telephone line interface (Telos "Link").

the acoustic coupling between the speaker and mic completes a feedback path which includes the hybrid. Clearly, the better the hybrid's isolation, the higher the feedback margin. In the second case, a feedback path exists from each active hybrid through all of the others which are conferenced to it. When the total gain exceeds unity, feedback results. The goal is to have the best possible trans-hybrid loss so that the maximum line-to-line gain may be achieved.

An auto-answer and disconnect function may be required for unattended operation. This circuit responds to a phone line ringing signal by activating the hybrid and de-activates the hybrid when the calling party hangs-up. As discussed in the section on CPC, a dial tone detector may be necessary to ensure reliable operation. The tone detectors are connected so as to respond to signals on the hybrid's separated telco *receive* audio signal. Were this not the case, and the detector was merely connected across the phone line, there would be a major problem when multiple lines are used together in a conference. Why? Because the tones would be conveyed to each line in use (through the intercom switching matrix) from every other line, causing *all* of the detectors to respond to the tones from all of the other lines as well as its own. When one line's interface gets a disconnect, all of the others would turn-off as well! Therefore, there is a critical requirement in this set-up that trans-hybrid loss must be sufficient in order to be certain that any cross-coupling is below the threshold of the tone detectors. The same situation applies with any DTMF detection which is used on a per-line basis.

Two-Wire Systems

These are the popular *party-line* systems such as made by RTS, Clearcom, and others. Here, the interface is going to require two hybrids. The hybrids are connected back-to-back so that the intercom hybrid's receive output is fed to the phone line hybrid's send input and vice-versa. Appropriate gain and processing stages are inserted in the four-wire path. This system is what telephone engineers call a *two-wire to two-wire repeater*.

High quality hybrids are required to prevent feedback. As should be evident from Fig. 38, the signals can feed around the loop and feedback could build-up. This happens when the combined trans-hybrid loss of the hybrids is not at least as great as the gain in the two amplifiers.

As telephone circuits have widely varying and unpredictable end-to-end transmission characteristics, interfacing intercom systems to phone lines without gain and without AGC is not likely to work very well.

BIBLIOGRAPHY

John C. Bellamy, *Digital Telephony*, Wiley-Interscience, 1982. Speech coding, digital phone networks, T-1 systems, etc.

Peter Bocker, *ISDN—The Integrated Services Digital Network—Concepts, Methods, Systems*, Springer-Verlag, 1988. A description of the technology and potential of ISDN from the European perspective where ISDN has had more presence than in the USA.

John L. Fike, George E. Friend, *Understanding Telephone Electronics*, Howard W. Sams & Co., 1984. An excellent tutorial introduction to telephone electronics, with emphasis on the telephone set. Unfortunately, not much on PBXs.

William A. Flanagan, *The Guide to T-1 Networking*, Telecom Library, 1990. A superbly written and very complete description of the technology and use of T-1 service. Highly recommended if you need to learn about T-1. Available directly from Telecom Library at (800) LIBRARY. Incidentally, the Telecom Library is a very useful source for all kinds of material on telephones, from tutorials on 1A2 key systems to the latest on ISDN and digital technology. Call for their free catalog.

William Stallings, *ISDN—An Introduction, Macmillan*, 1989. Accessible introduction to ISDN.

Teleconnect Magazine, monthly. Call (800) 999-0345 for subscription information. Edited by the quirky Harry Newton, this very useful (and entertaining!) source of current information reflects his character. Thinking of buying a new business phone system? This is the place to turn for a description of all of the available equipment. Special issues dedicated to key systems, PBXs, teleconferencing, T-1, etc.

Harry Newton, *Telecom Dictionary*, Telecom Library, (800) LIBRARY. If you're trying to get your ACD's ACIS algorithm to respond to DID and DNIS, you might need to call your ESP or the BOC. But before you do, you should know what all this stuff means! Look it up in this dictionary. Same guy, same quirky style as *Teleconnect Magazine*. Example: "APOLOGIZE—To lay the foundation for a future offense. . ."

Telecom Gear, monthly publication. Subscriptions free at (800) 322-5156. A "magazine" without any editorial content whatsoever—it's all advertising! Here's where people in the business buy telephone sets, PBXs, replacement circuit cards and accessories. You'll find the prices here to be much lower than from local interconnect dealers. Even though geared to the phone industry, most advertisers are willing to sell directly to broadcasters.

Telos Systems, *The Telos Telephone Q&A*, booklet, 1989. Available free of charge from Telos Systems, 1729 Superior Avenue, Cleveland, OH 44114, (216) 241-7225. Q&A format discussion of broadcast interfacing, hybrids, digital telephone techniques, etc. Helpful tips and information for working broadcast engineers.

6.4
Telco Audio Program Service

Skip Pizzi
Broadcast Engineering Magazine, Overland Park, Kansas

Broadcasters and telephone companies (telco) have been allied since the earliest days of both industries, and will likely remain so. The ability to distribute signals over a wide area is what broadcasters do best, and that merges nicely with telco's great ability to transport signals from point to point. This chapter will look at the opportunities for audio signal interconnections afforded to broadcasters by telephone companies and how broadcasters can implement these hookups. The telephone industry continues to undergo its quiet revolution from analog to digital distribution. The two methods will coexist for some time, so both are covered here.

For clarity, the term "circuit" will be used here to designate a telco audio path, supplanting the other commonly heard (and potentially confusing) synonyms of "loop," "private line," "leased line," "program circuit," and the like. Likewise, "telco" will be used to refer generically to all telephone companies, local, and long-distance. LDS (long distance service) will be used to refer generically to all common carriers providing service between local telcos. General comments regarding telco service in this chapter will be limited to program audio circuits. For standard dial-up telco information, see Chapter 6.3.

DEALING WITH TELCO

The best way to minimize problems with your telephone company is to establish a good working relationship with the appropriate personnel, and to understand as much as possible about the company's service and operation. If the station's staff comes across as friendly and knowledgeable, but also professionally firm and businesslike, things should go well. If possible, keep the station's liaison to telco limited to one staff member, and try to deal with the same person at telco as well.

For ordering digital circuits, a recently instituted option is the dataline brokerage service. Here, the station gives its time, place, and quality-of-service requirements to a third party, who books the line for the station at no charge. (Like travel agents, these services receive commissions from the telcos whose circuits they book.) This service can be especially helpful in long-distance applications, where two local telcos (one foreign to the station) and an LDS are involved. Again like travel agents, data-line brokers may also be able to book the service at a lower cost.

ANALOG CIRCUITS

Analog audio circuits are still offered by most local telcos under the schedule shown in Fig. 1. LDSs are phasing out such services, however. Costs for both service and installation continue to increase in most areas. Installation generally requires several hours of an experienced technician's work to equalize the line, as opposed to digital services, which take much less time and trouble to pass spec at installation. Therefore, analog circuit installation costs are often prohibitive, and digital alternatives may prove more economical.

On the other hand, some local analog circuits will provide wider bandwidth than what was ordered, whereas digital services typically cut off exactly as specified. Increasingly, though, analog services that have to pass between telco central offices (COs) or switching centers will make the trip bundled on an interoffice digital carrier anyway, whereupon excess bandwidth will probably be removed.

Obtaining And Testing Telco Lines

As Fig. 1 explains, analog circuits are available in a variety of bandwidths, and under temporary or permanent status. Check with your local telco to see which remain available, and for their rates. Installation

Class of Service	Approx. Bandwidth	Full- or Part-Time
Type 6002	200-3,500 Hz	PT
Type 6003	200-3,500 Hz	FT
Type 6004	100-5,000 Hz	PT
Type 6005	100-5,000 Hz	FT
Type 6006	50-8,000 Hz	PT
Type 6007	50-8,000 Hz	FT
Type 6008	50-15,000 Hz	PT
Type 6009	50-15,000 Hz	FT
Type 6010*	Dual 50-15,000 Hz	PT
Type 6030*	Dual 50-15,000 Hz	FT

*These classifications used only for long-distance (interLATA) service.

Figure 1. Classes of service for analog telco audio circuits.

charges for permanent lines may be much higher than those for temporary service, because telco may want to actually install new wiring rather than permanently occupy any pairs on its existing network cables.

Type 6008 and 6009 (15 kHz circuits) service may be ordered as a stereo pair, incurring a one-time installation surcharge for "stereo conditioning." (Some telcos also offer it for 6007/8 [8 kHz] service.) This insures that both lines are routed together throughout their runs, so that interchannel phase differences will be minimized. A third line can be ordered for backup, and this too should be included in the stereo conditioning. Although billed routing ("as the crow flies") may be a short distance, actual routing of the circuits may be rather longer and indirect, providing ample opportunity for phase differences to occur. (Approximately 5 μsec time difference occurs for every mile of path length difference.)

Lines should be ordered well in advance of your need for them. Check with your telco for their preferred lead time. Always specify a start date at least one business day earlier than the actual requirement, to allow time for your own tests on the lines to be performed. Check frequency response, signal-to-noise ratio (SNR), distortion, and headroom. For stereo pairs, check relative phase response and polarity.

Frequency response should be at least within ±3 dB of what was ordered. Be sure to check outside the passband, because response may not roll off but rather rise beyond the cutoff frequency. On occasion, in order to get a line to meet specifications, telco equalizers may be used to boost the extreme frequencies, and if done improperly, the response may indeed be flat to the cutoff frequency, but then rise for another octave or more before finally rolling off. This will result in audible consequences from the altered response and reduced headroom, particularly if a noise reduction system that preemphasizes high frequencies is used on the circuit.

The proper procedure for testing lines is to use sinewave test tones of 400 Hz or lower to be fed at the telco program operating level (POL) of +8 dBm. Frequencies above 400 Hz must be sent at the telco "test level" of 0 dBm to minimize crosstalk into other circuits via capacitive coupling of higher frequencies. To keep things simple, run all frequency response tests at 0 dBm.

For measuring noise, telcos use an unusual approach. They consider a noise level of −90 dBm to be "absolute quiet" or noise-free, and measure noise from that reference point. The unit used is dBrn (rn for reference noise). Therefore, a −50 dBm noise level would be called 40 dBrn by telco. If the telco's specified audio reference level of +8 dBm is used by the customer on this circuit, a 58 dB SNR is achieved in this case.

The greater the distance an analog circuit travels, the noisier it becomes. The wider a circuit's bandwidth, the quieter it needs to be. Although specs vary between telcos, noise specifications generally reflect those observations. A typical noise level for local 15 kHz circuits is 33 dBrn or lower, providing 65 dB or better SNR. Long distance analog circuit noise levels generally hit 40 dBrn (58 dB SNR) for 15 kHz circuits. Noise levels for a 5 kHz service is around 46 dBrn (52 dB SNR) for local and 56 dBrn (42 dB SNR) for long distance service. Again check with your local telco for its specifications, and always verify that they are met.

Some telcos offer a lossless or zero-loss option where the circuit acts as a unity gain device, in contrast to a standard circuit, which may exhibit up to 30 dB of loss. (Audio signals in copper cable drop at about 1 dB/mi from broadband resistive losses. Frequency-selective reactive losses occur at greater rates, but these are compensated for by receive-end equalization.) The additional telco amplifiers required by lossless lines may have a detrimental effect on the distortion and headroom performance, and may not do much to reduce the overall static noise floor, but they are useful in densely trafficked urban areas where crosstalk and impulse noise is prevalent.

Static or random noise is far less objectionable than those coherent noises sometimes found on telco circuits. Such noises can be caused by capacitive coupling between adjacent pairs in multipair cable, dial pulses and other switching, inadequate common-mode rejection, and carrier-beating (causing high-pitched "sings" or tones). Because of this, circuits should be carefully auditioned at the receive-end for a while after installation, without any audio fed from the send end. Check circuits again prior to each on-air use. Report any crosstalk or impulse noise problems to telco at the first sign of trouble.

Although there are usually no published specs for distortion or phase response on telco circuits, total harmonic distortion (THD) should be <0.25% on 15 kHz lines. For stereo pairs, relative phase response should be within 30° across the passband. Widely divergent frequency response between the two circuits in a stereo pair is a tip-off to check phase response carefully. When a spare third circuit is ordered, the two closest to each other in frequency and phase response should be designated as the main pair, assuming distortion and noise are equivalent across the three.

The maximum level guaranteed on telco circuits is

+18 dBm, only 10 dB of headroom above the +8 dBm reference level. A sensible alternative is to use +4 dBm as a reference level (most professional audio hardware uses this level anyway), thus allowing a more sensible 14 dB of headroom, at the expense of 4 dB less SNR, generally a worthwhile tradeoff.

Interfacing Procedures

Figs. 2 through 6 illustrate some do's and don'ts of audio interfacing to analog telco circuits. This is one area of today's audio where 600 impedance matching is still important. Telco equalizes its circuits for flat response under the conditions of a 600 source impedance and 600 termination. Because of the reactive

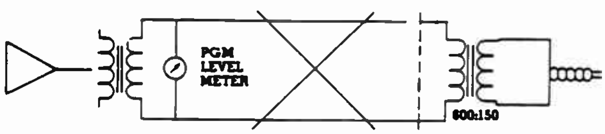


Figure 2. DON'T meter the signal across the input to the repeat coil (transformer at right) and DON'T feed the repeat coil directly from another transformer, if possible.

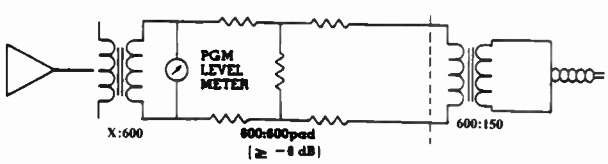


Figure 3. DO place a pad between the output device's transformer and the repeat coil; and DO place a level meter before the pad, calibrated for the voltage across the pad's output when terminated with a 600 ohm resistor.

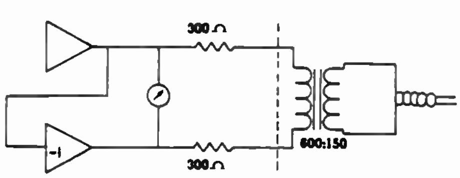


Figure 4. DO feed low-impedance sources (e.g. op-amps) through a 600 ohm differential balanced pad; and DO place the level meter before the pad, calibrated as Fig. 3.

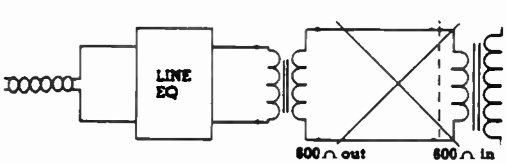


Figure 5. DON'T terminate the receive-end repeat coil (transformer at right) with another transformer, if possible. Its loading may vary with frequency from the true 600 ohm resistive termination used in line-up causing level and frequency response variations.

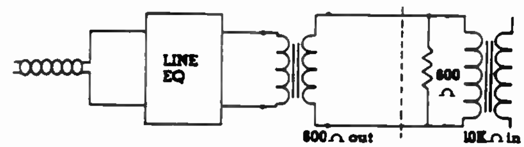


Figure 6. DO terminate the receive-end of a program circuit with a 600 ohm resistive load and bridge the load with a high-impedance transformer or active balanced input. Note: Common mode rejection of the transformer or active input must be considered if the distance from the repeat coil is great or is near other lines which induce cross-talk. Ideally, the secondary of the repeat coil should be resistively terminated. An active, balanced input circuit does this nicely, provided any RF is bypassed before the first stage of amplification. An alternative is the use of an input transformer with a high-impedance, bridging input. This allows the 600 ohm resistor, as shown in Fig. 6 to match the repeat coil. Another version, often seen on the input of broadcast line amps and modulation limiters is a 600 ohm h-pad; effectively the reverse of Fig. 3. Following the steps shown in Figs. 2-6 usually insures that transmission loss and frequency response closely match those of the phone company set up.

components in long wired paths (they act like transmission lines), varying these impedances will affect frequency response. Most contemporary audio equipment expects a bridging interface arrangement, so output source impedances are typically much lower than 600, and input impedances are much higher. Yet telco does not guarantee flat response without 600 conditions.

If a typical low impedance mixing console output is used to feed a telco circuit directly, it will generally cause a rising high frequency response to be received. This is not telco's responsibility, but rather the customer's. It is therefore essential that the device directly feeding the line have a 600 source impedance, and be capable of driving such a balanced load to at least +18 dBm across the passband. Many mixers—even expensive ones—are not designed to do this, so an appropriate outboard line amp or distribution amplifier (DA) fed by the mixer should be used.

A DA here will also provide isolation between the mixer and any other inputs, and more importantly, between all these devices and the telco circuit. Drive the telco circuit with its own DA output—no other inputs (even bridging ones) should be paralleled to it. DC voltages may appear on telco circuits, so this isolation is essential. For the same reason, a transformer-coupled DA is the better choice over a transformerless design in this application. High-quality transformers can be quite helpful on remotes, especially when interfacing with telco. On most 15 kHz circuits (and some others), though, telco repeat coils appear at each end, providing this isolation for you. In this case, active, balanced outputs can be used safely, but flat frequency response is only guaranteed if they are interfaced to the repeat coil in the manner shown in Fig. 4. A caveat here: remember that the active DA still has to drive a 600 balanced load to

+18 dBm across the passband, and many popular operational amplifiers do not have the current drive capability to do so. Do not conduct lengthy tests at +18 dBm to avoid crosstalk to other circuits.

The "H"-pad in Fig. 3 does not provide a 600 source to the line by itself. If the secondary of the output transformer on the left of the diagram were not 600, the pad would not set things right. The pad is there to provide isolation of the meter or a test oscillator from the repeat coil.

A simple test to determine the actual source impedance of a device's output is as follows. Place a bridging-input voltmeter across the device's output while feeding 1 kHz at reference level. (No other inputs should be connected.) Terminate the output with various resistances, and watch the meter. The device's source impedance is the value of the resistor that drops the level by 6.02 dB. Then verify that this level drop is consistent at other frequencies.

Note also that the meter on the transmit end shown in Figs. 3 and 4 should be used only for initial absolute level calibration, and not for relative levels in frequency response tests. For verification that consistent level is being transmitted at each frequency during response tests, and for realistic end-to-end results, put the oscillator further upstream, and use a more isolated meter for reference. This typically is done by feeding the oscillator into a properly telco-interfaced mixer or DA input (or using the mixer's inboard multifrequency oscillator, if it has one), and recalibrating for oscillator drift using the meter on the oscillator, mixer, or DA; not the downstream meter shown in the figures.

Failure to follow the procedures outlined above will result in poor frequency response, distortion, calls from telco warning of too high a level, too much noise, and general unhappiness with the service.

DIGITAL AUDIO CIRCUITS

As with much of the technological progress in our industry, the digital audio revolution that is now taking place in the broadcast world began at the telephone company. The transmission of data is nothing new to telco, but the high data rates required for digital audio transmissions had previously rendered the availability and cost of such service out of the practical range. Data compression, or bit rate reduction systems, have made possible broadcast applications of data transmission paths that previously were only useful for computer interconnection. Reductions from earlier data rates for digital audio transmissions of 4:1 or higher have become commonplace and greater compression ratios may become availble in the near future.

Data Compression

Although these data compression algorithms are viewed today as major breakthroughs, history will likely look upon them as natural evolutions, and consider the earlier linear pulse code modulation (PCM) systems as dinosaurs. While the straightforward nature of linear PCM may have been helpful in making the transition from analog systems, especially where bandwidth was cheap and available, it is an inefficient method for encoding digital audio. The resolutions of today's linear PCM systems are often overkill in terms of the actual needs of most listeners, and significant reductions in actual transmitted data can be achieved by applying data compression algorithms to the datastreams that linear PCM conversion produces. At present, linear PCM of as high a resolution as economically feasible is still a good idea for the original conversion of analog signals to the digital domain, and for any digital production or signal processing. But for signal delivery systems, data compression is an appropriate tool.

Earlier compression systems (again pioneered by telco) used a purely statistical or numerical analysis of the datastream's coding redundancies. Current systems acknowledge the limits of the human listener's sensory perception. The study of masking effects in the human auditory system has now become the starting point for algorithm design. Today's so-called perceptual coders are therefore based on psychoacoustic models, and owe their coding efficiencies to an appreciation of the audience's tolerances.

Data Rates

Like their analog counterparts, digital circuits come in various bandwidths. But rather than specifying cutoff frequencies of the audio passband, telco specifies digital circuits in terms of their data rates. Broadcasters need to think both in terms of kilohertz of bandwidth, and in kilobits or megabits per second (kbit/s or Mbit/s) of data transmission.

A data rate must be considered first in linear PCM terms, and then in its subsequent compressed form. In the linear mode, the data rate of a given signal is simply its sampling frequency (in Hertz) multiplied by its resolution (in bits/sample). For example, CD-quality audio uses 44.1 kHz sampling at 16 bit/sample resolution, requiring a 705.6 kbit/s data rate, per channel (stereo requires doubling that data rate to 1.411 Mbit/s), before adding any error correction overhead. A digital audio compression algorithm capable of 4:1 data rate reduction takes that linear PCM signal and reduces its resolution to an average of 4 bit/sample (while leaving its sampling frequency alone), therefore providing a 192 kbit/s data rate. Fig. 7 shows some other data compression ratios for audio, and their resultant data rates, at several common sampling frequencies.

Because the sampling rate is not changed by the digital compression system, frequency response and time-domain performance retain the same specifications typical of most linear PCM conversions. Delay is introduced by these codecs, however, and it is generally in direct proportion to the amount of data compression applied.

Classes Of Digital Service

Current telephone company installations and tariffs can provide a variety of services in most domestic and

RESOLUTION (av. bits/sample)	COMP. RATIO	OUTPUT DATA RATES (kbit/s) f_s = 48 kHz	f_s = 44.1 kHz	f_s = 32 kHz	f_s = 16 kHz
16	1:1	768	705.6	512	256
4	4:1	192	176	128	64
3	5.3:1	144	132.3	96	48
2.67	6:1	128	117.7	85.4	42.7
2	8:1	96	88.2	64	32
1.45	11:1	69.6	64	46.4	23.2

Figure 7. Data compression table showing range of compression ratios currently under consideration and their resultant output data rates at a variety of sampling rates (f_s). Audio bandwidth is approximately one-half o (f_s). Data rates shown are for a single audio channel (mono).

some international locations, with more new services continually being deployed in many cities. Fig. 8 shows current services and their data rates.

One of the most important differences to the broadcaster between analog and digital circuits is that most digital services are provided bidirectionally. This fact should not go unnoticed when making cost comparisons. Although interfacing hardware for return path channels must still be provided, their circuits require no separate costs or orders.

Note also that telco has always had its own insider vocabulary and set of acronyms, but since the introduction of digital services, this has expanded dramatically. Many of these are explained in the Glossary of Digital Telco Terms at the end of this chapter.

For overseas links, rough equivalents to each of the domestic services shown in Fig. 8 do exist outside the U.S., but their actual data rates differ. Format conversions are therefore required for international transmissions, but most LDSs can handle this for the broadcaster.

DS1 or T-1 Service

Digital audio transmission on DS1 (or T-1) lines has become widely available, and is often cheaper than standard analog circuits in both service and installation charges. (See Glossary for distinction between DS1 and T-1 nomenclature.) DS1 is a bidirectional 1.544 Mbit/s serial data link. The data rate calculations above show how DS1 can carry a linear PCM stereo audio signal, or several such compressed channels. (Figs. 24 and 25 in Chapter 6.3 may be helpful.)

DS1 service is extremely reliable. Its bit error rate (BER) of 10^{-9} (the probability of error reflected by the specification of no more than one erroneous bit in 10^9 transmitted) is the lowest of any available. By way of reference, IEEE and CCITT have both established 10^{-6} as the BER required for data customer satisfaction.

SERVICE OFFERED	DATA RATES (bits/sec)
Switched 56	56k
DS0	64k
DS1	1.544M
DS2	6.312M
DS3	44.736M

Figure 8. Current U.S. digital telephone data services and their data rates.

The data carried on a DS1 circuit is actually a multiplex of 24 data channels, or slots, of 64 kbit/s each. (An additional 8 kbit/s is reserved for sync data.) These individual 64 kbit/s slots are called DS0 channels. For standard telco T-carrier use, each DS0 carries a digital voice-grade circuit, using the so-called μ-Law (nonlinear 8-bit) algorithm on 8 kHz-sampled audio. (See Chapter 6.3 for more on this.) When a customer leases a DS1 circuit, it can be configured to carry any bandwidth channel that DS1 hardware is available for (3.5 kHz, 5 kHz, 7.5 kHz, 15 kHz) in any combination, up to the customer usable data limit. When a customer leases a full DS1, a telco may take one DS0 slot for framing and other overhead, in addition to the 8 kbit/s synchronization slot, leaving around 1.4 Mbit/s for customer data. Check with your telephone company for its exact rate.

A rack of coding and multiplexing hardware appears on each end of the DS1 line, usually as customer provided equipment (CPE), and the circuit can be reconfigured simply by changing the appropriate cards in the proper slots in the racks at both ends. These reconfigurations can be performed by the customer at any time, without telco involvement or notification.

Unlike the labor-intensive installation and equalization of an analog circuit, putting in a T-1 circuit has become as routine as a standard dial-up telephone service installation. This and the capacity gluts in some areas continue to lower costs for DS1 service. Customers' use of digital compression systems on DS1 channels will only increase this economy. Whereas a 15 kHz audio channel had in the past typically required six DS0 slots, current hardware implementing perceptual coding reduces this to two DS0. Further reductions will likely occur in the future.

Fractional T-1

In some areas, Fractional T-1 service is becoming available, generally for intraLATA (local) applications only. This service allows a customer to lease only the number of DS0 slots on a DS1 circuit that are needed for a particular application. Although installation charges will be about the same as for a full DS1, service charges

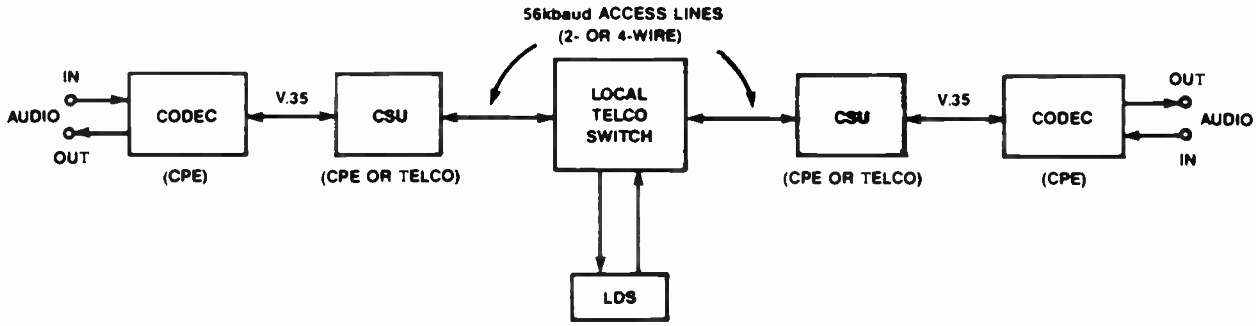


Figure 9. Block diagram of Switched 56 circuit path. (See *Glossary of Terms* for acronym definitions.)

may be drastically reduced for many remote audio applications.

Switched 56

Another recent telco digital offering of potential use to broadcasters is the Switched 56 service. This facility is now available in many metropolitan areas in the U.S., and from some long-distance carriers. It provides a bidirectional 56 kbit/s data path for use with dial-up terminals. Like plain old telephone service (POTS), a monthly service/network access fee is charged (which often includes at least some free local calling), with long-distance calls billed by the minute, at rates similar to regular dial-up long-distance service. In some cases, local calls are also billed for connect time. An installation fee is also typically charged for initiation of service. A switched channel service unit (CSU), the equivalent of a telephone instrument and data interface, is provided by the telephone company, or may be purchased by the customer. It allows voice or data interconnection, and dial-up routing of the data path to any other similarly equipped destination on the network. For broadcast use, additional codecs are required, as shown in Fig. 9, to feed wideband audio. These codecs are not available from telcos, and must be purchased by the customer. Codecs are available now in 7.5 kHz versions (using Consultative Committee of International Telephone and Telegraphy (CCITT) G.722 coding), with wider-bandwidth units implementing more sophisticated perceptual coding expected soon.

An obvious savings is possible with such a switched approach, since dedicated circuits need not be installed between a radio station and all of its remote sites. The station and each site need only to be wired for the switched service (with each line going to the CO only, and not end-to-end) and the station can dial up any remote site as needed. The station may require two or more lines for accessing multiple remote sites simultaneously or in quick succession.

CSUs are available in two-wire or four-wire versions, with two-wire types costing less. Unfortunately, the choice between two and four-wire operation is not up to the customer, but to the local telco serving the area. (See Fig. 10.)

Switched 56 is available to Europe and Japan, as well, although it is a 64 kbit/s service in most overseas locations. Conversion hardware is available for such an application. Like the higher speed channels mentioned above, no international standards exist for these switched services today, although they are currently under development.

Some telcos offer a similar unswitched service, in which a single DS0 channel can be leased on a monthly basis. For heavy point-to-point users, this may be more economical than a switched approach. It may also be available in some areas where switched service is not as yet. As single DS0s, these operate at 64 kbit/s, and their terminal hardware is less expensive because it need not accommodate switch signalling.

ISDN Service

The next level of service that U.S. telephone companies will be providing is the Integrated Services Digital Network (ISDN). Already available in a few locations, basic rate ISDN provides two 64 kbit/s paths ("B-channels") and one 16 kbit/s circuit ("D-channel"). This service is also referred to as 2B + D. Primary rate ISDN service will provide twenty-three 64 kbit/s B-channels, and one 64 kbit/s D-channel (23B + D); this is roughly equivalent to a full DS1 circuit (1.536 Mbit/s). The distinction between B- and D-channels

TELCO	SERVICE NAME	WIRE MODE
Ameritech[1]	Switched Digital Svc. (SDS)	2 or 4 wire[2]
Bell Atlantic	Switched 56	4 wire
Bell South	Accupulse	2 wire
NYNEX	Switchway	2 or 4 wire[2]
Pacific Telesis	Centrex IS[3]	2 wire
Southwestern Bell	MicroLink I	2 wire
USWest	Switchnet 56	4 wire
AT&T	Accunet 56	
US Sprint	VPN 56[4]	

NOTES: 1) Service varies from widely available (Illinois) to nonexistent (Ohio).
2) Wire mode varies within service area, depending on local swtich hardware.
3) Actually an ISDN service, providing two basic rate lines (four B-channels).
4) Interfaces only with Bell South or PacTel, and incompatible with AT&T.

Figure 10. Brand names under which Switched 56 service is marketed by RBOCs and LDSs. (Most are registered trademarks.) Wire mode of switch operation at RBOCs is also shown.

incorporates an important change from current switched networks (including most Switched 56 technology).

With ISDN, data paths are separated from the signalling paths used for all the logistics of call-directing and other switch control. This technique is referred to as out-of-band signalling, and is contrasted to the inband signalling in use with today's switched systems, wherein the switching is controlled by pulses or tones on the same path and within the bandwidth of the audio or data being transmitted. For digital transmission, the datastream interruptions that inband signalling demands will be eliminated with ISDN.

ISDN will be a bidirectional, customer-switched service, operating as a dial-up, billed-by-the-minute data network, allowing both circuit-switched and packet-switched operations. Its multichannel nature will allow simultaneous voice and data or other combinational applications. (Billing will be separate for each of the multiple lines, however, and identical routing for multiple lines is not guaranteed.) ISDN is eventually intended to replace the current dial-up telephone system, but most consider that a far future possibility. In fact, ISDN is characterized more as a concept than as an actual single service, since it is in effect the mating of several other separate services (some already available individually) with higher capacity switching hardware and new system architectures. While new fiber paths are being laid for more and faster intra- and intercity carriers, ISDN can continue to use existing network copper to the end user.

At present, some telcos offer ISDN service only to large, data-intensive businesses who order hundreds of lines to a single location, but acceptance of orders for single stations is on the horizon. When this ultimate stage is reached, and service is widespread, 2B + D ISDN may be the delivery method of choice for stereo 15 kHz broadcast audio plus communications to/from a remote site. Universal domestic deployment of ISDN is expected by the late 1990s. (See also Chapter 6.3.)

ISDN will certainly not be the last improvement in telephone service to become available. Telephone companies are already planning subsequent system evolutions, such as bandwidth on demand, in which the user will select the data rate required at the moment (and be billed accordingly) in a dynamic, switched network.

COMMUNICATION LINES

For remote broadcasts, circuits for transmission of program audio often must be complemented with communication lines. With analog circuits, these are usually standard dial-up phone lines, but the bidirectional nature of digital circuits usually allows backfeed and talk paths to be carried on the return side of the program lines.

Analog Applications

When using local analog circuits for program transmission from remote sites, one or more standard dial-up telephone lines can be ordered for backfeed audio and communications to the site if off-air monitoring is insufficient. Alternatively, an RF communication system (hand-held radio or cellular phone) could be used. If wired telephone service is chosen, the station should provide its own instrument, which may be equipped with visual-ring/bell-cutoff and headset attachment.

There are also some variations to consider over a standard seven-digit dial-up service. A communications phone line (comline) can be set up as an off-premises extension (OPX) of the station's private branch exchange (PBX) system, or the line can be a dedicated private line (PL) that only runs between the station and remote site, in an unswitched form. In the latter case, such a PL is normally equipped with automatic ringdown (PLAR), in which special hotline type phones are installed by telco, having no keypad or dial; the phone at one end rings whenever the phone at the other end is picked up. Services in this area (and their specific nomenclature) may vary among telcos. These services are also provided by a different department than the one that handles program audio circuits, so coordinate carefully. Typically, this part of telco is the same one that handles all dial-up installation and service calls from the general public, so lead time requirements and service call response time will often be longer than the audio circuit department's. (See also Chapter 6.1.)

Digital Applications

Using the return path on digital circuits for communications and backfeeding requires additional terminal hardware on each end. Channel configuration need not be the same as the circuit's other direction, and because narrowband (i.e., lower data rate) channels are usually all that are required, the number of communication channels on the return path can exceed the number of program audio channels on the transmit side, if necessary. Another possibility is the luxury of wideband backfeed and communications, such that the talent at a remote site could hear audio coming back in the same quality that the station was receiving from there.

For permanent interfacility hookups, a DS1 circuit could carry both audio transmissions between sites, and telephone service, such that both locations could have their phones connected to the same PBX. If the locations are far apart, a foreign exchange (FX) arrangement can be made, whereby one facility can place calls to the other facility's area without incurring toll charges. The PBX system is programmed to recognize area codes of the two cities involved (or special internal access codes), and it directs appropriate calls to FX lines on the DS1 rather than placing them as regular long-distance calls.

COST OF SERVICES

For local service, there is normally only one provider of telco circuits, analog or digital, which is, of course, the local telephone operating company. So there is little competitive choice in the matter for intraLATA service pricing. Check with your local telco frequently, though, to see whether analog or digital service is more economical, for both permanent and temporary applications. Analog services continue to increase in price, while digital services decline, at varying rates around the country, so it is recommended to keep a close eye on these changing rates in your area.

Long-distance service is another matter altogether. There is significant competition for digital service on the interLATA portions of such a hookup, so prices are kept low, and continue to drop. The low profitability of interLATA analog service has kept competition out of this market, creating the opposite effect. Shop around for the best deal on interLATA digital service, or use one of the dataline brokers mentioned earlier—the list of common carriers in this business is long, and continues to grow. Secondary services may also be offered, such as switching and monitoring.

In some cases, however, analog service may still be the only available method, so it will be helpful to understand how billing works. Fig. 11 shows how local and long distance analog circuits are billed. Digital circuits follow the same basic theories, but with the significant differences of bundling of multiple channels and bidirectional service, making actual costs per circuit vary with each application, but generally increasing the value to the user.

The importance to radio of interconnection, immediacy, and fidelity underscores the importance of a long-term relationship between broadcasters and telephone companies. Good awareness of and rapport with telco are essential to the daily work of the broadcaster. In today's context, that means keeping abreast with the changes in telecommunications that affect broadcasting as the digital revolution rolls along.

LEGEND

CM — Channel Mileage (per air-mile billing basis)
CO — Central Office (LEC switching center; also called SWC—Serving Wire Center)
CT — Channel Termination (flat rate billing)
IOC — Interoffice connection on LDS network (distance-sensitive)
LATA — Local Access and Transport Area (telco service zone)
LDS — Long-Distance Service (common carrier)
LEC — Local Exchange Company (local telco)
POP — Point of Presence [LDS's office in each LEC; also called SO (Serving Office)]

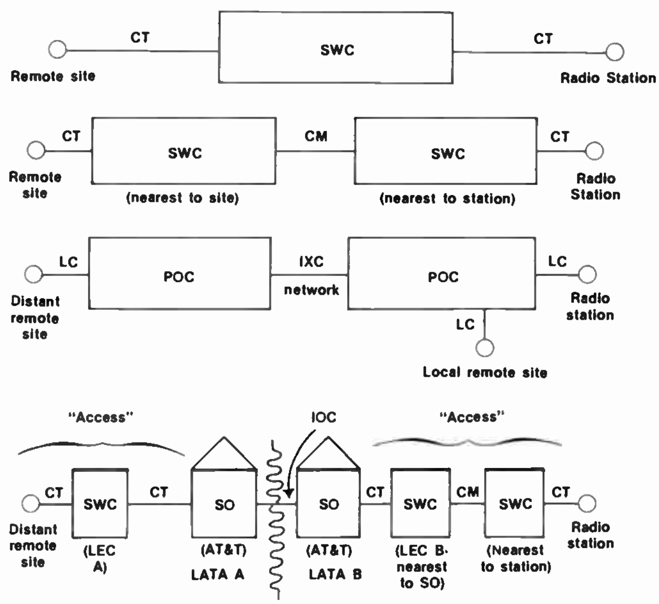


Figure 11. Billing methodology for analog circuits. In (a), the two ends of a local circuit are served by the same central office, or "rate center." In (b), two different central offices are involved. In (c), long-distance service is depicted. LEC connections to LDS are referred to as "access." Access shown in a LATA B will be more expensive because the radio station is in a different rate center of LEC B than the long-distance carrier's POP.

GLOSSARY OF TELCO TERMS

ADPCM—Adaptive Differential Pulse-Code Modulation. A form of digital coding more efficient than linear PCM because it only codes the difference between one sample and the next, instead of assigning a fully discrete value to each sample. It also adapts its coding to the signal values currently under process. Considered a form of statistical data compression.

AMI—Alternate Mark Inversion. The binary modulation code used by the telephone company for data and digital voice transmission. It uses RZ coding in an alternate bipolar scheme, with logical zeros corresponding to zero volts, and logical ones alternating between +3V and −3V. (The first logical 1 produces a +3V output, the next 1 produces −3V, the next +3V, and so on.) Self-synchronization is possible with this approach, but the number of continuous zeros must be limited.

Baud—Bits per second.

B-channel—In ISDN service, a channel designated for customer data transmission, uninterrupted by any signalling data.

Bellcore—Bell Communications Research. The R&D firm that feeds technology and standards to the RBOCs, and is funded by them. Formerly Bell Labs.

Carrier—In telco parlance, refers to a multiplexed digital interoffice signal, containing many individual calls or signals in a single cable or fiber.

CCITT—Consultative Committee of International Telephone and Telegraphy. The international standards-setting organization for telephone systems, established by the UN.

Codec—Coder/decoder. Any device which includes digital transmission/encoding and reception/decoding circuitry in the same chassis.

CPE—Customer Premise Equipment. Refers to any network interface hardware not provided by telco.

CSU—Channel Service Unit. Terminal hardware for a telco data line, either CPE or telco-provided. Also referred to as CSU/DSU (DSU = Data Service Unit) in T-1 applications. Interfaces unipolar NRZ computer-style datastreams to the RZ bipolar (AMI) telco data format. A switched CSU includes a keypad for call direction and other switch control.

D-channel—In ISDN service, a channel designated for signalling data only.

DDS—Dataphone Digital Service. The first telco data service in the U.S., originated in the mid-70s by AT&T.

DS0—Digital Service 0. A 64 kbit/s data channel.

DS1—Digital Service 1. A 1.544 Mbit/s data service usually configured as 24 DS0 channels plus an 8 kbit/s sync channel.

DS2—Digital Service 2. Four DS1 channels multiplexed together for transmission. Generally reserved for telco interoffice transmission, and not offered to customers directly.

DS3—Digital Service 3. Twenty-eight DS1 channels multiplexed together with additional control data, providing a data rate of 44.736 Mbit/s (generally quoted as 45 Mbit/s). Used for compressed NTSC and high-definition video distribution.

First-mile—Refers to the signal path between a program's origination site and its entry point to a common-carrier's network or a private satellite uplink. Usually a terrestrial RF link or a local telco circuit.

G.722—A CCITT standard for audio data compression. It uses two-subband ADPCM coding to put 7.5 kHz audio into 64 kbit/s.

InterLATA—Refers to telco service or rates between LATAs, or long-distance service.

IntraLATA—Refers to telco service or rates within a LATA, or local service.

ISDN—Integrated Services Digital Network. A new telco service designed to eventually replace POTS with flexible digital service. It will be offered in basic rate (2B + D) service, intended for home use, and primary rate (23B + D) service for business customers.

J.41—A CCITT standard for digital audio encoding. Using 14-11 PCM encoding (14 bits for lower level signals, 11 bits for higher level signals), it places 15 kHz audio on 384 kbit/s.

Last-mile—Refers to the short-haul signal path between a long-distance network terminal point (or private satellite downlink) and the customer's receive point. Usually a local telco circuit.

LATA—Local Access and Transport Area. The service area of a local exchange company (LEC).

LDS—Long Distance Service. A carrier of long-distance (inter-LATA) telecommunications, such as AT&T, MCI, US Sprint, and others.

LEC—Local Exchange Company. A local telco. Each RBOC contains one or more LECs. Also refers to independent, non-Bell local telcos.

Mark—The telco term for "high" level data pulse, usually corresponding to logical 1. (See Space.)

NRZ—Nonreturn-to-zero. The most basic form of binary modulation coding, in which logical 1s and 0s are directly represented by high and low levels respectively. Because no level transition occurs between continuous strings of like logical values, this form of modulation is not self-synchronizing, and requires an external bit clock output for synchronous operation.

Packet switching—A sort of data "party-line," in which data is transmitted in addressed bursts or "packets," occupying the transmission channel only for the duration of the packet, after which the channel is free for other packets to or from the same or other users. Many users can be interconnected to the same line, but data can be sent discretely to each destination.

PDN—Public Data Network. Telco data services, including both switched and leased lines.

POTS—"Plain Old Telephone Service." Refers to the public switched telephone network (PSTN).

PSTN—Public Switched Telephone Network. The standard dial-up phone system.

RBOC (or **BOC**)—(Regional) Bell Operating Company. The seven "Baby Bells" created when AT&T divested itself of its local telephone operations.

RZ—Return-to-zero. A form of digital modulation coding in which logical ones and zeros are directly represented by high and low levels respectively, but where coding output returns to low level following each high pulse.

Slot (or Timeslot)—Generally refers to a DS0 channel within a DS1 signal.

SMDS—Switched Multi-Megabit Data Service. A future high-speed switched data network, operating at DS1 to DS3 rates.

SONET—Synchronous Optical Network. An upcoming telco standard for ultrafast data transmission, operating at speeds of 150 Mbit/s to 2.4 Gbit/s.

Space—The telco term for "low" level data pulse, usually corresponding to logical 0. (See Mark.)

Switch—Generic name for any telco call routing and connection hardware.

Switched 56—A switched digital service offering 56 kbit/s data service on a dial-up network, available in an increasing number of areas.

TA—Terminal Adapter. The term for a CSU in an ISDN system.

Tariff—A schedule of services and their prices that a telephone company will provide to a given service area, subject to approval by the appropriate regulatory agency.

T-carrier—See Carrier.

T-1—The copper network and hardware used to carry DS1 service.

VSAT—Very Small Aperture Terminal. Refers to Ku-band satellite earth stations for fixed or portable use with dish diameters on the order of 1.5 meters or less.

V.35—An older (but still common in the U.S.) CCITT telco standard for low-speed data I/O to a CSU, with a unique multipin connector.

7.1 Broadcast Audio Measurement

Stanley Salek
Hammett & Edison, Inc., San Francisco, California
John Bisset
Multiphase Consulting, Falls Church, Virginia

AUDIO MEASUREMENT

The ability to quantify audio signals in terms of characteristics and qualities is paramount in audio engineering. This allows comparisons to be made with reference to established standards and requirements, including measurements relating to amplitude, frequency content, distortion, noise, and phase. The observation of such attributes allows a virtually complete characterization of an analog electrical audio system.

In the broadcast environment, audio measurements are used to gauge the overall quality of equipment such as amplifiers, tape recording systems, transcription equipment, mixing consoles, and other networks including the overall broadcast signal path. These types of measurements are detailed below. Other common audio measurements are of an acoustic nature and relate to the characteristics of microphones and speakers. These are not covered in this section.

Amplitude Measurement

The most basic of needs in audio measurement is to determine a value relating to the size, or amplitude, of an audio signal. Since an audio waveform is rapidly changing, methods have been developed to convert peak, root mean square (RMS), and average values of the changing waveform into corresponding proportional dc voltages that can be more easily observed.

There are specific cases where the peak value is the most direct measure of magnitude. It gives an indication of the largest excursions (either positive or negative) of an audio waveform. As shown in Fig. 1, the signal is applied to an absolute value circuit, which rectifies the waveform such that the output is all positive. A diode is then used to couple into C and R. These serve as memory and decay time elements, respectively, that can be adjusted in value to conform to the ballistics desired. Although the output is still changing with time, along with the input, the excursions corresponding to the peak values of the original waveform are much slower and more easily observed on metering devices. As the value of resistor R is increased, the decay time of the output is proportionally increased as well. If the resistor is completely removed, a peak hold circuit results.

Peak (actually peak-to-peak) functions can also be observed on an oscilloscope, although this technique is often impractical because of the difficulty in reading the random waveforms typical in most audio material. Storage oscilloscopes can perform a peak-to-peak hold function.

While there are many cases where the peak value is of considerable use, the RMS value of a signal is generally most meaningful since it gives indication of the energy content of the signal without regard to its waveform. In audio measurement, however, it is usually simpler to detect the average amplitude of given waveform and relate it to an associated RMS value, with reference to a sine wave. The RMS, or "root mean square" level can be defined as follows:

$$E_{rms} = \sqrt{\frac{E_1(t)^2 + E_2(t)^2 + .. + E_n(t)^2}{n}} \tag{1}$$

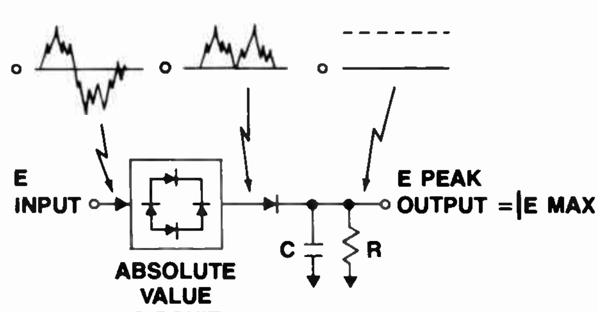


Figure 1. Peak value detection.

where E_1 through E_n are successive measurements over a total of n samples. As can be seen from its name, the value is computed by taking the average of n samples of E squared. Performing the square root function completes the calculation. This is also commonly referred to as "true RMS"[1] Fig. 2 shows how this is accomplished electrically.

Through the use of the absolute value circuit of Fig. 1 and the R/C configuration found in the RMS detector of Fig. 2, an average detector can be made. This is shown electrically in Fig. 3 and mathematically below:

$$E_{average} = \frac{E_1(t) + E_2(t) + .. + E_n(t)}{n} \qquad (2)$$

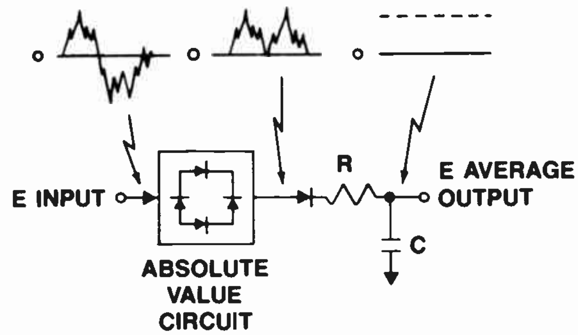


Figure 3. Average value detection.

In terms of audio perception, the average value of an audio signal is related to program material density, where the peak value described earlier relates to a maximum. Since the peak value defines the upper limit of allowable modulation in a transmission system, it is often technically desirable that the peak-to-average ratio be as low as possible to attain highest perceived loudness and signal-to-noise ratio. This may require compromising aesthetic goals and may not always be appropriate.

The decibel, or dB, is a unit for comparing relative levels of voltage or power signals in transmission systems. In broadcast audio systems, the most common representation of decibels is dBm. This is the value of a signal with reference to 1 mW into a 600 ohm load. The level in dBm of a signal can be found using the following relation:

$$dBm = 20 \log \left(\frac{E}{0.775}\right) \qquad (3)$$

where E is the voltage level to convert. The number 0.775 represents the voltage level reference of 0 dBm. Note that, strictly speaking, this formula is only true when the circuit impedance is 600 ohms. In practice, the formula is used typically without regard to the impedance level. Voltage levels obtained from the peak, RMS, and average circuits described above can be used for possible values of E. When this is done, some common types of metering can be synthesized to observe the activity of audio material.

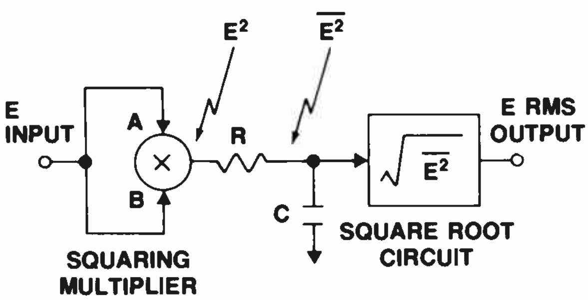


Figure 2. True RMS detection.

Metering

Two popular types of metering for the characterization of program audio are the standard volume unit indicator, commonly known as the volume unit (VU) meter, and the peak program meter (PPM). Although VU metering has been the more common in U.S. broadcast equipment, the standard PPM indicator is gaining popularity.

The VU meter (see Figs. 4A, 4B) was introduced in 1939 to serve as a standard program level indicating device.[2] Its original purpose was to be the reference between broadcasters (as well as other programming suppliers) and the telephone company. A VU meter is the combination of a bridge rectifier, a resistive attenuator, and an ammeter with an approximately voltage linear scale to produce an average responding ac voltmeter. The VU meter is calibrated such that it reads 0 VU across a circuit in which a sinusoid develops 1 mW in a resistance equal to the circuit impedance (i.e., 0.775 volts RMS across 600 ohms). This allows the meter to be powered directly by a 600 ohm program line, with the attenuator typically set to read 0 vu at +8 dBm.

Beyond reading continuous tones, the dynamic characteristics are set so that it will display 99% of its steady-state reading on a sine wave tone burst 300 milliseconds long, with a fall to 5% of the reading in 300 milliseconds. Essentially an average responding device, the VU meter will not respond to short duration program peaks. Therefore, levels normally should be set with a 10 dB margin (headroom) before the point of clipping.[3]

The peak program meter (PPM) is designed to read nearly the full peak value of the audio signal (See Figs. 4C, 4D). It uses a rectifier and an integrator, producing a fast rise and slow fall effect on the display device. Typical standards require the PPM to read −2 dB, ±0.5 dB of the steady state value for a tone burst of 10 mS, and take 2.8 seconds for the pointer to fall 20 dB.[4]

Typically, a PPM exhibits flat frequency response and is calibrated such that the nominal peak program level corresponds to a 0 dB meter reading near full scale (generally +16 dBm).

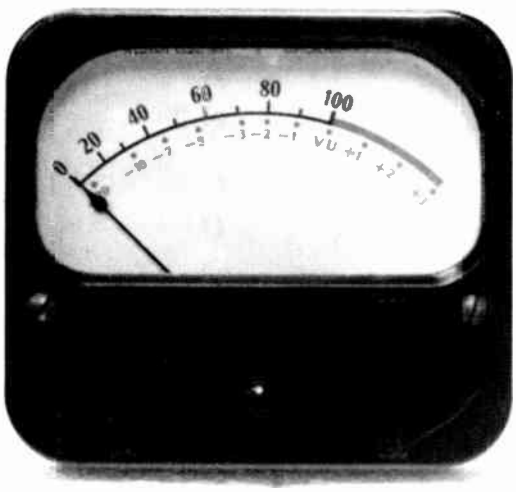


Figure 4A. VU Meter.

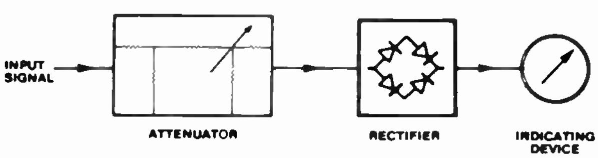


Figure 4B. Block diagram of the stages of a typical volume unit indicator.

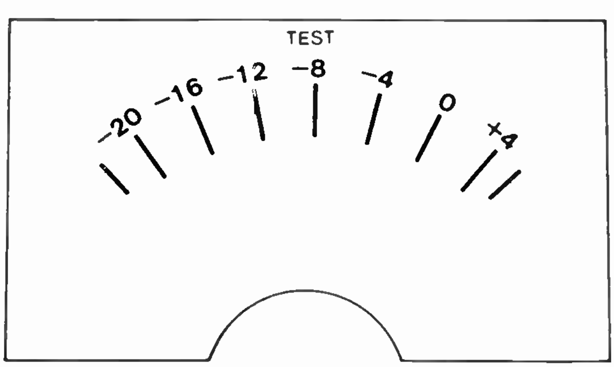


Figure 4C. Arrangement of a typical peak-program meter scale.

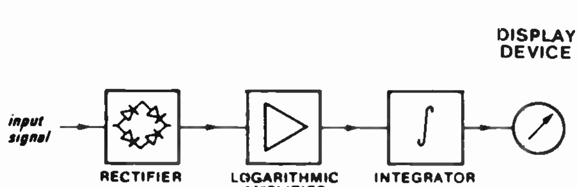


Figure 4D. Block diagram of the stages of a typical PPM.

Other types of metering devices have been developed to read wideband audio in a simultaneous mode. One such system consists of an LED bar graph display exhibiting peak program content. "Riding" upon this is a brighter display (utilizing the same display elements) corresponding to VU standards. Such an indicator allows continuous monitoring of program material compression and dynamic characteristics. Another system combines a VU movement meter with peak indicating LED flashers. A "hold" function is sometimes associated with these to allow an operator sufficient time to observe the approximate peak content.

FREQUENCY ANALYSIS

Amplitude analysis methods, as described in the previous section, are generally used to provide an indication of signal levels simultaneously over the entire audio range. It is sometimes more desirable, however, to be able to measure discrete frequencies in an audio system. This allows frequency response measurement as well as dynamic measurement of energy content throughout the audio spectrum.

Simply stated, frequency response is the capability of a device or system to pass or amplify equally, all frequencies within a specified range.[5] As far as audio in the broadcast environment is concerned, the range of interest is generally 50 Hz to 15 kHz or 20 kHz. Although few musical instruments produce fundamental frequencies greater than 4 kHz and the human voice much above 1 kHz, the reproducing device or system must be able to pass the harmonics that accompany the fundamental frequencies. Without adequate bandwidth or with uneven frequency response, an unnatural coloration of the perceived sound becomes evident. To solve this, a great deal of care is taken to construct amplifiers with very flat frequency response to high frequencies. Since the responses of series-connected amplifiers are additive, care must be taken to verify the flatness of each in a system.

Several methods are available to measure audio frequency response. They include discrete measurement and swept frequency methods. Parallel analysis and Fast Fourier transform (FFT) techniques can also be used.

The discrete frequency measurement method is uncomplicated and inexpensive. A simple measurement system consists of a low distortion audio frequency oscillator and a wideband ac voltmeter. The oscillator output is connected to the input of the device or system to be characterized. The voltmeter is used to observe the level at the output of the device, or at a desired intermediate point in a system.

The measurement is done by first setting the generator output level to a volume that represents the nominal input operating level of the device. Generally, a 400 Hz or 1000 Hz frequency is chosen initially in high fidelity audio systems. The output level is read on the ac voltmeter, and this quantity is noted as a zero dB relative reference. Provided the generator itself has a flat frequency response, measurements at frequencies through the audio band can be taken while recording the corresponding dB output levels with respect to the reference. A convenient and commonly used technique is to increment the frequency in a 1, 2, 5 sequence (i.e., 20 Hz, 50 Hz, 100 Hz, 200 Hz, etc.). This permits plotting the final response data on 4-cycle "LOG/LIN" graph paper, providing regularly spaced frequency increments horizontally. The logarithmic amplitude data are plotted along the horizontal axis, with the zero dB relative reference placed in a convenient position on the linear vertical axis.

Although the discrete frequency measurement technique is straightforward, it is also often tedious and time-consuming. Numerous frequency measurements must be made to insure adequate testing. This method is most usable in the response measurement of single ended devices that do not have a suitable input port for connection to an audio generator. Transcription equipment, such as turntables and compact disc players are examples of these types of devices. Test recordings supplied by the equipment manufacturer and other sources are used to provide the tones necessary for discrete frequency response characterization.

Swept Frequency

A faster and more efficient means of measuring frequency response is the swept frequency method. This process employs a sweep frequency generator as a signal source and measures response over the entire range of interest in one sweep. The detector for these measurements is most often a tracking type that follows the signal source and measures a narrow band of frequencies centered around the source frequency. Use of a tracking detector is a better guarantee that the amplitude measured is that of the tone generator and is not influenced by spurious tones, noise, or harmonics.

Devices specifically designed to conduct swept frequency measurements include wave analyzers and spectrum analyzers. Wave analyzers must be used with plotters to provide a hard copy data plot, while spectrum analyzers directly produce response images on a built-in display. A representative spectrum analyzer is shown in Fig. 5 while the one displayed in Fig. 6 is specifically intended for audio frequency use.

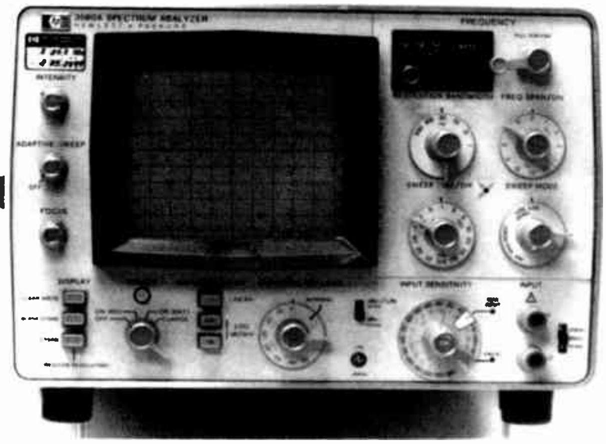


Figure 6. Spectrum analyzer intended for audio use.

Fig. 7 shows a typical setup for measuring amplifier frequency response using the swept frequency method. The signal source used to drive the test device is the tracking oscillator output of the analyzer. The device's output is terminated with an appropriate characteristic load impedance and connected to the analyzer input. Measurement of the frequency response is made by manually or automatically sweeping the analyzer across the frequency range of interest. A plotter connected to either type of analyzer or a scope-type camera used with the spectrum analyzer can provide a permanent record of the test device response characteristics.

The swept frequency method also can be used to measure the overall frequency response of multiple-head tape recorders. The tracking oscillator signal is connected to the record amplifier input and the output

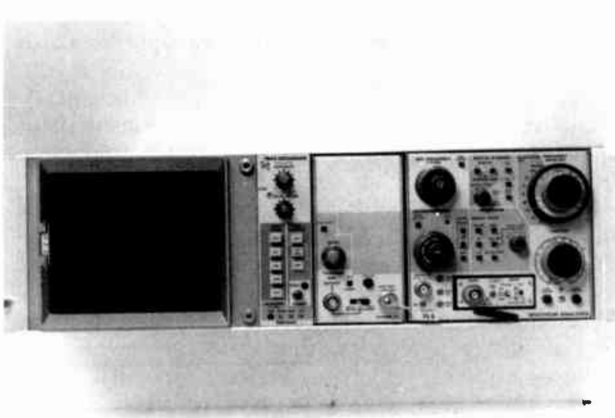


Figure 5. Spectrum analyzer suitable for audio and low frequency RF use.

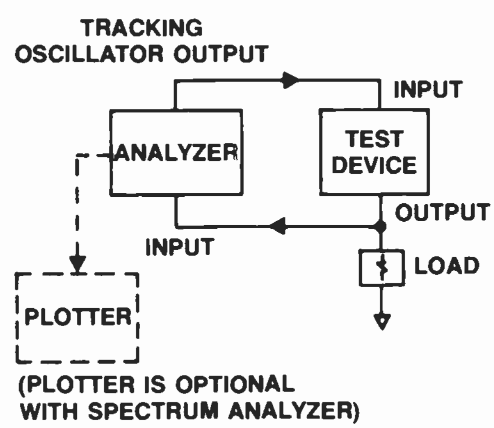


Figure 7. Frequency response measurement using swept method.

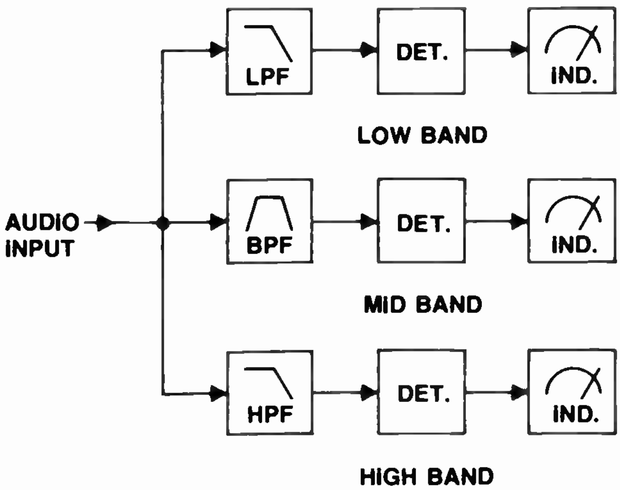


Figure 8. Basic real-time analyzer.

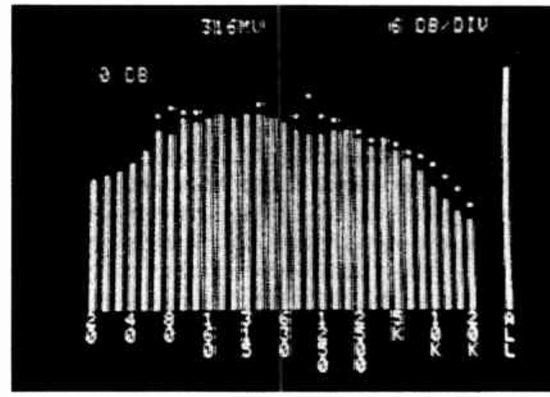


Figure 10. Typical RTA display.

is taken from the playback amplifier output. This allows recording and playback frequency response measurement to be done all at once and eliminates the necessity of synchronizing the analyzer sweep to the recorded signal. However, it should be noted that a time delay between input and output signals exists, caused by the physical displacement between record and playback heads. This often can be corrected simply by using a very slow sweep speed or a wider measurement bandwidth.

Real Time Analyzer

A real time audio analyzer (RTA) consists of a sequential collection of one octave or one-third octave filters having individual detectors and indicators at each output. The program audio is simultaneously fed to the inputs of all the filters. The output signal of each filter is proportional to the amount of energy occurring in that particular frequency band. This technique is also referred to as *parallel analysis*.

A simplified version of an RTA is presented in Fig. 8. As shown, it is intended to break the audio band into three sections using lowpass, bandpass, and highpass filtering. Signal detectors then are used to condition the audio for display on a suitable indicator, one set for each of the three bands. The detectors can be (and often are) the same peak, RMS or average circuits described earlier. Typical readout indicators are bar graph displays with dB-calibrated scales. When arranged side by side, the readouts provide a graphical presentation of amplitude versus frequency, just as the spectrum analyzer does. Unlike the spectrum analyzer, however, an RTA does not rely on a fixed sweep speed. An advanced audio analysis system with RTA function is shown in Fig. 9. Fig. 10 shows a typical RTA display produced by the same unit.

Figure 9. Advanced audio analyzer with RTA function and CRT display.

Parallel techniques using the RTA are often used for dynamic program material and room acoustics analyses.[6] This type of analyzer is also useful for measuring frequency response of audio devices when used in conjunction with a "pink" noise source. Pink noise has a constant mean squared voltage per octave of frequency. This makes it popular in audio work since it allows correlation between successive octaves by ensuring the same voltage amplitude is available as a reference. By connecting the pink noise source to the input of a device to be characterized, and the RTA to its output, a response curve can be displayed almost instantaneously.

Another device intended for real time program analysis is the *Audio Program Analyzer*, pictured in Fig. 11. It allows material from devices such as receivers, modulation monitors and audio processing equipment to be characterized in several ways. Peak-to-average ratios can be quickly determined, as well as maximum peak levels. Peak density can be measured; this is a parameter that can be related to audio processor effectiveness. The unit also contains a four band RTA, and provisions are made for monitoring stereo program material.

Although often prohibitive in cost, network and FFT analyzers are also exceptionally useful in audio frequency domain measurements. Network analyzers are swept analysis instruments, and are used to characterize two-port networks (i.e., devices having an input and output) as to frequency, phase, and delay responses. They are employed where substantial accuracy in the measurement of these parameters is required. RF subsystem and semiconductor device design have been the major application for network

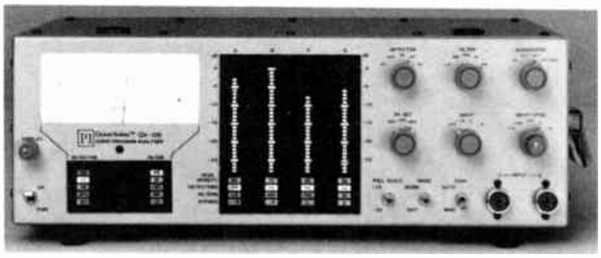


Figure 11. The QuantAural™ Audio Program Analyzer from Potomac Instruments permits the characterization of the effects of audio processors.

analyzers, although some newer generation equipment includes audio-frequency coverage. In the case of audio systems, network analyzers allow precision response measurement of amplifier and filter designs.

FFT (Fast Fourier transform) analyzers have the ability to convert a snapshot sampling of an audio or other time-varying source and mathematically transform the result into a display of the frequency components present. Because the conversion process is done by a specialized digital signal processing (DSP) microcomputer, an FFT analyzer often can produce a complete spectrum display as much as an order of magnitude faster than conventional swept spectrum analyzers. This is most helpful in low frequency measurement, where a swept analyzer would require a very slow sweep time to resolve closely spaced components. As their purchase costs have been decreasing, FFT analyzers are becoming increasingly popular for audio system analysis and measurement.

DISTORTION MEASUREMENT

When a two-port device is driven beyond its range of linear operation, or through areas of discontinuity, signal distortion occurs. As a result, additional frequencies appear at the output that were not present at its input. In cases where distortion becomes extreme, it can be identified through listening. Odd-order distortion (such as clipping distortion) can become audible at around 1.25%. Even-order distortion, characterized by a coloration of the program material, becomes audible at about 5%. Generally, systems with a wider frequency response capability need to maintain lower distortion levels to be acceptable. Since distortion is not always obvious to many people, techniques are available to measure its various types.

Distortion can be characterized in two basic ways: harmonic distortion and intermodulation distortion. While the two associated methods produce uncorrelated measurement values, each gives a quantitative result of the device's quality in terms of a single number. Although total harmonic distortion (THD) content is determined by only one method, intermodulation distortion (IMD) has several accepted measurement practices, including the SMPTE and CCIF methods. Transient intermodulation (TIM) distortion is another commonly measured type of IMD.

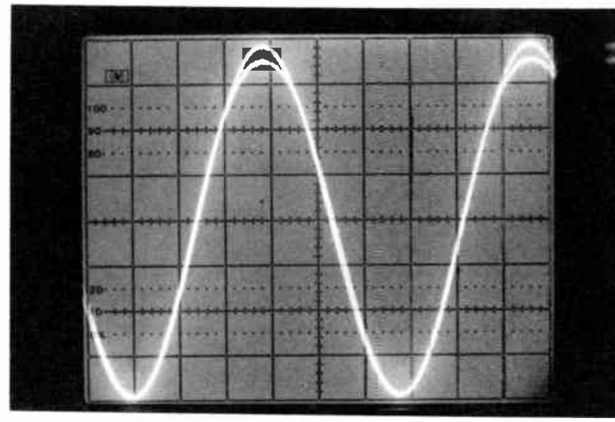


Figure 12. Comparing the distorted output with an undistorted 1 kHz tone with its undistorted component.

Harmonic Distortion

Harmonic distortion is a measure of individual harmonic amplitudes with respect to the amplitude of the fundamental frequency. In practice, harmonics greater than third order often add little to the resultant value because of their negligible amplitude. THD is defined as:

$$THD\% = 100\frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + .. + A_n^2}}{A_1} \qquad (4)$$

where A_2 through A_n are the amplitudes of the individual harmonics and A_1 is the amplitude of the fundamental.

As seen in Fig. 12, a 1 kHz sine wave with harmonic distortion shows only minor differences when it is overlaid with an undistorted signal (as viewed on an oscilloscope). The amplitude and slope errors do not lead directly to a numeric result. But when a spectrum photo of the same waveform is observed (Fig. 13), the

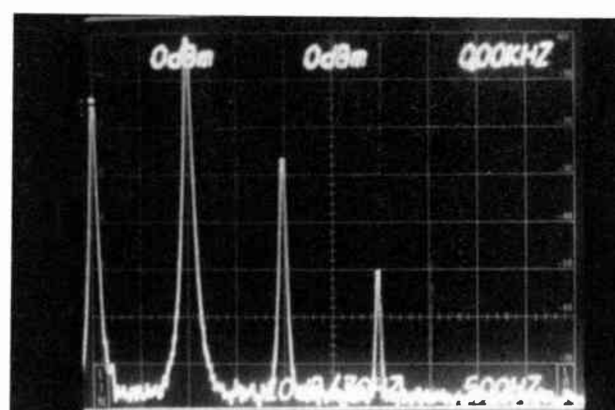


Figure 13. Measuring THD with a spectrum analyzer. (V:10 dB/div.; H:500 Hz/div.)

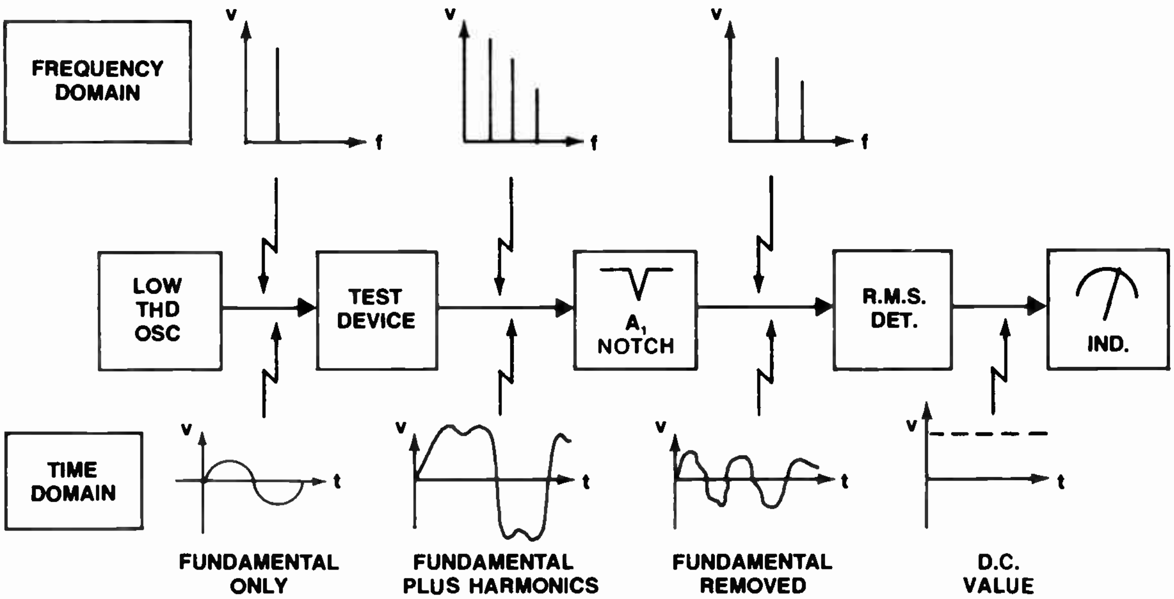


Figure 14. THD analyzer.

above relation can be applied. With the fundamental at 0 dBm (0.775V), the second harmonic at −26 dBm (38.8 mV), and the third harmonic at −50 dBm (2.5 mV), the harmonic distortion can be calculated:

$$THD\% = 100\frac{\sqrt{(0.0388)^2 + (0.0025)^2}}{0.775} = 5.0\% \quad (5)$$

Although spectrum analysis can produce accurate THD measurement results, a simpler and more cost effective procedure that produces a direct numeric quantity is more popular.

Fig. 14 shows the block diagram of a typical THD analyzer. An oscillator (with much less harmonic distortion than the device or system to be measured) is connected to the test device input. The distorted output signal of the device is filtered to remove A_l, the fundamental component. This produces a signal that, when RMS detected, is proportional to the THD imposed by the device being tested.

THD measurement is often conducted using the same 1,2,5 sequence of frequencies mentioned for discrete response measurement. The THD results can be plotted on the same graph to characterize the device under test on a single page. THD measurements may be taken over various input levels, but as the level is reduced, noise characteristics may affect the readings. In such cases, the spectrum analyzer method could produce more meaningful results.

Intermodulation Distortion (IMD)

The intermodulation method of measuring distortion uses a test signal composed of two sinusoidal signals of different frequencies (except for TIM measurement, to be covered shortly). After summation, they produce the effect of an amplitude modulated carrier when applied to a circuit having IMD. The intermodulation method is useful because the harmonic distortion of the signal sources do not affect the measurement.

The SMPTE (Society of Motion Picture and Television Engineers) method uses a low frequency (f_l) and a relatively high frequency (f_2) signal (usually 60 Hz and 7 kHz, respectively) that are mixed at a four to one amplitude ratio (see Figs. 15 and 16). This method involves the measurement of the relative amplitude of the modulation sidebands added to the higher frequency signal. For diagnostic purposes, it is often useful to determine even-order and odd-order distortions separately, although this is best done by spectrum

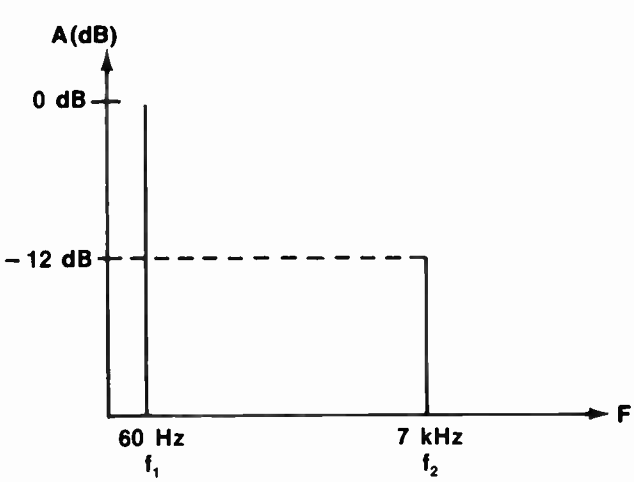


Figure 15. Spectrum of SMPTE IM input test signal ratios.

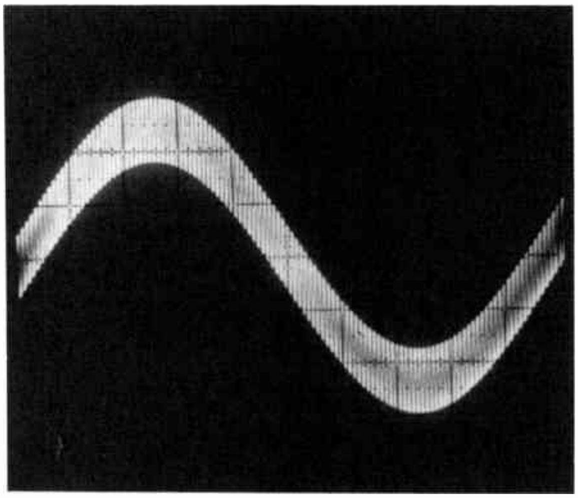


Figure 16. SMPTE IM test signal.

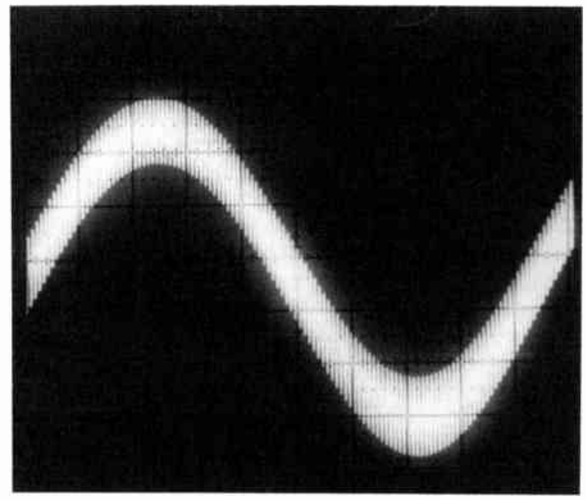


Figure 17. SMPTE IMD measurement—time domain output signal.

measurement techniques. Even-order distortion usually can be characterized by the ratio of the sum of the amplitudes of only the two second-order spurious frequencies, f_2-f_1 and f_1+f_2, to the amplitude of the carrier signal, f_2:

$$\underset{(second\ order)}{SMPTE\ IMD\%} = \left[\frac{A_{(f_2-f_1)} + A_{(f_1+f_2)}}{A_{f_2}}\right] \times 100 \quad (6)$$

In a similar manner, odd-order distortion can be characterized by the ratio of the sum of the amplitudes of the two third-order spurious frequencies, f_2-2f_1 and $2f_1+f_2$, to the amplitude of f_2:

$$\underset{(third\ order)}{SMPTE\ IMD\%} = \left[\frac{A_{(f_2-2f_1)} + A_{(2f_1+f_2)}}{A_{f_2}}\right] \times 100 \quad (7)$$

Fig. 17 shows the output signal of an amplifier with IMD as viewed on an oscilloscope. Note the elongated trough as compared to Fig. 16. As with THD, spectrum analysis can be used to determine the numerical amount of distortion present. Intermodulation sidebands can be seen around f_2 in the spectrum photo of Fig. 18. Second and third order distortion percentages for this example are calculated as follows:

$$A_{f_2} = -12dBm = 195\ mV$$

$$A_{(f_2-f_1)} = A_{(f_1+f_2)} = -38\ dBm = 9.76\ mV$$

$$\underset{(second\ order)}{SMPTE\ IMD\%} = \left[\frac{9.76 + 9.76}{195}\right] \times 100 = 10.0\%$$

$$A_{(f_2-2f_1)} = A_{(2f_1+f_2)} = -58\ dBm = 0.98\ mV$$

$$\underset{(third\ order)}{SMPTE\ IMD\%} = \left[\frac{0.98 + 0.98}{195}\right] \times 100 = 1.0\% \quad (8)$$

As is shown, the contribution of even-order distortion products is usually greater than that of the odd-order. To express the result as a single quantity, the vector sum of the two quantities is taken:

$$\underset{(total)}{SMPTE\ IMD\%} = \sqrt{(IMD\%\ Even)^2 + (IMD\%\ Odd)^2} \quad (9)$$

$$= \sqrt{10^2 + 1^2} = 10.05\%$$

As with THD, SMPTE IMD has a direct method of numeric solution, as shown in the block diagram of Fig. 19. The two test frequency oscillators are summed to produce the f_1+f_2 signal, which is then applied to input of the device to be tested. The distorted output signal is high-pass filtered to remove the f_1 fundamental

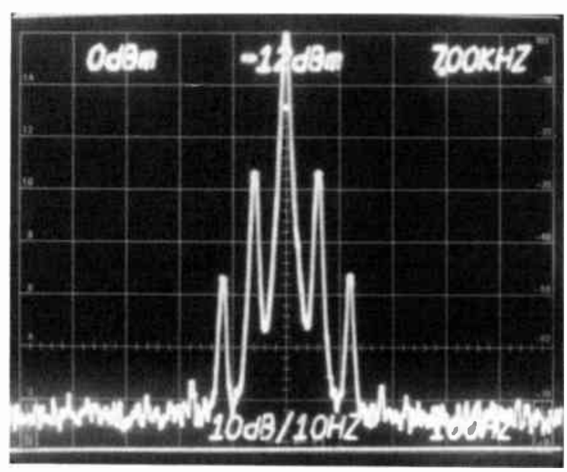


Figure 18. SMPTE IM distortion measurement. (V:10 dB/div.; H:100 Hz/div.)

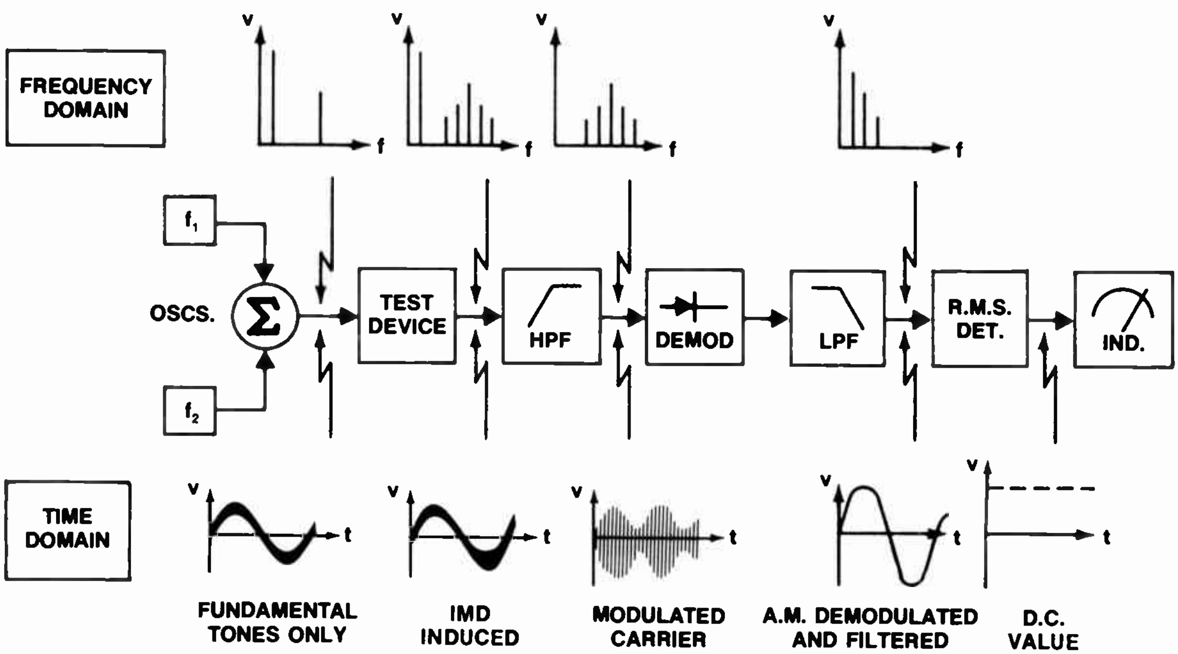


Figure 19. SMPTE IMD analyzer.

component, leaving only the amplitude modulated f_2 component. Using a standard AM demodulator and low-pass filter, the residual f_1 component is obtained. After RMS detection, a dc level proportional to the distortion is produced that can be viewed on a direct-reading indicator.

"Wow and flutter" is a term that describes a special case of IMD normally associated with analog tape recorders. It is caused by variations in tape velocity across the recording and/or reproducing heads, due to imperfections in the mechanical drive system. These variations result in frequency modulation of the recorded and reproduced signal. The frequency spectrum obtained is similar to that of the SMPTE IMD measurement method, except the f_1 low frequency signal is generated by fluctuations in tape speed and is not of any set amplitude.[5]

To measure wow and flutter, a test tape containing a prerecorded 3 kHz (or 3.15 kHz) tone is played. Using an audio spectrum analyzer, for example, the amplitude of the first sideband (A_m) with reference to the 3 kHz amplitude (A_o) is measured. The frequency of the flutter (F_m) also must be known. Then the following relation can be used to approximate the percentage of wow and flutter present:

$$WOW\ \&\ FLUTTER\,\% = \frac{2(A_m)(F_m)}{(A_o)(3\ kHz)} \times 100 \quad (10)$$

(peak-to-peak)

For example, if the first sideband (either upper or lower) amplitude is 10dB below the 3kHz maximum amplitude (0.316 volts with reference to an arbitrary 1 volt amplitude at 3 kHz) and is at a frequency of 4Hz, the wow and flutter would be 0.084%.

The CCIF intermodulation method uses a combination of two higher frequency sinusoidal signals (f_3, f_4) of equal amplitude. They are typically 1 kHz apart and found at 5/6 kHz, 14/15 kHz, or 19/20 kHz in many applications. One of the spurious frequencies generated is low in frequency while others are gathered around the two driving frequencies. Figs. 20 and 21 spectrally show the driving frequencies before and after passing through a test amplifier. As with SMPTE IMD measurement, the generated spurious products can be classified as even-order or odd-order. Even-order distortion is expressed as the ratio of the amplitude of the difference component ($f_4 - f_3$) to the sum of the two driving frequencies (f_3, f_4):

$$CCIF\ IMD\% = \left[\frac{A_{(f_4 - f_3)}}{A_{f_3} + A_{f_4}}\right] \times 100 \quad (11)$$

(second order)

Odd-order distortion is determined by calculating the ratio of the sum of the amplitude of the two third-order products, $2f_3 - f_4$ and $2f_4 - f_3$, to the sum of the amplitudes of the two driving frequencies, f_3 and f_4:

$$CCIF\ IMD\% = \left[\frac{A_{(2f_3 - f_4)} + A_{(2f_4 - f_3)}}{A_{f_3} + A_{f_4}}\right] \times 100 \quad (12)$$

(third order)

In the case of Fig. 21, the driving frequencies f_3 and f_4 are at 5 kHz and 6 kHz, even-order product at 1

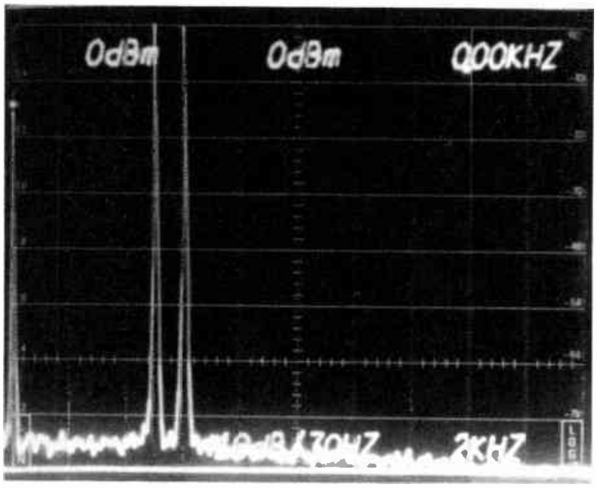


Figure 20. CCIF IM test signals. (V:10 DB/DIV.; H:2 KHZ/div.)

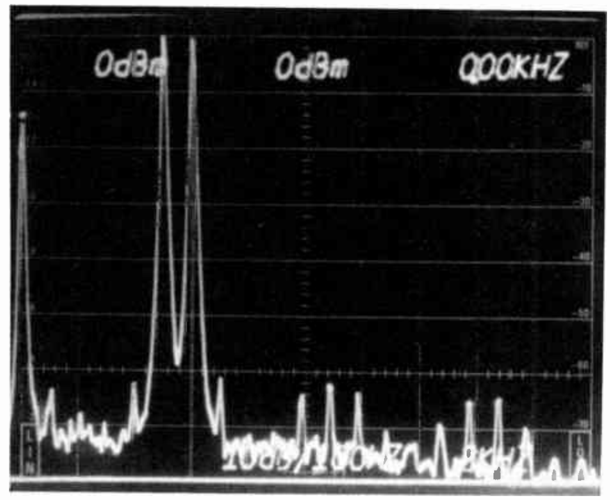


Figure 21. CCIF IM distortion measurement. (V:10 dB/div.; H:2 kHz/div.)

kHz, and third-order products at 4 kHz and 7 kHz, respectively. Distortion percentages for this example are calculated below:

$$A_{f_3} = A_{f_4} = 0\,dBm = 775\,mV$$
$$A_{(f_4 - f_3)} = -64\,dBm = 0.49\,mV$$
$$CCIF\,IMD\% = \left[\frac{0.49}{775 + 775}\right] \times 100 = 0.032\%$$
(*second order*)

$$A_{(2f_3 - f_4)} = -62\,dBm = 0.62\,mV$$
$$A_{(2f_4 - f_3)} = -61\,dBm = 0.69\,mV$$
$$CCIF\,IMD\% = \left[\frac{0.62 + 0.69}{775 + 775}\right] \times 100 = 0.085\%$$
(*third order*) (13)

It is a common practice for direct-reading metered analyzers to measure only the amplitude of the difference product ($f_4 - f_3$) with respect to the driving signal

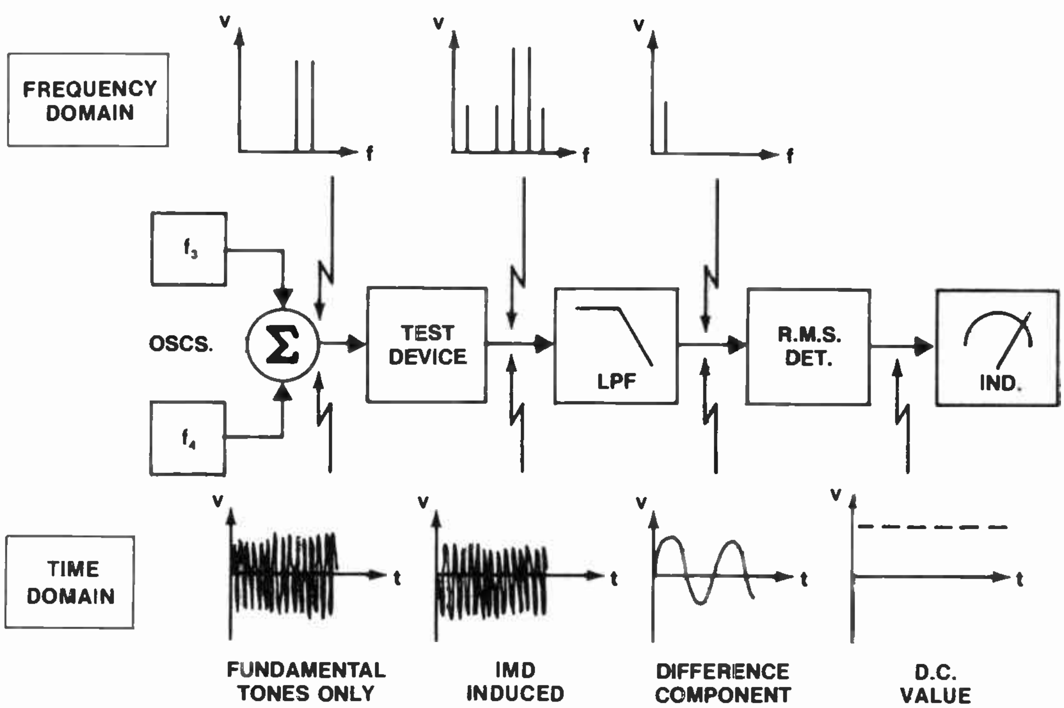


Figure 22. CCIF second-order IMD analyzer.

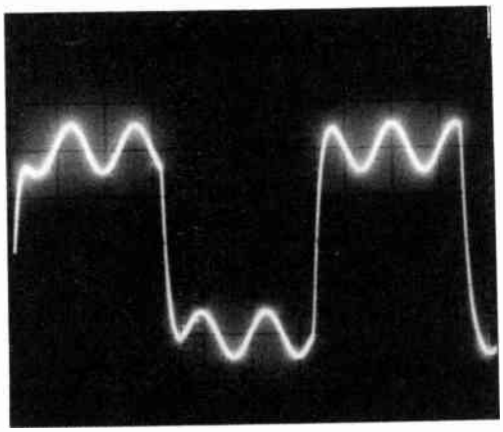


Figure 23. TIM test signal.

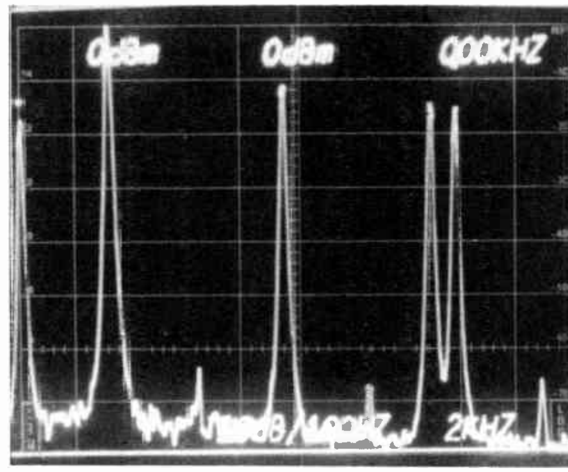


Figure 24. TIM test signal, spectral view. (V:10 dB/div.; H:2 kHz/div.)

amplitudes. A device that performs this task is called a CCIF second-order difference frequency distortion analyzer. Fig. 22 illustrates how the measurement is made.

Transient intermodulation (TIM) distortion is found only in amplifiers that utilize negative feedback. When this feedback is excessive, a fast-rising transient signal applied to the input of the amplifier can produce an internal overshoot that saturates the circuits in the amplifier.

The most popular procedure used to measure TIM distortion is called the sine-square wave method. The test signal employed uses a square wave (f_{sq}) to induce nonlinearity in the test device by saturating the amplifier's internal current, caused by its alternate rises and falls. Mixed with this square wave is a low level, high frequency sine wave (f_{si}), which is unrelated harmonically. As defined, the frequency of the square wave is 3.18 kHz and that of the sine wave 15 kHz, where the peak-to-peak amplitude ratio of the former to the latter is four to one.[7] Before summation, the square wave is low-pass filtered using a first order design having a cutoff frequency of 30 kHz. This reduces the harmonics outside of the band of interest that could damage the device being tested. The composite waveform produced is shown in Fig. 23 and spectrally in Fig. 24.

Using the test setup shown in Fig. 25, an amplifier can be measured for TIM. Mathematically, TIM distortion produced by the sine-square wave method is defined as:

$$TIM\% = \frac{\sqrt{A1^2 + A2^2 + \ldots + A9^2}}{A_{si}} \times 100 \qquad (14)$$

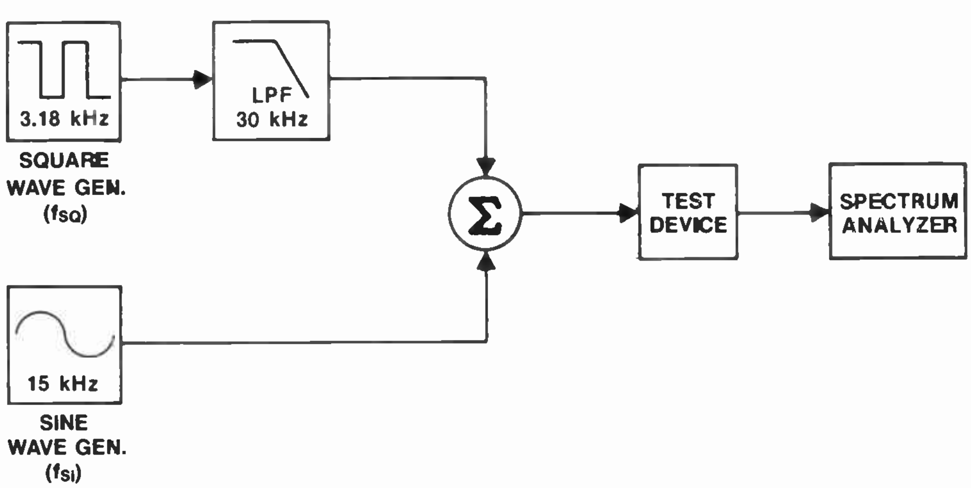


Figure 25. TIM distortion analysis.

where $A1$ through $A9$ represent the amplitudes of the distortion product present in the audio band and A_{si} is the amplitude of the sine wave. Values for the $A1$-$A9$ components are shown in the table below:

TABLE 1

Component	Relation	Frequency, kHz
A1	fsi— fsq	11.82
A2	fsi—2fsq	8.64
A3	fsi—3fsq	5.46
A4	fsi—4fsq	2.28
A5	|fsi—5fsq|	0.90
A6	|fsi—6fsq|	4.08
A7	|fsi—7fsq|	7.26
A8	|fsi—8fsq|	10.44
A9	|fsi—9fsq|	13.62

Fig. 26 shows the TIM distortion products produced by an amplifier that only displayed negligible THD, SMPTE IMD and CCIF IMD percentages. Using Formula 14, it can be determined that this amplifier is producing about 20% TIM distortion.

Other TIM measurement methods include a saw-tooth wave method that takes amplifier slew rate into account and a noise-square wave method, where the sine wave of the sine-square wave method is replaced by a narrow-band noise spectrum.[8] At the time of this writing, these have not been as commonly used as the sine-square wave method.

Filtering as part of distortion measurement is often useful to remove components that are of little interest and as a diagnostic aid.[9] This is especially true with THD measurement. A 20 kHz or 30 kHz high-pass filter placed in series with the output of the device being measured is useful for testing broadcast equipment. This is often an acceptable practice, since the harmonics produced outside the transmission bandwidth can be eliminated, producing a more realistic result.

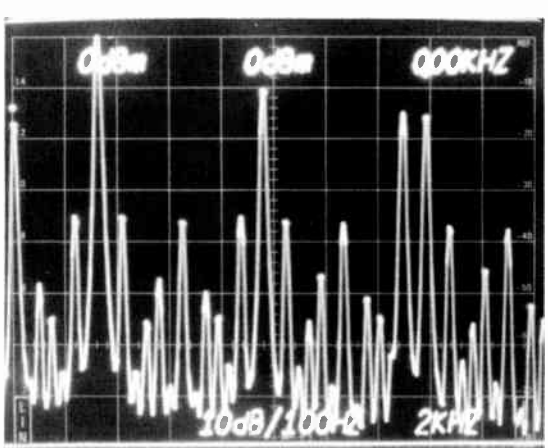


Figure 26. TIM distortion products. (V:10 dB/div.; H:2 kHz/div.)

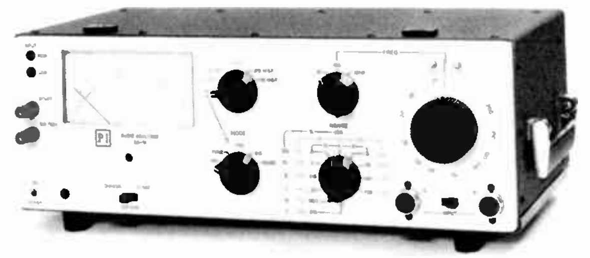


Figure 27. In addition to measuring frequency response and distortion, this audio analyzer measures phase as well as wow and flutter.

A high-pass filter also can serve as an important analytical aid. With a cutoff frequency in the 400 Hz range, it can be placed in series with the test device's output and used to determine the contribution of line frequency hum (60 Hz in the U.S.) to a THD measurement utilizing a fundamental frequency of 1 kHz or greater. Verification of adequate grounding used in the test setup can also be observed.

Care must be taken, however, when filtering is used in IM measurement. It is important to verify that in-band distortion components are not inadvertently removed. Also, extremely sharp cut-off filter designs may produced overshoot components that could affect measurement results.

A device capable of performing many of the distortion tests described above, as well as frequency response measurements, is shown in Fig. 27.

NOISE MEASUREMENT

In audio engineering, noise is a random energy distribution in which individual spectral components are not clearly resolved. Primary sources of noise in amplifiers are in the resistive circuit elements.[10] It is important to control noise in amplifiers as their gain increases to preserve a high signal to noise ratio, which is the ratio of the operating signal level to the noise level inherent in the amplifier itself.

To understand the origin of noise, we can model a passive resistive element as a noiseless resistor in series with a noise voltage generator, E_r:

$$E_r = \sqrt{4KTBR}\ (volts) \qquad (15)$$

where: K = Boltzman's constant (1.38×10^{-23} W-Sec/°K)
T = Temperature in degrees Kelvin
B = Noise bandwidth (Hz)
R = Resistance in ohms

Although noise bandwidth is not equivalent to an amplifier's 3 dB bandwidth, it can be related.

As can be seen from the equation, noise voltage is a physical phenomenon that can be worsened by an increase in any of the variable factors. Therefore, noise cannot be eliminated but it can be reduced. This is often done by proper selection of the resistive components used, because of an additional factor known as excess noise, which is proportional to the voltage drop across the resistor and related to the material from which it is made. Of the different available types, carbon composition resistors are prone to the most excess noise contribution while metal-film devices show the least.

At times, the actual spectral distribution of noise is of less importance than the noise voltage within a given bandwidth for comparison purposes. For audio frequencies, a 15 or 20 kHz bandwidth is of interest. With a low-pass filter in this range connected in series with an amplifier output, and the input of the amplifier grounded, an unweighted but band-limited noise measurement can be made. When the noise output level is obtained, it can be expressed as a ratio with a standard operating level and reference frequency. This produces an indication of the amplifier's signal-to-noise ratio (SNR).

When the gain of the amplifier is known, this same technique can be used to determine equivalent input noise voltage, i.e., the voltage of the noise that would be found at the input of the amplifier if the amplifier were completely noiseless.[10]

The measurement of a noise voltage quantity over a given frequency bandwidth in order to determine a signal-to-noise ratio does not provide a complete characterization. This is because the noise spectrum can occupy all or part of the same bandwidth.[11] For example, two amplifiers with identical SNRs can sound very different because one may have a uniform noise spectrum and the other may have most of the noise concentrated over a limited frequency range. Hence, the latter amplifier would sound "noisier" than the former. This has to do with the way the ear perceives the loudness of a signal that is uniform in amplitude across the audio band. To make comparative noise measurements more meaningful, several weighting filters have been used to alter noise spectra over the frequency band of interest.

"A" weighting is based on the inverse of early measurements by Fletcher and Munson of the ear's sensitivity at low sound pressure levels.[12] A more recently developed weighting curve utilizes the CCIR/ARM method, an updated scheme which places the zero dB reference at 2 kHz instead of 1 kHz.[13] It is believed that this method, which is based on the obtrusiveness as well as the levels of different kinds of noise, provides a more commercially acceptable result when used to characterize modern, wide-range audio equipment. Fig. 28 compares the two curves.

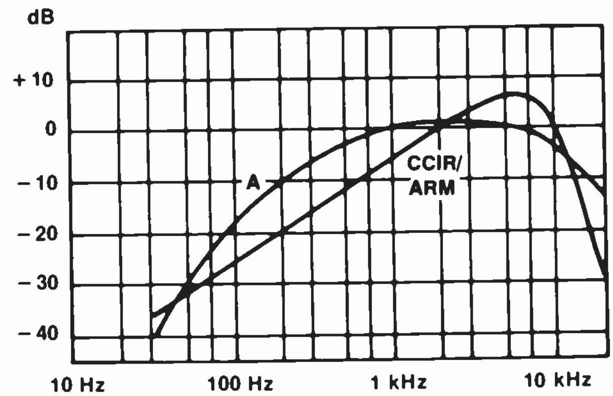


Figure 28. Noise measurement weighting schemes.

PHASE MONITORING AND MEASUREMENT

An L+R summation is the monophonic compatible signal for AM, FM, and TV stereo broadcasting. Separation information is transmitted via an L−R signal. Since these two signals are created through a summation and difference process of the original left and right channel stereophonic source, it is important that they be recombined properly at the receiver.[14] To accomplish this, amplitude and phase errors must be minimized in the transmission system. Phase measurement is important in accomplishing this task.

In a stereo program system, if left and right audio information is correlated but delayed in phase, the error would not be evident on a stereophonic receiver. A monophonic signal, however, will be degraded because of inexact summation. This problem is common to audio tape recording. Periodic azimuth adjustment or "phasing" of the heads is often done. A test tape containing a high frequency tone is played while azimuth is adjusted to minimize the difference signal between the two channels.

A phase meter can be used to simplify this task. A simple version would take phase and amplitude variances into account simultaneously by functioning as a two-input subtractor. When both characteristics are identical in each of the channels, the output becomes zero. A meter that measures only phase information compares the zero-crossing times of the two input signals, and the resulting time difference is used to generate a dc voltage proportional to the phase difference.[15] Phase detectors operating in this manner often limit the input signals in order to remove all amplitude information.

More popular, however, is the Lissajous figure method, involving the use of an oscilloscope in the X-Y mode. The patterns produced are shown in Fig. 29. An oscilloscope is connected such that the left channel audio causes an X-axis sweep and right channel audio produces a Y-axis sweep as shown in Figs. 29A and 29B respectively. When each channel contains the same program material, the pattern of Fig. 29C is produced. This is the L+R axis. If one of the channels is inverted, the pattern of Fig. 29D becomes evident. This is often called the L−R axis. Program material that follows this axis is said to be inverted in polarity because no sum or L+R information is present.

During alignment of tape reproducing equipment, when discrete tones are used, the patterns of Figs. 29E, 29F, and 29G are commonly seen when phase errors exist between the two channels. Stereo program

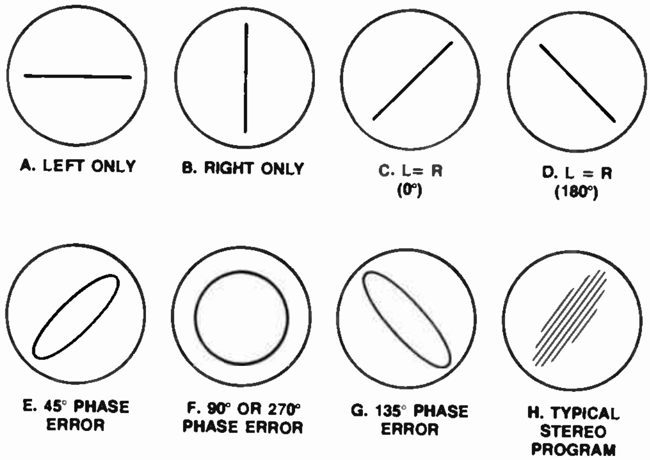


Figure 29. Interpretation of Lissajous patterns.

material, in unprocessed form, generally modulates the L+R axis while simultaneously deviating in the L−R direction to a lesser amount, as shown in Fig. 29H.

AUDIO TEST SIGNALS

A wide variety of audio generators are available to provide the necessary test signals for making the measurements discussed thus far. These signal generators range in complexity from the manually operated type as shown in Fig. 30, to computer-assisted and automated generators.

With the advent of the compact disc, precision test tones become available for all broadcasters.[16] Using the NAB Test CD and Test CD II, performance measurements can be made that formerly required equipment that was out of the reach of many stations. Both volumes contain tone pattern tests that can be utilized with a minimum of test equipment. The physical size of a CD versus a signal generator also lends itself to the wide variety of applications found in the broadcast setting. Except for the announcement tracks used to verify LEFT and RIGHT channel phasing and balance, the signals on the CDs are 100% digitally synthesized.[17]

Figure 30. A commonly found broadcast audio signal generator.

Figure 31. The NAB Test CD turns the compact disc player into a powerful signal generator.

Tracks 2 through 13 contain signals to permit the evaluation of the CD player performance. Tracks 3, 4 and 5 provide a reference (1001 Hz), low (40 Hz) and high (19999 Hz) sinusoidal tones for evaluating frequency response. By setting an oscilloscope to the X-Y mode and selecting the 19999 Hz signal of track 5, LEFT-RIGHT channel phase error can be evaluated. As mentioned previously in the "Phase Monitoring and Measurement" subsection, channel phasing is critical to monaural reproduction. Furthermore, since the CD player is being used as a signal generator, good engineering practice dictates that its performance be verified before using the player to test other pieces of equipment. When selecting a CD player, improved performance will be noted in players that employ dual D/A conversion and oversampling.

In addition to the CD performance tracks, both volumes of the NAB Test CD provide a variety of commonly used broadcast test signals. Although a spot frequency test was provided for the CD player in tracks 3 through 5, tracks 14 through 29 provide signals for more thorough measurements of frequency analysis. Beginning with a 400 Hz reference, these tones ascend in frequency from 20 Hz to 20 kHz. Not only do these tones provide the basis for measuring the frequency response of equipment, but using an Audio Analyzer set to measure total harmonic distortion (THD), these tones can be used to evaluate total harmonic distortion.

Evaluating the nonlinearities that occur when two frequencies are mixed is the basis of the intermodulation distortion (IMD) test as mentioned earlier in this chapter. Track 30 provides two types of SMPTE (Society of Motion Picture and Television Engineers) IMD signals. The first indexed signal consists of 60

Hz/7 kHz sine waves mixed at the typical 4:1 ratio, and is used in making tests through linear systems. The second index provides the same sine waves, but mixed in a 1:1 manner. When evaluating composite FM systems, the insertion of de-emphasis converts the 1:1 ratio back to 4:1, if pre-emphasis is not used.

For AM transmission systems, a CCIF (International Telephone Consultative Committee) IMD test is provided on tracks 31 through 36. These tone pairs could be used to evaluate the performance of the NRSC-1 bandstop filter.

A series of measurement tests for stereophonic performance are also included. Calibrated and indexed phase shifts between the left and right channels over a ±360° range can point to phasing integrity problems in tape machines or broadcast audio chains employing several pieces of processing equipment (track 45). Tracks 54 through 61 provide a variety of the most common pre-emphasis and de-emphasis curve functions. By playing a pre-emphasis track through its complementary de-emphasis network, a flat audio response (all amplitudes for each frequency are equal) should be measured.

Dynamic range can be measured using track 68. This track provides a 400 Hz tone which drops in amplitude by 5 dB every five seconds. The sweep begins a 0 dB and ends at −60 dB. Such a tone can also be used to calibrate level indicating devices.

Earlier in this chapter, the detrimental effects of noise were discussed. True random noise can be a valuable ally in measuring the performance of a system.[18] The NAB Test CD provides five different noise test signals in tracks 46 through 50. The white noise signal found on track 46 provides a constant output over the frequency of the noise, and this makes frequency response measurements very easy using a spectrum analyzer or wave analyzer. By passing white noise through a filter with a lowpass function that attenuates the noise at 3 dB per octave, the signal contains equal spectral power density in each octave, or doubling of frequency. Since pink noise contains equal spectral energy per octave, a flat frequency response is displayed as equal level readings in each band of a one-third or full octave real time analyzer (RTA). The frequency response of white noise and two commonly used weightings can be seen in Fig. 32.

Volume I of the test CD provides a number of tests for use with RF transmission equipment. Tracks 37 through 44 permit FM modulation monitor calibration. At certain modulation frequencies, the amplitude of the FM carrier goes to zero, as all the transmitted power is distributed at frequencies other than the carrier.

The tone tracks on Volume I insure a carrier null at 100% modulation for a given frequency deviation.

Using the triangular waves found on track 70, AM transmitter modulator linearity is quickly defined. As modulation levels are increased, nonlinearities appear as imperfect edges of the diamond-shaped scope trace.

For evaluating NRSC compliance, track 53 provides the ten minute pulsed-USASI noise test.

The second volume of the NAB Test CD includes a complete reel-to-reel and tape cartridge recorder alignment sequence. Organized in a practical order,

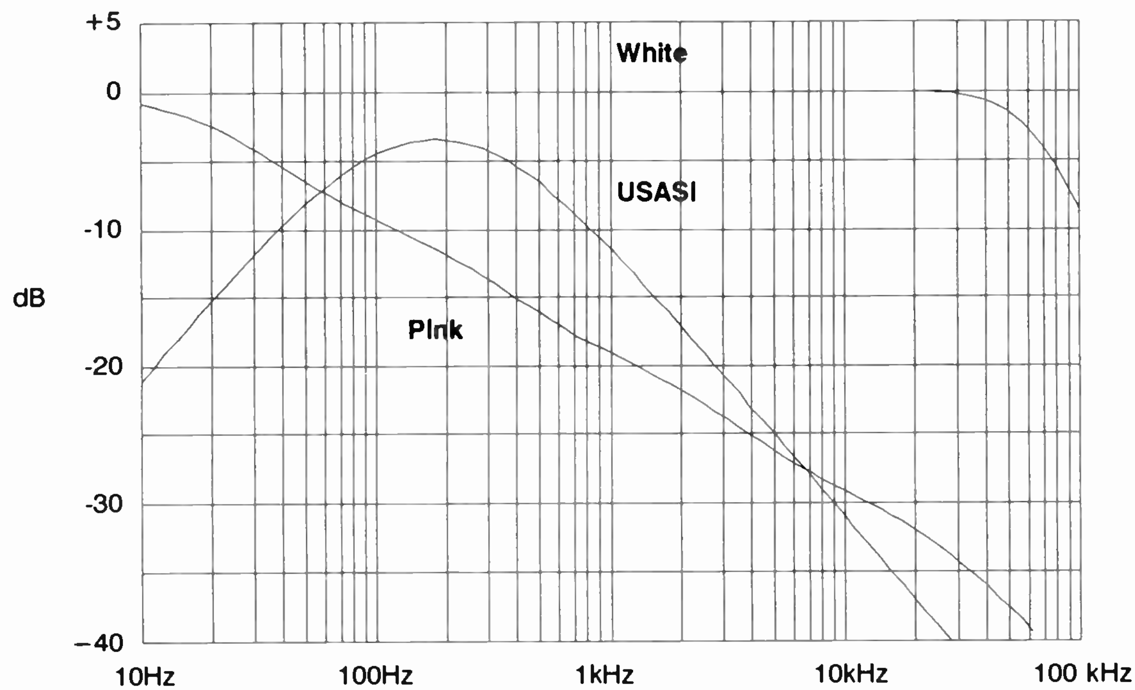


Figure 32. Response versus frequency for white noise and two popular noise weightings included in the NAB Test CD.

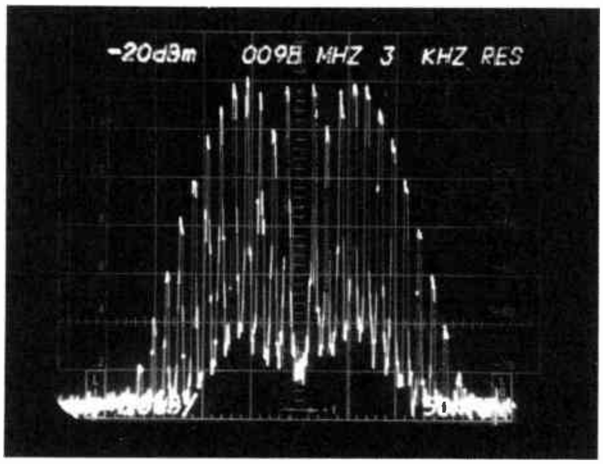


Figure 33. Spectral analysis of an FM transmitter modulated to ±75 kHz (13856.8 Hz, track 38) showing the Bessel carrier nulling effect.

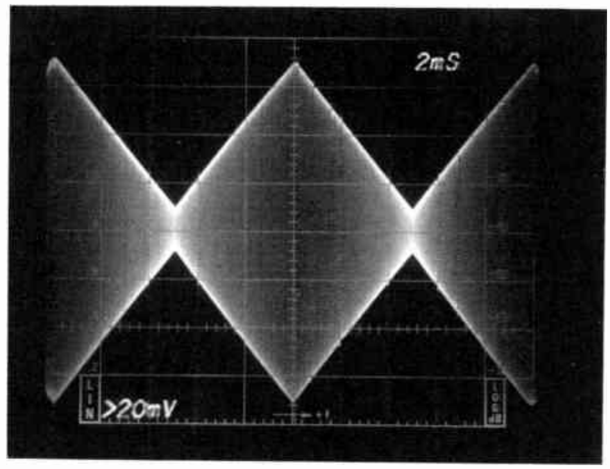


Figure 34. As the AM carrier is modulated, the edges of the diamond shape should be clearly defined. Track 70 provides the 100 Hz signal for making this test.

these tracks cover azimuth alignment, bias and equalization adjustments, as well as frequency response measurements.

Earlier in this chapter, the importance of proper left and right channel phasing was discussed. Volume II of the NAB Test CD approaches this important parameter by providing a unique means of determining proper head azimuth alignment. When a single high frequency tone is used, it is possible to align the head to one of many peaks. Only one peak will provide maximum signal output, but unless close attention is paid, incorrect adjustment can open a Pandora's box of problems.

In both the reel to reel and cartridge alignment sequences, the high frequency azimuth alignment tone is chopped and mixed with pink noise. When viewed on an oscilloscope set in the X-Y mode, the technician has the benefit of a tone for adjusting the head and "closing the loop" as well as the "fuzzball" of pink noise which will only collapse to the in-phase diagonal line at one point. Field tests of this new multiplexed tone/noise pattern appear to have made azimuth alignment virtually fool-proof.

Both volumes of the NAB Test CD include calibration test tones for verifying performance of a number of instruments. In Volume I, tracks 80 through 95 are used to test both VU (volume unit) and PPM (peak program meter) meters. The peak flasher circuits of FM modulation monitors can be checked using the tone bursts found on tracks 98 and 99. Phase meters can be calibrated by using the precision indexed phase shift signal located on track 45. Tracks 76 through 79 contain signals with a calculated amplitude of second harmonic distortion. These tones can be fed into a total harmonic distortion (THD) analyzer, discussed earlier in this chapter, for calibration verification.

Volume II of the NAB Test CD provides the test signals for calibrating wow and flutter meters. These tracks have specific amounts of flutter added to the fundamental test signal (either 3.15 kHz or 3 kHz).

With the price of over-sampled, dual converter consumer-grade compact disc players falling and the wide range of measurements that the Test CDs permit, the NAB Test CDs are fast becoming the industry measurement standard. When the Test CDs are coupled with a portable CD player, a powerful measurement and diagnostic tool is born—one which can be easily carried into the field to provide the user with a test tone generating system unequaled in size or cost.

REFERENCES

1. Graeme, J., Applications of Operational Amplifiers, McGraw-Hill, 1973 pp. 132–139, 202–203.
2. "A New Standard Volume Indicator and Reference Level," Proceedings of the IRE January 1940, pp. 1–17.
3. Harry, D., "Audio Program Analysis," as presented to the National Association of Broadcasters, April 1985.
4. IEEE Standard: Recommended Practice for Audio Program Level Measurement, DOC. G-2.1.2/13, 1988.
5. Hewlett Packard Co. "Using a Narrow Band Analyzer for Characterizing Audio Products," AN192, Oct. 1975.
6. Crown International Inc., BDP-2 The Audio Micro Computer, equipment manual, September 1981.
7. Skritek, P. and H. Pichler, "Extended Application of T.I.M. Test Procedures," preprint no. 1557, 64th AES Convention, Nov. 1979.
8. Takahashi, S. and S. Tanaka, "A New Method of Measuring Transient Intermodulation Distortion: A Comparison with the Conventional Method," preprint no. 1539, 64th AES Convention, Nov. 1979.

9. Tektronix, Inc., AA501 Distortion Analyzer, equipment manual, February 1984.
10. National Semiconductor Corporation, Audio/Radio Handbook, 1980, pp. 2.3–2.10.
11. Hewlett Packard Co., "Spectrum Analysis . . . Noise Measurements," AN150-4, April 1974.
12. National Association of Broadcasters, Technical Standards, 1990, Section 5, pg 23.
13. Dolby, R., D. Robinson and K. Gundry, "CCIR/ARM: A Practical Noise-Measurement Method," Journal of the Audio Engineering Society, vol. 27, March 1979, pp. 149–157.
14. Mendenhall, G., "The Composite Signal-Key to Quality FM Broadcasting," Broadcast Electronics.
15. Graeme, J., Designing with Operational Amplifiers, McGraw-Hill, 1977, pp. 76–77, 251–253.
16. Salek, S., "The NAB Test CD—Use and Applications", as presented to the National Association of Broadcasters, April 1989.
17. National Association of Broadcasters, The NAB Broadcast and Audio System Test CD Instruction and Applications Booklet, 1988.
18. Gardner, H. and J. Bisset, "Employ Some Noise", Delta Electronics, Inc.
19. Driscoll, F.F., Analysis of Electric Circuits, Prentice-Hall, Inc., 1973, pp. 39–49.
20. Hewlett Packard Co., "Spectrum Analysis . . . Distortion Measurement," AN150-11, Oct. 1976.
21. The Institute of Electrical and Electronics Engineers, IEEE Standard Dictionary of Electrical and Electronics Terms, Third Edition (1984).

7.2
Radio Receivers

Jon GrosJean
Woodstock Engineering, South Woodstock, Connecticut

PURPOSE OF THIS CHAPTER

This chapter aims to acquaint the broadcaster with the characteristics of typical receivers used by the majority of the listening public. There is a wide variety of receivers currently in use from pocket radios with very limited frequency response to component receivers which have specifications better than most transmitters and everything in between. This makes it difficult for the broadcaster to tailor the characteristics of the audio signal so that it sounds good on all types of receivers. Above all, the broadcaster must understand that basic physical laws and, except in the case of the most expensive receivers, the retail price determine how good the performance of the receiver will be.

OVERVIEW OF CURRENT RECEIVER TECHNOLOGY

Virtually all radios made today are superheterodyne type having the configuration shown in Fig. 1. The incoming signal beats with the tunable local oscillator to produce, among others, the intermediate frequency (IF) signal.

Intermediate frequencies usually used in radios are:

AM radios, except car: 455 kHz
AM car radios: 450 kHz
FM radios: 10.7 MHz

The local oscillator typically operates *above* the desired signal. That is, while the heterodyne or mixing process produces new frequencies at both sum and difference (and multiples) of the original frequencies, most radios designed for broadcast reception in the U.S. use the difference signal:

$$F_{osc} - F_{in} = F_{if}$$

Other mixing products and the input and oscillator frequencies are removed by the IF filter. It should be noted that any other signal that gets through the passband of the tuned input circuit, if any, will also mix with other signals (including F_{osc}, F_{in}, F_{if}, and

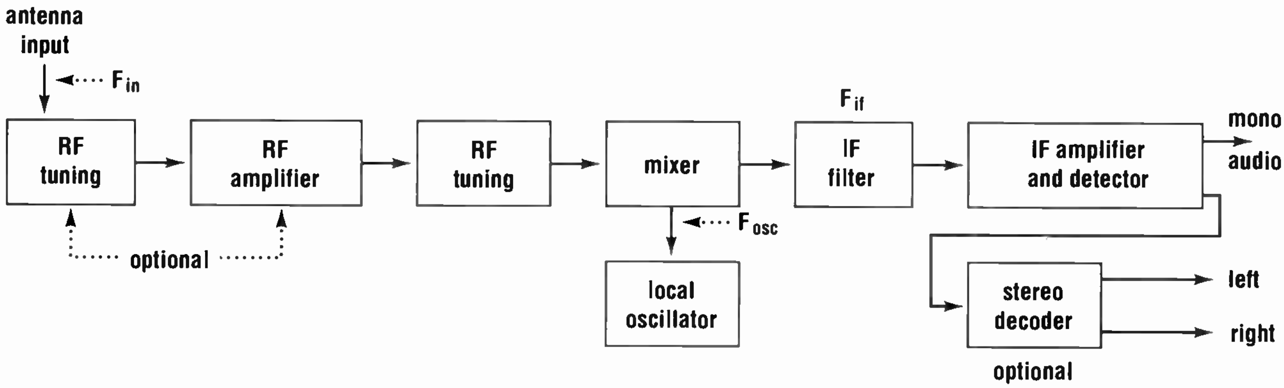


Figure 1. Block diagram of the most common, basic design for radio receivers.

other incoming signals and mixing products) to produce still more undesired signals. The implications of this are discussed below.

Not shown in this basic diagram is the feedback path from the detector to the radio frequency (RF) section used for automatic gain control (AGC), to compensate for variations in the level of the incoming signal. In FM radios that do not use synthesized local oscillator frequency, the oscillator is usually stabilized by another feedback circuit for automatic frequency control (AFC).

Interference Sources

Radio receivers are susceptible to certain types of interferences, commonly called spurious responses, that are important to radio broadcasters. Some of these are a result of the local oscillator in the superheterodyne design.

Image

In most receivers, the oscillator operates above the desired signal frequency. The *image* interference occurs at $F_{if} + F_{osc}$ (or $F_{in} + 2F_{if}$). That is, an undesired signal on the "other side" of the local oscillator will produce a beat frequency at the intermediate frequency. If a radio were tuned to a station transmitting on 570 kHz, the local oscillator would be operating at 455 kHz above, or 1025 kHz. If another station in the area were operating at 1580, it would beat with the 1025 kHz to produce 1580 − 1025 or 455 kHz.

This is mostly a problem in AM because the interfering frequencies are in the AM band. In FM receivers, the image frequencies are in the aircraft band.

Half-IF

The *half-IF* response occurs at $F_{in} - F_{if}/2$. This is mostly a problem in FM receivers, because the interference appears 5.35 MHz above the desired signal and, in the case of FM, the deviation of the undesired signal is doubled. This means that half-IF interference can occur in a large part of the FM band.

Oscillator 2nd Harmonic

The *oscillator 2nd harmonic* response is at $2F_{osc} - F_{if}$. This is mostly a problem in FM receivers because TV channel 13, at 213 MHz, will appear at 101 MHz.

IF Mixing

The *IF mixing* response occurs at $F_a - F_b = F_{if}$, where F_a and F_b are any two FM channels that are approximately 10.7 MHz apart. This has been a subject of controversy, sometimes called the "FM-IF taboo." Integrated circuit-balanced mixers virtually eliminate this problem.

Intermodulation

Intermodulation is simply the mixing of two signals to produce a mixing product at another frequency. This is also heterodyning, but the radio's local oscillator is not involved. For instance, two strong AM signals at 700 kHz and 800 kHz can produce a response at 1500 kHz. If the receiver is tuned to 1500 kHz, the listener could hear the combination of the 700 and 800 kHz signals. Older car radios with tuned inputs usually did not have this problem, but the newer digitally-tuned ones do not have a tuned circuit before the RF stage and are more susceptible to this problem. Nonlinearity in the RF stage causes the mixing product, which is received as a normal signal by the radio when it is tuned to 1500 kHz.

The *second-order intermodulation* response is not unique to superheterodyne receivers and is very important. It is due to nonlinearity in the RF section and is caused by the second harmonic of one RF signal mixing with another. It can be determined from $F_{in} = 2F_a - F_b$. For example, if the radio is tuned to 1000 kHz, strong signals at 1040 kHz (F_a) and 1080 kHz (F_b) will produce a signal at 1000 kHz. The same situation occurs in FM receivers. Receiver susceptibility is usually measured using two signals of equal strength, spaced four and eight channels above or below the desired frequency.

Crossmodulation

Crossmodulation is unique to AM radios because it is caused by amplitude modulation of a desired signal by an undesired signal. It is often measured using both a desired signal and an undesired signal modulated 80% spaced 40 kHz from the desired frequency. Receivers are usually most susceptible to this when the received desired signal is relatively strong; around 1 mV/m field strength.

Desensitization

Desensitization occurs when a strong signal, which does not produce a spurious response, causes a usable signal to become noisy. In most cases, spurious responses will appear before desensitization becomes noticeable, but many radios, especially car radios, have wide-band AGC systems which are designed to reduce the gain of the RF stage if there is a signal strong enough to cause spurious responses. The undesired signal may be at almost any frequency, but those within the band usually have more effect. As an example, a listener tuned to a local station at 1500 kHz may notice that it fades out during part of a journey which happens to take him/her near another station's transmitter at 1200 kHz.

Other Interference Sources

Interference is any undesired energy that degrades a desired signal. Sources of undesired energy can include not only other broadcast stations but also other transmissions, thermal and impulsive noise, all manner of electrical and electronic devices, and even the desired signal arriving by a different path (multipath).

Stronger signals will produce responses to higher harmonics of the local oscillator and incoming signal. For example, a strong paging signal at 432.5 MHz can mix with the fourth harmonic of an FM receiver local oscillator and be received at 100.1 MHz. There are also many other sources of interference in a normal home or business, and these are usually more of a problem with AM than FM reception. Computers can cause TV, FM, and AM reception problems, and as they tend to have higher and higher clock speeds, they may produce interference at higher frequencies. The harmonics of the deflection system of TV sets will usually make an AM radio almost unusable in the home if the signal strength is less than 1 mV/m.

In the last 10 to 15 years the increase in solid-state controls on electrical devices has greatly increased the possibility for AM radio interference. Light dimmers are often very bad, but there are low Radio Frequency Interference (RFI) types with internal filters available. Triac controllers and solid-state relays for motors can produce very strong signals on the power lines. Many of these problems can be fixed by simply putting power line filters on the offending device. With the proliferation of computers, these filters are commonly available. A line filter on a fluorescent lamp will completely eliminate the line conducted interference they produce.

Noise blankers are included in almost all FM car radios currently produced. They simply blank out the composite audio signal before the stereo decoder when a noise pulse is detected, and virtually eliminate ignition noise from other vehicles or even the one in which the radio is being used. ICs that perform the same function in AM have recently become available but since the cost of the AM tuner in a radio is so low, addition of a noise blanker significantly increases the cost. When properly designed in a radio, they can almost completely eliminate impulse noise caused by light dimmers, car ignitions, and some sparking of high voltage power lines. Lightning-generated noise cannot yet be eliminated because the pulse length caused by lightening is very long, and the blanker would have to blank out a long segment of the program material.

Almost all AM and FM receivers of a given category (automobile, home, portable, etc.) have virtually identical performance with respect to distortion, frequency response, spurious responses, and even power output. One exception to this is starting to occur because of the adoption of the National Radio Systems Committee (NRSC) AM receiver standard, but receivers complying with this standard will probably be limited to higher cost AM tuners because of the requirement for a dual bandwidth IF filter. The main reason for all the radios being the same is that compromises must be made when designing a radio. Most of these compromises are dictated by the range of signals which must be accommodated and the physical limitations of the devices being used. For example, strong-signal performance or spurious responses and distortion are mostly determined by the basic characteristics of the integrated circuits (ICs) or transistors, and audio frequency response is determined by the IF filter bandwidth. Many manufacturers of radios purchase integrated circuits, transistors, and ceramic IF filters from the same vendors, so it follows that most receivers will have approximately the same performance.

Double-Conversion Receivers

Many of the shortwave and some of the newer AM car radios have an additional mixer, oscillator, and lower frequency IF filter before the first RF as shown in Fig. 2. They are called dual conversion or sometimes (incorrectly) up-converter receivers.

This arrangement is usually used when the receiver covers a wide frequency range or to eliminate the requirement for tuning the RF section and is particularly suited to synthesized radios.

Typical frequencies for dual conversion receivers are 10.7 MHz and 450 kHz. To receive a station at 1000 kHz, the first oscillator would operate at 11.7 MHz so that the difference would be 10.7 MHz. The second oscillator would probably be a crystal operating

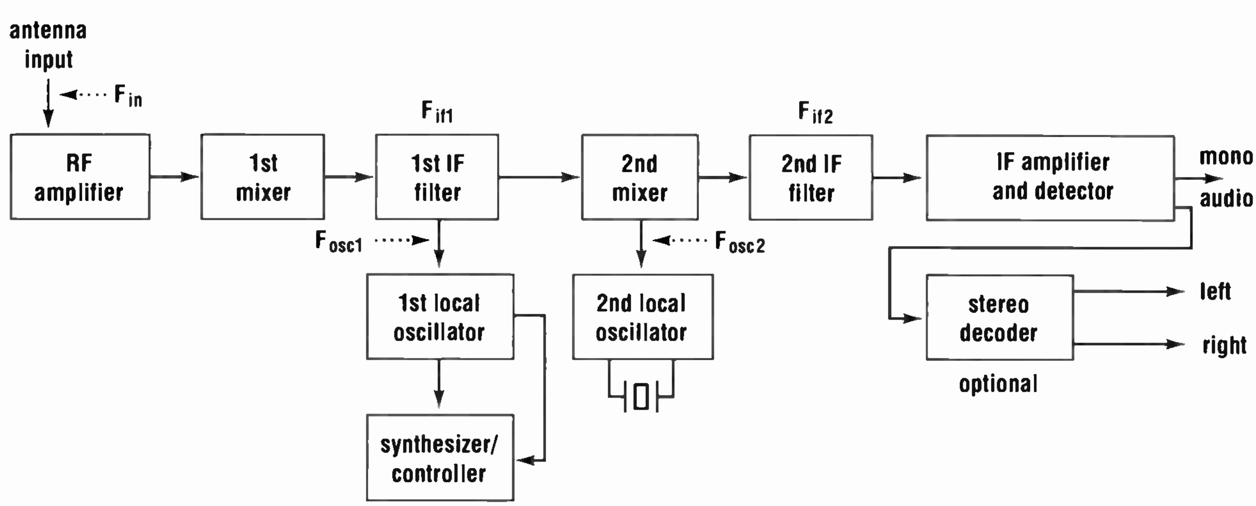


Figure 2. Block diagram of a double-conversion radio receiver.

at 10.250 MHz; note that this is *below* the first IF. The difference would be the 450 kHz second IF. The *image frequency* for the first IF would be 22.4 MHz. This could easily be removed with a low-pass filter.

If the receiver tunes the shortwave band from 150 kHz to 30 MHz, the first IF is often 45 or 70 MHz, so that it is outside the band being tuned. The addition of another mixer, however, results in another set of spurious responses and the possibility of beats being produced from mixing of harmonics of the two oscillators. The beat frequencies can be calculated from the equation:

$$F_{if2} = mF_{osc1} - nF_{osc2}$$

where m and n are integers from 0 to the desired harmonic number including positive and negative values.

FILTER DESIGN CONSIDERATIONS

All filters must be designed according to basic rules, and there are no exceptions. Filters made with coils, crystals, or ceramic resonators follow the same rules. Surface wave and digital filters are slightly different in that attenuation and group delay are not interrelated, but these types are relatively rare in consumer radios. Currently, it is not practical to make a digital IF filter operating at 455 kHz or 10.7 MHz, because the required digital processing speeds are too high for the current technology and the resulting integrated circuit too complex to be cost effective. Furthermore, when digital circuits operating at high speeds are put inside a radio, it is very difficult to keep the spurious radiated signals from interfering with the desired signal coming into the radio.

The bandwidth of a filter is determined by the number of elements or poles and the Q of each element. Each resonant element constitutes a pole or technically a pole pair. Very high Q elements like crystals can be used for designing very narrow filters such as those in communication receivers. Intermediate and low Q elements like ceramic filters and coils are good for designing filters with wider bandwidths such as those in AM and FM broadcast receivers.

Group Delay and Frequency Response

The group delay and response of a filter are related in that a filter with a sharp cutoff at the band edge has poor group delay. Prior to the days of stereo AM and FM, there was little concern for group delay, but it is now a very important factor in achieving low distortion and good separation. Unfortunately, it is not possible simply to add more elements to a filter to achieve a desired selectivity while preserving a desired 3 dB bandwidth. As the number of elements increases, the Q of each element must increase. For instance, if a four section AM IF filter has a 3 dB bandwidth of 15 kHz and a 40 dB bandwidth of 60 kHz, increasing the number of elements in the filter to six may increase the 60 kHz attenuation to 80 dB but may also reduce the 3 dB bandwidth to 12 kHz. In fact, it may not be possible to build the six-section filter because the Q of the elements may not be high enough. The result is an increase in the sensitivity at the expense of audio frequency response. Conversely, if more attenuation at 40 kHz is desired with a given number of elements, the 3 dB bandwidth must be decreased. The same situation occurs in an FM receiver when more adjacent channel rejection is desired. Simply adding more filters increases the selectivity, but instead of the audio frequency response changing, the distortion and stereo separation become worse.

Therefore, there is always a compromise between the signal bandwidth and selectivity. This compromise is determined by the number of poles the filter has and the Q of the resonators. There is also a compromise between the shape of the filter response and the desired distortion and stereo separation of the radio. It is not possible to build a filter with very high selectivity and have good group delay to the edges of the passband unless surface wave filters are used. This is the reason that an AM radio meeting the NRSC specification must have two switchable IF filters in order to have enough selectivity for nighttime reception, and why FM car radios with very high adjacent channel selectivity often have relatively high distortion. In FM SCA receivers, the group delay is even more important in determining the crosstalk between the main and SCA signal. Since the group delay must be very good, the selectivity cannot also be good, and so most FM SCA receivers have relatively poor or no adjacent channel selectivity.

It should be noted that the NRSC *AM Pre-emphasis and Audio Transmission Bandwidth Specifications* allow for modulating frequencies up to 10 kHz. Since the channel spacing is 10 kHz, the modulation components will appear in the adjacent channel, and it is not possible to filter out these components when the radio is tuned to the adjacent channel. In order to eliminate adjacent channel interference with receiver filtering, the modulating frequency must be less than 5 kHz.

POCKET RADIOS

Most pocket radios and so called personal, portable, or Walkman®-type radios use a single AM/FM integrated circuit which often contains all the circuitry for the whole radio, often including the audio amplifier. If the FM portion contains a stereo decoder, it may or may not be contained in the basic radio IC and the audio amplifier will be a separate IC. Their AM sensitivity is determined by the size of the ferrite antenna which will fit in the case. Since the case is small, the ferrite antenna is small and the AM sensitivity is usually poor. The FM antenna is usually the headphone cord or, if it has a speaker, a small whip, or telescoping, antenna. Unfortunately, ferrite antennas are directional, and using the earphone wire for an AM antenna does not work well. The AM audio response in these radios is determined by the AM IF filter and the ferrite antenna since it is a tuned circuit with a high Q. The IF filter usually does not contain many elements because of cost and size considerations, so, as was mentioned above, the 3 dB bandwidth must be made small in order to achieve enough adjacent or alternate channel selectivity.

The ferrite antenna in these radios presents some other unique problems. First, it is the only RF tuned circuit, and the image response is particularly important in AM receivers because it occurs in the AM band. For example, a receiver tuned to 600 kHz has an image response at 1500 or 1510 kHz depending on the IF frequency. Image frequencies for AM and FM are listed in a table at the end of this chapter. Therefore, in order to obtain enough image rejection and to have the best sensitivity, the loaded Q of the antenna must be kept high. At 600 kHz, an antenna Q of 100 will result in an audio 3 dB bandwidth of 3 kHz independent of the IF filter. This results in a 3 dB detected audio response of about 1.5 kHz. Therefore, in this type of radio, it is just about impossible to have good AM audio fidelity, sensitivity, and image rejection at the same time.

On the other hand, the FM performance of these radios may vary greatly depending on the circuit configuration. If the AM/FM IC does not contain the FM tuner section, discrete transistors are used for this part. The spurious response rejection, particularly the IF mixing and half-IF responses, are usually not very good. If they are good, the sensitivity is usually poor. If the FM tuner portion is in the receiver IC, it usually contains a balanced mixer, and the spurious response rejection may be very good. In this case, the sensitivity can be made better because of the better performance of the circuits in the IC. The audio quality of these receivers is not affected by the RF circuits, and it usually does not vary much between units. It is difficult to make any FM radio which has poor audio performance.

TABLE RADIOS AND COMBINATION UNITS (BOOM BOXES)

With a few exceptions, the radio tuner circuitry in these units is essentially the same as that in the pocket radios. The AM ferrite antenna is sometimes larger because the physical size of the radio is larger, so the AM sensitivity is often better. Unfortunately, the Q limitation mentioned above still exists, so the AM audio frequency response is usually not much different from that of the pocket radios. The only major difference from the pocket radios is that the audio amplifier power capability is higher. This is due to the higher power supply voltage and available current from larger batteries. Recently some higher quality table radios have become available. These use most of the same ICs which are used in car radios and so have very good spurious response rejection. The sound quality of these units can vary considerably from very good, to poor and very boomy. Most of the designers producing them have very little knowledge of acoustics.

TABLE 1

Representative specifications: pocket, table, and combination radios.

Sensitivity*: AM:	1 to 5 mV/m
mono FM:	5 to 20 μV/m
AM image response:	400 mV/m
FM half-IF response:	4 to 20 mV/m
FM oscillator 2nd harmonic response:	8 mV/m
IF mixing response:	10 to 100 mV/m
FM second-order intermodulation response:	4 to 10 mV/m
AM crossmodulation rejection:	500 mV/m

*RF signal for 30 dB audio signal-to-noise ratio.
For telescoping or whip antennas, the RF input level to the receiver is approximately equal to:
1.2 the field strength in μV/m, for FM
1.5 the field strength in μV/m, for AM

HOME TUNERS AND TUNER-AMPLIFIERS

The AM portion of most home stereo tuners either does not exist or has about the same performance as a table radio. They often have a ferrite antenna and an IF filter with a coil and single 2-pole ceramic filter. Recently, many receivers have come with a detachable loop antenna which can be moved to overcome the directional effect for a desired station. This type of antenna is the same as those used on the back of the old tube table radios, and has the advantage that the Q can be made lower with a corresponding loss of sensitivity so that an AM tuner with better audio fidelity can be made. If this is done, another RF tuned circuit and an RF amplifier stage is usually required, but the additional cost is often not too great for this type of receiver. With the advent of the NRSC standard for AM receivers, this is one of the more likely types of AM tuners to be improved by the addition of a second switchable IF filter for daytime Hi-Fi reception. A wire type of antenna could also be used, but this is less convenient for many users and it sometimes picks up more interference from household wiring than a loop antenna. The FM section of home Hi-Fi tuners received the most attention for several years, and some of the analog tuners used in the '60s and '70s were very good. Now, the good tuners are mostly reserved for the high end expensive sets, and the others have suffered from severe price competition. Many of the digitally tuned receivers now use car radio components, the characteristics of which are described in the paragraph on car radios. One problem with most digital tuners is that they cannot be offset when an adjacent channel is causing interference. Some models do allow offsets in 25 kHz increments, but the offset may not be stored in the memory selection. In recent years, the proliferation of FM stations as well as the use of heavy audio processing has resulted in the need for better selectivity, especially in digitally controlled tuners which can be tuned in only 200 kHz steps. This is accomplished by using ceramic IF filters with a narrower bandwidth or by using three filters instead of the normal two. As was mentioned in the section

TABLE 2

Representative specifications: home tuners and tuner-amplifiers.

Sensitivity* AM:	0.3 to 1 mV/m
mono FM:	3 to 5 μV/m
AM image response:	3 to 50 mV/m
FM half-IF response:	3 to 50 mV/m
FM oscillator 2nd harmonic response:	6 to 200 mV/m
FM IF mixing response:	20 to 200 mV/m
FM second-order intermodulation response:	8 to 30 mV/m
AM crossmodulation rejection:	500 mV/m

*RF signal for 30 dB audio signal-to-noise ratio.

on Filter Design, a narrower IF bandwidth in an FM receiver will produce more distortion and sometimes less stereo separation. If the tuner contains a stereo decoder which will blend on noise or multipath signals, the intermodulation distortion caused by the narrow IF filter will also cause the stereo decoder to blend on highly processed music because the deviation of the signal is at 100% most of the time.

In order to accommodate this situation, some higher priced FM tuners have a switchable IF bandwidth. In the wide position, the distortion is very low where it is specified in the specifications. In the narrow position, for which the distortion is often not specified, the selectivity is very good, but the distortion may be over 1%.

CAR RADIOS

Car radios have received by far the most attention in recent years. Additionally, they cover the widest range of performance parameters from very bad to very good. The manually tuned radios generally use a separate IC for FM and AM. The AM section is often quite good with respect to spurious responses and sensitivity because they have two tuned RF circuits, but since these are low cost radios, a minimum amount of IF filtering is used. In terms of audio fidelity, they are about the same as the pocket radio. Many times, they have a wide-band RF automatic gain control (AGC) circuit which will desensitize the receiver if there is a strong signal near the desired one. The only indication the listener will have of this is that the desired signal will fade out.

The FM section of the analog car radios is usually poor at best. Most use two or three transistors in the RF section. The trade-off in these tuners is between sensitivity and spurious response rejection, and in most cases, they are not very good at either. Since the tuners are subject to temperature drift because of the wide temperature range inside an automobile, they usually have an automatic frequency control (AFC) which cannot be turned off. This combined with limited IF selectivity causes the radio to jump from one station to the other when tuning and sometimes when it is not being tuned as one station fades out and another on an adjacent channel becomes stronger. These radios are becoming less common because the cost of the digital and corresponding voltage tuned circuitry has become comparable to the cost of a mechanical tuning mechanism.

Digitally tuned car radios have a wider range of performance parameters even though the components used may be very similar or the same. Most manufacturers now purchase the AM and FM tuners from tuner module manufacturers, but a few, both Japanese and American, still make their own tuners. In order to make a digitally tuned radio, some method of tuning the radio with a voltage must be used. Varicap diodes are the only choice, but on AM, this presents a problem because the relatively large capacitance of the cable from the antenna to the radio makes tuning the input to the RF stage impractical. The solution so far has been to use a low-noise junction field effect transistor (JFET) RF amplifier connected directly to the antenna input. Unfortunately, this means that all signals picked up by the antenna appear at the FET input with approximately equal amplitude, and undesired signals outside of the AM band may cause interference. The most common solution for this problem is what is called a wide band AGC system, but this causes another problem which most broadcasters or listeners are not aware of, and was referred to above under desensitization. One AM IC car radio used by about 90% of the world's manufacturers has a wide band AGC circuit which is not tuned. Sometimes it is more sensitive to signals in the shortwave band than to those in the AM band, so a desired station which is normally at a high enough level to give a 40 dB signal-to-noise ratio may completely disappear if another transmitter is in the vicinity of the radio. This AGC system is very good at preventing spurious responses by eliminating all signals which could cause a spurious response. A couple of U.S. manufacturers of car radios are aware of this problem, and make sure the wide-band AGC system is not too wide.

Since the input to the AM section of car radios is not tuned, a double tuned interstage RF filter is used. This arrangement is very easily made into a wide-band filter so that it does not limit the audio fidelity very much. Thus, in many current car radios, the audio frequency response is mostly determined by the bandwidth of the IF filter. Since car radios must operate over a wide range of signal strengths, it is very important to have enough adjacent channel selectivity, and the most common requirement is at least 40 dB. The limitations in the IF filter then dictate a 3 dB bandwidth of about 6 to 7 kHz and a resulting maximum 3 dB audio response of 3 to 3.5 kHz. The audio response falls off very rapidly above these frequencies because of the very high attenuation of the filter. Thus to make an AM car radio which will meet the NRSC specifications, a switchable IF bandwidth is required.

As was mentioned above, most car radio manufacturers purchase their FM tuner modules. They all use an integrated circuit balanced mixer, and most use a dual gate FET RF amplifier. This results in a tuner with very good half-IF rejection and usually good sensitivity. They, like the AM tuners, have a wide-band RF AGC circuit which is supposed to prevent spurious responses, mostly the second-order intermodulation. It is usually set so that two signals will produce a response which is slightly above the receiver noise level. This is difficult to set exactly, so one particular brand of receiver may have significantly better performance than another. Many receivers have what is called keyed AGC, so that if a desired signal falls on the intermodulation response from two other signals, it will reduce the wide-band AGC threshold level. The result is that the desired signal will be noisier, but the undesired signal will not be heard. It should be noted that every radio manufacturer has a favorite testing location in a particular city, so one brand of radio may work well at this location while another may not. The converse may be true in another location.

As many broadcasters know, the biggest problem with FM car radios is multipath reception and fading problems. Receiver manufacturers have been trying to find a solution to this for many years. So called soft mute and stereo blend are the most common approaches used today. Soft mute is simply a muting circuit which attenuates the detected audio signal at a predetermined slope and rate as the RF signal level decreases. The RF level at which mute starts varies significantly between manufacturers, and the speed of mute and recovery are very important in preventing noise bursts in the audio signal. The amount of mute also varies depending on the manufacturers preference, and common levels of interstation noise vary from −20 to −50 dB. Many sets also include a high-cut circuit which attenuates high frequencies an additional 10 to 15 dB at low signal levels to reduce audio noise.

The other technique for dealing with fading and multipath is to blend the stereo into mono as the signal strength decreases or multipath increases. FM stereo signal-to-noise ratio (S/N) is 22 dB worse than for mono at levels below about 300 μV, so the obvious solution for improving the S/N is to blend the stereo decoder to mono as the signal level decreases. The level at which blend starts vary considerably from one manufacturer to another about 1 mV to 50 μV RF input level. Some manufacturers feel quiet mono is better than noisy stereo, and others feel the opposite. Unfortunately, multipath signals cause much more distortion to the FM stereo than the mono signal, but a blend controlled by the RF signal level cannot distinguish between a good strong signal and a strong one with a lot of multipath. Thus these sets may be very distorted in the presence of a strong signal with a lot of multipath. A few receivers have provisions for blending under multipath conditions, and this may become more common in the future. The FMX(R) system is supposed to overcome some of the requirement for blending on weak signals by companding a quadrature modulated L−R signal so that the S/N is better.

Another problem with FM stereo decoders which is slowly improving is the interference caused by adjacent channel signals. Most stereo decoders operate by multiplying a 38 kHz square wave with the composite signal. The fifth harmonic of 38 kHz at 190 kHz beats with an adjacent channel signal at 200 kHz to cause a swishing sound in the background. Several newer FM stereo decoders utilize a switching waveform called a Walsh function which virtually eliminates the harmonics which beat with adjacent channel signals.

TABLE 3

Representative specifications: car radios.

Sensitivity* AM:	200 μV/m
mono FM:	3 to 10 μV/m
AM image response:	60 mV/m
FM half-IF response:	60 mV/m
FM oscillator 2nd harmonic response:	200 mV/m
FM mixing response:	200 mV/m
FM second-order intermodulation response:	8 to 20 mV/m
AM crossmodulation rejection:	60 to 500 mV/m
Desensitization level: AM:	10 to 500 mV/m
FM:	10 to 40 mV/m
RF input at which stereo blend starts:	1 mV to 100 μV.

*RF signal for 30 DB audio signal-to-noise ratio.
For car radio whip antennas, the RF input level to the receiver is approximately equal to:
1/2 the field strength in μV/m, for FM
1/5 the field strength in μV/m, for AM

SCA RECEIVERS

An SCA receiver is simply an FM receiver with the stereo decoder replaced by a filter and FM demodulator tuned to the desired SCA carrier frequency.

Since the deregulation of the FM channel from 53 to 110 kHz, commonly called the SCA band, it has been used for almost every imaginable application. Unfortunately, it suffers from a six to ten times sensitivity loss compared with the main channel because the SCA carrier modulates the main carrier at a maximum of 10%. At lower modulation levels, the loss is proportionally greater. Since the bandwidth of the SCA channel is usually much less than 15 kHz and the pre-emphasis is usually 150 μsec or more, the sensitivity loss is somewhat mitigated.

The most commonly-used SCA frequencies are 67 and 92 kHz. 57 kHz is used for a special phase-shift keyed (PSK) paging system and the radio data system (RDS) system. Both of these systems were developed in Europe where the channels are sometimes spaced 100 kHz, and higher frequency subcarriers are not used. 76 kHz is sometimes used, but the transmitter system must be very good because this is a harmonic of the stereo signal. Any distortion in the system will cause a reduced stereo signal to be at 76 kHz, and since the SCA signal is already at least 20 dB below the main channel, the distortion must be down 50 dB or more if it is not to cause interference.

The biggest problem in an SCA receiver is keeping the received signal strong enough for good signal-to-noise ratio, avoiding multipath, and keeping the main channel stereo signal out of the SCA channels. The signal coming out of the detector of an FM receiver contains the main channel up to 53 kHz at a level which is a minimum of 10 times (20 dB) higher than the SCA signal. Any nonlinearity in the receiver or transmitter will cause crosstalk between the main and SCA signals. Multipath signals will produce the same effect. If the SCA channel is being used for data, this is much less of a problem than when audio or music is being sent. In some cases, the SCA receiver may have less crosstalk than the transmitter.

The phase linearity of the transmitter and receiver are extremely important if the SCA quality is to be good. A good SCA receiver will have a main-to-SCA channel crosstalk of 50 dB. Because of the severe group delay requirements on the receiver IF filter, SCA receivers usually have no adjacent channel rejection because the IF bandwidth must be 250 kHz or more to obtain the good group delay. However, since they are essentially dual conversion receivers, the SCA filtering portion provides the required adjacent selectivity for adjacent channel signals.

In an SCA receiver, it is necessary to filter the SCA signal coming out of the detector before it is applied to the SCA demodulator. The deviation of the SCA signal is usually 5 kHz peak so the SCA signal filter must have good selectivity both to reject the main channel stereo signal and any other SCA carriers which may be in use. As was mentioned above, a tradeoff in group delay and selectivity must be made, and as a result of this, the SCA distortion may not be very good. One to two percent is typical.

Data signals can be sent using the SCA carriers by simply using the digital signal to deviate the SCA subcarrier. Some SCA generators have data inputs. The receiver simply has a comparator connected to the SCA FM detector, and the output of the comparator is the data signal. 1200 baud is a common rate because the bandwidth after the SCA detector can be reduced to improve the signal-to-noise ratio and, therefore, the SCA sensitivity. Sometimes frequency shifted tones from a telephone modem are used, but the allowable data rate is much lower and sensitivity is poorer.

TABLE 4

Table of AM and FM band frequencies with common oscillator frequencies and some spurious responses.

AM BAND

	with 455 Hz IF		with 450KHz IF	
desired	oscillator	image	oscillator	image
540	995	1450	990	1440
550	1005	1460	1000	1450
540	1015	1470	1010	1460
570	1025	1480	1020	1470
580	1035	1490	1030	1480
590	1045	1500	1040	1490
600	1055	1510	1500	1050
610	1065	1520	1060	1510
620	1075	1530	1070	1520
630	1085	1540	1080	1530
640	1095	1550	1090	1540
650	1105	1560	1100	1550
660	1115	1570	1110	1560
670	1125	1580	1120	1570
680	1135	1590	1130	1580
690	1145	1600	1140	1590
700	1155	1610	1150	1600
710	1165	1620	1160	1610
720	1175	1630	1170	1620
730	1185	1640	1180	1630
740	1195	1650	1190	1640
750	1205	1660	1200	1650
760	1215	1670	1210	1660
770	1225	1680	1220	1670
780	1235	1690	1230	1680
790	1245	1700	1240	1690
800	1255	1710	1250	1700
810	1265	1720	1260	1710
820	1275	1730	1270	1720
830	1285	1740	1280	1730
840	1295	1750	1290	1740
850	1305	1760	1300	1750
860	1315	1770	1310	1760
870	1325	1780	1320	1770
880	1335	1790	1330	1780
890	1345	1800	1340	1790
900	1355	1810	1350	1800
910	1365	1820	1360	1810
920	1375	1830	1370	1820
930	1385	1840	1380	1830
940	1395	1850	1390	1840
950	1405	1860	1400	1850
960	1415	1870	1410	1860
970	1425	1880	1420	1870
980	1435	1890	1430	1880
990	1445	1900	1440	1890
1000	1455	1910	1450	1900
1010	1465	1920	1460	1910
1020	1475	1930	1470	1920
1030	1485	1940	1480	1930
1040	1495	1950	1490	1940
1050	1505	1960	1500	1950
1060	1515	1970	1510	1960
1070	1525	1980	1520	1970
1080	1535	1990	1530	1980
1090	1545	2000	1540	1990
1100	1555	2010	1550	2000
1110	1565	2020	1560	2010
1120	1575	2030	1570	2020

AM BAND

	with 455 Hz IF		with 450KHz IF	
desired	oscillator	image	oscillator	image
1130	1585	2040	1580	2030
1140	1595	2050	1590	2040
1150	1605	2060	1600	2050
1160	1615	2070	1610	2060
1170	1625	2080	1620	2070
1180	1635	2090	1630	2080
1190	1645	2100	1640	2090
1200	1655	2110	1650	2100
1210	1665	2120	1660	2110
1220	1675	2130	1670	2120
1230	1685	2140	1680	2130
1240	1695	2150	1690	2140
1250	1705	2160	1700	2150
1260	1715	2170	1710	2160
1270	1725	2180	1720	2170
1280	1735	2190	1730	2180
1290	1745	2200	1740	2190
1300	1755	2210	1750	2200
1310	1765	2220	1760	2210
1320	1775	2230	1770	2220
1330	1785	2240	1780	2230
1340	1795	2250	1790	2240
1350	1805	2260	1800	2250
1360	1815	2270	1810	2260
1370	1825	2280	1820	2270
1380	1835	2290	1830	2280
1390	1845	2300	1840	2290
1400	1855	2310	1850	2300
1410	1865	2320	1860	2310
1420	1875	2330	1870	2320
1430	1885	2340	1880	2330
1440	1895	2350	1890	2340
1450	1905	2360	1900	2350
1460	1915	2370	1910	2360
1470	1925	2380	1920	2370
1480	1935	2390	1930	2380
1490	1945	2400	1940	2390
1500	1955	2410	1950	2400
1510	1965	2420	1960	2410
1520	1975	2340	1970	2420
1530	1985	2440	1980	2430
1540	1995	2450	1990	2440
1550	2005	2460	2000	2450
1560	2105	2470	2010	2460
1570	2025	2480	2020	2470
1580	2035	2490	2030	2480
1590	2045	2500	2040	2490
1600	2055	2510	2050	2500
1610	2065	2520	2060	2510
1620	2075	2530	2070	2520
1630	2085	2540	2080	2530
1640	2095	2550	2090	2540
1650	2105	2540	2100	2550
1660	2115	2570	2110	2560
1670	2125	2580	2120	2570
1680	2135	2590	2130	2580
1690	2145	2600	2140	2590
1700	2155	2610	2150	2600

(continued)

TABLE 4 (continued)

Table of AM and FM band frequencies with common oscillator frequencies and some spurious responses. (concluded)

FM BAND

desired	oscillator	1/2 IF	image
87.9	98.6	93.25	109.3
88.1	98.8	93.45	109.5
88.3	99.0	93.65	109.7
88.5	99.2	93.85	109.9
88.7	99.4	94.05	110.1
88.9	99.6	94.25	110.3
89.1	99.8	94.45	110.5
89.3	100.0	94.65	110.7
89.5	100.2	94.85	110.9
89.7	100.4	95.05	111.1
89.9	100.6	95.25	111.3
90.1	100.8	95.45	111.5
90.3	101.0	95.65	111.7
90.5	101.2	95.85	111.9
90.7	101.4	96.05	112.1
90.9	101.6	96.25	112.3
91.1	101.8	96.45	112.5
91.3	102.0	96.65	112.7
91.5	102.2	96.85	112.9
91.7	102.4	97.05	113.1
91.9	102.6	97.25	113.3
92.1	102.8	97.45	113.5
92.3	103.0	97.65	113.7
92.5	103.2	97.85	113.9
92.7	103.4	98.05	114.1
92.9	103.6	98.25	114.3
93.1	103.8	98.45	114.5
93.3	104.0	96.85	114.7
93.5	104.2	98.85	114.9
93.7	104.4	99.05	115.1
93.9	104.6	99.25	115.3
94.1	104.8	99.45	115.5
94.3	105.0	99.65	115.7
94.5	105.2	99.85	115.9
94.7	105.4	100.05	116.1
94.9	105.6	100.25	116.3
95.1	105.8	100.45	116.5
95.3	106.0	100.65	116.7
95.5	106.2	100.85	116.9
95.7	106.4	101.65	117.1
95.9	106.6	101.25	117.3
96.1	106.8	101.45	117.5
96.3	107.0	101.65	117.7
96.5	107.2	101.85	117.9
96.7	107.4	102.05	118.1
96.9	107.6	102.25	118.3
97.1	107.8	102.45	118.5
97.3	108.0	102.65	118.7
97.5	108.2	102.85	118.9
97.7	108.4	103.05	119.1
97.9	108.6	103.25	119.3

FM BAND

desired	oscillator	1/2 IF	image
98.1	108.8	103.45	119.5
98.3	109.0	103.65	119.7
98.5	109.2	103.85	119.9
98.7	109.4	104.05	120.1
98.9	109.6	104.25	120.3
99.1	109.8	104.45	120.5
99.3	110.0	104.65	120.7
99.5	110.2	104.85	120.9
99.7	110.4	105.05	121.1
99.9	110.6	105.25	121.3
100.1	110.8	105.45	121.5
100.3	111.0	105.65	121.7
100.5	111.2	105.85	121.9
100.7	111.4	106.05	122.1
100.9	111.6	106.25	122.3
101.1	111.8	106.45	122.5
101.3	112.0	106.65	122.7
101.5	112.2	106.85	122.9
101.7	112.4	107.05	123.1
101.9	112.6	107.25	123.3
102.1	112.8	107.45	123.6
102.3	113.0	107.65	123.7
102.5	113.2	107.85	123.9
102.7	113.4	108.05	124.1
102.9	113.6	108.25	124.3
103.1	113.8	108.45	124.5
103.3	114.0	108.65	124.7
103.5	114.2	108.85	124.9
103.7	114.4	109.05	125.1
103.9	114.6	109.25	125.3
104.1	114.8	109.45	125.5
104.3	115.0	109.65	125.7
104.5	115.2	109.85	125.9
104.7	115.4	110.05	126.1
104.9	115.6	110.25	126.3
105.1	115.8	110.45	126.5
105.3	116.0	110.65	126.7
105.5	116,2	110.85	126.9
105.7	116.4	111.05	127.1
105.9	116.6	111.25	127.3
106.1	116.8	111.45	127.5
106.3	117.0	111.65	127.7
106.5	117.2	111.85	127.9
106.7	117.4	112.05	128.1
106.9	117.6	112.25	128.3
107.1	117.8	112.45	128.5
107.3	118.0	112.65	128.7
107.5	118.2	112.85	128.9
107.7	118.4	113.05	129.1
107.9	118.6	113.25	129.3

7.3 Color Television

Part I: The Basic NTSC System*

FUNDAMENTALS

Color is a dimension that has been added skillfully to black-and-white television. To the engineering community as a whole it signifies one of the most dramatic technological achievements of this age.

Nearly every branch of science, including chemistry and psychology, contributes in some way to the reality of color television. Through chemistry, improved phosphors are continually being found for use in color picture tubes. Psychology enters into the selection of lighting arrangements and picture composition to obtain desirable interpretations by the viewer. But physics plays the leading role with intense application in optics and illumination as well as in the design of electronic circuitry and components for the complete television system.

Two specialized branches of physics, namely radio and television engineering, are responsible for the electronic techniques which make color television compatible with black and white (monochrome) television, marking what is probably the greatest technical advance in television in the past decade.

Compatibility

The compatible color system offers tremendous economic advantages to the home viewer as well as to the television broadcaster. Because of compatibility, color telecasts can be seen (in monochrome) on black and white television receivers without any changes or added devices. Also, color receivers can receive monochrome as well as color telecasts. Since compatible color is transmitted over the same channels as monochrome and within the same framework of standards, the television broadcaster can utilize his monochrome system as the transmitting nucleus when installing equipment to broadcast color. Moreover, he can utilize his color equipment to produce monochrome telecasts.

Another important advantage of the compatible color system is the part it plays in the conservation of the radio-frequency spectrum. Compatible color requires no additional space in the spectrum. However, it employs techniques which make much more efficient use of the standards originally set up for monochrome television.

A brief review of the fundamentals of monochrome television, particularly the areas wherein specialized color methods are employed, is presented in the next few paragraphs as an aid in describing the basic color concepts.

Television: A System of Communications

Basically, television is a system of communications consisting of the television station at one end of the system and the television receiver at the other.

Very simply, the function of the television station is to divide and subdivide the optical image into over 200,000 picture elements, each of different light intensity; convert these light elements to electrical equivalents; and transmit them in orderly sequence over a radio-frequency carrier to the television receiver.

Reversing the process at the receiver, these electrical signals are each converted to light of corresponding brightness and reassembled to produce the transmitted image on the picture tube.

Scanning

Picture elements to be transmitted in sequence are selected by a process of image scanning which takes

* This section of the handbook has been reproduced through the courtesy of RCA Corporation who kindly permitted us to use protions of their "Color Television, Manual for Technical Training."

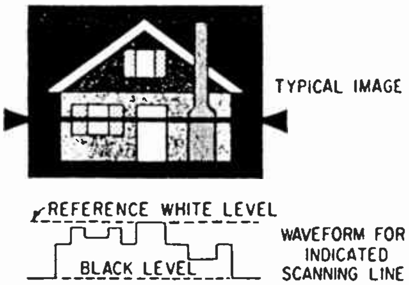


Figure 1. Typical image and camera output waveform produced by light and dark areas during one scan along line indicated by arrows.

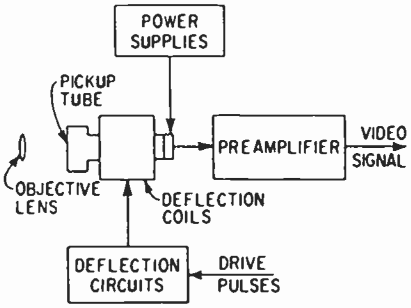


Figure 2. Block diagram of monochrome-camera circuits.

place in the television camera focused on the studio scene at the station. Within the camera, an electron beam in a pickup tube scans a sensitive surface containing an "electrical image" of the scene of action. The electron beam successively scans the image at great velocity, beginning at the upper left corner and continuing left to right in a series of parallel lines to scan the image completely. Movement of the electron beam, which can be controlled magnetically by vertical-and horizontal-deflection coils surrounding the tube, is analogous to that of the eye in reading a printed page. The speed of movement is such, however, that 30 complete image frames of approximately 500 lines each are scanned every second. Of course, at the receiver, an electron beam in the picture tube, moves with the same speed and in synchronism with the camera-tube beam so that the corresponding picture elements appear in the proper relative position on the television screen.

Owing to persistence of vision and the speed of scanning, these elements appear to be seen all at once as a complete image rather than individually. Thus, the impression is one of continuous illumination of the screen and direct vision.

Scanning standards have been established in this country to assure that all television receivers are capable of receiving programs broadcast by any television station within range. The scanning pattern adhered to by manufacturers in the design of television receivers and broadcast equipment consists of 525 lines with odd-line interlaced scanning. Interlaced scanning, effective in eliminating perceptible flicker, is a method whereby the electron beam scans alternate rather than successive lines. For example, the beam begins by scanning odd-numbered lines (1, 3, 5, 7, etc.) until it reaches the bottom of the image, whereupon it returns to the top of the image to scan the even-numbered lines (2, 4, 6, 8, etc.). Thus, each scan, or field, comprises only half of the total number of scanning lines, and two fields are required to produce the 525-line frame. Each field is completed in one-half the frame time. The vertical scanning frequency is 2 × 30 or 60 Hz, and horizontal scanning frequency is 30 × 525 or 15750 Hz.

Resolution and Bandwidth

The degree of resolution, or fine detail, that can be seen in a televised image depends upon the number of scanning lines used and the bandwidth of the transmitting and receiving system.

The relationship between resolution and bandwidth can be seen by considering the number of picture elements that can be transmitted each second.

The standard 6 MHz broadcast channel provides a video bandwidth of approximately 4.1 MHz (the remaining bandwidth being required for a vestigial sideband plus the sound signal). Since each cycle of a sine wave is capable of conveying two picture elements (one black and one white), the maximum rate at which picture elements can be transmitted is 4,100,000 × 2, or 8,200,000 per second. Since 30 complete frames are transmitted per second, the number of picture elements per frame would be 8,200,000 ÷ 30, or 273,333, if it were not for the retrace blanking problem, which requires interruption of the picture signal periodically by blanking pulses. Since the combination of horizontal and vertical blanking pulses requires nominally 25% of the total time, the maximum number of picture elements per frame is reduced in practice to 0.75 × 273,333, or approximately 205,000.

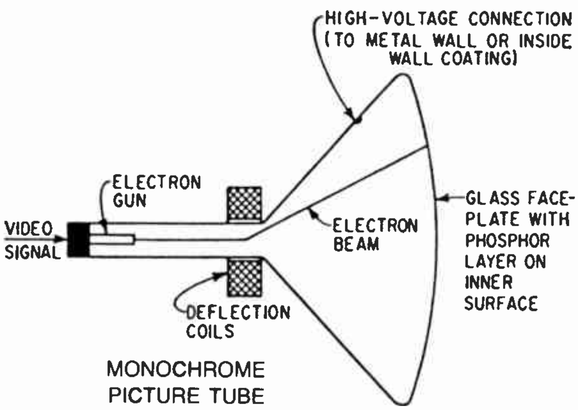


Figure 3. Diagram showing principal elements of the monochrome picture tube.

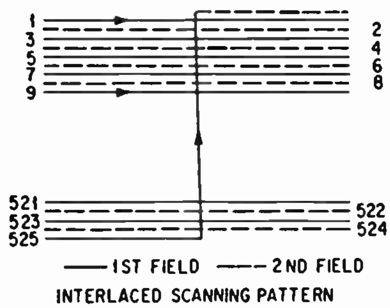


Figure 4. Diagram showing paths of the electron beam in both the pickup tube and kinescope to produce the interlaced scanning pattern.

Synchronizing

In addition to the picture information, or video signals, blanking and synchronizing signals are transmitted by the television station to control the intensity and movement of the scanning beam in the kinescope of the television receiver. Both these signals are in the form of rectangular pulses. Moreover, their polarity and amplitude are such that they are received as "black" signals and therefore do not appear on the receiver screen.

Blanking pulses eliminate the "retrace" lines which would otherwise appear between scanning lines and at the end of each field from the bottom of the picture to the top. Horizontal blanking pulses, transmitted at the end of each line, or at intervals of 1/15,750 sec, blank the beam during retrace periods between lines. Vertical blanking pulses, transmitted at the end of each field, or at intervals of 1/60 sec, blank the beam during the time required for its return to the top of the picture. Because the vertical retrace is much slower than the horizontal, the vertical blanking periods are longer than the horizontal blanking periods. Vertical blanking pulses are about 20 lines duration, while horizontal blanking pulses have a duration of only a small fraction of a line.

Synchronizing signals keep the scanning beam of the picture tube in step with that of the camera tube. These signals consist of horizontal and vertical pulses which are transmitted within the respective blanking periods. Although the sync pulses are of the same polarity as the blanking pulses, they are of greater amplitude ("blacker than black") and thus easily separated in the receiver and fed to the deflection circuits of the picture tube.

Since the vertical sync pulses are quite long compared with the horizontal sync pulses and the two are of the same amplitude, separation at the receiver is accomplished through frequency discrimination. Serrations, or slots in the vertical pulses, prevent loss of horizontal sync during the vertical blanking period.

The Monochrome Television System

The major equipment in a typical television station consists of the aural and visual units illustrated in the

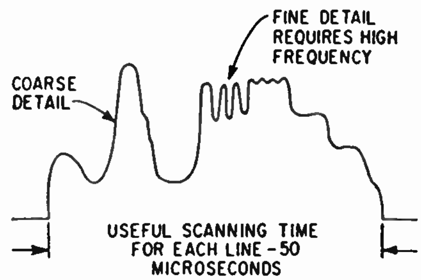


Figure 5. Diagram illustrating the relationship between picture detail and signal bandwidth.

block diagram of Fig. 6. In the visual channel, the video signal leaving the camera is passed through processing equipment which inserts the blanking and synchronizing signals and performs other functions such as aperture compensation and gamma correction. From the processing chain, the video signal is fed to a switching system which provides for selection from a number of video sources. The selected signal is then sent to the visual transmitter through coaxial cable or over a microwave relay link, depending upon the distance between the television studio and transmitter. In the transmitter, the composite video signal amplitude-modulates a carrier in the VHF or UHF range, which is radiated by the television antenna.

In the aural channel, the audio signal is fed from the microphone or other sound source through the switching system and to the aural transmitter. Frequency-modulated output from the aural transmitter is combined with the visual output and radiated from the same antenna.

The Radiated Picture Signal

Amplitude relationships between the synchronizing pulses and the tonal gradations from white to black in

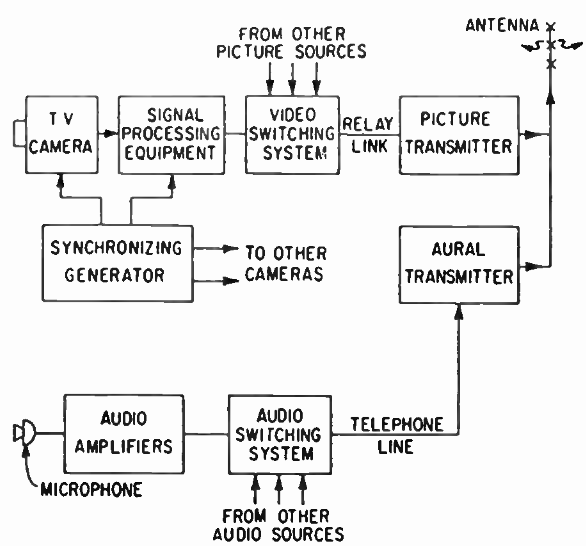


Figure 6. Simplified block diagram of the monochrome-television station.

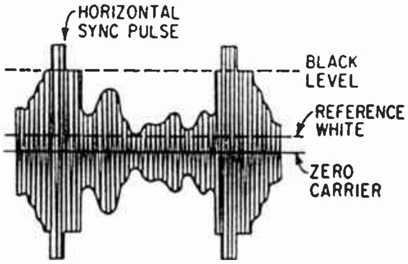


Figure 7. Waveform and radiated picture signal.

the picture are represented in the waveform of the radiated picture signal. From Fig. 7, it can be seen that modulation takes place in such a way that an increase in the brightness of the picture causes a decrease in carrier output power. Note that the reference-white line indicated on the sketch is relatively close to zero carrier level. Also, the synchronizing pulses are in the "blacker than black" region, representing maximum carrier power. Use of a widely different range of amplitude for the sync pulses makes it possible for home receivers to separate them by a simple clipping technique.

Receiver

The basic elements of the television receiving system are illustrated in Fig. 8. The radiated television signal is picked up by an antenna and fed to a tuner which selects the desired channel for viewing. Output from the tuner is passed through an intermediate-frequency amplifier which provides the major selectivity and voltage gain for the receiver. A second detector then recovers a video signal which is essentially the same as that fed to the visual transmitter.

The sound signal is usually taken off at the picture second detector in the form of a frequency-modulated beat between the picture and sound carriers. The sound signal is further amplified in a special intermediate frequency (IF) stage, detected by a discriminator or ratio detector, and applied to the speaker through an audio amplifier.

Picture output from the second detector is fed to two independent channels. One of these is the video amplifier which drives the electron beam in the picture tube and the other is the sync separator, or clipper, which separates the sync pulses from the picture information. The separated pulses are then used to control the timing of the horizontal and vertical deflection circuits. The high-voltage supply, which is closely associated with the horizontal deflection circuit, provides accelerating potential for the electron beam.

The Three Variables of Color

Color is the combination of those properties of light which control the visual sensations known as brightness, hue, and saturation. Brightness is that characteristic of a color which enables it to be placed in a scale ranging from black to white or from dark to light. Hue, the second variable of a color, is the characteristic which enables a color to be described as red, yellow, blue, or green. Saturation refers to the extent to which a color departs from white, or the "neutral" condition. Pale colors, or pastels, are low in saturation, while strong or vivid colors are high in saturation.

The monochrome system is limited to the transmission of images that vary with respect to brightness alone. Thus, brightness is the only attribute of a color which is transmitted over a monochrome-television system. To produce a color image, therefore, provision must be made for the transmission of additional information pertaining to all three of the variables of color. However, since the primary-color process can be employed, it is not necessary to transmit information in exactly the form expressed by the three variables.

Primary Colors in Television

Experiments have proved conclusively that virtually any color can be matched by the proper combination of no more than three primary colors. While other colors could be used as primaries, red, green, and blue have been selected as the most practical for color-television use. A few of the many colors that can be made by mixing lights of red, green, and blue are illustrated in Fig. 9. Red and green combined produce yellow, red plus blue give purple, and green plus blue gives cyan or blue-green. The proper combination of all three of the primary colors produces white, or neutral, as shown at the center of the illustration. By relatively simple optical means, it is possible to separate any color image into red, green, and blue, or RGB components, as shown by Fig. 10.

Generating RGB Signals

Major components of a color-television camera may have the block-diagram form shown in Fig. 11. Whereas the monochrome camera contains only one pickup

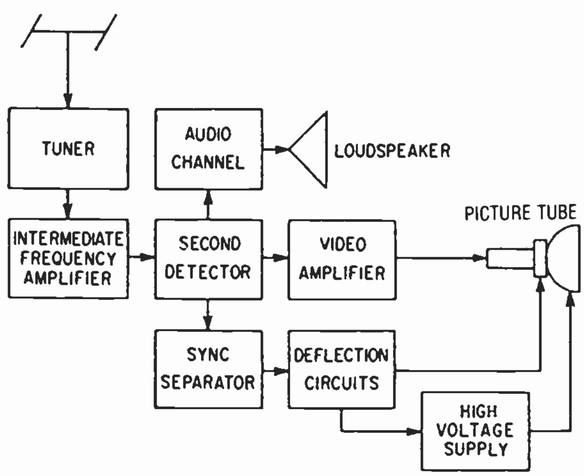


Figure 8. Block diagram of monochrome-television receiver.

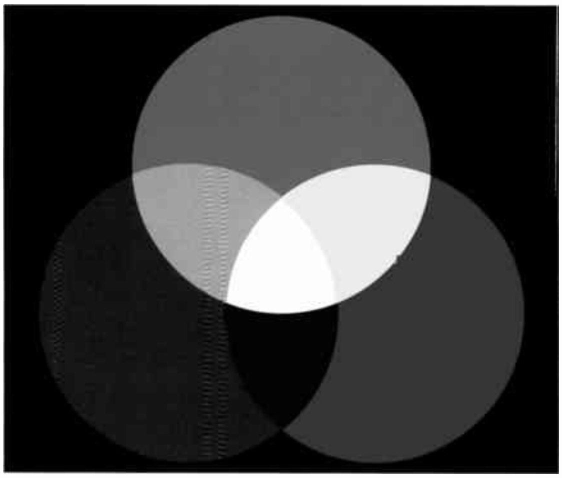


Figure 9. The primary colors of television are red, green, and blue. Virtually any color can be matched by combining proper amounts of these primaries. White is produced by a combination of all three.

Figure 10. Illustrating how a typical color image (upper left corner) can be separated by optical means into red, green, and blue counterparts.

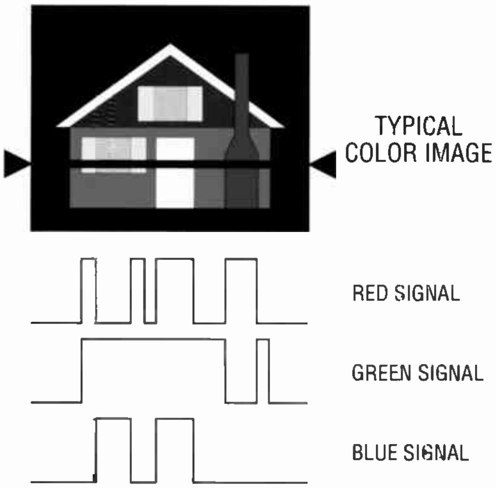


Figure 18. Typical color image and RGB waveforms.

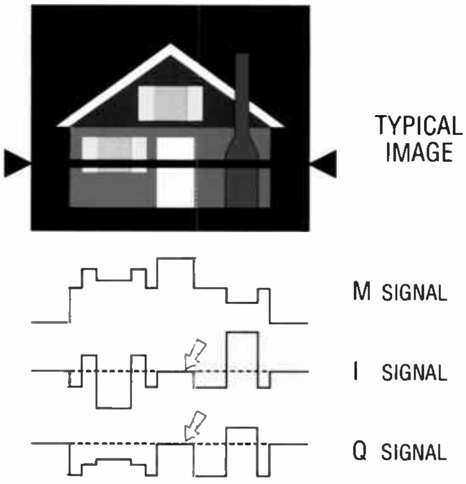


Figure 19. Typical color image and MIQ waveforms.

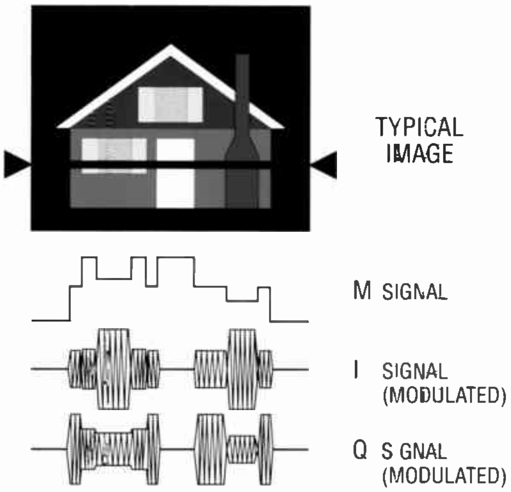


Figure 27. Typical color image and waveforms of the M signal and modulated I and Q signals.

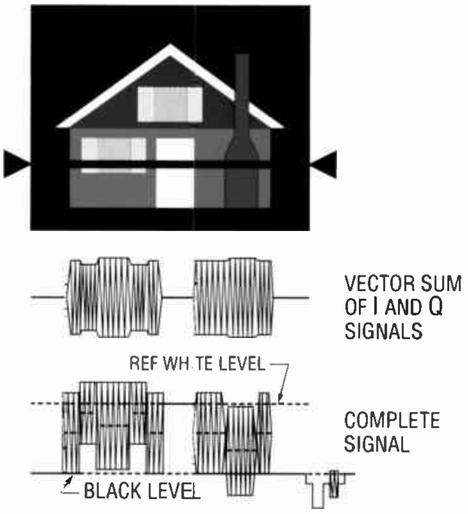


Figure 28. Typical color image.

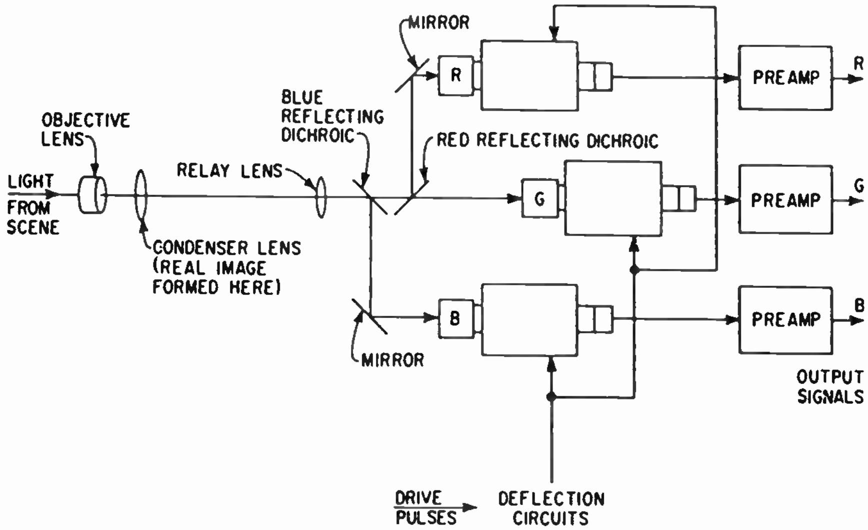


Figure 11. Simplified block diagram of the optical and electrical components of the color camera.

tube, or solid-state sensor, the color camera usually contains three separate pickup devices. An objective lens at the front of the camera forms a real image within a condensor lens which is located where the pickup device is usually mounted in a monochrome camera. A relay lens transfers this real image to a system of dichroic (color separating) mirrors or prisms which shunt the red and blue light to the red and blue pickup devices and permit the green to pass straight through to the green tube or sensor. In this manner, the three pickup devices produce three separate images corresponding to the RGB components of the original scene. These images are scanned in the conventional manner by common deflection circuits.

A single scanning line through the typical color image at the point shown (Fig. 18) produces three separate waveforms. It is important to note the correlation between these waveforms and the image at the top. The yellow shutters in the image, for example, must be produced by a mixture of red and green, and the blue signal is not required. Thus, at this interval of scanning the red and green signals are both at full value and the blue signal is at zero. The white door utilizes all three color signals. Of course, similar correlations can be seen for other parts of the image along the scanning line.

Displaying RGB Signals

RGB signals are displayed in color by the tricolor picture tube, the basic components of which are shown in the diagram of Fig. 12. Three electron guns produce three beams which are independently controlled in intensity by the red, green, and blue signals. These three beams are all made to scan in unison by deflection coils around the neck of the tube. The three beams converge at the screen owing to the magnetic field produced by a convergence yoke.

The phosphor screen of the color picture tube consists of an array of very small primary-color dots. Approximately ½″ behind the phosphor screen is an aperture mask which has one very small opening for each group of red, green, and blue phosphors. Alignment of this aperture mask and screen is such that each beam is permitted to strike phosphor dots of only one color. For example, all the electrons emitted by the red gun must strike red phosphor dots on the aperture mask; they cannot strike either the green or blue dots because of the "shadow" effect of the mask. Likewise, the beams emanating from the other two guns strike only green or blue dots.

In this way, three separate primary-color images are produced on the screen of the tricolor tube. But since these images are formed by closely intermingled dots too small to be resolved at the normal viewing distance, the observer sees a full-color image of the scene being televised.

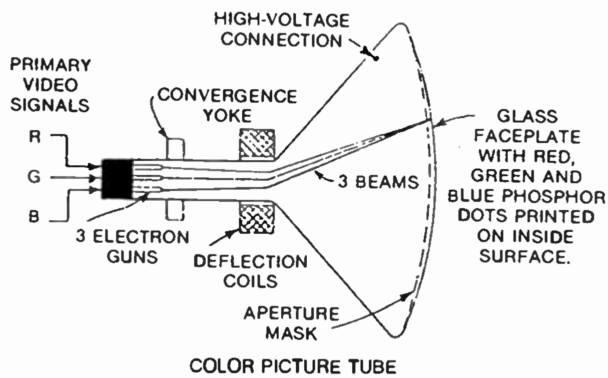


Figure 12. Diagram showing components of the three-gun picture tube.

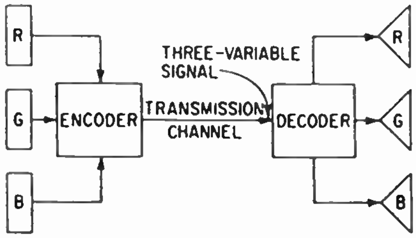


Figure 13. Encoding of the RGB signals provides a three-variable signal which can be transmitted over existing monochrome channels.

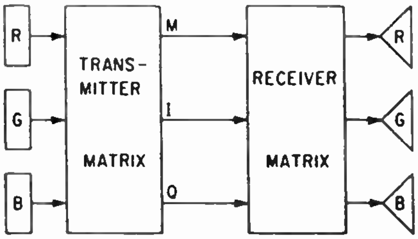


Figure 14. A part of the encoding process is the matrixing of R, G, and B signals to provide M, I, Q signals.

ELECTRONIC ASPECTS OF COMPATIBLE COLOR TELEVISION

To achieve compatibility with monochrome television, color television signals must be processed in such a way that they can be transmitted through the same channels used for monochrome signals, and they must also be capable of producing good monochrome pictures on monochrome receivers. Since color television involves three variables instead of the single variable (i.e., brightness) of monochrome television, an encoding process is required to permit all three to be transmitted over the one available channel. Likewise, a decoding process is required in the color receiver to recover the independent RGB signals for control of the electron guns in the color picture tube. Moreover, the process used must enable existing monochrome receivers to produce a monochrome picture from the color information.

Encoding and decoding processes used in compatible color television are based on four electronic techniques known as matrixing, band shaping, two-phase modulation, and frequency interlace. It is these processes which make the color system compatible with monochrome and enable the color system to occupy the existing 6 MHz channel.

Matrixing

Matrixing is a process for repackaging the information contained in the red, green, and blue output signals from a color camera to permit more efficient use of the transmission channel. The matrix circuits which perform this function consist of simple linear cross-mixing circuits. They produce these signals, commonly designated M, I, and Q, each of which is a different linear combination of the original red, green, and blue signals. Specific values for these signals have been established by FCC standards.

The M-signal component, or *luminance* signal, corresponds very closely to the signal produced by a monochrome camera, and therefore is capable of rendering excellent service to monochrome receivers. The M component is obtained by combining red, green, and blue signals in a simple resistor network (Fig. 15) designed to produce a signal consisting of 30% red, 59% green, and 11% blue.

The I and Q signals are *chrominance* signals which convey information as to how the colors in the scene differ from the monochrome, or "neutral" condition. The component I is defined as a signal consisting of 60% red, −28% green, and −32% blue. Minus values are easily achieved in the matrix circuits by use of phase inverters to reverse the signal polarity (see Figs. 16 and 17). The Q signal is defined as 21% red, −52% green, and 31% blue.

It can be seen that the quantities are related so that when red, green, and blue are equal, corresponding to the neutral condition, both I and Q go to zero. Thus, when the color camera is focused on an object having no color information, such as a monochrome test chart, the I-signal and Q-signal components are absent, leaving only the M component, or monochrome signal.

The matrix circuits, therefore, produce a new set of waveforms corresponding to the M, I, and Q components of the image. A comparison of the MIQ and RGB waveforms (Figs. 18 and 19) obtained from the image illustrates the correlation among the types of signals. It will be seen that the M signal remains in the region between black level and reference white. The I and Q signals, on the other hand, swing positive and negative around a zero axis.

Band Shaping

The eye has substantially less acuity in detecting variations in chrominance than it has for resolving differences in brightness. This important characteristic of human vision was considered in setting up the I and Q equations because it permitted a significant reduction

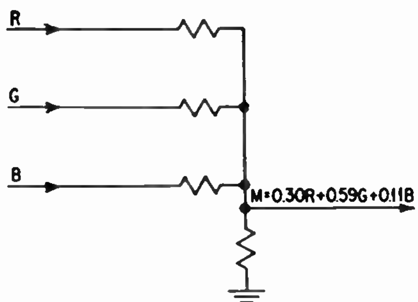


Figure 15. Diagram of resistance matrix circuit used to produce the M luminance signal.

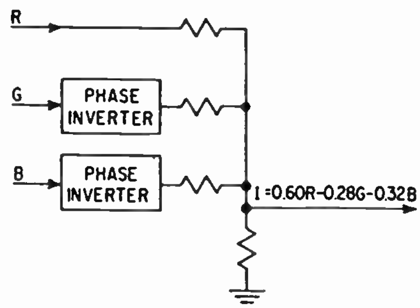


Figure 16. Diagram of I matrix showing phase inverters to produce minus green and blue quantities.

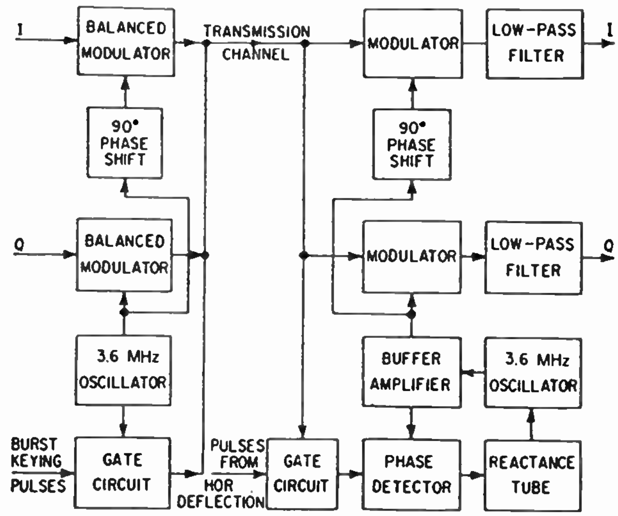


Figure 20. Simplified block diagram showing elements for transmitting and receiving the I, Q, and burst signals.

in the bandwidth of these signals through use of low-pass filters. A bandwidth of approximately 1.5 MHz was found to be satisfactory for the I signal, which corresponds to color transitions in the range extending from orange to blue-green. For color transitions in the range from green to purple, as represented by the Q signal, the eye has even less acuity and the bandwidth was restricted to only 0.5 MHz. The M-signal component, which conveys the fine details, must be transmitted with the standard 4 MHz bandwidth.

Two-Phase Modulation: Generation of Color Subcarrier

Two-phase modulation is a technique by which the I and Q signals can be combined into a two-variable signal for transmission over a single channel. This is accomplished by adding the sidebands obtained through modulation of two 3.6 MHz carriers separated in phase by 90. The resultant waveform is the vector sum of the components. Elements of the transmitting and receiving system are shown in Fig. 20. The two carriers, which are derived from the same oscillator, are suppressed by the balanced modulators. Thus, only the two amplitude-modulated sidebands, 90 out of phase, are transmitted. At the receiving end of the system, the I and Q signals are recovered by heterodyning the two-phase wave against two locally generated carriers of the same frequency but with a 90 phase separation and applying the resultant signals through low-pass filters to the matrix circuits. Typical signal waveforms are illustrated in Fig. 21.

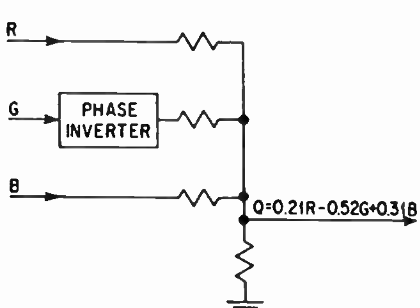


Figure 17. Diagram of the Q matrix showing phase inverter to produce required minus green signal.

The 3.6 MHz oscillator at the receiver must be accurately synchronized in frequency and in phase with the master oscillator at the transmitter. The synchronizing information consists of 3.6 MHz, "bursts" of at least 8 Hz duration transmitted during the "back-porch" interval following each horizontal sync pulse. The bursts are generated at the transmitter by a gating circuit which is turned on by burst keying pulses derived from the synchronizing generator. At the receiver, the two-phase modulated signal is applied to another gating circuit, known as a burst separator, which is keyed on by pulses derived from the horizontal deflection circuit. The separated bursts are compared in a phase detector with the output of the local 3.6 MHz oscillator. Any error voltage developed is applied through a smoothing filter to a conventional reactance tube or varactor which corrects the phase of the local oscillator.

FCC Standard phase relationships between the I and Q signals and the color synchronizing burst are shown

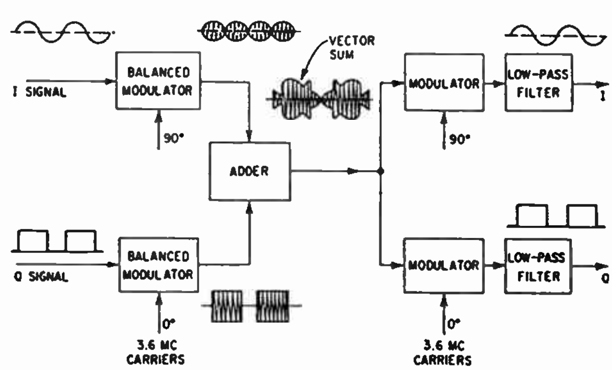


Figure 21. Representative waveforms of the separate I, Q signals and the vector sum of the suppressed carrier sidebands at the modulator output. Original I and Q signals are recovered by heterodyning in balanced modulators at receiver.

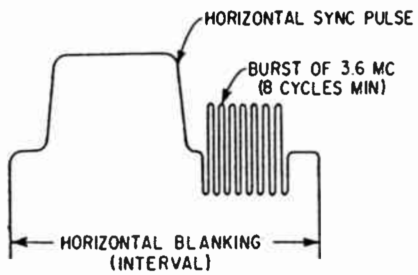


Figure 22. Diagram showing position of subcarrier burst during horizontal blanking interval.

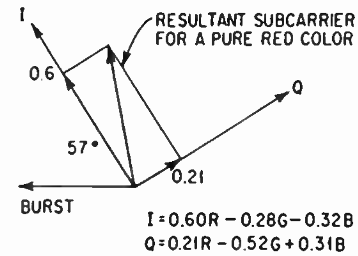


Figure 24. Vector diagram showing phase and amplitude of subcarrier for a pure red signal.

in the vector diagram of Fig. 23. The I and Q signals are transmitted in phase quadrature, and the color burst is transmitted with an arbitrary 57 phase lead over the I signal.

Several interesting properties of the two-phase modulated signal are illustrated by the vector diagrams which represent the resultant signal under known transmission conditions. For example, when a pure red color of maximum amplitude is being transmitted, the green and blue components are at zero and the I and Q signals have levels of 60% and 21%, respectively. When modulated upon their respective carrier, these signals produce the resultant shown in Fig. 24. The phase and amplitude shown are characteristic of pure red of maximum relative luminance. Fig. 25 is a composite vector diagram showing the phase and amplitude characteristics of the three primaries and their complementary colors. This composite diagram indicates that there is a direct relationship between the *phase* of the resultant two-phase modulated signal and the *hue* of the color being transmitted. There is also a relationship (although indirect) between the *amplitude* of the resultant signal and the saturation of the color being transmitted. If the phase of the resultant subcarrier and the level of the monochrome signal both remain constant, then a reduction in the amplitude of the subcarrier indicates a decrease in color saturation. The composite vector diagram also shows an interesting symmetry between complementary colors (colors are complementary if they produce a neutral when added together); the resultants for any two complementary colors are equal in amplitude but opposite in phase.

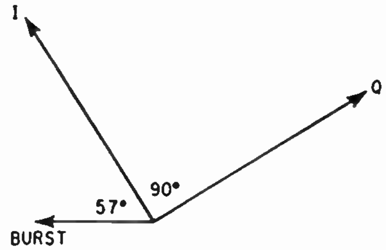


Figure 23. Diagram showing phase relationship of I, Q, and burst signals.

Frequency Interlace

Since the 3.6 MHz carriers, consisting of the I and Q sidebands, fall within the video passband as shown in the diagram of the television channel (Fig. 26), they become subcarriers and can be handled in many respects like unmodulated video signals. By use of *frequency interlace* it is possible to add the several components of the chrominance and monochrome signals together without causing objectionable mutual interference.

The significance of the straightforward addition of signal components made possible by frequency interlace may be brought out by a study of waveforms derived from a simple color image. Fig. 27 shows M, I, and Q signals after the latter two have been modulated upon 3.6 MHz subcarriers. Note that both the I- and Q-signal components are at zero during the scanning of the white door, a neutral area. Fig. 28 shows the vector sum of the I and Q signals and also the complete compatible color signal formed by adding together all the components, including synchronizing pulses and color-synchronizing bursts. The most significant fact about this signal is that it is still capable of providing good service to monochrome receivers, even though a modulated wave has been added to the monochrome-signal component. Although the modulated wave is clearly a spurious signal with respect to the operation of the kinescope in a monochrome receiver, its interference effects are not objectionable because of the application of the frequency-interlace principle.

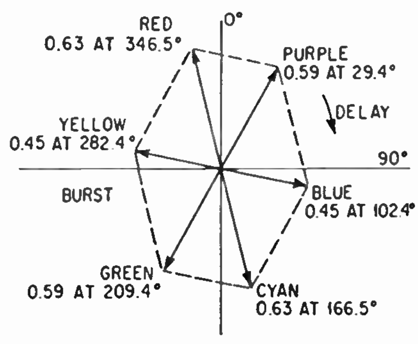


Figure 25. Composite vector diagram showing subcarrier phase and amplitude for each of six colors.

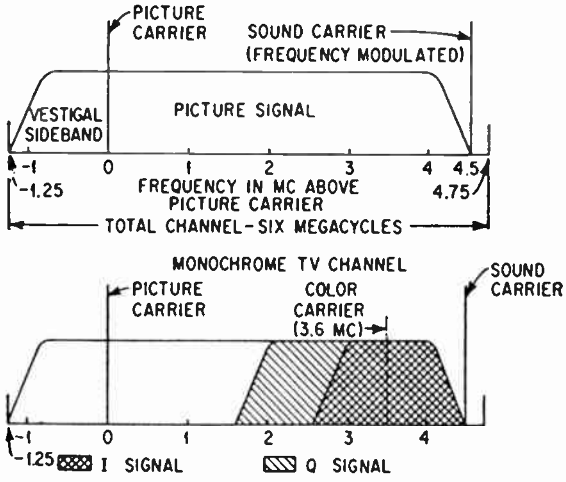


Figure 26. Diagram of television channel showing portions occupied by color and monochrome signal components.

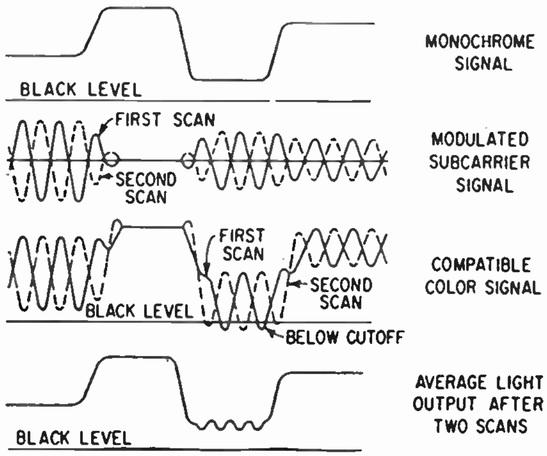


Figure 29. Waveforms showing superposition of modulated subcarrier on scanning signals, compatible color signal, and effect of subcarrier on average light output.

The frequency-interlace technique is based on two factors: a precise choice of the color subcarrier frequency and the familiar persistence-of-vision effect. If the color subcarrier is made an *odd multiple of one-half the line frequency*, its apparent polarity can be made to reverse between successive scans of the same area in the picture. Since the eye responds to the average stimulation after two or more scans, the interference effect of the color subcarrier tends to be self-canceling, owing to the periodic polarity reversals (see Fig. 29).

Color-Frequency Standards

The relationships among the various frequencies used in a compatible color system are illustrated in the block diagram of Fig. 30. The actual frequency of the color subcarrier, which has been referred to as 3.6 MHz is specified by FCC Standards as 3.579545 MHz or exactly 455 multiplied by ½ the line frequency.

In broadcast practice, the frequency of the color subcarrier provides a frequency standard for operation of the entire system. A crystal oscillator at the specified frequency provides the basic control information for all other frequencies. County stages and multipliers derive the basic frequencies needed in the color studio. A frequency of nominally 31.5 kHz required for the equalizing pulses which precede and follow each vertical sync pulse and for the serrations in the vertical sync pulse. A divide-by-2 counter controlled by the 31.5 kHz signal provides the line-frequency pulses at nominally 15.75 kHz needed to control the horizontal blanking and synchronizing waveforms. Another counter chain provides the 60 Hz pulses needed for control of the vertical blanking and synchronizing circuits.

The Overall Color System

The major functions performed in transmitting and receiving color are shown in the overall block diagrams of the transmitting and receiving systems (Fig. 31 and 32).

At the transmitting end, camera output signals corresponding to the red, green, and blue components of the scene being televised are passed through nonlinear amplifiers (the gamma correctors) which compensate for the nonlinearity of the picture tube elements at the receiving end. Gamma-corrected signals are then matrixed to produce the luminance signal M and two chrominance signals I and Q. The filter section establishes the bandwidth of these signals. The 4.1 MHz filter for the luminance channel is shown in dotted lines because in practice this band shaping is usually achieved by the attenuation characteristics of the transmitter and the filter is not required.

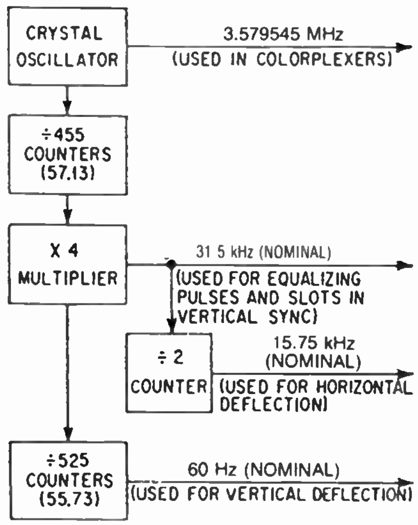


Figure 30. Block diagram showing relationship between various frequencies used in color television station.

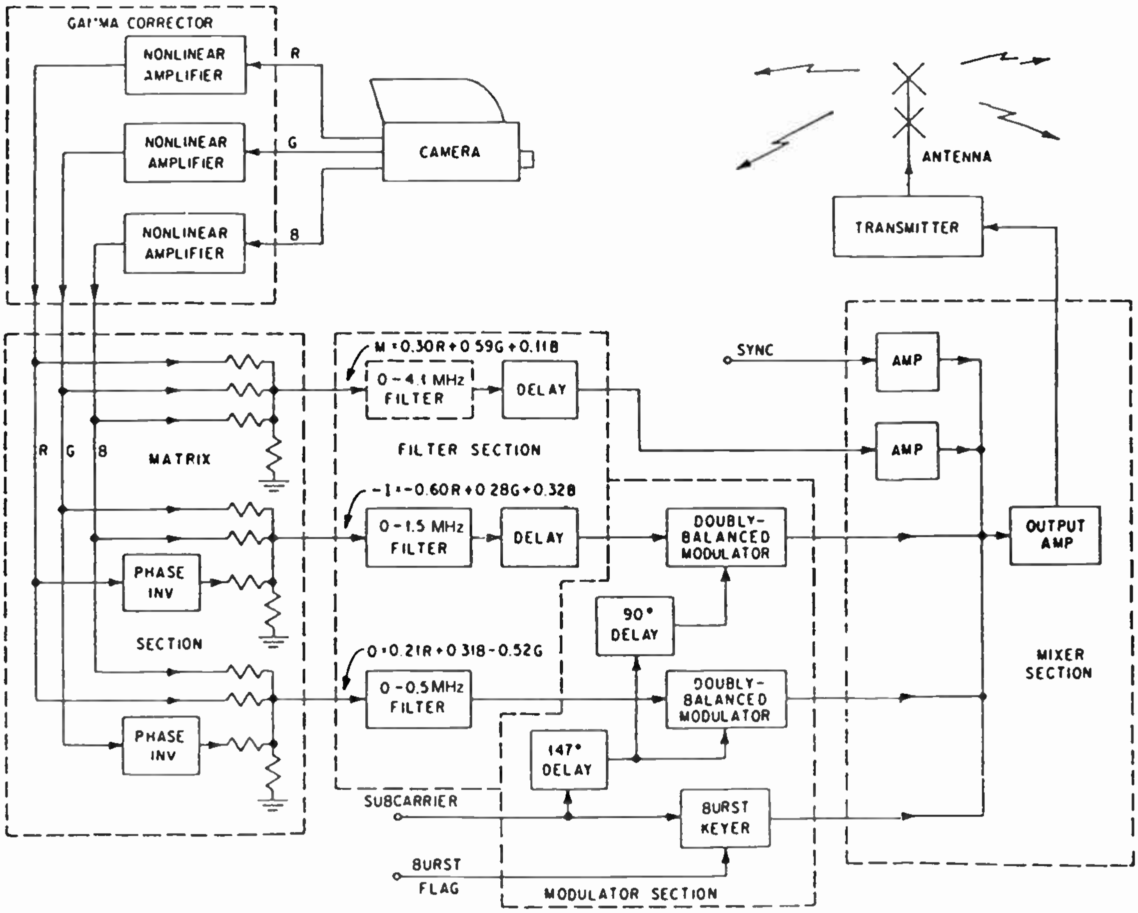


Figure 31. Block diagram showing major functions of color-transmitting system.

The bandwidths of 1.5 MHz and 0.5 MHz known for the I and Q channels, respectively, are nominal only; the required frequency-response characteristics are described in more detail in the complete FCC signal specifications. Delay compensation is needed in the filter section in order to permit all signal components to be transmitted in time coincidence. In general, the delay time for relatively simple filter circuits varies inversely with the bandwidth. The narrower the bandwidth, the greater the delay. Consequently, a delay network or a length of delay cable must be inserted in the I channel to provide the same delay introduced by the narrower band filter in the Q channel, and still more delay must be inserted in the M channel.

In the modulator section, the I and Q signals are modulated upon two subcarriers of the same frequency but 90 apart in phase. The modulators employed should be of double balanced type, so that both the carriers and the original I and Q signals are suppressed, leaving only the sidebands. Some sort of keying circuit must be provided to produce the color-synchronizing bursts during the horizontal blanking intervals. To comply with the FCC signal specifications, the phase of the burst should be 57 ahead of the I component (which leads the Q component by 90). This phase position was chosen mainly because it permits certain simplifications in receiver designs. Timing information for "keying in" the burst can be obtained from a "burst flag generator," which is a simple arrangement of multivibrators controlled by horizontal and vertical drive pulses.

In the mixer section, the M signal, the two subcarriers modulated by I and Q chrominance signals, and the color-synchronizing bursts are all added together. Provision is also made for the addition of standard synchronizing pulses, so that the output of mixer section is a complete color television signal containing both picture and synchronizing information. This signal can then be put on-the-air by means of a standard television transmitter, which must be modified only to the extent necessary to assure performance within the reduced tolerance limits required by the color signal. (Since the color signal places more information in the channel than a black-and-white signal, the requirements for frequency response, amplitude linearity, and uniformity of delay time are stricter).

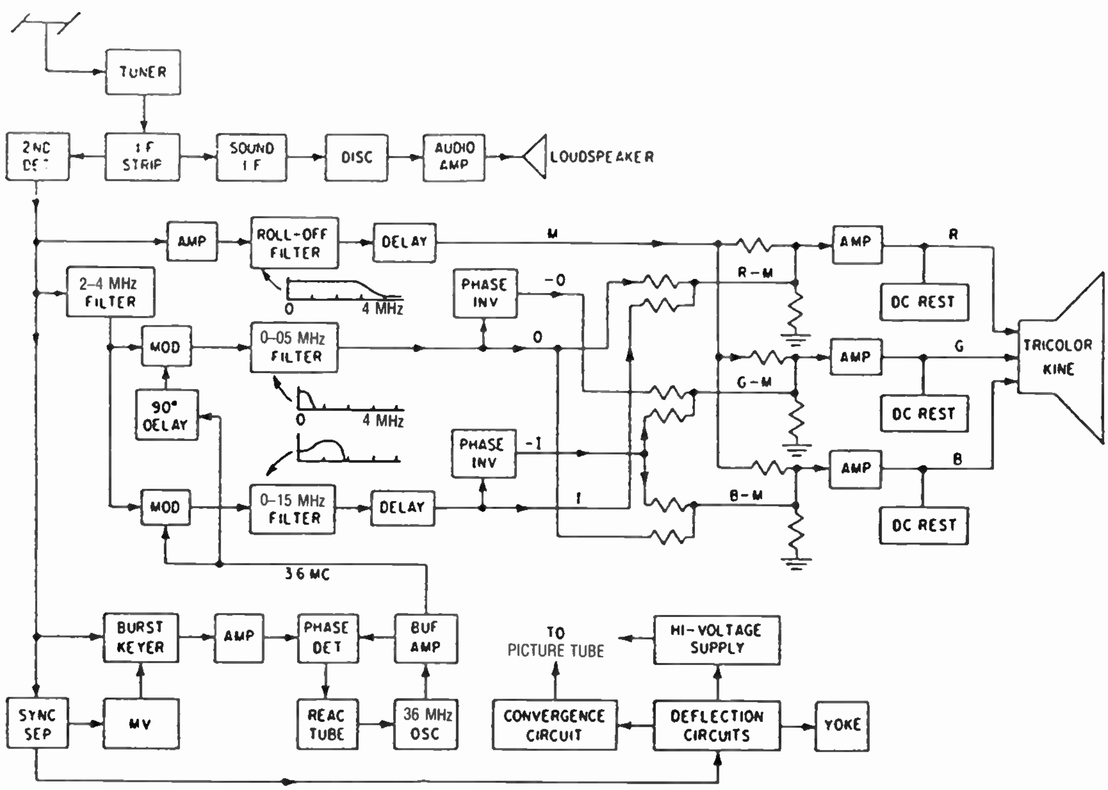


Figure 32. Block diagram showing major functions of color-receiving systems.

The Color Receiving System

In a compatible color receiver, the antenna, RF tuner, IF strip, and second detector serve the same functions as the corresponding components of a black-and-white receiver except that the tolerance limits on performance are somewhat tighter.

The signal from the second detector is utilized in four circuit branches. One circuit branch directs the complete signal toward the color kinescope, where it is used to control luminance by being applied to all kinescope guns in equal proportions. In the second circuit branch, a bandpass filter separates the high-frequency components of the signal (roughly 2.0 MHz to 4.1 MHz) consisting mainly of the two-phase modulated subcarrier signal. This signal is applied to a pair of modulators which operate as synchronous detectors to recover the original I and Q signals. It should be noted that those frequency components of the luminance signal falling between about 2 MHz and 4.1 MHz are also applied to the modulators and are heterodyned down to lower frequencies. These frequency components do not cause objectionable interference, however, because they are frequency-interlaced and tend to cancel out through persistence of vision.

The remaining two circuit branches at the output of the second detector make use of the timing or synchronizing information in the signal. A conventional sync separator is used to produce the pulses needed to control the horizontal- and vertical-deflection circuits which are also conventional. The high-voltage supply for the kinescope can be obtained either from a "flyback" supply associated with the horizontal deflection circuit or from an independent RF power supply. Many color kinescopes require convergence signals to enable the scanning beams to coincide at the screen in all parts of the picture area; the waveforms required for this purpose are readily derived from the deflection circuits.

The final branch at the output of the second detector is the burst gate, which is turned on only for a brief interval following each horizontal sync pulse by means of a keying pulse. This pulse may be derived from a multivibrator controlled by sync pulses, as illustrated in Fig. 32, or it may be derived from the "flyback" pulse produced by the horizontal output stage. The separated bursts are amplified and compared with the output of a local oscillator in a phase detector. If there is a phase difference between the local signal and the bursts, an error voltage is developed by the phase detector. This error voltage restores the oscillator to the correct phase by means of a reactance tube or varactor connected in parallel with the tuned circuit of the oscillator. This automatic-frequency control circuit keeps the receiver oscillator in synchronism with the master subcarrier oscillator at the transmitter. The output of the oscillator provides the reference carriers for the two synchronous detectors; a 90 phase shifter is necessary to delay the phase of the Q modulator by 90 relative to the I modulator.

There is a "filter section" in a color receiver that is rather similar to the filter section of the transmitting equipment. The M, I, and Q signals must all be passed through filters in order to separate the desired signals from other frequency components which, if unimpeded, might cause spurious effects. The I and Q signals are passed through filters of nominally 1.5 MHz and 0.5 MHz bandwidth, respectively, just as at the transmitting end. A step-type characteristic is theoretically required for the I filter, as indicated in Fig. 26, to compensate for the loss of one sideband for all frequency components above about 0.5 MHz. Actually, this requirement is ignored in many practical receiver designs, resulting in only a slight loss in sharpness in the I channel. A roll-off filter is desirable in the M channel to attenuate the subcarrier signal before it reaches the kinescope. The subcarrier would tend to dilute the colors on the screen if it were permitted to appear on the kinescope grids at full amplitude. Delay networks are needed to compensate for the different inherent delays of the three filters, as explained previously.

Following the filter section in the receiver there is a matrix section in which the M, I, and Q signals are cross-mixed to recreate the original R, G, and B signals. The R, G, and B signals at the receiver are not identical with those at the transmitter because the higher frequency components are mixed and are common to all three channels. This mixing is justifiable because the eye cannot perceive the fine detail (conveyed by the high-frequency components) in color. There are many possible types of matrixing circuits. The resistance mixers shown provide one simple and reliable approach. For ease of analysis, the matrix operations at the receiver can be considered in two stages. The I and Q signals are first cross-mixed to produce R-M, G-M, and B-M signals (note that *negative* I and Q signals are required in some cases), which are, in turn, added to M to produce R, G, and B.

In the output section of the receiver, the signals are amplified to the level necessary to drive the kinescope and the dc component is restored. The image which appears on the color kinescope screen is a high-quality full-color image of the scene before the color camera.

It should be made clear that the block diagram of Fig. 32 is intended only to illustrate the principles used in color receivers and does not represent any specific model now on the market. Design engineers of color receivers have shown great ingenuity in simplifying circuits, in combining functions, and in devising subtle variations in the basic process which have made possible significant cost reductions while maintaining excellent picture fidelity. The principles of compatible color television are firmly established, and it is to be expected that steady progress will be made in the practical application of those principles.

COLOR FIDELITY

Color fidelity as used herein, is the property of a color-television system to reproduce colors which are realistic and pleasing to the average viewer.

Although perhaps not apparent at first, color fidelity is analogous to high fidelity as applied to sound reproduction. Just as a high-fidelity audio system faithfully reproduces sounds reaching the microphone, the color-television system is capable of faithfully reproducing colors as seen by the television cameraperson. In fact, the color television system is capable of reproducing colors more accurately than techniques presently used in color printing and color photography.

Tests have shown, however, that color television pictures are generally more pleasing to viewers when deliberate modifications are made in the reproduced colors to compensate for the surroundings in which they are reproduced. The situation is similar to that experienced in the art of sound reproduction in the case of a symphony orchestra recorded at high sound levels in a large hall and reproduced at lower sound levels in a small room. In this case, a more pleasing effect is obtained if the ear's new environment is taken into consideration and the reproduction modified accordingly. Similarly, in color television, the changed environment of the eye must be considered and the reproduced colors modified accordingly.

Color fidelity, therefore, is a term used to indicate a color reproduction which pleases viewers aesthetically and persuades them that they are viewing a faithful reproduction of the original colors in the scene being televised.[1]

The following describes possible distortions in the color system and their effect on the picture and prescribes amounts or degrees of distortion that can be tolerated without adverse effects on picture quality.

Color System Analysis

Individual elements or areas of the complete color system are discussed in the following paragraphs with the aid of the diagrams shown in Fig. 33 through 37.

Fig. 33 is a theoretical color system in that it assumes linear camera tubes and a linear picture tube interconnected by a distortionless wire system. The only distortion that can result from this system is a flaw in colorimetry.

Fig. 34 introduces linearity correctors to compensate for color errors produced by nonlinearities in the transducers.

Fig. 35, 36, and 37 successively introduce the complexities of matrixing, band limiting, delay compensation, and the transmission system (shown dotted in Fig. 37). These diagrams, each representing a possible color system, introduce techniques used in compatible color television and permit the study of color distortions peculiar to each technique.

The systems diagrammed in Fig. 33 and 34 are described under "Possible Distortions in Transducers," and those in Fig. 35, 36, and 37 under "Possible Distortions in Encoding and Decoding Processes." The system shown in Fig. 37 is discussed under "Distortions in the Transmission System."

Characteristics of the Eye

To appreciate fully the significance of color fidelity, it is helpful to consider some of the characteristics of

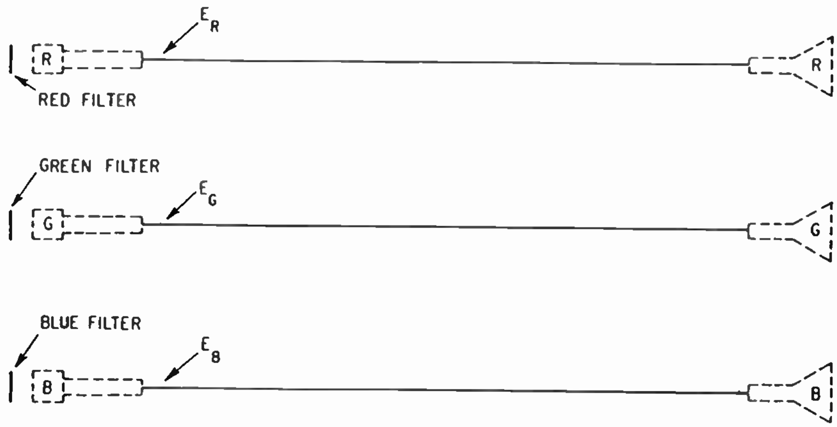


Figure 33. Diagram of a theoretical color system showing linear RGB pickup tubes and picture tubes interconnected by wire.

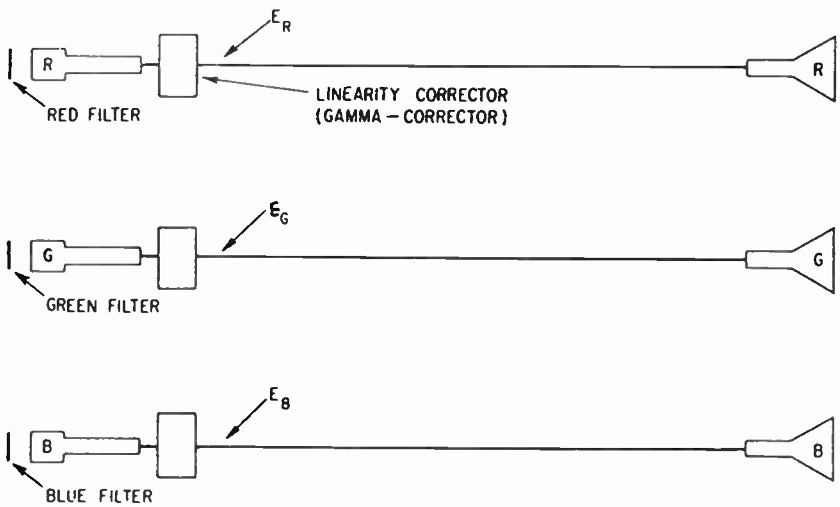


Figure 34. The basic color system shown with necessary linearity correctors to compensate for color errors introduced by the nonlinear transducers.

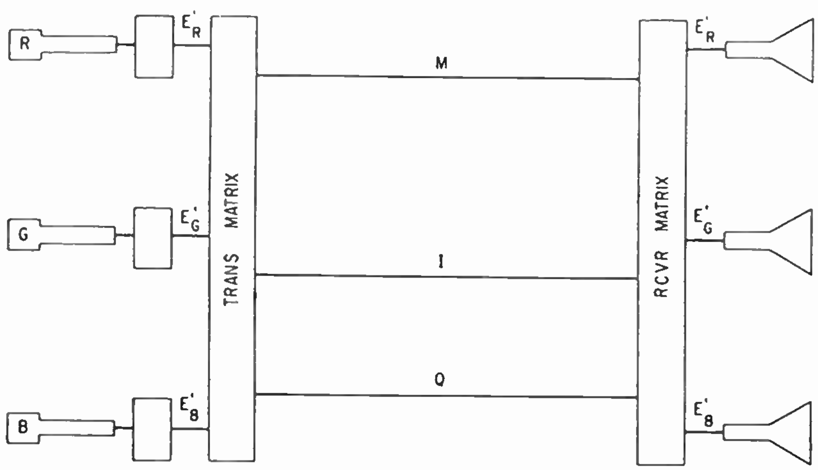


Figure 35. Diagram showing transmitter and receiver matrix functions in the color system.

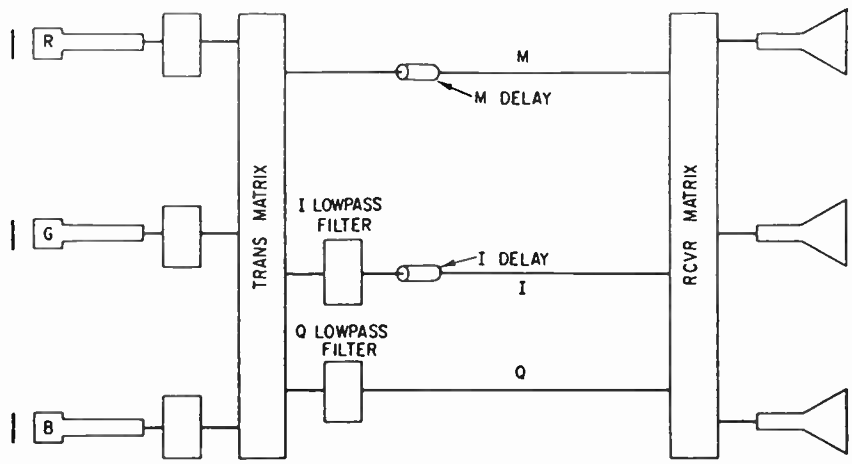


Figure 36. Basic color system with band limiting and delay compensation.

the eye associated with color perception and to analyze such terms as color adaptation, reference white, and primary colors and determine their relationship to a color-television system.

Color Adaptation

One amazing characteristic of the eye is the phenomenon known as color adaptation. It is this adaptation which enables one to describe accurately the color of an object under "white" light while viewing in nonwhite light. That is to say, recognition of color is surprisingly independent of the illumination under which an object is viewed. For example, if sunlight at high noon on a cloudless day is taken as "white" light, then, by comparison, the illumination from a typical 100-watt incandescent bulb is very yellow light. Yet it is known that an object viewed under sunlight looks very little if any different when viewed under incandescent light. Moreover, it is obvious to the observer, after a very few minutes in a room illuminated with incandescent lights, that the light is not yellow at all; it is really "white."

It is apparent, then, that the color seen by an observer is dependent upon the illumination to which that observer has been exposed for the past several minutes. This ambient illumination will have a marked effect on the choice of color to be called "white."

This phenomenon can cause a loss of color fidelity under certain conditions. Consider, for example, a theoretically perfect color system with the camera viewing an outdoor scene under a midday sun while the reproduced picture is being viewed in a semidarkened room, with what little light is in the room also being derived from the midday sun. Under these conditions, the ambient illuminations at both camera and receiver are identical, so a man standing alongside the camera and a man viewing the receiver would both see the same colors. Now, if a change in the weather at the

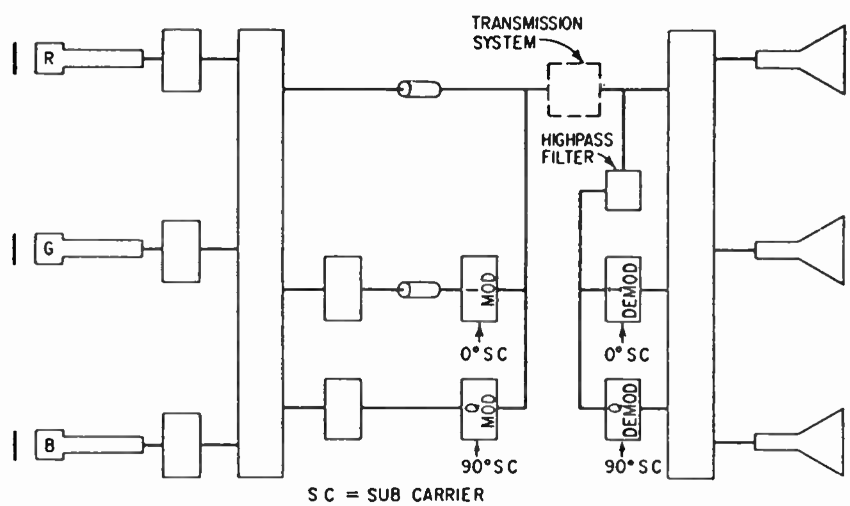


Figure 37. Basic color system showing all major elements, including the transmission system.

camera location should cause a cloud to cover the sun, the ambient illumination at the camera location would shift toward a bluer color. This shift would not disturb the viewer standing alongside the camera, because his eyes, bathed in the new ambient light, would rapidly adapt to the new viewing conditions and he would perceive the scene as being unchanged.

The man viewing the receiver would not be so fortunate. Assuming that he is far enough away that this same cloud would not affect his ambient, he would observe that everything on his screen had suddenly and inexplicably taken on a bluish cast, which he would certainly find most disturbing.

Such errors in color fidelity can be corrected by making the camera imitate the human eye in adaptation. The eye adapts to changes in ambient illumination by changing its sensitivity to a certain color. For example, if a light source changes from white to blue-white (as in the above example), the eye reduces its blue sensitivity until the light again appears to be white to the observer. Likewise, a camera operator can correct for the same situation by decreasing the gain of the blue channel of the camera or by attenuating the light reaching the blue camera tube. In this way, the camera is made to "color adapt," and the reproduced picture on a receiver loses in bluish cast.

Reference White

Although color adaptation can generate a problem such as the one just described, it also simplifies certain requirements. Specifically, it eases the requirement that white be transmitted as a definite, absolute color, for there clearly can be no absolute white when almost any color can be made to appear subjectively white by making it the color of the ambient illumination to which an observer's eye has adapted.

In color television, we take advantage of this characteristic in the following manner: A surface in the studio which is known by common experience to be white, for example, the Electronics Industry Association (EIA) Gray-Scale Chart of a piece of Neutracor white paper, is selected to be reproduced as white on a home receiver. The relative sensitivities of the three-color channels of the camera are then adjusted so that the camera adapts to this white regardless of the studio illumination. The home receiver can then be adjusted to reproduce the surface as any "white" which the home viewer prefers, depending upon his surroundings.

It has already been mentioned that the eye adapts readily to the illumination that surround conditions of an overcast day. This representative standard illumination has been adopted internationally as a base for the specification of the color of objects when they are viewed outdoors. This standard (Illuminant C) has been chosen to be the "standard-viewing-white" of the receiver. A slightly different illuminant (Illuminant D) has been proposed as more accurately representative of outdoor illumination and may replace Illuminant C in the near future.

The change in reference white between studio and home will inevitably produce errors in all reproduced colors, but the errors are small and, more important, tend to be subjectively self-correcting, so that any given object will produce the same color sensation whether viewed in relation to the studio reference white or the home reference white.

Consequently, a viewer may become familiar with an object such as a sponsor's packaged product and will recognize it on his television screen, under the fluorescent lighting of his supermarket, or under the incandescent lighting of his home and, furthermore, will note little difference in the colorimetric values of the package under the three conditions, even though the absolute colorimetric values would be appreciably different in the three situations.

Primary Colors

Of all the characteristics of the eye, there is perhaps none more fundamental to practical color television than that characteristic which allows us to choose certain colors called primary colors, and from these synthesize almost any other desired color by adding together the proper proportions of the primary colors. If it were noted for this characteristic, each hue in a color system would have to be transmitted over a separate channel; such a system would be too awkward to be practical. Because of the eye's acceptance of synthesized colors, it is possible to provide excellent color rendition by transmitting only the three primary colors in their proper proportions.

Possible Errors in Transducers

The block diagram of Fig. 33 shows a fundamental color-television system using red, green, and blue primaries and three independent transmission channels. The camera tubes and kinescopes are shown dotted to indicate that any inherent nonlinearities in these devices are to be disregarded, for the moment, in order to simplify the discussion of the colorimetry of the system.

The general plan is a system, such as Fig. 33, to provide the three picture tubes with red, green, and blue phosphors, respectively, and to allow the corresponding camera tubes to view the scene through an appropriate set of red, green, and blue filters. If a phosphor and a filter have the same dominant wavelength, that is, if they appear to the eye to be the same color, it might be mistakenly supposed that they would be colorimetrically suited to be used as a filter and phosphor set for the channel handling that color. Actually, the basis for choosing filters and phosphors is much more complex and is based on the shape of the response curve of the filter, plotted against wavelength, and the shape of the light-output curve of the phosphor, also plotted against wavelength. The following paragraphs will discuss briefly a technique which might be used to determine the required relationship between the phosphor curves and the filter curves.

The color characteristics of the phosphor are generally less easily changed than are filter characteristics; for this reason characteristics of phosphors are taken as the starting point, and characteristics of the filters are determined from them. A laboratory setup which could be used to determine these characteristics is shown in Fig. 38. In this figure, an observer (who must have "normal" vision) is viewing simultaneously two adjacent areas, one of which is illuminated by a source of single-wavelength light which can select any wavelength in the visible spectrum, the other of which is illuminated by a red picture tube, a green picture tube, and a blue picture tube. The phosphors of these kinescopes are the phosphors which are to be used in the color system. Starting at, say, the red end of the spectrum, a single-wavelength red is selected to illuminate the left-hand area, and the light from each of the three phosphors is varied until a color match is obtained between the left-hand and right-hand areas. The respective amounts of red, green, and blue lights needed to accomplish this match are recorded. Then another wavelength is chosen, the kinescope outputs varied to produce a match, and the new amounts of red, green and blue needed for a match are recorded. Similarly, points are obtained throughout the entire spectrum, and a graph is plotted showing the various required outputs versus wavelength. The shapes of these three curves (one for red, one for green, and one for blue) are the required shapes for the three camera-filter response curves. The resulting curves would in general resemble Fig. 39.

(To simplify the above discussion it was assumed that the camera tubes responded equally well to all wavelengths. In practice, camera tubes show higher output at certain wavelengths than at others. The filter-response curves derived by the above technique would have to be modified so that the combined response of filter and camera would be correct.)

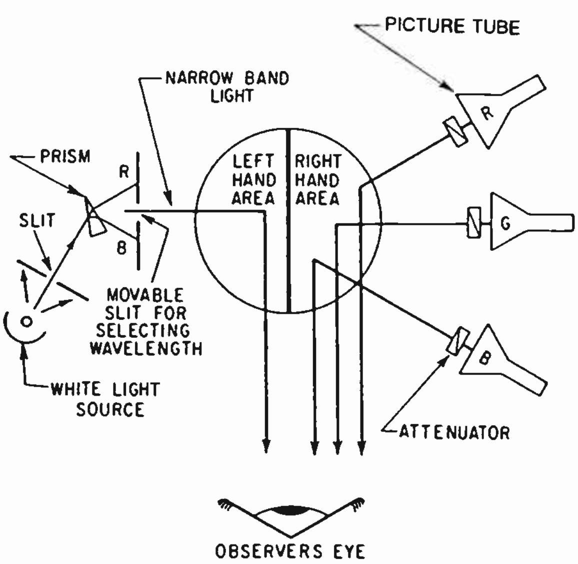


Figure 38. Diagram showing laboratory setup arranged to compare narrow-band light source and R, G, and B light produced by picture tubes to determine proper camera-filter color characteristics.

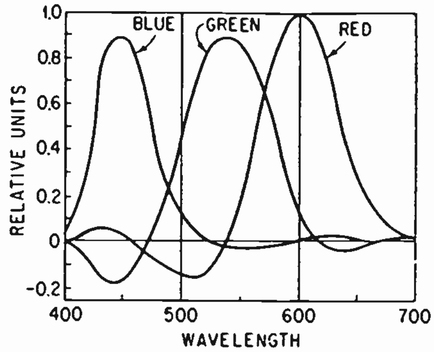


Figure 39. Curves showing relative quantities in camera output required to produce correct picture tube colors over the visible spectrum.

Certain practical difficulties could result in errors in the above procedure. For example, if the observer had any deviations from normality in his color-vision characteristics (as most people do), these deviations would result in "nonstandard" matches and, hence, improper camera-filter characteristics. Also, if the phosphors were contaminated in any way during their manufacturing process (as most phosphors are, at least to some small degree), the resulting phosphor characteristics would not be the proper ones and hence would give rise to improper camera-filter characteristics. The observer errors can be normalized out by standard colorimetric procedures, but phosphor errors represent a basic error which may possibly be present not only in the above experiment but also in varying degrees in a large number of receivers. Quality control of phosphor manufacture is sufficiently good, however, to make the net effect unnoticeable in home receivers.

A striking practical difficulty would also arise regardless of observer or phosphor errors. For most wavelengths, no combination of red, green, and blue picture tube outputs could be found which would produce a match. In order to obtain a match at these wavelengths, it would be necessary to move one or two of the kinescopes over to the other side so that they could add their light to the single-wavelength light being matched. This procedure can be described mathematically, for graphing purposes, by saying that adding light to the left-hand area is the same as subtracting light from the right-hand area. Therefore, the amount of light added on the left would be considered as a negative quantity and would result in a point below the axis on the graph. Since this condition would be found to exist for several successive wavelengths, the resulting graph would show one or more minor lobes below the axis. These are called negative lobes.

These negative lobes represent a need for filters with negative light-transmission characteristics at certain wavelengths. Simple attenuating filters cannot yield such a characteristic, much more elaborate means would be required.

It is theoretically possible to achieve these negative lobes with added camera complexity but it has been

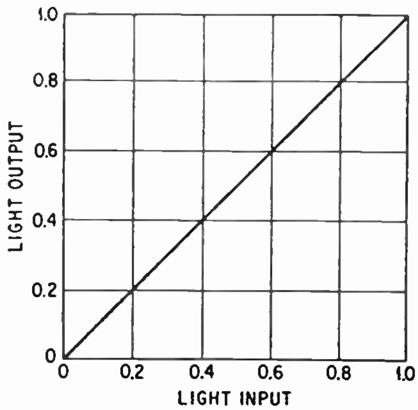


Figure 40. Curve showing light-transfer characteristics of a perfectly transparent piece of window glass.

shown that excellent color fidelity can be obtained by ignoring the negative lobes and using filters which yield the positive lobes only. Positive-lobe processes such as color photography have gained wide acceptance for years. Masking techniques which employ electrical matrixing have been introduced which can modify the spectrum characteristics of a color camera. These techniques can be used to help compensate for deficiencies in the color fidelity such as the lack of negative lobes.

Transfer Characteristics

A piece of window glass is perhaps the nearest approach to a perfect video system. For a piece of glass, the light output (to the viewer) is essentially identical with the light output (from the scene). This fact is shown graphically in Fig. 40. This plot could be called the "transfer characteristic" of a piece of glass, since it describes the way that light is transferred through the system.

If the window glass is replaced by a neutral-density filter which attenuates light 3 to 1, the transfer characteristic will then be given by Fig. 40. The difference between Fig. 40 and 41 can be described by these simple relationships:

For the glass:

Light output = light input

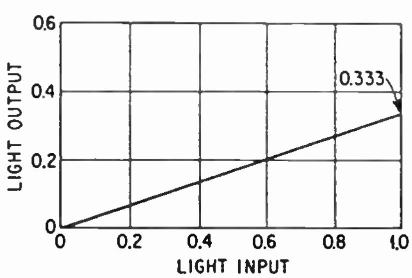


Figure 41. Curve showing transfer characteristic of a neutral density filter with 3-to-1 light attenuation.

For the neutral-density filter:

Light output = $k \times$ light input

where: $k = 1/3$ in this case

Both systems are linear; that is, doubling the light input of either will double its light output; tripling input will triple output; etc. A nonlinear system does not exhibit this simple proportionality. For example, consider a system described by

Light output = $k \times$ (light input)^2

Doubling the input to this system will quadruple its output; a threefold increase in input will result in a ninefold increase in output; etc. The transfer characteristic for this type of system is shown in Fig. 42. Note that the characteristic is definitely nonlinear; that is, it is not a straight line as were Figs. 30 and 40.

In television and photography, nonlinearity is more common than linearity. For example, an ordinary picture tube is a nonlinear device, having a transfer characteristic which can be approximated by the expression:

Light output = k (voltage input)$^{2.2}$

Camera tubes can be linear or nonlinear devices. For example, the characteristic of a vidicon is approximately

Current input = k (light input)$^{0.65}$

The general expression for nonlinear transfer characteristic can be given approximately as

Output = k (input)$^{\gamma}$

where the exponent is the Greek letter gamma.

Graphical Displays of Transfer Characteristics

Linear plots. The first reaction of any person asked to display two variables (like light input and light output) on a set of XY coordinates is to divide X and Y coordinates into equal increments and plot the variables in this manner. A typical result of such a plot

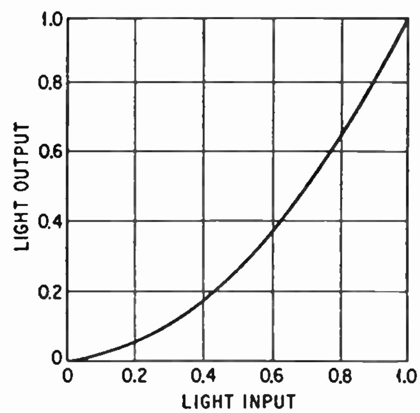


Figure 42. Curve showing a nonlinear transfer characteristic.

has already been described (Fig. 40 and 41). Such a plot has the advantage of showing at a glance the linearity of the device described by the variables. If the plot is a straight line, we say the device is linear; if curved, we say the device is nonlinear. Moreover, the slope of the line describes the attenuation (or gain) of the device. If the slope is unity (which occurs when the plot makes a 45° angle with the X axis), there is no attenuation; we are dealing with a very good piece of glass. For the neutral-density filter described above, which has the equation (light output) = ⅓ (light input), the line has a slope of one-third (see Fig. 41).

Such are the advantages of plotting transfer characteristics with equal-increment divisions of the X and Y axis. However, other advantages (very important ones) can be obtained by dividing up the X and Y coordinates logarithmically. Such a plot is called a log-log plot.

Log-Log plots. Consider a system which has a transfer characteristic given by $L_0 = (L_{in})^{2.2}$. If this equation is plotted on axes which are divided logarithmically, the resulting plot is the same as though the logarithm of both sides of the equation were plotted on equal-increment axes. Taking the logarithm of both sides, we obtain

$$\log L_0 = \log (L_{in})^{2.2}$$

Since $\log (L_{in})$ is the same as $2.2 \log(L_{in})$, then $\log L_0 = 2.2 \log L_{in}$.

Comparing the form of this equation with an earlier equation, light output = ⅓ light input, we can see that just as the attenuation, ⅓, was the slope of the earlier equation, so 2.2, the exponent, is the slope of the latter equation. We see then that the use of logarithmically divided coordinates yields a plot in which the exponent is given by the slope of the line. Therefore, this plot will show at a glance the magnitude of the exponent and will also show whether or not the exponent of the system is constant for all light levels. It also is advantageous in showing the effects of stray light.

Fig. 43A and 43B compare the two types of plotting for three types of transfer characteristics.

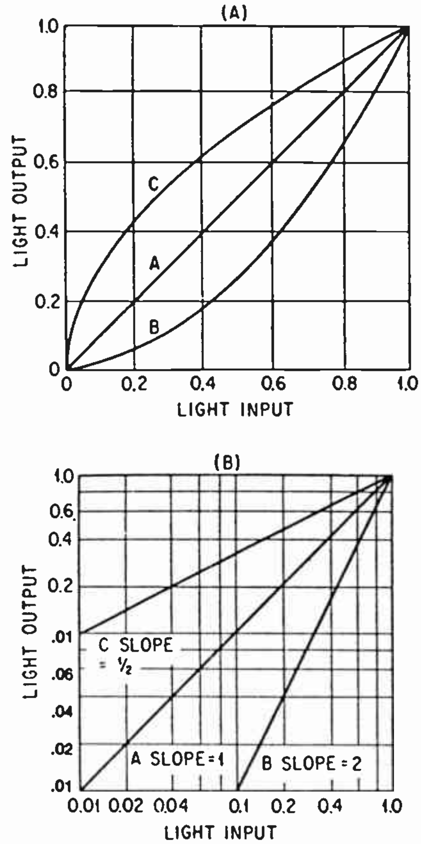


Figure 43. Graphs showing the curves obtained by plotting A, B, and C types of transfer characteristics on linear coordinated (A) and on log-log coordinates (B).

The Effect of a Nonlinear Transfer Characteristic on Color Signals

Effect of identical nonlinearities in each channel. In monochrome television, some degree of nonlinearity can be tolerated, but such is not the case for a color television system. It can be shown that a system exponent different from unity must inevitably cause a loss of color fidelity. For an example, consider a situation in which signals are being applied through linear amplifiers to the red and green guns of a perfectly linear (theoretical) picture tube. The green amplifier is receiving 1.0 volt; the red amplifier, 0.5 volt. If everything is perfectly linear, the proportions of the light output should be 1.0G + 0.5R = greenish yellow. However, if the kinescope has an exponent of 2.0, the light output will be $(1.0)^2G + (0.5)^2R = 1.0G + 0.25R$ = greenish yellow with an excess of green.

From the above specific case, it may be correctly inferred that in general, a system exponent greater than 1 will cause all hues made of the combination of two or more primaries to shift toward the larger or largest primary of the combination. Conversely, a system exponent less than 1 will shift all hues away from the largest primary of the combination.

In the above example, an exponent of 0.5 would yield $(1.0)^{0.5}G + (0.5)^{0.5}R = 1.0G + 0.707R$ = a greenish yellow which is a shade off a pure yellow.

In addition, the reader can correctly conclude that white or gray areas, in which all the primaries are equal, will not be shifted in hue by a nonunity exponent.

Effect of differing exponents in each channel. The preceding discussion assumed that all three channels (in Fig. 33) have the same exponent, whether in unity or not. In practical systems, however, there is always the possibility that the exponents of the channels may differ from one another. This situation will produce intolerable color errors if the differences become even moderately large. In general, the requirements for "tracking" among the light-transfer characteristics of the individual channels are even more stringent than the requirement for unity exponent.

Figs. 44A, 44B, 44C, and 44D show graphically the effects of unequal exponents in the three channels. In

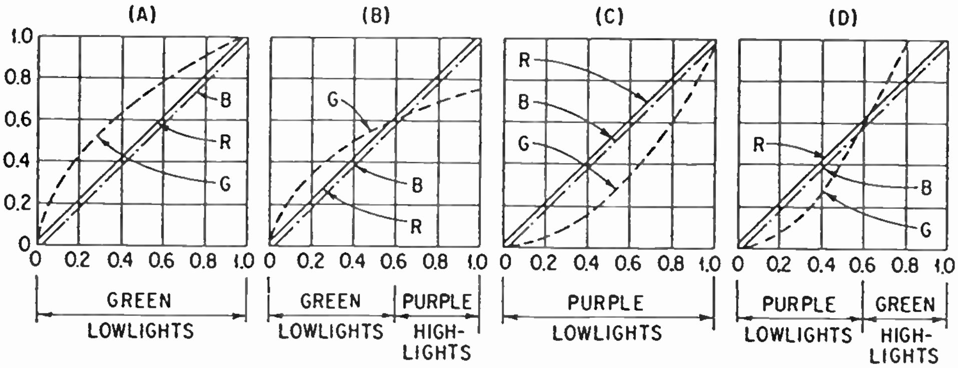


Figure 44. Linear plots showing graphically the effect of unequal exponents in the R, G, and B channels. In all four graphs the R and B exponents are taken as unity. In (A) and (B) the green exponent is taken as less than 1, and in (C) and (D), as greater than 1.

all four figures, the red and blue exponents are taken as unity; in Figs. 44A and 44B the green exponent is taken as less than 1, and in Figs. 44C and 44D, as greater than 1. In Fig. 44A, the transfer characteristics are shown for the system adjusted to produce peak white properly. It can be seen that the bowed characteristic of the green channel will cause all whites of less than peak value to have too much green. A gray-scale step tablet before the camera would be reproduced properly only at peak white; the gray steps would all have a greenish tinge. Relative channel gains could be readjusted to reproduce one of the gray steps properly (Fig. 43B), but then all highlight steps would be purplish while lowlight steps would still be greenish.

A green-channel exponent greater than unity would reverse the above results (Fig. 44C and 44D). With gains adjusted to reproduce peak white properly (Fig. 44C), lowlights would be purplish; with gains readjusted to provide proper reproduction for one of the lower steps (Fig. 44D), highlights would be green and lowlights purple.

The effect of stray light. If a picture tube is viewed in a lighted room, there will always be some illumination on the faceplate. Therefore, the eye will always receive some "light output" from the picture tube, regardless of the magnitude of the signal input voltage. Under this condition, a true black is impossible to obtain.

This condition is reflected in the transfer characteristic of the system. If, for example, the stray light were 5% of the peak highlight brightness of the picture, a linear plot of light output versus light input would have the entire transfer characteristic shifted upward by 5%. However, the most interesting change is found in the log-log plot, where, as seen in Fig. 45, the stray light causes a change in the slope in the lowlight regions. Since the slope is equal to the exponent, this change shows that stray light causes an effective exponent error in the lowlight regions of the picture and hence will cause color fidelity errors which will be most marked in lowlight regions.

These errors will be noted by an observer as improper hues and saturations, with the saturation errors (a "washing out" of the more saturated lowlight areas) being the more objectionable to a viewer.

Stray light is not the only cause of errors of this type. Similar effects will be noted whenever the kinescope bias ("brightness") is set too high, if camera pedestal is set too high, or if stray light enters the camera (whether through lens flare or any other source). In general, any condition which prevents the light output of the system from becoming zero when the light input is zero will cause errors similar to those caused by stray light.

Linearizing a system. It can be shown that a system using a vidicon with an exponent of 0.675 to drive a kinescope with an exponent of 2.2 will have an overall exponent given by the product $0.65 \times 2.2 = 1.43$, assuming that all devices in the system are linear. In general, the overall exponent of a system is the product of the exponents of the cascaded elements.

This knowledge provides an excellent tool for linearizing a system. For example, a system with an overall exponent of 1.43 could be linearized by inserting somewhere (in a video path) an amplifier having an exponent of 1/1.43 (= 0.7) so that the product becomes unity: $1.43 \times 1/1/43 = 1$.

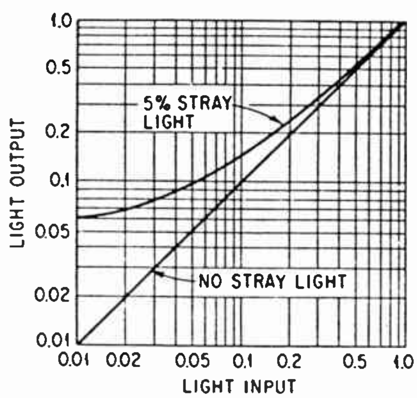


Figure 45. Log-log plot of system with stray light, illustrating change of slope in the low-light regions.

In Fig. 34, a nonlinear amplifier, or gamma corrector, is shown inserted in each of the three paths.

Possible Encoding and Decoding Distortions

The second of the two systems discussed in the preceding section bordered on being a practical system but still required three independent 4 MHz channels. A fortunate characteristic of the human eye (the inability to see colored fine detail) allows us to modify this requirement to one 4 MHz channel for monochrome fine detail and two much narrower channels for color information. Before this modification can be made, the red, green, and blue signals must be combined to form three other signals, usually called M, I, and Q, such that the M signal alone requires a 4 MHz channel, and the I and Q channels, which contain the color information, are confined to narrower channels. This rearrangement of red, green, and blue to form M, I, and Q is called matrixing and was described in the previous part. A system which uses a matrix is block-diagramed in Fig. 35. The illustration also shows that to recover the original red, green, and blue signals at the receiving end, a "rearranging" device is needed. This device is usually called the receiver matrix.

Matrixing alone offers no advantage unless steps are taken actually to limit the I-signal and Q-signal channels to the narrow bandwidths allowed. Fig. 36 shows a system employing such band shaping. The band-shaping filters themselves always introduce delay, which must be compensated for by placing delay lines in the wider band channels, as shown in the diagram.

To put both color and monochrome information in the spectrum space normally occupied by monochrome only requires that the color information overlap the monochrome. This overlap can be allowed for both I and Q signals, without incurring visible cross talk, if two techniques, known as frequency interlace and two-phase modulation are employed. A system using these techniques, which were described in the section on Electronic Aspects of Compatible Color Television, is block-diagramed in Fig. 37.

Possible Errors in the Matrixing Process

The entire matrixing process can be summed up in two sets of equations, the first set describing how the transmitter matrix takes in red, green, and blue and turns out M, I, and Q:

$$M = 0.30R + 0.59G + 0.11B$$

$$I = 0.60R - 0.28G - 0.32B$$

$$Q = 0.21R - 0.52G + 0.31B$$

and the second set describing how the receiver matrix takes in M, I, and Q and recreates red, green, and blue:

$$R = 0.94I + 0.62Q + M$$

$$G = 0.27I + 0.67Q + M$$

$$B = 1.11I + 1.7Q + M$$

Both matrices can therefore be considered as analogue computers which continuously compute the desired output from the given input. The coefficients in the above six equations are usually determined in the computers by precision resistors or, in the case of negative numbers, by precision resistors and signal-inverting amplifiers. The basic error that can occur, therefore, is a change in a resistor value or an amplifier gain, resulting in a change in one or more coefficient. In general, the resulting picture error resembles cross talk among the primary colors.

More specifically, the transmitter matrix can have two distinct types of errors. The first type involves the coefficients of the equation for M; the second type, the coefficients for I and Q. An error in an M coefficient will brighten or darken certain areas. In a monochrome reproduction of a color signal, such an error, if small, would not be noticed; if large, it would still probably be tolerated by the average viewer. In a color reproduction, however, even a small error would be objectionable. For example, a reduction of the red coefficient from 0.3 to 0.2 would cause a human face to be reproduced with an unnatural ruddy complexion and dark red lips.

Note that the sum of the M coefficients is 1. An error in one coefficient would change this sum, so that peak white would no longer occur as 1 volt. An operator could mistake this condition for a gain error and adjust either M gain or overall gain in an effort to obtain the correct peak-white voltage. Changing M gain would cause errors to occur in all M coefficients; changing overall gain would put errors in all coefficients. Although such an error is rare in well-engineered equipment, it is a possible source of color error which can be compounded by misdirected attempts at correction.

Note that the sums of the Q and I coefficients are each zero, which means that when R = G = B (the condition for white or gray), Q and I both equal zero. An error in a Q or I coefficient would cause color to appear in white or gray areas and, in addition, would cause general errors in colored areas resembling cross talk among the primaries. Controls are usually provided in the Q and I matrices, called Q white balance and I white balance, respectively, which allow the operator to adjust the sum of the Q or I coefficients by changing the value of one of the coefficients. If the coefficient controlled is the one in error, then adjusting white balance restores the condition that the sum of the coefficients is zero, that is, it removes the color from white and gray objects, but it does so by giving the controlled coefficient an error which just counteracts the error of a nonadjustable coefficient, so that two coefficients are wrong instead of one. Again, such an error is rare in well-engineered equipment, for the adjustable coefficient is usually the one in error. However, the possibility of an error compounded by adjustment should be kept in mind.

A far more likely cause of white-balance error is an error in input level, that is, a discrepancy between the peak white levels of input red, green, or blue. In such a case, an operator can still achieve white balance (Q and I = 0 for white input) but the entire system will be in error. The starting point for all investigations of the cause of white-balance errors should be the levels of the red, green, and blue colorplexer inputs.

In the receiver matrix, only one general type of error can occur instead of two as in the case of the transmitter matrix. This type of error, a general coefficient error, results in cross talk among the primary colors. For example, a change in the I coefficient for the red equation from 0.94 to 0.84 would yield about a 7% reduction in the peak red output available and would also result in unwanted red light output in green or blue areas at about 3½% of the green or blue level.

Gain Stability of M, I, and Q Transmission Path

In the system of Fig. 35, every gain device or attenuating device in the three transmission paths must maintain a constant ratio between its input and output in order to maintain the proper ratios among the levels of M, I, and Q at the input to the receiver matrix. A variation in the gain of one of these paths will result in a loss of color fidelity.

For example, a reduction in M gain must obviously cause a reduction in the viewer's sensation of brightness. Not quite so obvious are the effects of I and Q gain. Since these are color signals, their amplitude would be expected to influence the sensation of saturation, but the manner of this influence is not intuitively obvious until the factors which influenced the selection of I and Q compositions are recalled. It previously was pointed out that the eye has the greatest need for color detail in the color range from orange to blue-green (cyan) and the least in the range from green to purple. Hence I, the wider band signal, conveys mainly orange and cyan information, and Q, the narrower band signal, conveys principally the greens and purples. Therefore, a reduction in I gain could be expected to reduce the saturation sensation for colors in the orange and cyan gamut, leaving the greens and purples virtually unaffected. Conversely, Q gain will influence the greens and purples without causing much change in the appearance of orange and cyan objects.

Modulation and Demodulation

The system of Fig. 35, which introduced bandwidth limiting of the I and Q signals in accordance with the capabilities of the eye to see colored fine detail, is a fairly practical and economical system, except for the fact that three individual transmission channels are employed. If we are to have a compatible system, however, these three channels must be reduced to one through some multiplexing technique. The technique used has already been described, and a system employing this technique is block-diagrammed zero)in Fig. 37.

Possible Errors in Modulation

Burst phase error. Perhaps the most fundamental error in the multiplexing process would be an error in the phase of the main timing reference, burst. Since the entire system is based on burst phase, an error in burst phase will appear as an opposite error in every phase except burst, because the circuits will insist that burst phase cannot be wrong. The general result will be an overall hue error in the reproduced picture. This effect can be better visualized by referring to Fig. 46.

A phase error in burst produces the same result as holding burst phase stationary and allowing all other phases to slip around the circle an equal amount (but in a direction opposite to the burst phase error). Each color vector then represents a hue other than the one intended.

Burst amplitude error. In theory, the receiver circuits which extract timing information from the burst are insensitive to variations in burst amplitude as long as the burst is large enough to maintain a respectable signal-to-noise ratio and not so large that some type of clipping or rectification upsets the burst circuitry. But practical receivers always exhibit some degree of sensitivity depending mainly upon the error in the subcarrier oscillator in the receiver. If the free-running frequency of the receiver oscillator is very different from burst frequency—particularly if the difference is so great that the burst is in danger of losing control of the oscillator—then a fairly appreciable amplitude sensitivity will be noted. This sensitivity will take the form of a phase error, and the net result will be indistinguishable from a burst phase error, as discussed above.

Some receivers have a circuit which automatically adjusts the gain of the color-information channels so that the viewer always sees the proper saturations, regardless of errors which might tend either to "wash out" or to oversaturate the picture. Such a circuit, called an automatic chroma control (ACC), derives its control information from the amplitude of burst, which

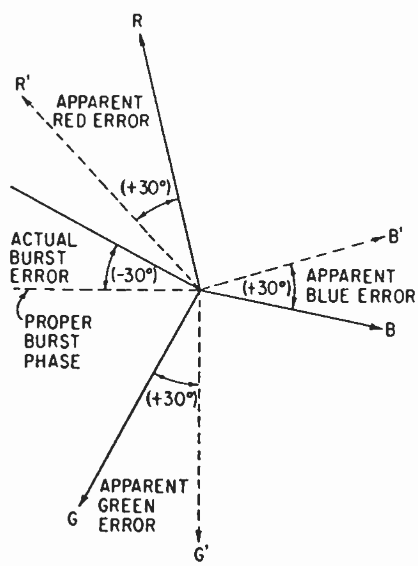


Figure 46. Vector diagram showing how error in subcarrier phase becomes an opposite error in all other phases.

is presumed to bear a constant ratio to the amplitude of chroma. Transmission distortions, for example, might decrease the amplitude of both burst and chroma, but since the ratios of their amplitudes would be preserved, an ACC receiver could automatically modify its chromachannel gain to compensate for the decreased chroma amplitude. However, if a color encoder error should cause burst alone to decrease in amplitude, the ACC circuits would increase chroma gain just as in the above case, with the result that a viewer would receive an oversaturated picture.

Two-phase modulation errors. The fidelity of color reproduction can be seriously affected if the phase separation of the Q and I subcarriers is not maintained at 90°. It can be shown that a "slip" in the angular position of the Q axis, for example, will result in cross talk of Q and I. The final result will be the same as cross talk among the primary colors.

Likewise, in a receiver, the phase relationship between the reference subcarriers must be maintained to avoid a similar error. Any deviation from the proper phase relationship will have a result similar to the above, that is, cross talk of I into Q or Q into I, with the net picture result resembling cross talk among all the primary colors.

Carrier unbalance. In a properly operating doubly balanced modulator, the carrier component of the signal is suppressed in the modulator circuit. If some error in components or operation causes this suppression to be imperfect, the carrier will appear in the output. This condition is known as carrier unbalance.

The effect of carrier unbalance can be evaluated by considering the unwanted carrier as a vector of constant amplitude which adds itself vectorially to every vector present in the colorplexer output. In general, such a vector will shift all vectors and hence all hues seen in the picture toward one end of the other of the color axis represented by the unbalanced modulator. For example, a positive unbalance in the I modulator would shift all colors toward the color represented by the positive I axis, that is, toward orange. A negative I unbalance would shift all colors toward cyan.

To visualize this effect, refer to Fig. 47, in which has been added to each color vector a small positive vector which is parallel to the I axis. This small vector represents the amount of carrier unbalance. The resultant vectors will all be rotated toward the positive I axis and changed in amplitude as well. Such changes represent errors in both hue and saturation.

Another error from carrier unbalance occurs in white and gray areas of the picture. In a normally operating colorplexer, a white (or gray) area in the scene causes the Q and I signals to become zero and thereby causes the modulator outputs to become zero. Hence, a white or gray area will normally appear in the signal as an interval of zero subcarrier amplitude. If one of the modulators begins to produce a carrier-unbalance vector, however, a white or gray area will become colored because of the subcarrier which will be added in this interval. Moreover, certain areas which are normally colored may have their subcarrier canceled by the

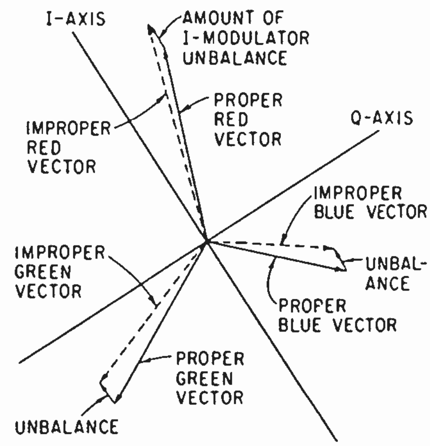


Figure 47. Vector diagram of subcarrier phase and amplitude with positive vectors added to represent carrier unbalance in the I modulator.

carrier-unbalance vector and become white. Such white-to-color and color-to-white errors are very objectionable.

Video unbalance. A doubly balanced modulator derives its name from the fact that it balances out or suppresses both the carrier (as described above) and the modulating video (Q or I). If, for any reason, the video suppression becomes less than perfect, the resulting condition is called video unbalance.

Video unbalance will cause unwanted Q or I video to appear in the modulator output, in addition to the desired sideband outputs. This unwanted video signal will be added to the luminance signal, thereby distorting the gray scale of the picture. For example, a slight positive unbalance in the Q modulator would slightly brighten reds and blues and slightly darken greens. A negative unbalance would have the opposite effect.

Subcarrier frequency error. The color subcarrier frequency is specified by the Federal Communications Commission to be 3.579545 MHz ±10 Hz. Deviations within this specified limit are of no consequence (provided they are slow deviations). Large deviations, however, can affect color fidelity. The effect does not usually become serious within the possible frequency range of a good crystal-controlled subcarrier source driving a properly designed receiver.

In receivers, the subcarrier timing information is extracted from the burst on the back porch and used to control the frequency of a subcarrier-frequency oscillator in the receiver. As long as the unlocked frequencies of the burst and the receiver oscillator remain the same, the locked phase relationship between the two will remain the same. But if either the burst frequency or the receiver-oscillator frequency becomes different (and the difference between them is not so large that lockup is impossible), then the locked error, which obviously cannot be a frequency error, manifests itself as a phase error. This error can become as large as ±90° before the AFC circuit can no longer hold the receiver oscillator on frequency. The frequency range over which this phase shift occurs depends upon the receiver design.

Possible Distortions in the Transmission System

Preceding sections have described the processes involved in the generation and display of a color television signal. Errors in these processes are not the only possible source of distortion; when the signal is transmitted over great distances, the transmission system itself may contribute errors. This section discusses parameters which specify the behavior of a transmission system and describes the effects that errors in these parameters can have on the reproduced picture.

This section is divided into two parts. The first relates to the parameters of a perfectly linear transmission system, while the second part discusses the additional parameters required to describe the nonlinearities that are inevitable in any practical system.

The Perfectly Linear Transmission System

A perfectly linear and noise-free transmission system can be described by its gain and phase characteristics plotted against frequency as the independent variable.[2] Typical plots are shown in Fig. 48 and 49, respectively. These two characteristics known, it is possible to predict accurately what effect the transmission system will have on a given signal.

Gain Characteristic

Fig. 48 is usually known as the frequency response or gain characteristic of the system. Ideally, it should be perfectly flat from zero to infinite frequency, but this, of course, is impossible to attain. An amplifier has a definite gain-bandwidth product, depending upon the transconductance of its active elements (tubes or transistors), the distributed capacity shunting these elements, and the types of compensation (peaking) employed. The bandwidth of a given combination of tubes, transistors, stray capacitances, and peaking networks can be increased only by decreasing its gain, or conversely, its gain can be increased only by decreasing its bandwidth. There is a limitation, therefore, to the actual bandwidth that can be obtained. For a given scanning standard, the bandwidth required in a monochrome-television system is determined by the desired ratio between the horizontal resolution and the

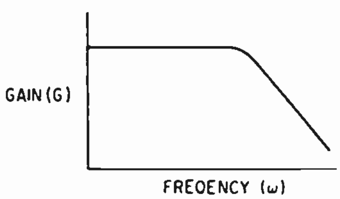


Figure 48. Typical curve showing a gain of a system plotted against frequency to determine its gain characteristic.

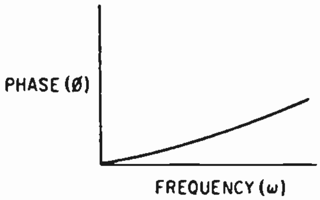


Figure 49. Curve showing phase characteristic of a system plotted versus frequency.

vertical resolution. Although nominally a 4.0 MHz bandwidth is required for the monochrome standards, the requirement can be relaxed to the detriment of only the horizontal resolution. The subjective result is a "softening" of the picture in proportion to the narrowing of the bandwidth (neglecting the influence of the phase characteristic in the vicinity of the cutoff frequency). As pointed out in preceding sections, the entire chrominance information of the color system is located in the upper 1.5 MHz of the prescribed 4.0 MHz channel; hence, any loss of response in this part of the spectrum can have a marked effect on the color fidelity of the reproduced picture.

One of the most serious forms of distortion inflicted on a color picture by bandwidth limiting is loss of *saturation*. Consider a case in which the bandwidth is so narrow as to result in no gain at the color subcarrier frequency. The output signal then contains no color subcarrier and hence reaches the color receiver as a monochrome signal, producing zero saturation. Nearly as poor results can be expected from an amplifier with response such that the gain at 3.58 MHz is one-half the low-frequency gain. Since the saturation depends chiefly on the amplitude of the subcarrier, the saturation will be correspondingly reduced. The resultant color picture will have a "washed out" look.

Loss of high-frequency response, which can be expected to contribute to loss of fidelity, is usually accompanied by phase disturbance, depending on the type of networks employed in the system. The intent in this section, however, is to treat each variable separately. Therefore, discussions are based on the effects of varying only one parameter of a system. It is suggested that the reader can determine the combined effect of two or more variables by comparing the results shown for the individual variables.

Phase Characteristic

An ideal system has a *linear* phase characteristic, as in Fig. 50A. Such a characteristic implies that all frequencies of a signal have exactly the same *time delay* in passing through this system, since the time delay is given by the phase angel divided by the (radian) frequency. It can be seen in Fig. 50 that if three frequencies are chose arbitrarily, then the corresponding phase angles must have values proportional to their corresponding frequencies (because of the geometric properties of a right triangle). To state it another way, if $\phi_1/\omega_1 = 0.2$ μsec, then $\phi_2/\omega_2 = 0.2$ μsec, and ϕ_3/ω_3 also equals 0.2 μsec. Plotting these three values and

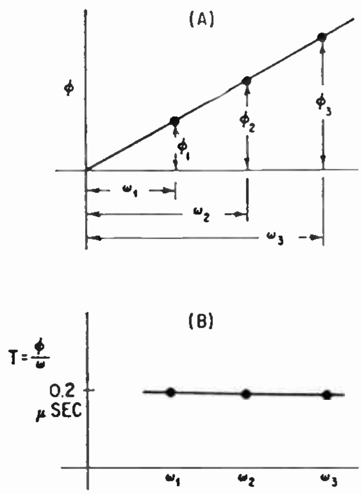


Figure 50. Curves illustrating a system with linear phase characteristics, which will give the same time delay for signals of all frequencies.

drawing a straight line through them as in Fig. 50A will show that the time delay for all frequencies is 0.2 µsec.

A signal is not distorted by delay as long as all parts of it are delayed by the same amount. However, when the phase characteristic is nonlinear (as in Fig. 51A), the time delays for all parts of the signal are no longer equal (see Fig. 51B). For example, if a complex waveform is made up of a 1 MHz sine wave and its third harmonic, these two components will suffer unequal delays in passing through a system having the characteristics of Fig. 50. The resultant distortion can be seen by comparing Figs. 52A, 52B, and 52C.

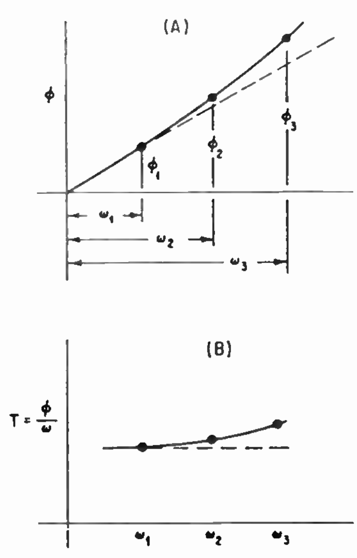


Figure 51. Curves showing the effect of nonlinear phase characteristic on time-delay characteristic.

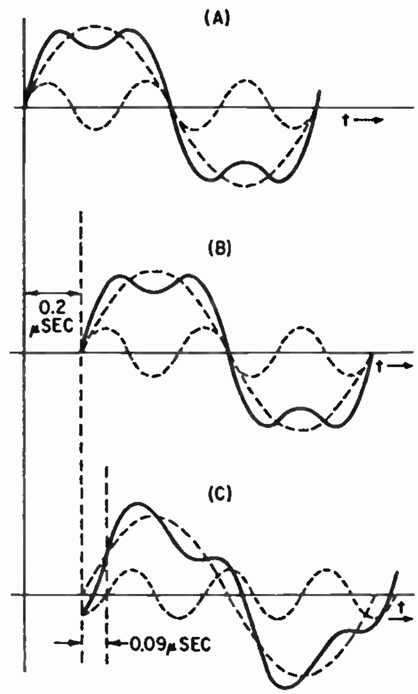


Figure 52. Curves showing that a complex wave (A) is not distorted by time delay (B) when both components (shown dotted) are delayed by the same amount. Unequal delays (C), however, cause distortion.

Such distortion is detrimental to both the luminance and chrominance of a composite signal. The luminance signal will have its edges and other important details *scattered*, or *dispersed*, in the final image. Such a transmission system is said to introduce *dispersion*. (Conversely, if a system does not scatter the edges and other high-frequency information, it is said to be dispersionless.) The effect of phase distortion on the chrominance information is of a rather special nature and can best be explained by introducing the concept of *envelope delay*.

Envelope Delay

In the preceding discussion, the time delays ϕ_1/ω_1, ϕ_2/ω_2 and ϕ_3/ω_3 were always determined by measuring the frequencies and the phases from $\phi = 0$ and $\omega = 0$. It might be said that the delay at zero frequency is commonly taken as the reference point for all other delays. This method is usually adequate for determining the performance of systems that do not carry any signals which have been modulated onto a carrier. But a carrier, with its family of associated sidebands (Fig. 53B), can be thought of as a method of transmitting signals in which the zero frequency reference is translated to a carrier frequency reference. This translation can be understood by referring to Fig. 53A and 53B. To calculate the delay of the carrier-borne signals *after* they have been demodulated, measurements of ϕ and ω must be referenced, not from zero frequency, but from *carrier* frequency.

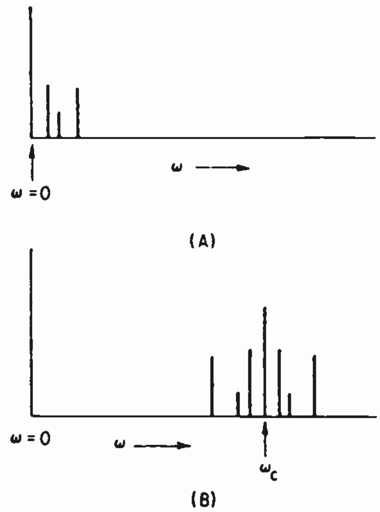


Figure 53. Sketch showing how a group of frequencies near $\omega = 0$ [sec. (A)] can be translated by modulation onto a carrier to a group of sidebands near ω_c – a carrier frequency (B).

In Fig. 54A, an impossible phase characteristic has been drawn to aid in further discussion of this subject. Such a characteristic, consisting of two perfectly straight lines, is never met in practice but makes a very simple system for developing the subject of envelope delay.

First, pass two frequencies ω_1 and ω_2 through this system. Let ω_1 be a carrier and ω_2 a sideband which might be, for example, 1000 Hz higher. If ω_1 and ω_2 fall on the characteristic as shown in Fig. 54A, the delay which the 1000 Hz will show after demodulation can be found by putting new reference axes (shown dotted) with ω_1, the carrier, at zero on these new axes. Now, when ϕ_s and ω_s are measured as shown, the time delay after demodulation is ϕ_s/ω_s. In this case, the delay of the 1000 Hz after demodulation is the same as it would have been had it been passed through the system directly.

Second, pass two other frequencies ω_3 and ω_4 through this system as redrawn in Fig. 54B. This time drawing in the new axes at ω_3, it can be seen that although ω_s is still 1000 Hz, ϕ_s is larger than ϕ_s. Therefore, it can be concluded that the time delay ϕ_s/ω_s for this second case is greater than for the first case. The 1000 Hz, when demodulated, will show a considerable error in timing.

Stressing the phrase "delay in a demodulated wave" should not be taken to mean that the demodulation process produces this delay or even make it apparent where it was previously not detectable. Any delay that a demodulated wave shows was also present when the wave existed as a carrier having an envelope. In short, the delay of the demodulated wave appears first as a delay of the envelope, hence the phrase *envelope delay*.

Envelope delay does not constitute a distortion. If a system such as the one shown in Fig. 54A introduces a delay of 0.2 μsec to the 1000 Hz wave (measured after demodulation), then the *envelope delay* of the system is 0.2 μsec. However, it was shown that a 1000 Hz signal passed directly through the system (without first being modulated into a carrier) would also suffer a delay of 0.2 μsec. As long as the envelope delay ϕ_s/ω_s is the same as the time delay ϕ_1/ω_1 the envelope delay introduces no timing errors. But in the second system (Fig. 54B) the demodulated 1000 Hz wave suffered a *larger* delay, say 0.29 μsec. A 1000 Hz signal passed directly through this system, however, would still be delayed only 0.2 μsec. Therefore, the second system has an *envelope delay* of 0.29 μsec and an *envelope delay distortion* of 0.09 μsec.

It is probably wise to point out that the time delay ϕ_3/ω_3 in Fig. 54B is considerably less than the 0.29 μsec estimated for the value of envelope delay. Although ϕ_3/ω_3 would be greater than 0.2 μsec (say, for example, that ϕ_3/ω_3 is 0.22 μsec), the value would be optimistic about the amount of timing error that would be shown by the demodulated 1000 Hz signal. The need for a knowledge of the envelope delay ϕ_s/ω_s of the system is therefore obvious.

Effect of envelope delay distortion on a color picture. A transmission system which exhibits envelope delay distortion will destroy the time coincidence

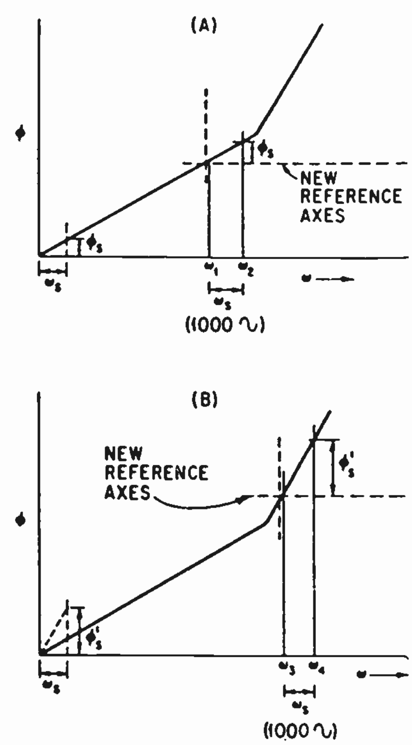


Figure 54. Idealized straight-line phase characteristics showing how a carrierborne 1,000 Hz signal can be delayed excessively when the carrier and sideband fall on a steeper portion of the phase characteristic.

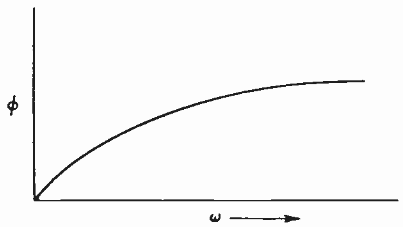


Figure 55. Phase characteristic of an RC network.

between the chrominance and luminance portions of the signal. This will result in misregistration between the color and luminance components of the reproduced picture. The following paragraph explains briefly how envelope delay distortion causes this error.

Any colored area in a reproduced picture is derived from two signals: a chrominance signal and a luminance signal. Since these two signals describe the same area in the scene, they begin and end at the same time. The chrominance signal arrives at the receiver as a modulated subcarrier; the luminance signal does not. Therefore, as shown above, the delay of the chrominance signal is determined principally by the envelope delay of the system and the delay of the luminance signal is determined principally by the ordinary time delay ϕ/ω If the two delays are not identical (that is, if there is envelope delay *distortion*), then the chrominance signal does not coincide with the luminance signal and the resultant picture suffers *color luminance misregistration* in a horizontal direction.

For example, in a system having the characteristic of Fig. 54B, the luminance signal is delayed by 0.2 μsec but the chrominance signal is delayed by 0.29 μsec. The error in registration then amounts to 0.09 μsec, or about 0.2% of the horizontal dimension of the picture, which is about 0.3″ on a 21″ (diagonal) picture.

Although the subject of compatibility is outside the scope of this part, it is worth noting in passing that envelope-delay distortion adversely affects compatibility, since it causes wideband monochrome receivers to display a misregistered dot-crawl image in addition to the proper luminance image.

General method for envelope delay. The specific cases described (Fig. 54A and 54B) made use of simple, idealized straight-line approximations to develop the concept of envelope delay. Practical circuits are not so simple. For example, a simple *RC* network has a ϕ versus ω plot as in Fig. 55. Finding the envelope delay of this curved-line plot will clarify what is meant by envelope delay.

Referring back to the plots of Fig. 53A and 53B, it can be seen that the characteristic of the plot that determines the value of envelope delay is its *slope*. The larger envelope delay, which was suffered by the ω_3-ω_4 pair (Fig. 54B), was a result of their lying on the steeper slope. The envelope delay of *any* system is equal to the slope of the phase versus frequency characteristic. If this characteristic is a curved line (as for the *RC* network, Fig. 55), then the slope is different at at every frequency.

The slope of a curved line can be found by the methods of the differential calculus or to a good approximation by breaking up the line into a number of straight-line segments, as in Fig. 56. If the slope of each of these straight lines is plotted against its corresponding frequency (that corresponding to the center of the line), the resulting curve will be approximately the envelope-delay characteristic.

Nonlinearities of a Practical Transmission System

It is important to emphasize that the effect of nonlinearities in a color television system depends upon whether these nonlinearities precede or follow the matrixing and modulation sections of the system. Nonlinearities in transfer characteristics detract from color fidelity; the same degree of nonlinearity after matrixing and modulation also affects color fidelity although in a different way. The purpose of the following paragraphs is to discuss how a nonlinear transmission system affects a *composite* color signal. It is assumed that all other nonlinearities in the entire system either are negligible or have been canceled by use of nonlinear amplifiers such as gamma correctors.

The major sources of nonlinearity in a transmission system are its amplifying devices. These devices (tubes and transistors)[3] have a limited dynamic range. For example, if too much signal is supplied to them, an *overload* results. The transfer characteristic of such a system can be sketched as in Fig. 57A.

Such a nonlinearity is one of three types commonly encountered in video transmission systems. These three types are:

- Incremental gain distortion
- Differential gain
- Differential phase

The paragraphs below will show that Type 2 is merely a special case of Type 1.

Incremental gain. The concept of the slope of a plot, developed in the discussion of envelope delay, will be useful here as well. Consider a plot as in Fig.

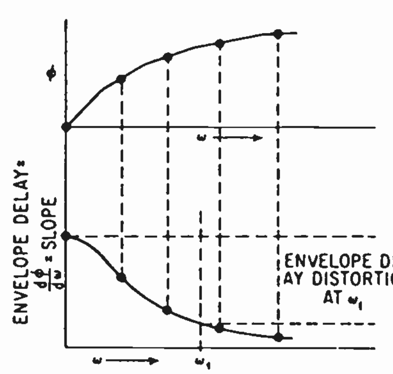


Figure 56. Graphs showing how a series of straight-line segments can be used to approximate the envelope delay characteristics (bottom).

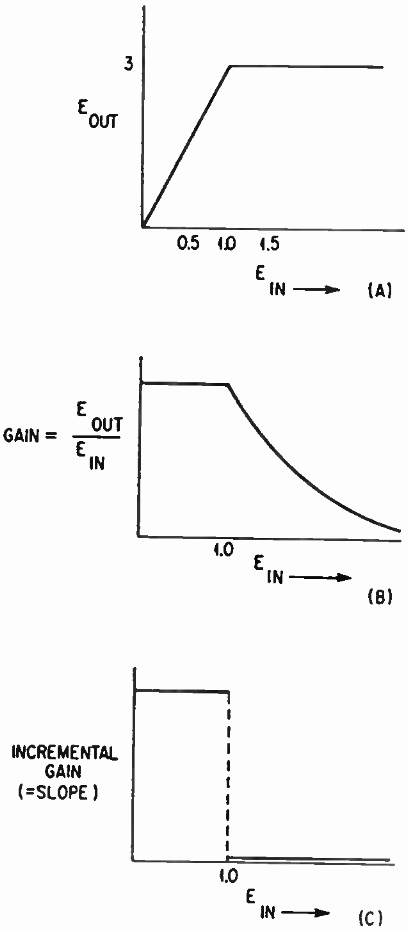


Figure 57. Idealized straight-line plots showing (A) output voltage of an amplifier versus input voltage (B) gain of the amplifier versus input voltage and (C) incremental gain of the amplifier versus input voltage. Curve (C) is the slope of curve (A).

57A which shows output voltage of an amplifier plotted against input voltage. Idealized straight-line plots are shown for simplicity. It can be seen that the amplifier has a maximum output of 3 volts for 1 volt input. Larger input voltages result in no more output; the amplifier *clips* or *compresses* when inputs larger than 1 volt are applied.

The gain of the amplifier is

$$\text{Gain} = \frac{E_0}{E_{in}} = \frac{3 \text{ volts}}{1 \text{ volt}} = 3$$

The gain is obviously constant below the clip point. For example, k an input voltage of 0.5 volt gives

$$\text{Gain} = \frac{1.5 \text{ volts}}{0.5 \text{ volt}} = 3$$

But at an input of 1.5 volts, the output is still 3 volts, so the "gain" is only 2. (The word "gain" is of doubtful use here because of the clipping involved.) vThe gain, defined as E_0/E_{in}, is plotted against E_{in} in Fig. 57B. It

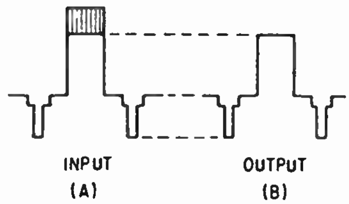


Figure 58. Extreme case of distortion resulting from passing signal at left (A) through the amplifier represented by Fig. 57. The output (B) has no color information remaining.

can be seen in this figure that the gain is constant only as long as the *slope* of Fig. 57A is constant.

It is useful, then, to establish a new term, called *incremental gain*, which will be defined as the *slope* of a plot such as Fig. 57A. For the particular plot of Fig. 57A, the slope is constant up to $E_{in} = 1$ volt and then suddenly becomes zero. The corresponding plot of slope versus E_{in} is shown in Fig. 57C.

The importance of incremental gain in color television can be assessed by applying the input signal shown in Fig. 57 to the distorting system of Fig. 57A. Before being applied to the distorting system, such a signal could be reproduced on a monochrome receiver as a vertical white bar and on a color receiver as a pastel-colored bar, say, for example, a pale green. After passing through the distorting system, the signal would still be reproduced as a white bar on the monochrome receiver with the only apparent error being a luminance distortion, that is, a slight reduction in brightness, which, for the magnitudes shown here, would probably pass unnoticed. The color receiver, however, would receive a signal completely devoid of any color information and would reproduce a white bar in place of the former pale-green one.

The less extreme case is shown in Fig. 59. For the system represented by this characteristic, the slope (incremental gain) does not become zero for inputs above 1 volt but instead falls to one-half its below − 1 volt value. The color signal of Fig. 59 would not lose

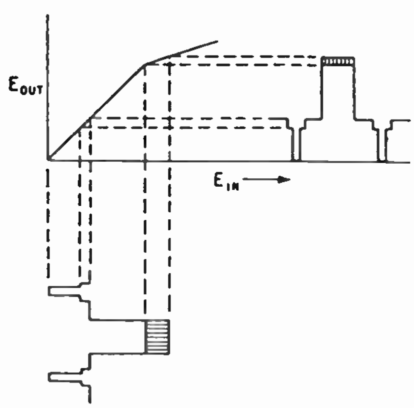


Figure 59. Diagram showng effect of incremental gain distortion of reducing amplitude of color portion of signal.

all color in passing through this system, but the amplitude of the subcarrier would become only one-half of its proper value. Since saturation is a function of subcarrier amplitude, the pale green of the undistorted reproduction would, in this case, become a *paler* green. The luminance distortion would also be less than in the extreme (clipping) case.

It can be seen, then, that unless the incremental gain of a system is constant, that system will introduce compression which will distort the saturation and brightness of reproduced colors. Usually, the error is in the direction of *decreased* luminance and saturation. For certain systems, however, exceptions can be found. For example, the effect that the system represented by Fig. 59 will have on a signal depends on the polarity of the signal. For the signal as shown, the usual *decrease* in luminance and saturation is exhibited. For an inverted signal, however, the subcarrier amplitude would not be reduced, but the luminance signal would still be diminished. The subjective result of this distortion would be an *increase* in saturation. The unusual behavior of this particular system is attributable to its peculiar transfer characteristic, which was drawn with curvature at one end only to simplify the discussion. Most practical system transfer characteristics exhibit curvature at both ends and therefore have an effect on the signal which is essentially independent of polarity.

Incremental gain can be measured in two ways, the first of which stems from its contribution to luminance distortion and the second, from its contribution to chrominance distortion.

In the first method, an equal-step staircase waveform such as shown in Fig. 60A is applied to the system to simulate a signal having equal luminance increments. If the system has constant incremental gain, the output will, of course, also have equal-step increments. But if the system does not have constant incremental gain, certain of the steps will be compressed, as in Fig. 60B. If the compression is as in the figure, the *incremental gain distortion* (IGD) is indicated by the distorted amplitude of the last step. Numerically, it can be stated as a percentage:

$$IGD = 1 - \frac{S_{distorted}}{S_{undistorted}} \times 100\%$$

where S is a step amplitude.

For example, if an undistorted step is 0.1 volt and the distorted one is 0.075 volt, then the incremental gain distortion would be 25%.

Using the other (chrominance distortion) technique, an input signal consisting of the step wave plus a small, high-frequency sine wave, as shown in Fig. 60C, is applied to the system. After the signal has passed through the system, it is fed through a high-pass filter which removes the low-frequency staircase. The incremental gain distortion then is indicated by the differences in the amplitude of the high-frequency sine waves (see Fig. 60D). In this case, the high-frequency sine wave associated with the top step is shown as having 75% of the amplitude of the sine waves associated with the lower steps, which are assumed to be undistorted. Again, the incremental gain distortion is 25%.

A most important point must be made regarding the equivalence of these two techniques. Certain systems which show incremental gain distortion when tested

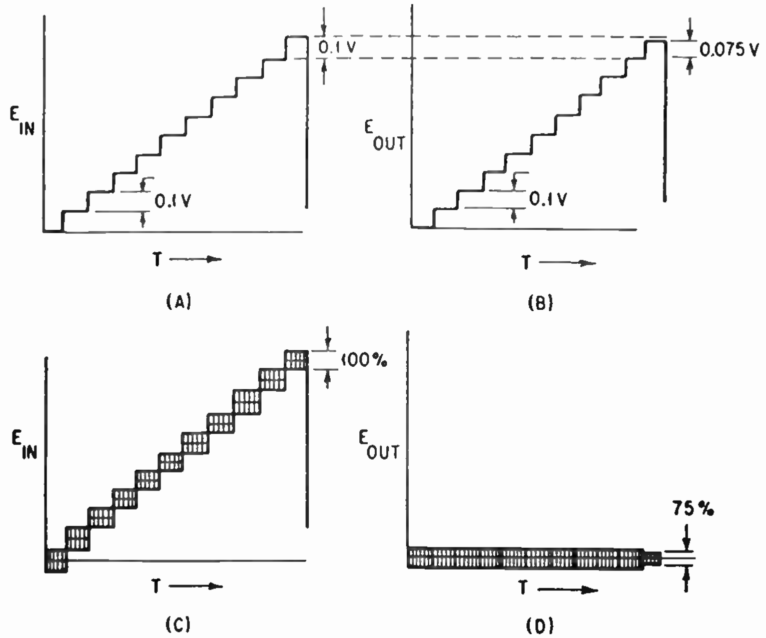


Figure 60. Diagrams showing two methods of measuring incremental gain distortion, namely, in (A) and (B) by its contribution to luminance distortion and in (C) and (D) chrominance distortion.

by the luminance-step technique may or may not show the same distortion when tested by the high-frequency and high-pass filter technique. Moreover, a system which shows distortion by the second technique may or may not show distortion by the first. In other words, the incremental gain distortion may be different for different frequencies. Such differences are frequently found in staggered amplifiers, feedback amplifiers, or amplifiers having separate parallel paths for high and low frequencies, such as might be found in stabilizing amplifiers.

A thorough test of a system, therefore, should include test of its incremental gain by both techniques. The staircase-plus-high-frequency waveform can be used to provide *both* tests by observing the system output (for this test waveform input) first through a low-pass filter and then through a high-pass filter. The first test will show low-frequency distortions; the second, high-frequency distortions.

Differential gain. On the basis of the above discussion of incremental gain distortion, the extremely important concept of *differential gain* can be presented merely as a simple definition. Differential gain is identical with incremental gain distortion when the latter is measured by observing ". . . the difference in the gain of the system for a small high-frequency sine-wave signal at two stated levels of a low-frequency signal upon which it is superimposed."[4] In other words, differential gain is a special form of incremental gain distortion which describes the IGD of a system for the superimposed high-frequency case only.

One of the reasons for selecting the high-frequency aspect of incremental gain distortion for the Institute of Radio Engineers (IRE) definition of differential gain was applied in Fig. 58, when the ". . . high-frequency sine wave . . ." of the definition was made equal to color subcarrier. This special case of differential gain explores the system gain linearity in the vicinity of this particularly important frequency. The definition of differential gain was purposely made in the broad terms of a ". . .high frequency sine wave . . ." to allow the greatest possible versatility in devising methods of measurement. In present color-television practice, however, the ". . . high frequency sine wave . . ." is always color subcarrier and the low-frequency signal mentioned in the definition is a 15750 Hz staircase, sine-wave, or sawtooth. The complete specifications for the signal presently used in this measurement will be found elsewhere in this article.

Another reason for emphasizing high-frequency IGD was implied previously by the sentence ". . . the signal . . . would . . . be reproduced. . . with the only apparent error being a luminance distortion . . . which, for the magnitudes shown here, would probably pass unnoticed." The magnitude shown was a 25% IGD, which is passing unnoticed, indicating that large incremental gain distortions usually cause no detectable luminance errors. Incremental gain distortion is almost too sensitive a tool to measure luminance distortions. For this purpose, simple gain distortion (compression) is more useful. Therefore, the luminance/distortion aspect of IGD was deliberately omitted from the definition of differential gain.

Incremental phase and differential phase. The phase characteristic sketched in Fig. 49 indicates that the system described by this plot will introduce a certain amount of phase shift for any given frequency. For example, it might be found that a certain system would introduce a phase shift of 60° at 2 MHz. If the system in question were perfectly linear, this 60° phase shift would be produced regardless of how the 2 MHz signal might be applied to the system.

It can be shown, however, that some systems, when presented with a signal of the type shown in Fig. 61, will introduce a delay *different* from 60°, depending on where the zero axis of the sine wave falls on the transfer characteristic of the system. For the case sketched in the figure a phase shift of 70° is drawn for the largest zero-axis displacement.

By analogy with the incremental gain and differential gain arguments above, it is possible to define three quantities which pertain to this type of distortion. These quantities are *incremental phase, incremental phase distortion, and differential phase*. It can also be shown that of the three, differential phase is the most important quantity.

Incremental phase is the least exact analogue, since it is not very similar in form to incremental gain. Incremental *gain* is a *slope*; incremental phase is simply the absolute value of phase shift. In the above system, the incremental phase was 60° or 70° (or somewhere in between), depending upon the location of the zero axis.

Incremental phase distortion, like its analogue *incremental gain distortion*, depends upon the magnitude of the error. It should be zero for a perfect system. In the system of Fig. 61, the 2 MHz signal with 70°

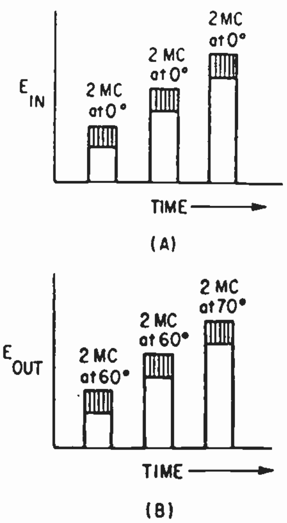


Figure 61. Graphs illustrating how a signal (A) may undergo different phase shifts (B) depending upon where the zero axis at the sine wave falls on the system transfer characteristic. This distortion is called differential phase.

incremental phase would be said to have 10 incremental phase distortion, so it is clear that the difference between two phases (one of which is assumed to be "correct") gives the incremental phase distortion.

As previously stated, *differential gain* is identical with *incremental gain distortion* for the superimposed high-frequency case only. Similarly, *differential phase* is identical with *incremental phase distortion*, but there is no need to limit the definition to the superimposed high-frequency case, since there is no other case which is meaningful for phase distortion. Without the superimposed sine wave, no phase measurement is possible. Therefore, differential phase is identical with incremental phase distortion. In practical work, the first two terms are seldom used, for the last, differential phase, has been found completely adequate to describe this aspect of a system.

In summary, the differential phase of a system is "the difference in phase shift through the system for a small high-frequency sine-wave signal at two stated levels of a low-frequency signal on which it is superimposed.[5]

It is important that the phrases "differential phase *distortion*" and "differential gain *distortion*" be avoided because differential phase *is* distortion as is differential gain, since they are defined as being identical with incremental phase distortion and incremental gain distortion, respectively. To add the word *distortion* to either is redundant. A sample of proper usage is "this amplifier has a differential gain of 1.5% and a differential phase of 0.5°."

Effect of differential phase on color picture. The phase of a subcarrier in a composite signal carries information about the *hue* of the signal at that instant. If the signal passes through a system which introduces differential phase, the subcarrier phase (and hence, the hue) at the output will become dependent upon the amplitude of the luminance associated with the hue, since it is the luminance signal which determines the location of the zero axis of the subcarrier. For example, a system introducing 10° of differential phase might be adjusted to reproduce properly a low-luminance hue such as saturated blue or a high-luminance hue such as saturated yellow, but *not both*. One or the other would have to be in error.

State of the Art

The preceding portions of this part have discussed in general terms the possible sources of color errors in a color television system. In no practical system can any of these errors be reduced to zero; therefore, anyone working with practical systems should know how nearly perfect any given parameter should be to be considered acceptable according to the present state of the art.

System Colorimetry

Talking qualitatively about colorimetric accuracy is one thing; assigning numbers and magnitudes is quite another. For the practical purposes of this part, however, we are spared the need to dig deeply into the quantitative aspects of colorimetry by one simple fact: at the present time, color errors attributable to phosphor errors, filter errors, and other basic colorimetric errors are generally small in comparison with other sources of error.

System Exponent

At the present state of the art, adjusting a system to precompensate for a kinescope exponent of 2.2 is not enforced by the Federal Communications Commission, since this parameter is not yet well established. Adjusting the system to precompensate for this median value, however, can be done with precision. A gamma corrector which uses four or five diodes to make a series of straight-line approximations to a 0.7 exponent can be made so as to have a maximum error of less than 2% of the peak signal amplitude. The exponents of the three channels can be made to match within 1% of the peak signal amplitude.

Matrix Coefficients

A high-quality matrix, such as would be found in a well-engineered color encoder or studio monitor, uses 0.5% precision resistors for all resistances which will influence the values of the coefficients, while inverters and amplifiers are either stabilized by feedback or made adjustable. Errors of greater than 1% are rare in such circuits.

While balance in the transmitter matrix, which is a special case of the subject of matrix coefficients, can usually be adjusted and held to a tolerance of the order of 0.5% of peak white.

Phase Accuracies

Adjustment of Q subcarrier, I subcarrier, and burst to within 1 of their proper relative phases is easily accomplished using standard commercial equipment and techniques. This accuracy is ten times that required by the Federal Communications Commission.

Subcarrier-Frequency Accuracy

Subcarrier frequency can be easily adjusted to within ±1 Hz the real limit on the accuracy of the adjustment being in the inherent accuracy of the standard used for frequency comparison. Long-term stability of well-engineered equipment should be easily within the required limits of ±10 Hz.

Transmission Characteristics

A single amplifier should have a gain characteristic with less than ±0.5 dB variation out to 8 MHz. Its envelope-delay error should be of the order of 0.001 μsec at 3.58 MHz, relative to 200 kHz. Differential gain of 0.5% and differential phase of 0.25 represent good performance.

Tolerable Color Errors

Sensitivity of the eye to color errors depends upon the manner in which two colors, the original and the reproduction, are compared. For example, if the two colors are placed side by side, the eye becomes a very sensitive indicator of color errors. However, if the comparison is made only by recollection or long term color memory, the eye is far more lenient in its requirements of perfect reproduction. Furthermore, if the reproduced color is one that the eye has not viewed before, the eye requires only that the color relayed to the brain be plausible, that is, that it be a reasonable color for the object.

Fortunately, side-by-side comparison of colors seldom, if ever, occurs in home viewing of color television. However, the system is frequently called upon to reproduce objects whose colors may be well known to the viewer, such as flesh tones or a sponsor's packaged product. Reproductions of these objects must be accurate enough to satisfy the viewer's recollection or color memory. If the system can satisfy the color memory of the viewer, the color-plausibility requirement will be easily met.

Investigations made to determine the sensitivity of the eye to color errors introduced by a deliberate shift of burst phase show that a shift of 10° or more produces perceptible change of hue. With color bar signals a burst phase shift of 3° can just be detected as a hue shift. With typical scenes a phase shift of 5° can be tolerated.

Tests have shown that the eye is much more tolerant of amplitude shifts in R, G, B components, which correspond to changes in color saturation, than it is of phase shifts or changes in hue.

One must distinguish between long-term adaptive errors in viewing of a color television picture and short-term differential color errors. In the first case the eye is quite tolerant of changes or shifts in color balance providing that no direct side-by-side comparisons are involved. Thus a viewer is reasonably well satisfied with color pictures in which white is reproduced within the range of 3,200° K to 9,500° K. As soon as he views two color TV pictures side-by-side at two different white balance conditions, there will be a much more critical reaction to color fidelity.

For this reason, it is important that color monitors in a broadcasting control room be adjusted to have the same effective white balance, the same color phasing, and the same peak brightness. Since such monitors are usually arranged in a row adjacent to each other, great care must be taken so that when the same picture signal is applied to all monitors, there is negligible difference in the color picture displays. Only then can the color monitors be useful in matching and comparing color balance of the various camera signal sources.

It is unusual to have more than one color receiver at a home viewing location at a given time. There the absolute color balance problem has little direct impact.

Control of short-term differential color errors is vitally important to the broadcaster. In any broadcast sequence, a given scene is generally viewed from different angles with several color cameras, at various magnifications, and the available video signals are selected from camera to camera to obtain program continuity. The eye views these color scenes in quick succession and is very critical of even small color differences, particularly with regard to skin tone rendition. Variations of the R, G, B or primary color components of 2% can be detected. Although the eye can easily adapt to any of the pictures in a few seconds, the viewer will find the abrupt color shifts very disturbing with switching transitions. Thus great care is taken with colorimetric tolerances in color cameras and with color-balancing procedures to provide color matching among cameras which will be precise.

A similar situation exists in the reproduction of color motion pictures. A feature movie having adequate color quality is usually shown in a sequence lasting 15 minutes or more, with the eye having adequate time to adapt to any discrepancies in color balance and skin tones. Commercials spliced into this feature program produce an instantaneous switch to a new and different skin tone balance without time for eye-adaptation. This transition to commercials and back to the feature can exhibit color mismatch to varying degrees, depending on the colorimetric control which has been exercised.

In fact, if the feature film is somewhat misbalanced, and intentionally "corrected" by appropriate use of R, G, B gains or "paint-pot" controls, the transition to the commercial will be more objectionable since the "correction" can then increase the misbalance, even for a "perfectly-balanced" commercial. Effort is going on in the industry to tighten up tolerances on skin tone rendition so that adequate performance can be obtained by purely routine operating methods.

Conclusion

This discussion of color errors indicates *possible* degradations in color fidelity and their probable sources. However, in a properly adjusted color TV system the picture quality is excellent. The various techniques now in development to improve picture quality within the framework of the NTSC system have assured a bright future for color TV.

THE COLOR ENCODER

The color encoder in the color television system performs the required encoding of the R, G, B signals from three-tube cameras or the R, G, B, and Y (luminance) signals from four-tube cameras into a single color video signal conforming to FCC specifications. It is the heart of the modern color television system and represents a most ingenious application of many elements of communication circuit theory.

Fig. 62 shows a block schematic of a basic color television system indicating the functions and major components of the color encoder.

A more detailed block diagram of the color encoder showing the matrixing, bandwidth-limiting and quadrature modulation functions is shown in Figs. 36 and 37.

Basic Functions

The principal operations and functions performed by the color encoder are:

1. Matrixing of R, G, B video signals to produce luminance and chrominance signals.

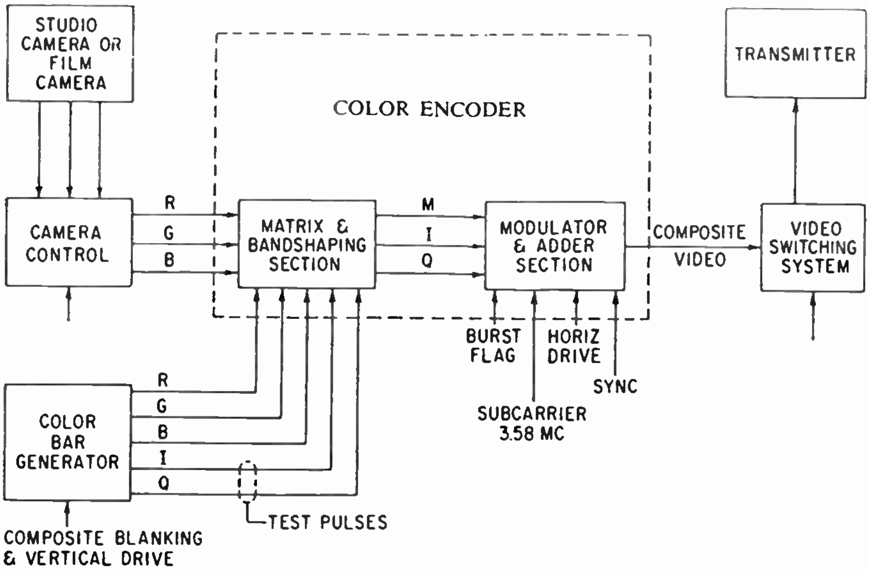


Figure 62. Basic color-television system showing functions and major components of the color encoder.

2. Filtering of the chrominance signals to obtain the required bandwidth.
3. Delay compensation to correct for band-limiting time delay.
4. Modulation of 3.58 MHz carriers by chrominance signals.
5. Insertion of color sync burst.
6. Addition of luminance and chrominance signal to form a complete color signal.
7. Optional addition of sync.

Design and system philosophy determines whether a color encoder is a separate unit or an integral portion of a modular assembly. Present solid-state equipment design tends toward the modular concept since it is generally easier to maintain, repair, update, and revise specific modular units or board assemblies without affecting the overall installation.

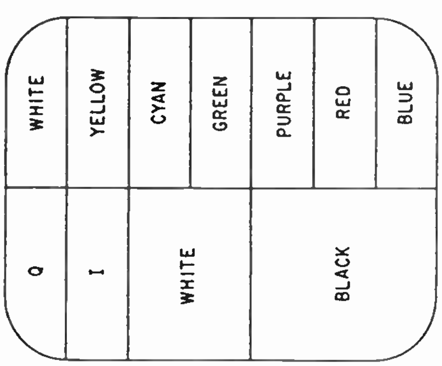


Figure 63. Diagram showing color monitor display of color and test bars electronically produced by RCA color-based generator.

The electrical color bar generator, which is generally provided for systems test and color encoder alignment, is available either as a separately contained unit or as a module in a complete operating assembly.

Color encoders of modern solid-state design are inherently stable and require only routine verification or adjustment. Set-up of a color encoder involves the use of color bars which are electrically generated waveforms of high precision. A color bar generator is capable of producing on a color monitor all of the signal bars illustrated in Fig. 63.

Colors at the top of this display pattern are arranged from left to right as white, yellow, cyan, green, magenta, red, and blue in their decreasing order of luminance. The lower portion of the pattern contains I, 100% White, Q, and black signal areas. The I and Q signals simplify subcarrier phase adjustments in the color encoder and the 100% white bar facilities white-balance adjustments. The specifications of the standard encoder color bar signal are given in EIA standard RS-189.

Waveforms

Fig. 64 shows the oscilloscope waveforms at a horizontal sweep rate of the color bar signals displayed on the television raster. Note that this is a composite representation of waveforms of the top and bottom areas of the raster. The color sync precedes the color bar pulse information.

Fig. 65 shows the various band-pass response characteristics of the luminance channel and of the "I" and "Q" channels of the color encoder.

A color encoder is set up and adjusted by using the calibrated color bars just described. The color encoder luminance gain is adjusted by using the 75% white bar as a reference. By switching off the luminance channel

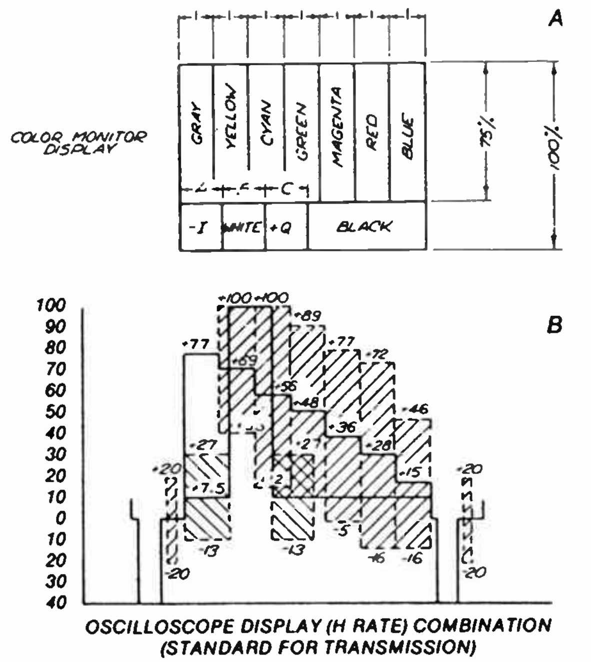


Fig. 64. (A) Color monitor display and (B) Oscilloscope display (H rate).

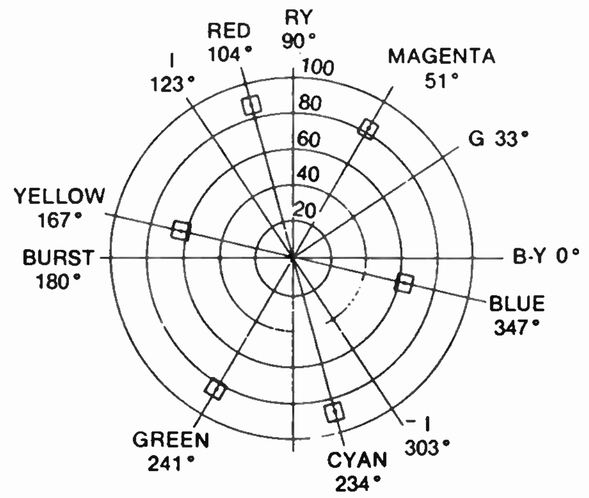


Figure 66. Vector relationship among chrominance components.

the appropriate I and Q waveforms are available to set the proper peak amplitudes and the 90° phase separation. Either a wide-band oscilliscope or a vectorscope can be used for display in a variety of specialized set-up procedures. The vector relationship of chrominance components is shown in Fig. 66.

Aperture Compensation

Aperture compensation is used in television systems to correct for the decrease in signal output at high frequencies caused by the finite-size limitations of the scanning spot or of equivalent optical lens aperture response. If one considers abrupt or black to white square-wave transitions at 400 TV lines, corresponding to 5 MHz video components, the video signal amplitudes from a Plumbicon or vidicon pick-up tube may be only 30% to 40% of the amplitude of low frequency transitions at 40 TV lines or 0.5 MHz. If the signal-to-noise ratio of the output video is good, aperture compensation to give practically 100% flat response at 5 MHz can be applied, producing subjectively sharper pictures.

Horizontal aperture correction is done by comparing the amplitude response of a given picture element with that from adjacent elements by the use of differential amplifiers and electrical delay lines. This difference, suitably amplified and of correct polarity is added to the signal being corrected which increases the sharpness of the transition.

Vertical aperture response can also decrease with increased line number and can similarly be improved by comparing the response of the picture elements on a given TV line with that of line elements preceding and following it. Differential amplifiers compare the video signals obtained from delay lines of a horizontal period (63.6 μsec) in duration with the picture elements of the TV line to be corrected.

Differences between these video responses are obtained from differential or comparison amplifiers. The difference signal is amplified and suitably added to the main signal to improve the vertical transition sharpness.

(a) (b) (c)

Figure 65. Waveforms showing response characteristics of colorplexer monochrome, I and Q channels. (*a*) Response of monochrome channel without aperture correction, marker at 8.0 Mhz (*b*) output if I filter, marker at 2.0 MHz; (*c*) output of Q filter, marker at 500 kHz.

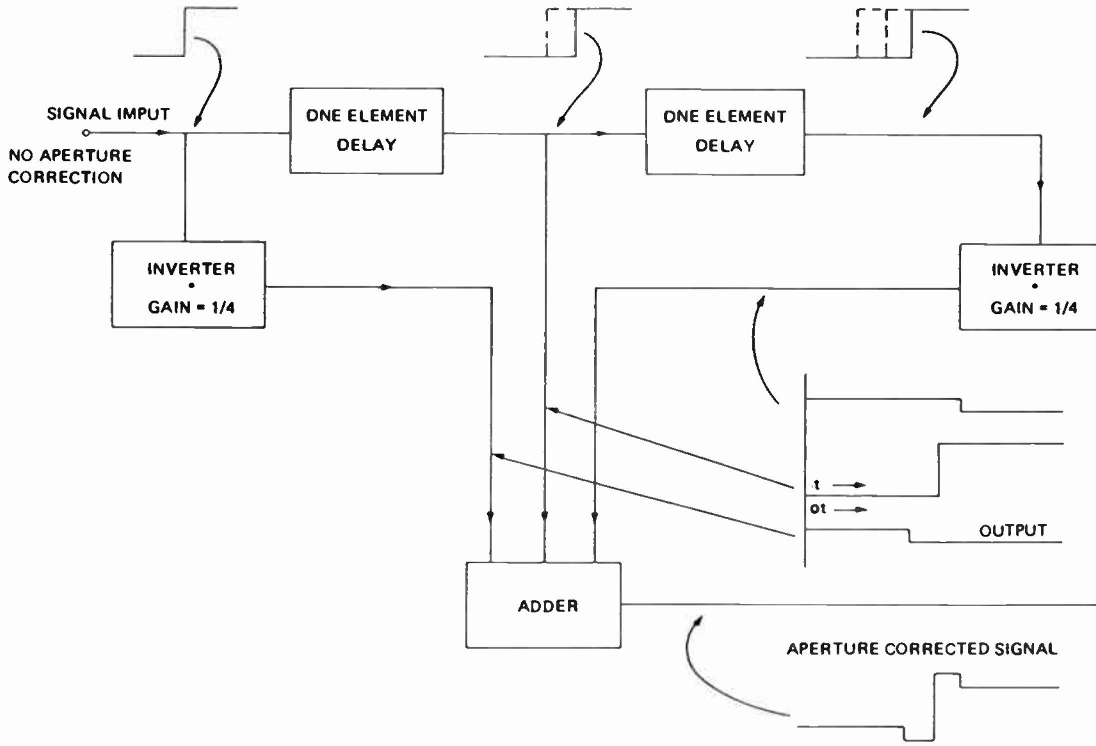


Figure 67. Generalized aperture corrector.

Judicious use of combined horizontal and vertical aperture correction or enhancement produces marked improvement in subjective picture sharpness. Since the luminance channel of a color system provides the sharpness information, it is generally used as the signal for aperture response improvement.

A block diagram of aperture compensation circuits is given in Fig. 67.

Factors Affecting Color Camera Performance

The following general principles, which outline procedures for the proper alignment and operation of color cameras, are directed toward three-tube color camera models and are presented to assist the station engineer in understanding the effect that each adjustment can have on the composite color picture. No attempt is made to present a step-by-step alignment procedure which, while basically the same for all cameras, will vary in detail depending on the manufacturer and the type of camera.

CAMERA ALIGNMENT

It is important to point out that color camera alignment should be made by viewing the proper test charts for a given adjustment or procedure. Such charts are useful for direct indication of the required camera adjustments. The practice of making an indiscriminate adjustment during a scene to "paint" a pleasing picture should be avoided in any operational procedure. Such an adjustment is usually successful for only isolated conditions and may easily produce errors in subsequent scenes. It is also important to note that certain controls in the color cameras when improperly set may give a false indication that other controls are misaligned. Therefore, maximum effort should be given to logical rigorous routine alignment of the controls before program time.

During operation a properly aligned camera should require no more than exposure control using the lens iris as an opening control and an occasional adjustment of pedestal or black level setting.

The Three-Tube Concept

A three-tube camera consists basically of an optical system which "sees" the scene being televised through a dichroic mirror or prism assembly, suitably separated into its red, green, and blue image components. These three red, green, and blue images are focused on the photosensitive layer of the pickup tube in each color channel. By synchronous scanning of the three pickup tubes, one obtains three independent video signals which differ only in their amplitude response to the three color images. Thus, we effectively obtain a red signal from the red tube, a green signal from the green tube, and a blue signal from the blue tube. With the optical and electrical adjustments available, these three pictures are superimposed or registered on each other within an accuracy of a picture element. In order to carry out this registry process one has access to individual horizontal and vertical size and centering controls, as well as to mechanical rotation of the individual yokes and to "skew" which provides for orthogonal deflection by means of electrical cross-coupling between horizontal and vertical deflection. Thus, in principle one can obtain three independent R, G, and B channels which are effectively superimposed in space at the pickup device and in time by virtue of the synchronous deflection process. One could apply these three signals to the red, green, and blue gun of the color kinescope to produce a replica of the scene being televised. However, certain procedures are necessary to obtain normalized and predictable camera behavior.

Signal-to-Noise and Sensitivity

One must set the gains of the individual video amplifiers in the R, G, B channels to a specific value. Then a given signal current from the Plumbicon will produce the required output level at the required signal-to-noise ratio depends almost entirely on the figure of merit of the external video amplifier. Nominal values of signal current are of the order of 300 to 400 nanoamps. These are obtained at exposures of approximately f:4 with 150 fc to 200 fc on an average scene. In contrast to the Image Orthicon camera where the signal-to-noise ratio is determined primarily by the signal-to-noise ratio of the image orthicon tube itself, there is a trade-off possible between sensitivity with a 6 dB decrease in signal-to-noise ratio or four times this sensitivity with a 12 dB decrease in signal-to-noise ratio. As long as the signal-to-noise ratio under standard conditions is excellent, in practice about 50 dB before gamma correction, one can tolerate such a degradation to obtain increased sensitivity and still achieve pictures which have adequate signal-to-noise.

Specular Highlights

The signal current magnitude is chosen so as to achieve a compromise between signal-to-noise ratio and the ability of the tube to discharge highlights. A standard procedure is to adjust for a factor of two reserve in the signal current by proper beam bias adjustments. In set-up, the normal scene lens exposure opening is deliberately increased by one f-stop, doubling the light to the Plumbicons® and the beam currents in the R, G, and B tubes are then adjusted to discharge just the picture highlights. The exposure is then restored to its "normal" setting. With this camera adjustment procedure any increase in peak brightness due to speculars or highlights in a scene which does not exceed this factor of two will be discharged effectively in the Plumbicon by the "available" beam current reserve which has been provided. If one attempts to use larger signal currents than 400 nanoamperes there may be limitations in the gun which cause loss or normal resolution and an inability to supply the required beam current reserve for satisfactory discharge of highlights.

When a camera is operated under conditions of specular highlights and there is motion in the scene, the presence of undischarged areas in the raster will give rise to false color halo effects, generally red, which are usually described as comet tails. This comet tail effect on motion is called "puddling" by British broadcasters. The two-to-one highlight beam reserve usually controls the comet tail effect satisfactorily.

Gamma Correction

Since the gamma of the Plumbicon® tube is essentially unity, gamma correcting amplifiers must be used to produce a pleasing picture display using modern color kinescopes. The effective gamma characteristic of the color kinescope has approximately a 2.2 exponent; thus gamma correction of $1/2.2$ or 0.45 is needed to obtain an overall gamma or transfer function of unity.

In order to obtain color tracking with changes of lighting or exposure, it is important that the transfer characteristics or gamma of the R, G, B channels be identical. This matching can be achieved by using techniques such as superimposing of the transfer characteristics waveforms on a display oscilloscope, using a standard input sawtooth, and adjusting the individual gamma circuits for the same power law and the individual black levels or capped lens references for zero. A direct check for transfer characteristic adjustment is to use the neutral EIA logarithmic gray scale chart[6] placed directly in the scene viewed by the camera. When the tube and gamma circuits are correctly adjusted to an overall gamma, which is the same for all three channels, and a 0.45 slope value is maintained, the color picture display of the EIA chart on the kinescope will be observed as neutral or shades of gray with no apparent color misbalance over the entire gray scale range.

Aperture Correction

The aperture response of Plumbicon® tubes of the 30 mm variety generally used for color TV broadcast is approximately 35% to 45% at 400 TV lines or 5 MHz as compared to a 100% reference response for low line-number transitions. For this reason it has been almost universal practice to aperture-correct or crispen the picture both horizontally and vertically by the use of omnidirectional aperture correction circuits. The response can be made effectively 100% of the time within the 5 MHz TV channel without noticeably deteriorating the signal-to-noise ratio. Such aperture correction techniques are described in the section on color encoder and shown in Fig. 66. Clamping, blanking addition, and clipping of the processing signal, followed accepted monochrome picture techniques, are performed on the three channels before they are ready to encode into the NTSC color encoded form adopted for transmission.

Color Matching Techniques

In a color television operation the color matching of the individual color cameras against each other is of prime importance. Ideally there should be *no* discernible color differences in the color TV pictures from all cameras when viewing the same subject. Experience has shown that by exercising tight control on the production tolerances of dichroic colorimetric components in the optical system and on the electronic components, one can achieve accurate color rendition from any cameras used on a given scene.

In practice each camera is aligned under normalized video gain conditions so as to obtain the required signal-to-noise performance and the same effective sensitivity. Then routine adjustment procedures to obtain the same gamma correction or transfer characteristics in the R, G, B are carried out.

The cameras now view an EIA logarithmic neutral gray scale under standard conditions. If the inputs to the color encoder are standardized and cameras have been well aligned, the gray scales will be reproduced on a color monitor over the complete brightness range as a neutral picture, since the subcarrier amplitude everywhere in the scene should be zero. Such a chart is a very sensitive indicator of small misadjustments and is generally used as a tool for vernier balancing of a color camera.

Any minor discrepancies in color rendition of the cameras used in a studio are corrected by very small changes in either R, B, or G gain provided by "paint pots."

Operationally it has been found that one camera control operator, using a single color monitor, can match four cameras more rapidly and accurately than four operators working independently.

Electronic masking devices such as the RCA Chromacomp® and the CBS Color Masking Processor permits color matching cameras to any degree of precision without upsetting white balance.

Flare in Pickup Tubes

Under certain conditions of scene content, an unwanted lift of black level or pedestal can occur in one or more of the color pickup tubes. The effect is due to light scattering in the photoconductive layer of the tube itself and is strongest in the red channel. Thus, for example, if a scene which is predominantly red is viewed by the camera, the red pedestal will rise by 3% or 4% producing a red cast in the picture. This can be corrected by manually resetting the red tube black-level control. Automatic circuits which are duty-cycle sensitive are often used to provide a good approximation to black level with changes in scene content without any operator attention. Flare in green is much less than in red and is quite negligible in the blue channel. Light scattering in optical components and lenses will also cause artificial lift of black level.

In a well-designed and well-aligned camera, color balance and color tracking are obtained automatically over a wide range of scene content and exposure.

A special opaque test pattern developed by BBC uses a "super-black" enclosure hole as a reference for black level setting in addition to the usual logarithmic gray scale for gamma checks. American broadcasters often use a square of clean black velvet as a "super-black" for flare-compensation circuit test and adjustment and as a solid black-level reference.

Lighting on the Scene

With Plumbicon® cameras the incident lighting required for studio-quality signal-to-noise picture performance generally approaches 250 fc for a lens opening of f:4. The contrast of the scene which the color camera must handle is the product of the incident light and the reflectance of the subject matter. Technically, uniform or flat lighting is easiest to handle since this limits the range to the reflectance of the scene components, generally restricted to a highest white of 60% reflectance and a lowlight of 2% to 3%, giving a range of 20 or 30 to 1 at most. The rendition in monochrome TV is as important as the rendition in color since many of the TV viewers still look at the picture in monochrome. It is therefore important to select scene materials and surfaces so as to obtain good monochrome separation in the gray scale as well as to provide colorful rendition in the final color picture. Flat lighting, as mentioned previously, is easiest to carry out, but becomes monotonous and boring from the standpoint of the producer. Any departures from flat lighting must be executed with caution. It is necessary to "fill-in" holes and deep shadows in lighting the scene to obtain results which are pleasing from the standpoint of signal-to-noise, range, and lag.

Specular or mirror reflections can be controlled by positioning of lighting and cameras or by "dull-spraying" of the surfaces responsible. Dimming is not an acceptable method of controlled scene lighting, since skin tone balance, which is the key to good performance, is very susceptible to changes in illuminant color temperature. Changes in scene lighting are generally provided by changing the total number of fixtures illuminating the set. Where skin tones are not involved, some liberty can be taken in dimming or fading.

Outdoor Broadcast Pickup

When color cameras view outdoor scenes, such as football and baseball games and other outdoor events, the subject matter and the illumination on the scene are no longer under the direct control of the broadcaster. Thus, for example, in the sunlight and in the shadows the incident illumination can vary 10 to 1, thereby increasing the effective scene range from 200 to 1 or more for a reflectance gamut of 20 to 1. In this case the broadcaster has an option of exposing for proper rendition of detail in the lowlights and compressing the highlights or adjusting for proper highlight exposure and crushing the dark portions of the scene. Fortunately, with mutiple-camera pickups used in sporting events, one can attempt to provide correct exposure for a camera scene with minimal overlap into underexposed or overexposed areas.

Specular Reflections

An annoying problem frequently met in outdoor pickup is specular reflection from shiny surfaces which can effectively direct an image of a light source or the sun itself into the pickup tube. Under such conditions there will be "tailing," "puddling," or "comet tail" effects during motion due to the fact that it is impractical in standard cameras to provide sufficient beam current to completely discharge such specular highlights. Usual practice is to provide a minimum of twice the normal peak signal reserve for beam current to take care of such specular highlights. The signals themselves, of course, are clipped electrically in the video circuits so as to avoid overload problems in transmission. New developements now underway show promise of providing relief from comet tail effects by providing a very high current discharge beam during horizontal retrace time.

Low-Light Pickup

A frequent color camera problem is the case of providing satisfactory results with insufficient or minimal light on the scene. In this case one trades signal-to-noise ratio in the camera for increased sensitivity. For example, with a reduction of 6 dB in signal-to-noise ratio, an effective gain of 2 in sensitivity can be obtained. Even a factor of 4 gain in sensitivity is quite possible with acceptable signal-to-noise performance. However, at lower values of scene lighting, lag on motion becomes a factor in obtaining satisfactory performance. Under these conditions "bias lighting" has been used experimentally in color cameras to provide increased sensitivity with reduced differential color lag on motion. A uniform "light level" applied to the red, green, and blue photocathodes of the Plumbicon® tubes so as to increase the dark current to about 8 nanoamperes provides a noteworthy improvement in build-up and decay lag performance, under low-light operating conditions.

Color TV Film Chains

It is a universal American practice to use photoconductive pickup tubes in the reproduction of color film. A powerful reason for this choice is that conventional high reliability intermittent pull-down motion picture film projectors for 16 mm and 35 mm film transport can be used. These are generally modified to convert the 24 frames per second motion picture standard to the 60 exposure fields per second required for nominal color TV standards, using the well-known 3 to 2 intermittent TV motion sequence. With a 3 to 2 intermittent film pull down, one motion picture frame is scanned by three television fields and the next picture frame is scanned by two television fields. Since each field lasts 1/60 second, the five fields take exactly 5/60 seconds or 1/12 second, which is exactly the same time as required to show two motion picture frames, 2/24 or 1/12 second. Thus, we have automatically the 24 frames to 60 field conversion needed for TV.

The use of photoconductive tubes with storage such as the vidcon or the Plumbicon® permits nonsynchronous system operation. The projector can be driven from the nominal 60 Hz house power supply even though the color TV field frequency is slightly less than 60 Hz. There is an effective tolerance of ¼ to ½ Hz in the power supply frequency before any disturbing "application bar" effects can be noticed due to the nonsynchronous operation of the projector with respect to the vertical scan rate. Experience has shown that this tolerance is entirely adequate for well-stabilized electrical power systems used in America.

In addition to 16 mm and 35 mm color film, 2" × 2" slides are used for program announcements, commercials, and special tests. It is standard practice to provide as many as 3 or even 4 different optical inputs into the same color TV film chain by the use of moving-mirror or fixed-prism multiplexing techniques. Any one of these sources can be selected for color transmission, thereby increasing the utilization of the equipment. Practically all modern color film chains use a field lens into which the image if projected. A typical film island is shown in Fig. 68 and a schematic of an optical multiplexer arrangement is shown in Fig. 69.

Network operations rely heavily on 35 mm color films for prime time programs. The local or regional stations use 16 mm color film and it is also used for news programs.

A publication titled "Color Television" contains reprints of important color TV technical papers from the *Journal of SMPTE*, and is an important reference for the background of some of the fundamental developments in color TV theory, equipment design, and practice. It also provides a reference appendix listing current standards and recommended practices for TV and motion pictures, and a comprehensive bibliography of color television papers published in the SMPTE Journal.

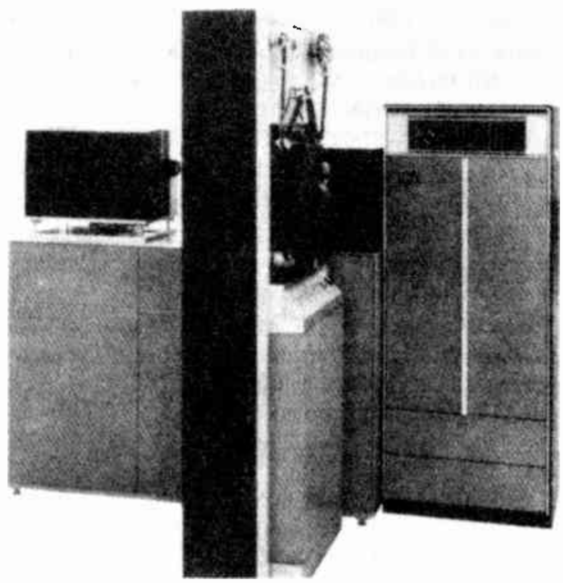


Figure 68. A typical film island.

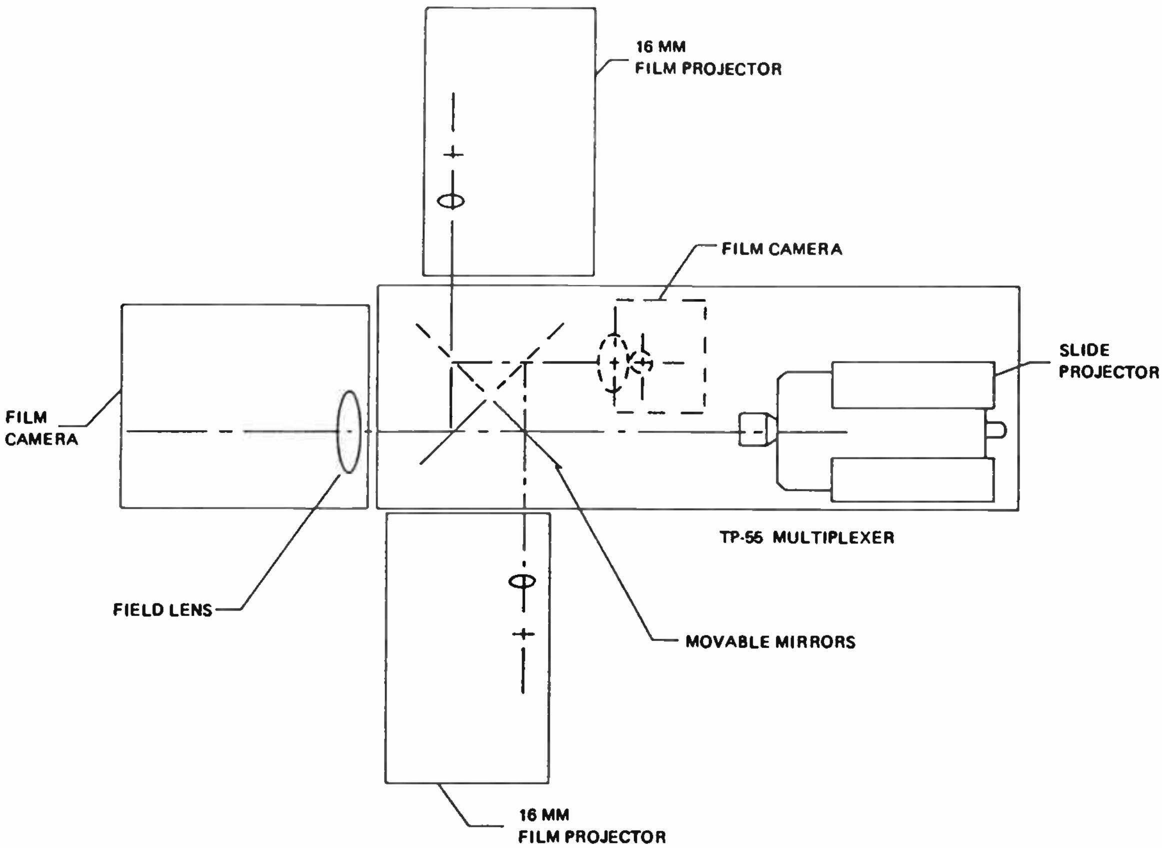


Figure 69. Schematic of an optical multiplexer.

COLOR TEST EQUIPMENT

The color television broadcast station relies heavily on specialized test and monitoring facilities in order to maintain adequate standards of performance and to ensure compliance with FCC regulations. In the early days of monochrome and color TV, the techniques and equipment were cumbersome and difficult to use on a routine basis. With the growth of the TV, test signals have become more sophisticated and yield much more useful information on the performance of monochrome and color TV systems than was previously available with a series of isolated-function measurement techniques.

A stable high-performance color monitor is an essential element of color test equipment. This, together with a vectorscope and a standard color bar generator for set-up and calibration serves as a means of evaluating performance.

The color monitor, vectorscope, and color bar generator find utilization in rapid routine day-to-day check of the television system adjustments.

Additional test equipment needed for color TV performance evaluation falls into two categories: (1) equipment to evaluate studio performance, and (2) equipment to evaluate microwave relay and transmitter performance.

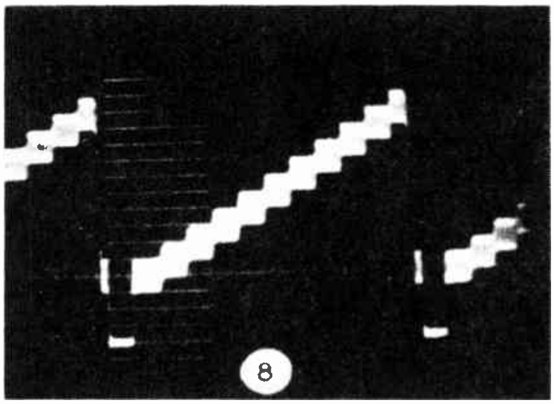
The important electrical characteristics to be measured in either category are:

- Linearity or differential gain
- Frequency response and differential phase performance
- Group delay characteristic
- Low frequency square-wave response

Evolutionary developments have followed the requirement that specific test waveforms be made available which are compatible with normal television signal systems and can be introduced easily without disabling or upsetting normal operating conditions. Measurements of such test waveforms, after passing through selected portions of the equipment or the complete system under evaluation, will give the required differential gain, phase and group delay information.

Stair-Step Generator

A modulated stair-step generator waveform is shown in Fig. 70. The signal conforms to IEEE standard IEEE 206. It consists of five 20-IRE-unit risers with subcarrier modulation on each transition. The ampli-



Fgure 70. Modulated stair-step generator waveform. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

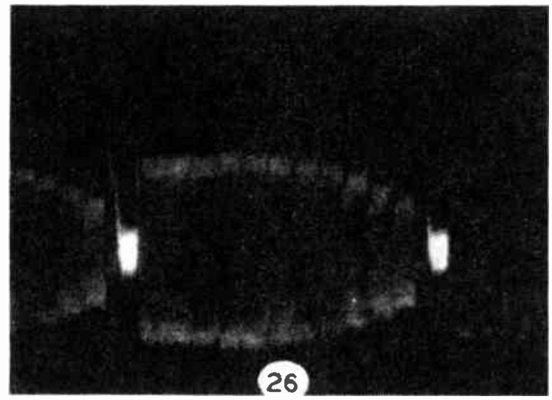


Figure 72. High pass filter output of modulated stair-step waveform showing large amount of differential gain error in amplifier under test. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

tude-linearity or differential gain response of an amplifier can be determined directly from oscilliscope measurements of the output wave display. By the use of a high-pass filter the differential gain characteristic can be displayed more graphically (Fig. 71, input), (Fig. 72, output) showing appreciable distortion. Differential phase measurements can be obtained by comparison of the subcarrier phase at each discrete level with phase of the color burst. Various oscillographic display techniques for precision phase measurements are available.

Sine-Squared Pulse and Bar

A second specialized waveform which is rapidly gaining popularity in color TV testing is the sine-squared pulse and bar with chrominance subcarrier modulation as shown in Fig. 73. It evolved from the monochrome sine-squared pulse and bar shown in Fig. 74. Use of this color test signal shows presence of differential gain distortions as in Fig 75 and delay distortions as shown in Fig. 76. Operationally the elegance of the method is in the direct-display presentation where distortion limits may be checked by reticle overlay techniques.

Another frequently used waveform is the multiburst signal, Fig. 77, which provides a series of selected frequency, constant-amplitude sine-wave electrical bursts of 0.5 MHz, 1, 2, 3, and 4 sweep signals which sequentially sample all frequencies in the video pass band. However, it is more convenient to use and to interpret in routine frequency response tests of broadcast equipment.

Vectorscope

The vectorscope[6] is a measurement instrument developed especially for color TV system tests and monitoring. Its essential feature is the polar or vectorial display of chrominance information in which the radial deflection is proportional to saturation of a color, and the angular position is equal to the phase angle of that color subcarrier with respect to the color burst. The 360° polar coordinate display corresponds to a com-

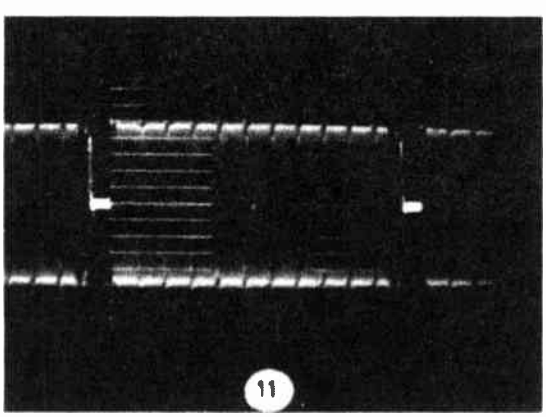


Figure 71. High pass filter output with modulated stair-step waveform input. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

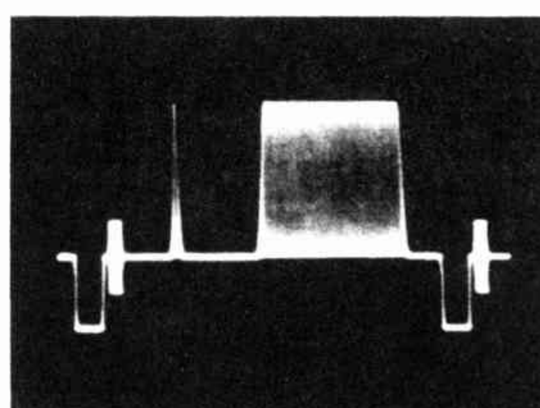


Figure 73. Combined luminance and chrominance sine-squared pulse and bar. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

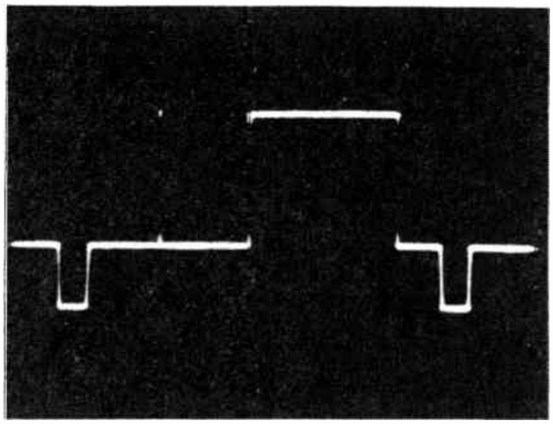


Figure 74. Monochrome sine-squared pulse and bar. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

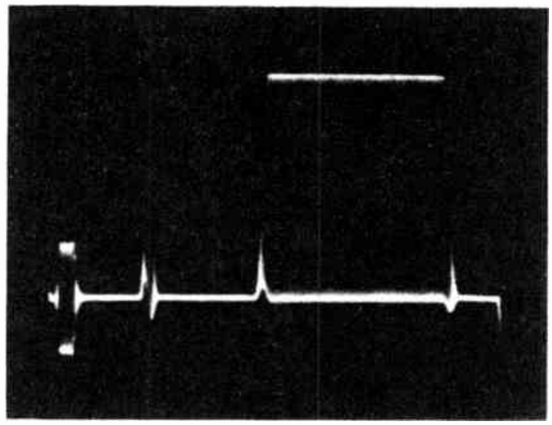


Figure 76. Delay inequality indicated by the combined luminance and chrominance sine-squared pulse and bar. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

plete cycle of color subcarrier or 280 nanoseconds in a time display. By convention, the color burst is normalized at 180°. If the color bar signal described in Figs. 64 and 66 is applied to the input to the vectorscope and the burst is normalized at 180°, the display shown in Fig. 79 is obtained on the graticule.

It is noted that for standard signal levels each color vector in the color bar sequence falls within its approximately marked box on the graticule. The outer boxes define the FCC maximum permissible errors of ±10° in phase and ±20% in amplitude. The inner boxes correspond to ±2.5 phase error and 2.5% amplitude error.

A feature of the vectorscope color bar technique is that it gives immediate reassurance on system performance with a color bar test signal display.

By alternating two signal sources at the input, one can obtain direct readings on differential phase and amplitude behavior of any selected picture sources.

Vertical Interval Reference and Test Signals

A development which has important long-range possibilities is the use of a special signal transmitted in a specific line of the vertical blanking interval. The vertical interval reference (VIR) signal, consists of a chrominance bar having the same phase as color burst, together with an appropriate luminance pulse and a black level interval. The vertical interval reference signal is added to the main video signal and is in fact a certification that at the time it is added all conditions are normal. If various distortions occur to this VIR, it can be corrected, with the expectation that the main signal will also be corrected. Thus more rigorous control and compensation of system errors is possible. A vertical interval test (VIT) signal is used to verify transmission conditions using multiburst, sine-squared or stair-step test signals. Such signals can be used for continuous monitoring of TV system performance, and in the future will probably find application in automatic control or correction of color system performance.

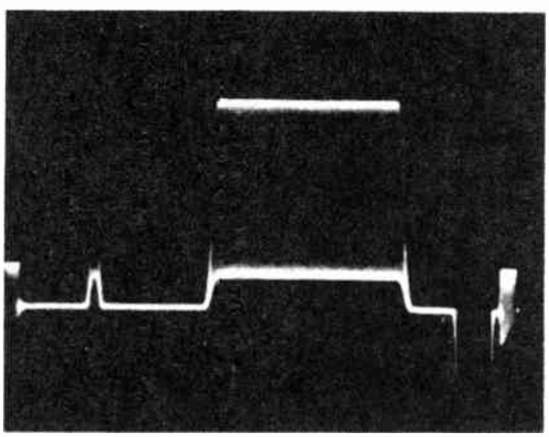


Figure 75. Gain inequality indicated by combined luminance and chrominance sine-squared pulse and bar. Compare with waveforms of Figure 73. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

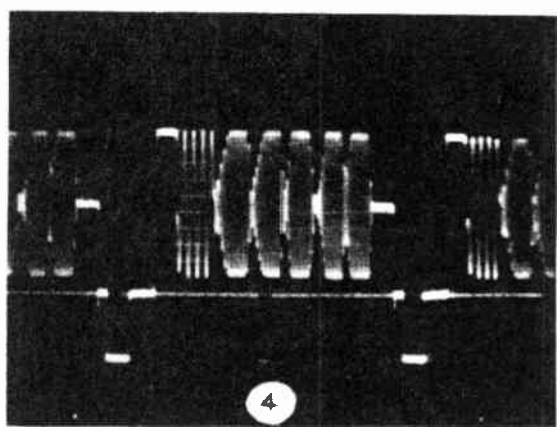


Figure 77. Multiburst test signal with burst at 0.5 MHz, 1, 2, 3, and 4 MHz. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

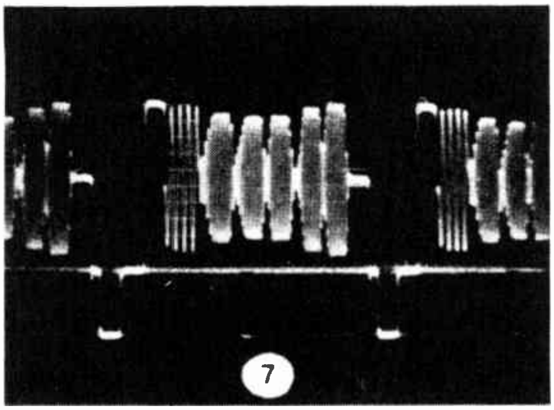


Figure 78. Multiburst output signal from amplifier having distortion. Compare with Figure 77. (Picture courtesy of Marconi Instruments, Division of English Electric Corp.)

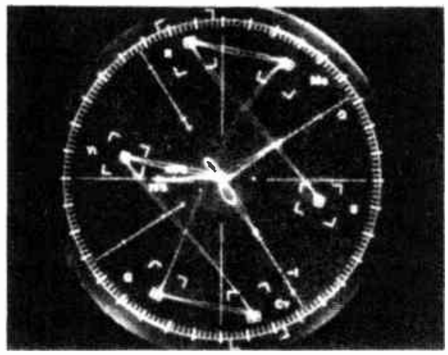


Figure 79. Vector display. Split field color bars 75 per cent amplitude 100 percent white reference, 10 per cent set up. Conforms to EIA specification RS 189. (Picture courtesy of Tektronix, Inc.)

Test Charts

There are available several pictorial charts which serve to optically generate special test signals useful for color camera alignment and system adjustment. These were developed by industry technical committees and are available as opaques from EIA or from equipment manufacturers for live cameras and as 2″ × 2″ slide from SMPTE for color TV film chains.

They are:

- EIA Resolution Chart
- EIA Linear Gray Scale Chart
- EIA Logarithmic Gray Scale
- EIA Registration Chart
- RCA Multiburst Chart
- SMPTE Resolution Slide
- Registration Slide
- Linearity Slide

The development of TV test signals and facilities is one which continually strives to increase the information to be obtained on systems performance, preferably on a continuous basis and without taking the system out of service. The VITs and VIRs concepts appear capable of providing a major step forward in test and measuring techniques.

ENDNOTES

1. A detailed discussion of colorimetry and perception, and how these factors affect the viewer, is presented in "Color Television Engineering" by John W. Wentworth, McGraw-Hill Book Company, Inc., New York, 1955.
2. If the filters in the system are of minimum-phase type, only one of the plots is needed, for either plot can be derived from the other for this type of filter. Almost all common interstage coupling networks are of the minimum-phase type.
3. FM systems can have nonlinearity as a result of *passive* networks, but this is not considered here.
4. From the definition of differential gain by IRE Subcommittee 23.4.
5. From the definition of differential phase by IRE Subcommittee 23.4.
6. Tektronics Model 520 vectorscope is widely used for these measurements.

REFERENCED TO FCC PART 73.699

	TIMING MEASUREMENTS	DEFINITION	RANGE MINIMUM	RANGE NOMINAL	RANGE MAXIMUM	EIA RS-170A TENTATIVE STANDARDS SPECIFICATIONS APPLICABLE TO STUDIO FACILITIES	NOTES
1	H SYNC WIDTH	THE HORIZONTAL SYNC PULSE IS MEASURED BETWEEN THE -4 IRE POINTS ON THE LEADING AND TRAILING EDGES OF THE WAVEFORM.	4.45 μ SEC, @-4 IRE	4.76 μ SEC, @-4 IRE	5.08 μ SEC, @-4 IRE	4.7 μ SEC ±0.1 μ SEC, @-20 IRE	FCC REQUIRES THAT THE HORIZONTAL SYNC PULSE MUST BE BETWEEN 4.45 AND 5.08 μ SEC.
2	FRONT PORCH	THE FRONT PORCH IS MEASURED BETWEEN BLANKING AND THE LEADING EDGE OF H SYNC. THIS COMPONENT IS MEASURED FROM THE +4 IRE LEVEL AT BLANKING TO THE -4 IRE LEVEL ON THE LEADING EDGE OF THE H SYNC PULSE	1.27 μ SEC, SEE NOTE	1.54 μ SEC, @-4 IRE	1.60 μ SEC, @-4 IRE	1.5 μ SEC ±0.1 μ SEC, @+4 IRE -20 IRE	THE FCC SPECIFIES THAT THE FRONT PORCH, MUST BE NO LESS THAN 1.27 μ SEC MEASURED FROM THE +4 IRE LEVEL AT BLANKING TO THE -4 IRE LEVEL ON THE LEADING EDGE OF H SYNC.
3	SYNC TO START OF VIDEO DURATION	THIS INTERVAL IS MEASURED FROM THE -4 IRE POINT ON THE LEADING EDGE OF H SYNC TO THE +4 IRE POINT ON THE TRAILING EDGE OF BLANKING.	9.22 μ SEC, SEE NOTE	9.4 μ SEC, SEE NOTE	9.61 μ SEC, SEE NOTE	9.4 μ SEC ±0.1 μ SEC, @-20 IRE +4 IRE	THE FCC SPECIFIES 0.145 H MINIMUM DURATION, FOR THIS COMPONENT.
4	SYNC TO END OF BURST DURATION	THIS SECTION IS MEASURED FROM THE -4 IRE POINT ON THE LEADING EDGE OF H SYNC TO THE TRAILING ZERO CROSSING OF THE LAST BURST EXCURSION EXCEEDING 50% OF BURST AMPLITUDE	7.07 μ SEC, SEE NOTE	7.50 μ SEC, SEE NOTE	7.94 μ SEC, SEE NOTE	7.80 μ SEC ±0.1 μ SEC, @-20 IRE TO LAST CYCLE OF BURST EXCEEDING 50% AMPLITUDE	THE FCC SPECIFIES 0.125 H MAXIMUM DURATION, FOR THIS COMPONENT.
5	H BLANKING	THE FCC DEFINES HORIZONTAL BLANKING AS MEASURED BETWEEN POINTS ON THE WAVEFORM AT +4 IRE UNITS, WITH A DURATION OF 10.49 μ SEC. THE MAXIMUM WIDTH SPECIFICATION FOR H BLANKING DEFINED BY THE FCC, IS 11.44 μ SEC MEASURED AT 90 IRE UNITS	10.49 μ SEC, @4 IRE	10.8 μ SEC, @20 IRE	11.44 μ SEC, @90 IRE	10.9 μ SEC ±0.2 μ SEC @+20 IRE	THE MAXIMUM WIDTH SPECIFICATION FOR H BLANKING IS MEASURED AT 90 IRE UNITS ABOVE BLANKING. IT IS INTERESTING TO NOTE THAT MANY VIDEO SIGNALS DO NOT REACH 90 IRE IMMEDIATELY AFTER BLANKING.
6	COLOR BURST WIDTH	THE COLOR BURST IS MEASURED FROM THE LEADING ZERO CROSSING OF THE FIRST BURST EXCURSION EXCEEDING 50% OF THE BURST AMPLITUDE TO THE TRAILING ZERO CROSSING OF THE LAST BURST EXCURSION EXCEEDING 50% OF THE BURST AMPLITUDE	8 CYCLES	9 CYCLES	10 CYCLES	9 CYCLES	FCC STANDARDS REQUIRE A MINIMUM OF 8 CYCLES OF COLOR BURST. THE NEW STANDARD WHICH IS NOW BEING USED IN THE INDUSTRY IS 9 CYCLES OF COLOR BURST.
7	BREEZEWAY WIDTH	THE BREEZEWAY IS DEFINED AS THE PERIOD BETWEEN THE TRAILING EDGE OF THE HORIZONTAL SYNC PULSE AND THE FIRST CYCLE OF COLOR BURST. THIS IS MEASURED FROM THE 4 IRE POINT ON THE TRAILING EDGE OF H SYNC TO THE LEADING ZERO CROSSING OF THE FIRST BURST EXCURSION EXCEEDING 50% AMPLITUDE	381 nSEC, SEE NOTE	600 nSEC, SEE NOTE	900 nSEC, SEE NOTE	600 nSEC ±100 nS, @-4 IRE TO FIRST CYCLE OF COLOR BURST EXCEEDING 50% AMPLITUDE	THE FCC SPECIFIES THAT THE BREEZEWAY MUST BE NO LESS THAN 381 μ SEC, MEASURED FROM THE TRAILING EDGE OF HORIZONTAL SYNC PULSE A -4 IRE POINT, TO THE FIRST CYCLE OF COLOR BURST.
8	SUBCARRIER FREQUENCY	THE FREQUENCY OF THE COLOR SUBCARRIER, 3.579545 MHZ, IS AN ODD MULTIPLE OF HALF-LINE FREQUENCY, I.E., HORIZONTAL LINE FREQUENCY/2. $\frac{15734.264\ HZ}{2}$ = 7867.132 FOR NTSC COLOR TELEVISION. (7867.132) (455) = 3.579545 MHZ	3.579535 MHZ	3.579545 MHZ	3.579555 MHZ	3.579545 MHZ, ±10 HZ	THE FREQUENCY OF THE COLOR SUBCARRIER MUST BE HELD WITHIN 10 HZ OF 3.579545 MHZ. NOTE. SHORT TIME DURATION OF BURST SIGNAL MAKES DIRECT FREQUENCY COUNTING INACCURATE
9	H SYNC RISE AND FALL TIMES	THE RISE AND FALL TIMES OF THE HORIZONTAL SYNC PULSE ARE MEASURED BETWEEN THE 10 AND 90% POINTS ON THE LEADING AND TRAILING EDGES OF THE WAVEFORM.	———	<250 nSEC, SEE NOTE	250 nSEC, SEE NOTE	0.14 μ SEC ±0.02 μ SEC MEASURED BETWEEN 10 AND 90% POINTS ON THE LEADING AND TRAILING EDGES OF THE PULSE	FCC REQUIRES THAT RISE TIME OF H SYNC BE LESS THAN 0.250 μ SEC ASSUMING 40IRE UNITS OF H SYNC AMPLITUDE. 10% CORRESPONDS TO -4IRE AND 90% CORRESPONDS TO -36IRE
10	SERRATION WIDTH	THE SERRATIONS ARE LOCATED IN THE VERTICAL SYNC PULSE AND MEASURED AT THE 4 IRE POINTS	3.81 μ SEC, -4 IRE SEE NOTE	4.45 μ SEC, @-4 IRE SEE NOTE	5.08 μ SEC, @-4 IRE SEE NOTE	4.7 μ SEC ±0.1 μ SEC, @-20 IRE	THE SERRATIONS IN THE V SYNC PULSE MUST BE BETWEEN 3.8 AND 5.08 μ SEC, MEASURED AT 4 IRE LEVEL. THE RISE AND FALL TIMES OF THE SERRATIONS MUST BE LESS THAN 250 nSEC.
11	RISE AND FALL TIMES OF EQUALIZERS AND SERRATIONS	THE RISE AND FALL TIMES OF THE EQUALIZERS AND SERRATIONS ARE MEASURED BETWEEN THE 10 AND 90% POINTS ON THE LEADING AND TRAILING EDGES OF THE WAVEFORMS	———	250 nSEC, SEE NOTE	250 nSEC, SEE NOTE	0.14 μ SEC ±0.02 μ SEC MEASURED BETWEEN 10 AND 90% POINTS ON THE LEADING AND TRAILING EDGES OF THE PULSE	FCC SPECIFIES RISE AND FALL TIMES OF EQUALIZERS AND SERRATIONS MUST BE LESS THAN 250 nSEC. THE 10% AND 90% POINTS CORRESPOND TO 4 IRE AND 36 IRE UNITS, RESPECTIVELY.
12	EQUALIZING PULSE WIDTH	THE EQUALIZING PULSES PRECEDE AND FOLLOW THE VERTICAL SYNC PULSE INTERVAL. THEY ARE MEASURED AT THE 4 IRE LEVEL	2.00 μ SEC, @-4 IRE	2.26 μ SEC, @-4 IRE	2.54 μ SEC, @-4 IRE	2.3 nS ± 0.1 μ SEC, @-20 IRE	THE TOLERANCE ON EQUALIZING PULSES, IS THAT THE AREA OF THESE PULSES MUST BE BETWEEN 45 AND 50% OF THE AREA OF THE HORIZONTAL SYNCHRONIZING PULSE.
13	V BLANKING WIDTH	VERTICAL BLANKING INTERVAL IS THE TIME BETWEEN THE LAST PICTURE INFORMATION AT THE BOTTOM OF ONE FIELD TO THE FIRST PICTURE INFORMATION AT THE TOP OF THE NEXT FIELD.	19 LINES	20 LINES	21 LINES	FIELD I = 20 LINES FIELD II = 19.5 LINES	IN TERMS OF TIME, VERTICAL BLANKING MUST BE GREATER THAN 1.17 mSEC, BUT LESS THAN 1.33 mSEC. NOTE. THE BROADCAST INDUSTRY NOW FAVORS A MAXIMUM OF 20 LINES OF VERTICAL BLANKING
14	V SYNC PULSE INTERVAL	THE VERTICAL SYNC PULSE SHOULD HAVE A DURATION EQUAL TO THREE HORIZONTAL LINES. THE SERRATIONS IN THE VERTICAL SYNC PULSE SHOULD BE BETWEEN 3.8 TO 5.1 μ SEC, MEASURED AT -4 IRE LEVEL.	———	3H	———	3H H = 1 HORIZONTAL LINE H = 63.55 μ SEC	FCC REQUIRES THE VERTICAL SYNC PULSE TO HAVE A DURATION OF THREE HORIZONTAL LINES IN TERM OF TIME, 3 H = 190.67 μ SEC. THE VERTICAL SYNC PULSE IS TO BE EXACTLY 3 HORIZONTAL LINES.

Developed and Prepared By:
MICHAEL J. SCHUMILA
Senior Member Engineering Staff
Systems Engineering
RCA American Communications
Princeton, New Jersey 08540

FCC time-base specifications for NTSC-M color television.

7.3 Color Television

Part II: Worldwide Color Television Standards—Similarities and Differences*

D. H. Pritchard
J. J. Gibson
RCA Laboratories, David Sarnoff Research Center, Princeton, New Jersey

INTRODUCTION

A simple, concise summary of the similarities and difference in the ever-changing color television system techniques and standards employed throughout the world is difficult to achieve, as evidenced by the efforts of the CCIR (International Radio Consultative Committee) in attempting to establish the elusive "universal" system. Nevertheless, it is hoped that this tutorial review and update may be useful for those who desire a conceptual overview of the technical situation.

The picture performance of a motion picture system in one location in the world is generally the same as in any other location. Thus, international exchange of film programming is comparatively straightforward.

Not so in the case of broadcast color television systems. The lack of compatibility has its origins in many factors such as constraints in communications channel allocations and techniques, differences in local power source characteristics, network requirements, pickup and display technology, and political considerations relating to international telecommunications agreements. The intent of this paper is to provide a tutorial review of the technical standardization characteristics pertinent to the problem of international exchange of images—not a system's performance comparison.

BACKGROUND

The most outstanding as well as the most controversial effort of the Eleventh Plenary Assembly of the CCIR, held in Oslo in 1966, was an attempt at standardization of color television systems by the contributing countries of the world. The discussions pertaining to the possibility of a universal system proved inconclusive. Therefore, the CCIR, instead of issuing a unanimous recommendation for a single system, was forced to issue only a report describing the characteristics and recommendations for a variety of proposed systems. It was, therefore, left to the controlling organizations of the individual countries to make their own choice as to which standard to adopt.

This outcome was not totally surprising since one of the primary requirements for any color television system is compatibility with a coexisting monochrome system. In many cases, the monochrome standards already existed and were dictated by such factors as local power line frequencies (relevant to field and frame rates) as well as radio frequency channel allocations and pertinent telecommunications agreements.

Thus, such technical factors as line number, field rate, video bandwidth, modulation technique, and sound carrier frequencies were predetermined and varied in many regions of the world. The ease with which international exchange program material may be accomplished is thereby hampered and is accomplished at present by means of standards conversion techniques, or "transcoders," with varying degrees of loss in quality.

On the other hand, these techniques have provided surprisingly good service in more recent years with the growing use of satellite relays coupled with the advances in digital signal processing technology in both the video and audio domains. In view of this rapidly changing situation and considering that more and more countries, particularly in Latin America (Region II), arrive at the point of deciding on a color system, it becomes apparent that a clear understanding of the implications of system variations has a high order of priority for those involved in international live television broadcast and film/videotape programming exchange.

* Reprinted wih permission of the Society of Motion Picture and Television Engineers. This was originally published in the SMPTE Journal.

The data quoted herein is referenced to the recent recommendations of the Fourteenth Plenary Assembly of the CCIR held in Kyoto, Japan in 1978. It should be recognized that the situation is of a continually shifting nature and future decisions can and no doubt will alter some of the details.

MONOCHROME COMPATIBLE COLOR TV SYSTEMS

In order to achieve success in the introduction of a color television system, it is essential that the color system be fully compatible with the existing black-and-white system. That is, monochrome receivers must be able to produce high-quality black-and-white images from a color broadcast, and color receivers must produce high-quality black-and-white images from monochrome broadcasts. The first such color television system to be placed into commercial broadcast service was developed in the United States. On 17 December, 1953, the Federal Communications Commission (FCC) approved transmission standards and authorized broadcasters, as of 23 January, 1954, to provide regular service to the public under these standards. This decision was the culmination of the work of the NTSC (National Television System Committee) upon whose recommendation the FCC action was based.[1] Subsequently, this system, now referred to as the NTSC system, was adopted by Canada, Japan, Mexico, and others.

That 26 years later, in 1980, these standards are still providing color television service of good quality testifies to the validity and applicability of the fundamental principles underlying the choice of specific techniques and numerical standards.

The previous existence of monochrome television standards was two-edged in that it provided a foundation upon which to build the necessary innovative techniques while simultaneously imposing the requirement of compatibility. Within this framework, an underlying theme—that which the eye does not see does not need to be transmitted nor reproduced—set the stage for a variety of fascinating developments in what has been characterized as an "economy of representation."[1]

The countries of Europe delayed the adoption of a color television system, and in the years between 1953 and 1967, a number of alternative systems that were compatible with the 625-line, 50-field existing monochrome systems were devised. The development of these systems was to some extent influenced by the fact that the technology necessary to implement some of the NTSC requirements was still in its infancy. Thus, many of the differences between the NTSC and other systems are due to technological rather than fundamental theoretical considerations.

Most of the basic techniques of NTSC are incorporated into the other system approaches. For example, the use of wideband luminance and relatively narrowband chrominance, following the teachings of the principle of "mixed highs," is involved in all systems. Similarly, the concept of providing horizontal interlace for reducing the visibility of the color subcarrier(s) is followed in all approaches. This feature is required to reduce the visibility of signals carrying color information that are contained within the same frequency range as the coexisting monochrome signal, thus maintaining a high order of compatibility.

An early system that received approval was one proposed by Henri de France of the Compagnie de Television of Paris. It was argued that if color could be relatively band-limited in the horizontal direction, it could also be band-limited in the vertical direction. Thus, the two pieces of coloring information (hue and saturation) that need to be added to the one piece of monochrome information (brightness) could be transmitted as subcarrier modulation that is sequentially transmitted on alternate lines—thereby avoiding the possibility of unwanted crosstalk between color signal components. Thus, at the receiver, a one-line memory, commonly referred to as a 1-*H* delay element, must be employed to store one line to then be concurrent with the following line. Then a linear matrix of the red and blue signal components (R and B) is used to produce the third green component (G). Of course, this necessitates the addition of a line-switching identification technique. Such an approach, designated as SECAM (SEquential Couleur Avec Memoire, for sequential color with memory) was developed and officially adopted by France and the USSR, and broadcast service began in France in 1967.

The implementation technique of a 1-H delay element led to the development, largely through the efforts of Walter Bruch of Telefunken Company, of the Phase Alternation Line (PAL) system. This approach was aimed at overcoming an implementation problem of NTSC that requires a high order of phase and amplitude integrity (skew-symmetry) of the transmission path characteristics around the color subcarrier to prevent color quadrature distortion. The line-by-line alternation of the phase of one of the color signal components averages any colorimetric distortions to the observer's eye to that of the correct value. The system in its simplest form (simple PAL), however, results in line flicker (Hanover bars). The use of a 1-H delay device in the receiver greatly alleviates this problem (standard PAL). PAL systems also require a line identification technique.

The standard PAL system has been adopted by numerous countries in continental Europe as well as in the United Kingdom. Public broadcasting began in 1967 in Germany and the United Kingdom using two slightly different variants of the PAL system (to be described shortly).

NTSC, PAL, AND SECAM SYSTEMS OVERVIEW

In order to properly understand the similarities and differences in systems used today, a familiarization with the basic principles of NTSC, PAL, and SECAM is required. As previously stated, because many basic techniques of NTSC are involved in PAL and SECAM, a thorough knowledge of NTSC is necessary in order to understand PAL and SECAM.

The same R, G, and B pickup devices and the same three primary color display devices are used in all systems. The basic camera function is to analyze the spectral distribution of the light from the scene in terms of its red, green, and blue components on a point-by-point basis as determined by the scanning rates. The three resulting electrical signals must then be transmitted over a band-limited communications channel to control the three-color display device to make the *perceived* color at the receiver appear essentially the same as the *perceived* color at the scene.

It is useful to define color as a psycho-physical property of light—specifically, as the combination of those characteristics of light that produces the sensations of brightness, hue, and saturation as shown in Fig. 1. Brightness refers to the relative intensity; hue refers to that attribute of color that allows separation into spectral groups perceived as red, green, yellow, etc. (in scientific terms, the dominant wavelength); and saturation is the degree to which a color deviates from a neutral gray of the same brightness—the degree to which it is "pure," or "pastel," or "vivid," etc. These three characteristics represent the total information necessary to define and/or recreate a specific color stimulus.

This concept is useful to communication engineers in developing encoding and decoding techniques to efficiently compress the required information within a given channel bandwidth and to subsequently recombine the specific color signal values in the proper proportions at the reproducer. The NTSC color standards define the first commercially broadcast process for achieving this result.

A preferred signal arrangement was developed that resulted in reciprocal compatibility with monochrome pictures and is transmitted within the existing monochrome channel as shown in Fig. 2. Thus, one signal (luminance) is chosen in all approaches to occupy the wide-band portion of the channel and to convey the *brightness* as well as the detail information content. A second signal, termed the chrominance signal, representative of the chromatic attributes of *hue* and *saturation*, is assigned less channel width in accordance with the principle that in human vision full three-color reproduction is not required over the entire range of resolution—commonly referred to as the "mixed-highs principle."

Another fundamental principle employed in all systems involves arranging the chrominance and luminance signals within the same frequency band without excessive mutual interference. Recognition that the scanning process, being equivalent to sampled-data techniques, produces signal components largely concentrated in uniformly spaced groups across the channel width, led to introduction of the concept of horizontal frequency interlace (dot interlace). The color subcarrier frequency is so chosen as to be an odd multiple of one-half the line rate (in the case of NTSC) such that the phase of the subcarrier is exactly opposite on successive scanning lines. This substantially reduces the subjective visibility of the color signal "dot" pattern components.

Thus the major differences among the three main systems of NTSC, PAL, and SECAM are in the specific modulating processes used for encoding and transmitting the chrominance information. The similarities and differences are briefly summarized in Fig. 3.

The following four sections discuss the basic color television systems in some technical detail, including some never actually implemented. For a summary and comparisons of system standards and specifications, the reader is referred to a later section.

THE NTSC COLOR SYSTEM

The importance of the colorimetric concepts of brightness, hue, and saturation comprising the three pieces of information necessary to analyze or recreate a specific color value becomes evident in the formation of the composite color television NTSC format.

The luminance, or monochrome, signal is formed by addition of specific proportions of the red, green, and blue signals and occupies the total available video

BRIGHTNESS (Luminance):

RELATIVE INTENSITY OF THE COLOR

HUE

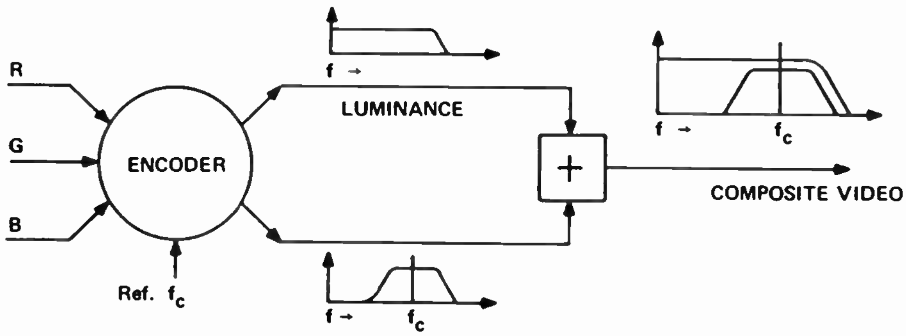
THE ATTRIBUTE THAT ALLOWS DESIGNATION IN TERMS OF RED, YELLOW, BLUE, etc. (Dominant wavelength).

SATURATION

DEGREE TO WHICH A COLOR DEVIATES FROM A NEUTRAL GRAY OF THE SAME BRIGHTNESS – PURITY, PASTEL, VIVIDNESS, etc.

Figure 1. Basic definition of "color."

- COMPATIBILITY WITH CO-EXISTING MONOCHROME SYSTEM.
- ENCODE WIDEBAND R, G, B COLOR PRIMARY SIGNALS.
  - WIDEBAND LUMINANCE (BRIGHTNESS)
  - NARROW-BAND MODULATION OF A COLOR SUBCARRIER (Hue and Saturation)



- SUBCARRIER FREQUENCY INTERLACE
  - ODD MULTIPLE OF ½ H TO REDUCE VISIBILITY OF CHROMINANCE INFORMATION SUBCARRIER.

Figure 2. Basic communications channel principles applied to color television.

bandwidth of 0–4.2 MHz. The NTSC, PAL, and SECAM systems all use the same luminance (Y) signal formation, differing only in available bandwidths.

The "Y" signal components have relative voltage values representative of the brightness sensation in the human eye. Therefore, the red (E'_R), green (E'_G), and blue (E'_B) voltage components are tailored in proportion to the standard luminosity curve at the particular values of the dominant wavelengths of the three color primaries chosen for color television. Thus, the luminance signal makeup for all systems, as normalized to white, is described by

$$E'_Y = 0.299\,E'_R + 0.587\,E'_G + 0.114\,E'_B \quad [1]^*$$

Fig. 4 also indicates the equations for the chrominance signal components. Signals representative of the

- ALL SYSTEMS:
  - THREE-PRIMARY ADDITIVE COLORIMETRIC PRINCIPLES
  - SIMILAR CAMERA PICK-UP AND RECEIVER DISPLAY TECHNOLOGY
  - WIDEBAND LUMINANCE AND NARROW-BAND COLOR
- COMPATIBILITY WITH CO-EXISTING MONOCHROME SYSTEM:

  INTRODUCES FIRST ORDER DIFFERENCES
  - LINE NUMBER
  - FIELD/FRAME RATES
  - BANDWIDTH
  - FREQUENCY ALLOCATION
- MAJOR DIFFERENCES IN COLOR TECHNIQUES
  - NTSC – PHASE AND AMPLITUDE QUADRATURE MODULATION OF INTERLACED SUBCARRIER
  - PAL – SIMILAR TO NTSC BUT WITH LINE ALTERNATION OF "V" COMPONENT
  - SECAM – FREQUENCY MODULATION OF LINE SEQUENTIAL COLOR SUBCARRIERS

Figure 3. General comparison of worldwide television systems.

LUMINANCE:

$$E'_Y = 0.299\ E'_R + 0.587\ E'_G + 0.114\ E'_B$$

(Common for all systems)

CHROMINANCE:

NTSC

$$E'_I = -0.274\ E'_G + 0.596\ E'_R - 0.322\ E'_B$$

$$E'_Q = -0.522\ E'_G + 0.211\ E'_R + 0.311\ E'_B$$

$$B\text{-}Y = 0.493\ (E'_B - E'_Y)$$

$$R\text{-}Y = 0.877\ (E'_R - E'_Y)$$

$$G\text{-}Y = 1.413\ (E'_G - E'_Y)$$

PAL

$$E'_U = 0.493\ (E'_B - E'_Y)$$

$$\pm E'_V = \pm 0.877\ (E'_R - E'_Y)$$

SECAM

$$D'_R = -1.9\ (E'_R - E'_Y)$$

$$D'_B = 1.5\ (E'_B - E'_Y)$$

Figure 4. Electronic color signal values for NTSC, PAL, and SECAM.

chromaticity information (hue and saturation) that relate to the differences between the luminance signal and the basic red, green, and blue signals are generated in a linear matrix. This new set of signals is termed *color-difference* signals and is designated as R − Y, G − Y, and B − Y. These signals modulate a subcarrier that is combined with the luminance component and passed through a common communications channel. At the receiver, the color difference signals are detected, separated, and individually added to the luminance signal in three separate paths to recreate the original R, G, and B signals according to the equations

$$\begin{aligned} E'_Y + E'_{(R-Y)} &= E'_Y + E'_R - E'_Y = E'_R \\ E'_Y + E'_{(G-Y)} &= E'_Y + E'_G - E'_Y = E'_G \\ E'_Y + E'_{(B-Y)} &= E'_Y + E'_B - E'_Y = E'_B \end{aligned} \qquad [2]$$

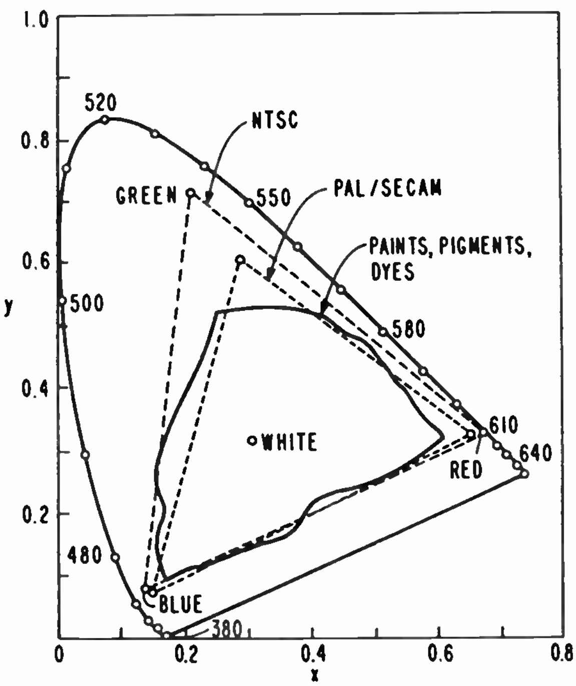
In the specific case of NTSC, two other color-difference signals, designated as I and Q, are formed at the encoder and used to modulate the color subcarrier, indicated in Fig. 4. The reason for the choice of I and Q signals is discussed later.

It may be noted that the B − Y, R − Y, and G − Y color signal modulation components are the same in NTSC, PAL, and SECAM.

Another reason for the choice of signal values in the NTSC system is that the eye is more responsive to spatial and temporal variations in luminance than it is to variations in chrominance. Therefore, the visibility of luminosity changes due to random noise and interference effects may be reduced by properly proportioning the relative chrominance gain and encoding angle values with respect to the luminance values. Thus, the "principle of constant luminance" is incorporated into the system standards.[1,2]

The voltage outputs from the three camera tubes are adjusted to be equal when a scene reference white or neutral gray object is being scanned for the color temperature of the scene ambient. Under this condition, the color subcarrier also automatically becomes zero. The colormetric values have been formulated by assuming that the reproducer will be adjusted for "Illuminant *C*," representing the color of average daylight.

Fig. 5 is a CIE chromaticity diagram (CIE = Commission Internationale de l'Eclairage) indicating the primary color coordinates for NTSC, PAL, and SECAM. It is interesting to compare the television available color gamut relative to that of all color paint, pigment, film, and dye processes.



		x	y
NTSC	R	0.67	0.33
	G	0.21	0.71
	B	0.14	0.08
PAL/SECAM	R	0.64	0.33
	G	0.29	0.60
	B	0.15	0.06
WHITE:	NTSC (ILL. C)	0.310	0.316
	PAL/SECAM (D6500)	0.313	0.329

Figure 5. CIE chromacity diagram-system comparison.

In NTSC color standards, the chrominance information is carried as simultaneous amplitude and phase modulation of a subcarrier chosen to be in the high frequency portion of the 0–4.2 MHz video band and specifically related to the scanning rates as an odd multiple of one-half the horizontal line rate as shown by the vector diagram in Fig. 6. The hue information is assigned to the instantaneous phase of the subcarrier. Saturation is determined by the *ratio* of the instantaneous amplitude of the subcarrier to that of the corresponding luminance signal amplitude value. For details of the derivation, the reader is directed to References 2, 3, and 4.

The choice of the *I* and *Q* color modulation components relates to the variation of color acuity characteristics of human color vision as a function of the field of view and spatial dimensions of objects in the scene. The color acuity of the eye decreases as the size of the viewed object is decreased and thereby occupies a small part of the field of view. Small objects, represented by frequencies above about 1.5 MHz to 2.0 MHz, produce no color sensation ("mixed-highs"). Intermediate spatial dimensions (approximately 0.5 MHz to 1.5 MHz range) are viewed satisfactorily if reproduced along a preferred orange-cyan axis. Large objects (0–0.5 MHz) require full three-color reproduction for subjectively pleasing results. Thus, the *I* and *Q* bandwidths are chosen accordingly and the preferred colorimetric reproduction axis is obtained when only the *I* signal exists by rotating the subcarrier modulation vectors by 33. In this way, the principles of "mixed-highs" and "*I*, *Q* color-acuity axis" operation are exploited.

At the encoder, the *Q* signal component is band-limited to about 0.6 MHz and is representative of the green-purple color-axis information. The *I* signal component has a bandwidth of about 1.5 MHz and contains the orange-cyan color axis information. These two signals are then used to individually modulate the color subcarrier in two balanced modulators operated in phase quadrature. The "sum products" are selected and added to form the composite chromaticity subcarrier. This signal in turn is added to the luminance signal along with the appropriate horizontal and vertical synchronizing and blanking signals to include the color-synchronization burst. The result is the total composite color video signal.

Quadrature synchronous detection is used at the receiver to identify the individual color signal components. When individually recombined with the luminance signal, the desired *R*, *G*, and *B* signals are recreated. The receiver designer is free to demodulate either at *I* or *Q* and matrix to form $B-Y$, $R-Y$, and $G-Y$, or as in nearly all present-day receivers, at $B-Y$ and $R-Y$ and maintain 500 kHz equiband color signals.

The chrominance information can be carried without loss of identity provided that the proper phase relationship is maintained between the encoding and decoding processes. This is accomplished by transmitting a reference "burst" signal consisting of eight or nine cycles of the subcarrier frequency at a specific phase

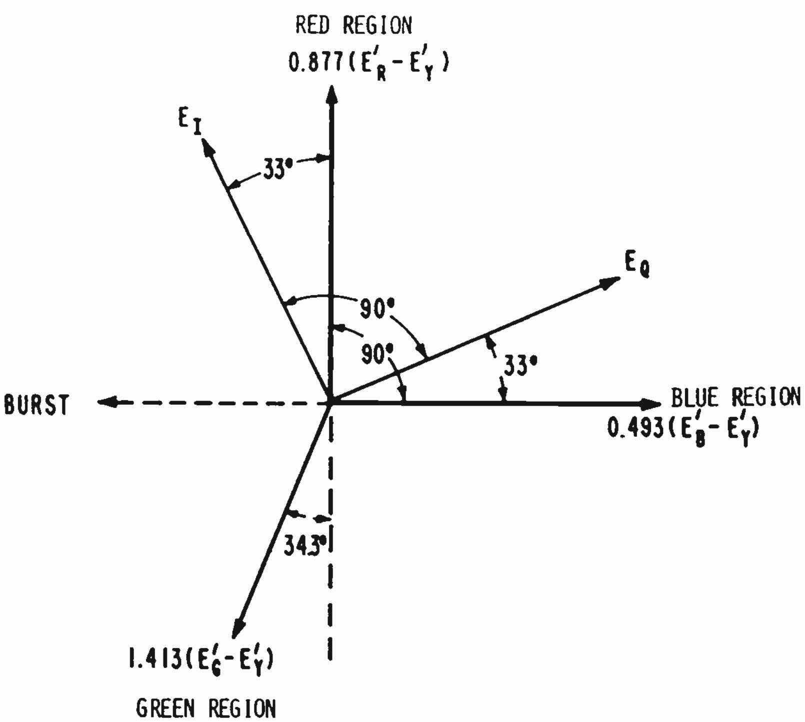


Figure 6. NTSC color modulation phase diagram.

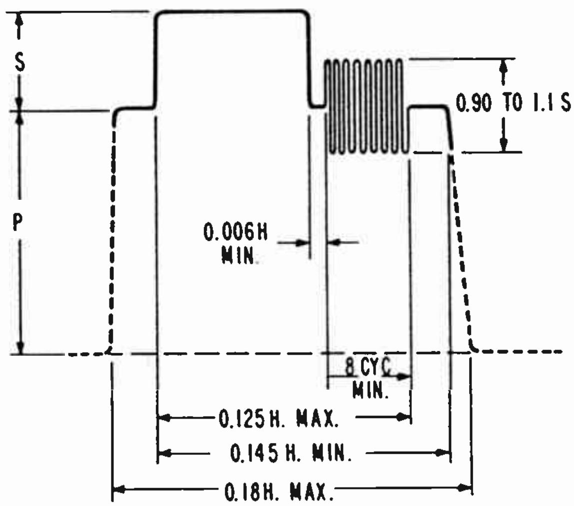


Figure 7. NTSC color burst synchronizing signal.

[$(B-Y)$] following each horizontal synchronizing pulse as shown in Fig. 7.

The specific choice of color subcarrier frequency in NTSC was dictated by at least two major factors. First, the necessity for providing horizontal interlace in order to reduce the visibility of the subcarrier requires that the frequency of the subcarrier be precisely an odd multiple of one-half the horizontal line rate. Fig. 8 shows the energy spectrum of the composite NTSC signal for a typical stationary scene. This interlace provides line-to-line phase reversal of the color subcarrier, thereby reducing its visibility (and thus improving compatibility with monochrome reception). Second, it is advantageous to also provide interlace of the beat-frequency (about 920 kHz) occurring between the color subcarrier and the average value of the sound carrier. For total compatibility reasons, the sound carrier was left unchanged at 4.5 MHz and the line number remained at 525. Thus, the resulting line scanning rate and field rate varied slightly from that of the monochrome values, but stayed within the previously existing tolerances. A good rule of thumb is that the difference is exactly one part in a thousand. The exact specifications and method of calculating the frequencies are shown in Fig. 9.

It is seen that the line rate is 15.734 kHz, the field rate is 59.94 Hz and the color subcarrier is 3.578545 MHz.

The NTSC system fundamentals have been reviewed in some detail since it was the first truly compatible system placed in commercial use and because the other systems subsequently proposed make use of most of the basic principles, differing mainly in the techniques of color encoding primarily to overcome early implementation difficulties.

PAL COLOR SYSTEM

Except for some minor details, the color encoding principles for PAL are the same as those for NTSC. However, the phase of the color signal, E_v = R−Y, is reversed by 180° from line-to-line. This is done for the purpose of averaging, or canceling certain color errors resulting from amplitude and phase distortion of the color modulation sidebands. Such distortions might occur as a result of equipment or transmission path problems.

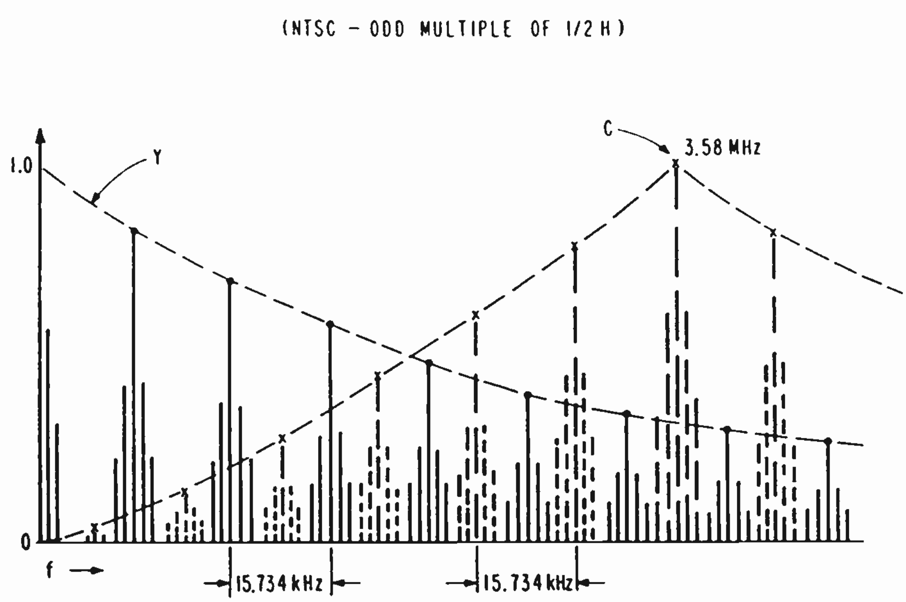


Figure 8. Luminance/chrominance horizontal frequency interface principle.

CHOICE OF EXACT FREQUENCIES

$$f_{LINE} = \frac{4.5 \times 10^6}{286} \text{ cps} \qquad f_{FIELD} = \frac{f_{LINE}}{525/2} \text{ cps}$$

$$= 15{,}734.26 \text{ cps} \qquad = 59.94 \text{ cps}$$

$$f_{SC} = \frac{13 \times 7 \times 5}{2} \times f_{LINE}$$

$$f_{SC} = \frac{455}{2} \times f_{LINE}$$

$$f_{SC} = \frac{455}{2} \times \frac{4.5 \times 10^6}{286} \text{ cps}$$

$$= \underline{3.579545} \text{ MHz}$$

Figure 9. Calculation of NTSC specific line, field, and color subcarrier frequencies.

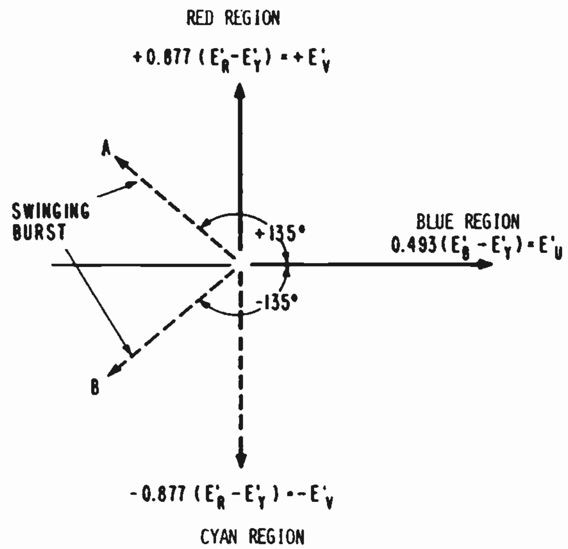


Figure 10. PAL color modulation phase diagram.

The NTSC chroma signal expression within the frequency band common to both I and Q is given by:

$$C_{NTSC} = \frac{B-Y}{2.03} \sin \omega_{sc}t + \frac{R-Y}{1.14} \cos \omega_{sc}t \qquad [3]$$

The PAL chroma signal expression is given by:

$$C_{PAL} = \frac{U}{2.03} \sin \omega_{sc}t \pm \frac{C}{1.14} \cos \omega_{sc}t \qquad [4]$$

where U and $\pm V$ have been substituted for $B-Y$ and $R-Y$ signal values, respectively.

The PAL employs equal bandwidths for the U and V color-difference signal components which are about the same as the NTSC I signal bandwidth (1.3 MHz at 3 dB). There are slight differences in the U and V bandwidth in different PAL systems due to the differences in luminance bandwidth and sound carrier frequencies as discussed later under the heading of Summary and Comparisons. Reference is made to the CCIR documents for specific details.

The "V" component was chosen for the line-by-line reversal process because it has a lower gain factor than U and therefore is less susceptible to switching rate ($\frac{1}{2} f_H$) imbalance. Fig. 10 indicates the vector diagram for the PAL quadrature modulated and line-alternating color modulation approach.

The result of the switching of the V signal phase at the line rate is that any phase errors produce complementary errors from V into the U channel. In addition, a corresponding switch of the decoder V channel results in a constant V component with complementary errors from the U channel. Thus, any line-to-line averaging process at the decoder, such as the retentivity of the eye (simple PAL) or an electronic averaging technique such as the use of a $1-H$ delay element (standard PAL), produces cancellation of the phase (hue) error and provides the correct hue but with somewhat reduced saturation; this error being subjectively much less visible.

Obviously, the PAL receiver must be provided with some means by which the V signal switching sequence may be identified. The technique employed is known as A B sync, PAL sync, or "swinging burst" and consists of alternating the phase of the reference burst by ±45° at a line rate as shown in Fig. 10. The burst is constituted from a fixed value of U phase and a switched value of V phase. Because the sign of the V burst component is the same sign as the V picture content, the necessary switching "sense" or identification information is available. At the same time, the fixed-U component is used for reference carrier synchronization.

Fig. 11 explains the degree to which horizontal frequency (dot) interlace of the color subcarrier components with the luminance components is achieved in PAL and may be summarized as follows: In NTSC, the Y components are spaced at f_H intervals due to the horizontal sampling (blanking) process. Thus, the choice of a color subcarrier whose harmonics are also separated from each other by f_H (as they are odd multiples of $\frac{1}{2} f_H$) provides a half-line offset and results in a perfect "dot" interlace pattern that moves upward. Four complete field scans are required to repeat a specific picture element "dot" position.

In PAL, the luminance components are also spaced at f_H intervals. Because the V components are switched symmetrically at half the line rate, only odd harmonics exist, with the result that the V components are spaced at intervals of f_H. They are spaced at half-line intervals from the U components which, in turn have f_H spacing intervals due to blanking. If half-line offset were used, the U components would be perfectly interlaced but the V components would coincide with Y and thus not be interlaced, creating vertical, stationary dot patterns.

For this reason, in PAL, a ¼-line offset for the subcarrier frequency is used as shown in Fig. 11. The

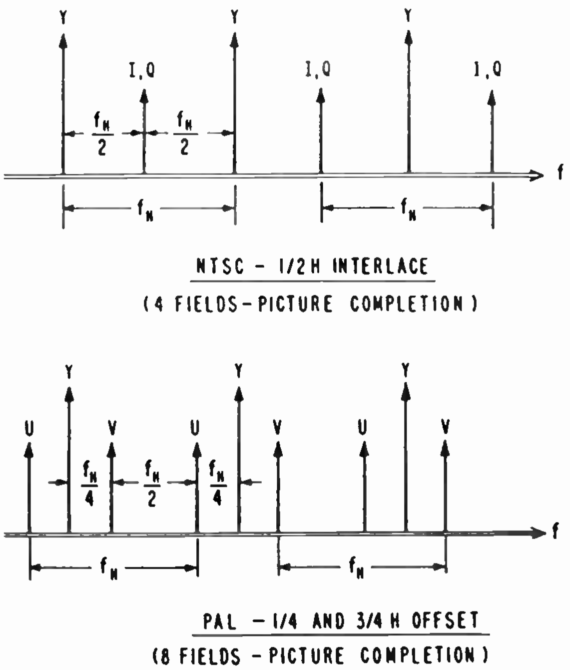


Figure 11. NTSC and PAL frequency interface relationship.

expression for determining the PAL subcarrier specific frequency for 625-line/50-field systems is given by

$$f_{SC} = \frac{1135}{4} f_H + \frac{1}{2} f_V \qquad [5]$$

The additional factor $\frac{1}{2}f_v = 25$ Hz is introduced to provide motion to the color dot pattern thereby reducing its visibility. The degree to which interlace is achieved is therefore not perfect, but is acceptable, and eight complete field scans must occur before a specific picture element "dot" position is repeated.

One additional function must be accomplished in relation to PAL color synchronization. In all systems, the burst signal is eliminated during the vertical synchronization pulse period. Because, in the case of PAL, the swinging burst phase is alternating line-by-line, some means must be provided for ensuring that the phase is the same for the first burst following vertical sync on a field-by-field basis. Therefore, the burst reinsertion time is shifted by one line at the vertical field rate by a pulse referred to as the "meander" gate. The timing of this pulse relative to the A versus B burst phase is shown in Fig. 12.

The transmitted signal specifications for PAL systems include the basic features discussed above. Although description of a great variety of receiver decoding techniques is outside the scope and intent of this paper, we should here review at least briefly the following major features: "Simple" PAL relies upon the eye to average the line-by-line color switching process and can be plagued with line beats called Hanover bars caused by the system nonlinearities introducing visible luminance changes at line rate. "Standard" PAL employs a 1-*H* delay line element to separate *U* color signal components from *V* color signal components in an averaging technique coupled with summation and subtraction functions. Hanover bars can also occur in this approach if imbalance of amplitude or phase occurs between the delayed and direct paths.

For an excellent discussion of the variety of other decoder approaches such as Chroma Lock, Super

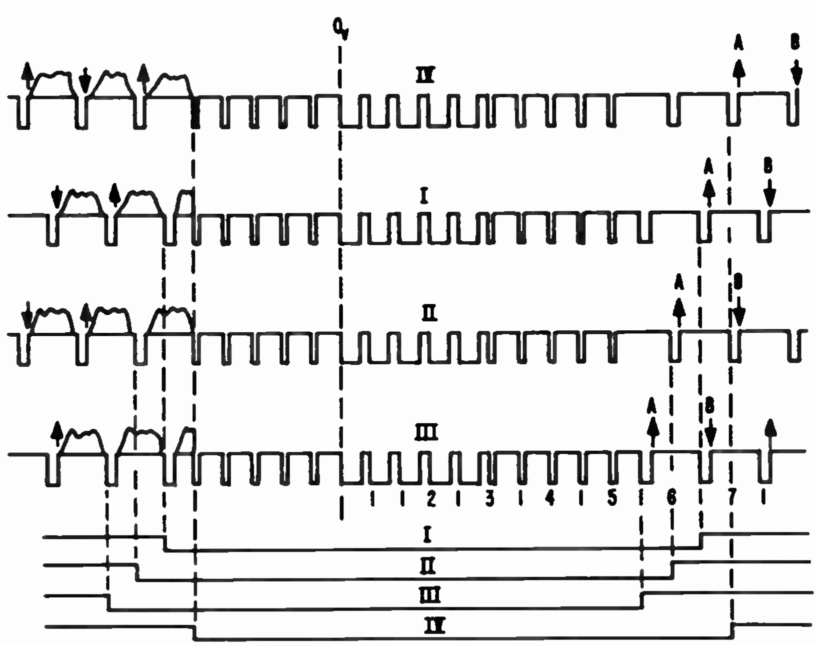


Figure 12. PAL "meander" burst blanking gate timing diagram for *B*, *G*, *H*, and *I* PAL.

PAL, New PAL, or PAL (not to be confused with N(PAL)), readers are referred to Vol. 2 of *Colour Television* by Carnt and Townsend.[4]

In a PAL system, vertical resolution in chrominance is reduced as a result of the line averaging processes. The visibility of the reduced vertical color resolution as well as the vertical time coincidence of luminance and chrominance transitions differs depending upon whether the total system, transmitter through receiver, includes one or more averaging (comb filter) processes.

Thus PAL provides a similar system to NTSC and has gained favor in many areas of the world, particularly for 625-line/50-field systems.

SECAM COLOR SYSTEM

The "optimized" SECAM system, called SECAM III, is the system adopted by France and the USSR in 1967. The SECAM method has several features in common with NTSC such as the same (E'_Y signal and the same $E'_B - E'_Y$ and $E'_R - E'_Y$ color-difference signals. However, this approach differs considerably from NTSC and PAL in the manner in which the color information is modulated onto the subcarrier(s).

First, the R − Y and B − Y color difference signals are transmitted alternately in time sequence from one successive line to the next; the luminance signal being common to every line. Since there is an odd number of lines, any given line carries R − Y information on one field and B − Y information on the next field. Second, the R − Y and B − Y color information is conveyed by frequency modulation of different subcarriers. Thus, at the decoder, a 1-H delay element, switched in time synchronization with the line switching process at the encoder, is required in order to have simultaneous existence of the B − Y and R − Y signals in a linear matrix to form the G − Y component.

The $R - Y$ signal is designated as D'_R and the B − Y signal as D'_B. The undeviated frequency for the two subcarriers, respectively, is determined by

$$f_{OB} = 272 f_H = 4.250000 \text{ MHz}$$
$$f_{OR} = 282 f_H = 4.406250 \text{ MHz} \qquad [6]$$

These frequencies represent zero color difference information (zero output from the FM discriminator), or a neutral gray object in the televised scene.

As shown in Fig. 13, the accepted convention for direction of frequency change with respect to the polarity of the color difference signal is opposite for the D_{OB} and D_{OR} signals. A positive value of D_{OR} means a decrease in frequency whereas a positive value of D_{OB} indicates an increase in frequency. This choice relates to the idea of keeping the frequencies representative of the most critical color away from the upper edge of the available bandwidth to minimize the instrumentation distortions.

The deviation for D'_R is ±280 kHz and D'_B is ±230 kHz. The maximum allowable deviation, including preemphasis, for D'_R = −506 kHz and +350 kHz while the values for D'_B = −350 kHz and +506 kHz.

Two types of preemphasis are employed simultaneously in SECAM. First, as shown in Fig. 14, a conventional type of preemphasis of the low-frequency color difference signals is introduced. The characteristic is specified to have a reference level break-point at 85 kHz (f_1) and a maximum emphasis of 2.56 dB. The expression for the characteristic is given as

$$A = \frac{1 + j(f/f_1)}{1 + j(f/3f_1)} \qquad [7]$$

A second form of preemphasis (Fig. 14) is introduced at the subcarrier level where the amplitude of the subcarrier is changed as a function of the frequency

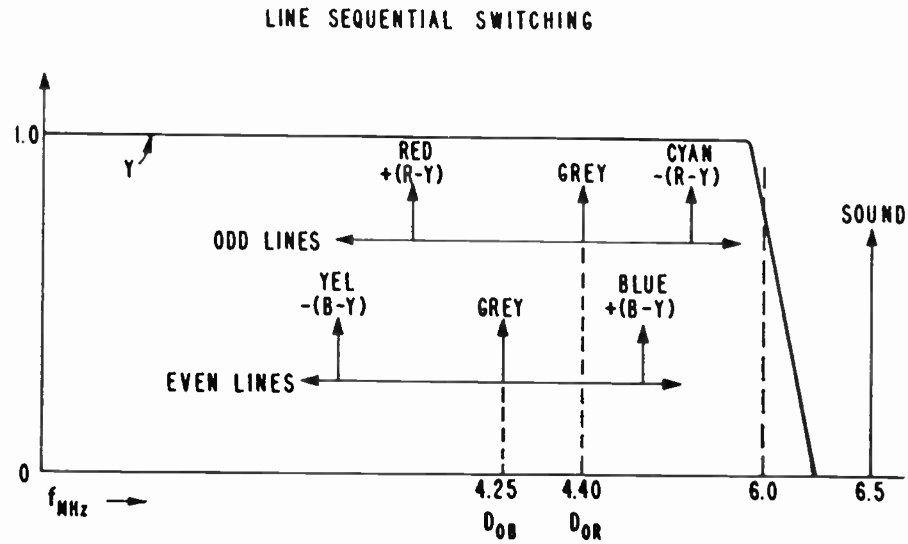


Figure 13. SECAM FM color modulation system.

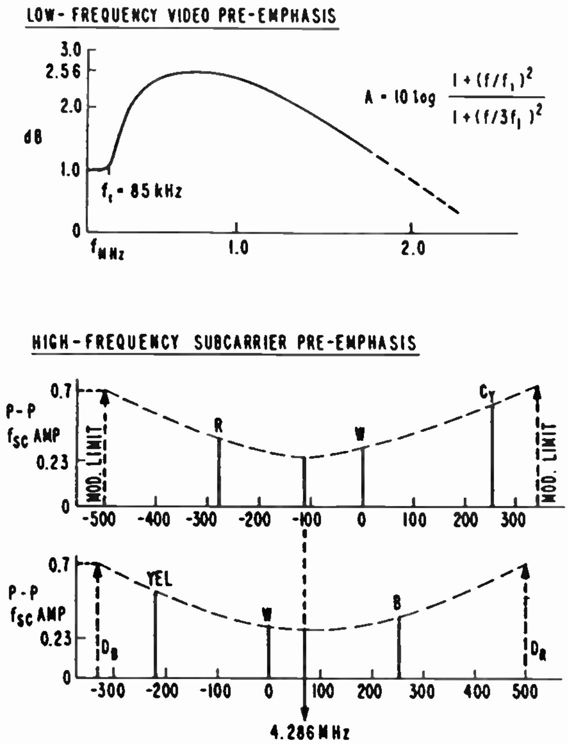


Figure 14. SECAM color signal pre-emphasis.

deviation. The expression for this inverted "bell" shaped characteristic is given as

$$G = M_0 \frac{1 \times j16\left(\frac{f}{f_c} - \frac{f_c}{f}\right)}{1 + j1.26\left(\frac{f}{f_c} - \frac{f_c}{f}\right)} \qquad [8]$$

where: f = 4.286 MHz and 2M = 23% of the luminance amplitude (100 IRE).

This type of preemphasis is intended to further reduce the visibility of the frequency modulated subcarriers in low luminance level color values and to improve the signal-to-noise ratio (SNR) in high luminance and highly saturated colors. Thus, monochrome compatibility is better for pastel average picture level objects but sacrificed somewhat in favor of SNR in saturated color areas.

Of course, precise interlace of frequency modulated subcarriers for all values of color modulation cannot occur. Nevertheless, the visibility of the interference represented by the existence of the subcarriers may be reduced somewhat by the use of two separate carriers, as is done in SECAM. Fig. 15 indicates the line-switching sequence in that at the undeviated "resting" frequency situation, the two-to-one vertical interlace in relation to the continuous color difference line-switching sequence produces adjacent line pairs of f_{OB} and f_{OR} signals. In order to further reduce the subcarrier "dot" visibility, the phase of the subcarriers (phase carries no picture information in this case) is reversed 180° on every third line and between each field. This, coupled with the "bell" preemphasis, produces a degree of monochrome compatibility considered subjectively adequate.

As in PAL, the SECAM system must provide some means for identifying the line-switching sequence between the encoding and decoding processes. This is accomplished, as shown in Fig. 16, by introducing alternate D_R and D_B color identifying signals for nine lines during the vertical blanking interval following the equalizing pulses after vertical sync. These "bottle" shaped signals occupy a full line each and represent the frequency deviation in each time sequence of D_B and D_R at zero luminance value. These signals can be thought of as fictitious green color that is used at the decoder to determine the line-switching sequence.

During horizontal blanking, the subcarriers are blanked and a burst of f_{OB}/f_{OR} is inserted and used as a gray level reference for the FM discriminators, to establish their proper operation at the beginning of each line.

Thus the SECAM system is a line sequential color approach using frequency modulated subcarrriers. A special identification signal is provided to identify the line-switch sequence and is especially adapted to the 625-line/50-field wideband systems available in France and the USSR.

It should be noted that SECAM, as practiced, employs amplitude modulation of the sound carrier as opposed to the FM sound modulation in other systems.

Additional Systems of Historical Interest

Of the numerous system variations proposed over the intervening years since the potential development of the NTSC system, at least two others, in addition to PAL and SECAM, should be mentioned briefly. The first of the these was ART (Additional Reference Transmission) which involved the transmission of a continuous reference pilot carrier in conjunction with a conventional NTSC color subcarrier quadrature modulation signal. A modification of this scheme involved a "multiburst" approach that utilized three color bursts, one at black level, one at intermediate gray level, and one at white level, to be used for correcting differential phase distortion.

Another system, perhaps better known, was referred to as NIR or SECAM IV. Developed in the USSR (NIR = Nauchni Issledovatelskaia Rabota or Scientific Discriminating Work), this system consists of alternating lines of (1) an NTSC-like signal using an amplitude and phase modulated subcarrier and (2) a reference signal having "U" phase to demodulate the NTSC-like signal. In the linear version the reference is unmodulated, and in the nonlinear version the amplitude of the reference signal is modulated with chrominance information.

To the author's knowledge, neither ART nor NIR were ever implemented or used for commercial broadcast. Still, they are of theoretical and historical interest; details may be found in Ref. 3.

FIELD	LINE #	COLOR	SUBCARRIER Ø	
ODD (1)	n	f_{OR}	0°	
EVEN (2)	n + 313	f_{OB}		180°
ODD (3)	n + 1	f_{OB}	0°	
EVEN (4)	n + 314	f_{OR}		0°
ODD (5)	n + 2	f_{OR}	180°	
EVEN (6)	n + 315	f_{OB}		180°
ODD (7)	n + 3	f_{OB}	0°	
EVEN (8)	n + 316	f_{OR}		180°
ODD (9)	n + 4	f_{OR}	0°	
EVEN (10)	n + 317	f_{OB}		0°
ODD (11)	n + 5	f_{OB}	180°	
EVEN (12)	n + 318	f_{OR}		180°

Note: • 2 frames (4 fields) for picture completion.
• Subcarrier interlace is field-to-field and line-to-line of same color.

Figure 15. Color vs. line and field timing relationship for SECAM.

SUMMARY AND COMPARISONS OF SYSTEMS STANDARDS AND SPECIFICATIONS

History has shown that it is apparently impossible to obtain total international agreement on "universal" television broadcasting standards. Even with the first scheduled broadcasting of monochrome television in 1936 in England, the actual telecasting started using two different systems on alternate days from the same transmitter. The Baird system was 250 lines (noninterlaced) with a 50 Hz frame rate while the EMI (Electric and Musical Industries) system was 405 lines (interlaced) with a 25 Hz frame rate.

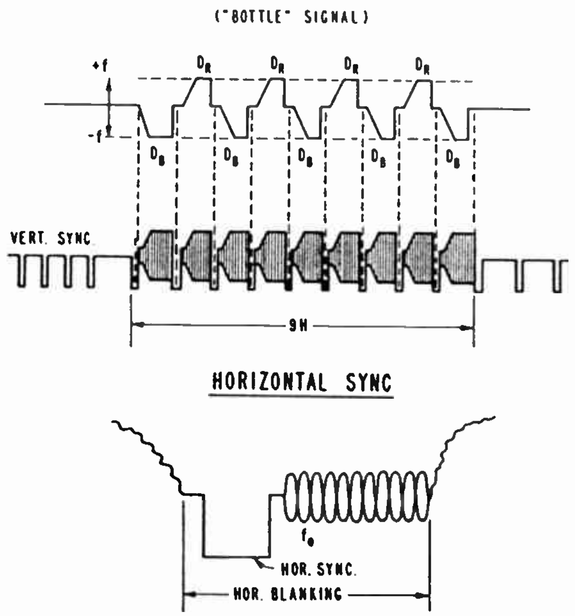


Figure 16. SECAM line identification signal.

These efforts were followed in 1939 in the United States by broadcasting a 441 line interlaced system at 60 fields per second (the Radio Manufacturers Association (RMA) system). In 1941, the NTSC initiated the present basic monochrome standards in the U.S. of 225 lines (interlaced) at 60 fields per second, designated as system M by the CCIR. In those early days, the differences in power line frequency were considered as important factors and were largely responsible for the proliferation of different line rates versus field rates as well as the wide variety of video bandwidths. However, the existence and extensive use of monochrome standards over a period of years soon made it a top-priority matter to assume reciprocal compatibility of any developing color system.

The CCIR documents[5] define recommended standards for worldwide color television systems in terms of the three basic color approaches—NTSC, PAL, and SECAM—as shown in Fig. 17. The variations—at least 13 of them—are given alphabetical letter designations; some represent major differences while others signify only very minor frequency allocation differences in channel spacings or the differences between the VHF and UHF bands. In 1978, at least 98 countries were listed as either employing or considering one or more of the proposed systems in monochrome and/or color format.

The key to understanding the CCIR designations lies in recognizing that the letters refer primarily to *local monochrome standards* for line and field rates, video channel bandwidth, and audio carrier relative frequency. Further classification in terms of the particular

THREE BASIC SYSTEMS
- NTSC
- PAL
- SECAM

THIRTEEN VARIATIONS OR SUBSYSTEMS:
A*, M, N, C*, B, G, H, I, D, K, KI, L, E*

SYSTEMS A (405 LINES), C (625 LINES) AND E (819 LINES) NOT RECOMMENDED FOR NEW SERVICE*

98 COUNTRIES LISTED BY CCIR AS EMPLOYING ONE OR MORE SYSTEMS

Figure 17. CCIR worldwide color television designations.

color system then adds to NTSC, PAL, or SECAM as appropriate. For example, the letter "M" designates a 525-line/60-field, 4.2 MHz bandwidth, 4.5 MHz sound carrier monochrome system. Thus, M(NTSC) describes a color system employing the NTSC technique for introducing the chrominance information within the constraints of the above basic monochrome signal values. Likewise, M(PAL) would indicate the same line/field rates and bandwidths but employing the PAL color subcarrier modulation approach.

In another example, the letters "I" and "G" relate to specific 625-line/50-field, 5.0 MHz or 5.5 MHz bandwidth, 5.5 or 6.0 MHz sound carrier monochrome standards. Thus, G(PAL) would describe a 625-line/50-field,5.5 Mhz bandwidth, color system utilizing the PAL color subcarrier modulation approach. The letter "L" refers to a 625-line/50-field, 6.0 MHz bandwidth system to which the SECAM color modulation method has been added (often referred to as SECAM III).

System E is an 819-line/50-field, 10 MHz bandwidth, monochrome system. This channel was used in France for early SECAM tests and for system E transmissions.

Some general comparison statements can be made about the underlying monochrome systems and existing color standards:

1. There are four different scanning standards: 405-lines/50-fields, 525-lines/60-fields, 625-lines/50-fields, and 819-lines/50-fields.
2. There are six different spacings of video-to-sound carriers, namely 3.5 MHz, 4.5 MHz, 5.5 MHz, 6.0 MHz, 6.5 MHz, and 11.15 MHz.
3. Some systems use FM and others use AM for the sound modulation.
4. Some systems use positive polarity (luminance proportional to voltage) modulation of the video carrier while others, such as the U.S. (M)NTSC system, use negative modulation.
5. As previously discussed, there are also differences in the techniques of color subcarrier encoding represented by NTSC, PAL, and SECAM, and of course, in each case there are many differences in the details of various pulse widths, timing, and tolerance standards.

It is evident that one must refer to the CCIR documents for accurate information on the combined monochrome/color standards. Fig. 18 presents a comparison of the relative bandwidths, color subcarrier frequencies, and sound carrier spacing for the major color systems used in the world today.

The signal in the M(NTSC) system occupies the least total channel width, which when the vestigial sideband plus guard bands are included, requires a minimum radio frequency channel spacing of 6 MHz. The L(III) SECAM system signal occupies greater channel space with a full 6 MHz luminance bandwidth. Signals from the tow versions of PAL also occupy greater channel widths (though less than SECAM signal) and vary in vestigial sideband width as well as color and luminance bandwidths. NTSC is the only system to incorporate the *I,Q* color acuity bandwidth variation. PAL minimizes the color quadrature phase distortion effects by line-to-line averaging, and SECAM avoids this problem by only transmitting the color components sequentially at a line-by-line rate.

Figs. 19–24 summarize, in "organization chart" form, the CCIR designations for NTSC, PAL, and SECAM basic system identifications and characteristics. In Fig. 19, M(NTSC) identified the system used in the United States, Canada, Japan, Mexico, the Philippines and several other Central American and Caribbean area countries. The N system may be implemented in color either in the NTSC or the PAL format. At present, many Latin American countries are in the process of adopting one or other version of this approach.[3] Fig. 20 provides a summary of the PAL systems. PAL

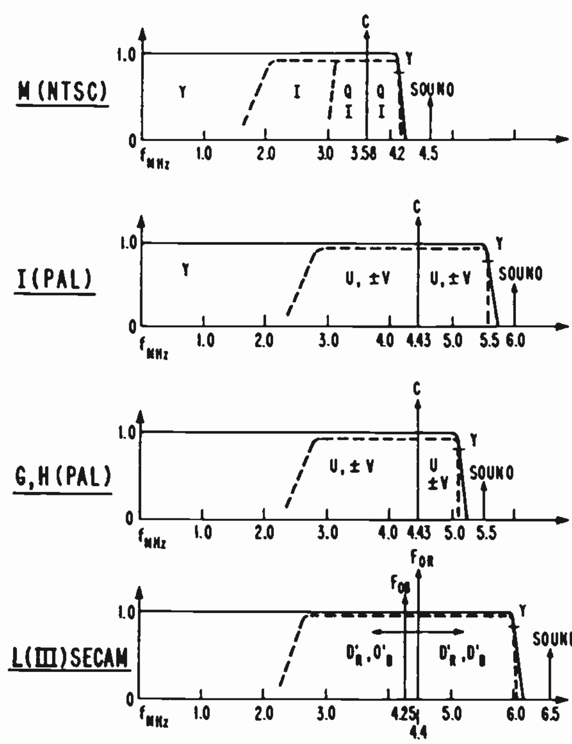


Figure 18. Bandwidth comparison between NTSC, PAL, and SECAM.

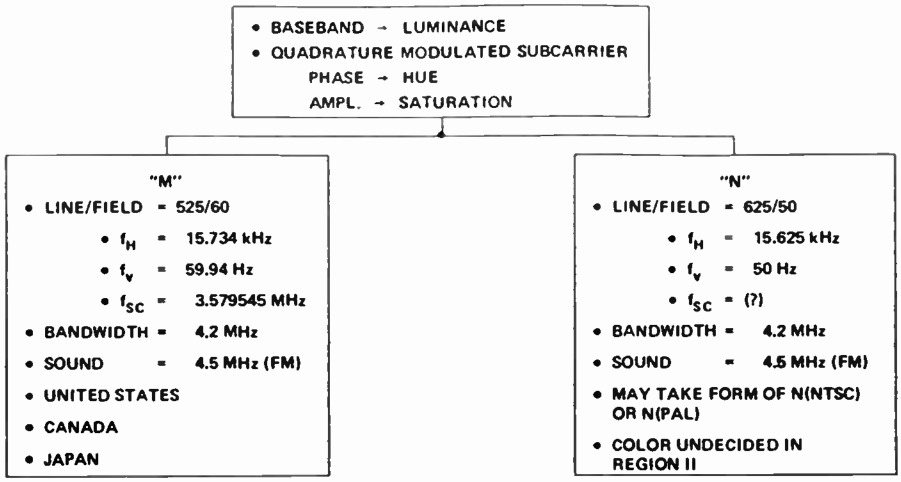


Figure 19. CCIR designation for NTSC system—summary.

systems in one or another of the 625-line formats are predominately used in Continental Europe, the United Kingdom, some African countries, and China. An "M" (525-line) version of PAL has been in use in Brazil.

Fig. 21 summarizes the SECAM III system, which is in use primarily in France and the USSR. The SECAM IV system, as a proposal,[1] almost gained favor in 1966 as a universal European approach but to the authors' knowledge has never been used for normal broadcasting. The E system, mentioned in connection with early SECAM tests in France, is limited to monochrome broadcasts and is slowly becoming extinct even in this application.

Fig. 22 (Systems Comparison Summary—Part I) provides a "summary-at-a-glance" of the major color television system general characteristics as presently practiced, whether it be monochrome only or including the addition of chrominance information. Fig. 23 (Part II) characterizes the fundamental features relating to the differences between NTSC, PAL, and SECAM in the critical areas of color encoding techniques. Similarly, Fig. 24 (Part II) indicates the color encoding line-by-line color sequence operation for the three systems.

The information conveyed in these last seven charts (Figs. 18–24) highlights the technical equalities and differences among the systems and attempts to show some kind of order as an aid to understanding the existing worldwide situation. It serves as well to point out the difficulties of entertaining the notion of a "universal" system.[7]

Comments on International Exchange of Images

The international exchange of images in broadcast television in the face of the variety of standards is difficult. It should be remembered that all TV systems, both monochrome and color, can be operated from movie film. Special television camera chains have been developed and manufactured that are capable of

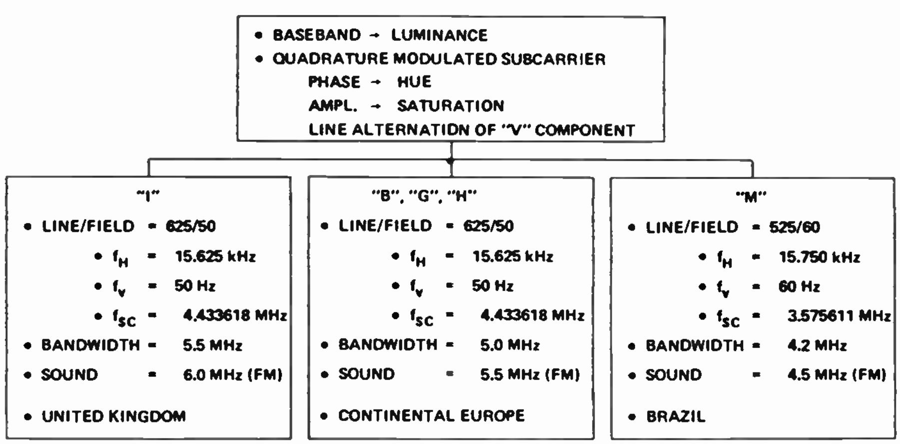


Figure 20. CCIR designation for PAL system—summary.

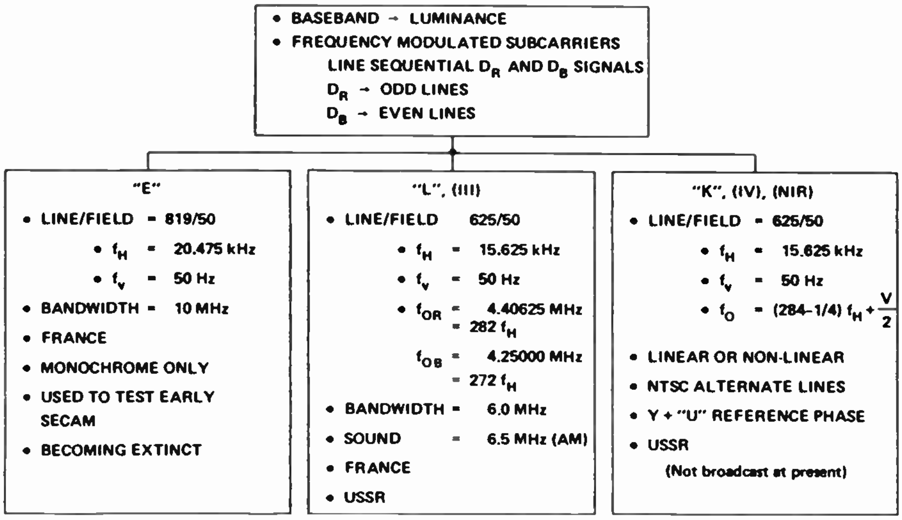


Figure 21. CCIR designation for SECAM system—summary.

	NTSC	PAL	SECAM
TV SYSTEM	M	G, I	L
FIELD RATE (f_v Hz)	59.94	50	50
TV LINES	525	625	625
LINE RATE (f_H kHz)	15.734	15.625	15.625
LUMA BANDWIDTH (MHz)	4.2	(5.0) (5.5)	6.0
SOUND (MHz)	4.5 (F 3)	(5.5) (6.0) (F 3)	6.5 (A 3)
VERTICAL INTERLACE	2:1	2:1	2:1
GAMMA	2.2	2.8	2.8
WHITE	ILL. "C" (D6500)	D6500	D6500

Figure 22. General system technical summary—Part I.

	NTSC	PAL	SECAM
COLOR SUBCARRIER (MHz)	3.579545	4.433618	4.250000 = f_{OB} 4.406500 = f_{OR}
f_{SC} MULTIPLE OF f_H	$\frac{455}{2} f_H$	$\frac{1135}{4} f_H + \frac{f_v}{2}$	272 f_H = f_{OB} 282 f_H = f_{OR}
CHROMA ENCODING	PHASE & AMP. QUAD. MOD.	PHASE & AMP. QUAD. MOD. (LINE ALTERNATION)	FREQUENCY MODULATION (LINE SEQUENTIAL)
COLOR DIFFERENCE SIGNALS	I, Q, (1.3 MHz) (0.6 MHz)	U, ± V (1.3 MHz) (1.3 Mhz)	D_R (f_{OR}) (> 1.0 MHz) D_B (f_{OB}) (> 1.0 MHz)
COLOR BURST PHASE	–(B-Y)	U and ± V	f_{OR} AND f_{OB} 180° PHASE SWITCH EVERY 3rd LINE AND EVERY FIELD
COLOR SWITCH IDENT.	NOT REQUIRED	SWINGING BURST ± 45°	9 LINES OF D_R AND D_B DURING VERTICAL INTERVAL
ADDITIONAL SIGNALS	NONE	"MEANDER" GATE $f_{H/2}$	$f_{H/2}$, $f_{H/4}$, f'_v, $f_{v/2}$

Figure 23. Chrominance encoding systems comparison—Part II.

	LINE (N)	LINE (N + 1)	LINE (N + 2)	LINE (N + 3)
NTSC:				
CHROMA:	I, Q	I, Q	I, Q	I, Q
BURST PHASE:	−(B-Y)	−(B-Y)	−(B-Y)	−(B-Y)
PAL				
CHROMA:	U, + V	U, − V	U, + V	U, − V
BURST PHASE:	−U + V = + 135°	−U − V = +225°	−U + V = + 135°	−U − V = +225°
SECAM: (FM)				
CHROMA:	D_R ± 280 kHz	D_B ± 230 kHz	D_R ± 280 kHz	D_B ± 230 kHz
BURSTS:	(D_R DEVIATION = $^{+350\text{ kHz}}_{-500\text{ kHz}}$)		(D_B DEVIATION = $^{+500\text{ kHz}}_{-350\text{ kHz}}$)	

CHROMA SWITCH IDENT. LINES DURING VERTICAL INTERVAL

LINE #:	7	8	9	10	11	12	13	14	15
	320	321	322	323	324	325	326	327	328
INDENT SIGNALS:	D_R	D_B	D_R	D_B	D_R	D_B	D_R	D_B	D_R

(NOTE: Phase reversed 180° every 3rd line and every field).

Figure 24. Line-to-line chroma signal sequence comparison—Part III.

operating at 655-lines and 48-field rate: the field rate purposely being made to be compatible with the 24-frame rate motion picture standards.

It is comparatively straightforward to exchange television program material by tape, microwave, or satellite between areas employing the same scanning rates: the video bandwidths are, of course, not equivalent but the differences do not result in major image degradation. Electronic standards converters have been developed and used for converting between 50-fields/s and 60-fields/s systems.

The direct exchange of color television programs between the three major systems is obviously more complex. Special transcoding systems have been developed to translate color subcarrier frequencies between similar color systems having different scanning rates. More complex transcoders are possible which translate from one color technique to another, although always at the cost of some degradation of resolution or reduction of performance. Even simultaneous translation between different scanning rates and different color systems, such as between 525-line NTSC and 625-line PAL, has been accomplished.

As previously stated, the advent of satellite worldwide television relay, coupled with recent advances in digital processing of television signals, has given new importance to standards conversion relative to the exchange of program material on an international basis. Thus, the intent of this worldwide color systems standards review is to highlight the similarities as well as the major differences for those who desire an overview of the related television concepts and standards. A thorough understanding of those concepts and standards by many people is essential if effective international exchange of programming is to grow.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the aid and suggestions provided by A. Lind of the RCA Broadcast Products Division, and are especially grateful for the assistance and information provided by R. Mills and E. Rutishauser of the Laboratories RCA Ltd., Zurich.

ENDNOTES

* From Equation [1]:
The signal of Equation [1] would be exactly equal to the output of a linear monochrome camera tube with ideal spectral sensitivity if the red, green, and blue camera tubes were also linear devices with theoretically correct spectral-sensitivity curves. In actual practice, the red, green, and primary signals are deliberately made nonlinear to accomplish gamma correction (adjust the slope of the input/output transfer characteristic). The prime mark (′) is used to denote a gamma-corrected signal.

** The IEEE Standard Dictionary of Electrical and Electronics Terms notes that in constant-luminance signal and no control of luminance is provided by the chrominance signal. Noise signals falling within the bandwith of the chrominance channel produce only chromaticit variations, which, if they are coarse-structured, are subjectively less objectionable than correspondingly coarse-structured luminance variations.

REFERENCES

1. J. W. Herbstreit and J. Pouliquen, "International Standards for Color Television," *IEEE Spectrum*, March 1967.
2. Donald G. Fink (Ed.), *Color Television Standards*, McGraw-Hill Book Co., New York, 1955.
3. D. H. Pritchard, "U.S. Color Television Fundamentals—A Review," *SMPTE Journal*, 86:819/828, November 1977.
4. P. S. Carnt and G. B. Townsend, *Colour Television—Volume 1: NTSC; Volume 2: PAL and SECAM*, ILIFFE Bookes Ltd. (Wireless World), London, 1969.
5. "CCIR Characteristics of Systems for Monochrome and Colour Television—Recommendations and Reports," Recommendations 470-1 (1974–1978) of the Fourteenth Plenary Assembly of CCIR in Kyoto, Japan, 1978.
6. C. J. Hirsch, "Color Television Standards for Region 2," *IEEE Spectrum*, February 1968.
7. J. Roizen, "Universal Color Television: An Electronic Fantasia," *IEEE Spectrum*, March 1967.

7.4 Color Television Receivers

John W. Wentworth
Wentworth Training Services, Lovell, Maine

INTRODUCTION

Purpose of This Chapter

This chapter is designed to provide a general interpretation of modern color television receivers, with emphasis on those characteristics of special interest or concern to broadcast engineers or technicians, especially those who wish to develop full understanding of television as a communications medium. Although broadcast station people are not directly responsible for what happens within television receivers, they should be aware of both the limitations and opportunities introduced by these important instruments which handle their output signals before they are delivered to the ultimate recipients. It is self-evident, of course, that a television receiver serves as the final stage in each chain of technological devices that links the cameras and microphones in a television studio or remote pickup point to the eyes and ears of the home audience.

General Design Considerations for Television Receivers

Many of the technical *concepts* in television receivers are similar to those found in various categories of studio equipment, but the *requirements* for successful home receiver designs are quite different from those which apply to broadcast equipment. As consumer products, color television receivers are typically manufactured in relatively large quantities and are sold at prices that are held quite low by competition and other marketing considerations. It follows that receiver designers are trained to give very serious attention to *cost-effectiveness* in their designs, but cost savings must never compromise such matters as product safety and consumer acceptance.

Receiver design engineers are neither expected nor permitted to rely solely upon their own technical judgments; they must become intimately familiar with the requirements imposed by such organizations as the FCC, the Consumer Product Safety Commission, and the Underwriters Laboratories (and similar organizations in other countries if the receivers are to be sold in the international market). Some of the more important receiver characteristics controlled by regulatory agencies are:

- Spurious radiation of X-Rays or electromagnetic waves
- Unreasonable susceptibility to EM interference
- Shock and fire hazards
- Explosion or implosion hazards
- Mechanical hazards (such as broken handles, sharp edges or structural collapse)
- Conservative ratings of components and materials
- UHF/VHF comparability
- Placement of required warning labels

Individual television receiver manufacturers typically place additional constraints on their design engineers to assure good *manufactureability* (i.e., compatibility with the company's manufacturing facilities), *serviceability*, *documentation*, and *consumer acceptance*. Some of the more important consumer acceptance factors considered by most manufacturers are:

- Reliability
- Human engineering (ease of use)
- Styling
- Performance under adverse conditions
- Performance with nonstandard signals (such as those delivered by cable TV systems, VCRs, and video games)
- Price and operating costs

The matter of *performance with nonstandard signals* provides some of the more challenging problems faced by contemporary television receiver engineers. The low-cost modulators used in typical VCRs and video games do not reliably conform to FCC broadcast specifications with respect to the precision of the carrier frequencies, the depth of modulation, or the ratios of sync pulses to video signals; but consumers still expect these devices to produce reasonable pictures on their receivers. Cable TV systems commonly use a wide range of nonbroadcast channels in addition to the regular VHF broadcast channels, leading to significant new requirements for the design of "cable-ready" tuners. Consumer-grade VCRs do not, in general, provide the time-base stability required for normal broadcast signals, but receiver designers have successfully learned how to design deflection systems that provide stable pictures from VCR signals without compromising the performance in response to conventional broadcast signals.

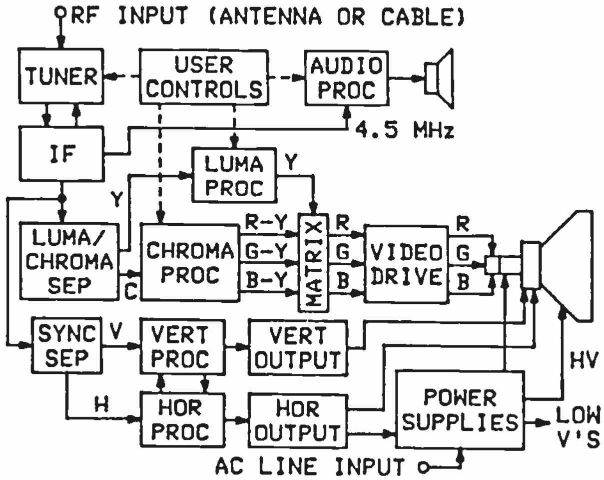


Figure 1. Functional block diagram for a color television receiver.

An Introductory Block Diagram

Most of the technical content in this chapter will be presented in a series of block diagrams, waveform sketches, or other graphics supplemented by explanatory notes. This approach provides a compact format that should be suitable both for the initial study of unfamiliar concepts and for future reference purposes. The emphasis here will be on the electronic processing functions which convert the very low-level modulated signal available at a receiver's input to the useful pictures and sounds delivered to the viewer.

The *tuner* function enables the user to select one channel from the many normally available at the *RF input*. Like most radio receivers, television receivers utilize the superheterodyne principle: that is, the selected channel is frequency-converted to a fixed-band intermediate frequency (IF) signal of sufficient bandwidth to contain both the picture and sound signals. Bidirectional arrows are shown between the *tuner* and the IF functional block, because most modern tuners employ a feedback signal from the IF stages to accomplish frequency-synthesis tuning and/or automatic fine tuning (AFT). These details will be discussed later.

The IF functional block provides most of the *gain* and *selectivity* required to amplify the desired signal and discriminate against all others. As drawn here, the IF function also includes the *video detector* which yields the baseband signal from which both picture information and synchronizing pulses are derived. The sound signal is delivered from the IF function in the form of a 4.5 MHz frequency-modulated wave, developed as the beat between the picture and sound carriers within the IF function, using a principle known as *intercarrier sound*.

The *audio processor* function includes features for demodulating the 4.5 MHz sound signal, processing the resulting audio baseband signal, and ultimately driving one or more loudspeakers. In some modern receivers, the audio processor is capable of decoding stereo signals and handling an optional *second audio program* (SAP) signal that may be broadcast by subcarrier techniques on the audio carrier.

The video signal from the IF function is separated into *Synchronizing*, *Luminance* and *Chrominance* components, customarily abbreviated to *Sync*, *Luma* (or Y) and *Chroma* (or C). The SYNC signal is further separated into *horizontal* and *vertical* components, each controlling the corresponding deflection subsystem.

In modern receivers using count-down deflection control systems (which will be described later), the *vertical processing* and *horizontal processing* functions are closely interrelated, as represented here by the bidirectional arrows between the two functional blocks. The *vertical output* and *horizontal output* stages work on quite different principles, however, because of the significant differences in frequency and power requirements.

The *power supplies* in most modern receivers are based on switch-mode principles, and are closely related to the the horizontal output system.

In most receivers of current design, the chroma (or subcarrier) portion of the video signal is processed through filters and synchronous detectors to yield R-Y, G-Y and B-Y color difference signals, which are then combined with the Y or luma signal in a *matrix* to produce red, green and blue signals suitable for application to the electron guns of the color picture tube through *video drivers*. In most low-cost receivers, the luma processing consists primarily of low-pass filtering, augmented by a trap at the color subcarrier frequency; but many "top-of-the-line" receivers now make use of *comb filters* for more complete separation and enhancement of the luma and chroma signals.

The *user controls* required for modern receivers are much simpler than those provided for earlier color television receivers, because an extensive list of automatic control features have been introduced to eliminate many of the traditional manual adjustments. The principal user controls relate to *channel selection*, sound *volume*, picture *brightness* and *contrast*, and color *saturation* and *hue*. As indicated by dotted lines in the block diagram, these controls interact primarily with the tuner function and the audio, luma, and chroma processors. Many modern receivers include provision for *remote control* of many or all of these functions.

TELEVISION TUNERS

Summary of Tuner Implementation Requirements

The input to a television tuner can be derived from an antenna system, a coaxial cable distribution system, or an accessory device such as a VCR or a video game system. Although we have already indicated that there will be no comprehensive coverage of receiving antenna systems in this chapter, the reader should be reminded that the signal normally delivered to the input terminals of a television tuner may have many imperfections. Because only a tiny amount of broadcast signal energy is actually intercepted by a typical home receiving antenna, the signal level at the tuner input is usually low enough to make noise a matter of real concern, and the signal desired for viewing at any given time is frequently contaminated by unwanted signals in adjacent or nearby channels, or even within the desired channel. *Multipath* reception can produce serious ghost images displaced from the main signal, and dynamic disturbances such as airplane flutter or severe weather conditions can cause random variations in signal strength.

In areas where mountains, tall buildings, steel towers, and other sources of signal reflection create serious multipath conditions, it is often necessary to use highly directional receiving antennas, typically mounted on rotators, to permit reception of reasonably clean pictures. One key reason for the growing popularity of cable television systems is the possibility of providing relatively clean signals of nominally equal strength on many different channels. The designers of *cable-ready* tuners must now consider, however, the very real possibility that there will be signals of approximately equal strength in adjacent channels (a problem that is minimized in the strictly broadcast environment by mileage separations between stations on adjacent channels).

The basic implementation requirements for television tuners (in the American market) may be summarized as follows:

1. Frequency coverage from 54 MHz to 806 MHz. (The range once extended to 890 MHz, but new designs are no longer required to cover channels 70 through 83. The 88–174 MHz and 216–470 MHz gaps in the broadcast television spectrum cannot be ignored by tuner designers because these gaps are now used in cable TV systems for many supplementary channels.)
2. Good noise factor (i.e., minimum added noise).
3. High signal gain (typically 35–40 dB).
4. Effective automatic gain control (typically implemented by using a feedback signal from the IF function).
5. Minimum spurious responses (from image frequencies, cross modulation, etc.).
6. Tuning stability (typically implemented through a combination of frequency synthesis and automatic fine tuning).
7. Acceptance of nonstandard signals (from devices such as VCRs and video games).
8. Compliance with Safety and Regulatory Guidelines.

Review of Heterodyne Conversion Fundamentals

Most courses in basic electronics make it clear that *amplitude modulation, amplitude demodulation*, and *heterodyne conversion* are all variants of the same basic process. In each case, an intelligence-bearing signal of some sort is electronically multiplied by a single-frequency carrier signal. Mathematical analysis shows that such analog multiplication results in the generation of sum and difference frequencies. In the case of basic amplitude modulation, these sum and difference signals form a pair of sidebands on either side of the carrier frequency; the original baseband and carrier signals may be eliminated from the modulator output, if desired, either by filtering or by the use of balanced modulator circuits.

The normal objective of demodulation is to move a modulated signal back to its original baseband. This is done by multiplying the set of sidebands by a carrier at the original carrier frequency and in the same relative phase. In ordinary AM broadcasting, the carrier required for demodulation is transmitted right along with the sidebands; in other applications, the carrier for demodulation may have to be *regenerated* at the receiver. Both sum and difference signals are generated in demodulation, but only the two difference signals resulting from each of the two sidebands are useful; the unwanted sum signals (consisting of a pair of sidebands around the second harmonic of the carrier) are normally rejected by a low-pass filter at the demodulator output.

In the great majority of existing television receivers, the tuners are based on *single conversion*. That is, the incoming signal (including both picture and sound components) is moved by a single heterodyne conversion process from its incoming frequency band to a standard IF range, which is 41 MHz to 47 MHz for receivers used in the United States. This process is illustrated by Fig. 2.

Within a single-conversion type of TV tuner, the incoming signal (including both picture and sound components) is multiplied by the output of a local oscillator in a *mixer* circuit (typically implemented by a simple diode arrangement), yielding both sum and difference signals. Only the difference signals, falling within the 41–47 MHz IF band, are wanted; the sum signals are rejected by filtering.

In the process of tuning the receiver, the local oscillator frequency is adjusted to a value exactly 41

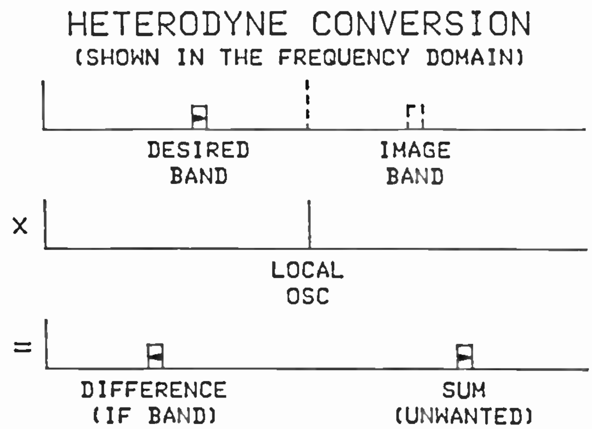


Figure 2. Spectrum sketches illustrating the process of heterodyne conversion in TV tuners.

MHz above the upper limit of the desired channel. As noted by the arrow points within the band sketches in this figure, the spectrum of the IF band is a mirror image of the original incoming band. In other words, the picture carrier that is 1.25 MHz above the lower limit of the spectrum at the output of a TV transmitter will be 1.25 MHZ below the upper limit of the IF spectrum. Likewise, the sound carrier that is placed 4.5 MHz above the picture carrier at the transmitter will be found 4.5 MHz *below* the picture carrier in the IF Band.

The basic heterodyne conversion process as used in a television tuner has intrinsic sensitivity to an *image band* of frequencies located from 41 MHz to 47 MHz *above* the local oscillator frequency, since the difference frequencies between this image band and the local oscillator also fall within the IF band. Any signals falling within this image band must be greatly attenuated before reaching the mixer to avoid objectionable interference; this is normally accomplished by using a tuned RF amplifier between the antenna input and the mixer.

A Simplified Tuner Block Diagram

To keep the diagram in Fig. 3 simple, the *band partitioning* arrangements required in practical tuners to handle the wide range of incoming frequencies have been omitted; a typical partitioning scheme will be shown later.

Most modern receivers are tuned by electrical rather than mechanical means. In the scheme shown in Fig. 3, *varactors* (or variable-capacitance, voltage-controlled diodes) are used as part of the tuned circuits which precede and follow the RF amplifier, and also as part of the tuned circuit for the local oscillator. A single tuning voltage (developed outside the tuner module) is used to control all varactors simultaneously.

The optional *balun* at the RF input may be required to convert from a *balanced* input (typically at 280 or 300 ohms) to the *unbalanced* configuration used internally to provide good shielding. The *isolation box* provides safety features to assure that no shock hazard is presented by the input terminals, bypass filters to assure no direct pickup within the receiver of interfering co-channel signals, and a high-pass filter that prevents the entry of interfering signals within the IF band.

Bandpass filters (BPF) are used both before and after the RF amplifier stage to improve the selectivity (i.e., the rejection of unwanted signals outside the desired channel, especially those in the image band). Nominally flat response across the desired 6 MHz channel is provided by the double-tuned BPF at the output in association with a single-tuned filter at the input. The bandpass filter following the IF buffer provides rejection of the unwanted sum signals at the mixer output; much more precise shaping of the IF spectrum is provided by later features within the IF functional unit.

The *noise figure* of a television tuner is determined primarily by the RF amplifier, although close attention must be given to noise considerations in mixers and post-amplifiers as well. The sources of noise that must be considered by tuner designers include:

- Source thermal noise (which varies with bandwidth, resistance, and the absolute temperature)
- Input impedance thermal noise
- 1/f noise (excess noise that can be introduced even by passive components)
- Shot noise
- Galactic (or cosmic) noise
- Interfering signals

In very simple terms, good noise performance at a tuner input requires an amplifying device with very high gain and very low input capacitance. The high gain is needed to raise the useful signal as promptly as possible to levels well above those of the various noise sources, while the low capacitance is needed to maintain the high level of gain over a wide bandwidth. Historically, the most popular circuit configuration for low-noise applications has been the *cascode* amplifier.

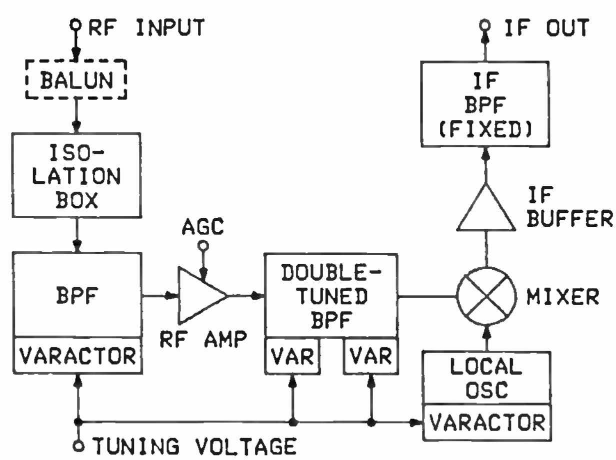


Figure 3. Simplified block diagram of a varactor-tuned single conversion tuner.

which has been implemented with vacuum tubes, bipolar transistors, and field-effect transistors. The RF amplifiers in some tuners of recent design use *dual-gate Mosfets* which provide a new way of realizing the same benefits traditionally associated with the cascode configuration.

The RF amplifier is designed to accept an automatic gain control (AGC) feedback signal developed in the IF functional block. In essence, this control voltage serves to reduce the gain of the RF amplifier, in the presence of very strong signals, to prevent overloading the RF amplifier itself or the subsequent IF stages.

Band Partitioning

It is not possible for a single set of varactor-tuned resonant circuits to cover the entire range of frequencies used for broadcasting and cable television, so some type of band partitioning is a practical requirement. Different partition boundaries are used in the products of various manufacturers, but a typical set of tuning bands is shown in the following table:

Band	Range (MHz)	Name	TV Channels
1	54–88	VHF-Low	2 thru 6
2	90–120	Mid-Band	A-1 thru A-5
3	120–174	Mid-Band	A thru I
	174–216	VHF-High	7 thru 13
4	216–300	SuperBand	J thru W
	300–402	HyperBand	W + 1 thru W + 17
5	470–806	UHF	14 thru 69

Practical means for implementing this partitioning scheme are shown in Fig. 4.

The LO/K output line carries a signal at the selected local oscillator frequency divided by a factor "K" (which is 256 for Bands 4 or 5, and 64 for Bands 1, 2, or 3). This signal is used for frequency synthesis purposes, which will be discussed later.

Q6 and Q8 are isolation or buffer amplifiers; Q9 functions as a buffer amplifier for Bands 4 or 5, but as a mixer for Bands 1, 2, or 3.

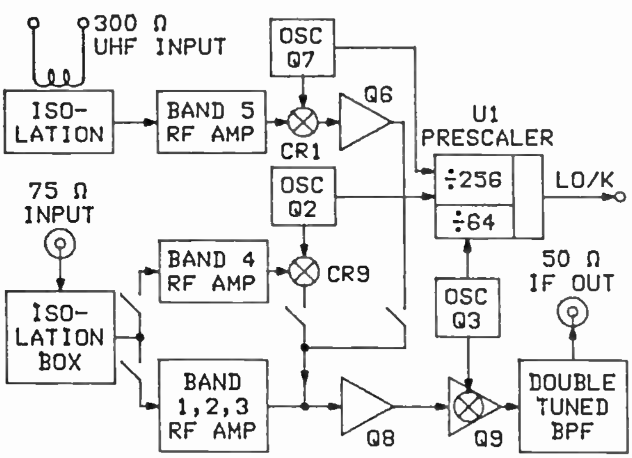


Figure 4. Simplified block diagram of a typical television tuner partitioned into five bands.

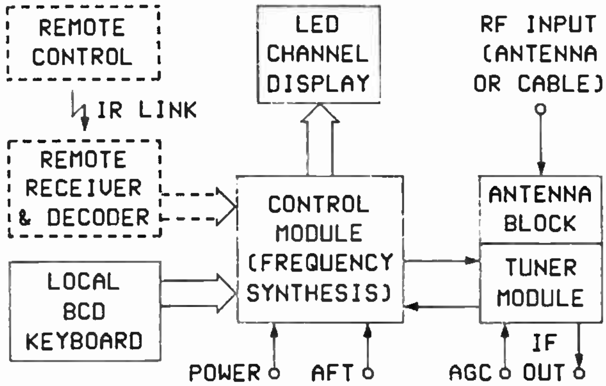


Figure 5. Simplified block diagram of a modern frequency-synthesis tuning system. (IR = Ingra Red; BCD = Binary Coded Decimal)

This complete tuning system can be controlled by only six signal lines from the related control module: one digital control line for each of the five bands, and a single analog *tuning voltage* applied simultaneously to the many varactors in the oscillators and RF amplifiers. In addition, an AGC line provides a feedback signal from the IF function.

Channel Selection and Fine Tuning

The diagram in Fig. 5 shows a typical modern arrangement for implementing user control of the tuning process. As stated in the notes for Fig. 4, there are actually six control lines extending from the *control module* to the *tuner module*, but only one feedback signal (carrying the LO/K signal) goes the other way.

The user selects a desired channel number and enters it through a binary-coded decimal keyboard. The control module (1) decodes this information (typically supplemented by the current setting of a cable vs. antenna input selector switch), (2) displays the selected channel number on an LED panel (or, in some receivers, through a momentary on-screen digital display), and (3) sets up the required configuration of the control lines to the tuner module.

As suggested by the dotted-line blocks in Fig. 5, a growing fraction of modern color receivers are equipped with remote control units, many of which provide remote control of power on/off, audio volume, and other receiver functions as well as channel selection. Such remote controls are typically linked to a *receiver/decoder* unit within the receiver through an infrared link.

The frequency synthesis technique (which will be explained more fully with the next diagram) works well with broadcast signals which are required to have precisely-controlled carrier frequencies, and which require only integer settings (in MHz) of the local oscillator frequencies. In cable systems, however, these conditions do not necessarily apply, so provision is usually made in modern control modules to "pass through" to the tuner module a tuning voltage derived

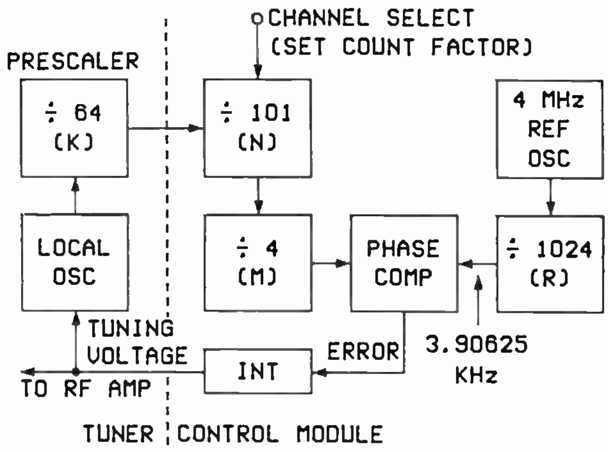


Figure 6. Simplified block diagram illustrating the basic phase-locked loop operation in a typical frequency synthesis tuner.

from an automatic fine tuning (AFT) circuit in the IF functional block. The changeover from frequency synthesis to AFT is normally made by the cable vs. antenna selector switch.

For any of the regular broadcast channels, the count factor at N is set to a number equal to the required oscillator frequency in MHz. The value of 101 selected for this illustration corresponds to Channel 2 (41 MHz higher than the upper channel limit of 60 MHz).

The product of K and M is always 256. When the prescaler is set to 256 (for Bands 4 and 5 as defined in the table presented earlier), M is simultaneously set to unity.

The *phase comparator* develops an analog error voltage proportional to the difference between the counted-down local oscillator frequency and the correct nominal value of 3.90625 KHz. This error signal passes through an *integrator* and is used as the *tuning voltage* for the *local oscillator* and the RF amplifiers within the tuner. The combination of the detector characteristic, the integrator, and the loop gain in the circuits which handle the tuning voltage provides an effective "pull-in" range of about 1.25 MHz.

Nonbroadcast signal sources (such as cable signals, VCR output signals, and video game signals) may provide picture carriers at nonstandard frequencies. In this event, an AFT system (not shown on this simplified diagram) will initiate a search for the correct tuning position by adjusting the count factors in such a way as to offset the local oscillator by 1, 2, or 3 MHz from the nominal value for selected channel in whichever direction yields a proper IF signal.

The Dual-Conversion Option

Significant engineering attention has been directed in recent years to *double conversion* techniques as a means of overcoming some of the limitations of the conventional single-conversion method, especially with respect to band partitioning problems and the difficulty of providing adequate RF selectivity, especially for UHF channels. A promising opportunity for the development of double conversion tuners has been provided by an international agreement that no broadcast stations will be assigned to channel 37 (608 to 614 MHz), thus protecting a band important to radio astronomy. At the time of this writing, few receivers using double-conversion tuners have actually reached the U.S. market, but the concept now being intensively studied is shown in Fig. 7.

Note that this basic scheme uses two IF bands, the first in the unused band from 608–614 MHz and the second in the conventional range of 41–47 MHz. A tuner for this system employs two mixers and two local oscillators, the second of which uses fixed tuning at 567 MHz.

The frequency of the second local oscillator is placed *below* the first IF band so that there is no second inversion in the sequence of frequencies in the final IF band, thus permitting the use of conventional IF filters and integrated circuits.

Advantages of this double conversion approach may be summarized as follows:

1. Reduced requirement for RF selectivity.
2. Easier rejection of local oscillator and image frequencies.
3. Relatively uncritical filter for first IF.
4. Easier band partitioning.
5. Less amplitude and group delay distortion in passband.

On the other hand, the double conversion approach is subject to the following drawbacks, which have limited its application to date:

1. More complex and expensive (added IF stage, second oscillator and mixer, higher-frequency prescaler required for frequency synthesis channel selection).
2. High stability required for second oscillator.

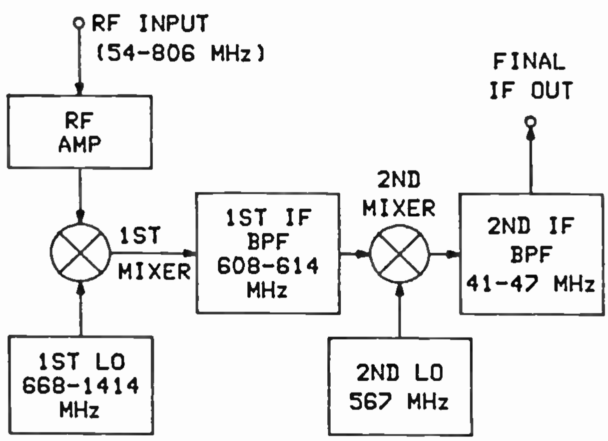


Figure 7. Simplified block diagram of a double-conversion tuning system.

3. Tracking problems made worse by wide separation of local oscillator and RF amplifier frequencies.
4. Number of spurious beat possibilities increased by the presence of two local oscillators.

As a practical matter, it has been found necessary to use a *doubly-balanced* circuit for the first mixer, and several types of traps are needed to avoid interference from strong broadcast signals near the Channel 37 band. In spite of the design problems involved, successful double-conversion tuners have actually been built and tested, but their widespread use awaits further progress in cost reduction.

IF (INTERMEDIATE-FREQUENCY) SECTION

Summary of IF Section Objectives

The basic purpose of the IF section of a television receiver is to refine the *selectivity*; i.e., to shape the intermediate-frequency response characteristic in such a way as to provide distortion-free handling of the desired signals and rejection of all others. The basic processes are illustrated by the spectrum sketches in Fig. 8.

The frequency X in the top-most sketch may be as low as 54 MHz (for channel 2) or as high as 800 MHz (for channel 69).

As explained previously, frequency inversion of the channel structure at the tuner output results from the use of a local oscillator frequency placed above those of the picture and sound carriers.

Proper demodulation of the vestigial-sideband picture signal requires that the picture carrier be placed at about the 50% response point (or the −6 dB point) on a slope that is nominally linear over a range of 0.75 MHz on either side of the carrier frequency.

It is a common design practice to place the color subcarrier at about the −6 dB point on a somewhat steeper slope at the low end of the IF characteristic, but this frequency response is modified by a later *chroma peaker* circuit after the video signal has been

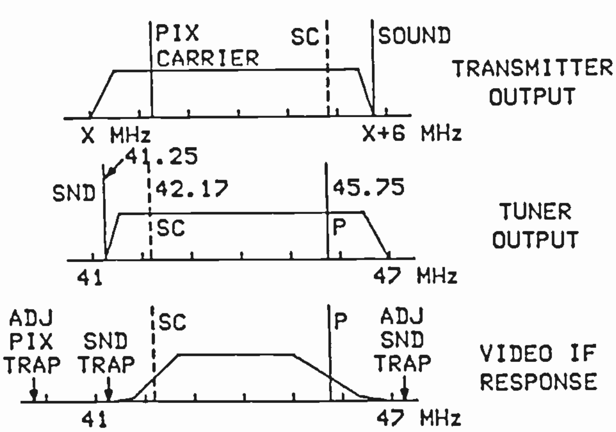


Figure 8. Spectrum sketches showing the standard U.S. broadcast channel, tuner output signal, and "idealized" IF response curve.

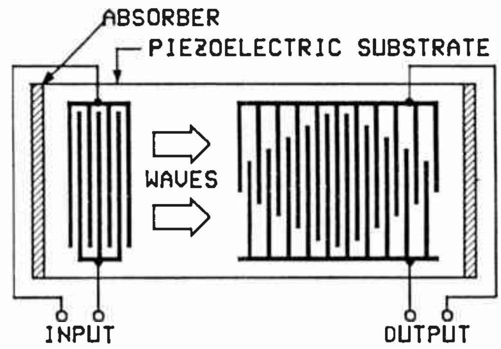


Figure 9. Simplified sketch showing the working principle of a surface acoustic wave filter.

demodulated. Before the chroma information is demodulated from its subcarrier, the overall frequency response for the chroma channel must be made nominally flat over a range extending at least 0.5 MHz on either side of the subcarrier.

Trap circuits (or band-stop filters) are required to provide extra attenuation at frequencies corresponding to adjacent-channel picture or sound carriers; this is particularly important for modern receivers that are frequently used with cable systems that provide strong signals on virtually all available channels.

The desired channel sound signal must be recovered, of course, but a trap at the sound carrier frequency is required in the picture channel to prevent audio contamination of the picture signal. In most receivers, the final sound trap is placed in the video domain after demodulation, since the intercarrier sound IF signal at 4.5 MHz is often derived from the output of the video detector.

IF Applications of SAW Filters

Prior to the practical refinement of surface acoustic wave (SAW) filters and related integrated circuits in the late 1970s or early 1980s, the factory alignment of the IF amplifiers in television receivers typically required ten to 12 tuning adjustments, many of which interacted with each other. Modern designs of IF systems based on SAW filters have greatly reduced the need for such factory adjustments.

SAW filters exploit the bilateral property of the piezoelectric effect. Variable electrical fields set up between the interdigital fingers of an input transducer create surface acoustic waves (involving deformation of a crystal material). These waves are propagated down the substrate, and a fraction of their acoustic energy is converted back into electrical energy by the output transducer, also of the interdigital type.

Surface waves are actually propagated in both directions from the input transducer, but those traveling to the left (as drawn here) are converted to harmless heat energy by absorbing material placed on the surface at the left end of the substrate. Similar absorbing material is placed beyond the output transducer to prevent reflections and multiple trips for the desired waves.

Shaping of the frequency and phase (or group delay) characteristics of a SAW filter is accomplished by varying the effective lengths and spacings of one or both transducers, as shown here. The design equations become very complex if both transducers are varied, so it is common practice to use uniform spacing and uniform "finger" lengths for one transducer or the other. In actual practice, both transducers have many more "fingers" than those shown in this simplified sketch.

The center frequency of a SAW filter is determined by the velocity of surface-wave propagation along the substrate (a property of the piezoelectric material) in association with the mean wavelength, which is determined by the physical dimension of each full cycle in the transducer structure along the direction of propagation. For simple filters or delay lines, it is normal design practice to make both the widths of the individual fingers and the spacing between them (in the direction of propagation) equal to one-quarter wavelength at the desired center frequency, but the width and spacing may be varied slightly to control group delay characteristics.

Because SAW filters can be designed to have specific phase shift or group delay characteristics as well as precise amplitude-versus-frequency characteristics, it has become common practice to "tailor" their specifications to compensate for other less-than-perfect components within the RF or IF sections of a receiver. It is particularly noteworthy that receiver engineers can now match with relative ease the envelope delay characteristic cited in the FCC specifications (for which broadcast transmitters are required to provide precompensation).

The advantages of SAW filters in the IF section of a television receiver may be summarized as follows:

- Flexibility in shaping the IF response
- Excellent rejection of unwanted signals
- Independent control of amplitude and group delay
- Small physical size
- High repeatability (no adjustments required)
- Low manufacturing cost

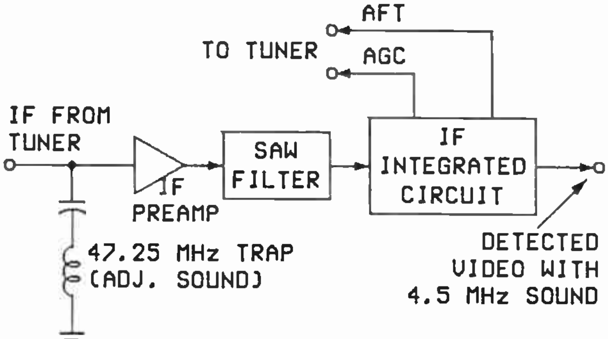


Figure 10. Simplified block diagram of a complete IF system incorporating a SAW filter.

As suggested by Fig. 10, it is now common practice to place most of the active electronics for a receiver IF system within a single integrated circuit, which will be shown in a little more detail later. This diagram applies to a "standard" color television receiver, not to one with more advanced "top of the line" features.

The IF *preamplifier* provides a signal level appropriate for driving the SAW filter, and also provides further isolation from the tuner output. It is common practice to use a trap circuit at 47.25 MHz ahead of this preamplifier to minimize any intermodulation distortion within the preamplifier from a sound carrier signal in the channel immediately below the one selected by the tuner.

Considered as a single functional block, a typical IF integrated circuit receives a properly-shaped but low-level IF signal at its input and delivers at its principal output a base-band video signal with a superimposed FM sound signal centered at 4.5 MHz. Two supplementary outputs provide feedback signals to the tuner for automatic gain control (AGC) and automatic fine tuning (AFT) purposes.

Typical Integrated Electronics for IF Systems

For purposes of this handbook, it is not practical to provide detailed descriptions of specific integrated circuits (IC); since these vary significantly with time, the techniques adopted by different manufacturers, and the "feature levels" (and the subsequent price tags) of different receivers. We shall limit ourselves here, therefore, to simplified block diagrams showing first a typical IF integrated circuit for a basic receiver and later a more elaborate IC featuring a separate sound channel for a more advanced receiver. (Fig. 11)

Because the required IF selectivity is provided by the SAW filter which precedes this IC, no interstage tuning adjustments are required for the three-stage IF amplifier.

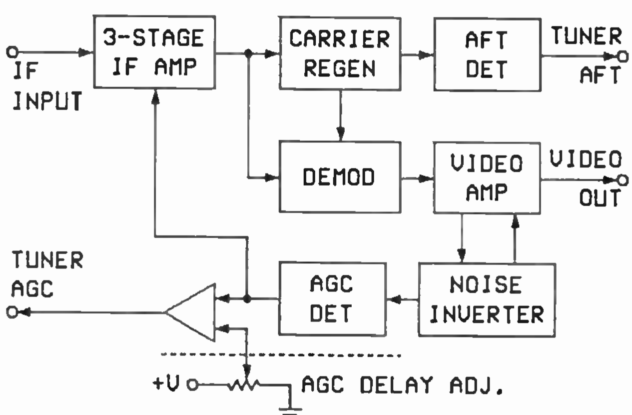


Figure 11. Simplified block diagram for an IF integrated circuit suitable for a standard-feature color television receiver.

This IC uses a synchronous detector for demodulating the video signal, which provides much lower distortion than a simple envelope detector (of the diode rectifier type). Such a synchronous detector (based on the same principles as those used in NTSC color decoders) requires a regenerated carrier, which is provided by circuits in the block marked *Carrier Regen* which remove the modulation from a sample of the incoming picture signal.

In the *AFT detector* block, the frequency of the regenerated carrier is compared with a stable reference at 45.75 MHz to produce an automatic fine tuning signal that is fed back to the tuning system. As a general rule, this feedback signal is actually needed only when the receiver is used with nonbroadcast signals that do not have the carrier predictability and stability required by FCC broadcast standards.

In this design, the *video out* signal also includes the sound signal in the form of a superimposed 4.5 MHz subcarrier frequency-modulated by the audio program signal. In subsequent separation circuits, a band-pass filter is used in the audio channel to reject everything except the 4.5 MHz wave, while a band-stop filter (or *trap*) centered at 4.5 MHz is inserted in the video channel for the sync and picture signals.

The *noise inverter* greatly reduces problems resulting from impulse noise that may be present along with the desired signal. Separate circuit features are used to detect either white-going or black-going noise "spikes" that extend well outside the normal dynamic range and to clamp the signal to "safe" values during these noise pulses. These processes will be explained more fully in Fig. 12 below.

The *AGC detector* develops a voltage proportional to the sync level from the video demodulator, which should be a constant under broadcast transmission conditions. (Keep in mind that a negative modulation characteristic is used for television broadcasting, causing the tips of sync pulses to be at the highest carrier level, independent of picture content.) As will be shown in more detail in Fig. 13, this voltage is used as a negative feedback signal to adjust the gains of both the IF and RF amplifiers in such a way as to hold the output of the video detector essentially constant in spite of great variations in the strength of the incoming signal.

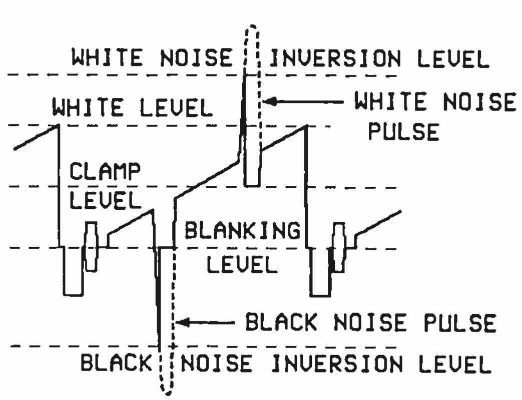


Figure 12. Waveform sketch illustrating the operation of noise inversion circuits.

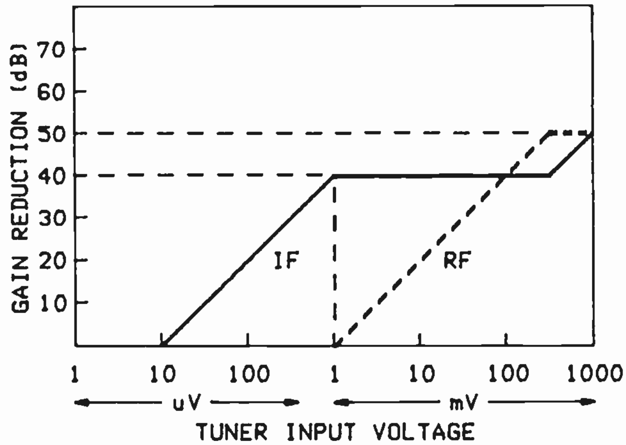


Figure 13. Plot of AGC gain reduction versus tuner input level for a typical receiver, showing both tuner (RF) and IF characteristics.

The term delay related to the external *AGC delay adjustment* applies not to the time domain but to the control voltage domain. To allow the IF system to work with signals that are well above noise levels, the gain reduction required for AGC is first applied to the IF amplifier. Only when the incoming signal is strong enough to produce an AGC voltage above a threshold established by the AGC delay potentiometer will a gain-reduction signal be applied to the tuner RF amplifier as well. Refer to Fig. 13 for more details on this point.

The *noise inverter* circuit block in Fig. 12 consists essentially of level detectors which sense noise pulses which extend above the white noise inversion level or below the black noise inversion level, supplemented by shaping circuits for the detected pulses.

Black-going noise pulses which pass the black noise inversion threshold are used to clamp the video signal to the blanking level for the duration of the noise pulses.

White-going noise pulses which exceed the white noise inversion threshold are used to clamp the video signal to a mid-range gray value for the duration of the noise pulses.

The narrow residual "spikes" extending up or down to the noise inversion levels are far less visible and troublesome than the uncorrected noise pulses.

Typical modern receivers have enough gain to provide a full-contrast picture for any tuner input voltage greater than about 10 microvolts, so the onset of automatic gain reduction is usually set at about this point.

For signals less than about 1 millivolt, the gain reduction is applied solely to the IF amplifier so that the best possible signal-to-noise ratio can be provided by the tuner.

For signals above 1 millivolt, the problem of intermodulation beats arising from excessive levels at the tuner/IF interface generally becomes more serious than the signal-to-noise problem, so for the next 50 dB of input range the IF gain is held constant and gain reduction is applied to the RF stage in the tuner. If the signal becomes strong enough to exceed the AGC range of the tuner, further gain reduction may be applied to the IF amplifier.

This diagram is simplified in the sense that it does not show how gain reduction is applied progressively to the several stages within the IF amplifier, beginning with the highest-level stage.

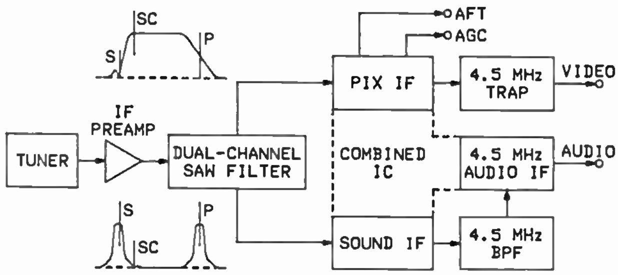


Figure 15. Simplified block diagram of a quasi-parallel IF system using a dual-channel SAW filter.

Quasi-Parallel IF Systems

The relatively recent changes in the FCC regulations authorizing the transmission of stereophonic sound signals with television programs have stimulated interest in the design of more advanced IF systems that handle audio and video signals in separate channels. Laboratory tests and field trials have shown that the performance characteristics of conventional intercarrier systems (such as those illustrated by Figs. 10 and 11) are often not good enough for the satisfactory recovery of stereo audio signals and related auxiliary audio services. Many of the first receivers featuring stereo sound incorporated IF sections based on the concepts illustrated by the following diagram.

In the system in Fig. 14, the video circuits are essentially the same as those shown in previous illustrations, but the audio is recovered from a parallel system that uses relatively narrow-band tuned circuits at the picture and sound carrier frequencies, plus a conventional IF integrated circuit (the same as in the video channel) to develop an improved 4.5 MHz sound IF signal from which the complex baseband signals can be successfully recovered.

In the system shown in Fig. 15, a trap at 42.17 MHz (which is the color subcarrier frequency in the IF spectrum) can be used in association with the sound carrier bandpass filter at 41.25 MHz, and the picture carrier bandpass filter at 45.75 MHz can be made narrow enough to effectively remove all video modulation. Many video-to-audio intermodulation hazards are eliminated by such blocking of the video sideband frequencies from the audio IF channel.

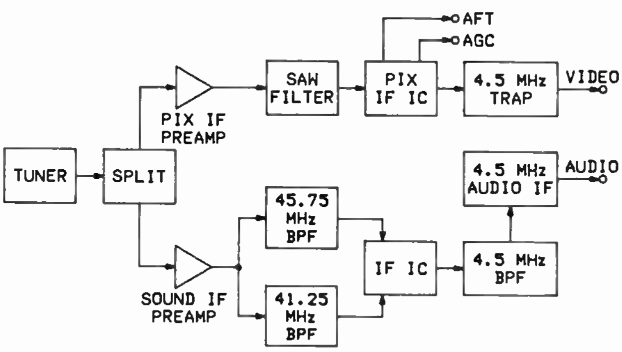


Figure 14. Simplified block diagram of a quasi-parallel IF system suitable for the delivery of stereophonic sound signals.

Compared to the system of Fig. 14, the system shown in Fig. 15 provides advances both in performance and in reduced manufacturing costs.

The dual-channel SAW filter has a single input transducer but separate output transducers (placed side-by-side on the substrate) for the picture and sound IF signals.

The part of the SAW filter used for the video IF channel can provide a deeper sound notch and a flatter response in the vicinity of the color subcarrier, since there is no need to recover any part of the audio signal from the video channel.

Likewise, the part of the SAW filter used for the audio channel can have optimum shaping for its intended purpose, including a notch at the color subcarrier frequency.

The circuitry within the sound IF integrated circuit can be simplified relative to its video counterpart, and manufacturing economies can be realized by producing both ICs on a single chip. It is also quite possible to include many of the 4.5 MHz audio IF circuits in the same integrated circuit for still greater savings in board space and manufacturing costs.

AUDIO DEMODULATION AND PROCESSING

Basic Monaural Systems

Although there is growing interest in advanced audio features for color television receivers, the great majority of the instruments in use at the time of this writing fall into the "no frills" category with only basic monaural audio output. The audio requirements for such receivers may be summarized as follows:

1. "Clean" demodulation of the audio signal from the sound carrier, with good sensitivity, high signal-to-noise ratio, and low distortion.
2. Apply de-emphasis (i.e., attenuation of the higher audio frequencies) as required by FCC standards.
3. Adequate power output for comfortable listening in the typical viewing environment, plus an adequate range of volume control.

4. Bass and treble tone control appropriate to the intended receiver application. (This sometimes means a fixed tone setting with no user controls.)
5. Immunity to turn on/turn off transients (clicks and pops).

Under current design practices, most of the circuits required for audio demodulation and processing are contained within a single integrated circuit, although certain bulky components that do not lend themselves to economical integration are normally mounted separately. A typical "no frills" arrangement is shown in Fig. 16.

Following the point where the 4.5 MHz audio IF signal is separated from the video signal, its spectrum is further shaped by an external bandpass filter before it is applied to the differential input terminals of the sound integrated circuit.

Within the IC, amplitude variations in the audio IF signal are removed by a four-stage limiter, and demodulation is accomplished by a quadrature detector (which requires an external tank circuit).

Volume control is implemented through an *electronic attenuator*, which is based on a variant of the multiplier or balanced modulator circuits used at various other points within the receiver. In essence, the audio signal is multiplied by a DC control voltage derived through an external network which includes the user's volume control and a point for applying the *mute* signal which silences the audio output during control operations (such as changing channels).

De-emphasis of the audio signal is accomplished by the combination of an internal resistor and an external 0.01 uF capacitor to ground. An external DC-blocking capacitor of about 0.047 uF value is also required to couple the audio signal back to the IC for the power amplifier function.

The integrated audio output amplifier is protected against overload conditions by internal thermal sensing and automatic shutdown circuits. In typical table-model applications, the desired output is usually in the range of one to two watts (at about 1.5% total harmonic distortion), but some versions of the audio processing integrated circuit can provide up to about five watts of output power if the circuit itself is mounted on a suitable heat sink.

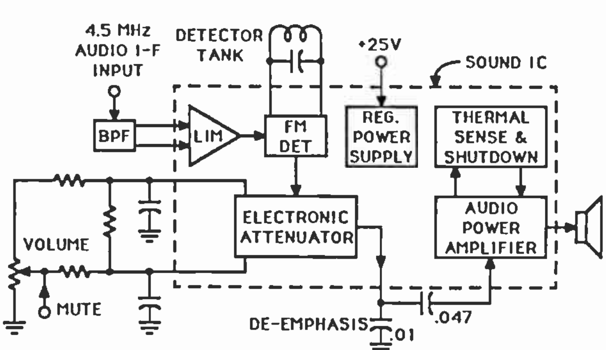


Figure 16. Simplified block diagram of a basic monaural sound system for a television receiver.

Advanced Audio Systems for "Full Feature" Receivers

The authorization of stereophonic sound for television broadcasting, plus the growing public interest in high-fidelity audio systems in general, has stimulated efforts to design and produce receivers with audio features that go well beyond the basic requirements cited above. Some of the additional audio requirements for "full feature" receivers are summarized in the following list:

1. Bandwidth of the audio demodulator increased to about 100 kHz for proper recovery of the encoded stereo and optional second audio program signals.
2. Decoding of composite stereo signals to produce left and right signals.
3. Apply dBx expansion as required by the BTSC (Broadcast Television Systems Committee) stereo signal specifications. (NOTE: "dBx" is a registered trademark of dBx, Inc.)
4. Maintain stereo channel balance, separation, and tracking.
5. Recover optional second audio program (SAP) (sometimes referred to as a separate audio program).
6. Provide switching for selection of alternative audio sources.
7. Provide more versatile control of bass and treble levels.
8. Provide automatic tone compensation for volume control.
9. Adjust processing to suit characteristics of non-broadcast signals.
10. Apply optional dynamic noise reduction (DNR: a registered trademark of National Semiconductor Corp.).

We cannot provide detailed descriptions of these many options in this brief chapter, but we shall attempt to describe the basic BTSC stereo signal specifications and present a simplified block diagram for a suitable decoding arrangement. For more information on TV stereo sound, see Chapter 3.5 "TV Transmitters," and Chapter 3.6 "Multichannel Television Sound."

The techniques used for BTSC stereo are generally similar to those used for stereo broadcasting in the FM radio band, but there are significant differences in the processing features incorporated in the standards.

The spectrum sketch shown here applies to the input to the sound modulator in a broadcast transmitter, or to the output of the audio demodulator in a receiver.

For compatibility with existing monaural television broadcast services, the main channel (within the baseband audio spectrum from nominally 0 kHz to 15 kHz) is occupied by the "left plus right" (L+R) signal, processed only with the standard 75 microsecond pre-emphasis and modulating the main carrier to the same

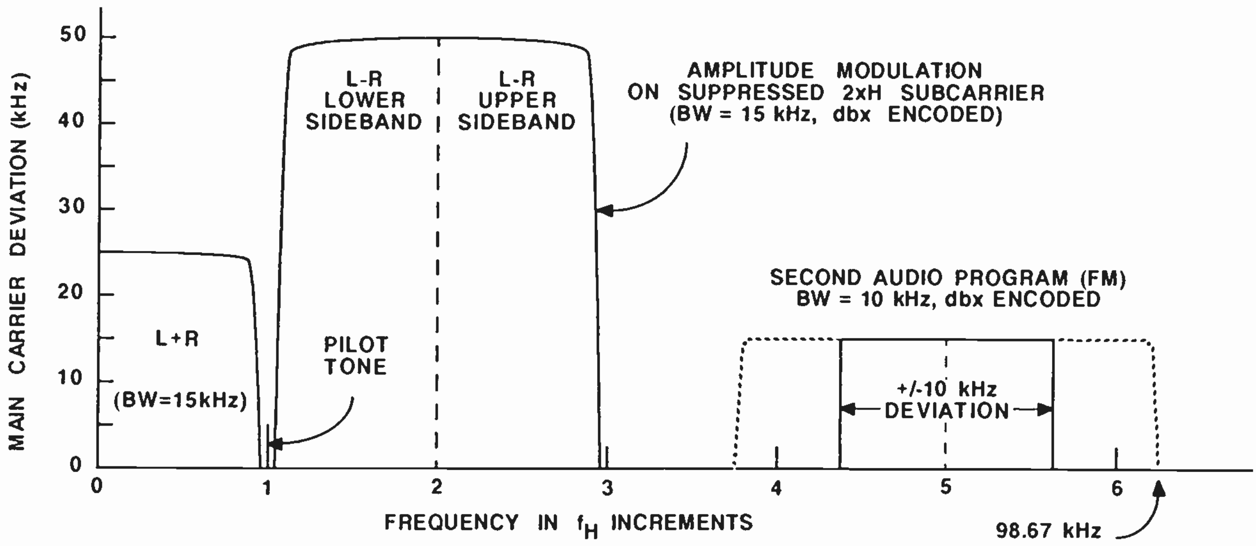


Figure 17. Spectrum sketch showing the major components of a composite audio signal under the BTSC standards for Stereo Television Broadcasting.

maximum of ±25 kHz specified in the older monaural standards.

The "left minus right" (L − R) signal is transmitted in the form of a pair of symmetrical sidebands corresponding to amplitude modulation of a suppressed subcarrier at twice the horizontal frequency. This signal has the same 15 kHz bandwidth as the L + R signal, but is compressed by the dBx technique prior to modulation. The level of this modulated L − R signal is such as to produce a maximum main carrier deviation of ±50 kHz.

A *pilot tone* at the horizontal frequency (15.734 kHz) is added to the L + R signal and transmitted at the level which produces ±5 kHz deviation of the main carrier. This pilot tone is used both to indicate the presence of a stereo signal and to control the regeneration of a subcarrier at 2 × H for synchronous demodulation of the L − R signal.

A second audio program (SAP) of 10 kHz bandwidth is frequency-modulated on a subcarrier at five times the horizontal frequency (78.67 kHz). This signal is also subject to dBx encoding (or compression) before modulation. The maximum deviation permitted for the subcarrier is ±10 kHz, and the resulting modulated wave is added to the other signals in the composite "mixture" at a level which produces ±15 kHz deviation of the main carrier.

Under the "rule of thumb" that the bandwidth required for an FM signal is nominally equal to the deviation range plus a double set of sidebands, the modulated SAP signal may occupy a band extending from about 58.67 kHz to 98.67 kHz. (The same result is predicted by Carson's Rule, which states that the half bandwidth required for an FM signal is the peak deviation plus the bandwidth of the modulating signal.)

The term "dBx encoding," which applies to both the L − R signal and the SAP signal, refers to a specific set of audio processing techniques developed and promoted by dBx, Inc. and adopted by the Broadcast Television Systems Committee for inclusion in the standards for stereophonic television broadcasting. In contrast to the simple preemphasis used as the principal audio processing technique for monaural television broadcasting, dBx encoding is a combination of three noise-reduction techniques, specifically designed to provide the best possible psychoacoustic masking with program audio signals of varied spectral content transmitted through the subcarrier channels of the BTSC television audio system. The three techniques are (1) fixed preemphasis, (2) compression, and (3) variable preemphasis. Each technique predistorts the transmitted signal in a manner that requires a complementary decoding process at the receiver to restore normal transmission conditions for the overall system with a significant reduction in the perceived noise level.

In addition to the features shown in Fig. 17, the BTSC standards permit an optional "professional" channel for data or voice-grade signals in the form of a low-level FM or FSK (frequency shift keyed) subcarrier signal centered at 6.5 × H. This optional feature of the BTSC standards is designed for use by the television broadcasters for operational purposes and is not intended for reception by the public.

The various operations performed in the system shown (Fig. 18) are essentially the inverse of those used in encoding the BTSC stereo signal at the transmitter, but a few special techniques are introduced to minimize receiver costs. Much of the processing is accomplished within two integrated circuits: one for the basic stereo decoding, and another for dBx decoding.

The wideband composite audio signal from the sound IF system is buffered by a preamplifier within the IC and split out into two channels. A 50 kHz low-pass

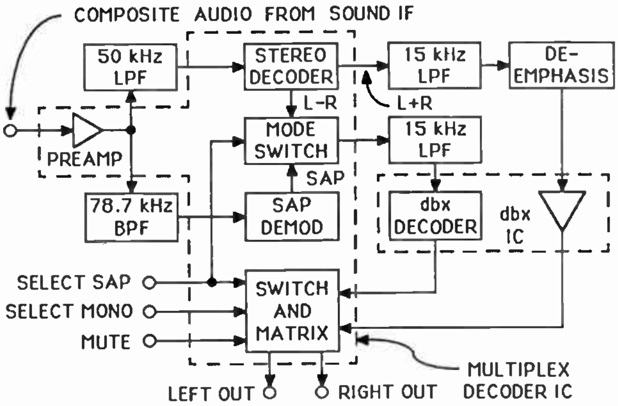


Figure 18. Simplified block diagram for a BTSC stereo decoding system.

filter selects the part of the spectrum containing both the baseband L + R signal and the modulated L − R signal for the primary audio program, while a band-pass filter centered at 78.7 kHz selects the part of the spectrum containing the second audio program (SAP) in the form of a frequency-modulated wave.

Circuits within the *stereo decoder* functional block recover baseband L + R and L − R signals, the latter requiring a carrier regenerator (controlled by the pilot tone) and a synchronous detector. Likewise, circuits within the SAP demodulator functional block provide the limiter and FM demodulator needed to recover the baseband second audio program.

To permit sharing of the relatively complex and expensive dBx decoding circuits, a *mode switch* is used to select either the L − R signal or the SAP signal for subsequent processing. This mode switch is controlled by a *select second audio program* line from the user control subsystem.

The two 15 kHz low-pass filters which limit the frequency response of the L + R and L − R channels must be carefully matched to preserve good stereo separation. The matching requirement does not apply when the second audio program has been selected, since the output will not then be in stereo.

Processing for the L + R signal consists only of de-emphasis and a bit of delay compensation to keep the L + R channel matched to the L − R channel. The L − R signal (or the SAP signal, should it be selected) is decoded or expanded in accordance with the dBx techniques specified in the BTSC standards to compensate for the dbx encoding applied at the transmitter.

In the *switch and matrix* part of the IC, the final left and right audio signals are assembled in accordance with the user's desires as expressed through the *select SAP* and *select Mono* control lines, subject to certain automatic overriding conditions based on detection of the pilot tone and a valid SAP signal. The *mute* control line is used to suppress clicks and pops during certain control operations, such as channel selection.

Dynamic Noise Reduction (DNR)

Dynamic noise reduction (or DNR) is a proprietary audio noise-reduction technique licensed to the National Semiconductor Corporation and implemented by integrated circuit packages produced by that company. Although it is based on psychoacoustic masking, it is distinctly different from the standard preemphasis/de-emphasis technique and the dBx encoding/decoding technique in that it is not based on complementary circuits at the transmitting and receiving ends of a system. It may be used within a receiver alone to provide subjective improvement in audio signals from many different sources, including broadcast channels, audio or video cassette recorders, and video disc systems.

In essence, dynamic noise reduction is a technique for varying the frequency response of an audio channel (or a pair of stereo channels), over a range of about 1 kHz to 30 kHz (as indicated by the family of curves at the left), by an automatic circuit which senses the amount of signal energy in the upper part of the audio band through a control channel with frequency response. (See Fig. 19.) Subjective noise performance is improved because the bandwidth is permitted to

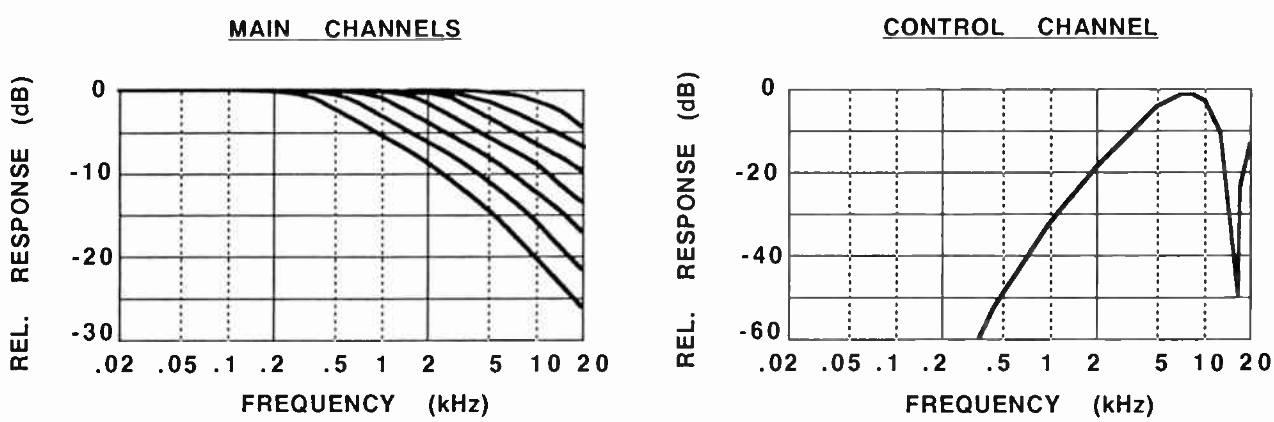


Figure 19. Frequency response curves for a dynamic noise reduction system.

extend to the high frequencies only when there is enough signal energy at these frequencies to provide "masking" for the noise.

The nominal frequency response of the control channel is designed to match the sensitivity of the typical human ear to noise in the presence of a tone, but a notch at 15.734 kHz is needed to prevent the fixed pilot tone or crosstalk from horizontal deflection circuits within the receiver from affecting the control setting.

DNR control circuits are designed to widen the bandwidth very rapidly (in less than 1 millisecond) when a signal with significant high-frequency energy is detected, but to reduce the bandwidth more gradually, over a period of about 60 milliseconds. These adjustments permit the "ambience" of music to pass through a system while keeping the "noise floor" quite imperceptible.

Since not all audio signals can be improved by dynamic noise reduction, it is customary to provide a user switch to enable or disable the system.

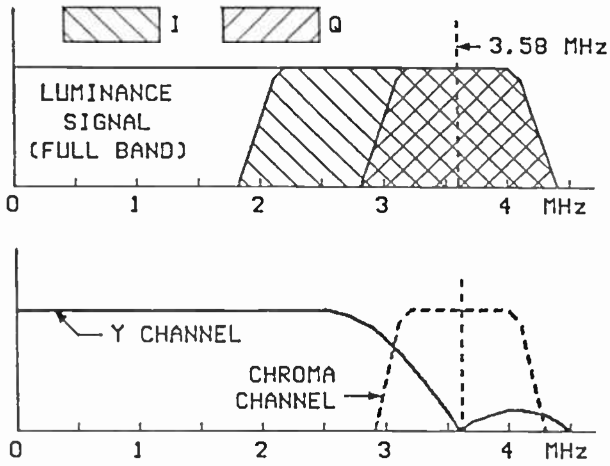


Figure 20. Sketches of the spectrum of an NTSC video signal and typical response curves for chroma and luminance channels using the notch filter separation method.

BASIC LUMINANCE AND CHROMINANCE PROCESSING

The Chroma/Luma Separation Process

Because the luminance and chrominance signals in NTSC color television share the same video spectrum, some special problems are involved in separating the two categories of signals. In separating chrominance from luminance, it is desirable to (1) minimize the presence of dots (resulting from the subcarrier) in large areas, (2) minimize the dot crawl phenomenon at the boundaries of colored areas, and (3) preserve as much luminance resolution as possible. In separating luminance from chrominance, it is desirable to minimize the cross color effects resulting from the presence of luminance frequency components in the vicinity of the color subcarrier frequency.

Because of the frequency interlace technique incorporated in the NTSC color signal specifications, part of the luminance/chrominance separation process is psycho-physical. Thanks to the deliberate choice of a color subcarrier frequency at an odd multiple of one-half the line scanning frequency, some of the effects of interfering signals in both the luminance and chrominance channels tend to be averaged out by the persistence of vision effect over periods of two or more frames, but the nonlinear response characteristics of practical picture tubes limits the effectiveness of such psychophysical filtering. Thus, it is normal practice to employ electronic filters in the *luma/chroma separator* portion of a receiver. In traditional designs and in the lower-cost receivers still being manufactured today, the filters consist essentially of a bandpass filter for the chroma channel and a notch filter in the luminance channel, as shown in Fig. 20. In more deluxe receivers, it has now become common practice to use comb filters, which will be discussed in a later section of this chapter.

Although a notch filter in the luminance channel can remove most of the subcarrier energy (and thus the large-area dot patterns), it leaves residual dot-crawl effects on edges and reduces horizontal resolution in the luminance channel.

Because the response characteristic at the high end of the video spectrum is usually rolled off to some extent by the IF spectral response (as shown in Fig. 8 earlier), a chroma peaker (that is, a circuit whose response rises with frequency in the vicinity of the color subcarrier) is customarily used in association with a bandpass filter to provide a flat-topped characteristic for the chroma signal in the overall receiver system.

Use of a relatively simple bandpass filter to recover chrominance information provides no discrimination against luminance components in the same band, and thus leaves the system subject to cross color (or false color) effects for certain types of luminance signals that contain strong signal components within the chrominance band. These effects are particularly noticeable in pictures containing picket fences or other strong vertical stripes.

A chroma separator based on a simple bandpass filter does not utilize the additional chrominance bandwidth available in the I signal channel.

In spite of the limitations cited above, many millions of receivers using the principles shown here have won consumer acceptance. (Even more serious bandwidth restrictions are found in most consumer-grade VCRs, but these also have won consumer acceptance.)

An Introduction to Luma/Chroma Processing

In most modern receivers, most of the luminance and chrominance processing circuits are contained with a single integrated circuit, a typical example of which is shown in Fig. 21. Each of the functional

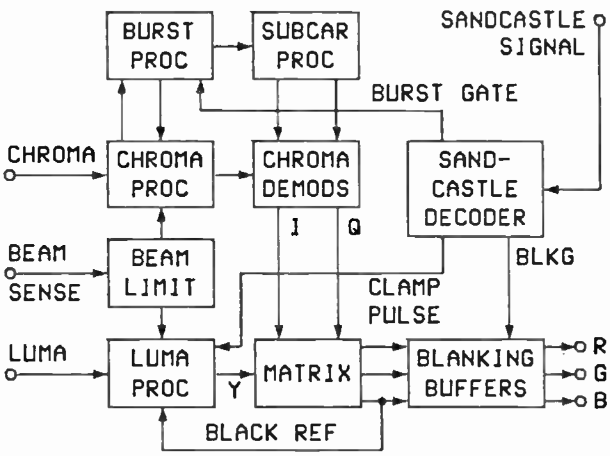


Figure 21. Simplified block diagram of a typical luma/chroma integrated circuit.

blocks will be more fully described in the somewhat more detailed diagrams to follow.

The *sandcastle signal* is a composite timing signal containing horizontal blanking, vertical blanking, and burst gating pulses, generated in the pulse and deflection sections of the receiver. The horizontal and blanking pulses correspond to the retrace periods for the corresponding deflection systems, while the burst gating pulses are usually generated by clipping the output of a simple ringing circuit that is shock excited by the trailing edges of horizontal synchronizing pulses. Use of this composite signal helps to reduce the number of pins required for the integrated circuit.

The sandcastle signal is internally decoded to provide composite blanking, a burst gating signal and a clamp pulse (with the same timing as the burst gate).

A *beam sense* signal, proportional to the total beam current in the picture tube, is used to drive an automatic over-ride circuit, which will reduce the levels in both the chroma and luma channels simultaneously if the beam current passes a threshold indicating an over-driven condition, thus protecting the picture tube from damage.

The chroma signal is handled through a series of four functional blocks, each performing processing operations that will be shown more fully in later diagrams. In brief summary, the first block provides for saturation control, automatic beam-current limiting and driving power for an external high-pass filter (not shown in this simplified diagram). The second block separates the color synchronizing burst and uses it to control a subcarrier oscillator and an *automatic chroma control circuit*. The subcarrier processor provides the carriers for the subsequent demodulators and also provides an *automatic flesh-tone corrector* circuit. The final block contains the demodulators which yield I and Q chrominance signals.

The *luma proc* functional block provides black level stabilization (through clamping), contrast control (with automatic beam current limiting), and brightness control.

The *matrix* functional block cross-mixes the Y, I, and Q signals to produce red, green, and blue video signals, which then pass through blanking amplifiers which effectively turn off the picture tube beams during retrace periods.

Chrominance and Subcarrier Processing

The small numbered blocks in Fig. 22 (and in a few subsequent diagrams as well) represent IC pin numbers used for connecting external components or signal lines. Such pin numbers are valid only for the specific type of IC selected for illustration, but their use is convenient for tutorial purposes in the following notes.

Following the first chroma amplifier, gated bursts are applied to two separate detectors. The *burst-phase detector* is a quadrature-phase detector producing an output proportional to the direction and magnitude of any phase error between the incoming bursts and the −90° output of the local voltage controlled oscillator (VCO). The phase reference for the *burst level detector* is delayed 90° by external components connected between Pins 11 and 12, so it serves as an in-phase detector yielding an output pulse proportional to the amplitude of the incoming bursts. (Literally speaking, this detector is set to be 180° out of phase with the VCO, which means only that the pulses produced are opposite in polarity from those that would be produced by a true 0° reference.)

Both detectors produce pulse-type signals which are converted to steady control voltages by *sample and hold* circuits (which require external storage capacitors).

The center frequency of the voltage controlled oscillator is stabilized by an external 3.58 MHz quartz crystal connected between Pins 13 and 11. The tight specification on the subcarrier frequency in the FCC

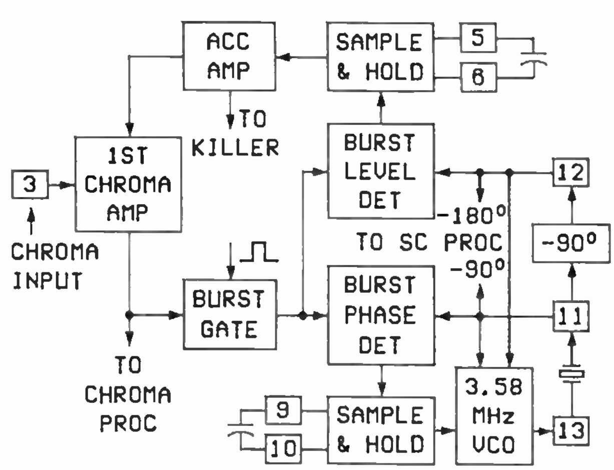


Figure 22. Partial block diagram of a typical luma/chroma IC, showing the first chroma amplifier, burst processing circuits, and burst-controlled oscillator.

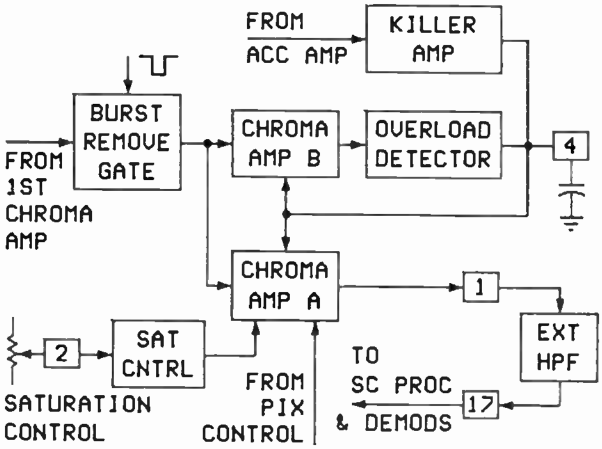


Figure 23. Partial block diagram of a typical luma/chroma IC, showing the major chroma processing circuits.

Standards makes this simple crystal stabilization technique practical; the VCO does not need a wide pull-in range.

The automatic chroma control (ACC) amplifier provides a gain-control signal for the first chroma amplifier, thus closing a feedback loop that tends to maintain a uniform level of burst at the burst gate. If the burst/chroma ratio conforms to industry standards, the chroma level fed to the subsequent chroma processing circuits is also held at its correct value.

The *burst remove gate* in Fig. 23 is similar to the *burst gate* of Fig. 22, except that the polarity of the keying pulse is reversed. Thus, the burst remove gate passes the picture chroma signal, but blocks the bursts.

Because the burst/chroma ratio of the incoming signal is not always correct, the automatic chroma level control based on the burst level is not always sufficient. The absolute level of the chroma signal is measured by an *overload detector* circuit (which requires an external filter capacitor at Pin 4), and a feedback signal from this detector reduces the gains of Chroma Amp A and Chroma Amp B if the absolute level exceeds normal limits.

The *killer amplifier* clamps the error voltage from the overload detector to a value which completely turns off Chroma Amplifiers A and B if the incoming burst level falls to less than about 5% of its normal value. Under these conditions, the picture will be displayed in black and white only.

Manual control of the gain of Chroma Amp A is provided by an external user control that may be labeled either saturation or color. Yet another electrical adjustment of the gain of Chroma Amp A is derived from the user control labeled either *picture* or *contrast*, which also controls the gain of the luminance channel.

The external high-pass filter between Chroma Amp A and the subsequent processing and demodulator circuits is needed both to provide a DC block (to re-establish proper bias conditions for the internal stages) and to provide a bit of phase compensation for the dynamic flesh-tone correction circuit shown in Fig. 24.

A control voltage from the user control labeled *hue* or *tint* is applied to a phase-shift circuit which can vary the phase of the subcarrier applied to the I demodulator over a range of about ±57° around the nominally correct position. A *limiter* is used to remove any incidental subcarrier amplitude changes resulting from the operation of this hue control circuit.

In the normal home environment, a user setting of the hue or tint control on a color receiver is usually made on the basis of subjective evaluation of human skin tones in the picture. The human visual system has wide tolerances for many other colors (as long as they remain plausible), but skin tones that are either greenish or purplish are generally considered quite objectionable. A dynamic flesh-tone correction circuit is included in the typical luma/chroma IC to minimize the need for frequent adjustment of the user's hue control.

If it were not for amplitude limiters at both the chroma and carrier inputs, the *flesh phase detector* would produce an output signal roughly equivalent to that of the I demodulator. It happens that normal skin tone colors fall rather close to the I axis on the NTSC color vector diagram: more accurately, about 7° closer to red. Because of the limiters, the flesh detector output responds to phase information only, and this output serves as the video input to a special flesh-tone correction modulator for which the carrier input is a sample of the chroma signal from which the amplitude modulation has been removed by a limiter. This special modulator has its bias adjustment offset in such a way that only positive excursions in the output of the flesh phase detector will produce phase-shifting vectors which are added to the main carrier signal enroute to the I and Q demodulators. These phase-shift vectors will tend to pull the reproduced colors toward the flesh-tone axis, as indicated by Fig. 25.

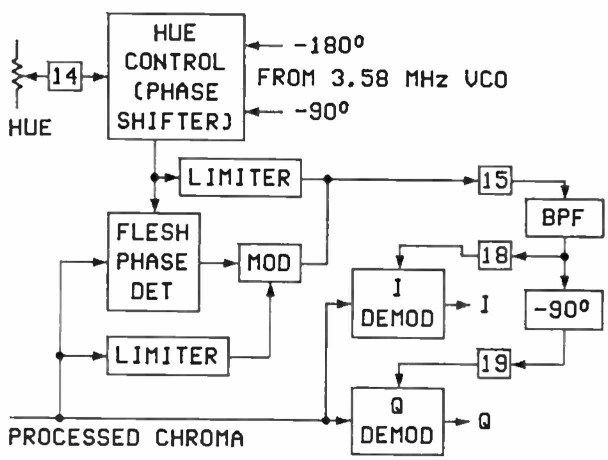


Figure 24. Partial block diagram of a typical luma/chroma IC, showing the subcarrier processing circuits and the chroma demodulators.

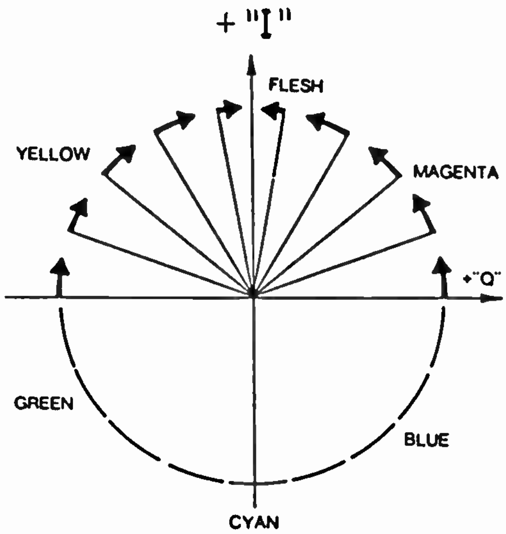


Figure 25. Vector diagram illustrating the operation of the dynamic flesh-tone correction circuit in a typical luma/chroma, IC.

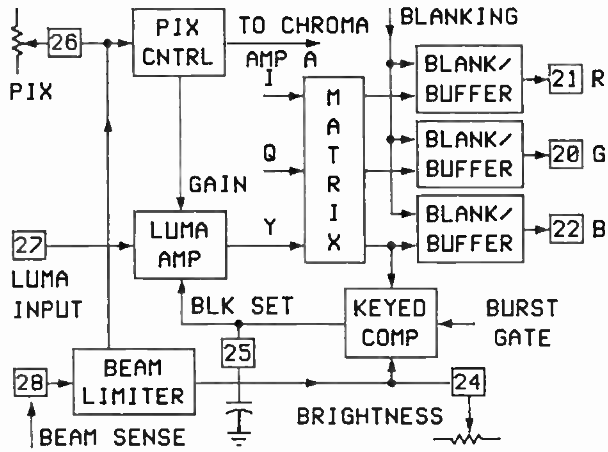


Figure 26. Partial block diagram of a typical luma/chroma IC, showing the luminance processing circuits, matrix circuit, and blanking/buffer output amplifiers.

An external 3.58 MHz bandpass filter connected between Pins 15 and 18 removes any harmonics that may be generated by the limiter and flesh-tone correction circuits. An external 90° delay network between Pins 18 and 19 establishes the correct phase relationship between the I and Q demodulators, which provide the baseband chrominance signals.

Luminance Processing and Matrixing

The heart of the luminance processing circuits is an amplifier which features a keyed clamp and electronically variable gain control. (See Fig. 26.)

Manual control of the gain of the luminance amplifier is provided by the user control labeled *picture* or *contrast* (shown in this diagram as PIX). The same control will simultaneously adjust the gain of the chroma channel as well, so that color saturation (which is controlled by the chrominance/luminance ratio) remains constant.

DC restoration (or black-level stabilization) for the video portion of the receiver is provided by a feedback clamp which compares the level in the *blue* output channel with a reference voltage set by the user's *brightness* control during the burst gating period. The resulting feedback signal, filtered by an external capacitor at Pin 25, controls the DC bias of the luminance amplifier. Although the sampling is done only in the blue channel, all three primary-color output channels are effectively stabilized because their relative levels are accurately controlled within the matrix amplifiers.

An automatic control voltage proportional to the picture tube beam current is brought in to Pin 28 for the *beam limiter*, which prevents overdrive of the picture tube. For moderate overload conditions, the beam limiter overrides the manually-set picture or contrast control to reduce the signal levels simultaneously in both the luminance and chrominance channels. For greater degrees of overload, the beam limiter will also override the manual brightness control to keep the picture tube operation within safe limits.

The objective of the matrix functional block is to cross-mix the Y, I, and Q signals to produce red, green, and blue video signals. A more detailed diagram would show that the matrixing is accomplished in two stages: the I and Q signals are first matrixed to form R-Y, G-Y, and B-Y signals, which are then combined with Y to yield the primary-color signals.

The *blanking/buffer* amplifiers use the composite blanking signal decoded from the sandcastle input to remove the synchronizing pulses and to hold the output levels during horizontal and vertical retrace periods well below the picture tube's cutoff levels to assure the invisibility of retrace lines.

COMB FILTER APPLICATIONS

Brief Summary of Background Theory

As noted in the commentary associated with Fig. 20, the traditional approach to the separation of chrominance and luminance signals involves a number of practical compromises, including some limitations on luminance resolution, residual dot-crawl effects at the edges of colored areas, and problems with visible cross color effects resulting from certain types of luminance signals within the chroma channels. Such problems can be greatly reduced by the use of comb filters in the circuits which separate luminance and chrominance signals. The principles on which comb filters are based were known to the NTSC pioneers who prepared the color television signal specifications, but their practical implementation had to await the development of low-cost ultrasonic and charge-coupled devices.

Stated quite simply, comb filters are useful in color television systems because the spectral energy of video signals resulting from scanning in a repetitive pattern tends to be concentrated in lines representing harmonics of the line, field, and frame frequencies, with particularly strong spectral lines found in packets around the harmonics of the line scanning frequency. The color subcarrier frequency was deliberately chosen so that it falls midway between two line-frequency harmonics, and the chrominance video signals modulated on this subcarrier produce spectral packets that tend to fall midway between the packets representing the spectral energy of the main luminance signal. Thus comb filters with teeth at line frequency intervals offer an attractive means of separating luminance and chrominance signals more effectively than simple notch and bandpass filters. These concepts are illustrated by the simplified spectrum sketches in Fig. 27.

The sketch at A represents a typical luminance signal in which the signal energy tends to be concentrated around harmonics of the horizontal frequency, which are identified by the numbers below the horizontal scale. The sketch is simplified in that the discrete spectral lines may be only 29.97 Hz apart.

The sketch at B represents a small section of the chrominance signal, which is modulated on a pair of suppressed subcarriers at 3.58 MHz (or 455/2 times the horizontal frequency). Note that the sideband energy is bunched at line-frequency intervals, and thus tends to fall within the gaps of the luminance spectrum.

The combined spectrum is shown at C. Note that the two spectra are frequency interlaced both with respect to the line-frequency bunches and with respect to the individual spectral lines.

The concept shown here applies rigorously only to still images. The spectral lines of signals developed from images in motion become frequency-modulated by the motion itself, and thus their positions are not as easily predicted. In general, however, there is strong statistical correlation from line to line and from frame to frame in nearly all television signals, so comb filters which exploit the frequency interlaced properties of NTSC signals work quite well (but not perfectly) in most practical applications.

Basic Comb Filter Systems for Luma/Chroma Separation

The heart of a comb filter for television applications is a device capable of delaying a signal by precisely one horizontal period (about 63.5 microseconds) which can be used in association with analog *adder* and *subtractor* circuits, as shown in Fig. 28.

When a video signal that has been delayed by 1H is added to an undelayed signal, a null appears in the combined frequency response characteristic at the frequency fH/2, since the delayed and undelayed signals exactly cancel each other at this frequency. As shown in the sketch labeled Y above, an additional null appears at each odd multiple of half the line frequency thereafter, producing a comb-like response curve in which the nulls are spaced at horizontal-frequency intervals. (Only a small part of the video spectrum is shown in the sketch.) Phasor analysis techniques may be applied to show that the teeth of the comb are shaped like half sine waves. Such a response curve can do a good job of recovering luminance signal components while attenuating chrominance signal components.

When a video signal that has been delayed by 1H is subtracted from the undelayed signal, the frequency response available at the subtractor output has a similar comb-like shape, but the null points occur at zero frequency and all even multiples of one-half the horizontal frequency, as shown in sketch C above. In the high-frequency region near the color subcarrier, this configuration is very effective for recovering the frequency-interlaced subcarrier sidebands while rejecting the luminance components.

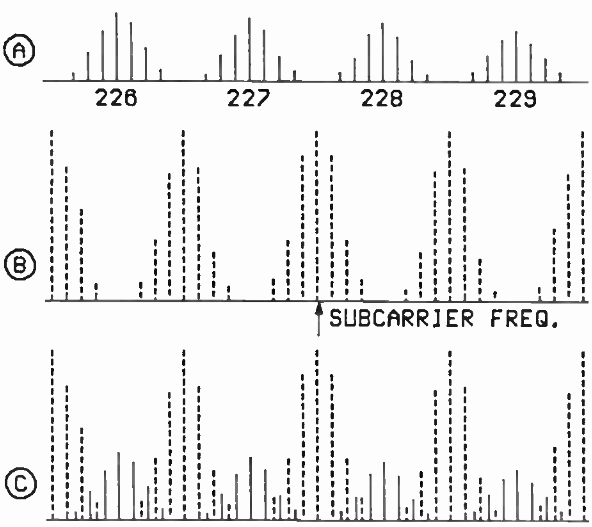


Figure 27. Simplified sketches of a video spectrum, showing only a small zone near the subcarrier frequency.

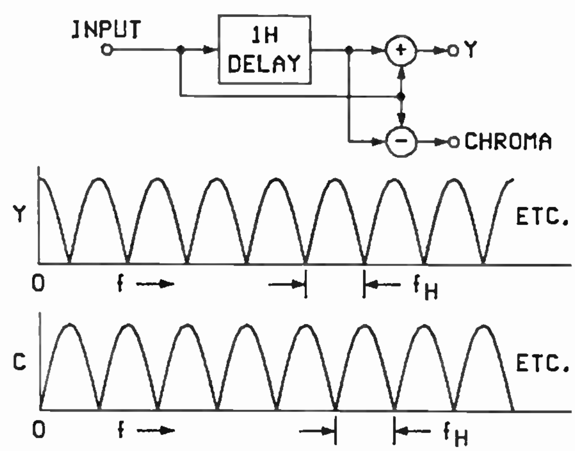


Figure 28. Simplified block diagram and output spectrum sketches for a comb filter luma/chroma separation system using a single 1H delay element.

Somewhat more elegant comb filter systems incorporating two 1H delay elements are commonly used in color television broadcast equipment (for such applications as luminance aperture compensation for cameras and chroma/luma separators for professional-quality video recorders). The key advantage of such systems is the creation of wider nulls between the peaks of response in the comb, but these more sophisticated filters are seldom found in home receivers because of the higher costs involved.

Comb filter separation of luminance and chrominance is based on the assumption that most of the content of both the luminance and chrominance signals is repeated from line to line, but that the phase of the subcarrier on which the chrominance is modulated reverses from one line to the next (because the subcarrier frequency is an odd multiple of ½ the line frequency). Fig. 29 shows only a couple of cycles of the modulated chrominance signal, greatly stretched out for clarity.

When two successive lines are added, the luminance components will reinforce each other, while the chrominance components will cancel.

When two successive lines are subtracted, the chrominance components are reinforced while the luminance components are canceled.

In typical-case picture signals with small changes from line to line, the comb-filter separation of luminance and chrominance is less perfect than illustrated here, but still much more satisfactory than with simple low-pass and bandpass filters.

Practical Implementation of Comb Filters

Practical 1H delay elements can now be fabricated with either ultrasonic or charge-coupled device (CCD) technology. Interest in such devices in mass production quantities was stimulated by the PAL and SECAM color television systems in the international market; these broadcasting systems require at least one such device in every receiver. The ultrasonic devices used in many PAL or SECAM receivers are based on technology generally similar to that in SAW filters, but they often use the bulk properties of piezoelectric materials rather than the surface properties.

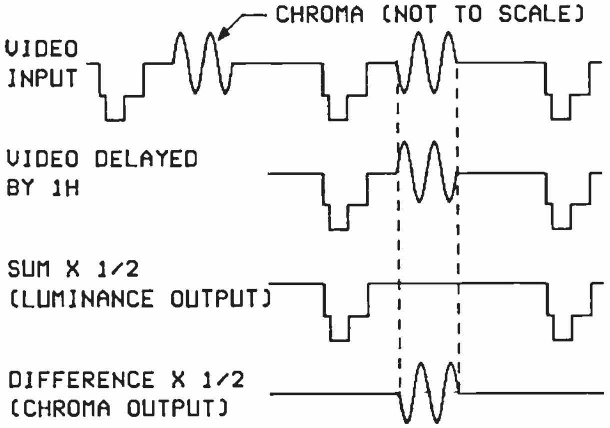


Figure 29. Simplified comb filter waveform sketches illustrating chroma/luma separation in the time domain.

Among American receiver manufacturers, CCD versions of 1H delay devices are more popular than ultrasonic devices. Ultrasonic devices are relatively bulky, they usually require modulator/demodulator arrangements to place the signal within the limited bandwidth of the device itself, and meticulous design engineering is required to avoid problems with spurious responses and temperature stability. CCD devices, on the other hand, can be designed to work directly with baseband video signals, their timing can be controlled by reference to the subcarrier of the signal actually being handled, and they can be manufactured in packages compatible with those that house ordinary integrated circuits. The working principles involved in typical CCD delay elements are shown in Fig. 30.

The incoming video signal is *sampled* at a three-times-subcarrier rate (10.7 MHz) to produce a time series of *charge packets*, which are then propagated down a very long charge-coupled device (usually fabricated in serpentine form) clocked at the same 10.7 MHz rate. The CCD device is nominally equivalent to a shift register capable of handling variable charges, which are converted back to an analog video signal by sample-and-hold circuits and filters at the output end of the device.

The CCD delay devices in most current designs employ a two-phase clocking system, requiring two sets of special gate electrodes per delay stage. Charges are constrained to move in one direction only by using two separate polysilicon gates for each electrode, one of which is placed over a region where an ion-implanted barrier has been created in the underlying N-channel. Readers needing more comprehensive understanding of the charge-movement process and the techniques used to form the charge packets at the input, and to recreate a normal video signal at the output, should consult the literature supplied by the current manufacturers of charge-coupled delay devices.

Because the NTSC signal specification set the color subcarrier frequency at 455/2 times the horizontal frequency, there are 3 × (455/2) or 682.5 cycles of the 10.7 MHz clocking signal in one horizontal period, and 682.5 CCD stages will provide the desired 1H delay. The extra half-cycle is not a practical problem, since the physical structure is already divided into half-cycle segments; the output is simply derived from a stage at a clock phase opposite from that used at the input stage.

The type of CCD delay element described in Fig. 30 is shown in Fig. 31 in the more complete context of a luma/chroma separation system, which includes an arrangement for restoring and enhancing vertical detail in the luminance system.

The *gain control amplifiers* serve to match the gains of the two channels feeding each summation point. In practice, the gain controls are set for optimum nulling or attenuation of the undesired components in each channel.

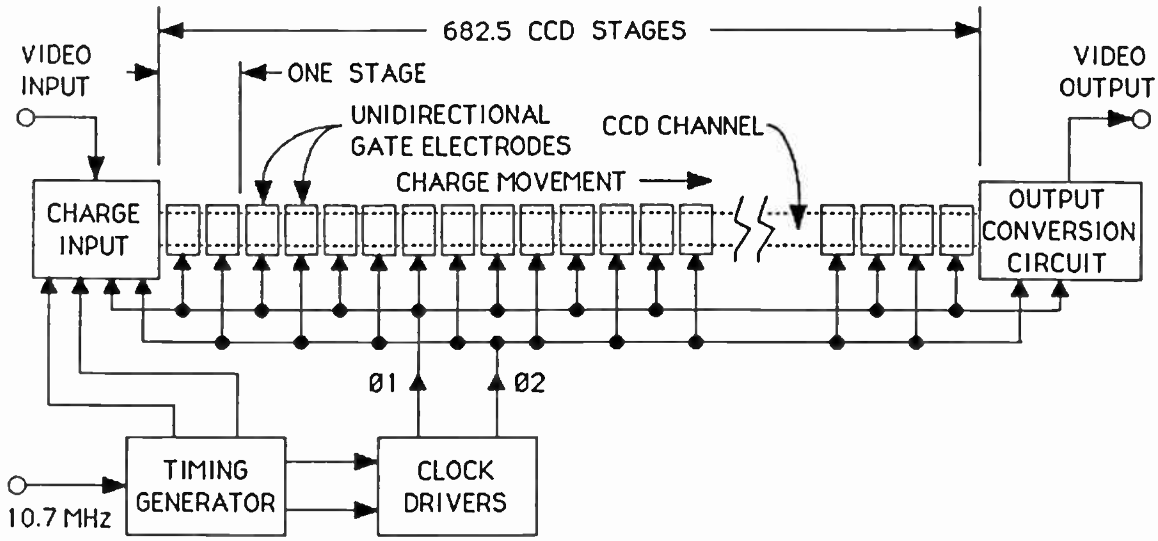


Figure 30. Simplified block diagram showing the basic principles involved in a charge-coupled device (CCD) delay line.

The summing point just above the chroma (C) gain control amplifier actually serves as a *subtractor* because of the analog inverter shown at the bottom of the diagram. The inverter which follows this summing point reverses the subtraction, yielding the direct signal minus the delayed signal instead of the other way around.

The output of the subtractor is separated into two signals by filters: chrominance information is recovered through a bandpass filter, and a vertical detail signal that should be added back to the luminance channel is recovered through a low-pass filter.

The processes of vertical detail restoration and enhancement (which involves the *nonlinear processor* shown in Fig. 31) will be described in the next subsection of this chapter.

Vertical Detail Restoration and Enhancement in Comb Filter Systems

At first glance, the inclusion of this topic in this brief broad brush treatment of television receiver technology may appear unwarranted, since we are dealing here with a design detail that is actually found only in relatively few top of the line receivers. This author believes, however, that broadcast engineers and technicians should give special attention to receiver features which not only correct picture defects that might arise from practical problems within the receiver itself but which might actually enhance the signals delivered by broadcasters.

One problem in simple comb-filter systems is a blurring of vertical detail resulting from the addition of pairs of successive lines that are not truly identical.

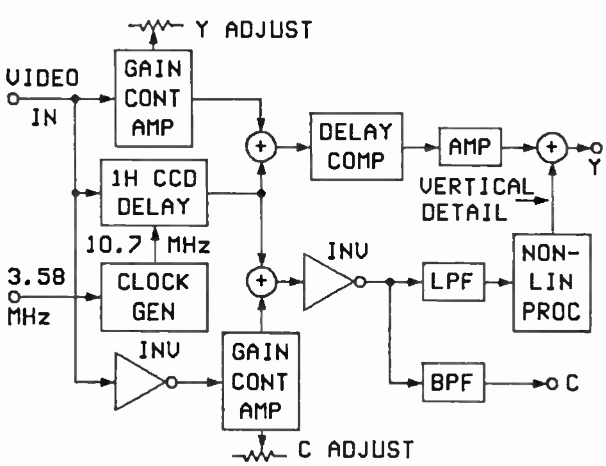


Figure 31. Simplified block diagram of a CCD-based comb filter system for luminance and chrominance separation.

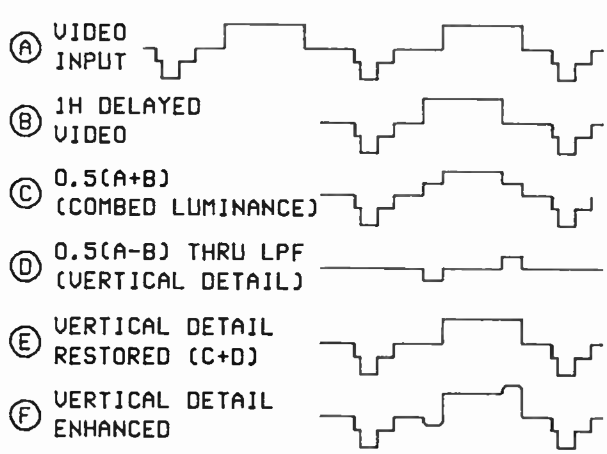


Figure 32. Simplified waveform sketches illustrating the processes of vertical detail restoration and enhancement in comb filter luma/chroma separation systems.

(Keep in mind that two adjacent lines in time are separated by two lines in space because of line interlace.) In the simplified waveform sketches in Fig. 32, the problem of degraded vertical detail is illustrated by an image feature whose luminance pulse for the second line is appreciably to the right of the position for the first line.

The basic combed luminance signal shown at C suffers from degraded vertical detail in that the edges of the luminance pulse are different from those in both the direct and delayed signals; the edges are clearly less distinct.

A basic correction signal to restore the degraded vertical detail can be produced by passing the direct minus delayed signal (i.e., the same signal from which chrominance information is recovered) through a low-pass filter and adding the result to the combed luminance (as shown at D and E). The corrected signal is a closer approximation to the true luminance component of the undelayed signal for the second line as shown at A.

Laboratory experiments and field trials with modern receiver circuits have shown that a worthwhile improvement in subjective picture sharpness can be achieved by increasing the level of the vertical detail signal slightly beyond the value required for restoration, thus achieving a moderate degree of enhancement. The same experiments have shown, however, that the degree of enhancement that can be effectively used is increased if the vertical detail signal used for enhancement is passed through a special nonlinear processing amplifier. The portion of the signal used for restoration should not have nonlinear processing.

This characteristic describes the experimentally-determined optimum performance for the enhancement channel within the nonlinear processor shown in Fig. 31.

A vertical detail signal normally has no DC component because it is produced by a comb filter that has a null at DC (and at all harmonics of the line frequency). Thus the signal will tend to swing both plus and minus (by nominally equivalent amounts) with respect to its own axis.

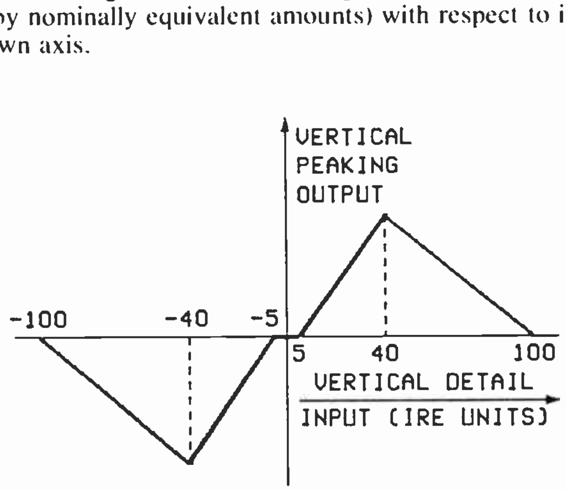


Figure 33. Transfer characteristic for a typical nonlinear processing amplifier for vertical enhancement signals.

There are three distinct zones in each half of the input voltage domain:

1. Zone I (sometimes called the coring zone): there should be no output for input signals of less than 5 IRE units in either direction.
2. Zone II: there should be positive gain for signals in the range of 5 to 40 IRE units in either direction.
3. Zone III: there should be negative incremental gain for signals greater than 40 IRE units, designed to yield zero output for signals of 100 IRE units.

The coring zone assures that low-level noise and unwanted drift signals are not amplified.

The gradual reduction in enhancement level for Zone III assures that the enhanced output signal will remain within the normal 100-IRE-unit range between the blanking and reference white levels.

COLOR PICTURE TUBES AND VIDEO DRIVERS

Color Television Display Systems

The great majority of color television receivers now in service use one of three basic types of display systems:

1. Direct-view picture tubes of the shadow-mask type. A tube in this category uses a cluster of electron guns (usually mounted in a horizontal row in recent designs) to produce three separate beams to excite red, green, and blue phosphors which are deposited in the form of either dots or vertical stripes on the inside surface of the tube's glass faceplate. Spaced about 0.5 inch or so behind the faceplate is a shadow mask which contains either one round hole for each cluster of three dots (one each in red, green, and blue) or a column of elongated vertical slots for each group of three stripes (again one each in red, green, and blue). This entire screen system is direction-sensitive. Each of the beams passes through the deflection plane in the scanning system at a slightly different point in space, so each beam approaches the shadow mask at a slightly different angle. The electron optics and the scanning system within the tube are set up in such a way that each beam can illuminate only the dots or stripes of its assigned color: electrons within each beam that might otherwise land on a phosphor of the wrong color are intercepted by the shadow mask. The dots or stripes are too fine to be resolved separately at normal viewing distances, so the viewer perceives an additive mixture of red, green, and blue from each area of the screen.
2. Direct-view picture tubes of the Trinitron type. (Trinitron is a registered trademark of the Sony Corporation.) Tubes in this category have a single electron beam that is time-shared between red, green, and blue video signals in synchronism with switching signals applied to a wire grid structure

mounted behind the screen, which consists of fine vertical phosphor stripes deposited on the inside surface of the glass faceplate. A set of three phosphor stripes (one each in red, green, and blue) is provided for each vertical slot in the wire grid structure, and the switching signals applied to the grid deflect the beam just enough to cause it to land on a phosphor stripe of the right color each time the beam current itself is switched from one primary channel to another.

3. Projection-television in which three small but very bright displays on separate red, green, and blue picture tubes are projected optically by suitable lens-and-mirror systems to produce three registered images on a common screen. Receivers using this projection principle are usually more expensive that the direct-view types, but they can produce larger images for somewhat greater audiences.

The following is a very brief discussion of direct-view displays of the shadow mask type, along with a brief description of a typical video driving circuit for such displays. The major features of a typical shadow-mask color picture tube are shown in Fig. 34.

Proper alignment between the phosphor stripes and the slots in the shadow mask is assured by using the mask itself as an optical template in the multi-step screen manufacturing process. To expedite precision manufacture of the screen/mask subassembly, the faceplate end of the tube is made separately from the funnel. The two sections are eventually joined at the seal line shown in the sketch by coating the surfaces to be joined with a glass frit (a material with a melting point somewhat below that of the structural glass components) and applying heat.

The three sections of the electron gun have separate cathodes (for independent control of the red, green, and blue beams), and the electrostatic lenses needed to define the beams are kept separate near the cathode end of the structure, but common lens structures may be used near the high-voltage end of the gun structure.

The *convergence device* near the exit end of the electron guns is needed to assure that the beams converge properly along the center axis of the tube so that the red, green, and blue spots appear to be properly superimposed. (Even when well focused, the beams produce scanning spots large enough to cover multiple stripes of the designated colors.) In many modern receivers, this device takes the form of a band of magnetic material in which the required magnetic pattern can be permanently set by an automatic jig in the factory. Should field service ever require removal of the deflection yoke, this simple band must be replaced with a more complex adjustable magnetic device.

In modern receivers, proper convergence all the way out to the edges of the picture is maintained by (1) proper shaping of the magnetic fields produced by the

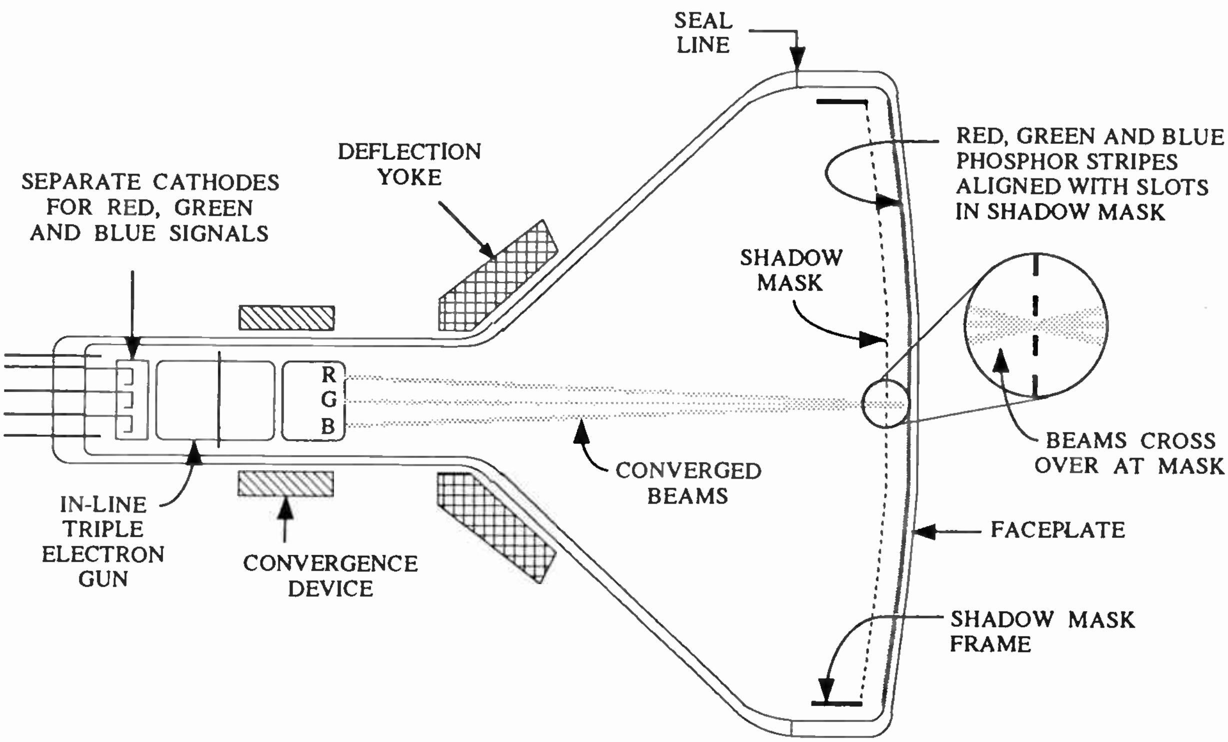


Figure 34. Simplified cross-section view of a typical shadow-mask color picture tube (as seen from above), showing also the major external components.

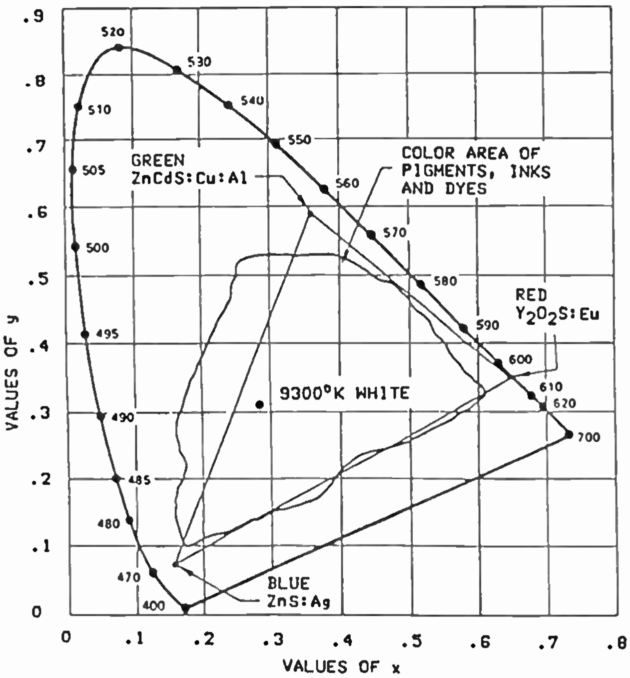


Figure 35. CIE chromaticity diagram showing the gamut of colors reproduced with a typical color picture tube phosphor set in comparison with the gamut of common pigments, inks, and dyes.

deflection yoke, (2) precise placement and sealing of the yoke to the tube during manufacture, and (3) proper shaping of the deflection waveforms to compensate for the greater distances from the deflection plane to the screen when scanning near the outer edges.

At least 20 different phosphor compounds have been used for consumer-type picture tubes throughout the history of color television, but the set shown in Fig. 35 is representative of current practice.

The red phosphor shown is yttrium oxysulfide:europium, the green phosphor is zinc cadmium sulfide:-copper:aluminum, and the blue phosphor is zinc sulfide:silver.

In modern phosphor sets, the range of reproducible colors has been slightly compromised (by moving the chromaticity points for the red and green phosphors closer to each other) to gain the important advantage of greater screen brightness for whites and other low-saturation colors.

From a practical point of view, the current color television primaries do quite well in reproducing most of the range of colors that are found in nature and that can be produced with conventional inks, pigments, and dyes. The blue-green colors that fall outside the color TV primary triangle are of relatively little practical importance.

Even though the primaries now used in actual practice differ significantly from those specified in the NTSC standards for the guidance of camera designers, most receiver engineers acknowledge that the colorimetric properties of signals as broadcast should have consistency: they do not expect camera designers to make adjustments to follow revised practices at the receiving end. Should a particular phosphor set fail to win consumer acceptance, receiver designers themselves can make compensating adjustments in matrix circuits or elsewhere to improve the subjective appearance of their pictures.

Video Drivers

The basic requirements for color picture tube drivers may be summarized as follows:

1. Amplify video signals to the levels required for cathode drive of the picture tube (typically 40 V to 100 V peak-to-peak).
2. Provide red, green, and blue gain and bias controls with the ranges required to provide proper RGB beam currents.
3. Provide bias stability. Some full feature receivers now provide automatic kinescope bias (AKB).
4. Provide adequate video bandwidth.

As in other receiver functional blocks, great variations in video drivers can be observed between different manufacturers and even between the standard and deluxe product lines of a single manufacturer. For purposes of this chapter, we shall limit our attention to a representative example of a driving circuit from a recent standard receiver.

In the context of the receiver from which Fig. 36 was taken, the objective of the green driver circuit is to raise the level of the green video signal from a nominal 4 V peak-to-peak (p-p) at the output of the luma/chroma IC to about 40 V p-p at the picture tube cathode. (The red and blue driver circuits in the same receiver need output voltages in the range of 50–60 V p-p because of lower phosphor efficiencies.) Gain may be varied over a suitable range by the DRIVE control, which varies the amount of degeneration (a form of negative feedback) at the emitter of Q1.

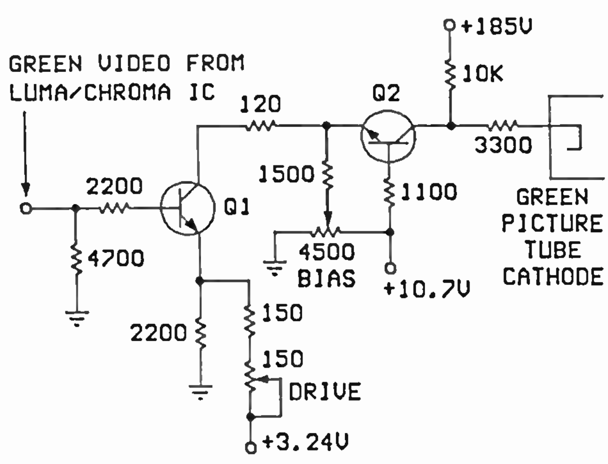


Figure 36. Simplified schematic diagram of a typical green video driver.

The basic circuit configuration shown here is a cascode; a common-emitter stage driving a common-base stage. This configuration provides good bandwidth, partly because the capacitive loading of the output is well isolated from the input. There is no significant Miller effect from the base-to-collector capacitance of Q1 because the load for Q1 is the very low input impedance of Q2.

This circuit is direct-coupled throughout, and the BIAS control sets the operating points for both Q2 and the picture tube cathode.

In the receiver from which this example was derived, Q1 is mounted on the main circuit board and Q2 is mounted on a small board as part of the kinescope socket assembly. Because the interface between Q1 and Q2 is a low-impedance point, stray capacitive loading of the moderately long leads required to reach the socket board is not a serious practical problem, nor is there any significant risk of excessive electromagnetic radiation from high-level signals on these leads.

DEFLECTION AND POWER SUPPLY SYSTEMS

Sync Separation

In essence, a *sync separator* is a specialized clipping circuit designed to recover the synchronizing pulses from a composite video signal, discarding the actual picture information. The sync separators used in color television receivers are generally more complex than those found in most categories of broadcast equipment because of the many hazards to which the signal has been subject before reaching the sync separator terminals. A typical example of the sync separator found in a modern receiver is shown in Fig. 37.

The active device in this sync separator is transistor Q305, which functions as a switch between cut-off and saturation levels. The 4.5 V p-p video signal available from the detector is at a sufficiently high level that the voltage range from the tip of sync to blanking is more than enough to drive the transistor from cut-off to saturation.

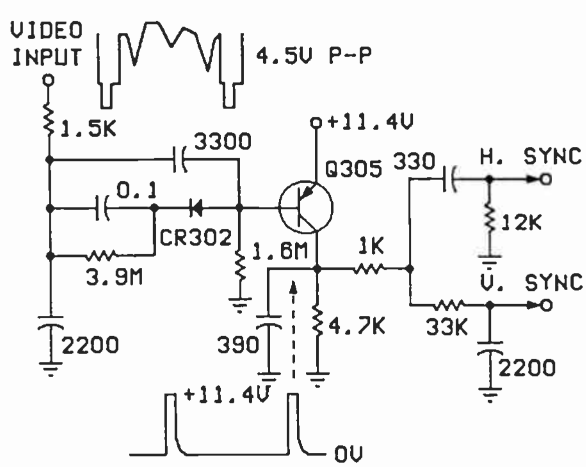


Figure 37. Schematic diagram for a typical dual time constant receiver sync separator circuit.

The bandwidth of the incoming video signal is limited to about 500 kHz by the integrator consisting of the 1.5 K resistor and 2,200 pf capacitor at the input. This bandwidth limitation improves the noise performance of the circuit, and assures that the color synchronizing burst will not affect the sync separation process. (Noise protection is also provided by the noise inverter circuits in the IF functional block, as explained in the notes for Fig. 12.)

The video signal is clamped at the sync tip level by the action of the 3300 pf coupling capacitor in association with the base-to-emitter junction of Q305. At its most negative excursions (i.e., at the sync tips), the video signal causes base current to flow, and this current charges the capacitor enough to make the absolute voltage at the base approximately + 10.8 volts (one diode drop below the emitter at 11.4 volts) during sync tip time. The time constant at this base (established by the coupling capacitor in association with the 1.6 M resistor to ground) is large enough that the bias is sustained during the subsequent line interval, causing the more positive portions of the video signal (including all of the picture information) to rise above the cut-off level for the transistor. On the other hand, this time constant is kept short enough to assure reliable recovery of horizontal sync pulses even with rapid fluctuations in signal level, especially the airplane flutter resulting from moving objects in the propagation path between transmitting and receiving antennas.

A second (and somewhat larger) time constant for the clamping action is formed by the 0.1 μf capacitor shunted by a 3.9 M resistor. The DC clamping reference for this time constant is one diode drop less positive because CR302 is connected in series with the base lead. This longer time constant is effectively switched in only during the vertical sync interval, and assures reliable recovery of vertical sync pulses.

The separated sync pulses at the collector of Q305 swing between cut-off and saturation levels (0 V and 11.4 V, respectively). The rounding of the waveform as it moves from saturation to cut-off results both from charge storage within Q305 and from the deliberate roll-off in frequency response introduced by the 390 pf capacitor across the 4.7 K load resistor (again provided for noise immunity reasons).

The circuit includes a simple decoder (consisting of elementary differentiator and integrator networks) to separate the horizontal and vertical components of the composite sync signal. The horizontal sync output actually consists of differentiated pips from all leading and trailing edges of the sync waveforms.

Typical Circuits for Deflection Control

As in the case of many other functional blocks within the receiver, a great many of the circuits required to implement modern horizontal and vertical deflection systems are housed within a single integrated circuit. Using simplified block diagrams, we shall describe

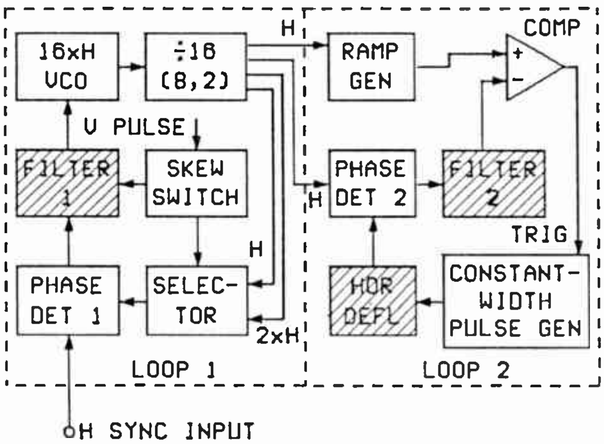


Figure 38. Simplified block diagram for the major horizontal circuits in a typical deflection IC.

most of the major features in a typical such IC, suitable for use in a standard receiver (as opposed to a deluxe model), but one which will accept nonstandard input signals, such as those from VCRs or video games. Our simplified diagrams will also show how the integrated circuits relate to those components or circuits which cannot reasonably be housed within the IC package, including the high-power output stages and the deflection coils themselves.

The horizontal control circuits in Fig. 38 are organized into two phase-locked loops (PLLs), identified here as Loop 1 and Loop 2. In essence, Loop 1 provides immunity from disturbances in the incoming sync signal, while Loop 2 provides compensation for horizontal timing variations that might result from variable loads on the power supply, which is closely coupled to the horizontal output system. Circuit functions shown within the shaded blocks are external to the deflection IC.

The phase detector in Loop 1 compares the incoming horizontal sync pulses with the output of a 16 × H voltage controlled oscillator (VCO) as observed through a ÷ 16 counter and a digital selector. The external filter associated with this detector requires a relatively long time constant (or a narrow bandwidth) to provide good noise immunity.

The *skew switch* in Loop 1 is controlled by a pulse from the vertical decoder (to be shown in Fig. 39) that identifies the vertical sync period. This skew switch does two things: (1) it operates a digital selector to choose 2 × H pulses instead of H pulses during the vertical interval so that every leading edge of the equalizing and vertical sync pulses can contribute to useful gain at the detector output, and (2) it lowers the time constant (or widens the bandwidth) of filter 1 during the same interval. This is particularly helpful when the receiver is connected to a consumer-type VCR. Such machines frequently have abrupt timing errors (frequently called skew errors) at the head-to-head switching points (near vertical sync) resulting from improper tape tension, and the shorter time constant permits more rapid recovery from these timing errors.

The phase detector of Loop 2 compares a horizontal-rate output of Loop 1 with a timing feedback signal from the horizontal deflection subsystem. Filter 2 has a relatively short time constant (or wide bandwidth) so that this PLL can respond quite rapidly to timing changes resulting from the variable load on the horizontal deflection output, which also supplies much of the power for the receiver. A major component in this variable load is the picture tube beam current, which varies with picture content. Obviously, it is desirable for the scanning pattern to remain fixed in size and position, independent of picture content. Loop 2 helps to maintain this desirable condition.

The combination of the horizontal *ramp generator* and *comparator* within Loop 2 form an automatic phase shifter to adjust the timing of the trigger pulses fed to the *constant-width pulse generator* in such a way as to maintain consistent timing between the two signals compared by *phase detector 2*.

Key details of the *horizontal deflection* circuit and the closely related integrated high-voltage transformer (IHVT) system will be shown in later diagrams.

As in the previous figure, the shaded blocks in Fig. 39 represent circuit functions external to the IC.

To improve vertical noise immunity, the vertical sync input to the mode control circuit is clamped to ground (or logic low) at all times except during a window extending from a count of 464 to a count of 592 as determined by a decoder associated with the 10-stage vertical counter, which is clocked by 2 × H pulses derived from the noise-immune horizontal system. Any valid vertical sync signal (including one from any reasonable nonstandard video signal) will occur during this window. The window will be closed earlier than a count of 592 if a *reset* pulse is applied to the counter: this will normally occur at a count of 525.

If there should be no incoming vertical sync signal (if the receiver is not tuned to an active channel, for

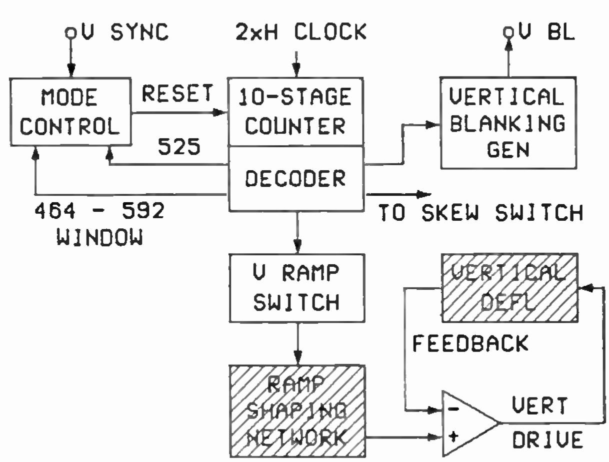


Figure 39. Simplified block diagram of major vertical control circuits within a typical deflection IC.

example), the vertical system will free run at a count of 592, yielding a vertical frequency a little slower than normal. This behavior is very similar to that of a traditional vertical deflection system, where the free-running frequency of the injection-locked vertical oscillator is normally set a little below the normal value.

The arrangement of digital circuits within the mode control block is such that after two consecutive vertical sync pulses coincide with count of 525 pulses from the decoder, the control unit steers the count of 525 pulses directly to the reset terminal of the ten-stage counter, but continues to compare the incoming vertical pulses with the internally-generated reset pulses. Only if the coincidence is NOT found for seven consecutive vertical intervals will the circuit use the incoming pulses to reset the counter. This arrangement provides excellent noise immunity for the vertical system. Several pulses can be missing or distorted without disturbing the continuity of the vertical scanning action, but a necessary adjustment (as when the receiver is tuned to a different channel) can be accomplished in a fraction of a second.

Separate outputs from the vertical decoder control the *vertical blanking generator*, the *vertical ramp switch* and the *skew switch* (in the horizontal subsystem).

External RC components are used in a *ramp shaping network* to form a ramp (or sawtooth) signal of the appropriate shape to produce a current waveform through the vertical deflection coils that is of optimum shape to produce a linear scanning action. (Because of geometric factors in the electron optics of the tube, the optimum shape is not a strictly linear sawtooth.) The vertical drive signal applied to the external vertical deflection circuit is derived from a differential amplifier that compares the applied ramp signal to a feedback signal proportional to the yoke current; this amplifier modifies the drive signal in such a way as to minimize any discrepancy between the compared signals. Only the final output stage of the vertical output, the vertical coils themselves, and the current-sampling resistor are external to the IC.

In many modern receivers, the *horizontal rate electronic switch* shown in Fig. 40 takes the form a common-emitter transistor switch which is driven from cut-off to saturation by a pulse applied to its base. In the system under discussion here, the base would be driven by the output of the *constant-width pulse generator* shown in Fig. 38. The output transistor itself must be of a type that can tolerate abnormally high inverse voltages when cut off.

The battery symbol VCC represents the regulated power supply which provides B+ power for the output transistor. In the system under discussion, this horizontal output power is derived from a rectified 60 Hz power supply that has been well regulated by an silicon controlled rectifier (SCR) triggered at a horizontal rate under the control of a pulse-width regulator included in the deflection integrated circuit.

Energy is fed into the yoke while the electronic switch is closed during approximately the second half

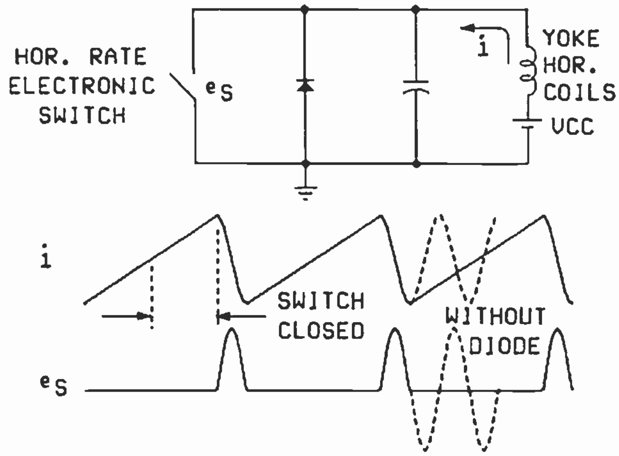


Figure 40. Simplified schematic diagram and related waveforms for a basic horizontal deflection output circuit.

of each horizontal period. The current builds up at a rate determined by the L/R time constant of the system.

When the switch is opened, the magnetic field around the yoke starts to collapse, and the self-induced voltage in the coils proceeds to build up a charge in the capacitor. During a period equal to a half-cycle at the resonant frequency of the LC combination, the energy from the magnetic field is transferred first to the electrostatic field in the capacitor, and then back to the magnetic field in the opposite polarity. (A small fraction of the energy is lost, of course, because of resistance in the system; that is why a certain amount of energy must be drawn from the regulated power supply in each cycle.) There is a 90° phase relationship between the current through the coils and the voltage across them during this resonant period.

During the second half of the resonant period after the opening of the switch, the damper diode across the switch becomes forward-biased by the self-induced voltage in the coils, causing the yoke current to change gradually in accordance with the system's L/R time constant. If the diode were removed, both the current and voltage waveforms would continue to oscillate until the switch is closed again, as suggested by the dotted lines in the waveform sketches.

In real life, of course, the challenge facing the designer of a practical horizontal output system is to produce a current waveform through the coils of exactly the right shape to produce linear deflection of each horizontal line. The task is not a simple one, but the performance of millions of TV receivers verifies that the problems are, indeed, solvable.

Power Supply and Protection Systems

The power supply system for a typical modern receiver is so closely related to the horizontal deflection system that it is not really possible to discuss one subsystem without mention of the other. The ultimate source of power for most receivers is, of course, a 60 Hz power line. In a typical modern receiver, this power is initially processed through a relatively simple (and conventional) rectifier/filter combination, and a few circuits related to starting up the receiver are powered directly from this preliminary source. The horizontal deflection subsystem, which also supplies power for most of the critical circuits within the receiver, receives its own power from the original rectified source through an advanced type of switch-mode regulator (using an SCR) which is triggered by pulses at the horizontal rate.

In most modern tranistor-based horizontal output systems, the high-voltage transformer used for focus, screen, and regulator supplies is operated in parallel with the deflection yoke, and thus is an important part of the load on the horizontal output stage. A typical arrangement is shown in Fig. 41.

The B+ source shown near the bottom center of this diagram is the switch-mode regulated supply referred to in the previous paragraph. Note that the primary winding of the 1HVT is connected in parallel with the horizontal deflection yoke (and thus forms a part of the resonant circuit described in connection with Fig. 40).

The high-voltage winding is divided into sections, and rectifier diodes are connected between these sections. This technique helps to keep the peak inverse voltage on each diode within reasonable limits, and the windings themselves are carefully designed and encapsulated to minimize the risk of insulation breakdown. (The high voltage potential is typically at about 25 kV.) The filter for the HV supply consists of the capacitance between conductive coatings inside and outside the funnel portion of the picture tube. An intermediate tap on the high-voltage winding provides the FOCUS supply for the electron guns. Although not shown in this simplified sketch, the low end of the HV winding is actually returned to ground through a simple network which includes a current-sampling resistor that permits monitoring of the beam current.

The picture tube filaments are heated by AC energy at the horizontal frequency.

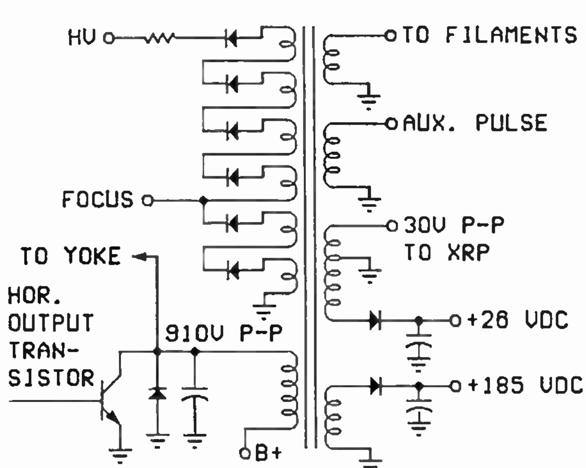


Figure 41. Simplified schematic diagram of an integrated high voltage transformer (IHVT), its driver, and related power sources for a modern receiver.

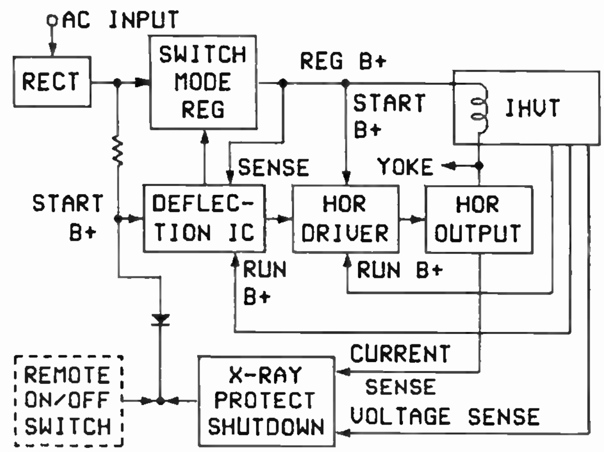


Figure 42. Simplified block for a typical receiver power supply system, showing start-up and protection features.

The 30 V p-p pulses to the X-ray protection (XRP) circuit also serve as the timing feedback pulses for the horizontal deflection IC.

Simple rectifier/filter combinations produce the DC supplies at +26 V and +185 V. (The latter is used for the video drivers.)

Not shown in detail on Fig. 42 are various other auxiliary functions, such as the turn-off pulses for the SCR used in the primary B+ regulator.

A simple bridge rectifier with a filter capacitor is connected directly across the AC input source to produce Raw B+ at about 150 volts. Through a dropping resistor, this source provides Start B+ for the horizontal oscillator system within the deflection IC. This oscillator will operate at its own free-running frequency even when there is no incoming sync signal. (As we have noted previously, the oscillator is literally at 16 times the horizontal frequency, but an internal ÷16 counter, also supplied from the Start B+ power source, produces horizontal-frequency output.)

Once the horizontal oscillator starts, trigger pulses become available for the SCR which functions as the switch mode regulator, allowing the Regulated B+ to appear, which serves as a start-up power source for the horizontal driver. (Regulated B+ is at about 118 volts.)

Once the horizontal driver becomes active, the horizontal output can begin to control the flow of pulses from the Regulated B+ source through the primary winding of the 1HVT (as well as through the horizontal deflection yoke). Rectifiers associated with appropriate windings on the 1HVT then produce the various supply voltages needed throughout the receiver, including the Run B+ sources for the horizontal driver and most of the circuits within the deflection IC. In the case of the horizontal driver, the Run B+ is at a somewhat higher voltage than the Start B+, and a diode automatically disconnects the latter when the former appears.

A small sampling resistor in series with the emitter of the horizontal output transistor produces a feedback voltage proportional to the current through the primary winding of the IHVT, and a simple rectifier driven by one of the secondary windings produces another feedback signal proportional to the high voltage output. The *X-ray protection shutdown* circuit monitors both of these feedback signals, and if either a current overload or an over-voltage condition is detected, the protection circuit triggers an SCR at its output which will pull down the Start B+ source for the horizontal oscillator, thus stopping the oscillator and shutting down the receiver's entire power system.

In a receiver equipped with remote control, the Raw B+ source remains operative at all times as long as AC power is applied to the receiver, and supplies the modest energy requirements of the remote-control receiver. The remote on/off switch function controls most of the receiver's power supply in the same manner as the X-Ray protection circuit.

CONCLUDING SUMMARY

This chapter was prepared to provide a brief survey of color television receiver technology for those broadcast engineers and technicians who are interested in what happens to their broadcast signals after they enter the homes of the viewing audience. There is truly a wealth of technical detail and many examples of clever engineering techniques in a modern television receiver, and progress is generally so rapid that trying to define the technology at any given moment is very much like aiming at a moving target. Truly fundamental principles do not change, however, and an engineer with a reasonable understanding of any one of today's receivers will have an excellent platform on which to build understanding of future developments in this field.

7.5 Closed Captioning Systems

Amnon M. Salomon
Gerald Freda
National Captioning Institute, Falls Church, Virginia

INTRODUCTION

Closed captioning is the depiction of the audio portion of a television program as text displayed on a television screen with the aid of a decoder that may be internal or external to a television receiver. Closed, as opposed to open, captioning means that the captions do not normally appear as part of the broadcast television picture. The viewer must have the proper equipment and must select the captioning mode. Closed captioned programs are compatible with other programs in that the addition of the captioning signal does not interfere with the regular audio or video signals. Digital data to create captions are transmitted with the television program on line 21 in the vertical blanking interval.

Open captioning refers to captions that are transmitted as part of the picture rather than as a separate signal. Such captions are visible to all viewers; an example would be the English subtitles in a foreign-language film.

The transmission of a data service in a broadcast signal directly to the home represents the first time a digital service has been broadcast for public use. The data rate has been set relatively low compared to a high speed teletext service in order to insure that the data is received with few errors in spite of noise, interference, and other impairments found in the terrestrial broadcast environment.

Captioned TV enables viewers to read the dialogue and narration of programs. The technique provides access to the entertainment, educational, and informational benefits of television for viewers who are deaf or hard-of-hearing. Other audiences can use captioning as well, including people learning English as a second language.

The captions produced by the closed captioning system generally appear in the lower portion of the television screen. The size of the characters varies in proportion to the size of the television screen. On a 19-inch screen, for example, they are about ½-inch high. The captions are easily visible—typically white letters against a black background—and usually do not obstruct essential parts of the picture (Fig. 1).

Closed captioning may be added in real-time to a live program or it may be added later as part of post production or distribution. It may be done in-house or by a captioning service.

A Brief History Of Closed Captioning

Closed captioned television technology was developed by the Public Broadcasting Service (PBS) during the period 1973–1979 with funding support from the Federal Government (Department of Health, Education and Welfare). Field test transmissions were conducted on all aspects of caption generation, encoding, decoding, and display features of the service. PBS and others provided captioning services for several years after approval of the system by the FCC in 1976.

The closed captioning service was launched in March of 1980 by the newly created National Captioning Institute (NCI), in cooperation with the ABC, NBC, and PBS networks, with 16 hours per week of captioned programming. The first consumer product containing the decoding feature, called "TeleCaption®(Trademark of the National Captioning Institute)," was sold by Sears, Roebuck and Co.

The original target audience for captioning included deaf or hard-of-hearing people (about 20 million). The market has since expanded to include people learning English as a second language and those learning to read, especially students with reading disabilities.

Since 1980, hundreds of thousands of decoders have been sold, and most are still in use. Consumer-style caption decoders with a variety of special features are now widely available to the public.

In 1991, the availability of captioned programming exceeded 400 hours per week as described below.

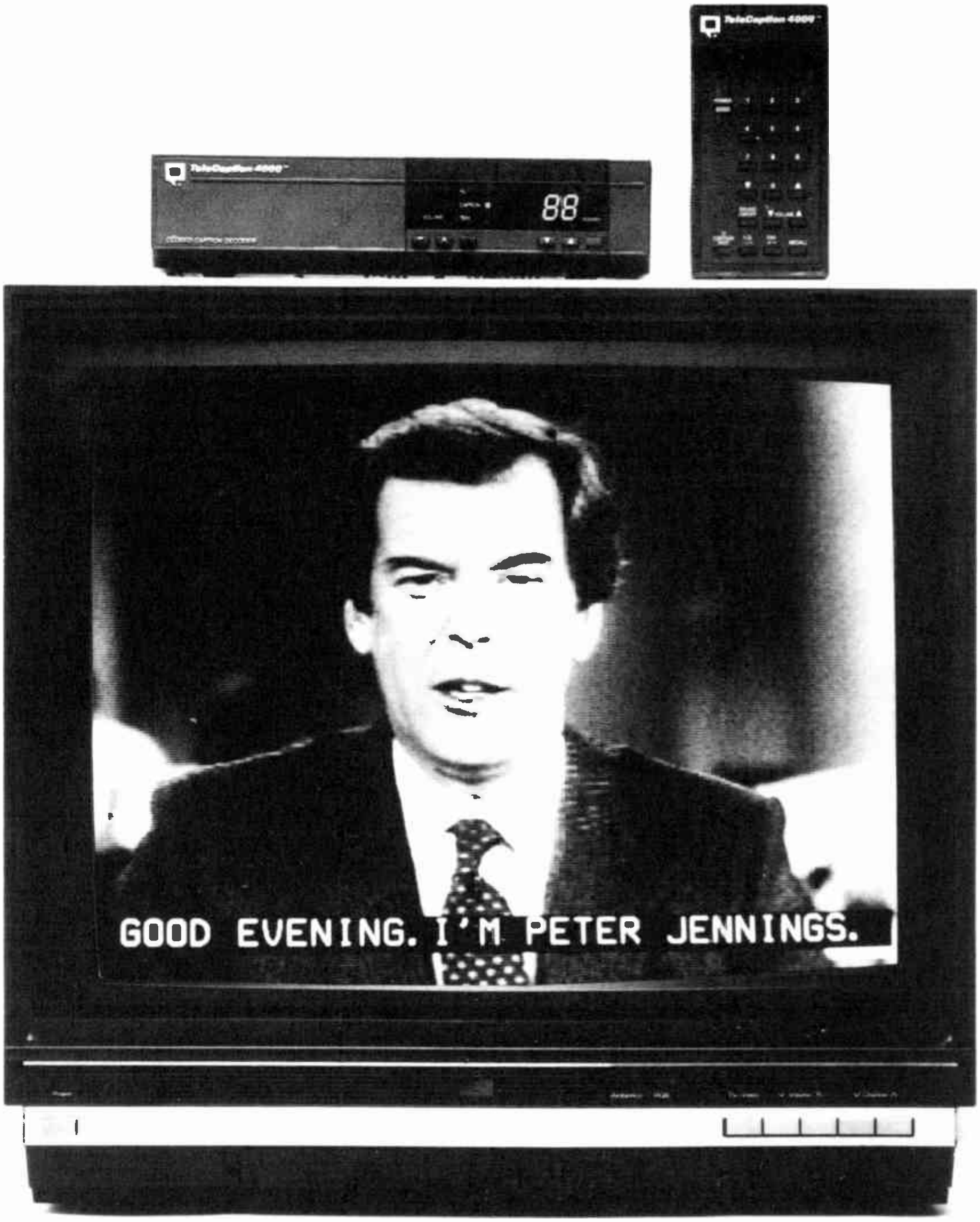


Figure 1. Digital closed caption data on video line 21 are processed by a decoder and displayed as text across the lower portion of the picture. (Courtesy of ABC News and the National Captioning Institute.)

- 215 hr/week of Network TV (ABC, CBS, NBC, FOX, and PBS)
- 170 hr/week of Cable TV (basic, subscription, pay per view)
- 85 hr/week of syndicated programming
- As of 1991, nearly 3,000 home video movies have been captioned
- News programming on more than 100 local TV stations: 80% use newsroom computers, 20% use real-time systems

FCC Standard

The line 21 captioning data signal is protected from interference from any other VBI service, test signal, or spillover from active video under FCC Rules and Regulations, Section 73.682(a)(22), adopted in 1976. These rules also established the transmission standards for captioning and list the uses of the data channel.

Excerpts from rules associated with the Television Decoder Circuitry Act of 1990 that amended Part 15 of the FCC Rules (Radio Frequency Devices, which relate to television receivers) are provided below.

Television Decoder Circuitry Act Of 1990

The United States Congress passed the Television Decoder Circuitry Act of 1990 (Pub. L. 101–431, 104 Stat. 960 (1990)). This act requires that, effective July 1, 1993, all television receivers with picture screens 13 inches or greater must be equipped to display closed-captioned television transmission.

In 1990 the Federal Communications Commission adopted Rules (Report & Order) to implement the provisions of the Act which take effect in 1993. As with all rules, these are subject to revision and should be checked against the most recent version.

Excerpts from FCC Report and Order

General Docket 91–1: "In the Matter of Amendment of Part 15 of the Commission's rules to implement the Provisions of the Television Decoder Circuitry Act of 1990." The Order was adopted April 12, 1991 and released to the public on April 15, 1991. The Order becomes effective July 1, 1993.

The Order amends Part 15 of the FCC's Rules. A new Section 15.119 is added to read as follows:

Section 15.119 Closed Caption Decoder Requirements for Television Receivers

(a) *Effective Date.* Effective July 1, 1993, all TV broadcast receivers with picture screens 13 inches or larger in diameter shipped in interstate commerce, manufactured, assembled, or imported from any foreign country into the United States shall comply with the provisions of this section.

(b) *Transmission Format.* Closed caption information is transmitted on line 21 of field 1 of the vertical blanking interval of television signals, in accordance with Section 73.682(a)(22) of this Chapter.

(c) *Operating Modes.* The television receiver will employ customer-selectable modes of *TV* and *Caption*. A third mode of operation, *Text*, may be included on an optional basis. The Caption and Text Modes may contain data in either of two operating channels, referred to in this document as C1 and C2. The television receiver must decode both C1 and C2 captioning, and must display the captioning for whichever channel the user selects. The TV Mode of operation allows the video to be viewed in its original form. The Caption and Text Modes define one or more areas (called "boxes") on the screen within which caption or text characters are displayed. . . .

(d) *Screen Format.* The display area for captioning and text shall fall within the SAFE TITLE AREA as defined by SMPTE Recommended Practice #27.3—1989. This display area will be further divided into 15 character rows of equal height and 32 columns of equal width, to provide accurate placement of text on the screen. Vertically, the display area begins on line 43 and is 195 lines high, ending on line 237 on an interlaced display. All captioning and text shall fall within these established columns and rows. The characters must be displayed clearly separated from the video over which they are placed. In addition, the user must have the capability to select a black background over which the captioned letters are displayed. . . .

(l) *Compatibility with Cable Security Systems.* Certain cable television security techniques, such as signal encryption and copy protection, can alter the television signal so that some methods of finding line 21 will not work. In particular, counting of lines or timing from the start of the vertical blanking interval may cause problems. Caption decoding circuitry must function properly when receiving signals from any cable security system that was designed and marketed prior to April 5, 1991. Further information concerning such systems is available from the National Cable Television Association, Inc., Washington, DC, and from the Electronic Industries Association, Washington, DC.

(m) *Labelling and Consumer Information Requirements.* The box or other package in which the individual television receiver is to be marketed shall carry a statement in a prominent location, visible to the buyer before purchase, which reads as follows:

This television receiver provides display of television closed captioning in accordance with Section 15.119 of the FCC Rules.

Receivers that do not support color attributes or text mode, as well as receivers that display only upper-case characters . . . must include with the statement, and in the owner's manual, language indicating that those features are not supported.

The complete text of the FCC Report and Order may be obtained from the U.S. Government Printing Office. The text was also published in the Federal Register Vol. 56, No. 114, p.27200. The FCC Rules & Regulations are contained in the Code of Federal Regulations, Part 47, Telecommunications. See Chapter 1.7, "Source Guide: Broadcast Standards and Information," for information on contacting the U.S. Government Printing Office and other organizations mentioned in this chapter.

DISPLAY FORMAT

The line 21 captioning system has two independent channels, C1 and C2, each of which contains a captioning subchannel and a text subchannel. Caption data for most programming are carried in channel one, leaving channel two available for other applications such as caption data for another language. The captioning channel can operate in either of two different modes, "pop-on" and "roll-up." Within each of these modes, paint-on can also be used.

The pop-on mode is used when the captions are prerecorded. Captions are loaded in advance of the time they are to appear and, on command, are displayed. While the caption in the first memory is displayed, a second memory may be loaded with the next caption for display upon the next command, and so on, alternately loading and displaying memories.

The roll-up mode is used for real-time captioning. From one to four rows at the bottom of the screen display the real-time captions with the text moving from bottom row upwards to the top-most of the 4 rows after which the text is erased. The roll-up mode distinguishes the captions as being real-time.

The paint-on mode is currently used within the text service. The text is displayed as it is received from left to right beginning at the top of the screen and working down to row 15 after which the text scrolls up.

Captions and text may have any of several attributes in addition to block monochrome characters. Attributes include upper and lower case, six different colors, italics (or slanted text), underline, and flash. Attributes are determined by the editor at the time of encoding the captions and text.

The caption display on the television screen consists of 15 rows with up to 32 characters per row as shown in Fig. 2. A switch in the decoder permits the viewer to select captions or text. In the captioning mode, a maximum of four rows is used onto which each caption pops on when prerecorded captions are received and rolls up when the captions are live. The four rows can appear anywhere on the 15 displayable rows which occupy most of the screen area.

The first row starts at line 43 in each field. Each row occupies 13 lines of a field scan or 26 lines of the 525 lines. When the decoder is in the text mode the screen is nearly covered by the 15 rows of text, which scroll upward.

Each row of characters is displayed within a black surround box to enhance the readability against the normal video background. The box extends one character position to the left of the initial character in each row and one character to the right of each row. Partial or segmented rows also conform to this black box surround.

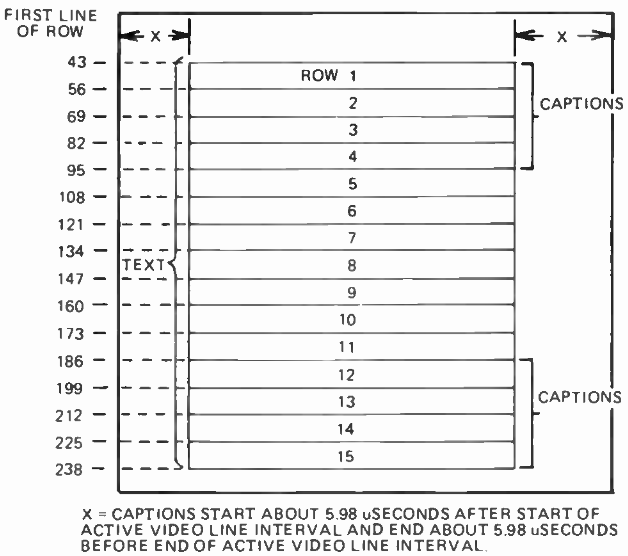


Figure 2. Caption row positions.

CAPTION DATA ENCODERS

Encoding is the process of inserting the caption data into the vertical blanking interval (VBI) of the television signal. The data contains the caption text in addition to positional instructions and display attributes (such as color and italics). The encoder is placed in the video path of the program to be captioned. There are two versions of the line 21 VBI encoder.

Smart Encoder

The smart encoder is used to insert caption data into line 21 of fields 1 or 2 of the vertical blanking interval. The smart encoder receives data through an RS-232 serial data port or via telephone line from an internal modem. Through the use of the smart encoder, locally produced caption and text services may be added to an already closed-captioned video program or if noncaption text is contained on line 21. If captions are present on the incoming program, noncaption text may be interleaved into the gaps between the captions by the smart encoder.

Smart encoders are used in the process of creating real-time captions, live-display captions, and captions from newsroom computers (prompters).

Simple Encoder

The simple encoder generates the line 21 data signal, to be inserted on a video signal, from caption data received at the RS-232 serial input. This encoder is used mainly for off-line or prerecorded captions. The simple encoder cannot add captions to a video signal already containing line 21 data.

To create the line 21 caption signal, after the editor has produced the captions and they are stored in the computer, the videotape with program video and timecode information is fed through the simple encoder. The time-code information is used by the encoder to trigger the transfer of data (captions) from the captioning computer. The special software in the computer processes the data and matches each caption to the appropriate time-code. A microprocessor in the encoder processes the RS-232 data and encodes two characters of caption data onto each field 1 of line 21 of the program video. The output of the encoder may then be recorded on a second videotape recorder for playback at a later time.

METHODS OF CREATING CAPTION DATA

Several techniques are employed to create caption data based on whether the broadcast is live or prerecorded. There are also variations within the live and prerecorded captioning technologies.

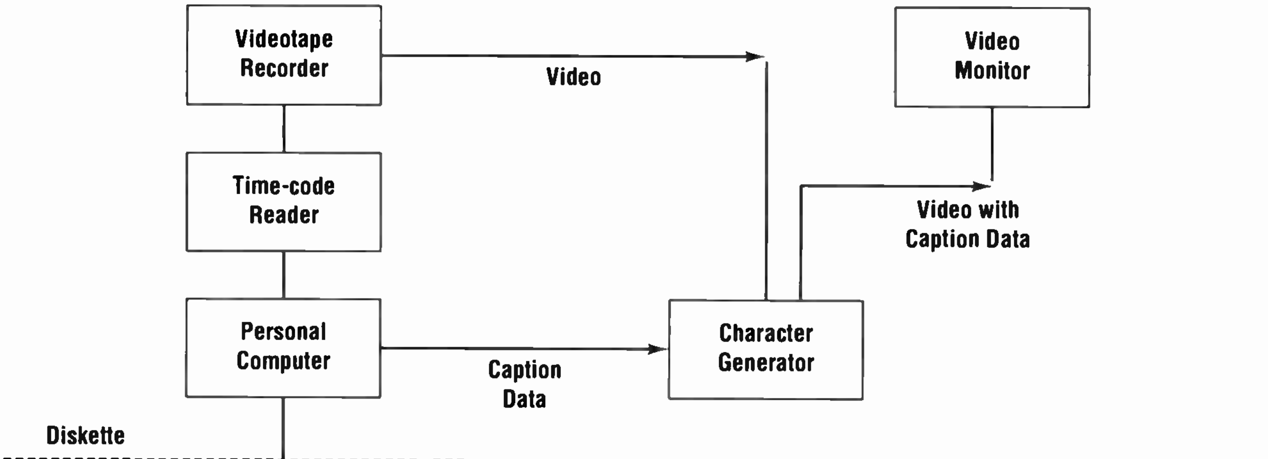
Prerecorded Captioning

Prerecorded captioning (off-line or nonlive captioning) involves the preparation of closed captions for programs that have been recorded prior to their telecast. The captions are also created in advance, and a closed captioned videotape version of the program is made. There are several steps in this process.

A captioning facility receives a time-coded videotape copy of a program master. A caption editor reviews each scene, listens to the spoken words, and types the captions using a standard keyboard on a personal computer (PC) (see Fig. 3). Several quality control checks are performed, including spelling, grammar,

Figure 3. Captioning prerecorded materials is done by an editor working at a standard computer keyboard. (Courtesy of National Captioning Institute.)

Figure 4. Off-line captioning system. Caption creation.



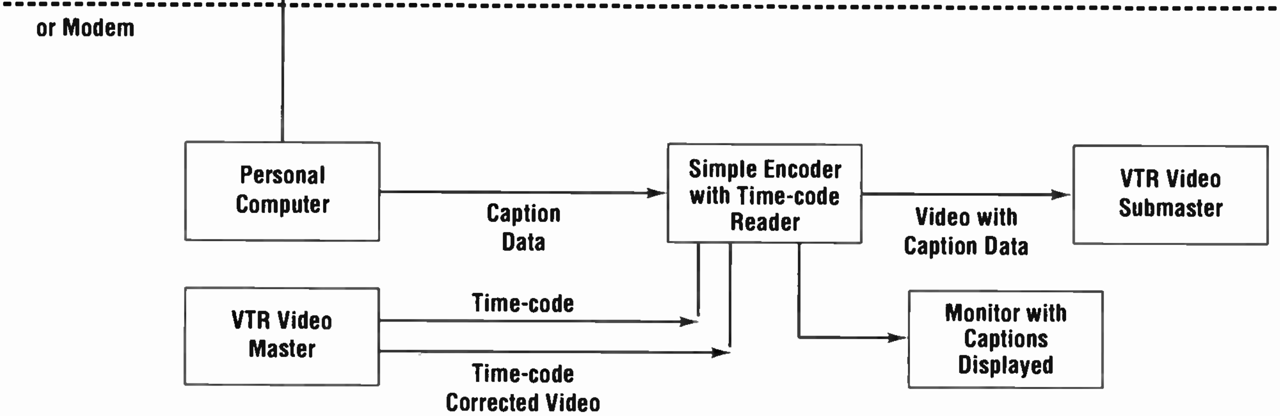


Figure 5. Off-line captioning system. Caption encoding.

syntax, timing, and screen positioning of the captions. The captions are then stored with their corresponding time-codes on a diskette. Fig. 4 shows a block diagram of a captioning facility.

The recorded program and data from the diskette are then merged, using the time-codes to trigger the captions, to produce a closed captioned sub-master video tape recording, with the captions encoded on line 21 (see Fig. 5). The encoded version of the program is returned to the producer and is used when the program is aired.

Captions which are produced in advance are displayed on a caption-equipped receiver as pop-on captions, each one complete and timed to coincide with the spoken dialogue on the screen.

The equipment necessary for creating prerecorded or off-line captions includes:

- PC with caption creation software
- Video monitor
- Audio monitor
- Character generator
- Time-code reader
- Videotape player
- Simple caption encoder
- Videotape recorder

Using this equipment, the caption editor can view the program, listen to the audio, enter the captions along with the appropriate timing and placement information, and review the program as it would appear through a decoder. It takes approximately 20–25 staff hours to caption one hour of television programming.

Live Captioning

Live captioning involves the addition of caption data to the television signal at the time of a live transmission or broadcast. Examples of live captioned programming include news programs, sporting events, news conferences, and special bulletins or reports. There are three versions of live captioning.

Real-Time Captions

For real-time captions, a specially trained caption editor listens to the audio of a live television program and, using a stenotype machine similar to the ones used by court reporters, keys in words using a special shorthand code corresponding to syllables or phonetic

Figure 6. Real-time captioning is done by an operator using a stenotype keyboard similar to those used by court reporters. (Courtesy of National Captioning Institute.)

codes (see Fig. 6). Instead of keying in individual letters, groups of keys can be pressed down simultaneously. Each group of keys, a stroke, produces different phonetics. Using this kind of machine saves time, since a word such as "institute" may be keyed-in using just three strokes (one for each syllable) rather than nine (one for each letter).

With a stenotype machine, the real-time caption editor can key-in up to 260 words per minute. The phonetic codes are translated into English words by the captioning computer which has been programmed with the phonetic codes and caption editor dictionary. (see Fig. 7). From the computer, the words and caption control codes are sent via a data circuit to a caption encoder where they are encoded into the television signal on line 21 as real-time captions (see Fig. 8). A television receiver with a decoder then displays these words as part of a continuous multi-line (1 to 4 lines) scroll (caption) at the lower portion of the television screen. It takes up to three seconds from the time a word is spoken to the time it appears on the screen as a caption.

Newscasters often speak faster than the captions can appear on the TV screen and may occur faster than the viewer can read. To keep up with what is being said, and to control the display rate, the real-time caption editor must, at times, edit words or delete phrases. The real-time caption editor must also realize that the computer dictionary may not contain a particular spoken word, and substitute a different word, one the computer will understand and at the same time convey the meaning to the caption viewer. Alternatively, if the computer does not have the word, the captions are transmitted phonetically spelled out.

The equipment needed for creating and monitoring real-time captions includes:

- Special stenotype machine
- Personal computer with translation software
- Video monitor
- Audio monitor
- Smart caption encoder
- Live video program signal on which line 21 will be encoded with the captioning signal

Figure 7. Live captioning system. Caption creation.

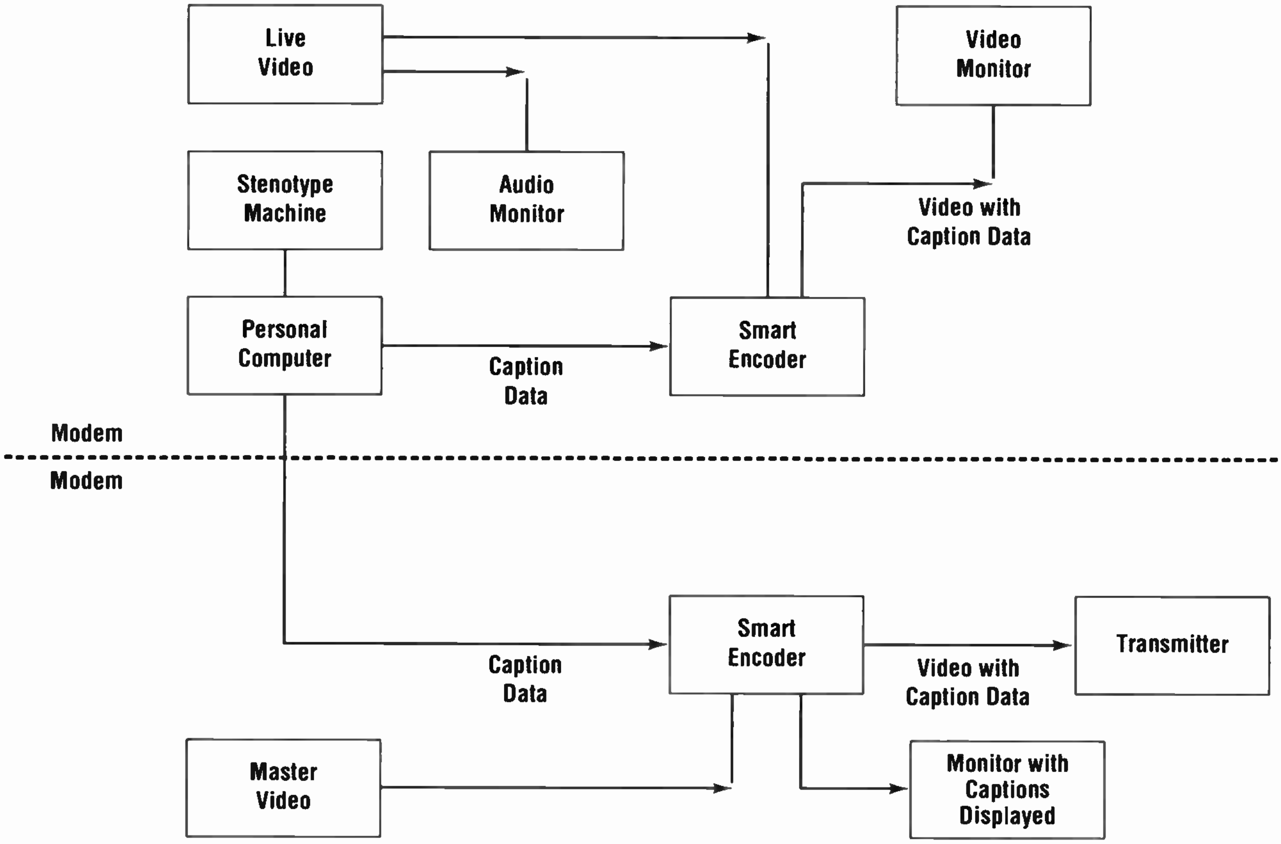


Figure 8. Live captioning system. Caption encoding.

The real-time captioning equipment need not be located where the live program is taking place. It is only necessary for the caption editor to hear and see the live program material (even over-the-air) and, using the stenotype machine, enter the captions into the captioning computer. The computer transmits the data through a modem (a telephone or data line) to the caption encoder which must be placed in the program line of the live video signal to be captioned. Thus, any live program may be captioned as long as the caption editor has access to the live program audio signal and a caption encoder is in the live program circuit. Of course, it is desirable to be able to see the results of the captioning process and the same over-the-air signal can be used for the monitor.

For real-time captioning of programs which may not be broadcast at the time of captioning or for which an over-the-air signal is not available, a microwave feed of the encoded signal may be used as the program monitor.

Because of the investment required for a captioning facility and the high level of skill needed by the real-time caption editor, stations and production facilities may find it more practical to employ the services of a captioning agency where captioning equipment and personnel are more readily available and the cost is based on an hourly basis.

Live-Display Captions

Live-display captioning is used when an accurate script is available in advance of a televised live event such as a speech or newscast. The scripted words are converted to captions by an editor and PC and stored on a computer disk. When the live event is televised, the editor manually calls (triggers) each caption from the disk which is then processed by the computer and transmits the data to a caption encoder. The editor transmits a display command when each caption is to be displayed. With live-display captioning, the words are timed to appear on the television receiver as they are being spoken. If last-minute changes are made in the script it may be necessary to switch to real-time captioning.

Newsroom Computer Captions

The third method of live captioning is the generation of captions through the use of newsroom computer equipment. Many television news rooms convert news stories from their word processors into data for use on a prompting machine for the news reporter to read

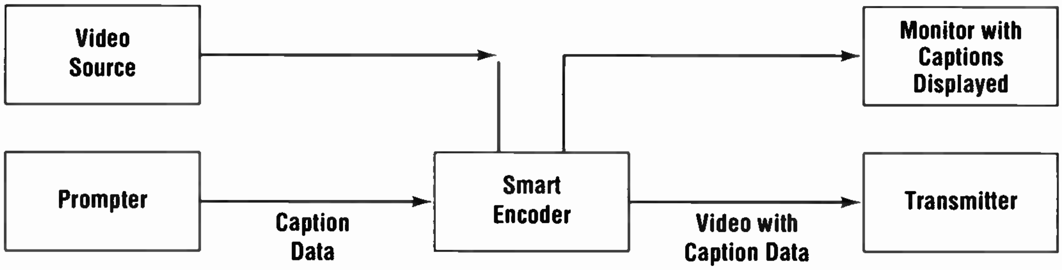


Figure 9. Captioning in an electronic newsroom.

when on-the-air. At the same time the reporter reads the script from the prompter, the computer controlling the prompter passes the data through a serial interface to a caption encoder, through which the live program signal passes, that inserts the captions into line 21 of the video signal (see Fig. 9). The result is the scroll or roll-up display on the viewer's screen.

There are some dangers in this process. It is not as automatic as might be assumed. Some newsrooms add cues to the prompting text designed for the on-air personnel, some of which might be inappropriate to appear as captions. Program breaks or changes of scene may require changes in the prompting operation that could interrupt the captioning process. Also, it is desirable for the operators and program producers to have an off-air monitor to insure the integrity of the captioning process. Finally, it may be necessary for an operator to signal the captioning encoder to add captions at the beginning of the local program and to revert to the pass-through mode at the end.

The equipment needed for electronic newsroom captioning includes:

- Prompting system with data output
- Captioning software for the prompting system
- Smart caption encoder

Automatic Live Encoding

When production schedules are tight, an alternate means of producing a captioned submaster is available. Automatic live caption encoding makes use of the same caption creation techniques used in pre-recorded captioning, but a different method is used to trigger the data into line 21 of the television signal. Fig. 10 shows a block diagram of the system. In this case, the captioned data are loaded into the PC and the internal clock within the PC is used to trigger the captions as opposed to using time-code from the program video tape. A manual trigger is used to start the transmission of data between the PC and the smart encoder.

The display of automatic live caption encoding on the television receiver is the pop-up display mode, the same as used for prerecorded captions.

Text Service

Information transmitted as part of the text service is usually independent of the program with which it is transmitted. Captions do not use all the available data capacity of the line 21 captioning system, and the additional capacity may be used to transmit nonprogram-related information to the television viewer. Such text may consist of program notes, news, weather, sports, farm and financial reports, and other information. Unlike the captions, the text service appears as full-screen rolling text (without pictures) on the television screen.

The text service is accessed in the decoder by a switch that selects either captioning or text. There is a text channel for each of the two caption channels (C1 and C2).

The data rate of the text service is quite low compared to the data rates of various teletext services that have been proposed for use in the United States and is in widespread use in Europe. The text may pause occasionally when the higher priority captions are being transmitted.

The equipment necessary for creation of a line 21 text service includes a personal computer with appropriate text generation software, an information

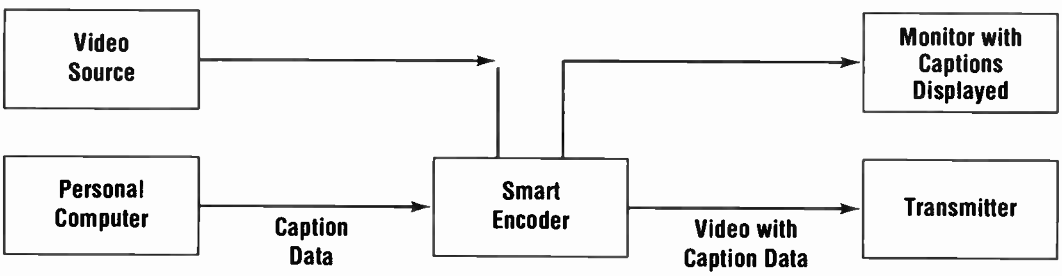


Figure 10. Block diagram of an automatic live encoding system.

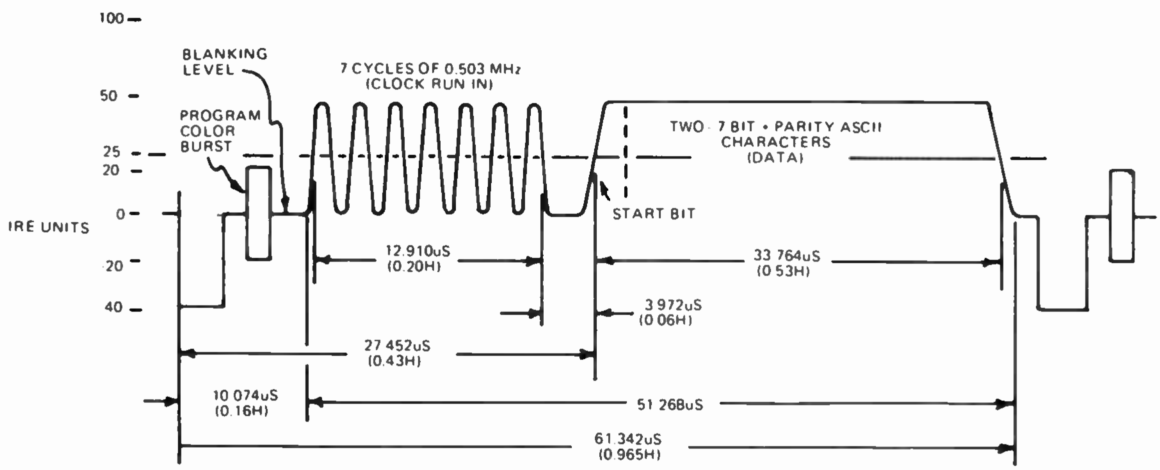


Figure 11. Line 21 field 1 data signal format.

source, and a smart encoder along with appropriate monitors. The text service can, under certain conditions, operate completely unattended if desired.

DATA TRANSMISSION FORMAT

Captions associated with a television program are transmitted as an encoded composite data signal during line 21 on field 1 of the standard NTSC video signal shown in Fig. 11. The signal consists of a clock run-in signal, a start bit, and 16 bits of data corresponding to two bytes of 8 bits each (7 bit ASCII code plus one parity bit). Therefore, transmission of actual data is 16 bits every 1/30 of a second or 480 bits per second (bps). The data stream also contains encoded information (control codes) which provides the instructions for formatting and changing the attributes of the characters to be displayed.

The clock run-in consists of a seven-cycle sinusoidal burst which is frequency and phase-locked to the caption data clock. The frequency of 32 f_H (0.503496 MHz = 32 × 15,734.26 Hz), which is twice that of the data clock, provides synchronization for the decoder clock. The clock run-in signal is followed by the equivalent of two data bits at logical zero level, then a logical one start bit. The last two cycles of the clock run-in, the two logical zero bits, and the logical one start bit constitute an eight-bit frame code signifying the start of data as shown in Fig. 12.

The seven bit ASCII transmitted data are coded in a nonreturn-to-zero (NRZ) format. An eighth bit is added to each character to provide odd parity for error detection.

The sequence of identification, control, and character code transmission is shown in Figs. 13 and 14. Each caption transmission is preceded by a preamble code, which consists of a nonprinting character followed by a printing character to form a row address and display color code. Both characters of all control codes are transmitted within the same field of line 21 and twice in succession to ensure correct reception of control information. A transmitted caption may be interrupted by a mid-caption control code between two words in order to change display attributes such as color or italics. At the completion of a caption transmission, an end-of-caption control code is sent.

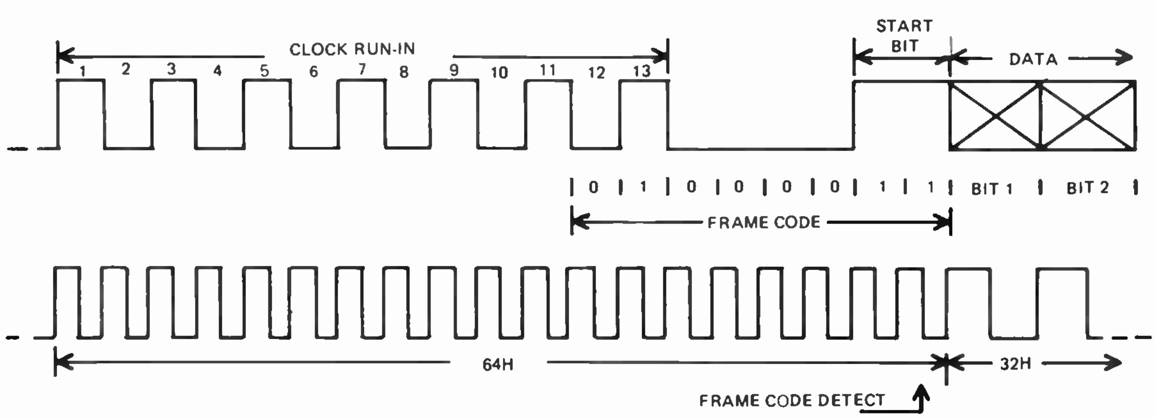


Figure 12. Line 21 data.

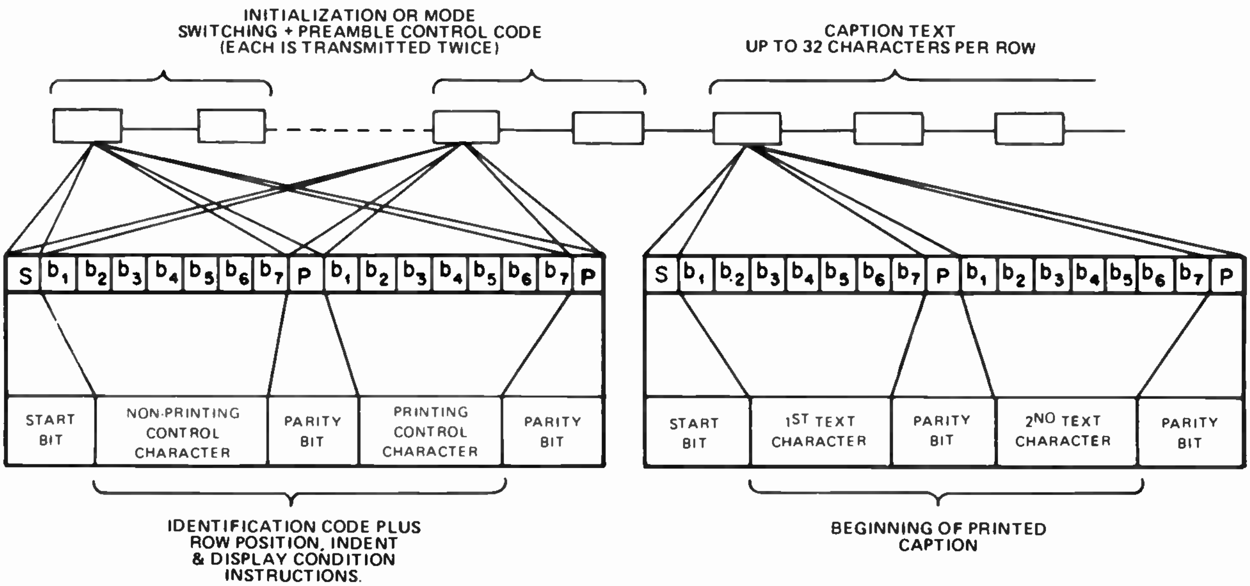


Figure 13. Caption row preamble format.

The first character of a control code is a nonprinting ASCII character (0000000 through 0011111). Codes 0000000 through 0010000 are not used. This is followed by a printing character (0100000 through 1111110). All characters that are received after a set of valid control codes are interpreted and loaded into the decoder memory as printing characters. Character codes with bad parity result in an all-ones code being written into memory; this causes display of a white box (the delete symbol) in place of the desired character which, of course, was in error.

The data rate is 480 bits/sec or 60 bytes per second (8 bits per byte including parity). At an average of 5 bytes or letters per word, a maximum word rate of 12 words per second or 720 words per minute could be achieved. In practice, the word rate is somewhat less than this, about 600 per minute, due to the time required for transmission of control codes. Because most speech is much slower than this, there is adequate time for a second captioning channel and a text service.

LINE 21 TECHNICAL ADVISORY

Special precautions must be undertaken to ensure the continuous and correct passage of the line 21 signal throughout the facility. Unless a station is actively involved in producing captioned program materials, the line 21 signals may not normally be monitored at

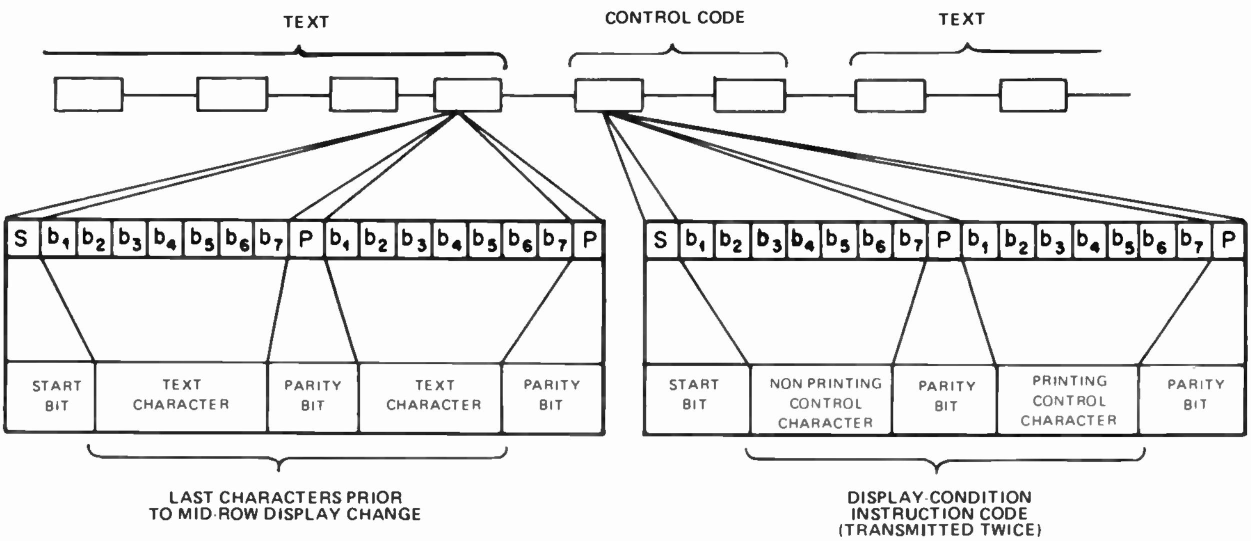


Figure 14. Mid-caption display. Condition change format.

a television facility to the same extent as the video and audio signals. With the passage of the Decoder Circuitry Act, millions of caption receivers will be in use within a few years. Therefore, it is prudent for stations to install captioning monitoring facilities and to routinely check the path of the line 21 signal.

The captioning data signal is contained in the NTSC television signal vertical blanking interval on line 21, field 1 (field 2 is not normally used for captioning but could be used for text and signalling data). Unless the caption data signal is transmitted intact and on line 21, field 1, a viewer's decoding circuitry will not operate properly.

The captioning signal waveform is detailed in Section 73.682(a)(22) and Fig. 17 of Section 73.699 Subpart E of the FCC Rules and Regulations and conforms to the standard television synchronizing waveform for color transmission given in Fig. 6 of Section 73.699.

The captioning data signal is a part of the television program, compared to vertical interval test signals which are not. Regular observation and monitoring is essential in order to verify the presence and proper location of the data signal. Equipment throughout a video facility should be routinely checked and, if necessary, adjusted to pass the line 21, caption data signal.

Maintenance personnel as well as control room operators should be made aware of the importance of the line 21 signal and adjust all applicable equipment accordingly.

The following video processing equipment may adversely affect the line 21 signal:

Video Processing Amplifiers. Certain types may delete the line 21 signal, move it to another line, exchange the fields or partially blank the signal.

Time Base Correctors (TBC). Some may blank the line 21 signal or advance or delay line 21 by a field or a line.

Video Tape Machines. Certain types contain a TBC which may blank or move line 21.

Video Switchers or Digital Store Devices. May blank or move the line 21 signal.

Frame Store or Frame Synchronizers. May blank or distort the line 21 signal.

VITS Generators or Inserters. Should be checked to determine if they are programmed to pass the line 21 signal correctly.

Digital Effects Generators. May delete or distort the line 21 data signal.

A waveform monitor with line select capability is required to view the line 21 caption data to determine that it is on the correct line and field. For example, a Tektronix 1480 series; VM 700A, 1700 Series, 1780R, or Hewlett-Packard equivalent with line selector will easily locate and display the signal.

The line 21, field 1 data signal may be partially, but inadvertently, blanked. This can occur when the active video line is advanced or delayed relative to horizontal sync. The decoding system is more tolerant of this problem when the data signal is early and causes part of the clock run-in sinusoidal burst to be blanked. On the other hand, if the data signal is delayed relative to horizontal sync, loss of the last data bit (parity bit) may occur producing decoding errors which in turn will cause the characters to be displayed as white boxes. In an extreme case, too many errors will cause the decoder to disable all closed captioning functions.

Some video processing amplifiers have sensitive vertical phase blanking and horizontal blanking controls. These may also adversely affect the line 21 signal.

Video processing amplifiers are frequently used in the following configurations:

- Output of a satellite receiver
- Input or output of an intercity relay link (ICR)
- Input or output of a studio-to-transmitter link(STL)
- Input to a transmitter
- Input to a videotape machine
- Output of a studio or master control switcher
- Remote truck video processor
- Utility proc amp (equipment racks)
- Maintenance shop test equipment

All signal paths, backup transmitters and their VITS generators and processing amplifiers, and tape-delayed broadcast installations should also be routinely checked to insure passage of the line 21 signal.

It is important to note that if the overall signal level is allowed to drop substantially below the normal 50 IRE level, captions could become garbled.

Special Cautionary Notes

Time Compression: The Lexicon® time compression process is incompatible with the line 21 closed captioning data signal. During the Lexicon compression, the video playback heads may not read every field and the line 21 data is not transmitted in its entirety, resulting in garbled or missing captions.

Locally Originated Captioning: Programs and commercials prior to, during, and after locally originated captioned programs may also be closed captioned. Therefore, the local smart encoder must be disabled, preventing pass-through, until the completion of a network telecast. The encoder must be enabled immediately following the locally captioned program to avoid stripping line 21 data from network programs which follow and pass through the caption data.

Newsroom Computers: Newsroom computer captioning typically uses the smart encoder in a newswire/real-time mode. When this mode is invoked, incoming line 21 data are automatically turned off, preventing any previously recorded caption information from being passed through and displayed. The newswire/real-time mode must be properly exited at the conclusion of local live captioning by issuing an "end of message" command to restore or enable the line 21 data.

Other systems may use the smart encoder pass-through mode rather than the newswire/real-time mode. Pass-through offers more complete control of encoder and decoder functions. In this mode, incoming line 21 data can be specifically turned off and must be specifically turned back on (enabled) to allow line 21 caption information passing through the smart encoder to be displayed. The encoder must be specifically issued an enable command to turn on line 21.

A smart encoder reset command will always return the encoder to the default enabled condition but will delete any user-stored articles, such as text channel information.

Network Origination/Local Origination: If a local station uses the master control routing switcher for signal distribution during a program originating from the network and then switches to local origination and a studio controller prior to the end of the network feed, the line 21 signal should be checked to insure its integrity.

Using Captioned Excerpts: Excerpts of programs, received or recorded off-air or from external satellite or network feeds, for use in other programming, may contain line 21 closed captions. In these cases, the captions should be stripped if the excerpted material audio will be edited or replaced because the captions may no longer be germane. A time base corrector or video processing amplifier can be used to remove or move the caption data signal from line 21.

Broadcast Decoder: In addition to consumer models, a professional broadcast type decoder is available for stations to use to monitor the line 21 data signal on the network, video tape machines, or the broadcast signal. The broadcast model accepts a standard video signal with captioning data on line 21 and produces a video output with the captioning data decoded and, using an internal character generator, displays the data as open captions on a standard video monitor.

Many stations have also installed Telecommunication Devices for the Deaf (TDD) in order to communicate with the deaf and hard-of-hearing television viewers.

Emergency Messages: Broadcasters which elect to transmit emergency messages are required to provide a readable version of the aural emergency message in order to inform hearing impaired viewers of impending weather or other emergencies (see Section 73.1250(h) of the FCC Rules). One means of doing this is to insert a moving line of words (crawl) with the message across the lower portion of the screen. If a program, over which the message is superimposed, is captioned the lower portion of the viewer's screen may be covered with the captions and the emergency message may become obscured or unreadable. Therefore, procedures should be established to stop or interrupt the captions or text transmission during the emergency message or to incorporate the message into the captions or text material.

Passage Through Other Distribution Systems: The line 21 captioning signal was designed to be robust and able to withstand a significant amount of degradation before errors appear in the text. The low data rate and placement of the signal at the very beginning of the active portion video of the signal contributes to this robustness. As a direct result, the captioning data signal will easily pass through:

- Master antenna (MATV) systems
- Community antenna (CATV) systems
- Satellite circuits and receivers
- Translators, multipoint distribution systems (MDS) and multichannel multipoint distribution systems (MMDS)
- Set-top converters (RF and baseband)
- Consumer videocassette recorders (VCR)

Most set-top receiving converters have outputs on channel 2, 3, or 4. A stand-alone captioning decoder can be inserted between virtually any converter and any television receiver with little degradation to picture and sound quality.

SUMMARY

Closed captioning is one of a series of signals associated with television transmission in addition to video, audio, stereo, separate audio program (SAP) channel, and vertical interval test and data. Broadcasters traditionally carefully monitor their own audio and video signals. With the addition of the new program-related signals, additional monitoring facilities and procedures must be established. The technology and practice for generating closed captions has evolved to the point where it is possible for individual television stations and production facilities to utilize either a captioning service or their own system for captioning programming.

The National Captioning Institute, Inc. provided most of the material contained in this chapter. Further information may be obtained by calling 1-800-533-9673 (Voice) or 1-800-321-8337 (TDD).

7.6 Weather Radar

James Block
Kavouras, Inc., Minneapolis, Minnesota

DEVELOPMENT OF WEATHER RADAR

Radar was developed during the 1930s as a method of detecting and tracking air and sea vessels. World War II saw rapid development of radar to meet military needs. Although the use of radar for detection of precipitation was not originally intended, it soon became apparent that radar could be used in meteorology. After the war ended, meteorologists began using surplus radars for their work. During the late 1950s the U.S. Weather Bureau, now called the National Weather Service (NWS), began construction of a network of weather radars around the country for surveillance of precipitation and severe storms. Weather radar technology has evolved to the point where quantitative measurements can be determined from its digital image format. Sophisticated radars are operated by meteorological services throughout the world.

OPERATIONAL WEATHER RADAR THEORY

Radar is an echo sounding system. The range of a target is determined by measuring the time for a pulse of transmitted electromagnetic energy to travel to the target and back again after reflection. A highly directive antenna is used to send pulses of radio frequency (RF) energy. A slightly diverging beam is often created in order to sample a significant volume of the atmosphere. Most of the energy is focused along the axis perpendicular to the face of the antenna. As with all directional antennas, some energy is radiated as side lobes. Side lobes typically account for 5% of the total radiated power, but result in the production of most of the ground clutter pattern. Ground clutter may make it impossible to monitor precipitation within several miles of the radar. Fig. 1 shows the relation of side lobes to the main beam.

The same antenna that is used for transmission is also used for reception. The signal is sent in pulses, and during the time between the transmitted pulses the antenna is connected to the radar receiver. Electromagnetic waves scatter upon striking a target and only a small portion of the original energy returns to the radar antenna. The amount of energy received depends upon the target's size and composition. All water particles within the radar beam return a minute fraction of the energy, but the combined effect of many water or ice particles allows a sufficiently strong signal to be returned.

The time interval between pulses must be such that signals from desired targets within the radar beam have time to be detected at the receiver before the next pulse is transmitted. The frequency at which the pulses

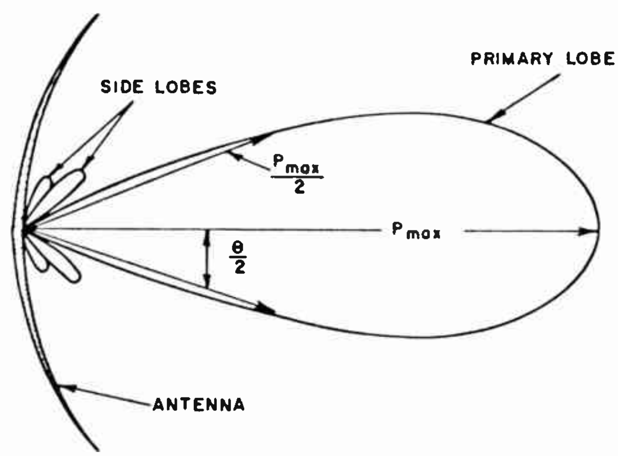


Figure 1. Cross-section of a radar beam from a paraboloid reflector. Note that the angles as drawn are much larger than would be the case in reality and that only the first and second side lobes are shown. (Courtesy U.S. Weather Radar Manual, 1967)

are sent is called the pulse repetition frequency (PRF). The PRF is one factor which determines maximum range. The higher the PRF, the shorter the maximum radar range will be. High PRFs allow for more target sampling which provides for a more accurate definition of the target.

The signal received by the radar is directly proportional to the size of the antenna. For meteorological purposes, the larger the antenna, the better. Limitations on size are imposed by engineering and economic considerations.

Range Effects

The power received by the radar is inversely proportional to the square of the target's range. This is referred to as range attenuation. Two targets of equal intensity will appear to have different strengths if they are at different ranges. Modern weather radars have means for compensating for this effect. Swept gain, range normalization, or sensitivity time control (STC) are three features that can make equal intensity targets appear of the same strength.

Due to the earth's curvature, the radar beam rises above the earth's surface under normal propagation conditions. This results in the beam overshooting the tops of distant precipitation. This phenomenon imposes another limit on how far away precipitation can be detected.

Fig. 2 shows how the beam curves with respect to the earth under normal conditions. A temperature inversion will deflect the beam downward and cause abnormally large ground clutter patterns to occur, making precipitation detection difficult.

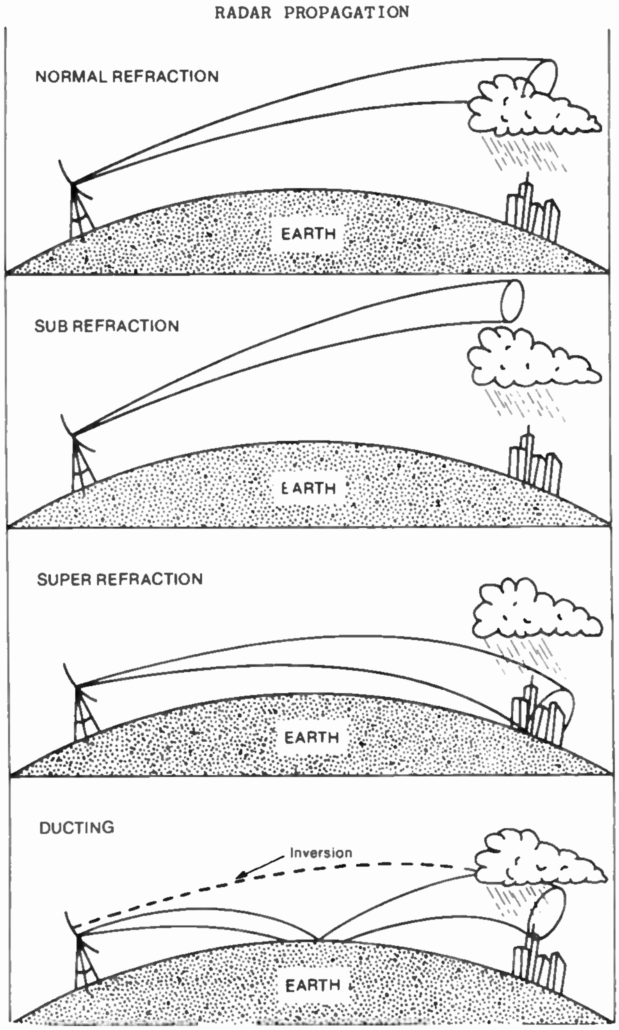


Figure 2. Refraction.

SPACE LOSS

The natural loss of signal power between a transmitter and a receiver has three components: free-space attenuation, absorption due to the molecular atmosphere, and fading due mostly to rain. All of these power loss phenomena are frequency dependent. Free space attenuation increases as the square of the frequency. Attenuation due to the atmosphere is generally negligible, but rain attenuation can be a factor above 1000 MHz, and increases with the frequency.

X-band (10 GHz, 3 cm) weather radars have become commonplace because of their low cost and small size. Because of the serious space loss that occurs at X-band, these radars are seldom used for quantitative measurement. The most effective radar band for use where rainfall rates are moderate to high is S-band (3 GHz, 10 cm). The disadvantages of S-band radars are their size, weight, and resulting high cost of purchase and installation.

C-band weather radars (6 GHz, 5 cm) offer a good compromise. Engineering problems and costs are smaller than with S-band radars due to the smaller antenna, while space loss problems are not as serious as with X-band radars.

DOPPLER RADAR

Conventional weather radars are noncoherent radars: that is, on a pulse-to-pulse basis, the transmitted frequency is not completely stable. Such radars are able to detect changes in precipitation intensity and location, and to map the size and relative motion of the echoes.

In order to measure the velocity of raindrops within the echo volume, it is necessary to utilize a radar with a precise transmitter frequency and a receiver system sensitive to small changes in return signal frequency caused by the movement of the precipitation. This type of radar is known as a coherent radar, or more commonly as a Doppler radar. Higher power radars (25 kilowatts to 1 megawatt) are known as pseudocoherent systems, where a modern coaxial magnetron is the transmitting device. Here, the receiver is phase and frequency locked to the transmitted pulse, and pulse-to-pulse echo phase shift is employed to measure wind speed. A Doppler radar can measure the wind velocity component along the radial axis of the radar beam, but not the component perpendicular to the beam.

An important limitation to the capabilities of Doppler radar is determined by the wavelength used and the PRF at which the radar is operating. Doppler radar measures the wind velocity by comparing the frequency of the returned signal with that which was transmitted. The radial velocity (V) is expressed as:

$$V = \frac{fl}{2} \qquad \text{(Eq. 1)}$$

where: V is the radial wind velocity
f is the frequency change
l is the radar wavelength

As Table 1 shows, short wavelength radars are most useful for measuring wind velocity because their frequency shifts are larger.

TABLE 1

Doppler shift frequencies in Hertz for various radar wavelengths and target velocities.

V (M/sec)	Wavelength (cm) 1.8 K-Band	3.2 X-Band	5.5 C-Band	10.0 S-Band
0.1	11	6	4	2
1.0	111	62	36	20
10.0	1,111	625	364	200
100.0	11,111	6,250	3,636	2,000

Also note from the above table that in order to measure small velocities, small frequency shifts need to be accurately measured. Thus it is important that the radar signal be very stable.

In order to measure radial wind velocity with a pulsed radar, the returned pulse must be compared with a stored version of the transmitted pulse. The difference in frequency between the two pulses can be used to determine the radial wind velocity. This method decreases in reliability as the frequency difference diminishes. If the difference in phase between the two pulses is measured, better results are obtained for small velocity values because the phase changes are more pronounced than the frequency changes. Magnetron-based coherent-on-receive radars avoid this problem by only measuring phase shift, thereby maintaining constant accuracy regardless of wind speed.

There are two limitations that using a pulsed radar imposes. The first is the unambiguous maximum range as shown below.

$$Rmax = c/2PRF \qquad \text{(Eq. 2)}$$

where: Rmax is the maximum unambiguous range of the radar
c is the speed of light
PRF is the pulse repetition frequency

The maximum radar range decreases as the PRF increases. This decreased distance results from the shorter "listening time" between pulses.

The other limitation is the highest velocity that can be measured without ambiguity.

$$Vmax = lPRF/4 \qquad \text{(Eq. 3)}$$

where: Vmax is the maximum unambiguous velocity
l is the radar wavelength
PRF is the pulse repetition frequency

Increasing the PRF allows the radar to measure higher velocities. From eqs. 2 and 3 it can be seen that choosing a high PRF allows for a greater Vmax but a lower Rmax. This interdependence of range and velocity is known as the "Doppler Dilemma", because a compromise must be made between long range and high velocity measurement. Since Vmax is directly related to the wavelength of the radar, short wavelength radars are limited in their ability to measure high velocities.

Each sampling volume within the radar beam returns reflected energy from millions of particles with many different speeds and directions. The final radial velocity which is displayed by the radar set is the average of the velocities of the largest drops within the radar beam volume. The range of velocity extremes measured over a specified number of pulses is called the variance or spectrum width. This parameter is a measure of the turbulence or shear within the beam volume. Higher spectrum width indicates greater turbulence.

Doppler radars are becoming more common in the operation of weather offices as the displays become less complicated and more user friendly.

TYPICAL RADAR APPLICATIONS

The selection of a weather radar set involves a tradeoff between range, performance, features, and cost. A typical conventional S-band radar similar to the NWS WSR-74 series offers a nice balance between the operational meteorological limitations, but costs in excess of $250,000. Adding Doppler processing to an existing conventional radar can add $150,000 to the original cost. There are several manufacturers of weather radars who offer more affordable conventional and Doppler radars.

An example of a low power, turbulence-only system is the Doppler radar manufactured by Rockwell Collins. This radar was originally designed for use in aircraft, and later the pedestal was modified for service as a ground-based radar. This system is a C-band Doppler radar, with a beamwidth of five degrees, and a peak power of 200 watts. This combination of wavelength, low peak power, and wide beamwidth has allowed for a relatively small flat plate antenna and a reasonable cost. This system does not actually process and display raw velocity information, but rather just looks at the spectrum width of the returned echoes, and determines, according to user-defined thresholds, whether the return is turbulent or not. This turbulent return is painted on the display as a single additional color overlaying the reflectivity. Many of these radars are now in use in the broadcast industry.

An example of a high power system for direct measurement of velocity and direction is that manufactured by Enterprise Electronics (EEC), operating at 250,000 watts in the C-band. Broadcast meteorologists have used this system with a 12-foot parabolic antenna for detection and analysis of mesocyclones. Wind speeds can be displayed in different colors for four ranges each, of negative and positive velocities and one of precipitation with no wind. Six other colors are assigned to precipitation intensity.

Any radar, regardless of type, radiates RF energy, and therefore must be licensed by the FCC. The radar vendor should review radar site selections and can advise on FCC procedures. See Chapter 1.1, "FCC Organization and Administrative Process" in this *Handbook*.

NEXRAD

The NWS is deploying a new generation of Doppler weather radars, called NEXRAD, to replace the current network. NEXRAD is a 500 kw S-band Doppler radar being built by Unisys for the NWS. The NWS has designated this radar the WSR-88D, and it is designed to replace all the current WSR-57 and WSR-74 radars by 1996. NEXRAD will run in one of two automatic volume scanning modes, completing one volume scan every five minutes in precipitation mode, and completing one volume scan every ten minutes in clear-air mode. This radar can determine reflectivity out to 250 miles, and velocity and reflectivity out to 125 miles.

NEXRAD has three moments, or raw observed products: reflectivity, velocity (toward or away from the radar), and spectrum width. From these, several derived products are produced. These include one-hour and three-hour rainfall accumulation, vertically integrated liquid, echo tops, various composite products, and velocity azimuth display winds. Most of these products can be updated with each volume scan. Thus, where radar remote displays used to receive one product (base reflectivity) every two to three minutes, with NEXRAD, up to thirteen products is available every five minutes.

RADAR EQUIPMENT

The form of weather radars has changed little over the years, but substantial progress has been made in the electronic circuitry. This has resulted in excellent stability and reliability. Display devices have also made advances with color digital images instantly available for display, animation, or printing.

The basic weather radar unit consists of four components: The *antenna*, for sending and receiving the signal; the *pedestal/servo amplifier*, for driving the antenna array: the *transmitter/receiver*, for generating and processing the signal; and finally, the *display and operator control*, from which the radar is operated and the information is displayed.

The size of the antenna equipment is dependent on the wavelength chosen. S-band radars require the largest antennas. The antenna should be mounted on a tower or on the roof of a building where a good radar horizon is available. The waveguide, which connects the transmitter/receiver to the antenna, causes a power loss in both directions which depends on its length. It is therefore practical to minimize the distance between the antenna and the transmitter/receiver. Functions of the transmitter/receiver can be remoted back to the display controls so that the operator need not be near the unit itself.

Some additional equipment is recommended to ensure proper operation of the radar. A compressor-dehydrator is often used to pump dry air into the waveguide because dampness in the waveguide could cause a reduction in signal strength, or could cause arcing. Since power outages sometimes accompany severe weather, a standby power generator is often provided.

Radomes provide protective housings for antennas and are highly recommended. A radome will prevent damage to the antenna from precipitation, wind, falling ice, and pollutants. The radome also allows a lighter antenna to be used and eliminates the problem of uneven antenna rotation in high winds. Radomes should be large enough to allow maintenance personnel inside.

Weather radars generally use a conventional plan position indicator (PPI) which displays the echoes from a bird's eye perspective on a horizontal plane centered on the radar antenna. If the antenna is stopped and held at a given bearing, then tilted up and down, the display is known as an RHI or range-height indicator. This allows a vertical profile of precipitation to be made to determine storm structure and echo tops.

The data displayed on the radar screen is often processed by a digital video integrator and processor (DVIP). This can be best described as an intensity contouring process which continuously averages logarithmic video signals from the radar receiver in both range and azimuth. Quantitative measurements of precipitation can be made through this process with each displayed level of intensity related to an actual rainfall rate. Weather radars often show six levels of precipitation intensity, making it easy to locate the heavier rainfall. In more modern weather radars, a radar video processor (RVP) performs the intensity contouring process while simultaneously calculating the radial speed and direction of movement of precipitation, employing the Doppler principle.

SITING A RADAR

Finding the proper location for a weather radar is an important consideration. There is probably no perfect location for a radar site, but the following rules are important:

1. The horizon should be free and clear. The top of a building or hill often gives a clear view in all

directions, but siting a radar high above other objects allows the side-lobes to extend over a larger area. This will increase the ground clutter pattern of permanent echoes. A high site has best performance if the antenna operates with its main beam elevation below 0 degrees.
2. Permanent echoes should be minimized. It is best to have surrounding obstructions as close to 0 degrees elevation as possible to minimize clutter.
3. Maximum range should be sought. Operation with the antenna at higher elevation angles will eliminate some of the clutter but will reduce the range.
4. The ideal site would be centered on a low saucer with the antenna looking just over the rim. Since the topography may not allow this, consideration should be given to making this saucer by topping nearby trees and keeping them trimmed.
5. Availability of facilities such as power, buildings, and access roads is necessary to operate and maintain the unit. Proximity to the location where the radar will be operated is also helpful to reduce cabling costs.
6. Towers located near the radar site will not cause problems of signal blockage as long as they do not subtend an angle more than half that of the radar beamwidth.
7. Lightning protection is necessary, since a lightning strike can cause considerable damage to the antenna and associated equipment.

MAINTENANCE OF THE RADAR

Preventive maintenance of the radar should be performed on a routine basis in accordance with the manufacturer's recommendations. Modern radars have a relatively short operational check list, but these checks can provide early warnings of potential problems. Other suggested maintenance includes cleaning slip rings and greasing motor parts in the antenna system on a six month schedule.

Receiver performance should be checked weekly. Some modern units do a self-test each time they are initialized. A weather radar system should have a life expectancy of 10 to 15 years of constant use. Purchasers of weather radars should inquire into the record of similar equipment already in use as well as to the availability of spare parts.

If there is a complete set of spare parts available at the site, most repairs to a radar unit are typically completed within two to three hours. Replacement of mechanical components such as antenna drive motors or gear trains may take considerably more time.

RADAR REMOTING

An alternative to purchasing and operating a radar is to receive the actual images from a remote radar, operated by someone else. The most extensive and comprehensive network of weather radars is operated by the National Weather Service (NWS). There is a basic network of 77 S-band radars covering the continental United States east of the Rocky Mountains, with a 120 nautical mile (nm) radius circle of coverage. There are an additional 52 C-band radars used to supplement the S-band network, which only operate when precipitation is actually detected within 120 nm of the radar (see Figs. 3 and 4).

Radar remoting was developed in the late 1970s. Kavouras Inc. of Minneapolis, Minnesota, developed and patented the first commercial transmitter and receiver which allowed NWS radar data to be viewed remotely. Kavouras, Enterprise Electronics, Accu-Weather, and WSI are among the companies that provide weather data to broadcasters that is based on national radar networks.

Radar remoting has proven to be a very cost-effective means of providing top quality radar imagery at reasonable prices. Not only are data from any one NWS radar timely and quality controlled, but images are available from the entire country for the cost of a long distance phone call.

RADAC S-Band Color Weather Radar Network Map

The Radac Network transmits from the mammoth National Weather Service 500,000 watt S-band radar sites strategically located across the country. The realtime weather data is displayed in a choice of 60 nautical miles, 120 nautical miles, 180 nautical miles, or 240 nautical miles.

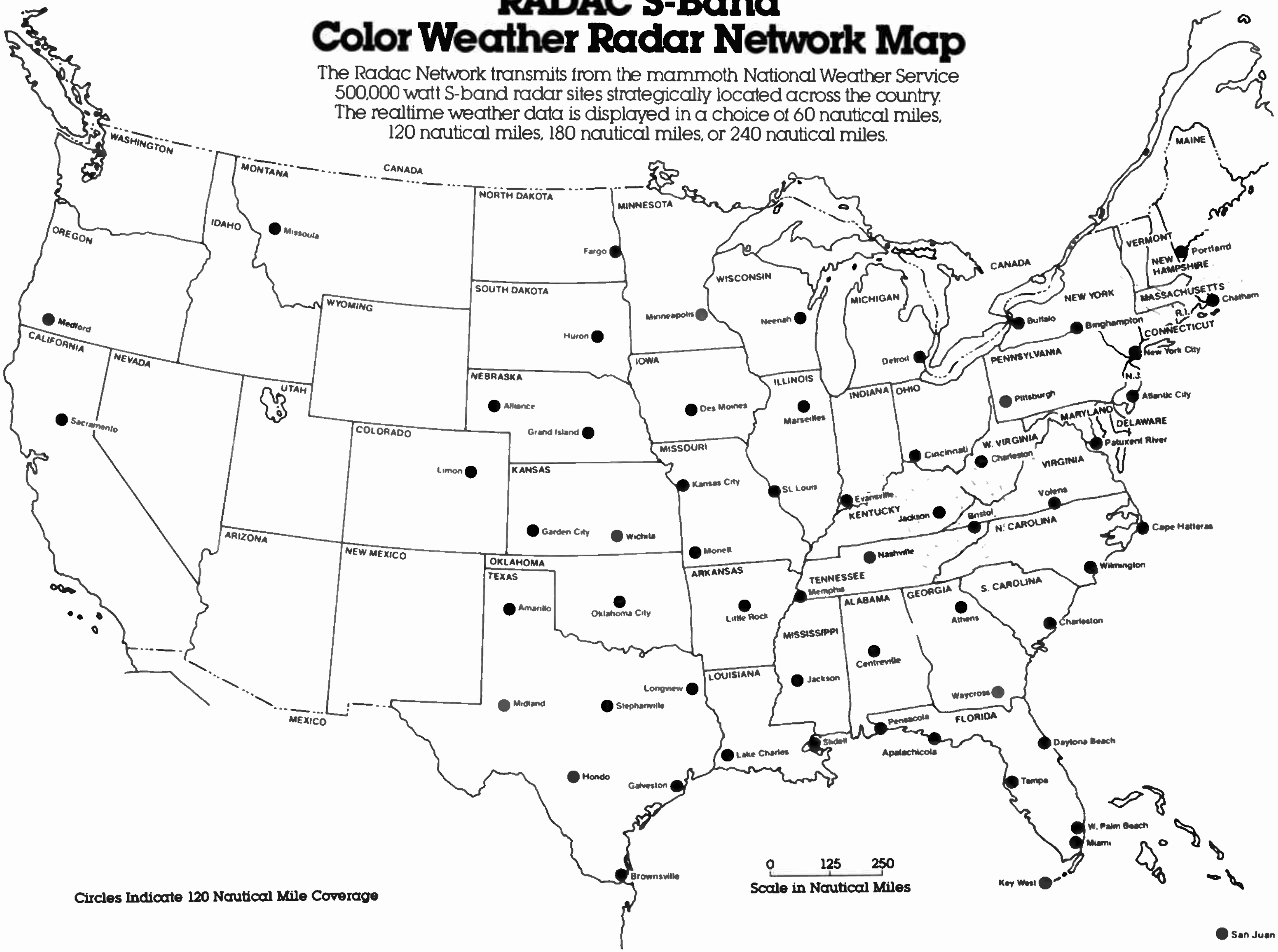


Figure 3. RADAC S-band weather radar network map.

RADAC C-Band Color Weather Radar Network Map

The Radac Network also transmits from the National Weather Service supplemental C-Band radar sites throughout the country. Less powerful than the S-Band, all C-Band data should be confirmed by area S-Band sites. The realtime data is displayed in a choice of 60 nautical miles, 120 nautical miles, 180 nautical miles or 240 nautical miles.

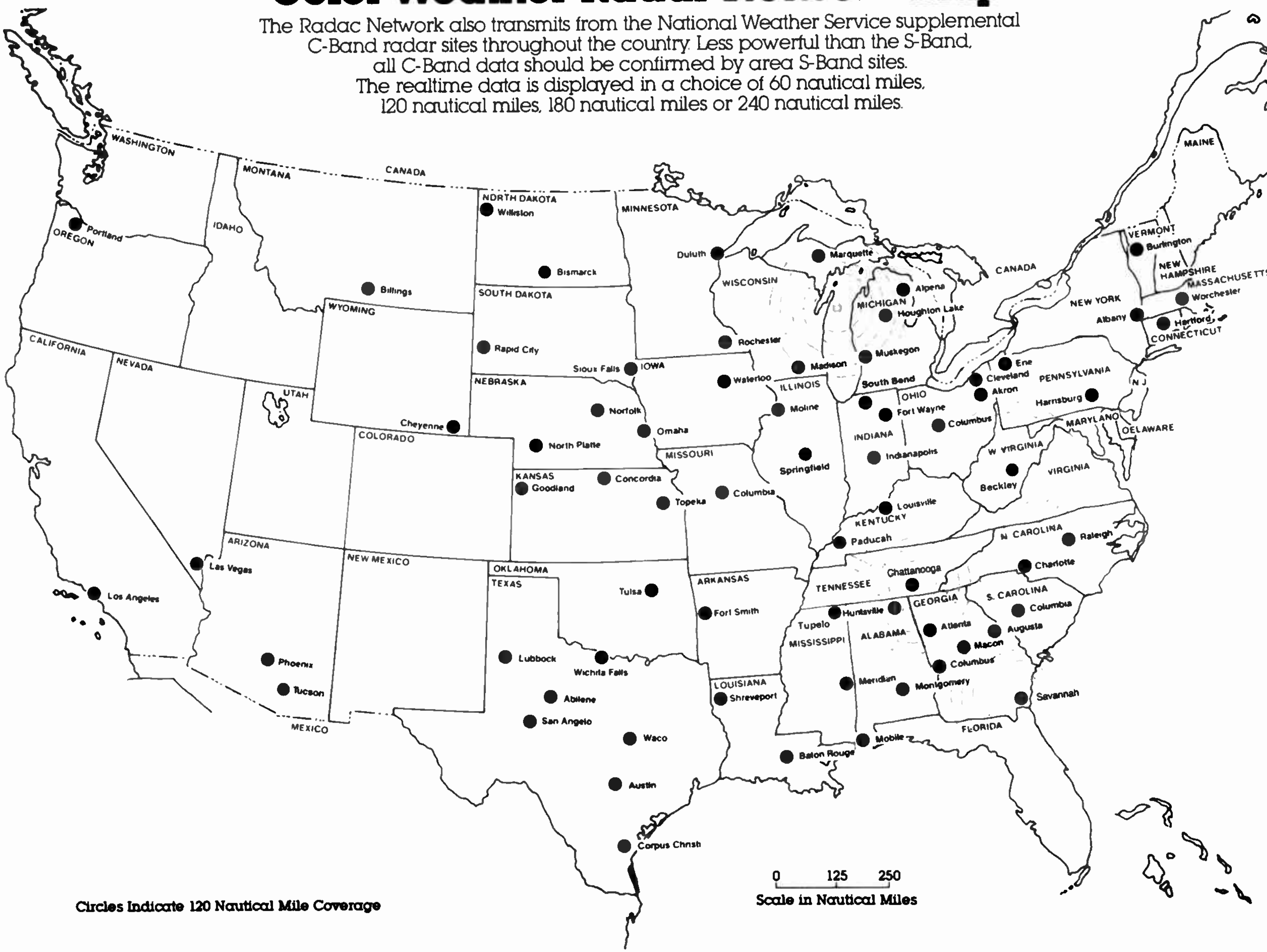


Figure 4. RADAC C-band weather radar network map.

7.7 Distance and Bearing Calculations

Dane E. Ericksen, P. E.
Hammett & Edison, Inc., San Francisco, California

INTRODUCTION

The current FCC Rules for distance calculations use two methods: "flat-earth" and "spherical." The flat-earth method assumes the distance to be the hypotenuse of a right triangle whose sides are determined by the difference in latitude and longitude of the starting and ending points multiplied by the length per degree of latitude and longitude at the mid-latitude of the two points, as shown in Fig. 1. The spherical method uses conventional spherical trigonometry to determine the distance. Section 73.208 of the FCC Rules now requires that the flat-earth method be used for distances up to and including 475 km. Distances greater than 475 km must be calculated using the spherical-earth method, which becomes more accurate than the flat-earth method for large distances. Section 73.208 is silent on how azimuths are to be calculated, nor does it specify the earth radius to be used for the spherical-earth case.

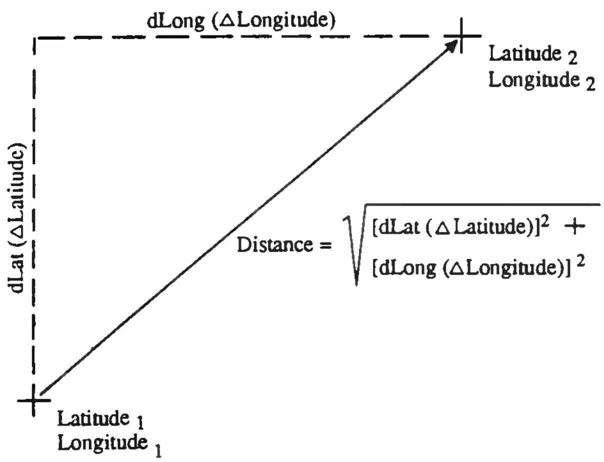


Figure 1. Flat-earth method.

FCC FLAT-EARTH METHOD

In FCC Docket 80–90[1], formulas were substituted for the tables previously used for determining the length of a degree of latitude or of longitude as a function of latitude. However, the coefficients adopted in Docket 80–90 truncated to only two terms the trigonometric series used to generate the tables and adjusted the coefficients by a factor of (1.609/1.609347) because of the Docket 80–90 decision to define the conversion factor from U.S. statute miles to kilometers as 1.609,[2] rather than the value of (5,280 feet/mile)(1200/3937 meters/foot)(1/1,000 kilometers/meter), or 1.609347219 km/mi (approximately)[3].

In the Second Report and Order to Docket 86–144[4], the FCC corrected these problems by adopting the full-precision, nontruncated trigonometric series for the arc length formulas given in the 1966 edition of U.S. Naval Hydrographic Office Publication Number 9, also known as "H.O. 9," or the "American Practical Navigator," or simply "Bowditch," after Nathaniel Bowditch (1773–1838), its original author. These trigonometric series are based upon a binomial theorem expansion[5] of an ellipsoid model of the earth corresponding to the Clarke spheroid of 1866, upon which topographic maps in the United States are currently based.[6]

The trigonometric series defining the length of one degree of latitude and one degree of longitude for the Clarke spheroid of 1866 are:

$$\text{dLat} = 111.13209 - 0.56605\cos(2L) + 0.00120\cos(4L)\ldots \quad [1]$$

$$\text{dLong} = 111.41513\cos(L) - 0.09455\cos(3L) + 0.00012\cos(5L)\ldots \quad [2]$$

where *dLat* is the length in kilometers of one degree of latitude at latitude *L* and *dLong* is the length in kilometers of one degree of longitude, again at latitude L.

The latitude, L, is taken as the mid-latitude of the two points between which the distance is to be calculated, as follows:

$$L = (\text{Latitude}_1 + \text{Latitude}_2)/2 \qquad [3]$$

where Latitude_1 and Latitude_2 are the latitudes of the starting and ending points. Similarly, Longitude_1 and Longitude_2 are the longitudes of the starting and ending points.

The distance between two points is then given by the Pythagorean theorem:

$$D = \sqrt{[(dLat)(Lat_1 - Lat_2)]^2 + [(dLong)(Long_1 - Long_2)]^2} \qquad [4]$$

Plots showing how the length of 1° of latitude and longitude vary with latitude are given in Figs. 2 and 3.

Canadian Method

In August 1987,[7] the Canadian government adopted the truncated and fudged arc-length formulas which had been implemented by the FCC in Docket 80–90. Namely,

$$\text{dLat} = 111.108 - 0.566\cos(2L) \qquad [5]$$

$$\text{dLong} = 111.391\cos(L) - 0.095\cos(3L) \qquad [6]$$

The Canadian government has not yet adopted the corrected, more accurate formulas which were implemented by the FCC in Docket 86–144. Nor was the flat-earth versus spherical-earth break point changed from 350 km to 475 km. This, then, is the source of current discrepancies between the U.S. and Canadian distance calculation methods. Although the differences between the two methods will usually not be significant when the calculated distance is rounded to the nearest kilometer (for FM) or to the nearest one-tenth kilometer (for TV), one should always check to see whether there is a difference between roundings. For calculations involving Canadian stations, the Canadian (Docket 80–90) version of the Clarke spheroid formulas is controlling.

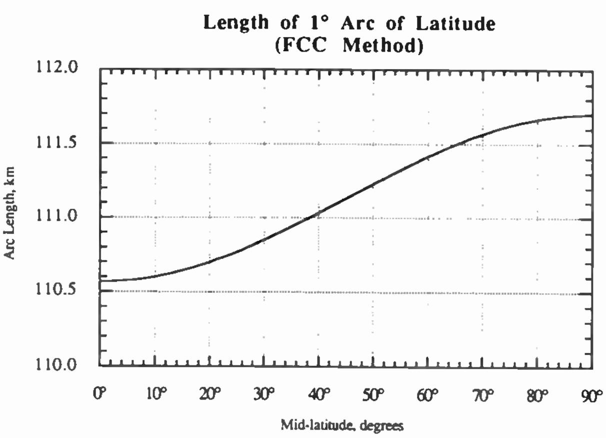


Figure 2. Length of 1° arc of latitude. (FCC method.)

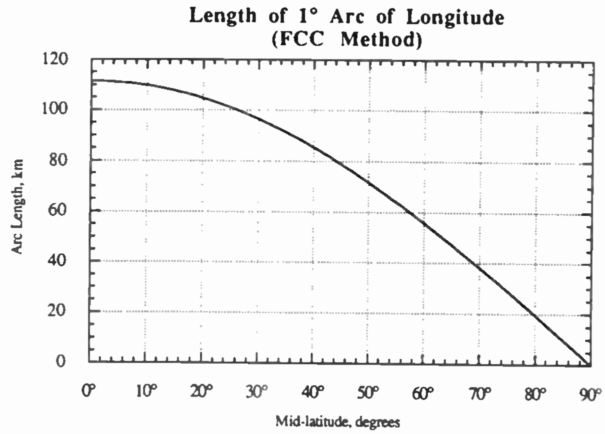


Figure 3. Length of 1° arc of longitude. (FCC method.)

Mexican Method

The U.S.-Mexican FM Agreement of 1972 has not been amended to update its distance calculation method from a look-up table to formulas. The agreement is only partially in metric units, and the primary units of the look-up table are in miles, not kilometers. This raises the question of what factor should be used when converting from statute miles to kilometers. It is unclear whether the approximate, truncated 1.609 factor specified in Docket 80–90 applies to a treaty obligation. While Mexico (and Canada) have adopted the international conversion factor of 1.609344 km/mile (exactly), the U.S. Metric Law of 1866 still defines the meter as equaling 39.37 inches, and no law has been passed by Congress to supercede it.[8] This is the source of the 1.609347219 km/mi (approximately) conversion factor in the United States. It is probably best to use the international 1.609344 factor where a treaty obligation with Mexico (or Canada) is involved. Readers are nevertheless cautioned to discuss this issue with their legal counsel or with the FCC staff if the exact conversion factor becomes critical.

The look-up table for lengths of one-degree arcs of latitude and longitude for mid-latitudes between 22° and 37° (given in Annex V, Part 2, of the U.S.-Mexican FM Agreement) is based upon Table 6 of the 1958 edition of H.O. Publication No. 9, "Length of a Degree of Latitude and Longitude". That table is, in turn, based upon formulas 1 and 2. Annex V stipulates that linear interpolation is to be used if the mid-latitude falls between the tabulated mid-latitude values. As shown in Figs. 4 and 5, there is virtually no difference in arc length between the two methods. However, care should be used if the calculated distance by the FCC method is close to a treaty minimum. There will still

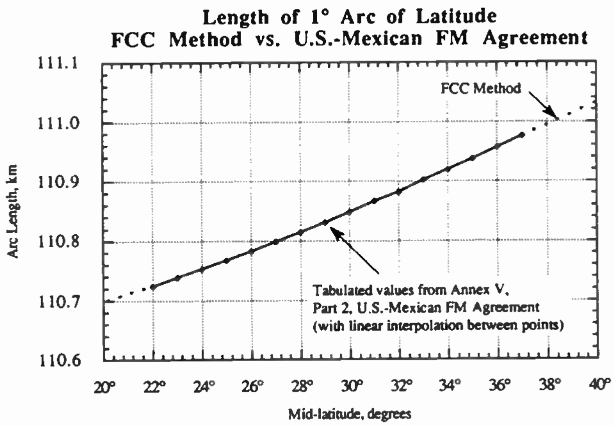


Figure 4. Length of 1° arc of latitude. (FCC method vs. U.S.-Mexican FM Agreement.)

be very slight differences in the arc lengths to be used when performing the flat-earth distance calculation method. This could cause a change when the required rounding to the nearest mile (not kilometer) specified in the U.S.-Mexican FM Agreement[9] is performed.

SPHERICAL-EARTH METHOD

The formula for the spherical-earth distance, or great-circle distance, is:

$$D = K*\arccos[(\sin Lat_1)(\sin Lat_2) + (\cos Lat_1)(\cos Lat_2)(\cos(Long_2 - Long_1))] \quad [7]$$

The constant K is in km/degree and is determined by the radius of the sphere being modeled. The FCC has never defined the earth radius to be used for spherical-earth calculations. The example given in Section 73.185(i) of the FCC Rules suggests an earth radius of 6,365 km (K = 111.090 km/degree). A 6,373 km radius (K = 111.230 km/degree) is implied by the 5,280 mile 4/3-earth radius given in Section 73.684(c)(1) of the FCC Rules. This 4/3-earth radius was also used in FCC Report No. R-6410, "Elevation and Depression Angle Tables," September 15, 1964. An earth radius of 6,367 km (K = 111.125 km/degree) can be deduced from the 1,852 meter definition of a nautical mile.[10] Finally, an earth radius of 6,371 km (K = 111.195 km/degree) corresponds to the mean radius of the Clarke spheroid of 1866.

AZIMUTH CALCULATIONS

Because the FCC Rules are silent on how azimuth, or bearing, calculations are to be performed, both the flat-earth and spherical methods are commonly in use. The flat-earth method determines azimuth using the arctangent of the right triangle defined in the FCC method. The spherical-earth method uses standard spherical trigonometry. The flat-earth method will be in error by up to two degrees at distances approaching 500 km, whereas the spherical-earth azimuth will be correct within about 0.1 degrees at such distances.[11] Fig. 6 shows why this is so. Azimuths determined using the arctangent of the right triangle defined by the FCC flat-earth method assume that lines of longitude are parallel, whereas they are not. Thus, the spherical-earth method is more accurate, even when the FCC flat-earth method is used to determine distance. This is why the forward and back (reciprocal) azimuths using the spherical-earth method will generally not be exactly 180° apart, whereas the forward and back azimuths using the flat-earth method are always 180° apart.

The formula for determining azimuth by the spherical-earth method is:

$$C = \cos^{-1}\left[\frac{\sin Lat_2 - \sin Lat_1 \cos(D/K)}{\sin(D/K)\cos Lat_1}\right] \quad [8]$$

$$\textit{if}\ \sin(Long_2 - Long_1) < 0,\ Azimuth = C \quad [9]$$

$$\textit{if}\ \sin(Long_2 - Long_1) \geq 0,\ Azimuth = 360° - C \quad [10]$$

The ratio (D/K) is the great circle arc length in degrees, and is obtained from formula 7. It should be noted that the Canadian Department of Communications Rules *do* specify that azimuth is to be calculated using the spherical-earth method and, further, that azimuths are to be rounded to the nearest degree.[12] It should also be noted that the radius of the sphere is irrelevant for azimuth calculations.

CLARKE SPHEROID VERSUS WGS ELLIPSOID

The original FCC distance tables were based upon the Clarke spheroid (or ellipsoid) of 1866, with a major-axis radius of 6,378.2064 km and ellipticity[13] of 1/294.98. The current edition of the American Practical Navigator[14] now bases Table 6, "Length of a Degree of Latitude and Longitude," on the World Geodetic

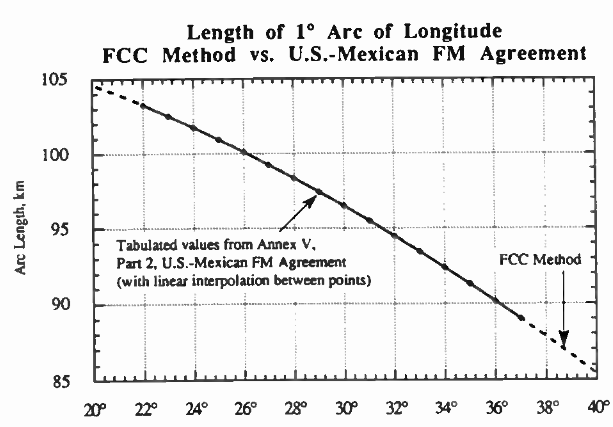


Figure 5. Length of 1° arc of longitude. (FCC method vs. U.S.-Mexican FM Agreement.)

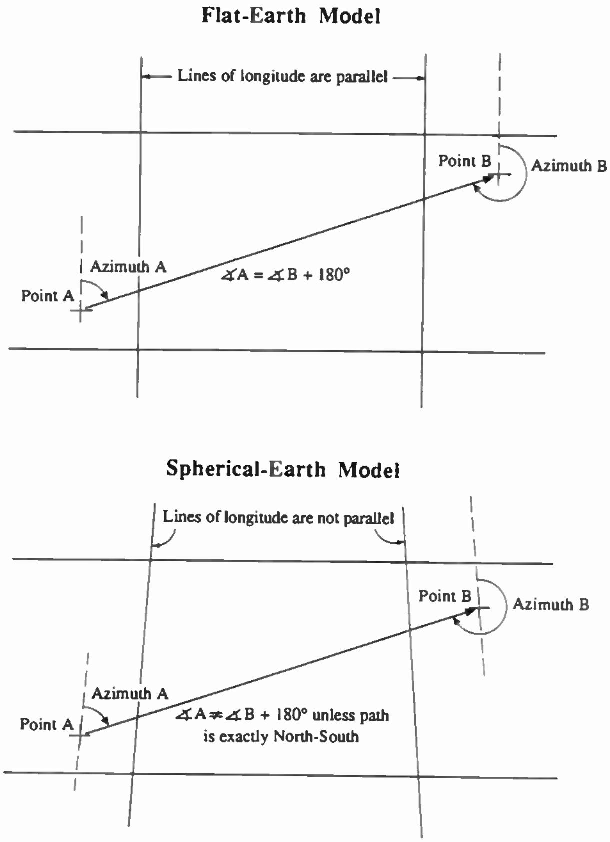


Figure 6. Azimuth calculations.

System (WGS) ellipsoid of 1972. The coefficients for *dLat* and *dLong* for this ellipsoid are:

$$dLat = 111.13292 - 0.55982\cos(2L) + 0.001175\cos(4L)\ldots \quad [11]$$

$$dLong = 111.41282\cos(L) - 0.0935\cos(3L) + 0.000118\cos(5L)\ldots \quad [12]$$

The difference between the Clarke 1866 and WGS 1972 trigonometric series is trivial when coordinate data bases are maintained only to the nearest second and distances are rounded to the nearest kilometer or ⅒ km. Over a 0° to 90° arc of latitude, at 1° increments, the RMS difference between the two arc length formulas is only 0.0041% for latitude and 0.0033% for longitude. For this reason the FCC decided not to amend the formulas given in Section 73.208(c) of its Rules from the Clarke 1866 coefficients and to WGS 1972 coefficients.

ROUNDING PRACTICES

There continues to be an inconsistency in rounding practices between the FCC FM and TV Rules. For FM distance calculations, Section 73.208(c)(7) specifies rounding to the nearest kilometer. For TV distance calculations, Section 73.611(d) specifies rounding to the nearest one-tenth kilometer. The FCC was asked to eliminate this inconsistency between the FM and TV Rules in the Docket 86–144 comments, but no action was taken. The Docket 86–144 Report and Order did not discuss this point, so it is not known whether the failure to adopt a uniform rounding policy was inadvertent or deliberate. Persons performing distance calculations must, therefore, be mindful to the different rounding criteria between the FM and TV Rules. The AM Rules and the Broadcast Auxiliary Rules are silent on distance rounding practices. The Broadcast Auxiliary Rules also do not prescribe how distances are to be calculated, even though portions of those Rules are now based on minimum distance requirements. Calculations regarding distances between TV Translator and LPTV stations, should, presumably, use the flat-earth method specified in the FCC's FM Rules.

NAD27/NAD83 DATUMS

The Federal Geodetic Committee is in the process of converting all maps in the United States from the 1927 North American Datum ("NAD27"), which is based upon the Clarke spheroid of 1866, to the 1983 North American Datum ("NAD83"). NAD83 differs from NAD27 in that it is referred to the earth's center of mass, making it fully compatible with satellite systems for position determination. The FCC has indicated it will eventually convert to NAD83 to maintain accuracy and consistency with other government agencies. NAD83 is based upon the Geodetic Reference System ellipsoid of 1980 ("GRS 1980"), with a major axis of 6,378.135 km and an ellipticity of 1/298.26. These are the same parameters as for the WGS 1972 ellipsoid, and the trigonometric series for the one degree arc length formulas are identical to those for WGS 1972.[15]

Current 7.5-minute topographic quadrangle maps published by the Geological Survey specify NAD27 as the reference datum, in the lower left corner of the map. At some point in the future, the Geological Survey will begin to issue maps based on the NAD83 datum. Initially, dual-sided (NAD27/NAD83) maps will be published.[16]

In order to prevent intermixing of numerical information using two different map datums, the FCC has stated[17] that the following procedures will be in effect until further notice:

1. Applicants must continue to furnish coordinates based on NAD27.
2. Applicants who have filed applications with coordinates based on the NAD83 datum must provide NAD27 coordinates.
3. Until further notice, the FCC will continue to specify NAD27 coordinates in its databases, authorizations, notifications, forms, and rules.
4. The FCC will issue further guidance on the conversion to NAD83 as more information becomes available.

CONVERSION ALGORITHMS BETWEEN DATUMS

Two conversion algorithms presently exist for conversion of NAD27 geographical coordinates to NAD83 geographical coordinates. By far the simpler algorithm is one created by the National Geodetic Survey (NGS)[18]. Although it uses polynomials through the 13th order, it is still quite feasible to implement on a programmable calculator or computer. The primary advantage of the NGS algorithm is that it is valid for the entire contiguous United States. The second algorithm is one developed by the United States Geological Survey (USGS). The USGS algorithm uses 10th order polynomials, but requires a multi-volume table of polynomial coefficients—one set of coefficients for each of the approximately 55,000 7.5-minute series topographic maps published by the USGS.

The NGS algorithm is reported to have an accuracy of approximately ±0.04 seconds in latitude and ±0.06 seconds in longitude; this is based on the NGS description that the RMS errors for both latitude and longitude conversion are ±1.3 meters.[19] In contrast, the USGS algorithm is reported to have conversion accuracies on the order "a few decimeters."[20] This corresponds to accuracies of about ±0.006 seconds in latitude and ±0.010 seconds in longitude, but has the disadvantage of being much more burdensome to implement. Either a reference volume of coefficients or a computer with a very large memory would be required to implement the USGS algorithm.

REFERENCES

1. Docket 80–90 *Report and Order*, May 26, 1983.
2. Docket 80–90 *Report and Order*, May 26, 1983, page 29, footnote 35.
3. ANSI/IEEE Standard 268–1982, "Metric Practice," page 31, note 14.
4. Docket 86–144 *Second Report and Order*, September 10, 1987.
5. Personal correspondence between the author and Mr. Adam W. Mink, Chief, Hydrography and Navigation Department, Defense Mapping Agency, Washington, DC 20315–0030. May 24, 1986.
6. "Maps For America," Second Edition, 1981, page 238. Published by the U.S. Department of the Interior, Geological Survey National Center, Reston, VA 22092.
7. Broadcast Procedure No. 13 (BP-13), Issue 2, Broadcasting Regulation Branch, Department of Communications, Government of Canada. Effective August 6, 1987.
8. Personal correspondence between the author and Mr. Robert B. McEwen, Acting Assistant Division Chief for Research, National Mapping Division, Geological Survey, U.S. Department of the Interior, Reston, VA 22092. May 8, 1986.
9. Annex V, Part 1, Section B5, U.S.-Mexican FM Agreement of 1972.
10. H.O. No. 9, 1981 Edition, Volume 2, page 862, defines a nautical mile as one minute of any great circle of the earth. In 1929, the International Hydrographic Bureau proposed a standard length of 1,852 meters (exactly), which is known as the International Nautical Mile. A nautical mile of 1,852 meters implies an earth radius of 6,366,707 km ((1,852 meters/minute x 60 minutes/degree x 360 degrees/circumference)/2!À).
11. Azimuth errors are referenced to values obtained from Andoyer-Lambert formulas. Andoyer-Lambert formulas model the earth as a true ellipsoid and are used extensively in Loran computations. The complexity of Andoyer-Lambert formulas do not warrant their routine use for FCC calculations.
12. Broadcast Procedure No. 13 (BP-13), Section 4.3.
13. Ellipticity, or flattening, f, is defined as f = (1 - b/a), where a is the equatorial radius and b is the polar radius.
14. The *American Practical Navigator* is now published in two volumes by the Defense Mapping Agency, Hydrographic Center, Combat Support Center, Attention: PMSR, Washington, DC 20315–0020; telephone (301) 227–2157. Volume I (DMA stock number NVPUB9V1) is dated 1984; Volume II (DMA stock number NVPUB9V2 is dated 1981. As of June 1990, each volume is $16.25, postpaid.
15. The major axis and ellipticity for WGS1972 and GRS1980 are identical, according to NOAA Technical Memorandum NOS NGS-16, "Determination of North American Datum 1983 Coordinates of Map Corners (Second Prediction," by T. Vincenty, National Geodetic Survey, January 1987. Because the binomial theorem expansion for an ellipsoid model starts with only two constants, the major axis dimension and the ellipticity, the trigonometric series for two ellipsoid models of the earth with the same major axis and ellipticity values must also be identical.
16. "Implementing North American Datum 1983 for the National Mapping Program (Ashaway Quadrangle)," U.S. Department of the Interior, Geological Survey, Reston, VA 22092. Undated.
17. FCC Public Notice "FCC Interim Procedure for the Specification of Geographic Coordinates," March 14, 1988.
18. Defense Mapping Agency Technical Report, "World Geodetic System 1984, Its Definition and Relationships with Local Geodetic Systems," DMATR 8350.2, September 30, 1987.
19. Ibid, at Table 7.6, page 7–31.
20. Personal correspondence between the author and Ms. Elizabeth Wade, Chief, Horizontal Network Branch, National Geodetic Survey, National Oceanic and Atmospheric Administration, Rockville, MD 20852. September 27, 1988.

7.8 The Emergency Broadcast System

Frank M. Lucia
The Federal Communications Commission, Washington, District of Columbia
Dane E. Ericksen, P.E.
Hammett & Edison, Inc., San Francisco, California

SYSTEM ORGANIZATION

The Emergency Broadcast System (EBS) uses the facilities and personnel of the entire communications industry (broadcast stations and networks, telephone companies, national press services, and national cable programmers) on a voluntary, organized basis to establish an emergency broadcasting network. This network is operated by the industry under government regulations and procedures and in a manner consistent with national security requirements. Broadcast station licensees participating in the EBS have been issued Emergency Broadcast System Authorizations by the Federal Communications Commission (FCC). Under peacetime conditions, presidential broadcasts are handled by existing nongovernment radio and television facilities. Under conditions that would call for the activation of the EBS the normal flow of communications could be disrupted, altered, or destroyed.

A national level Emergency Action Notification (EAN) will be released only upon presidential authority. Release of the EAN constitutes the notice to all broadcast stations, participating radio and television networks, and national cable programmers of an emergency situation. Both the Emergency Action Notification and Termination are transmitted only over the EBS. Upon activation of the national level EBS, the White House Communication Agency (WHCA), which is responsible for providing all communications for the president under all conditions, will deliver the presidential messages to selected originating points. From these points, the presidential messages or broadcast will be distributed to participating EBS stations and cable systems via the nationwide emergency action notification network.

The FCC has been assigned the overall responsibility for the development of the EBS. The FCC ensures effective coordination between that agency, the Federal Emergency Management Agency (FEMA), and other government and nongovernment agencies concerned. The EBS Advisory Committee (EBSAC) has been organized to advise and assist the FCC and other appropriate authorities. It is to study and submit recommendations for emergency communications policies, plans, systems, and procedures for all licensed and regulated communications in order to provide continued emergency communications service during emergency situations.

Key in development of the EBS is the EBSAC. This committee, with the assistance of special working groups, provides the FCC with continuing advice and recommendations to ensure a workable EBS. Members are responsible for providing advice and assistance in programming guidance, production, and other operations of the national level interconnecting facilities and systems voluntarily participating in the EBS.

A State Emergency Communications Committee (SECC) has been organized in each of the 50 States, American Samoa, Guam, Puerto Rico, Virgin Islands, and the District of Columbia. The function of the SECC is to prepare coordinated operational emergency communications plans, systems, and procedures for their areas. State plans must be consistent with approved national level plans.

An Operational Area Emergency Communications Committee, which functions as a subcommittee of the SECC, has been organized within geographical operational (local) areas designated by SECC and state authorities. An operational (local) area may include one or more communities; portions of two or more states may be included in borderline situations. The function of an Operational Area Emergency Communications Committee is to develop operational emergency communications systems, plans, and procedures for inclusion in the state EBS plans for use during local level day-to-day emergency situations.

Again, participation in the EBS by FCC licensees is on a voluntary basis. The FCC sends new licensees an EBS authorization and a letter requesting their voluntary participation in the EBS. Licensees subsequently receive an appropriate EBS checklist, an operational area mapbook, relevant state and local EBS plans, a special instruction card to be posted at AP/UPI teletypewriter machines, and the EBS Rules and Regulations.

Participating stations that remain on the air during a national level emergency situation must carry presidential messages "live" at the time of transmission. Activities of the national level EBS will preempt operation of the state or operational (local) area EBS.

National programming and information which is not broadcast at the time of original transmission will be recorded locally by the Common Program Control Station (CPCS) for broadcast at the earliest opportunity consistent with operational (local) area requirements.

SYSTEM OPERATION

National Level Activation and Termination (Fig. 1)

Implementation

The Emergency Action Notification (EAN) will be released at the national level upon request of the White House. When the White House directs activation of the national level EBS, the White House Communications Agency, after a series of interim steps, issues instructions to implement the EBS. The EAN message is then

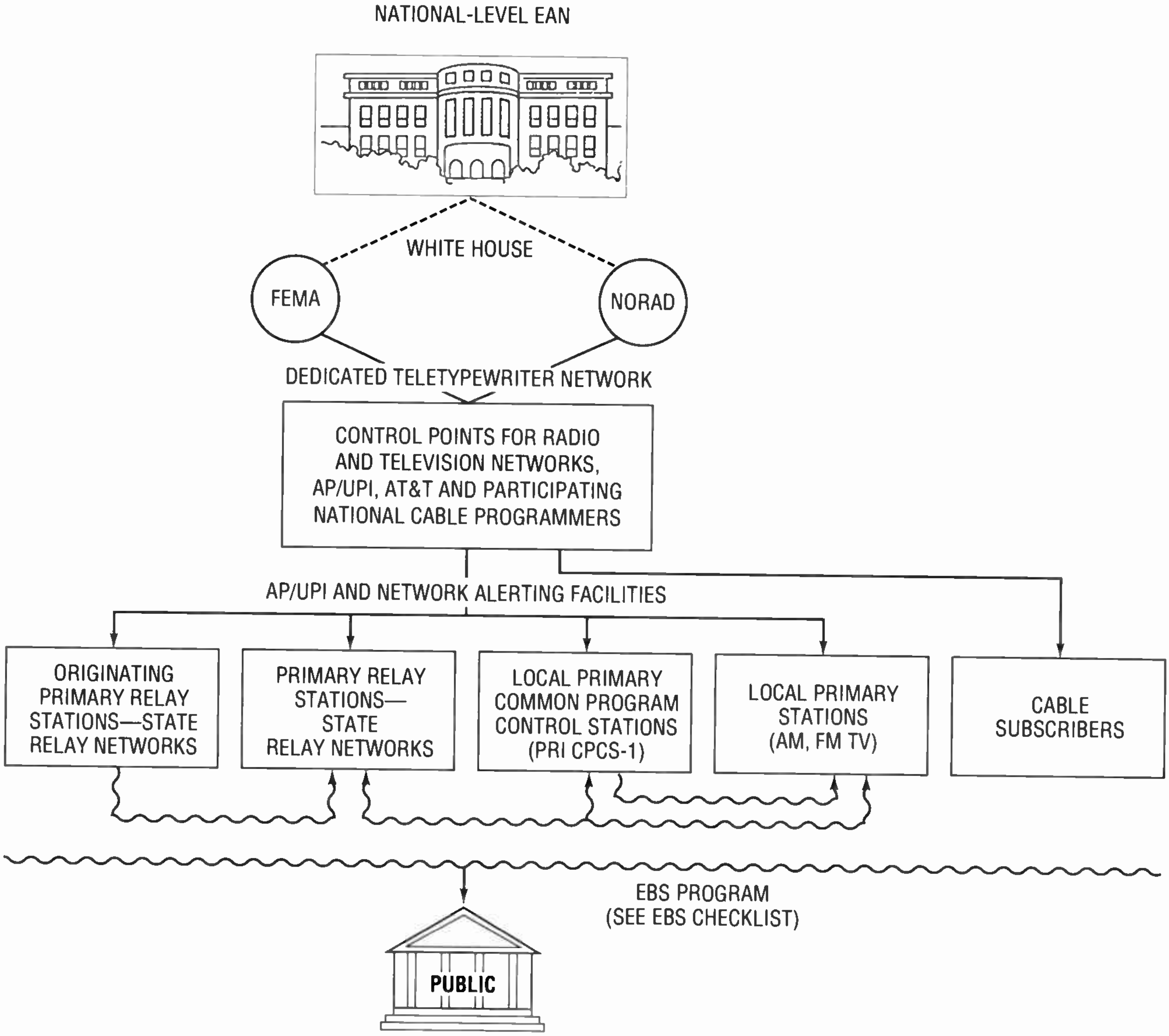


Figure 1. National level activation and termination.

released and is eventually received by the nation's broadcast stations. A dedicated network connecting the radio and television broadcasting networks and wire services transmits these messages. Following the EAN, there is a pause to allow time for the broadcasting networks to transmit a message over their internal alerting facilities, alerting stations that normal programming will be preempted by the EBS. Simultaneously, the EAN is transmitted over the respective Radio Wire Teletype Networks to alert those stations equipped to receive this service. Broadcast stations not equipped to receive either network alerting information or Radio Wire Teletype Network transmissions must rely on receipt of the EAN via off-the-air EBS monitoring from other stations. Receipt of the EAN by any one of the methods discussed above is sufficient for the broadcast stations to commence emergency actions.

Station Responsibility

Upon receipt of the EAN certain actions are taken by all stations. These actions are:

1. Authenticate the EAN.
2. Discontinue normal programming and broadcast a special announcement alerting the public to the fact that important instructions are forthcoming.
3. Transmit the Attention Signal.
4. Broadcast the message which informs the public of the fact that an emergency situation exists, that some stations will remain on the air, and that additional news and information will follow. Those stations required to cease transmitting will so inform their listeners.
5. Television stations must ensure that emergency information is transmitted both aurally and visually if the EBS has been officially activated. If the EBS has not been officially activated, then TV stations may transmit emergency information visually only, if desired. TV stations may use any method of visual presentation which results in a legible message conveying the essential emergency information. Methods which may be used include, but are not necessarily limited to, slides, open electronic captioning, manual methods such as hand printing, or mechnical printing processes.

The actions taken by the broadcast stations following transmission of the attention signal depend on the station designation. Stations fall into the categories of:

- Primary Station (Primary)
- Primary Relay Station (PRI Relay)
- Common Program Control Station (CPCS)
- Originating Primary Relay Station (Orig. PRI Relay)
- Nonparticipating Station (Non-EBS)

Participating stations monitor the CPCS or PRI Relay station in their area or radio and television broadcasting networks for further instructions and broadcast emergency programming when it becomes available. Nonparticipating stations remove their carriers from the air and monitor for the termination of the emergency. Should it become apparent that the CPCS or PRI Relay stations may not be able to provide appropriate emergency program feed, other participating stations may elect to assume the duties by providing program feed.

When the national level EBS is no longer needed, the emergency action termination message is transmitted. Broadcast stations receive the notification in the same manner as activation. The common carriers then return the networks to normal configuration, and broadcast stations resume normal programming in accordance with their regular station authorization.

System Tests

Periodic teletype test transmissions. The wire services will separately conduct test transmission to AM, FM, and TV broadcast stations on their Radio Wire Teletype Networks, a maximum of twice a month on a random basis at times of their choice. The subscribing broadcast stations enter the date and time of receipt of these tests consistently in their logs or records.

Closed circuit tests. These tests of the EBS will be conducted on a random or scheduled basis not more than once a month and not less than once every three months but only after FCC approval. Scheduled closed circuit tests will be conducted at a time selected by the White House, the EBS Advisory Committee representatives, and the FCC Defense Commissioner. The Closed Circuit Test Activation Message is disseminated to the various radio stations by the internal alerting facilities of the radio networks and the Radio Wire Teletype Networks to all subscribers.

The common carriers do not add participating independent stations to any of the radio networks during a closed circuit test, unless ordered by the FCC.

During a closed circuit test, broadcast stations do not interrupt programming and do not broadcast the message. The radio stations are required to:

- Monitor the radio network for the test program
- Check the wire services teletype
- Authenticate the message
- Record the time of the test consistently in station logs or records

Because of the limited time available for the closed circuit test, the termination of the test will occur on the following closing cue as it appears in the test of the program:

> "This concludes the closed circuit test of the Emergency Broadcast System."

Weekly transmission tests. All radio and TV stations are required to conduct a weekly test of the attention signal. This must be done once a week at a random day and time between 8:30 am and local sunset.

Note: Activation of the EBS or coordinated state/local tests may be substituted for the weekly test during the week the activation or coordinated test occurred.

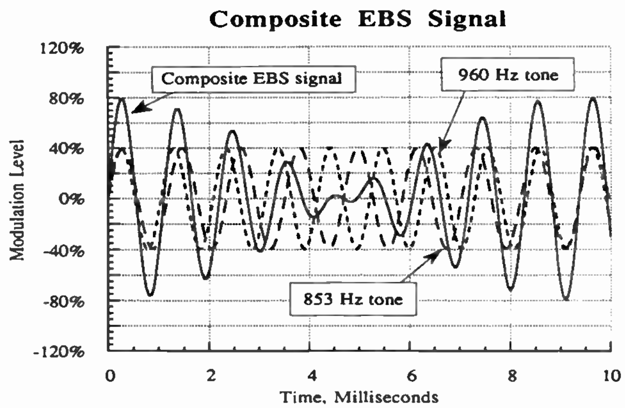


Figure 2. Composite EBS tone showing phase relation of the two component frequencies.

EBS Attention Signal Modulation Levels

The Commission's Rules require that the 853 Hz and 960 Hz tones which make up the attention signal each modulate the transmitter no less than 40% and that the modulation levels be within one decibel of each other. Because both tones must be simultaneously present to create the attention signal, and because the tones are not harmonically related, the modulation level of both tones in combination must therefore attain at least 80% modulation. Figs. 2 and 3 show the relationship between the 853 Hz and 960 Hz tones and the composite attention signal. For the currently prescribed 20 to 25 second attention signal duration, there will be at least 2,000 occurrences where the two tones are in phase and a peak modulation of 80% results if each tone individually modulates the transmitter 40%. Thus, the peak flasher of a properly operating pre-1983 type approved modulation monitor should easily detect the peak modulation of the attention signal.

Operators are cautioned that, because of limiters or other audio processing equipment which may be in the audio chain after the EBS encoder, it is entirely possible that a single EBS tone, by itself, will modulate the transmitter 40%, but that both tones in combination will modulate the transmitter substantially less than 80% due to nonlinear effects of downstream audio processing equipment. Operators should ensure that the composite modulation of the attention signal is at least 80%, in addition to checking that the modulation of each individual tone is at least 40% and that the levels are within 1 decibel of each other. This may require the bypassing of downstream audio limiters when the EBS encoder is activated.

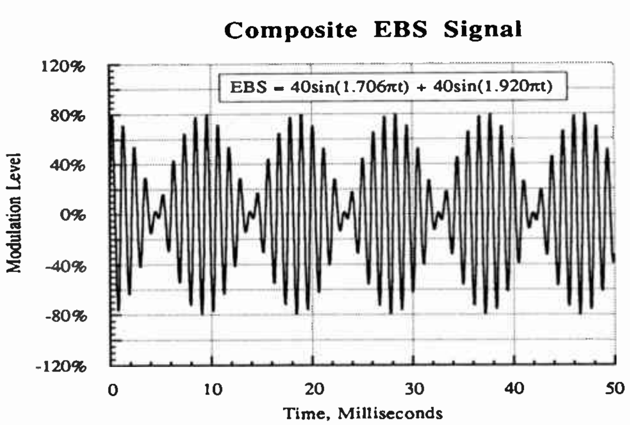


Figure 3. Composite EBS tone showing variation of modulation level of composite tone.

State Level EBS Operation (Fig. 4)

Implementation

Upon receipt of a state level EAN, all broadcast stations, including stations operating under equipment or program test authority, may, at the discretion of management, conduct operations in accordance with the provisions of the state level EBS Plan. Day-to-day emergencies posing a threat to the safety of life and property which could cause activation of the state level EBS include tornadoes, hurricanes, floods, tidal waves, earthquakes, icing conditions, heavy snows, widespread power failures, industrial explosions, and civil disorders. In most instances the state level EAN will be released from the state Emergency Operations Center (EOC). Common carrier or remote pickup units (RPU) are used to provide communications from the EOC. An FCC EBS authorization is not required for a broadcast station to participate in the operation of the state level EBS. Receipt of the state level EAN will

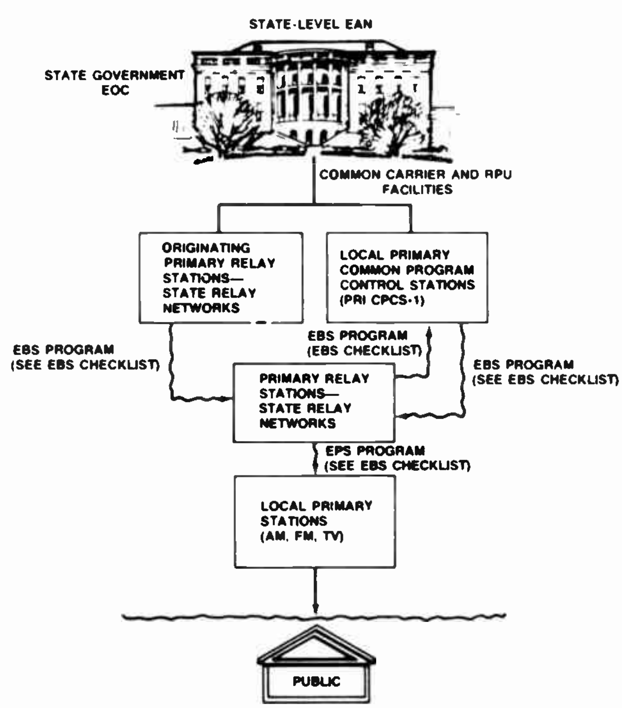


Figure 4. State level EBS operation.

be through off-the-air EBS monitoring or as otherwise stipulated in the state EBS plan.

Station Responsibility

Actions to be taken by the broadcast stations after receipt of the EAN are as follows:

1. Monitor the State Relay Network (PRI relay stations) for receipt of any further instructions from the originating PRI relay station.
2. Monitor the primary station designated as the CPCS-1 for your operational (local) area for receipt of any further instructions.
3. Discontinue normal program operation and broadcast the alert message.
4. Transmit the attention signal.
5. Broadcast the state level EBS EAN message.
6. Upon completion of the above transmission, resume normal programming until receipt of the cue from the CPCS for the operational (local) area, or primary relay station for the state EBS network. Then begin broadcasting the common state level program.
7. Television stations must ensure that emergency information is transmitted both aurally and visually if the EBS has been officially activated. If the EBS has not been officially activated, then TV stations may transmit emergency information visually only, if desired. TV stations may use any method of visual presentation which results in a legible message conveying the essential emergency information. Methods which may be used include, but are not necessarily limited to, slides, open electronic captioning, manual methods such as hand printing, or mechnical printing processes.

Upon receipt of notification of the termination of the state level EBS, participating broadcast stations will resume regular operations in accordance with the station authorization.

Operational (Local) Area EBS

Implementation

Upon receipt of an operational (local) area EAN, all broadcast stations, including stations operating under equipment or program test authority, which are voluntarily participating, may, at the discretion of management, conduct operations in accordance with the provisions of the state EBS plan. Day-to-day emergencies posing a threat to the safety of life and property which could cause activation of the operational (local) area EBS include tornadoes, hurricanes, floods, tidal waves, earthquakes, icing conditions, heavy snows, widespread power failures, industrial explosions, and civil disorders. The operational (local) area EAN will be released from the local government Emergency Operations Center (EOC) or a National Weather Service (NWS) office. Common carrier or remote pickup units are used to provide communications from the EOC. An FCC EBS authorization is not required for a broadcast station to participate in the operation

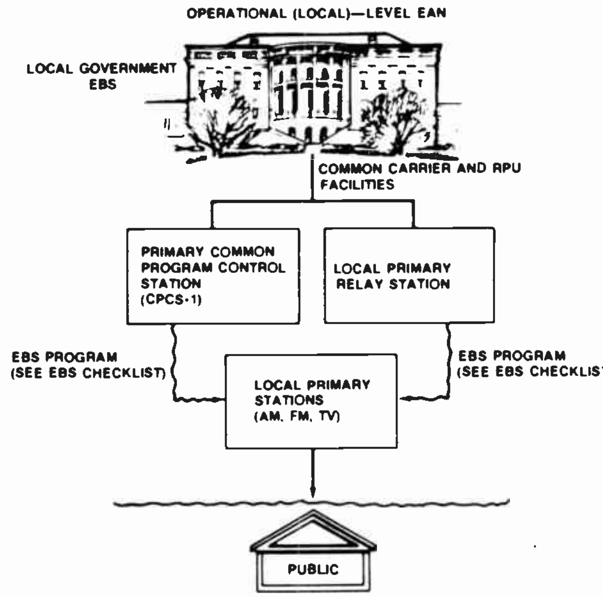


Figure 5. Operational (local) area EBS.

of the operational (local) area EBS. Receipt of the operation (local) area EAN will be through the means of off-the-air monitoring or as otherwise stipulated in the state EBS plan.

Station Responsibility

Actions to be taken by the broadcast stations after receipt of the EAN are as follows:

1. Monitor the primary station designated as the CPSC-1 for the operational (local) area for the receipt of instructions.
2. Discontinue normal program operation and broadcast the alert message contained in the EBS checklists.
3. Transmit the attention signal.
4. Broadcast the operational (local) area EBS EAN message.
5. Upon completion of the above transmission, resume normal programming until receipt of the cue from the CPCS-1 for the operational (local) area. Then begin broadcasting the common program.
6. Television stations must ensure that emergency information is transmitted both aurally and visually if the EBS has been officially activated. If the EBS has not been officially activated, then TV stations may transmit emergency information visually only, if desired. TV stations may use any method of visual presentation which results in a legible message conveying the essential emergency information. Methods which may be used include, but are not necessarily limited to, slides, open electronic captioning, manual methods such as hand printing, or mechnical printing processes.

Upon receipt of the termination of the operational (local) area level EBS, participating broadcast stations will resume regular operations in accordance with the station authorization.

State and Local Tests

Tests of implementing procedures developed at the state and local level may be conducted on a day-to-day basis as indicated in the state EBS operational plans. Coordinated tests of EBS operational procedures for an entire state or operational area may be conducted in lieu of the weekly transmission tests.

EMERGENCY BROADCAST SYSTEM (EBS) CHECKLIST FOR PARTICIPATING STATIONS

The following is a condensation of the EBS checklist, and contains simplified instructions for each type of situation. The checklist contains step-by-step procedures for stations to follow when receiving a national level emergency action notification. Also the termination procedures to be followed by all licensees at the conclusion of the national level emergency. The checklist provides detailed procedural information as to what action must be taken when in receipt of national level closed circuit tests, the periodic teletype tests, and the weekly transmission tests. Lastly, step-by-step instructions are included for state and local level procedures and tests. More detailed information may be found, concerning these simplified instructions, by referring to Part 73 of the Commission's Rules and Regulations.

National Level Instructions

When EBS is activated by the White House, all broadcast stations must take the following actions. Stations should record all emergency broadcasts (including notification) and log or record all significant events.

Activation Procedures: All Stations

1. RECEIVE EMERGENCY ACTION NOTIFICATION.

(This is the notice which activates EBS. Presidential messages may be received almost immediately.)

Notification by one of the following methods is sufficient to commence action:

- AP and UPI TELETYPE—Preceded and followed by a line of "X"s and an alarm.
- RADIO-TV NETWORKS—Affiliates only. Preceded by network alerting signal. Continue to monitor for further instructions.
- EBS MONITOR RECEIVER—Preceded by two-tone attention signal. Step 5 for message format. Continue to monitor for further instructions. Your EBS monitoring assignment is specified in the state EBS operational plan or map book.

EMERGENCY ACTION NOTIFICATION MESSAGE—AP AND UPI SUBSCRIBERS/NETWORK AFFILIATES ONLY

> "This is an Emergency Action Notification requested by the White House. The AUTHENTICATOR WORDS for this notification are (—). All broadcast stations follow procedures in the EBS Checklist for national level emergency. The president of the United States or his designated representative will shortly deliver a message over the EMERGENCY BROADCAST SYSTEM. The authenticator words are (—)."

2. AUTHENTICATE NOTIFICATION.

AP and UPI Subscribers/Network Affiliates Only

Compare authenticator words in notification with words on current EBS authenticator list *(contained in the red envelope inside the front cover of the EBS Checklist)*. Take no further action if words do not match.

Different activation and termination words are provided for each date. Words are effective 12:01 a.m., Washington, DC, time.

3. DISCONTINUE NORMAL PROGRAMMING AND BROADCAST ANNOUNCEMENT.

(TV Stations shall display an appropriate EBS slide and transmit all following announcements visually and aurally in the manner required by Section 73.1250(h) of the FCC Rules. Foreign language stations report all announcements in foreign language prior to English.)

> "We interrupt this program; this is a national emergency. Important instructions will follow."

4. TRANSMIT ATTENTION SIGNAL.

Broadcast the two-tone attention signal (See Sections 73.906 and 73.940 of the Rules) for 20 to 25 seconds.

Note: Noncommercial educational FM broadcast stations of 10 watts or less, and low power TV stations, are exempt from having to transmit the two-tone attention signal.

Primary Stations Only (Including PRI CPCS Stations)

5. BROADCAST ANNOUNCEMENT.

> "This is an Emergency Action Notification. All stations shall broadcast this Emergency Action Notification message. This station has interrupted its regular program at the request of the White House to participate in the Emergency Broadcast System. During this emergency, some stations will remain on the air broadcasting news and official information to the public in assigned areas. This is station (call letters). We will remain on the air to serve the (operational area name) area. *If you are not in this area, you should tune to other stations until you hear one broadcasting news and

information for your area. You are listening to the Emergency Broadcast System serving the (operational area name) area. Do not use your telephone. The telephone lines should be kept open for emergency use. The Emergency Broadcast System has been activated to keep you informed. I repeat. . ." (Repeat Announcement)

*Stations designated primary relay or originating primary relay should include the following sentence at this point in this announcement: "We will also be serving as a program distribution and relay channel to other stations."

6. MONITOR FOLLOWING SOURCES FOR FURTHER INSTRUCTIONS AND BROADCAST EMERGENCY PROGRAMMING AS SOON AS AVAILABLE:

Sources

- Common program control station for operational area
- Radio-TV network lines
- Affiliates & nonaffiliates serviced by participating communications common carriers
- Primary relay station of state relay network
- Any other source that may be available

Priorities

Record lower priority programming for rebroadcast at earliest opportunity.

First: presidential messages—must be carried "live"

Second: operation area (local) programming

Third: state programming

Fourth: national programming and information

7. USE THIS STANDBY SCRIPT UNTIL EMERGENCY PROGRAMMING AVAILABLE—LATER AS FILLER.

"We interrupt our program at the request of the White House. This is the Emergency Broadcast System. All normal broadcasting has been discontinued during this emergency. This is station (call letters). This station will continue to broadcast, furnishing news, official information and instructions, as soon as possible, for the (operational area name) area. If you are not in the (operational area name) area, tune to a station furnishing information for your area. I repeat. . ." (Repeat as needed.)

8. MONITOR FOR EMERGENCY ACTION TERMINATION.

Same methods as for notification—Upon receipt, proceed to termination procedures.

Termination Procedures: All Stations

1. RECEIVE EMERGENCY ACTION TERMINATION.

(Same methods as for notification)

EMERGENCY ACTION TERMINATION MESSAGE—AP/UPI SUBSCRIBERS/NETWORK AFFILIATES ONLY

"This is an Emergency Action Termination. The authenticator words for this termination are (—). All stations follow the EBS Checklist for termination procedures. The authenticator words are (—)."

2. AUTHENTICATE TERMINATION.

Compare authenticator words in termination with words on current authenticator list. Do not initiate termination if words do not match. *(Red envelope contained in pocket on inside front cover.)*

3. BROADCAST ANNOUNCEMENT.

"This concludes operations under the Emergency Broadcast System. All broadcast stations may now resume normal broadcast operations." (Repeat announcement.)

4. RESUME NORMAL PROGRAMMING.

(In accordance with regular station authorization.)

National Level Tests

National interconnecting arrangements and facilities (networks, key stations, AP/UPI, AT&T) will be tested periodically. Procedures for tests which affect all stations are described below.

Closed Circuit Tests (Radio Network Affiliates and AP/UPI Subscribers)

DO NOT INTERRUPT PROGRAM—DO NOT BROADCAST TEST MESSAGE

Notification Methods

- RADIO NETWORKS: Affiliates only—Preceded by network alerting signal.
- AP and UPI TELETYPE—Preceded and followed by a line of "X"s and an alarm.

CLOSED CIRCUIT TEST ACTIVATION MESSAGE—AP/UPI

SUBSCRIBERS/RADIO NETWORK AFFILIATES ONLY

"This is notification of a closed circuit test of the EMERGENCY BROADCAST SYSTEM. The test program will begin at ______, Washington, DC, time. Radio stations do not broadcast this message and do not broadcast the audio program.

The test authenticator words are _____. This message authorizes a closed circuit test of the Emergency Broadcast System. Broadcast stations monitor radio network lines for closed circuit test program. All stations follow procedures in the EBS checklist for closed circuit test. The test authenticator words are _____."

Action by Station

1. MONITOR RADIO NETWORK FOR TEST PROGRAM.
2. CHECK AP AND UPI TELETYPE.
3. AUTHENTICATE TEST MESSAGE.

Compare authenticator words with test words printed on outside of EBS Authenticator List Envelope (red envelope). If words do not match, take no further action.

4. RECORD TIME TEST RECEIVED IN STATION LOG OR RECORDS.

Termination Methods

- RADIO NETWORKS: Affiliates only—receive following aural Closing Cue: "*This concludes the closed circuit test of the* Emergency Broadcast System."
- AP and UPI TELETYPE: Preceded and followed by a line of "X"s and an alarm.

Closed Circuit Test Termination Message (AP and UPI Subscribers Only)

"This is an EBS Closed Circuit Test Termination. The authenticator words for this termination are (—). The closed circuit test was terminated at (date and time), Washington, DC, time. The authenticator words are (—)."

Action by Station:

1. RADIO NETWORK AFFILIATES MONITOR NETWORK FOR CLOSING CUE.
2. AP AND UPI SUBSCRIBERS CHECK TELETYPE FOR CLOSED CIRCUIT TEST TERMINATION MESSAGE.
3. AUTHENTICATE TEST TERMINATION MESSAGE.

Compare authenticator words with test words printed on outside of EBS authenticator list envelope (red envelope). If words do not match, take no further action.

4. RECORD TIME TEST TERMINATION MESSAGE RECEIVED IN STATION LOG OR RECORDS.

Periodic AP and UPI Test Transmissions (AP and UPI Subscribers Only)

DO NOT INTERRUPT PROGRAM—DO NOT BROADCAST TEST MESSAGE

Notification Method

- AP and UPI TELETYPE—Preceded and followed by a line of "X"s and an alarm.

Periodic Teletype Test Message

"This is a test of the emergency action notification procedures. If this were not a test, you would receive an emergency action notification message containing authenticator words. This is a test of the emergency action notification procedures. All stations follow procedures in EBS checklist for periodic teletype tests."

Action by Station

1. Record time test received in station log or records.

(No authentication provided.)

Weekly Transmission Test of the Attention Signal and Test Script

ALL RADIO AND TV STATIONS MUST PERFORM THIS TEST A MINIMUM OF ONCE A WEEK AT RANDOM DAYS AND TIMES BETWEEN 8:30 A.M. AND LOCAL SUNSET UNLESS THEY HAVE PARTICIPATED IN ONE OF THE FOLLOWING ACTIVITIES DURING THE TEST WEEK PERIOD: (1) A STATE OR LOCAL EBS ACTIVATION OR (2) A COORDINATED STATE OR LOCAL EBS TEST.

Action by Station

1. DISCONTINUE NORMAL PROGRAMMING.
2. BROADCAST ANNOUNCEMENT.

(TV stations shall display an appropriate EBS slide and transmit all following announcements visually and aurally in the manner described by Section 73.1250 of the FCC Rules. Foreign language stations repeat all announcements in foreign language prior to English.)

"This is a test. This station (optional—insert stations call sign) is conducting a test of the Emergency Broadcast System. This is only a test."

3. TRANSMIT ATTENTION SIGNAL.

Broadcast the Two-Tone Attention Signal. (See Sections 73.906 and 73.940 of the Rules)

4. BROADCAST ANNOUNCEMENT.

"This is a test of the Emergency Broadcast System. The broadcasters of your area in voluntary cooperation with federal, state, and local authorities have developed this system to keep you informed in the event of an emergency. If this had been an actual emergency (optional—stations may mention the types of emergencies likely to occur in their area), the attention signal you just heard would have been followed by official information, news, or instructions. This station (optional—insert station call sign) serves the (operational area name) area. This concludes this test of the Emergency Broadcast System."

5. RESUME REGULAR PROGRAMMING.
6. RECORD TIME TEST CONDUCTED IN STATION LOG OR RECORDS.
7. RECORD TIME TEST RECEIVED VIA EBS MONITOR/RECEIVER.

Your EBS monitoring assignment is provided in the appropriate EBS state plan or EBS map book. You should receive a test or activation each week. If you do not, you are responsible for determining why and making appropriate notations in your records or logs and/or taking corrective action.

State and Local Level Instructions

These procedures may be amended or altered as set forth in state and/or local EBS plans.

1. ACTIVATION.

- STATE LEVEL—A request for activation may be directed to the originating primary relay station by the governor, his designated representative, the National Weather Service, the state civil defense or State Office of Emergency Services.
- LOCAL LEVEL—A request for activation may be directed to the Common Program Control Station (CPCS-1) by the Weather Service, local civil defense, local government or public safety officials.

Note: The EBS may be activated at the state or local level by any AM, FM, or TV broadcast station, at management's discretion, in connection with day-to-day emergency situations posing a threat to the safety of life and property. (See Section 73.935 of FCC Rules.)

2. AUTHENTICATION.

The originating primary relay station and/or the common program control station will authenticate request for activation according to the state or local EBS Plan. National weather service requests received via NOAA weather radio and/or weather wire do not need authentication.

3. IMPLEMENTATION.

A. Record emergency program material. (Optional.)

B. Broadcast the following announcement:

"We interrupt this program because of a (state/local) emergency. Important information will follow."

C. Transmit the EBS attention signal for 20 to 25 seconds.

D. Broadcast the following announcement:

"We interrupt this program to activate the (name of state or operational area) Emergency Broadcast System at the request of (activating official) at (time)."

E. Broadcast emergency program material from (A) above.

Note: TV stations participating in the state or local level EBS shall display an appropriate EBS slide and transmit all announcements visually and aurally in the manner required by Section 73.1250(h) of the FCC Rules. Foreign language stations may transmit emergency announcements in the foreign language prior to broadcasting the announcements in English.

4. TERMINATION.

A. Upon receipt of the termination notice from activating official, make the following announcement:

"This concludes operations under the (name of state of operational area) Emergency Broadcast System. All broadcast stations may now resume normal broadcast operations."

B. Record emergency operation in station operating or program log. Send brief summary to FCC. (Optional.)

State and Local Tests

Tests of implementing procedures developed at the state and local levels may be conducted on a day-to-day basis as indicated in state and local plans. Coordinated tests of EBS operational procedures for an entire state or operational area may be conducted in lieu of the weekly transmission tests of the attention signal and test script required by Section 73.961(c) of the Rules.

EMERGENCY BROADCAST SYSTEM (EBS) CHECKLIST FOR NONPARTICIPATING STATIONS

National Level Instructions

When EBS is activated by the White House, all broadcast stations must take the following actions. Stations should record all emergency broadcasts (including notification) and log all significant events in the event these records are required at a later date.

Activation Procedures—All Stations

1. RECEIVE EMERGENCY ACTION NOTIFICATION (EAN).

(This is the notice which activates EBS. Presidential messages may be received almost immediately.)

Notification by one of the following methods is sufficient to commence action:

- AP and UPI TELETYPE: Preceded and followed by a line of "X"s and an alarm.
- RADIO-TV NETWORKS: Affiliates only. Preceded by network alerting signal. Continue to monitor for further instructions.

- EBS MONITOR RECEIVER: Preceded by two-tone attention signal. (Step 5 for message format.) Continue to monitor for further instructions. Your EBS Monitoring assignment is specified in the state EBS operational plan or mapbook.

EMERGENCY ACTION NOTIFICATION MESSAGE—AP UPI SUBSCRIBERS/NETWORK AFFILIATES ONLY

> "This is an Emergency Action Notification requested by the White House. The authenticator words for this notification are (—). All stations follow procedures in the EBS Checklist for national level emergency. The President of the United States or his designated representative will shortly deliver a message over the Emergency Broadcast System. The authenticator words are (—)."

2. AUTHENTICATE NOTIFICATION—AP AND UPI SUBSCRIBERS/NETWORK AFFILIATES ONLY.

Compare authenticator words in notification with words on current EBS authenticator list. Take no further action if words do not match.

Different activation and termination words are provided for each date. Words are effective 12:01 a.m., Washington, DC, time.

3. DISCONTINUE NORMAL PROGRAMMING AND BROADCAST ANNOUNCEMENT.

(TV stations shall display an appropriate EBS slide and transmit all following announcements visually and aurally in the manner required by Section 73.1250(h) of the FCC Rules. Foreign language stations repeat all announcements in foreign language prior to English.)

> "We interrupt this program: this is a national emergency. Important instructions will follow."

4. TRANSMIT ATTENTION SIGNAL.

Broadcast the two-tone attention signal (See Sections 73.906 and 73.940 of the Rules).

Note: Noncommercial educational FM broadcast stations of 10 watts or less, and low power TV stations, are exempt from having to transmit the two-tone attention signal.

5. BROADCAST ANNOUNCEMENT.

> "This is an Emergency Action Notification. All stations shall broadcast this Emergency Action Notification message. This station has interrupted its regular program at the request of the White House to participate in the Emergency Broadcast System. During this emergency, some stations will remain on the air broadcasting news and official information to the public in assigned areas. This is station (call letters). We will be leaving the air. You should now tune to other stations until you hear one broadcasting news and information for your area. This station will not be broadcasting news and information for your area. Do not use your telephone. The telephone lines should be kept open for official use. The Emergency Broadcast System has been activated to keep you informed. I repeat. . ." (Repeat announcement.)

6. REMOVE CARRIER FROM AIR.
7. MONITOR FOR EMERGENCY ACTION TERMINATION.

Same methods as for notification. Upon receipt, proceed to termination procedures.

Termination Procedures: All Stations

1. RECEIVE EMERGENCY ACTION TERMINATION.

(Same methods as for notification.)

EMERGENCY ACTION TERMINATION MESSAGE—AP/UPI SUBSCRIBERS/NETWORK AFFILIATES ONLY

> "This is an Emergency Action Termination. The authenticator words for this termination are (—). All stations follow the EBS Checklist for termination procedures. The authenticator words are (—)."

2. AUTHENTICATE TERMINATION.

Compare authenticator words in termination with words on current authenticator list. Do not initiate termination if words do not match. (Red envelope contained in pocket on inside front cover.)

3. BROADCAST ANNOUNCEMENT.

> "This concludes operations under the Emergency Broadcast System. All broadcast stations may now resume normal broadcast operations." (Repeat announcement.)

4. RESUME NORMAL PROGRAMMING.

(In accordance with regular station authorization)

STATE AND LOCAL LEVEL INSTRUCTIONS

Note: All the above national level tests and state and local level instructions also apply to nonparticipating stations.

EBS Monitoring Requirements

All stations can obtain their monitoring assignments from the FCC.

Over-The-Air EBS Station Monitoring

In order to insure the effectiveness of the EBS, all licensees must have installed, and, in operating condition, equipment capable of transmitting and receiving the attention signal. The receiving equipment must be maintained in operating condition, including arrangements for human listening or automatic alarm devices and shall be installed at the designated transmitter control point and/or studio location in such a way that it enables the broadcast station staff, at normal duty locations, to be alerted instantaneously upon the receipt of the attention signal and to immediately monitor the emergency programming.

The attention signal consists of the simultaneous transmission of two audio tones for not less than 20 seconds nor more than 25 seconds. The characteristics of the two-tone signaling system are as follows:

Encoder

Function. To give an alert by demuting a monitoring receiver at the station receiving the signal. The monitor is continuously tuned to the sending station for EBS information.

Signal. Two simultaneous audio frequencies, 853 Hz and 960 Hz, each ±0.5 Hz.

Harmonic Distortion. Not to exceed 5% of each tone at encoder output.

Minimum level of modulation. Each tone must be capable of modulating the transmitter to not less than 40 percent, with all equipment ordinarily used in the audio line between the encoder and transmitter. To assure this, the specification further says that the output at each audio tone shall be at least +8 dBm into a 600 ohm load. The unit shall allow calibration of each tone separately.

Time period. On activation, the two tones shall be generated for not less than 20 seconds nor more than 25 seconds.

Operating temperature. All foregoing specifications to me maintained in ambient temperature 0°C to 50°C.

Humidity. All specifications to be maintained up to 95% relative humidity.

Supply voltage variation. Operation must be within tolerances with supply voltage from 85% to 115% of the rated value.

Testing conditions. Must maintain the frequency, distortion, and time period specifications in a minimum RF field of 10 V/m at a frequency in the AM broadcast band, and with a minimum RF field of 0.5 V/m in either the FM or TV frequency band.

Indicator device. A visual and/or aural indicator must show clearly that the device has been activated.

Switch guard. The activation switch must have protection which will prevent accidental operation; this must include remote control switches.

Decoder

The decoder must be activated only on *simultaneous detection* of the two audio tones, 853 Hz and 960 Hz. This simultaneous reception must demute the monitoring receiver. The additional capability of activating an external alarm is *not required*, but has obvious value.

To prevent a false response, the decoder must have a time delay not less than eight seconds, and not more than 16 seconds, and a bandwidth such that there is no response to tones that vary more than ±5 Hz from each of the frequencies, 853 Hz and 960 Hz.

The decoder must have a reset switch, for returning the receiver to a muted state after activation.

The decoder must maintain all the foregoing specifications in ambient temperature from 0°C to 50°C.

For the convenience of operating personnel, the EBS decoder is in most cases muted for normal periods of operation. Such devices are, of course, activated upon receipt of the alert signal.

Possible Shortening of EBS Attention Signal

In September of 1989, the NAB petitioned the FCC to revise various aspects of the EBS Rules. FCC Rulemaking Number 7188 was assigned. Included in the petition was a proposal to shorten the duration of the EBS attention signal from its present 20 to 25 seconds to as short as eight seconds. The rationale was a concern that the current 20-second minimum duration attention signal causes "audience tuneout." In October 1989, the FCC authorized a six-month field test of a shortened EBS attention signal in the Central Michigan EBS Operational Area. A detailed report of the field test results was presented to the FCC in the Fall of 1990. The FCC found the results favorable and issued a Notice of Proposed Rulemaking for a shortened attention signal in October 1991.

It should be noted that the central Michigan field tests authorized by the FCC included permission to make circuit modifications to type accepted encoders and to certified decoders. Until such time as the FCC acts on the NAB petition, stations are cautioned that no modifications to existing EBS encoders or decoders should be made.

EBS Compliance and Satellite-Based Remote Control Systems

A service now being offered to broadcasters is a satellite-based remote control system, where the operator on duty is located at a 24-hour central facility which may be thousands of miles from the client station. The interconnect from each client station to the central monitoring location is accomplished by small-aperture Ku-band satellite uplinks. As long as such systems include the capability of relaying real-time audio from each client station, so the operator on duty at the central facility can monitor the audio of an unsquelched EBS receiver and determine the nature of the incoming message, compliance with the EBS rules would appear to be possible. However, any system which relies solely on the availability of the conventional switched telephone system to relay audio between the locally maintained EBS receiver and the central monitoring facility may not be in compliance; as such, a system assumes the availability of the switched telephone network at the very time when such service is most at risk. For example, a weather emergency such as a tornado could interrupt local telephone service at the very time the central monitoring facility has detected the unsquelching of the local EBS receiver and is trying to establish a telephone interconnect to ascertain the nature of the incoming EBS message.

Time is of the essence in such situations; a delayed tornado alert may be little better than no alert. Even where local telephone lines are not physically damaged, incoming calls may not be possible if the local telephone company has given outgoing calls priority for the duration of the local emergency, or if the number of calls simply overwhelms the capability of the switching network. This is exactly what happened during the October 1989 San Francisco earthquake. Many broadcasters found their dial-up remote control circuits useless because of the unavailability of a dial tone.

FCC Amended EBS Rules

In February 1989, the FCC amended Section 73.932 of the EBS Rules to make it clear that an entry must be made in the station log regarding failure to receive weekly EBS tests, regardless of the reason. The FCC Rules have long required that a broadcast station ascertain the reason for failure to receive a weekly EBS test but, until this rule change, only required a log entry if the investigation revealed the reason to be a defect in the station's own EBS equipment. The new rule, which became effective February 15, 1989, makes it clear that a log entry is required even if the reason for the nonreceipt of an EBS test was due to problems at the station sending the EBS test. Stations are cautioned that, under the 1895 memorandum between the FCC's Mass Media Bureau and the FCC's Field Operations Bureau, the lack of a covering log entry may result in a $300 Notice of Apparent Liability.

Broadcasting of Embellished EBS Tests Discouraged

Occasionally stations have embellished the wording or format specified in the FCC Rules and in the EBS checklist for weekly EBS tests by adding music, sound effects, singing, or a combination of all three. This is not a wise practice. The FCC issued a public notice in 1979 discouraging such activities. That notice is now codified in Section 73.4097 of the FCC Rules and as such is an enforceable FCC policy.

Use of Recorded Attention Signal Prohibited

The same FCC notice cited above also cautioned that the use of recorded EBS tones for the generation of the attention signal is not acceptable. The attention signal tones must always originate from an approved (type accepted) EBS encoder. The reason for this prohibition is that some audio cartridge machines have insufficient stability to ensure that the recorded tones are maintained within ±0.5 Hz of 853 and 960 Hz, as required by Section 73.940(a) of the FCC Rules. This prohibition only pertains to the recording of the 853 and 960 Hz tones as a source of the attention signal and is not intended to preclude prerecorded announcements for tests or alerts. Stations should, therefore, ensure that, when recording an incoming actual EBS alert for retransmission, the attention signal is obtained only from the station's local EBS encoder, and not from the attention signal which may have been recorded as part of the incoming EBS message.

Use of Alternative EBS Texts

It is now permissible to substitute station call letters everywhere the "this station" phrase is called for in the EBS checklist. It is also permissible to mention the types of possible emergencies likely to occur in a station's operational area. Other nonstandard EBS texts have been authorized by the FCC on a marketwide basis, upon specific request. However, stations should be careful not to use nonstandard wording (i.e., wording other than that appearing in the EBS checklist) without prior permission from the FCC.

EBS Authorization Must Be Posted

Participating stations are issued an EBS authorization, which becomes part of the station authorization and must be posted along with the station license. The EBS authorization is typed on an 8.5 × 11 inch document, FCC Form 392. The EBS authorization specifies the station call letters, location, and the authorization's effective date. All station personnel should know whether the station is a participating or a nonparticipating station. The correct EBS checklist (participating or nonparticipating) must be posted at the station control point.

Special Reception Techniques

In certain instances, either where a station is at a critical distance from another station or where receiving conditions are generally difficult by reason of directional operation, atmospherics, etc., extraordinary means are sometimes necessary to ensure the reception of the alert signal from another station. The Federal Communications Commission has provided information concerning the use of special receiving antennas which are applicable for such purposes. The application of this information, in many situations, will enable satisfactory reception of stations which ordinarily are completely unintelligible.

Shielding and Filtering

In cases where applicable, it is advisable to locate the antenna as far from a source of interference (or radio transmitter antenna) as possible and use coaxial or shielded transmission line to the receiver input terminals. Appropriate filtering and bypassing of power leads in order to eliminate the possible effect of high RF fields on the reception of the desired signal may be required. The antenna transmission line to the receiver may have to incorporate trap circuits to filter out the local transmitter radiation. For low-impedance receiver inputs, the circuit may be a parallel resonant trap connected in series with the inner conductor of the coaxial. For high-impedance receiver inputs, a series resonant trap may be connected directly across the input terminals, and tuned to the frequency of the local interfering station.

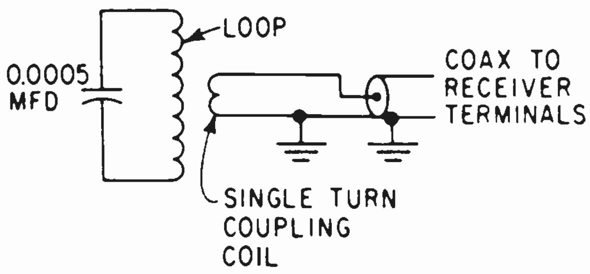


Figure 6. Simplified schematic of a loop antenna with construction details. ***Loop dimensions:*** **13 turns of No. 20 wire wound on a 12-in. square form, with ⅛ in. spacing between turns. The coupling coil is a single turn wound inside the loop and spaced approximately 1 in. from the loop. The mechanical mounting is optional, but for best results the loop should be mounted in the clear and away from metallic objects. Experiments with various turns in the coupling coil and using a coaxial cable from the loop to the receiver might prove beneficial. Shielding the loop is important in direction-finding work but is not important for ordinary reception.**

Loop Antenna

When the desired and undesired stations are on the same or very near the same frequencies, a directional antenna is often necessary. Such an antenna is also necessary in areas of weak signal strength from the desired station. A simple type of such antenna is the loop antenna, which is of maximum effectiveness only when the signals under consideration are within the ground-wave area. When this antenna is used, the null (obtained broadside to the plane of the loop) is oriented toward the undesired station. Maximum voltage is induced in the loop when the plane of the loop points toward the desired station. Since the antenna is bidirectional, a station on the back side may produce sufficient signal to cause interference, in which case a more elaborate directional receiving antenna may be required. (Fig. 6)

Beverage, or Wave, Antenna

A very effective receiving directional antenna is the beverage type. The theory of this antenna has been well covered in texts for many years and is not included here. Among the desirable properties of the beverage, or the wave, antenna for the reception are:

1. It delivers a strong signal over the entire standard broadcast band.
2. When terminated, it is unidirectional.
3. The antenna is low in cost, has long life, and is usually easy and simple to erect provided the space is available.

NATIONAL ADVISORY COMMITTEES

Introduction

National Advisory Committees have been organized (Fig. 7) to advise and assist the Federal Communications Commission, and other appropriate authorities. The committees' functions are to study and submit recommendations for emergency communications policies, plans, systems, and procedures for all FCC licensed and regulated communications in order to provide continued emergency communications services under conditions of crisis or war. In addition, they consider the adaptation and use of the systems, arrangements, and interconnecting facilities set forth in approved Operational Plans on a voluntary, organized basis during national, state, and operational (local) situations posing a threat to the safety of life and property. Included also are those conditions constituting a state of public peril or disaster. Such use of these capabilities during emergency situations is in accordance with the Commission's emergency and preparedness responsibilities, as defined in Sections 1, 4(o), 301, 308(a), and 706 of the Communications Act of 1934, as amended, and Executive Order 12472.

Organization

The National Security and Emergency Preparation Advisory Committee

Chair
Vice Chair
Executive Secretary

The Emergency Broadcast System Advisory Committee

Chair
Vice Chair
Executive Secretary

Members of the committees are appointed for a term not exceeding two years by the Federal Communications Commission, subject to appropriate security clearance when warranted. Membership is restricted to officers and employees of nongovernment Federal Communications Commission licensees (communications industry), subject to formal waiver when it is deemed in the public interest, convenience, and necessity. ("Nongovernment," as used herein, excludes federal government but includes state and local government agencies not licensed by the Federal Communications Commission.) Since all appointees serve at the pleasure of the Commission, any appointment may be terminated without cause. Such termination will be effective upon receipt of written notification from the Commission.

The Executive Secretary serves as the official correspondent for the Advisory Committees.

Functions and Responsibilities

The committees are concerned with operational emergency communications policies, plans, systems, and procedures to fulfill stated requirements under a broad range of emergency contingencies posing a threat to the safety of life and property. The principal

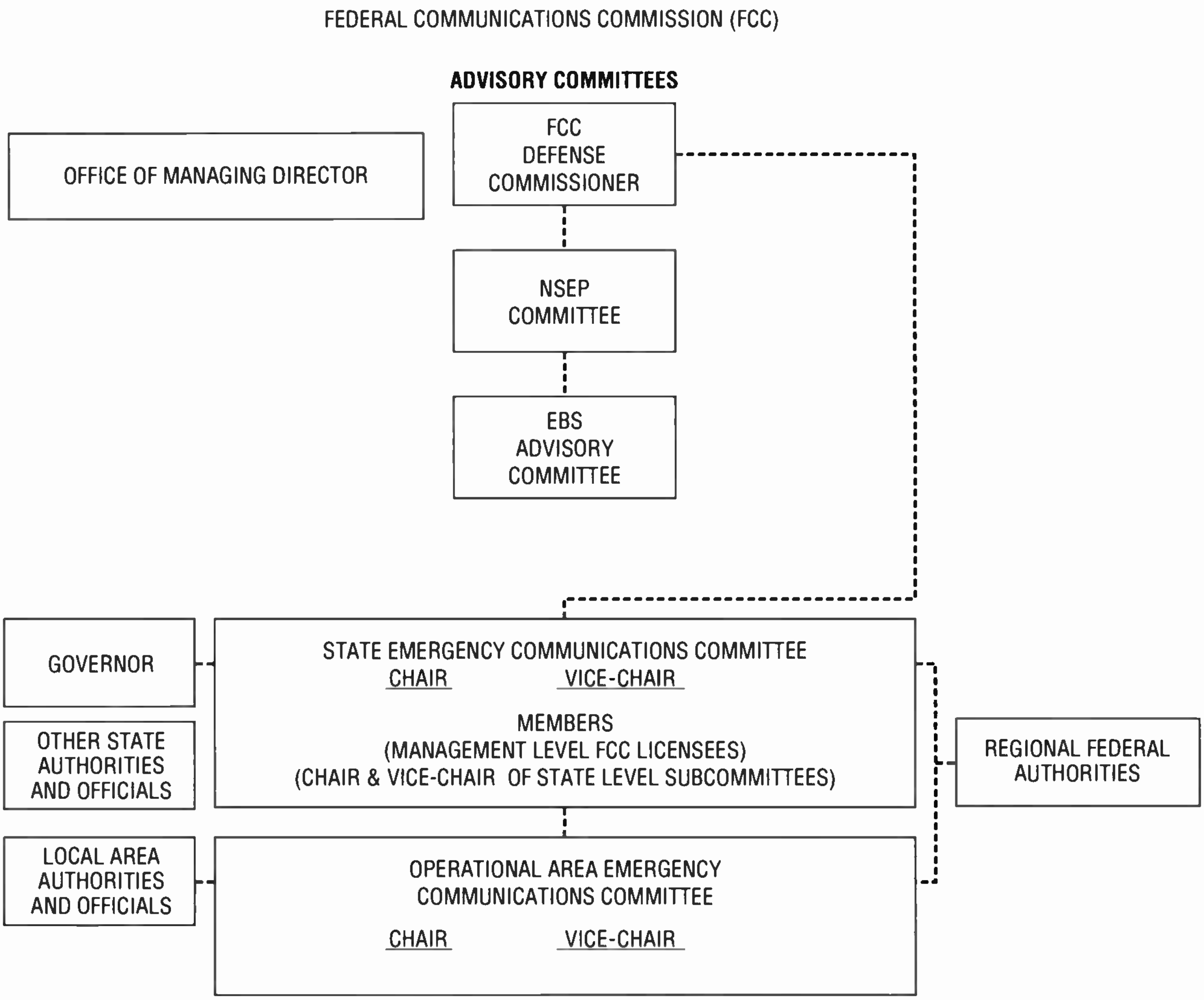


Figure 7. Industry Advisory Committee organization chart.

functions and responsibilities include but are not limited to:

1. Studying and submitting recommendations to the Federal Communications Commission concerning operational emergency communications.
2. Providing advice and recommendations through the Federal Communications Commission to appropriate federal, national, state and local authorities and organizations to enhance emergency communications operations.
3. Maintaining liaison with the nongovernment communications industry.
4. Maintaining liaison with all Working Groups, and ad hoc committees to coordinate and assist in the planning for the utilization of nongovernment communications facilities during emergencies.
5. Coordinating with the Federal Communications Commission in the establishment of authentication procedures for use during emergencies.
6. Advising the Federal Communications Commission concerning industry opinion relative to any proposed test or exercise of emergency communications systems, plans, and procedures. Also assisting, observing and evaluating the effectiveness of such activities.
7. Evaluating proposals for the development and use of operational emergency communications systems, plans, and procedures.
8. Encouraging studies and research directed towards the improvement of existing and development of new systems, plans, and policies which will improve the overall effectiveness of emergency communications.
9. Maintaining liaison with State Emergency Communications Committees.

Procedures

Detailed procedures with respect to operation, management, and functioning of the Committees, and working groups are published in separate documents and are available from the office of Managing Director, Executive Secretary, Industry Advisory Committees, Federal Communications Commission, Washington, DC, 20554.

REFERENCES

FCC Rules and Regulations, Subpart G, Emergency Broadcast System.

Broadcasting Emergency Information, FCC Rules Section 73.1250.

Detailed state EBS plans.

Operational (local) area EBS plans.

EBS checklists.

7.9 AC Power Systems

Jerry Whitaker
Technical Writer, Beaverton, Oregon

INTRODUCTION

Every electronic installation requires a steady supply of clean power in order to function properly. Recent advances in technology have made the question of ac power quality even more important as microcomputers are integrated into a wide variety of electronic products. The high-speed logic systems prevalent today can garble or lose data because of power supply disturbances or interruptions.

The ac powerline into a facility is the lifeblood of any operation. It is also, however, a frequent source of equipment malfunctions and component failures. The utility company ac feed contains not only the 60 Hz power needed to run the facility, but also a variety of voltage sags, surges, and transients. These abnormalities cause different problems for different types of equipment.

UTILITY POWER SYSTEM

The details of power distribution in the United States vary from one utility company to another, but the basics are the same. (See Fig. 1.) Power from a generating station or distribution grid enters an area substation at 115 kV or higher. The substation consists of switching systems, step-down transformers, fuses, circuit breakers, reclosers, monitors, and control equipment. The substation delivers output voltages of approximately 60 kV to subtransmission circuits, which feed distribution substations. The substations convert the energy to approximately 12 kV and provide voltage regulation and switching provisions that permit patching around a problem. The 12 kV lines power the pole- and surface-mounted transformers, which supply various voltages (generally 208–240 V 3-phase) to individual loads.

Fuses and circuit breakers are included at a number of points in the 12 kV distribution system to minimize fault-caused interruptions of service. Ground-fault interrupters are also included at various points in the 12 kV system to open the circuit if excessive ground currents begin to flow on the monitored line. Reclosers may be included as part of over-current protection of the 12 kV lines. The recloser will open the circuit if excessive currents are detected, and reclose after a preset length of time. The recloser will perform this trip-off reset action several times before being locked out.

In some areas, the actions of circuit breakers, pole-mounted switches and reclosers are controlled by two-way radio systems that allow status interrogation and switching of remotely located devices from a control center. Some utilities use this method sparingly, but others make extensive use of it.

Depending on the geographic location, varying levels of lightning protection are included as part of the ac power system design. Most service drop transformers (12 kV to 208 V) have integral lightning arresters. In areas of severe lighting, a shield wire is strung between the top insulators of each pole, attracting lightning to the grounded wire, and not the hot leads.

Power Factor

Capacitor banks are placed at various locations in the 12 kV distribution system. The number and location are determined by the load distribution and power factor of the circuit. The capacitors improve the short-term line voltage regulation (in the millisecond range) and reduce transient activity on the line. Surges are reduced because the capacitor presents a high impedance to the 60 Hz line voltage frequency, and a low impedance to a high frequency transient. The capacitors are placed on the line to keep the power factor as close to unity as possible. Transient suppression is simply a by-product.

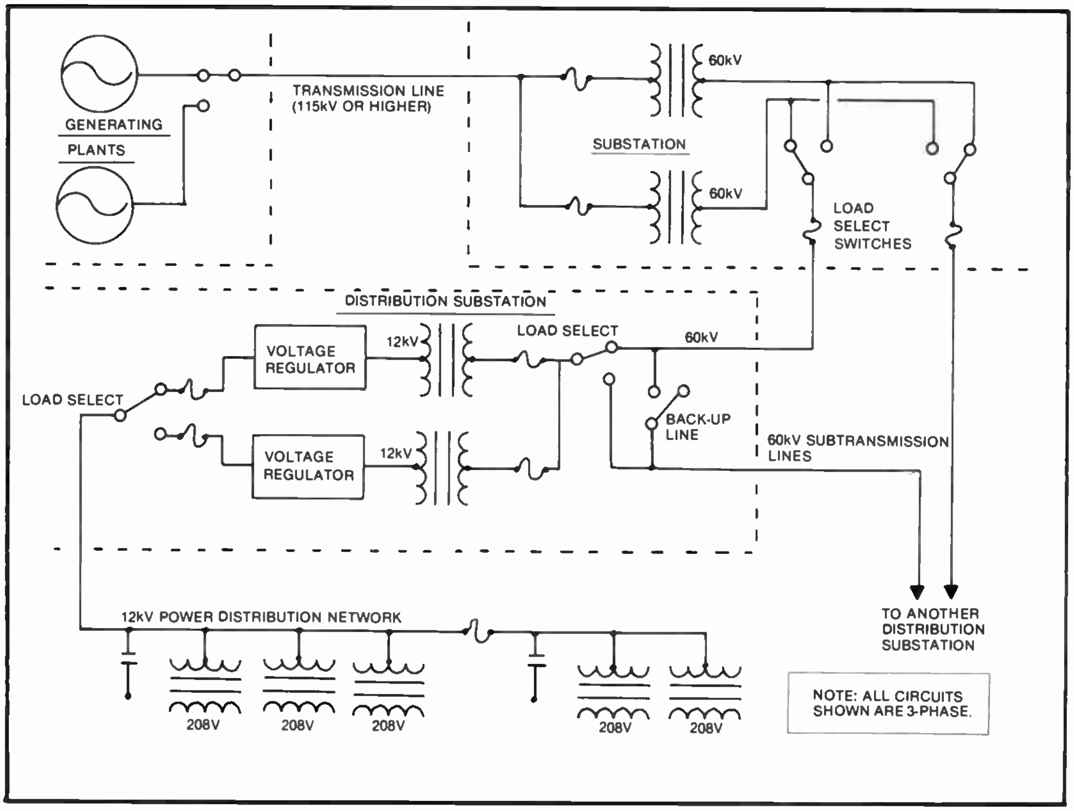
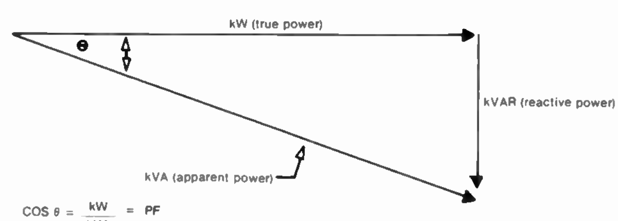


Figure 1. Simplified block diagram of a basic utility company power distribution system. The devices shown as fuses could be circuit breakers or reclosers, which function as automatic-resetting circuit breakers. All circuits shown are 3-phase. The capacitors perform power factor correction duty.

Power factor (PF) is defined as the ratio of true power to apparent power, generally expressed as a percentage. Reactive loads (inductive or capacitive) act on power systems to shift the current out of phase with the voltage. The cosine of the resulting angle between the current and voltage is the power factor.

A utility line that is looking into an inductive load (which is most often the case) is said to have a lagging power factor, while a line feeding a capacitive load has a leading power factor. (See Fig. 2.) A poor power factor will result in excessive losses along utility company feeder lines because more current is required to drive a load with a low power factor than the same load with a power factor close to unity (100%). For example, a motor requiring 5 kW from the line is connected to the utility service entrance. If it has a power factor of 86%, the apparent power demanded by the load will be 5 kW divided by 0.86, or more than 5.8 kW. The true power is 5 kW, and the apparent power is 5.8 kW. The same amount of work is being done by the motor, but the closer the power factor is to unity, the more efficient the system will be.

To keep the power factor as close as possible to 100%, utility companies place capacitor banks at various locations in the 12 kV distribution system, offsetting the inductive loading (lagging power factor) of most user equipment. The idea is to create an equal amount of leading PF in the system to match the load's lagging PF. When balanced, the power factor is 100%. In practice, this is seldom attainable because loads are switched on and off at random times, but utilities



$\cos\theta = \dfrac{kW}{kVA} = PF$

$\sin\theta = \dfrac{kVAR}{kVA}$

θ = the phase angle, a measure of the net amount of inductive reactance in the circuit.
kW = the true power that performs the "real work" done by the electrical circuit (measured in kilowatts).
kVA = the apparent power drawn by a reactive load (measured in kilowatts).
kVAR = the kilovolt-amperes-reactive component of an inductive circuit. The kVAR component (also known as the *phantom power*) provides the magnetizing force necessary for operation of inductive loads.

Figure 2. The mathematical relationships of an inductive circuit as they apply to power factor (PF) measurements. Reducing the kVAR component of the circuit causes θ to diminish, improving the PF. When kW is equal to kVA, the phase angle is zero and the power factor is unity (100%).

routinely, through much effort, maintain a power factor of approximately 99%. To accomplish this, capacitor banks are switched automatically to compensate for changing load conditions.

Utility Company Interfacing

Most utility company connections are the standard delta-wye type shown in Fig. 3. This transformer arrangement is usually connected with the delta side facing the high voltage and the wye side facing the load. This configuration provides good isolation of the load from the utility and somewhat retards the transmission of transients from the primary to the secondary. The individual 3-phase loads are denoted by Z-1, Z-2, and Z-3. They carry load currents as shown.

When using a wye-connected system, it is important that the building's neutral lead be connected to the midpoint of the transformer windings, as shown in Fig. 3. The neutral line provides a path for the removal of harmonic currents that may be generated in the system because of rectification of the secondary voltages.

In some parts of the country, an open-delta (or V-V) arrangement is used by the utility company to supply multi-phase power to customers. (See Fig. 4.) Users often encounter problems when operating sensitive 3-phase loads from such a connection because of the system's poor voltage regulation characteristics during varying load conditions. The open-delta configuration is also subject to high third harmonic content and transient propagation. The three loads and their respective load currents are shown in the diagram.

Other primary power connection arrangements are possible, such as wye-to-wye or delta-to-delta. Like the delta-to-wye configuration, they are not susceptible to the problems that can be experienced with the open-delta, or V-V service.

The open-delta system can develop a considerable imbalance between the individual phases in either

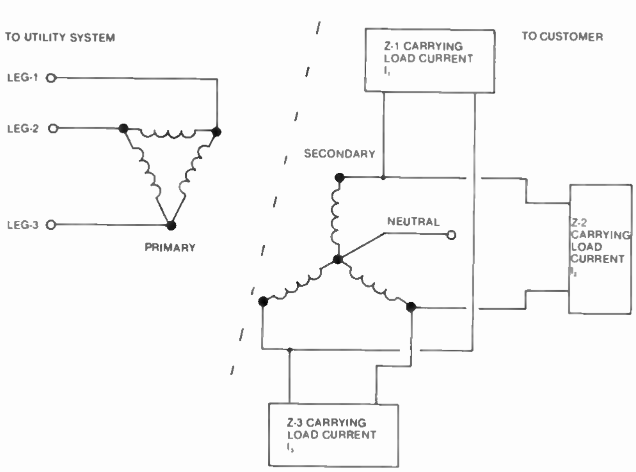


Figure 3. The delta-wye transformer configuration for utility company power distribution. This common type of service connection transformer provides good isolation of the load from the 12 kV distribution system line.

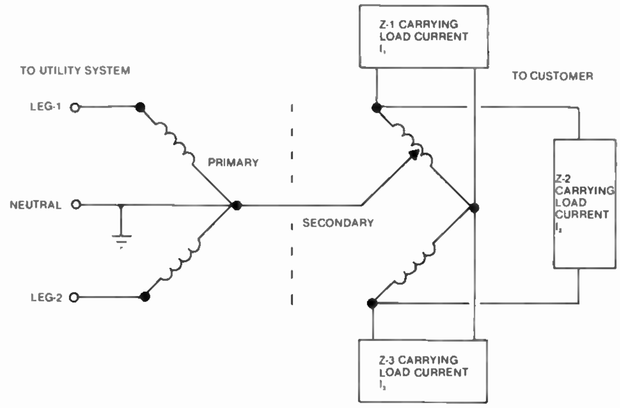


Figure 4. The open-delta (or V-V) utility company service connection transformer. Use of this configuration is not recommended because of its poor voltage regulation under varying load, high third harmonic content and transient disturbance propagation.

voltage or phase or both. Such an occurrence can introduce a strong 120 Hz ripple frequency in 3-phase power supplies, which are designed to filter out a 360 Hz ripple. The effects of this 120 Hz ripple can be increased noise in the supply and possible damage to protection devices across power supply chokes. Depending on the loading of an open-delta transformer, high third harmonic energy can be transferred to the load, producing transients of up to 300% of the normal voltage, which severely strain rectifiers, capacitors, and inductors in the power supply, as well as adding to the output noise of the supply.

The phase-to-phase voltage balance of a utility company line is important at most types of facilities, not only because of the increased power supply ripple it may cause, but also because of the heating effects that may result. Even simple 3-phase devices such as motors should be operated from a powerline that is well balanced, preferably within 1%. Studies have shown that a line imbalance of only 3.5% can produce a 25% increase in the heat generated by a 3-phase motor. A 5% imbalance can cause a 50% increase in heat, which is potentially destructive. Similar heating can also occur in the windings of 3-phase power transformers used in transmission equipment. Phase-to-phase voltage balance can be accurately measured over a period of several days with a slow-speed chart recorder. Phase-to-phase imbalance is generally caused by large single-phase power users on the 12 kV distribution line. Uneven currents through the utility company distribution system result in uneven line-to-line voltages at the customer's service drop entrance.

TRANSIENT DISTURBANCES

Transient overvoltages come in a wide variety of forms, from a wide variety of sources. They can, however, be broken down into two basic categories: those generated through natural occurrences and those

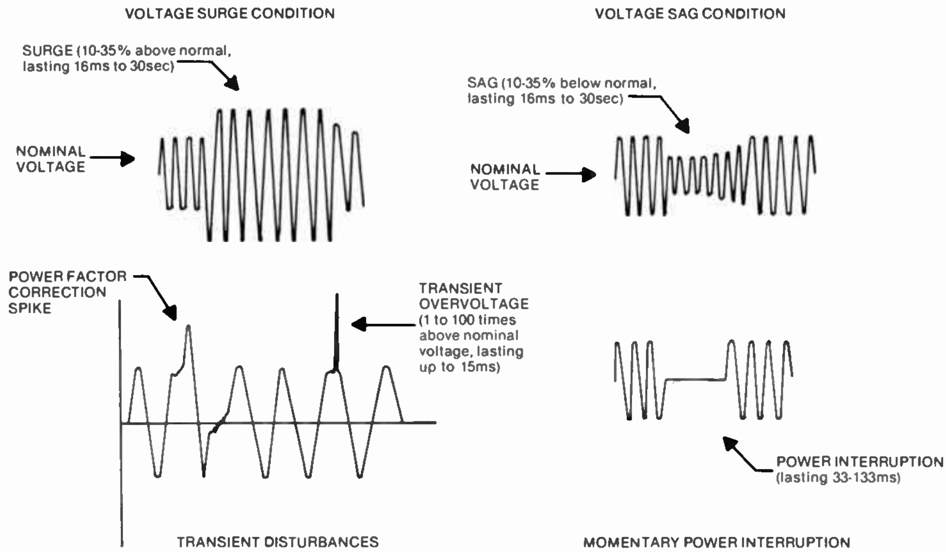


Figure 5. The four basic classifications of short-term power-line disturbances.

generated through the use of equipment, either on-site or elsewhere. Transient disturbances are what headaches are made of. Spikes, surges, or power bumps can take your equipment down and leave a complicated and expensive repair job. Ensuring that the equipment at a facility receives clean ac power has always been important. But now, with microcomputers being integrated into a wide variety of electronic products, the question of ac power quality is more critical than ever. The high-speed logic systems prevalent today can garble or lose data because of power supply disturbances or interruptions.

Classifications

Fig. 5 shows the four major classifications of short-term ac voltage disturbances. The generally accepted definitions for these disturbances are:

1. Voltage surge. An increase of 10% to 35% above the normal line voltage for a period of 16 ms to 30 sec.
2. Voltage sag. A decrease of 10% to 35% below the normal line voltage for a period of 16 ms to 30 sec.
3. Transient disturbance. A voltage pulse of high energy and short duration impressed upon the ac waveform. The overvoltage pulse may be one to 100 times the normal ac potential and may last up to 15 ms. Rise times can measure in the nanosecond range.
4. Momentary power interruption. A decrease to zero voltage of the ac powerline potential, lasting from 33 ms to 133 ms. (Longer-duration interruptions are considered power outages.)

Voltage surges and sags occasionally result in operation problems for equipment on-line, but automatic protection or correction circuits generally take appropriate actions to ensure that there is no equipment damage. Such disturbances can, however, garble computer system data if the disturbance transition time (the rise or fall time of the disturbance) is sufficiently fast. System hardware also may be stressed if there is only a marginal power supply reserve or if the disturbances are frequent.

Momentary power interruptions can cause a loss of volatile memory in computer-driven systems and place severe stress on hardware components, especially if the ac supply is allowed to surge back automatically without soft start provisions. Successful system reset may not be accomplished if the interruption is sufficiently brief.

Although voltage sags, surges, and momentary interruptions can cause operational problems for equipment used today, the possibility of complete system failure because of one of these mechanisms is relatively small. The greatest threat to the proper operation of electronic equipment rests with transient overvoltage disturbances on the ac line. Transients are difficult to identify and difficult to eliminate. Many devices commonly used to correct sag and surge conditions, such as ferroresonant transformers or motor-driven autotransformers, are of limited value in protecting a load from high-energy, fast rise-time disturbances on the ac line.

In the computer industry, the majority of unexplained problems resulting in disallowed states of operation actually are caused by transient disturbances on the utility feed. With the increased use of microcomputers in the broadcast industry, this warning cannot be ignored. Because of the high potential that transient disturbances typically exhibit, they not only cause data and program errors, but also can damage or destroy electrical components. This threat to electronic equipment involves sensitive integrated circuits and many other common devices, such as capacitors, transformers, rectifiers, and power semiconductors. Fig. 6

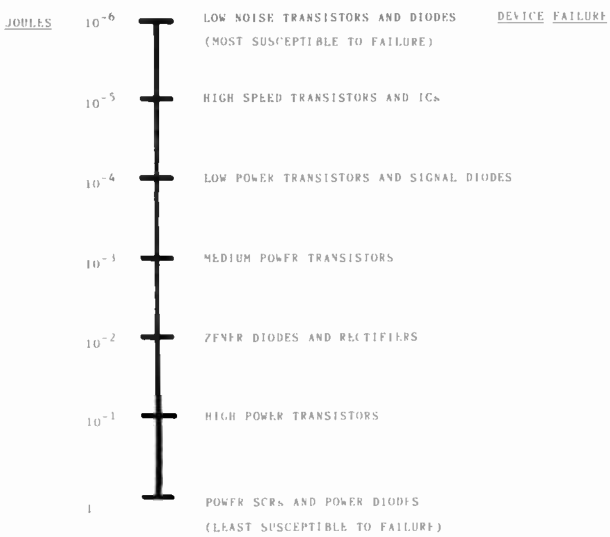


Figure 6. Estimate of the susceptibility of component failure as a result of transient overvoltages.

illustrates the vulnerability of common components to high-energy pulses. The effects of transient disturbances on electronic devices are often cumulative, resulting in gradual deterioration and, ultimately, catastrophic failure.

Standards of Measurement

It is difficult to assess the threat posed by transient disturbances without a guideline on the nature of spikes in ac power systems. To this end, a working group of the Institute of Electrical and Electronic Engineers (IEEE) has suggested two waveforms, one unidirectional and the other oscillatory, for measuring and testing transient suppression components and systems in ac power circuits with rated voltages of up to 277 V line-to-ground. The guidelines also recommend specific source impedance or short-circuit current values for transient analysis.

The voltage and current amplitudes, waveshapes, and source impedance values suggested in the IEEE guide (ANSI/IEEE standard C62.41–1980) are designed to approximate the vast majority of high-level transient disturbances. They are not, however, intended to represent worst case conditions: a difficult parameter to predict. The timing of a transient overvoltage with respect to the powerline waveform is also an important parameter in the examination of ac disturbances. Certain types of semiconductors exhibit failure modes that are dependent on the position of a transient on the 60 Hz ac system sine wave.

Fig. 7 shows the ANSI/IEEE representative waveform for an indoor-type spike (for 120 V to 240 V ac systems). Field measurements, laboratory observations and theoretical calculations have shown that the majority of transient disturbances in low-voltage indoor ac power systems have oscillatory waveshapes, instead of the unidirectional wave most often thought to represent a transient overvoltage. The oscillatory nature of the indoor transient waveform is caused by the natural resonant frequencies of the ac distribution system. Studies by the IEEE show that the oscillatory frequency range of such disturbances extends from 30 Hz to 100 kHz, and that the waveform changes depending upon where it is viewed in the power distribution system. The waveform shown in Fig. 7 is the result of extensive study by the IEEE and other independent organizations of various ac power circuits. The representative waveshape for 120 V and 240 V systems is described as a 0.5 μs/100 kHz ring wave. This standard indoor spike has a rise time of 0.5 μs and then decays while oscillating at 100 kHz. The amplitude of each peak is approximately 60% of the preceding peak.

Fig. 8 shows the ANSI/IEEE representative waveform for an outdoor-type spike. The classic lightning overvoltage pulse has been established as a 1.2 μs rise/50 μs decay waveshape for a voltage wave and a 8 μs rise/20 μs decay waveshape for a current wave. Accordingly, the ANSI/IEEE standard waveshape is defined as 1.2/50 μs open-circuit voltage (voltage applied to a high-impedance device), and 8/20 μs discharge current (current in a low-impedance device). The outdoor waveshapes, while useful in the analysis of components and systems, are not meant to represent all transient patterns seen in low-voltage ac circuits. Lightning discharges can cause oscillations, reflections, and other disturbances in the utility company power system that can appear at the service drop entrance as decaying oscillations.

Power System Protection Alternatives

Most utility companies make a good faith attempt to deliver clean, well-regulated power to their customers. Most disturbances on the ac line are beyond the control of the utility company. Large load changes imposed by customers on a random basis, power-factor correction switching, lightning, and accident-related system faults all combine to produce an environment in which tight control over ac power quality is difficult to maintain.

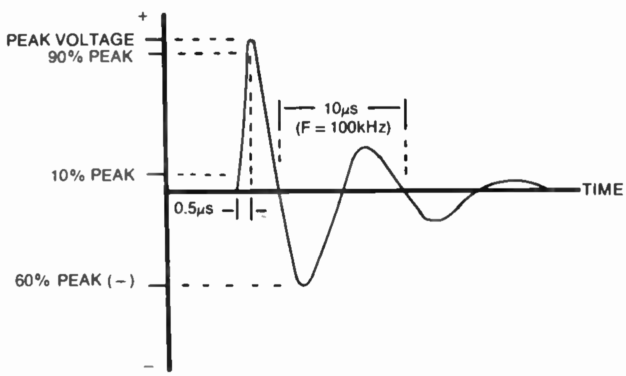


Figure 7. The suggested IEEE indoor-type transient overvoltage test waveform (0.5 μs/100 kHz ring wave, open-circuit voltage).

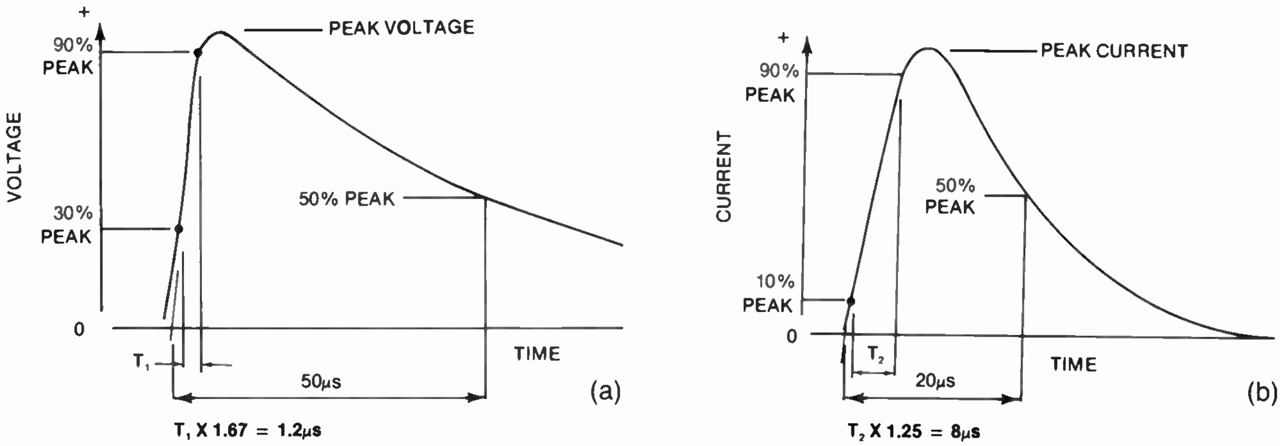


Figure 8. The unidirectional waveshape for outdoor-type transient overvoltage test analysis based on ANSI Standard C62.1: (a) the open-circuit waveform; (b) the discharge current waveform.

Therefore, the responsibility for ensuring ac power quality must rest with the users of sensitive equipment. The selection of a protection method for a given facility is as much an economic question as it is a technical one. A wide range of powerline conditioning and isolation equipment is available. A logical decision about how to proceed can be made only with accurate, documented data on the types of disturbances typically found on the ac power service to the facility. This data can be gained from a power quality survey, available from a number of consulting firms and power conditioning companies. The typical procedure involves installing a sophisticated voltage monitoring unit at the site for several weeks. During that time, data are collected on the types of disturbances the load equipment is likely to experience. The monitoring unit must be a high-speed system that stores disturbance data and delivers a printout on demand. Conventional analog chart recorders are too slow and lack sufficient sensitivity to accurately show short duration voltage disturbances. Chart recorders can confirm the presence of long-term surge and sag conditions but provide almost no useful data on transients.

The protection equipment chosen must be matched to the problems that exist on the line. Using inexpensive basic protectors may not be much better than operating directly from the ac line. Conversely, the use of a sophisticated protector designed to shield the plant from every conceivable power disturbance may not be economically justifiable. Purchasing the transient-suppression equipment is only one element in the selection equation. Consider the costs associated with site preparation, installation, and maintenance. Also consider the operating efficiency of the system. Protection units that are placed in series with the load consume a certain amount of power and, therefore, generate heat. These items may not be significant, but they should be taken into account. Prepare a complete life cycle cost analysis of the protection methods proposed. The study may reveal that lower long-term operating expense of one system outweighs the lower purchase price of another.

The amount of money a facility manager is willing to spend on protection from powerline disturbances generally depends on the engineering budget and how much the plant has to lose. Spending $50,000 on system-wide protection for a top-25 market television station is easily justified. At smaller operations, justification may not be so easy.

The susceptibility of electronic equipment to failure as a result of ac disturbances has been studied by many organizations. A bench mark study was conducted by the Naval Facilities Engineering Command (Washington, DC). The far-reaching program, directed from 1968 to 1978 by Lt. Thomas Key, identified three distinct categories of recurring disturbances on utility company power systems. As shown in Table 1, it is not the magnitude of the voltage, but the duration of the disturbance, that determines the classification. In the study, Key found that most equipment failures caused by ac line disturbances occurred during bad weather. According to a report on the findings, the incidence of thunderstorms in an area may help to predict failures. The type of power transmission system used by the utility company was also found to affect the number of disturbances observed on power company lines. For example, an analysis of utility system problems in Washington, DC, Norfolk, VA, and Charleston, SC, demonstrated that underground power distribution systems experienced one-third fewer failures than overhead lines in the same areas. Based on his research, Key developed the recommended voltage tolerance envelope shown in Fig. 9. The design goals illustrated are recommendations to computer manufacturers for implementation in new equipment.

Assessing the Lightning Hazard

As identified in the Naval Facilities study, the extent of lightning activity in an area significantly affects the probability of equipment failure caused by transient

TABLE 1
Types of voltage disturbances identified in the key report.

Parameter	Type 1	Type 2	Type 3
Definition	Transient and oscillatory overvoltage	Momentary undervoltage or overvoltage	Power outage
Causes	Lightning, power network switching, operation of other loads	Power system faults, large load changes, utility company equipment malfunctions	Power system faults, unacceptable load changes, utility equipment malfunctions
Threshold (Note 1)	200–400% of rated rms voltage or higher (peak instantaneous above or below rated rms)	Below 80–85% and above 110% of rated rms voltage	Below 80–85% of rated rms voltage
Duration	Transients 0.5–200 μs wide and oscillatory up to 16.7 ms at frequencies of 200 Hz to 5 kHz and higher	From 4–6 cycles, depending on the type of power system distribution equipment	From 2-60 sec if correction is automatic; from 15 min to 4 hrs if manual

Note 1: The approximate limits beyond which the disturbance is considered to be harmful to the load equipment.

activity. The threat of a lightning strike to a facility is determined in large part by the type of installation and its geographic location. The type and character of the lightning strike are also important factors. The keraunic number of a geographic location represents the likelihood of lightning activity in that area. Fig. 10 shows the isokeraunic map of the United States, which estimates the number of lightning days per year in various areas of the country. On the average, 30 storm days occur per year across the continental United States. This number does not fully describe the lightning threat because many individual lightning strikes occur during a single storm.

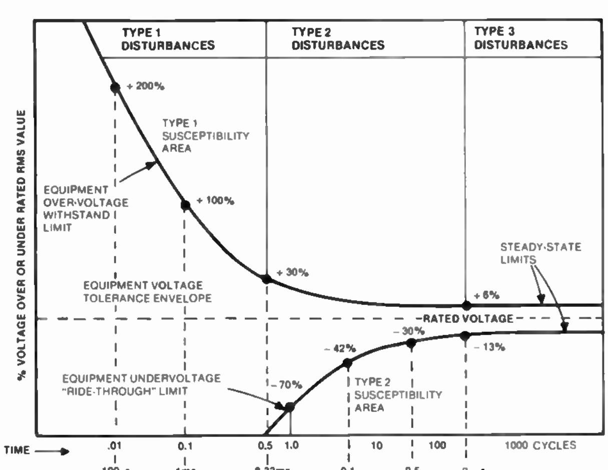


Figure 9. The recommended voltage tolerance envelope for computer equipment. This chart is based on pioneering work done by the Naval Facilities Engineering Command. The study identified how the magnitude *and* duration of a transient pulse must be considered in determining the damaging potential of a spike. The design goals illustrated in the chart are recommendations to computer manufacturers for implementation in new equipment.

The mechanical structure of a facility has a significant effect on the lightning threat to equipment operation. Higher structures tend to collect and even trigger localized lightning strikes. Because storm clouds tend to travel at specific heights above the earth, conductive structures in mountainous areas more readily experience lightning activity. The plant exposure factor is a function of the size of the facility and the isokeraunic rating of the area. The larger the physical size of an installation, the more likely it will be hit by lightning during a storm. The longer a transmission line (ac or RF), the more lightning strikes it will likely receive. The relative frequency of power problems is seasonal in nature, as shown in Fig. 11. Most problems are experienced during June, July, and August. These high problem rates can be traced to increased thunderstorm activity and heavy, unpredictable air-conditioning loads during the summer months.

Protection Using the Systems Approach

Equipment in a facility can be protected from transient disturbances in two basic ways: the systems

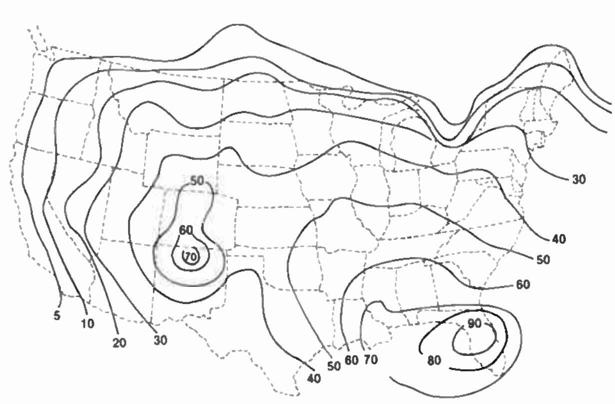


Figure 10. The isokeraunic map of the United States, showing the approximate number of lightning days per year.

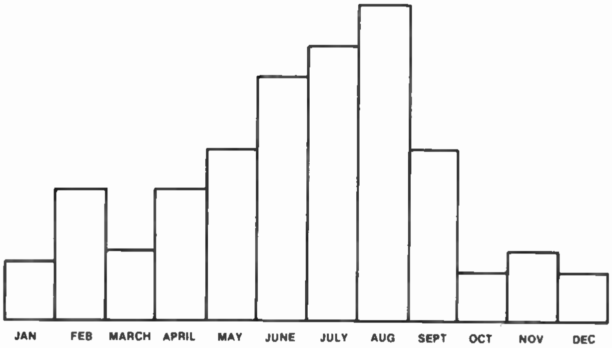


Figure 11. The relative frequency of power problems in the United States, classified by month.

approach or the discrete device approach. Table 2 outlines the major alternatives available for the systems approach to transient suppression.

Table 3 lists the relative benefits of each protection method. Information on lightning protection for the facility itself is covered in Chapter 2.2, "Lightning Protection for Broadcast Facilities" in this *Handbook*.

Uninterruptible Power System (UPS)

Fig. 12 shows a basic UPS system built around a rectifier-inverter combination. AC from the utility is rectified to continuously charge a bank of batteries, which powers an inverter. The closed-loop inverter provides both voltage and frequency regulation. The output of the inverter is generally a sine wave, or pseudo sine wave (a stepped square wave). If the utility voltage drops or disappears, current is drawn from the batteries. When ac power is restored, the batteries are recharged. Many UPS systems incorporate a standby diesel generator that starts as soon as the utility company feed is interrupted. With this arrangement, the batteries are called upon to supply the operating current for only 30 seconds or so, until the generator gets up to speed.

Motor-Generator Unit (MGU)

As the name implies, a motor-generator consists of a motor powered by the ac utility supply that is mechanically tied to a generator, which feeds the load. Transients on the utility line will have no effect on the load when this arrangement is used. Adding a flywheel to the motor-to-generator shaft will protect against brief power dips (up to ½ sec. on many models). Other features include output voltage and frequency regulation, ideal sine wave output, elimination of common-mode and transverse-mode noise, elimination of utility company power factor correction problems, and true 120 degrees phase shift for 3-phase models. The efficiency of a typical MGU ranges from 65% to 89%, depending on the size of the unit and the load.

High-Performance Isolation Transformer

Transients, as well as noise (RF and low-level spikes) can pass through transformers, not only by way of the magnetic lines of flux between the primary and the secondary, but through resistive and capacitive paths between the windings as well. Increasing the physical

TABLE 2

Types of systemwide protection equipment available to facility managers and the ac line abnormalities that each approach can handle.

System	Type 1	Type 2	Type 3
UPS System and Standby Generator	All source transients; no load transients	All	All
UPS System	All source transients; no load transients	All	All outages shorter than the battery supply discharge time
Secondary Spot Network (Note 1)	None	None	Most, depending on the type of outage
Secondary Selective Network (Note 2)	None	Most	Most, depending on the type of outage
Motor-Generator Set	All source transients; no load transients	Most	Only brown-out conditions
Shielded Isolation Transformer	Most source transients; no load transients	None	None
Suppressors, Filters, Lightning Arresters	Most transients	None	None
Solid-State Line-Voltage Regulator	Most source transients; no load transients	Some, depending on the response time of the system	Only brown-out conditions

Note 1: Dual-power feeder network.
Note 2: Dual-power feeder network using a static (solid-state) transfer switch.

TABLE 3

Approximate costs of systemwide protection equipment installation and operation.

System	Strong Points	Weak Points	Technical Profile
UPS System and Standby Generator	Full protection from power outage failures and transient disturbances; ideal for critical DP and life-safety loads	Hardware is expensive and may require special construction; electrically and mechanically complex; noise maybe a problem; high annual maintenance costs	Efficiency 80–90%; typical high impedance presented to the load may be a consideration; frequency stability good; harmonic distortion determined by UPS system design
UPS System	Completely eliminates transient disturbances; eliminates surge and sag conditions; provides power outage protection up to the limits of the battery supply; ideal for critical load applications	Hardware is expensive; depending on battery supply requirements, special construction may be required; noise may be a problem; periodic maintenance required	Efficiency 80–90%; typical high impedance presented to the load may be a consideration; frequency stability good; harmonic content determined by inverter type
Secondary Spot Network (Note 1)	Simple; inexpensive when available in a given area; protects against local power interruptions; no maintenance required by user	Not available in all locations; provides no protection from area-wide utility failures; provides no protection against transient disturbances or surge/sag conditions	Virtually no loss, 100% efficient; presents low impedance to the load; no effect on frequency or harmonic content
Secondary Selective Network (Note 2)	Same as above; provides faster transfer from one utility line to the other	Same as above	Same as above
Motor-Generator Set	Electrically simple; reliable power source; provides up to 0.5 sec power-fail ride-through; completely eliminates transient and surge/sag conditions	Mechanical system requires regular maintenance; noise may be a consideration; hardware is expensive; depending on m-g set design, power-fail ride-through may be less than typically quoted by manufacturer	Efficiency 80–90%; typical high impedance presented to the load may be a consideration; frequency stability may be a consideration, especially during momentary power-fail conditions; low harmonic content
Shielded Isolation Transformer	Electrically simple; provides protection against most types of transients and noise; moderate hardware cost; no maintenance required	Provides no protection from brown-out or outage conditions	No significant loss, essentially 100% efficient; presents low impedance to the load; no effect on frequency stability; usually low harmonic content
Suppressors, Filters, Lightning Arresters	Components inexpensive; units can be staged to provide transient protection exactly where needed in a plant; no periodic maintenance required	No protection from Type 2 or 3 disturbances; transient protection only as good as the installation job	No loss, 100% efficient; some units subject to power-follow conditions; no effect on impedance presented to the load; no effect on frequency or harmonic content
Solid-State Line Voltage Regulator	Moderate hardware cost; uses a combination of technologies to provide transient suppression and voltage regulation; no periodic maintenance required	No protection against power outage conditions; slow response time may be experienced with some designs	Efficiency 92–98%; most units present low impedance to the load; usually no effect on frequency; harmonic distortion content may be a consideration

Note 1: Dual-power feeder network.
Note 2: Dual-power feeder network using a static (solid-state) transfer switch.

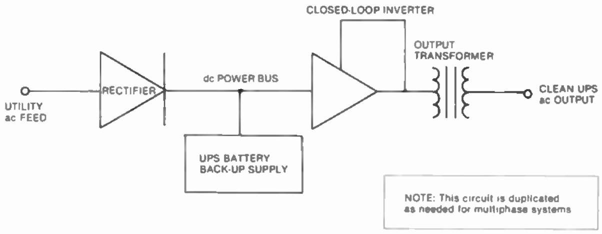


Figure 12. Block diagram of an uninterruptible power system (UPS) using ac rectification to float the battery supply. A closed-loop inverter draws on this supply and delivers clean ac power to the protected load.

separation between the primary and secondary windings will reduce the resistive and capacitive coupling. However, it will also reduce the inductive coupling and decrease the power transfer.

A better solution involves shielding the primary and secondary windings from each other to divert the primary noise current to ground. This approach leaves the inductive coupling basically unchanged. The concept can be carried a step further by placing the primary winding in a shielding box to shunt noise current to ground and reduce capacitive coupling between the windings. One application of this technology is shown in Fig. 13, in which transformer noise decoupling is taken a step further by placing the primary and secondary windings in their own wrapped foil box shields. The windings are separated physically as much as possible for the particular power rating, and placed between Faraday shields. This gives the transformer high noise attenuation from the primary to the secondary, as well as from secondary to the primary. The interwinding capacitance of a typical transformer using this technology is 0.005 pF. Common mode noise attenuation is generally in excess of 60 dB.

Solid-State Line Voltage Filter

A variety of proprietary techniques are used to implement solid-state line voltage filters. While the technology is not usually discussed by manufacturers, most rely on a combination of series and shunt elements constructed of semiconductor-based devices. The shunt element clamps voltage excursions, and the series element counteracts the iceberg effect of many transients. The iceberg effect is the negative voltage excursion that typically follows a positive voltage pulse. Performance varies greatly between types of units, as does cost. The insertion loss is usually small.

Regulation may also be accomplished through the use of a silicon-controlled rectifier (SCR) bank feeding an isolation transformer. A filter network within the system attenuates common-mode and transverse-mode noise. SCR-based systems are effective in compensating for voltage surge and sag conditions. They are relatively ineffective, however, in removing transient disturbances.

Selective or Secondary Spot Networks

AC power backup based on a selective or secondary spot network relies on the probability that most utility company failures are localized events that do not affect large geographic areas. The customer contracts with the local utility company to bring two separate service drops into the facility. The service drops are fed from different area distribution substations, and usually are routed via different physical paths. If power service fails on one line, an automatic transfer switch defaults to the secondary line. This approach does not protect a facility from transient disturbances or surge and sag conditions. It will, however, protect the facility from power outages that are localized in nature.

Discrete Transient Protection Devices

Discrete devices are less expensive and usually provide less protection compared with a sophisticated systems approach. It is unrealistic for a user to expect a group of discrete transient suppressors to do the job of a much more expensive systems design. However, properly applied discrete devices can prevent equipment damage from all but the most serious transient disturbances. The key to achieving this level of performance lies in understanding and properly applying discrete protection devices.

The performance of discrete transient suppression components available to engineers has greatly improved within the past ten years. The variety of reasonably priced devices now available makes it possible to exercise tight control over unwanted voltage excursions. Much of the credit for transient suppression work goes to the computer industry, which has been dealing with the problem for nearly three decades. Transient suppression hardware can be divided into three primary categories:

1. AC filters: The simplest type of ac power-line filter is a capacitor placed across the voltage source. The impedance of the capacitor forms a voltage divider with the impedance of the source, resulting in the attenuation of high frequency transients. This simple approach has definite limitations in spike-suppression capability and may introduce unwanted

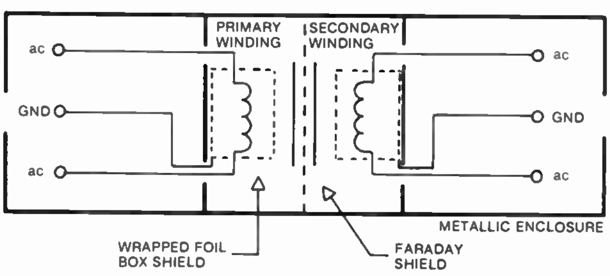


Figure 13. The shielding arrangement used in a high performance isolation transformer. The design goal of this unit is high common-mode and transverse-mode noise attenuation.

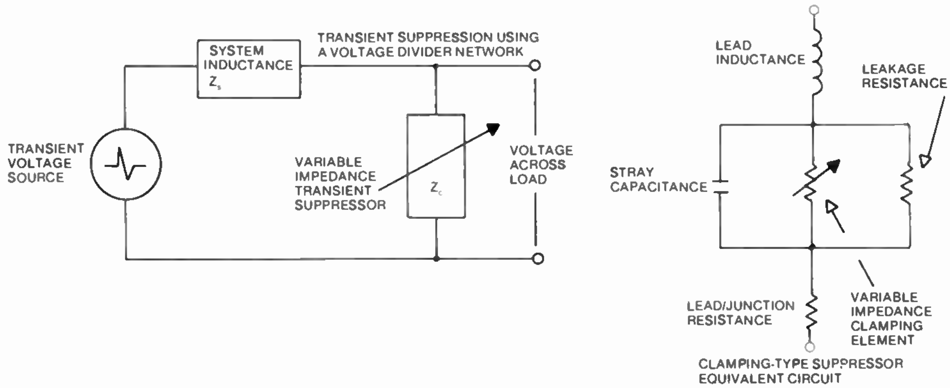


Figure 14. The mechanics of transient suppression using a voltage clamping device.

resonances with inductive components in the ac power distribution system. The addition of a series resistance will reduce the undesirable resonant effects, but it also will reduce the effectiveness of the capacitor in attenuating a transient disturbance.

2. Crowbar devices: Crowbar protection components include gas tubes (also known as spark gaps or gas gaps) and semiconductor-based active crowbar protection circuits. Although these devices and circuits can shunt a substantial amount of transient energy, they are subject to power-follow problems. Once a gas tube or active crowbar protection circuit has fired, the normal line voltage and the transient voltage are shunted to ground. This power-follow current may open protective fuses or circuit breakers if a method of extinguishing the crowbar clamp is not provided.
3. Voltage clamping components: Voltage clamping devices are not subject to the power-follow problems common in crowbar systems. Clamping devices include selenium cells, zener diodes, and varistors of various types. Zener diodes, using improved silicon rectifier technology, provide an effective voltage clamp for the protection of sensitive electronic circuitry. On the other hand, power dissipation for zener units is usually somewhat limited (compared with other suppression methods).

Selenium cells and varistors are different in construction, but act similarly on a circuit exposed to a transient overvoltage. Fig. 14 illustrates the variable nonlinear impedance exhibited by a voltage clamping device. The voltage divider network established by the source impedance (Z_s) and the clamping device impedance (Z_c) attenuates voltage excursions at the load. It should be understood that the transient suppressor depends upon the source impedance to aid the clamping effect. A protection device cannot be fully effective in a circuit that exhibits a low source impedance because the voltage divider ratio is reduced proportionately. A typical voltage-versus-current curve for a voltage clamping device is shown in Fig. 15. When the device is exposed to a high-voltage transient, the impedance of the component changes from a high standby value to a low-conductive value, clamping the voltage at a specified level.

Selecting a Protection Device

Selecting a transient suppression device for a particular application is a complicated procedure that must take into account the following factors:

1. The steady-state working voltage, including normal tolerances.
2. The transient energy to which the device is likely to be exposed.
3. The voltage clamping characteristics required in the application.
4. Circuit protection devices (such as fuses or circuit breakers) present in the system.
5. The consequences of protection device failure in a short-circuit mode.

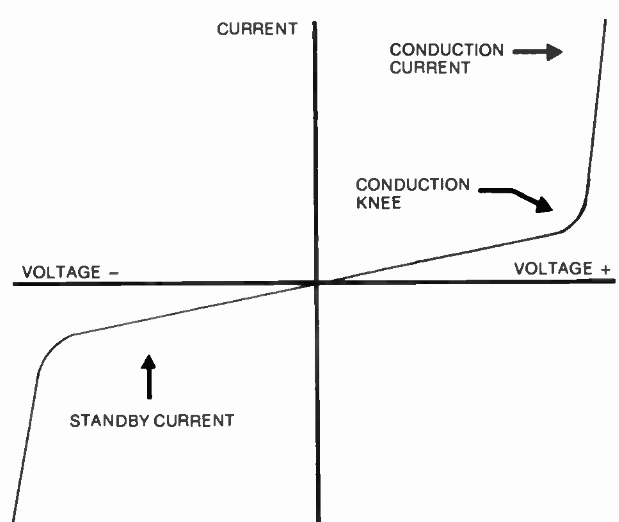


Figure 15. The voltage-versus-current curve for a bipolar voltage clamping device. The component is designed to be essentially invisible in the circuit until the applied positive or negative potential reaches or exceeds the *conduction knee* of the device.

6. The sensitivity of load equipment to transient disturbances.

Most transient suppression equipment manufacturers offer detailed application handbooks. Consult such reference data whenever planning to use a protection device. The specifications and ratings of suppression components are not necessarily interchangeable from one manufacturer to another. Carefully weigh the addition of transient suppression devices to a piece of equipment or ac power distribution system. Make allowances for operation of the circuit under all anticipated conditions.

STANDBY POWER SYSTEMS

When utility company power problems are discussed, most people immediately think of blackouts. The lights go out and everything stops. With the facility down and in the dark, there is nothing to do but sit and wait until the utility company finds the problem and repairs it. This generally takes only a few minutes. There are times, however, when it can take hours. In some remote locations, it can even take days. Blackouts are without a doubt the most troublesome utility company problem that a facility will have to deal with. Statistics show that power failures are, generally speaking, a rare occurrence in most areas of the country. They are also short in duration. Studies have shown that 50% of blackouts last six seconds or less, and another 35% are less than 11 minutes long. These failure rates are not usually cause for concern to commercial users, except where computer-based operations and broadcast stations are concerned.

A facility that is down for 11 minutes (or even five minutes) can suffer a significant loss of audience or data that can take hours (or perhaps days) to rebuild. Coupled with this threat is the possibility of extended power service loss due to severe storm conditions. Many broadcast and communications relay sites are located in remote, rural areas, or on mountaintops. Neither of these locations is well known for power reliability. It is not uncommon in mountainous areas

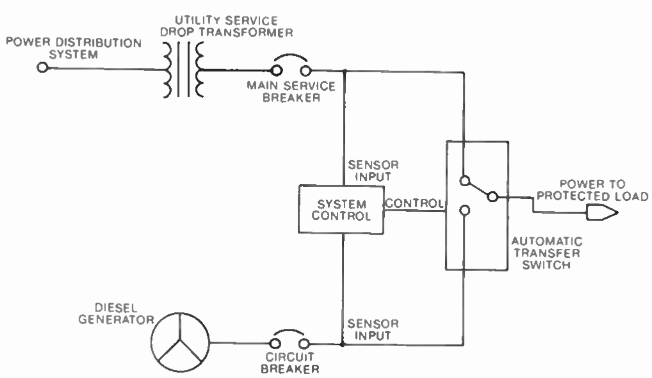


Figure 16. The classic standby power system using an engine-generator unit. This system protects a facility from prolonged utility company power failures.

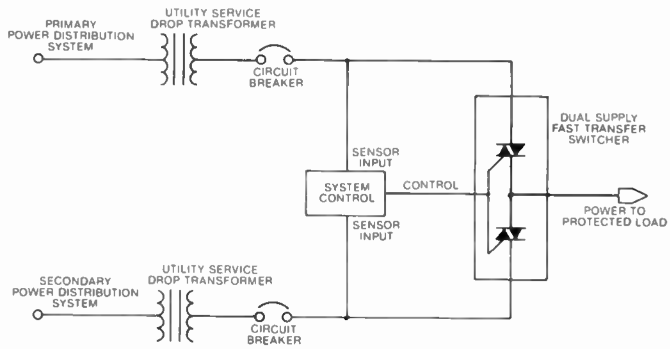


Figure 17. The dual feeder system of ac power loss protection. An automatic transfer switch changes the load from the main utility line to the standby line in the event of a power disruption.

for utility company service to be out for days after a major storm. Few operators are willing to take such risks with their business and choose to install standby power systems at appropriate points in the equipment chain. The cost of standby power for a facility can be substantial, and an examination of the possible alternatives should be conducted before any decision on equipment is made. Management should clearly define the direct and indirect costs and weigh them appropriately. This cost-versus-risk analysis should include the following:

1. Cost of purchasing and installing standby power hardware.
2. Exposure of the system to utility company power failure.
3. Alternative transmission or production methods available to the facility.
4. Direct and indirect costs of lost up-time because of blackout conditions.

Standby System Options

The classic standby power system using an engine generator is shown in Fig. 16. An automatic transfer switch monitors the ac voltage coming from the utility company line for power failure conditions. Upon detection of an outage for a predetermined period of time (generally one to ten seconds), the standby generator is started and once up to speed, the load is transferred from the utility to the local generator. Upon return of the utility feed, the load is switched back and the generator is stopped. This basic type of system is widely used in industry and provides economical protection against prolonged power outages (five minutes or more).

In some areas, usually metropolitan centers, two utility company power drops can be brought into a facility as a means of providing a source of standby power. This approach was discussed in "Selective or Secondary Spot Networks" above. As shown in Fig. 17, two separate utility company service drops are brought into the plant and an automatic transfer switch changes the load to the backup line in the event of a

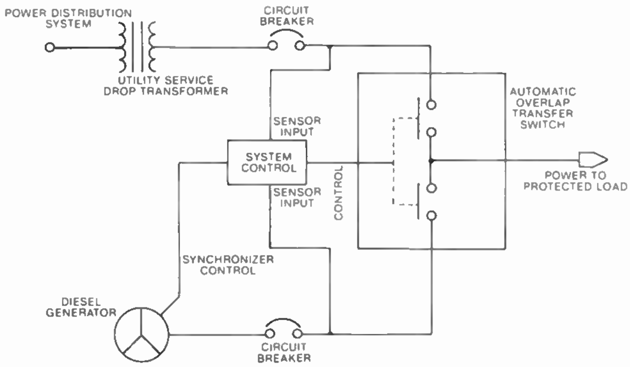


Figure 18. The use of a diesel generator for emergency power and peak power shaving applications. The automatic overlap transfer switch changes the load from the utility feed to the generator instantly so that no disruption of normal operation is encountered.

main line failure. The dual feeder system provides an advantage over the auxiliary diesel arrangement because the power transfer from main to standby can be made in less than a second. Time delays are involved in the diesel generator system that limit its usefulness to power failures lasting more than several minutes. The dual feeder system is primarily limited to urban areas. Rural or mountainous regions are not generally equipped for dual redundant utility company operation. Even in urban areas, the cost of bringing a second powerline into a facility can be high, particularly if special lines must be installed for the feed.

Fig. 18 illustrates the use of a backup diesel generator for both emergency power and peak power shaving applications. Commercial power customers can often realize substantial savings on utility company bills by reducing their energy demand during certain hours of the day. An automatic overlap transfer switch is used to change the load from the utility company system to the local diesel generator. The changeover is accomplished by a special transfer system that does not disturb the operation of load equipment. This application of a standby generator can provide financial return to the facility regardless of whether the unit is ever needed to carry the load through a commercial power failure.

A more sophisticated power control system is shown in Fig. 19, where a dual feeder supply is coupled with a motor-generator to provide clean, undisturbed ac power to the load. The motor-generator will ease the transition from the main utility feed to the standby, often making a commercial power failure unnoticed by on-site personnel. A motor generator will typically give up to ½ second of power fail ride-through, more than enough to accomplish a transfer from one utility feed to the other.

Choosing a Generator

The generator rating of a standby power system should be chosen carefully, keeping in mind any anticipated future growth of the plant. It is good practice to install a standby power system rated for at least 25% greater output than the peak facility load. This headroom gives a margin of safety for the standby equipment and allows future expansion of the facility without worry about overloading of the system. The type of generator chosen should also be given careful consideration. Generators rated for more than 100 kW power output are almost always diesel-powered. Smaller generators are available that use diesel, gasoline, natural gas, or propane as the fuel source. The type of power plant chosen is usually determined primarily by the environment in which the system will be operated.

For example, a standby generator that is located in an urban area office complex may be best suited to the use of an engine powered by natural gas, because of the problems inherent in storing large amounts of fuel. State or local building codes may place expensive restrictions on fuel storage tanks and make the use of a gasoline- or diesel-powered engine impractical. The use of propane is usually restricted to rural areas. The availability of propane during periods of bad weather (which is when most power failures occur) must also be considered.

UPS System

An uninterruptable power system is an elegant solution to power outage concerns. The output of the UPS inverter may be a sine wave or pseudo sine wave (really a stepped square wave). Sine wave units are usually more complicated and expensive than stepped square wave inverters. The sine wave systems, however, offer guaranteed compatibility with load equipment. Stepped square wave inverters may require some amount of compatibility testing before installation.

UPS systems are available in the rectifier-inverter configuration or in a less expensive switching version, shown in Fig. 20. In the normal mode, the utility company ac line is connected directly to the UPS output terminals. If the utility supply should drop

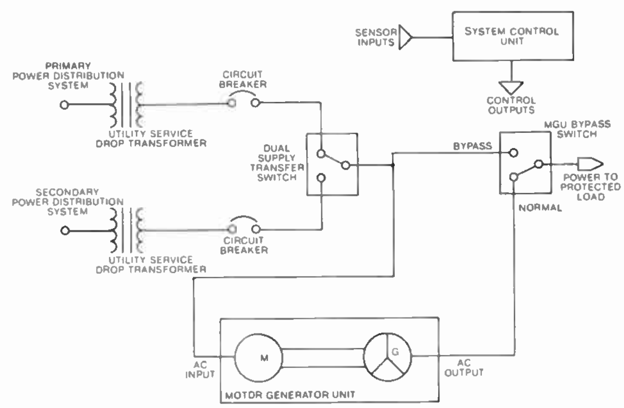


Figure 19. The dual feeder standby power system using a motor-generator unit to provide power fail ride-through and transient disturbance protection. Switching circuits allow the MGU to be bypassed, if necessary.

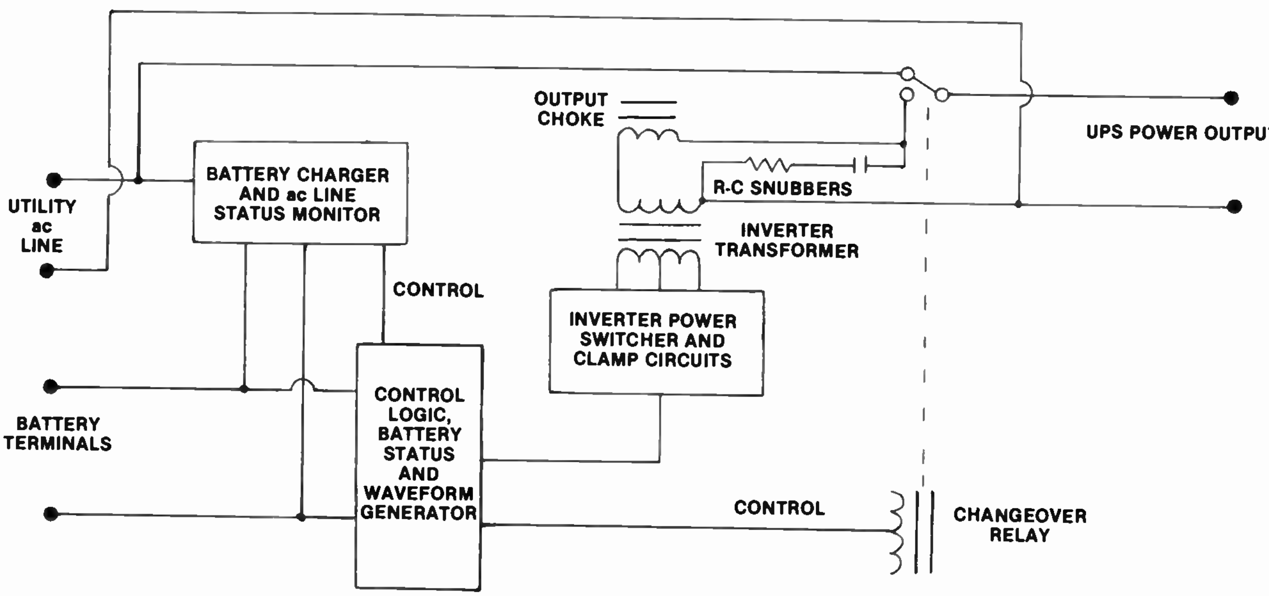


Figure 20. Block diagram of a load switching inverter/UPS system.

below a present level (or disappear completely) the inverter will start and feed the load from the battery supply. This transfer is generally accomplished in less than one ac cycle. No disturbance of the load is experienced in most cases. When the utility feed is restored, the UPS system will switch back to the primary power source and turn off the inverter. An internal charging circuit then recharges the system's battery bank. When shopping for a UPS system, consider the following points:

1. Power reserve capacity for future growth of the facility.
2. Inverter current surge capability (if the system will be driving inductive loads, such as motors).
3. Output voltage and frequency stability over time and with varying loads.
4. Required battery supply voltage and current. Battery costs vary greatly, depending upon the type of units needed.
5. Type of UPS system (rectifier-inverter or switching) required by the particular application. Some sensitive loads may not tolerate even brief interruptions of the ac power source.
6. Inverter efficiency at typical load levels. Some inverters have good efficiency ratings when loaded at 90 percent of capacity, but poor efficiency when lightly loaded.
7. Size and environmental requirements of the UPS system. High power UPS equipment requires a large amount of space for the inverter/control equipment and batteries. Battery banks often need special ventilation and ambient temperature control.

Critical System Bus Protection

A facility seeking standby power capability should consider the possibility of protecting key pieces of equipment from power failures with small, dedicated, uninterruptible power systems. Small UPS units are available with built-in battery supplies for microcomputer systems and other hardware. If cost prohibits the installation of a system-wide standby power supply (using generator or solid-state UPS technology), consider establishing a critical load bus that is connected to a UPS system or automatic transfer generator switch. This separate power supply is used to provide ac to critical loads, thus keeping the protected systems up and running. The concept is illustrated in Fig. 21. Unnecessary loads are dropped in the event of a power failure.

A standby system built on the critical load principle can be a cost-effective answer to the power failure

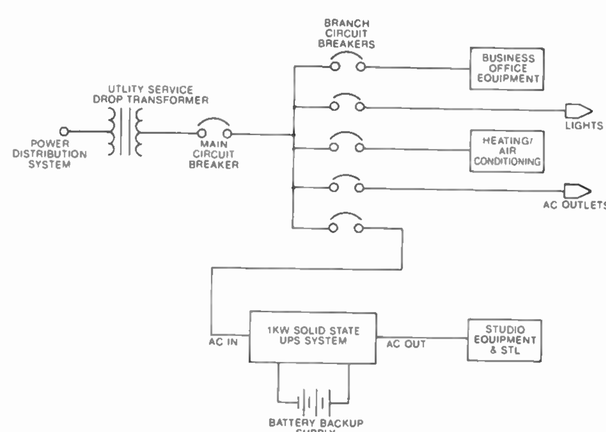


Figure 21. Application of the critical load power bus concept. In the event of a power failure, all equipment necessary for continued operation is powered by the UPS equipment. Noncritical loads are dropped until commercial ac returns.

threat. The first step in implementing a critical load bus is to determine accurately the power requirements for the most important equipment. Typical power consumption figures can be found in equipment instruction manuals. If the data are not listed in the equipment manual or available from the manufacturer, they can be measured using a wattmeter.

BIBLIOGRAPHY

"Voltage Transients and the Semiconductor," The Electronic Field Engineer, pp. 37–40, March/April 1979.

"Noise Suppression Reference Manual," Topaz Electronics, San Diego.

"General Electric MOV Varistor Data and Applications Manual," General Electric Company, Auburn, NY.

"A Review of Transients and Their Means of Suppression," Application Note AN-843, Motorola Semiconductor Products, Phoenix, AZ.

"The Susceptibility of the Open Delta Connection to Third Harmonic Disturbances," Harris Corporation application note, Quincy, IL.

Key, Lt. Thomas, "The Effects of Power Disturbances on Computer Operation," IEEE Industrial and Commercial Power Systems, conference paper, Cincinnati, OH, June 7, 1978.

Martzloff, F. D., "The Development of a Guide on Surge Voltages in Low-Voltage ac Power Circuits," Proceedings of the 14th Electrical/Electronics Insulation Conference, IEEE, Boston, October 1979.

Hill, Mark, "Computer Power Protection," Broadcast Engineering Magazine, Intertec Publishing, Overland Park, KS, April 1987.

7.10 Low Power TV Transmission Systems

Robert M. Unetich
ITS Corporation, McMurray, Pennsylvania

Additional material contributed by Peter Tannenwald
Arent, Fox, Kintner, Plotkin & Kahn, Washington, District of Columbia

INTRODUCTION

Conventional television broadcasting is built on the concept of high-powered single-channel stations with wide area coverage. This service is supplemented by two lower-powered terrestrial television technologies: *low power television/translators* operating on conventional TV channels, and *MMDS/ITFS* operating in the 2500–2690 MHz band.

Low Power Television and TV Translators

Low power television stations (LPTV) and television translators operate on a secondary basis on unused VHF and UHF television broadcast channels. Like conventional TV stations, they are licensed on a single-channel basis and utilize the same NTSC format, so their signals are receivable on consumer television receivers without any need for an adaptor. LPTV stations originate programming; but may also rebroadcast other television or LPTV stations, and may take feeds from satellites or microwave systems. TV translators rebroadcast the signal of a conventional television station. Licensees may change their status between LPTV and translator at will, simply by notifying the FCC.

TV translators were first authorized in 1956 as a way to bring television service to remote areas, or areas shielded by terrain obstacles. Program origination was forbidden until 1982, when the LPTV service was created as a way to introduce new local programming services to markets that might otherwise not enjoy them. There are some 4,000 LPTV stations and translators, several hundred of which are in Alaska. In addition, the FCC authorizes conventional TV stations to operate on-channel boosters to fill in coverage gaps. Very few boosters are in operation, as there are not many situations where it is possible to transmit on-channel without risking interference to the primary station signal.

LPTV stations and TV translators generally use 1-watt or 10-watt transmitters at VHF and 1 kW transmitters at UHF. Very high gain antennas are commonplace, resulting in UHF effective radiated power levels in excess of 100 kW, although with correspondingly narrow vertical apertures. Coverage generally extends from five to 30 miles from the transmitter, but is highly dependent on terrain and other obstacles, and whether or not other interfering signals are present. Only the 74 dBu contour is protected by the FCC.

Applications for new LPTV stations and TV translators and for major changes in existing stations may be filed only during designated filing "windows," usually one week in length, which are announced by the FCC approximately once each year. Minor change applications, defined as those which involve no new frequency or new service area, may be filed at any time. Mutually exclusive applications are subjected to lotteries rather than comparative hearings. During the pendency of the advanced television systems (ATV) rule making, applications for new stations are prohibited within 100 miles of 36 designated markets. Major changes are permitted within the major markets.

LPTV stations and translators may be displaced at any time by new full-power television stations. Displaced stations may file applications to change channel without waiting for an application filing window, and are able in that way to obtain priority over new applicants in applying for any available spectrum. If no new channel can be found that does not cause interference, the displaced station must go dark.

An LPTV station or translator may be built anywhere that it will not cause interference to other stations. Thus the allocation system is similar to AM radio, except that interference analyses are done solely on the basis of a special FCC computer program. Directional antennas are common. Special terrain shielding showings will be entertained by the FCC only in situations where no mutually exclusive applications are filed.

LPTVs and translators are governed by Part 74 rather than Part 73 of the FCC's Rules. Operating requirements are generally more lenient than for conventional television. Unattended operation is permitted when rebroadcasting another television station by direct translation of the input signal, or when rebroadcasting a satellite or terrestrial microwave feed other than an STL. At all other times, an operator must be on duty at the transmitter or an authorized remote control point. Section 74.780 of the FCC's Rules lists which sections of the broadcast rules in Part 73 are applicable to LPTV stations and translators.

Design considerations for LPTV stations and translators are similar to those for full-power stations, except that the economics are much more modest. Equipment is smaller and less expensive and requires much less power. A transmitter site at the highest available elevation is desirable, unless a lower elevation is needed to avoid causing prohibited interference to another station. Antenna design is critical, because extremely high gains and directional patterns are so common. Studio equipment may be high-grade consumer type, such as Super-VHS, when dictated by economic considerations.

MMDS and ITFS—"Wireless Cable"

The Multichannel Multipoint Distribution Service (MMDS) and the Instructional Television Fixed Service (ITFS) are multichannel technologies that transmit in the 2500–2690 MHz band and so require an adaptor (downconverter) at each receiving location in order to be received on a consumer-type television receiver. The technology is microwave-based, although operation is omnidirectional rather than point-to-point, except where directionalization is required to avoid interference to another station. Transmitter power levels have traditionally been 10 watts, but the FCC now authorizes operation at 100 watts subject to an interference showing. Because of the area of the spectrum in which these services operate, propagation is generally line-of-sight, and an unobstructed view of each receiving location is more important than power levels.

There are 32 MMDS and ITFS channels available, each 6 MHz wide, with one extra channel in major markets. Originally, MMDS was known as "MDS" and was a single-channel common carrier service at 2150 MHz ("Channel 1"), which is still available. In major markets, a second 6 MHz channel is available at 2162 MHz ("Channel 2"); in smaller markets, the second channel is restricted to only 4 MHz and was intended for nonvideo use ("Channel 2A"). The 31 channels at 2500–2690 Mhz are divided into groups known as Groups A through H, each with four 6 MHz channels, except for the H group, which has only three. The A and B groups are interleaved, as are the C and D, E and F, and G and H groups.

Originally, the A through G groups were reserved for ITFS and were used by educational institutions for "closed circuit" instruction. The H group was allocated for private use in the Operational Fixed Service. Entrepreneurs then began developing the idea of using this spectrum for pay movies and other video services, and the FCC came to perceive the technology to have the potential to compete with multichannel cable television. Thus the term "wireless cable" was born. The FCC reallocated the E and F groups, and most recently the H group, to MMDS. It also gave MMDS licensees the option to operate privately rather than as common carriers, so that the licensee of the station and the video programmer may be the same entity instead of being separated as they were in the common carrier mode.

MMDS operators are also permitted to lease unused time on ITFS channels, and restrictions against holding licenses for more than one MMDS channel group have been eliminated; so MMDS operators can potentially put together a relatively large number of channels to compete with wired cable service. Using a technique known as "channel mapping," MMDS operators can automatically switch consumer receivers from one channel to another, so that the consumer perceives that he or she is receiving full-time commercial service and is unaware that each channel is being diverted to instructional use during the 20 hours a week required by the FCC.

MMDS stations are given 45 dB protection from interference within a 15-mile radius of the transmitter, as are ITFS stations during hours when they are being leased to MMDS operators. During hours of instructional operation, ITFS stations have no fixed interference protection radius, but individual receive sites that are reported to the FCC and specified on the station license are protected no matter how far they are from the transmitter. Precise channel offset may be used to reduce co-channel protection to as little as 28 dB, but only if the licensees of both stations involved consent.

The MMDS and ITFS services are based on the use of all channels, including first adjacent channels, even though the consumer-type television receivers on which programming is displayed are not designed to reject strong first adjacent channel interference. Some adjacent channel protection can be built into downconverters. However, the principal technique to avoid adjacent channel interference is collocation of transmitters, as the FCC's interference rules provide for a 0 dB first adjacent channel protection ratio which can always be achieved if both stations are at the same site. The FCC permits one licensee to force another to collocate at the newcomer's expense, subject to certain conditions.

MMDS and ITFS applications may be filed at any time. Most channels have already been applied for in major markets. Conflicting MMDS applications are placed in lotteries. Conflicting ITFS applications are evaluated by a point system designed to select the most desirable applicant without a trial-type hearing.

Even though most MMDS stations now operate on a private rather than a common carrier basis, MMDS

TABLE 1
Typical signal coverage in miles.

Transmitter Power (watts)	Coverage in Miles			
	1	10	100	1,000
VHF	6	22	N/A	N/A
UHF	4	16	22	28
2.5 - 2.7 GHz	3	12	20	N/A

is still licensed by the FCC's Common Carrier Bureau and is regulated under Part 21 of the Rules. ITFS is licensed by the Mass Media Bureau and is licensed under Part 74.

SYSTEM DESIGN CONSIDERATIONS

System Design

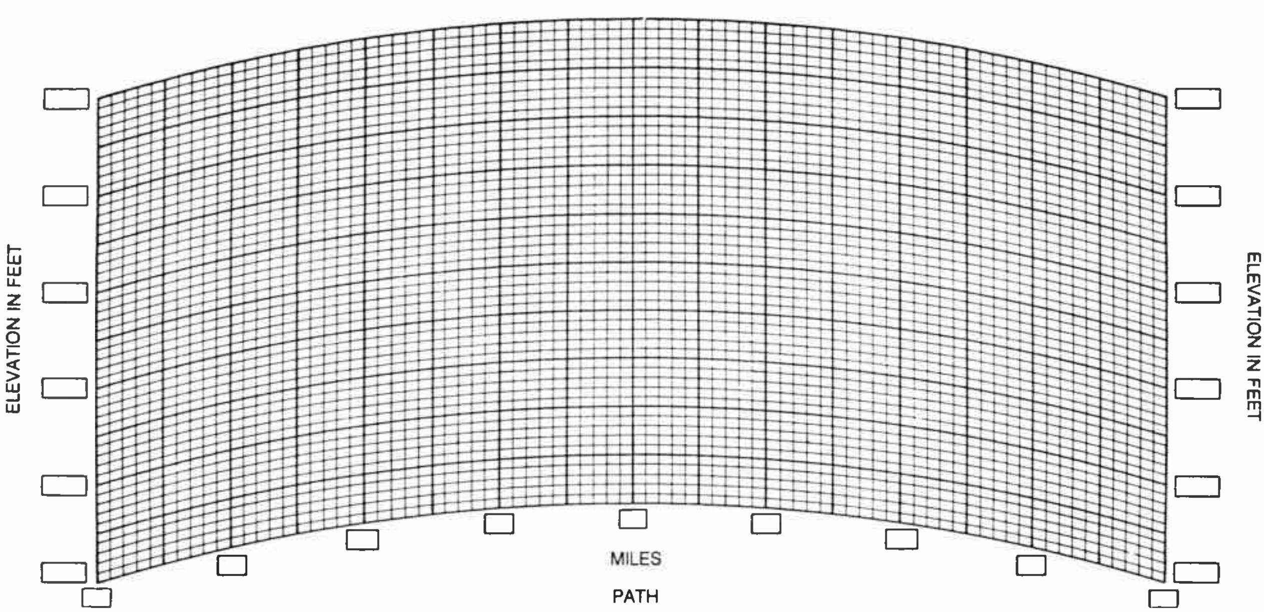
System design begins with the need for coverage of a geographical area by a television signal source. This source may consist of a single channel of NTSC video, a group of channels, or a multichannel array of encoded channels.

Although differences exist in the propagation characteristics in VHF, UHF, and S-band signals, coverage is primarily "radio line-of-sight." Coverage is often more limited by terrain than the transmission power selected. It is generally most effective to choose the highest available site with suitable facilities and select the lowest transmitter power that will provide the required coverage. The usefulness of a particular site is often influenced by the difficulty in delivering input signals to the site, and the availability of a building, AC power, and tower space. Table 1 indicates typical coverage areas achieved with the nominal output powers most commonly selected in these "low power" services.

Table 1 assumes minimal terrain roughness and a reasonable transmitter site. Conditions affecting a particular system can substantially alter the coverage areas achieved. Figures in the table were derived from FCC coverage curves and extrapolations to S-band.

Propagation Characteristics

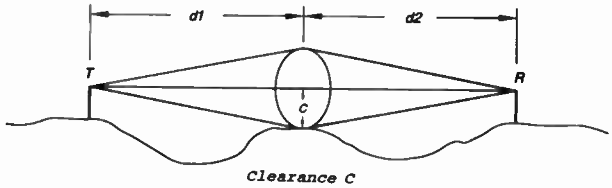
In a TV RF system, coverage beyond visible line-of-sight can be achieved through two mechanisms: refraction and diffraction. Refraction is the bending of electromagnetic waves that occur as a signal propagates through different densities of the atmosphere. This effect causes signals to be extended beyond visible line-of-sight. While the effect is somewhat frequency dependent (in the frequency range that these services operate in) it can be approximated by considering the earth to be of a larger radius. Special graphical techniques, such as 4/3 earth radius graph paper (see Fig. 1), can be used to estimate the propagation between particular points. This approach is useful in determining if sufficient clearance is available at a translator site, and it is useful in determining the potential for successful coverage of a particular area to be served. While variations in atmospheric condi-



MAJOR SCALE DIVISIONS

ELEVATION	PATH
250 feet	5 miles
1,000 feet	10 miles

Figure 1. 4/3 earth radius path profile.



where:

$C = 0.6F_1$

and:

F_1 = first Fresnel zone radius

$F_1 = 31.6\sqrt{\frac{\lambda d_1 d_2}{d}} = 548\sqrt{\frac{d_1 d_2}{fd}}$

d_1 = distance to near end of path (km)

d_2 = distance to far end of path (km)

$d = d_1 + d_2$ = total path length (km)

λ = wavelength (m)

f = frequency (MHz)

Equation 1. Path clearance.

tions can modify the nominal value of refraction, the 4/3 earth radius approach is sufficient in most U.S. domestic applications to estimate propagation.

A cross section of a path between sites shows a series of concentric circles known as Fresnel zones. Six-tenths of the first Fresnel zone clearance is desirable to achieve propagation with near free-space path loss. The first Fresnel zone size is calculated in Equation 1.

Diffraction is an additional bending of the wave and occurs as signals approach an obstruction. This effect varies according to the smoothness of the obstructing surface relative to the wavelength of the propagating signal. The most dramatic effect is known as "knife edge" diffraction. This is a common phenomena preventing visible line-of-sight into an area shadowed by a hilltop, but allowing coverage of signal into the shadowed area. Signal levels in the "diffraction region" fall off rapidly. Nevertheless, it is a common method of receiving signals at a site behind a structure. Loss of signal in the diffraction region is shown in Fig. 2. This figure also shows the results of path clearance in excess of 0.6 First Fresnel zone.

Receiver Sensitivity

A factor that may be frequency dependent is receiver sensitivity. S-band and UHF receivers are often less sensitive than VHF receivers. Receiver sensitivity can be determined by measuring system noise figure. Noise figure is defined in Equation 2.

The degradation of the signal-to-noise ratio as a signal is processed by a system is affected by the noise contributed by the system. A low noise figure indicates that a system adds little thermal noise to the signal.

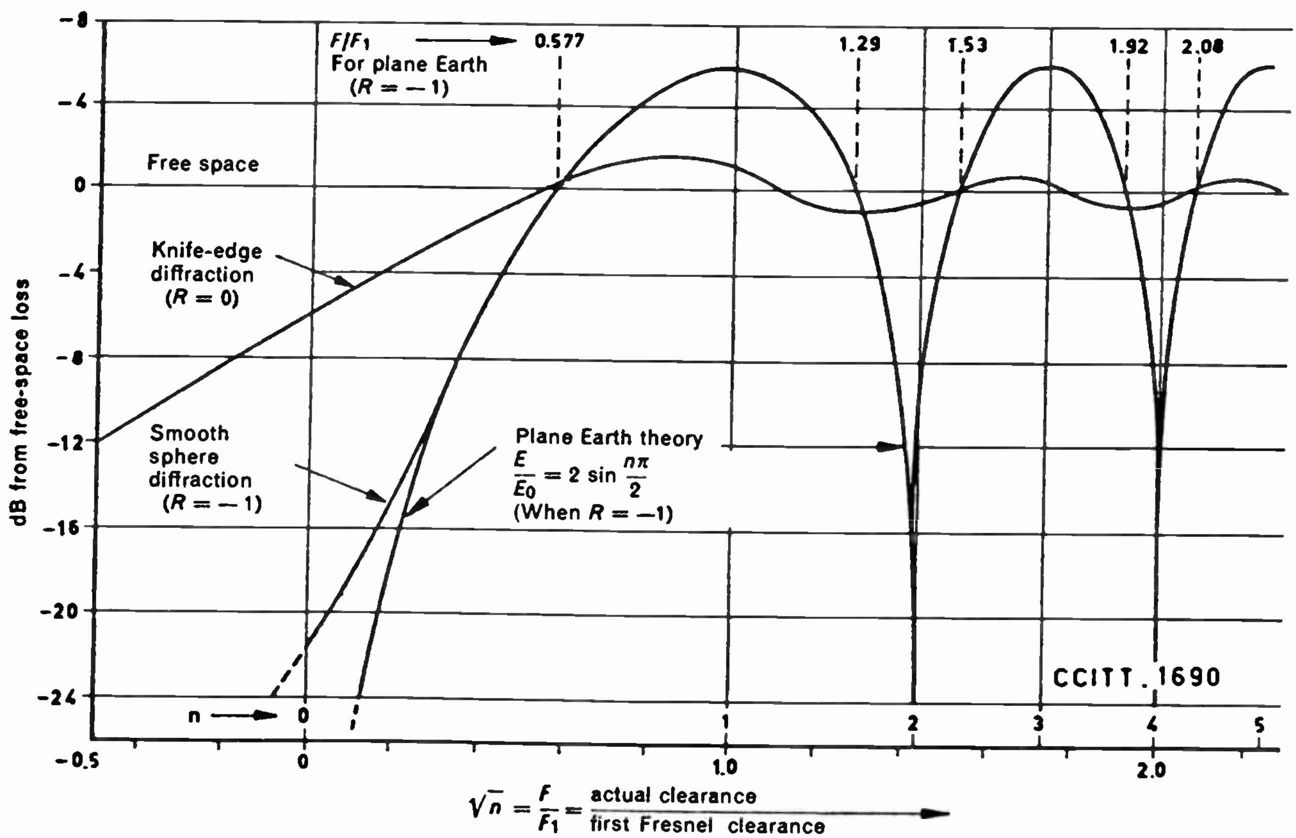


Figure 2. Path clearance.

$$\text{Noise Figure} = 10 \log_{10} (\text{NF})$$

where:

$$NF = \frac{S/N^{input}}{S/N^{output}}$$

and:

S/N is the ratio of Signal Power to Thermal Noise Power

Equation 2. Noise figure.

Lower noise figures are more easily achieved at lower frequencies because of available electronic device technology. Recently, however, gallium arsenide field effect transistors (FETs) are available that provide S-band noise figures comparable to the best noise figures available at VHF, with relatively little cost difference. Noise figures on the order of 1 dB to 7 dB are typical in these services.

An additional noise source often exists at VHF. In low band VHF channels, noise other than normal thermal noise may be present due to atmospheric or electrical discharges. This fact limits the usefulness of the low band VHF channels in some translator and LPTV applications. In these cases it may be found that the dominant noise source may limit the ability to receive low-band VHF channels at a translator site or at ultimate viewer locations. For this reason it is common for high-band VHF channels to be selected over low-band VHF, when possible, for translators and LPTV services.

Output Power

FCC Rules establish standard power limits for the low power services. Stations operating in the VHF band have a transmitter output power (TPO) limit of 10 watts east of the Mississippi and 100 watts west of the Mississippi. UHF stations allow 1 kW maximum output power level. Circularly polarized (CP) operation allows the doubling of transmitter power (up to a maximum of 2,000 watts), although the availability of effective CP receive antennas is limited. In the ITFS and MMDS bands, a TPO of 10 is a nominal power rating. Fifty watts and 100 watts are also common powers in this service.

Boosters are often limited by system considerations. The classic *on-channel repeater* or *booster* receives an input channel or channels and, simply amplifies and rebroadcasts the signal into a shadowed area with a directional antenna pattern. This requires a high gain amplifier and a high degree of isolation from the transmit-to-receive antenna to avoid feedback. Systems of this type are generally limited to 1 watt or less per channel, but operate with increased power if special shielding techniques are employed. Recent licensing activity that allows the feeding of boosters via FM microwave eliminates this feedback problem.

Another limitation of transmitter power is the regulation defining protected coverage areas. The normal protected areas for the services are shown in Table 2. These figures were in effect in 1990, but may have changed. Only the most recent set of regulations should be used in preparing a license application or operating a booster.

Channel Selection

The channel selection process involves complying with the interference protection requirements of the Rules. LPTV is a "secondary service" that may use a channel by accepting the risk that if the channel is a full service allocation, or if it would be affected by a full service station built on a different allocation, the LPTV station is required to select a new channel.

Because translators depend upon output to input frequency separation to achieve isolation, it is advisable to provide several channels of separation between a translator input and translator output channel, although a minimum guardband of one empty channel may be used.

Multichannel Systems

Multichannel combining techniques may be employed to allow the transmission of more than one

TABLE 2
Protected coverage.

SERVICE	PROTECTED CONTOUR OR REQUIRED SEPARATION	FCC RULES
LPTV / TRANSLATORS CH. 2-6	62 dBu Field Strength Boundary	Part 74.707 (a)(1)(i)
LPTV / TRANSLATORS CH. 7-13	68 dBu Field Strength Boundary	Part 74.707 (a)(1)(ii)
LPTV / TRANSLATORS CH. 14-69	74 dBu Field Strength Boundary	Part 74.707 (a)(1)(iii)
ITFS	50 Miles Separation	Part 74.903 (b)
MMDS (OMNI PATTERN)	15 Miles Radius	Part 21.902 (d)
OFS	50 Miles Separation	Part 94.63

channel from a single transmit antenna. The standard technique involves using selective channel filters coupled through a directional device into a common transmission line. Important considerations involve insertion loss per channel, power handling capability, channel combiner size/cost, and antenna system bandwidth. Channel combiners for the ITFS/MMDS band are available that efficiently couple up to 16 alternate channels onto a single transmission line. The transmission of adjacent channels can be accomplished through the use of separate transmit antenna systems or broadband hybrid couplers.

Antennas

Transmit antennas often concentrate power in a desired direction. This concentration of power is calculated as antenna gain and is a measure of the effectiveness of radiation of an antenna compared to a simple dipole antenna or a point source radiator. There are two important parameters that fundamentally determine antenna gain: vertical beam width and horizontal beam width (also known as azimuth).

Power is concentrated in a narrow beam vertically by the use of stacked radiators. The signal emanating from this array adds in-phase in the direction of the main beam and cancels in other directions, effectively increasing the power transmitted in the main radiation pattern of the antenna. The azimuth pattern of the antenna is determined by the placement of the radiating elements. These radiating elements are normally arrays of dipoles or arrays of slots in a waveguide. Both types of antennas are widely used from VHF through S-band.

While a variety of antenna patterns is available, the most common patterns are the omnidirectional and cardioid. The omnidirectional attempts to radiate power equally in all horizontal directions. A cardioid pattern is shown in Fig. 3.

This basic pattern and variations are available with antenna gains from 10 dB to 18 dB depending on the number of bays of radiating elements utilized. Vertical beamwidth narrows with increased horizontal gain as seen in Fig. 4.

Additional considerations in antenna selection and specifications are beam tilt and null fill. Beam tilt involves radiating the horizontal pattern of the antenna slightly away from the horizon. Beam tilt provides coverage into terrain below the antenna. It is common that − ½° or − 1° of beam tilt is specified to accomplish this objective. Null fill provides a small amount of radiation in the direction of an antenna null to provide some degree of coverage in the area of signal cancellation.

Directional antennas constructed with parabolic or cylindrical reflectors are occasionally used for transmission, but are commonly used for receive purposes at repeater sites and at ITFS/MMDS receive locations. This type of antenna is very directional compared to standard transmitting antennas, and gains in excess of 20 dB are common. Horizontal beam widths less than 6° can easily be achieved. A common receive antenna with specifications is shown in Fig. 5.

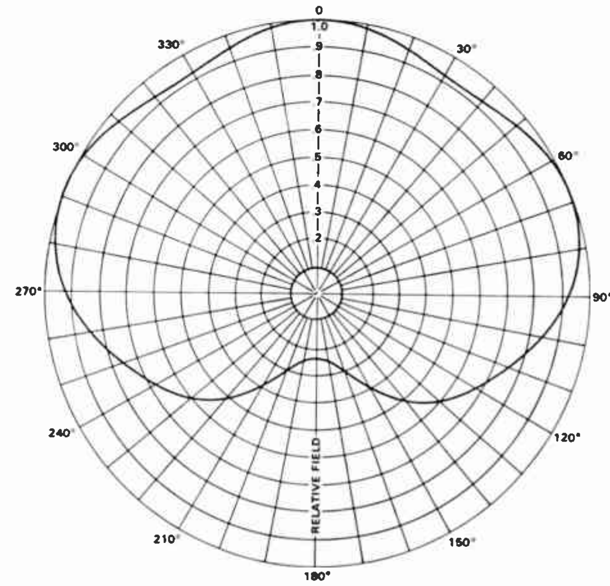


Figure 3. Cardioid antenna pattern. (Courtesy Andrew Corporation.)

Antennas have limitations in terms of power handling and bandwidth. Antenna manufacturers offer optional versions of antennas at higher costs that provide special capabilities.

Site Requirements

Transmitter sites are normally located on hilltops, in tall buildings or at convenient tower locations. In addition to the obvious requirements for the receiving and transmission hardware, a site should be selected to provide the desired coverage with the available antenna mast and selected antenna. Site requirements can be subdivided into three categories: environmental considerations, AC power requirements, and grounding and transient protection.

Environmental Considerations

Electronic equipment is designed to operate over a normal range of temperatures and within a reasonable range of moisture and dust. Maintaining room temperatures and equipment cleanliness below the upper limits can improve the reliability of a system. A good rule of thumb concerning equipment reliability is the common rule applied to chemical reactions that states that most reactions double in speed for each 10°C rise in temperature. Since many detrimental environmental effects, such as contact corrosion, are chemical reactions by nature, the speed of chemical reactions will affect equipment longevity. Corrosion, metal migration in transistors, and the degradation of electrolytic capacitors closely follows this general rule. Maintaining low temperatures at a site will usually enhance reliability. Most equipment operates comfortably in ambient temperatures as low as 0°C. Temperatures in excess of 35°C (95°F) will normally be outside the range the

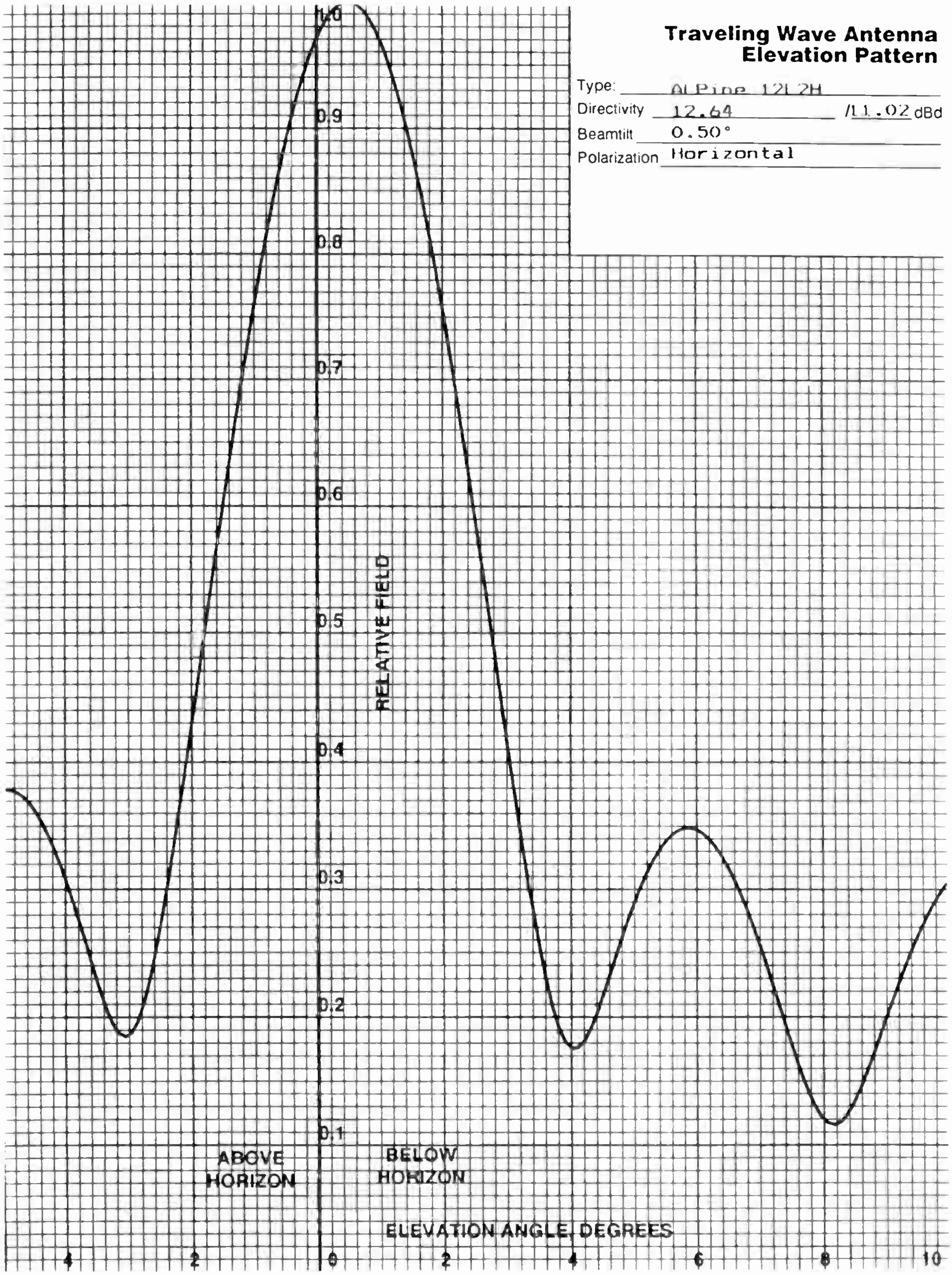
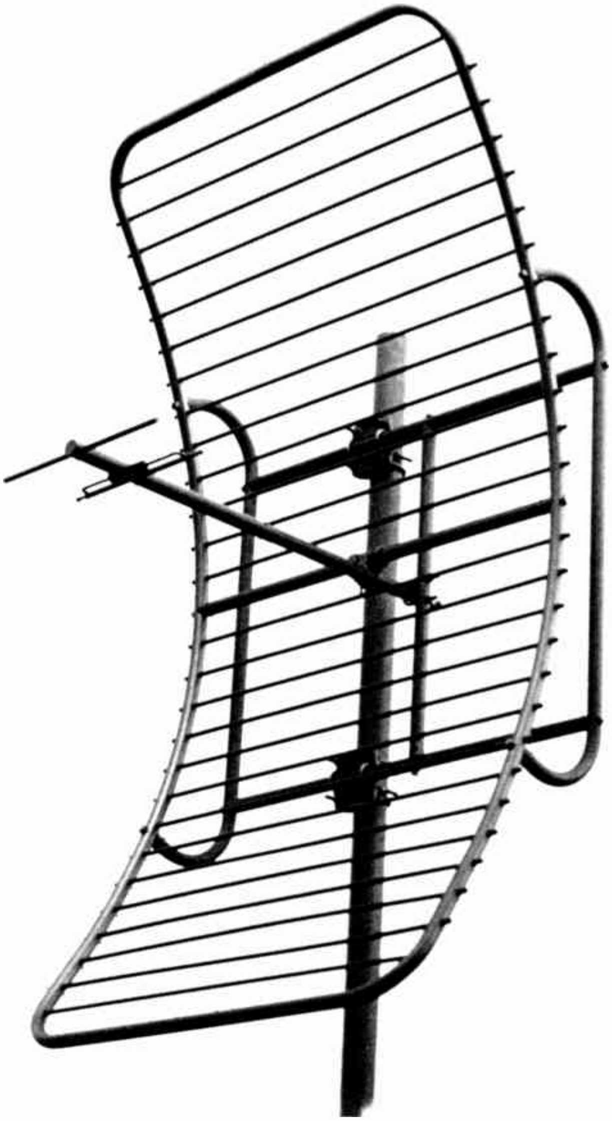


Figure 4. Twelve-bay antenna vertical beam width. (Courtesy Andrew Corporation.)



Cylindrical Parabolic Grid Reflector Antenna

Band: UHF
Gain: 18 dB nominal
Front-to-Back Ratio: 20 dB minimum
VSWR: 1.2:1
Power Rating: 100 Watts

Figure 5. Directional UHF antenna. (Courtesy Scala Electronic Corporation.)

equipment was originally tested in and should be avoided.

It is advisable to provide filtered air to a building and a method of easily removing and cleaning the building air intake filters. The cleanliness of air filters and metal heatsink surfaces will affect cooling system efficiency. Most air filters are washable and a supply of water with detergent should be available at a site. If equipment is exhausted directly to the outside, it is important that a substantial supply of unrestricted building intake air is provided. Most transmitter cooling systems are not designed to overcome substantial air pressure drops.

Another site consideration is moisture. Equipment normally operates above room temperature. Under these conditions condensation of water on the equipment is unlikely. However, in some instances it may be necessary to provide additional site heating to maintain equipment dryness. It may also be desirable to prefilter or route inlet air through baffles or ductwork to eliminate water droplets. Exhaust air should be directed away from the direction of prevailing wind. Exhaust air screens that prevent birds and animals from entering the building are necessary, but these screens may become clogged and adversely affect equipment cooling if not regularly cleaned. Automatic motor controlled louver systems can be utilized to allow the equipment to maintain a reasonable room temperature in cold weather. Motorized louver systems may malfunction and result in equipment overheating. Louver systems that rely on air pressure to open must be used with caution, since additional pressure drop is incurred.

AC Power Requirements

Site requirements for power must take into account transmitter demands, cooling/heating requirements, test equipment, and lighting needs. A 1 kW UHF transmitter typically draws more than 25 amps of AC current at 230 VAC. A 1 kW UHF transmitter will produce about 4,000 watts of exhaust heat that must be vented from the room or cooled with an air conditioning system. This converts to 13,600 BTUs per hour of air conditioning requirement.

In a multichannel system it is advisable to provide an AC shut-off per cabinet. This allows the power to be shut down to a particular cabinet to allow maintenance while other cabinets continue to operate. Wall circuit breakers are thermally operated devices. They should be sized substantially above the equipment requirements to allow for in-rush transients and thermal heating of the building. It is common to run conduit directly to cabinets for higher power systems. Lower power transmitters can often be plugged into standard AC outlet strips.

Grounding and Transient Protection

Transmitter sites require an effective earth ground to minimize AC hum problems, provide AC line transient protection and allow a safe path for lightning induced current. A site ground should bond together the ground provided from the power company with a grounding rod associated with the tower. A copper-clad steel grounding rod buried in the earth and connected to the tower, building structure, and AC power line ground is standard practice. It is also advisable to add a ground strap from equipment cabinets to the site ground. This is often constructed with flat copper strapping or heavy copper braid. Transmission line grounding kits are available for all standard transmission line cables.

These kits provide a low resistance method of connecting the ground system to the main transmission line.

Transient protection is necessary on both the AC line and RF transmission lines. Most modern transmitters incorporate some form of internal transient protection. The most common method is the installation of transient devices across the AC line and from each AC line connection to chassis ground. These transient protectors are often of the *metal oxide varistor* (MOV) type. These varistors are available in various voltage ratings. Because of the wide variation of power line voltages often present at transmitter sites, it is advisable to choose the highest rating varistor that will provide the transient protection required. For example, it is often desirable to use a 150 volt MOV instead of a 135 volt MOV to minimize transient protector damage from a momentary high-line condition. The higher voltage MOV will typically provide the necessary momentary transient protection. By connecting MOV's from line-to-line and line-to-chassis ground, common mode and unbalanced transients are clipped. Large MOVs can be installed at the AC power box to provide additional protection. More sophisticated transient protectors that insert inductive elements in series with the AC line prior to transient protection are also available.

RF input and RF output transmission lines should incorporate proper grounding kits on the outside shield of the cable. DC return paths should be provided in receiver and transmitter circuitry to reduce the danger of damage from static build-up of the RF transmission line center conductor. Harmonic filters that employ shorted quarter wave trap sections provide a low resistance DC return path to ground and further enhance the RF transmission line transient protection.

Long transmission lines that run between a transmitter site and a tower can result in lightning induced transient damage. Improved bonding between the site and the tower may be necessary. Enhanced transient capability of the receiver or transmitter circuit through the use of spark gap capacitors and low voltage transient protectors can also enhance reliability. Contact the equipment manufacturer to solve specific transient damage problems.

Encoding and Decoding

LPTV and MMDS services often depend upon encoding techniques to prevent signal pirating. ITFS requires this capability to facilitate leasing excess time to an MMDS system operator. ITFS systems occasionally require signal security to avoid broadcasting sensitive video programming. Encoding has historically been done using cable television encoding techniques. These techniques range from sync suppression to interfering carrier insertion or time-delay switching. This encoding can be accomplished at baseband video or at an *intermediate frequency* (IF).

Audio information may also be encrypted. A common technique is *audio spectrum inversion*, although other subcarrier schemes have been implemented. The Rules allow for the encoding of signals as long as "significant" picture degradation does not occur. Encoding systems cause some perceptible amount of change to the signal.

Addressing information is often sent as vertical interval or subcarrier data. Numerous variations on this basic system design exist. In most cases the demodulation/remodulation process reduces resolution and affects other signal parameters but this technique is able to produce picture quality delivered to the television viewer that is generally acceptable.

HARDWARE ANALYSIS

Describing the range of hardware required to cover VHF through S-band requires a general approach. Most television transmitters can be divided into three sections: A modulator, an upconverter and a final amplifier. The modulator processes incoming video and audio, and performs the basic modulation process at IF. The upconverter heterodyne converts the signals to the desired output frequency and provides additional processing such as level control and precorrection. The final amplifier raises the signal level to the output power desired. Peripheral equipment often includes encoding and multichannel combining equipment.

Heterodyne translators are similar except that the modulator is replaced by a receiver assembly that converts the input signal to IF. Boosters have two standard configurations: the simple on-channel amplifier and the video fed transmitter.

A full RF transmission system requires suitable transmission lines, filters, antennas, and other monitoring components. Receive hardware is subdivided into downconversion equipment and decoding hardware. The following descriptions are meant to provide basic technical information for this hardware.

Transmitters

A modulator performs the function of producing a vestigial sideband amplitude modulated picture carrier and a frequency modulated aural carrier at convenient intermediate frequencies. Signal level during demodulation determines picture brightness, and it is therefore necessary to maintain accurate modulation levels that are independent of average picture levels. This is accomplished with a video clamp circuit that holds either sync tip or black level to a specific voltage level. Other video processing can also be accomplished prior to modulation. This includes differential phase correction and video sync stretch. It is also desirable to add white clipping and sync spike clipping prior to modulation to prevent overmodulation or incorrect output levels at the output of the modulator. Vestigial sideband filtering is accomplished with a suitable surface acoustic wave (SAW) filter after modulation. These filters provide stable accurate characteristics with sharp roll-off and reasonably flat delay variation across the band.

Aural modulation is usually frequency modulation of an IF carrier. Standard IF frequencies chosen in most modulators are 45.75 MHz for visual and 41.25 MHz for aural. These frequencies correspond to the standard IF of television sets and other common test equipment, such as demodulators. With the advent of BTSC stereo television it became desirable to have an FM modulator capable of wide deviation with low distortion. It is also desirable to phase lock the center frequency of the FM carrier to the visual IF oscillator to maintain accurate visual/aural frequency separation. This can be accomplished with a very narrow band phase-locked loop that maintains the average center frequency of the FM carrier. FM modulators normally have a pre-emphasized audio input for standard baseband audio, and a wideband unbalanced input without pre-emphasis for composite audio input signals. Some modulators provide a separate filtered input for incorporating high-frequency subcarriers.

Fig. 6 illustrates the functional block diagram of a typical modulator/upconverter amplifier assembly. IF processing is performed to linearize the RF characteristics of the output amplifier array. This linearization is a combination of amplitude and phase precorrection. Amplitude precorrection is accomplished with attenuators that are bypassed at specific signal levels by the cut-in of biased diodes. This can be accomplished with little phase variation, and the sharpness of the amplitude response characteristic can be easily adjusted to match the opposite characteristics of the amplifier.

Phase correction, or more accurately, incidental carrier phase modulation (ICPM) correction can be accomplished by demodulating a sample of the IF visual carrier. The demodulated video is used to drive a phase modulator in line with the IF signal. By adjusting the amplitude and shape of the video signal it is possible to provide accurate compensation for phase errors that occur in the output amplifier of the transmitter.

An additional function that usually takes place at IF is automatic level control. A control loop is provided to maintain a constant IF or RF output level. Drift is compensated by an adjustment of a voltage controlled attenuator at IF.

Upconversion occurs in a double balanced mixer at low level. Filtering and amplification raise the level to the amplitude necessary for detection and metering.

The last section of most transmitters is final amplification. This can be accomplished with either transistors or vacuum tubes. Current transistor designs provide up to 1 kW of output in UHF and 100 watts of output in S-band. Because of the higher cost of solid-state UHF hardware, it is also common to use vacuum tubes at the 100 watt and 1,000 watt levels. Figs. 7A and 7B illustrate typical 100 watt UHF solid-state transmitters. Fig. 8 compares solid-state and vacuum tube UHF amplifiers.

There is a parallel array of similar transistor stages in the solid-state design. Because of the lower gain of these stages relative to tubes, it is necessary to provide additional driver stages. The vacuum tube amplifier accomplishes high gain and high power in a single device.

The solid-state design has the added advantage of providing a degree of reliability because of the parallel nature of the design. The failure of an individual device or individual portion of the array will not take the

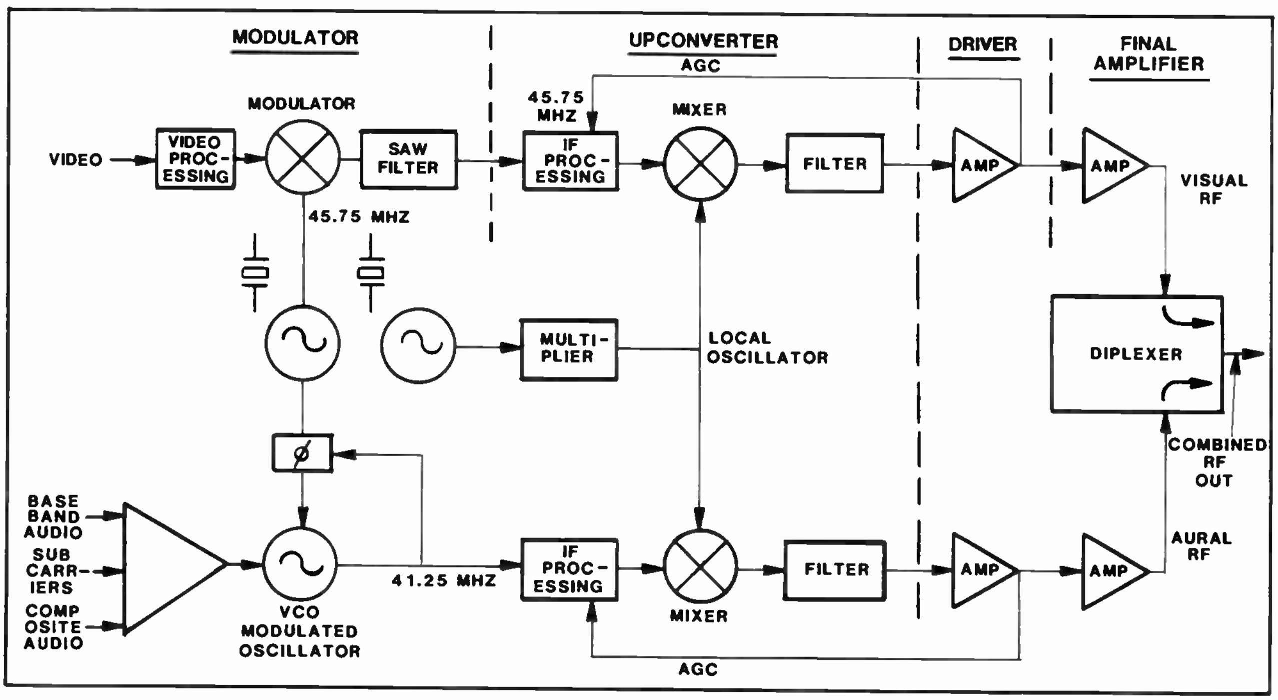


Figure 6. Television transmitter block diagram.

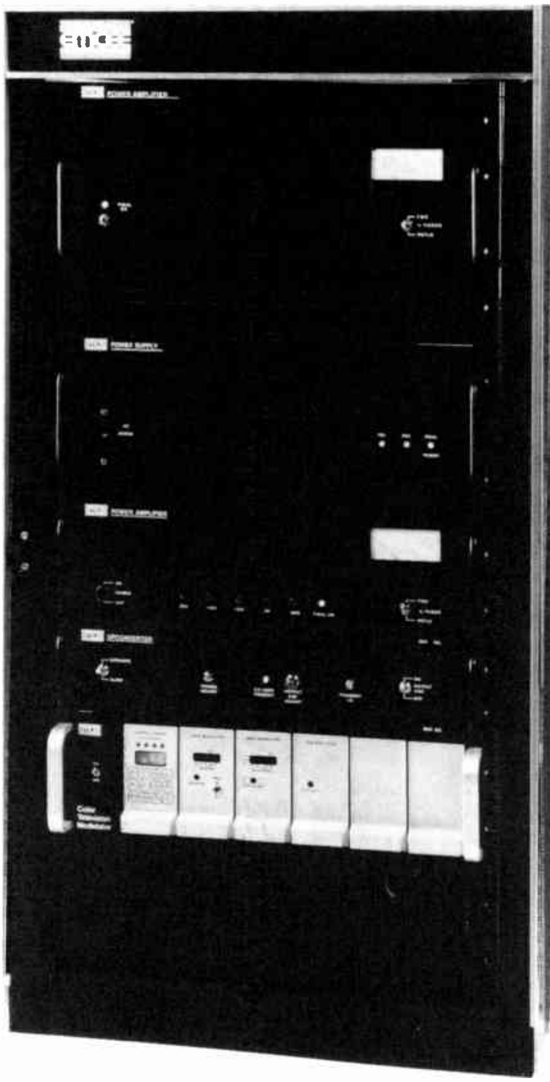


Figure 7A. 100 W UHF solid-state transmitter. (Courtesy EMCEE Broadcast Products.)

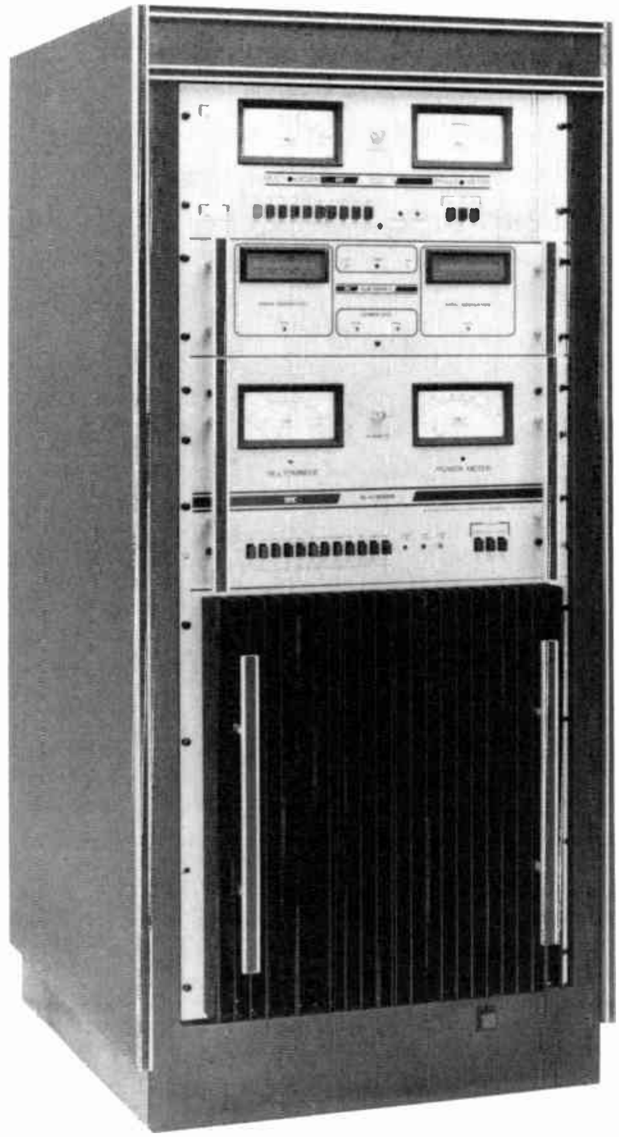


Figure 7B. 100 W UHF solid-state transmitter. (Courtesy Television Technology Corporation.)

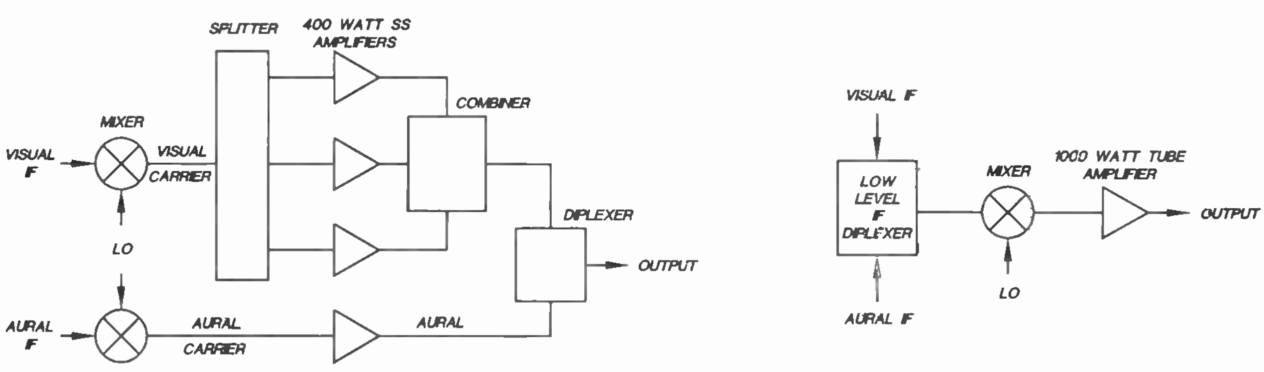


Figure 8. Solid-state and vacuum tube UHF amplifiers.

transmitter off the air. The vacuum tube transmitter, on the other hand, has all signal flowing through a single device. The solid-state transmitter shown in Fig. 8 is externally or high-level diplexed. This is done to avoid in-band intermodulation distortion products. The vacuum tube provides sufficient linearity to operate in a low level internally multiplexed mode. The extra driver stages and aural amplifier chain generally results in the solid-state array being initially more expensive than the vacuum tube design. The vacuum tube design requires tube replacement at periodic intervals. Complete 1 kW tube and solid-state models are shown in Figs. 9A and 9B, respectively.

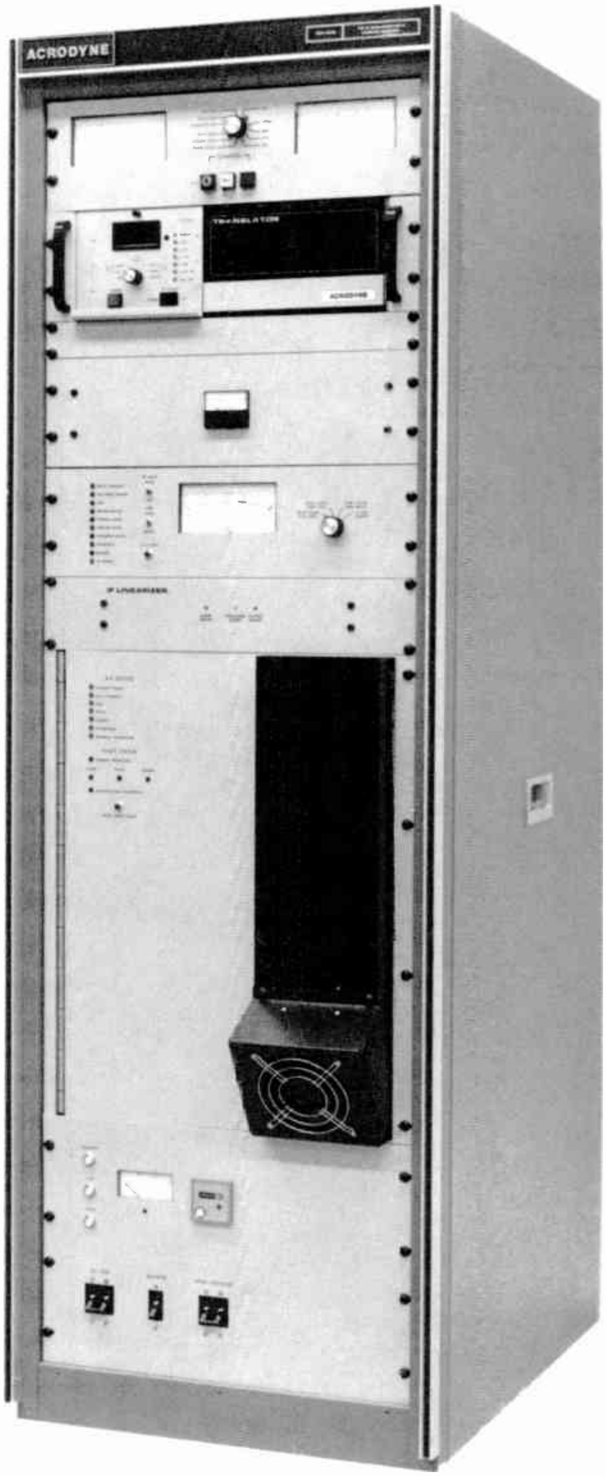


Figure 9A. Vacuum tube 1 kW UHF LPTV transmitter. (Courtesy Acrodyne Industries, Inc.)

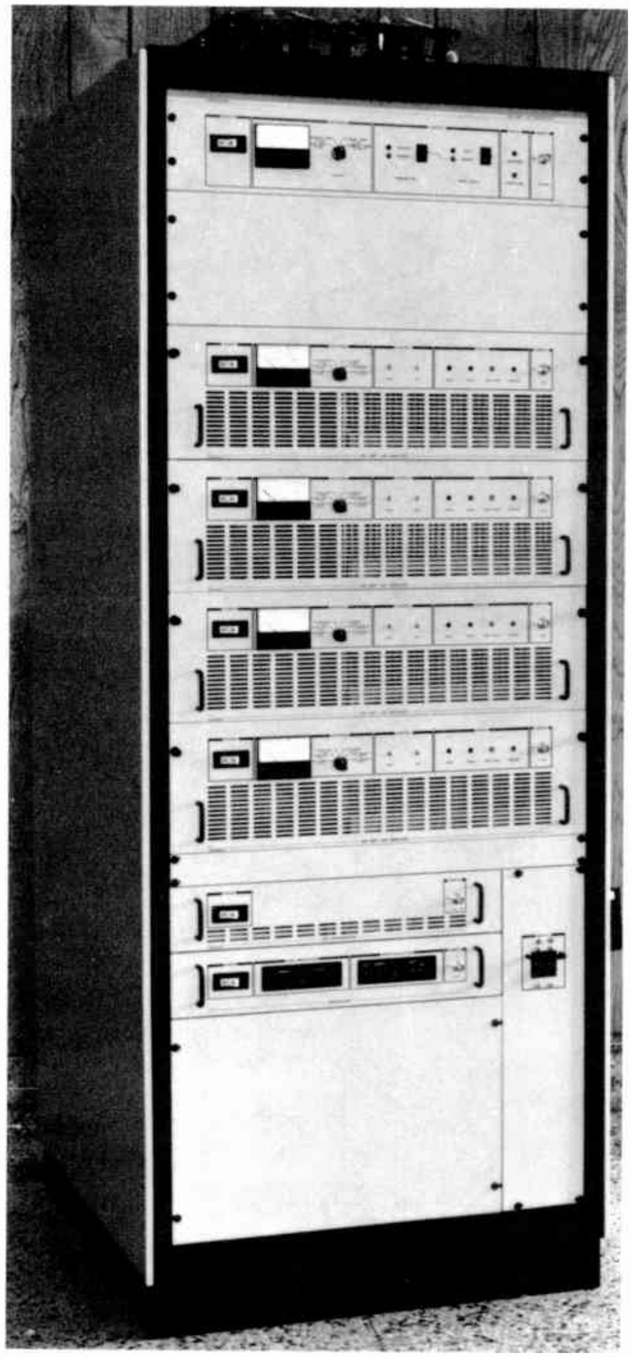


Figure 9B. Solid-state 1 kW UHF LPTV transmitter.

Back-up Transmitters

Full service broadcast systems and low-power television systems require a high degree of reliability to be successful services. While improved equipment design and the incorporation of solid-state components in place of vacuum tube stages enhances reliability, the increased dependability of a system is shown by an increase in mean time between failure (MTBF) and mean time to repair (MTTR). Major enhancement in the dependability of a system is most easily achieved through some degree of system redundancy. The ultimate redundancy is a complete backup transmitter identical to the main transmitting unit. This is a common operating method for full service broadcast stations. Low power systems often employ a back-up transmitter with reduced output power capability and a remote control or automatic switching mechanism to allow the transfer of the antenna system to the back-up unit. Multichannel systems have recently begun to employ a frequency agile back-up transmitter that has the capability of operating on any one of a group of selected channels. Fig. 10 illustrates a 28-channel frequency agile ITFS/MMDS transmitter that includes microprocessor control and instant channel switching. This frequency agility is achieved through the use of a phase-locked loop local oscillator and a UHF intermediate frequency that allows the upconversion to the output channel while maintaining the rejection of local oscillator products. This product employs broadband GaAs FET transistor circuitry and other features to insure operation on any standard ITFS/MMDS channel. Waveguide switching systems that allow the unattended operation of RF output switching are also available for multichannel backup applications.

Control from a remote point can be done with a simple dial-up remote control system. Software packages are now available to allow full remote control and operation of transmitters through standard terminals and desktop computers. Automatic dialing is included in these systems to alert the control site of a failure.

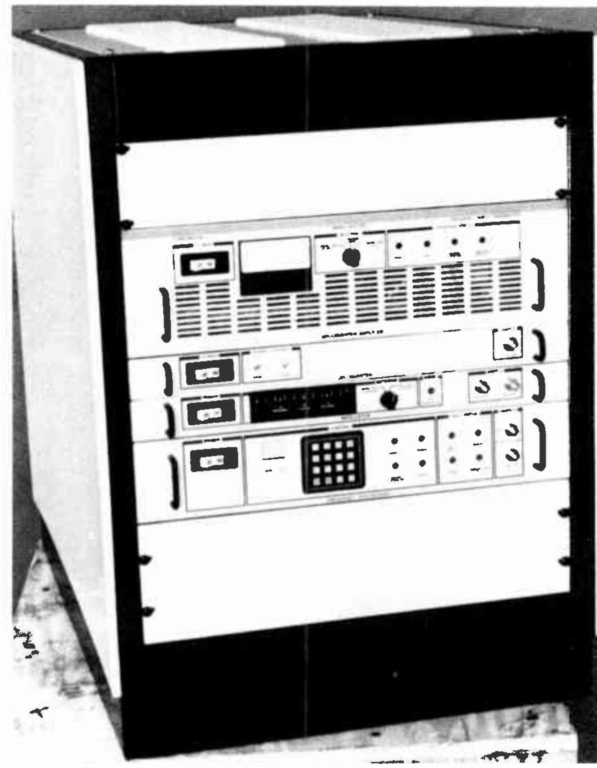


Figure 10. ITFS/MMDS frequency agile transmitter.

Translators

A heterodyne translator can be configured by replacing the modulator with a suitable UHF or VHF receiver. By using the same intermediate frequency and operating level of the modulator, it is possible to accomplish this change in a simple manner. The receiver needs to provide the necessary gain and filtering to avoid interference from other channels, and to avoid interference from the output of the translator. It is common to add a preamplifier at the receive antenna to reduce the system noise figure to the lowest possible level.

SAW filters are normally used in modern translators to provide filtering of adjacent channel products. As in transmitters, SAW filters provide extremely sharp selectivity with relatively constant delay characteristics. When the output amplifier is linear enough to allow operation with a multiplexed IF signal, SAW filter bandwidth is chosen to pass the visual and aural carriers. In the case of externally diplexed output amplifiers, it is necessary to provide a signal separation function in the receiver. Receivers have been developed to split the visual and aural IF carriers to allow separate processing in the following upconverter and amplifier assemblies. The splitting of the aural carrier is complicated by the fact that the incoming signal may arrive with a carrier frequency error. In addition, the incoming signal may contain stereo or other subcarrier information. These problems can be resolved with a phase-locked loop that tracks the incoming frequency and aural IF SAW filters that allow the full aural spectrum to pass.

Translator receivers require a very wide automatic gain control range. This is necessary to allow the use of the receiver at sites with very small input signals or at other sites with large input signals. Some designs incorporate switchable attenuators that adjust the automatic control range of the receiver to the expected range of the input signal.

Boosters

Boosters can be licensed as microwave fed transmitters. In this configuration the hardware is essentially identical to a video fed transmitter.

The earliest boosters predate television translator service. The first documented boosters were installed

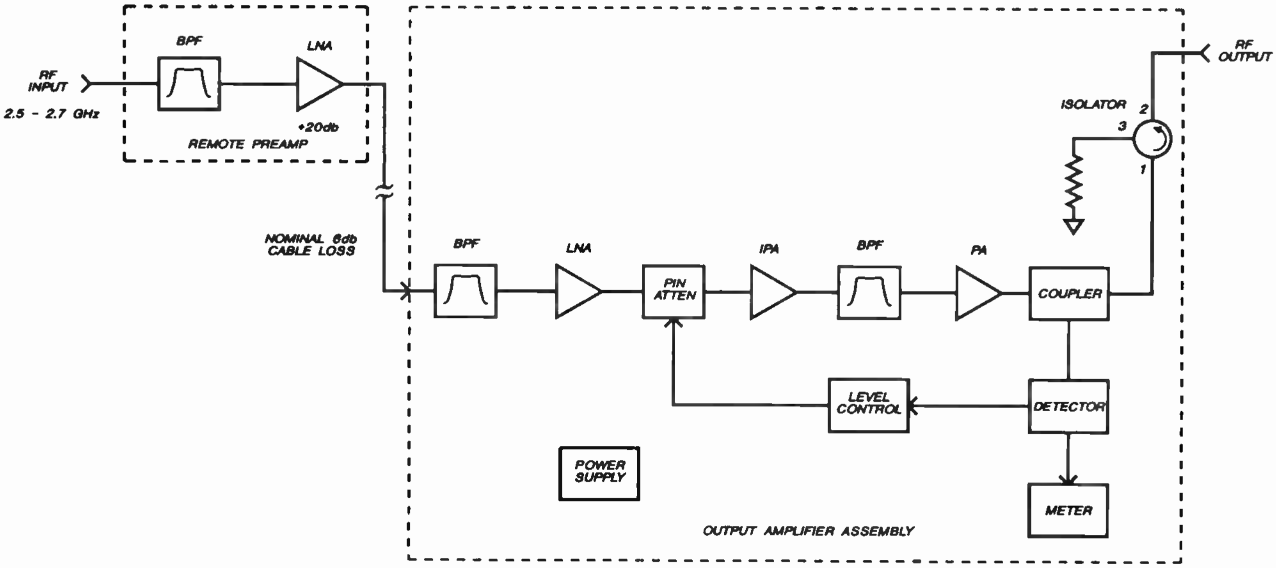


Figure 11. Multichannel booster block diagram.

in 1948 in Oregon. This booster and others which followed were simple on-channel amplifiers connected to receive and transmit antennas. In this configuration, if the antenna isolation substantially exceeds the amplifier gain, the system is stable and provides useful signal coverage. Common booster gains typically begin at 40 dB. A signal at the output of the receive antenna on the order of −40 dBm (approximately 2 mv) will provide about 0 dBm or 1 mw of output power to the transmit antenna. With a directional transmitting antenna it is possible to serve an area within several miles of the booster site. Multichannel boosters have been used in the ITFS/MMDS band for some years. The functional block diagram of an advanced booster design is shown in Fig. 11.

In this design, a remote preamplifier is mounted at the receive antenna. A coax cable connects the preamplifier to the final amplifier assembly. This allows physical separation of the gain blocks, improves the stability of the system, and reduces the system noise figure. The final amplifier includes a power detection circuit and an automatically controlled attenuator. This feedback control loop adjusts the gain of the amplifier to maintain the output power within the capabilities of the output amplifier. The selection of the proper filters in this design enable it to work in a narrow band or multichannel environment. Open loop gain exceeds 80 dB and the output amplifier has up to 10 watts of peak envelope power capability. This allows relatively high power levels to be transmitted. Pending booster regulation changes are expected to simplify the licensing of boosters. Boosters are performing an important function of both filling in shadowed areas and extending the coverage of low power systems in other countries. A photograph of a typical MMDS booster is shown in Fig. 12.

Figure 12. Multichannel booster.

Transmission Lines, Filters, and Antennas

Both coaxial and waveguide transmission lines can be used in S-band. At VHF and UHF frequencies coaxial transmission lines vary in diameter from ½" to 3⅛". The efficiency of transmission line varies with frequency and size. A table illustrating the insertion loss of various types of coaxial cable at different frequencies is shown in Table 3.

TABLE 3
Transmission line losses.

BAND	APPROXIMATE LOSS PER 100 FEET IN DECIBELS				
	1/2" FOAM	7/8" FOAM	1 5/8" AIR	3" AIR	ELLIPTICAL WAVEGUIDE
LOW BAND VHF	0.5	0.25	0.15	0.1	N/A
HIGH BAND VHF	2	1	0.6	0.4	N/A
UHF	1	0.5	0.3	0.2	N/A
S-BAND	4	2.5	1.25	N/A	0.4

At S-band it is possible to use elliptical or rectangular waveguide. Waveguide losses are substantially lower than coaxial transmission line losses, but the cost of waveguide is higher. Flexible elliptical waveguide is often used to feed transmit antennas in this frequency range. Rectangular waveguide is widely used in channel combining networks to minimize losses. WR340 waveguide and WR284 waveguide are common rectangular waveguide sizes for this frequency range. EW17 and EW20, manufactured by Andrew Corporation, are typical elliptical waveguides for this frequency range. It is possible to eliminate all coaxial components in a channel combiner and transmission line network. Suitable waveguide components, such as E-plane and H-plane elbows, directional couplers, and filters, are available for this purpose.

In the VHF and UHF frequency range there is often a harmonic and trap filter assembly feeding the transmission line. In the S-band it is also common to include a multichannel combining network. Notch filters are often required in low level diplexed assemblies to trap out-of-band spurious emissions.

In a multichannel environment it is necessary to couple several channels onto a common transmission line. This can be accomplished with a multiplexing filter. While this filter can take several forms, a typical *directional filter* is shown in Fig. 13. This allows a specific channel to be directionally coupled to the output. This filter rejects signals of other frequencies which pass over the filter unattenuated. A diagram of a multichannel directional filter is shown in Fig. 14.

Transmitting antennas are normally of the *slot* or stacked dipole design. A typical LPTV antenna with directive elements attached is shown in Fig. 15.

Side mounting can affect the pattern of a transmitting antenna. Fortunately, there are computer programs available that can predict the impact on the pattern and allow the consideration of this effect in system design. See Fig. 16.

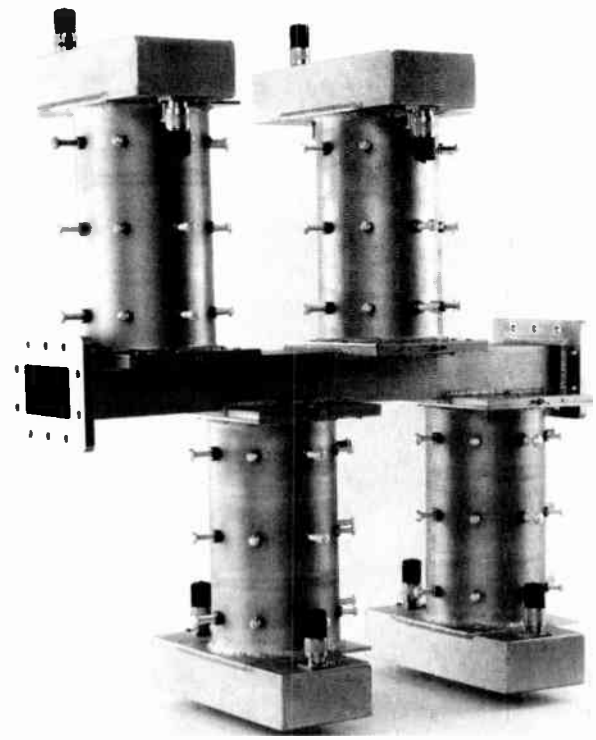


Figure 13. Typical directional filter.

S-Band Receiving Equipment

Hardware description of S-band systems is incomplete without a discussion of receiving equipment. A typical receiving system is shown in Fig. 17. In this system, a directional receive antenna couples signal into a broadband block downconverter which heterodyne converts the signal to VHF superband. The signal is then decoded and remodulated onto a low-band VHF channel.

The Receive Antenna

Receive antennas vary from 12 dB gain corner reflectors to 36 dB gain parabolic dish designs. A short

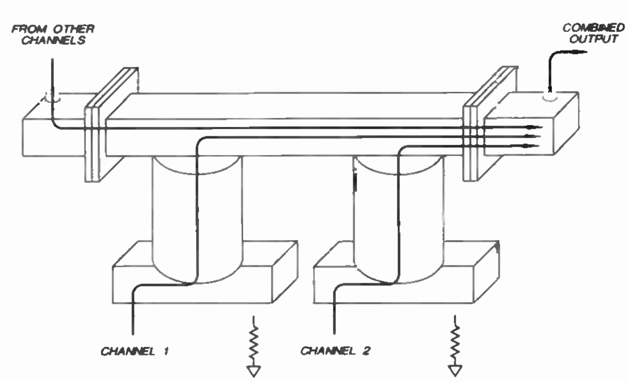


Figure 14. Two channel combiner.

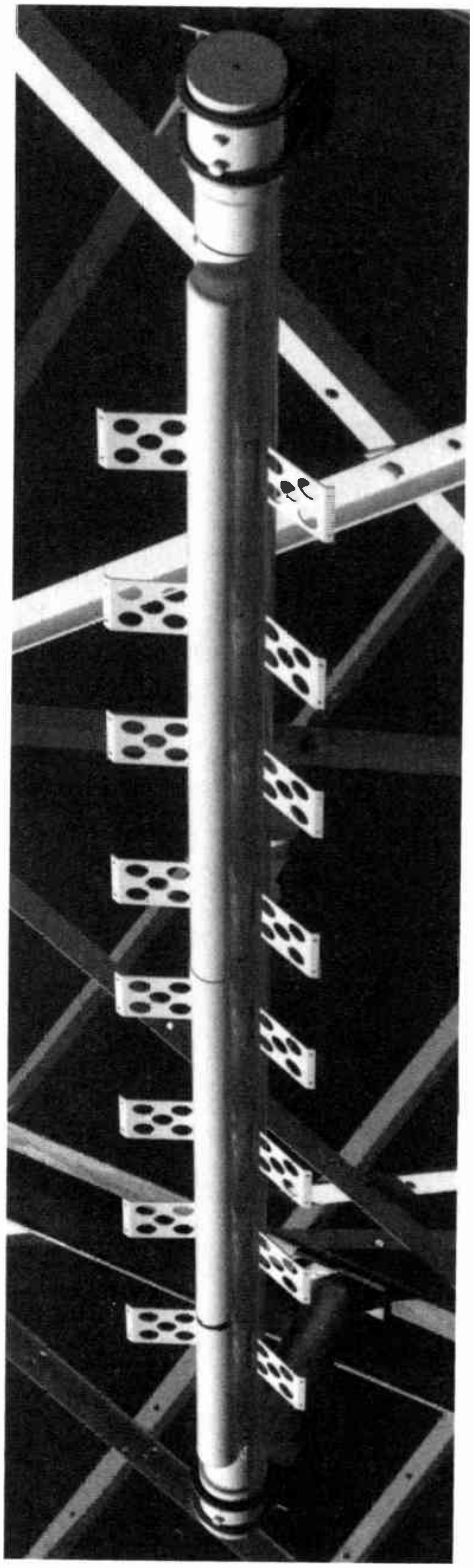


Figure 15. UHF slot array transmitting antenna. (Courtesy Andrew Corporation.)

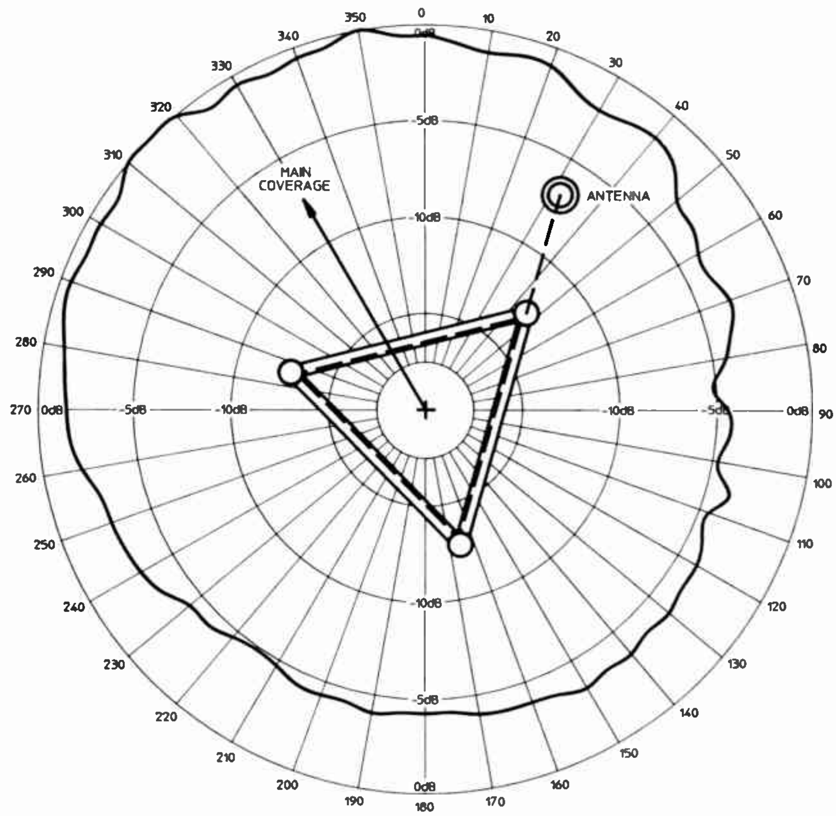


Figure 16. Calculated pattern of side mounted antenna. (Courtesy Dielectric Communications.)

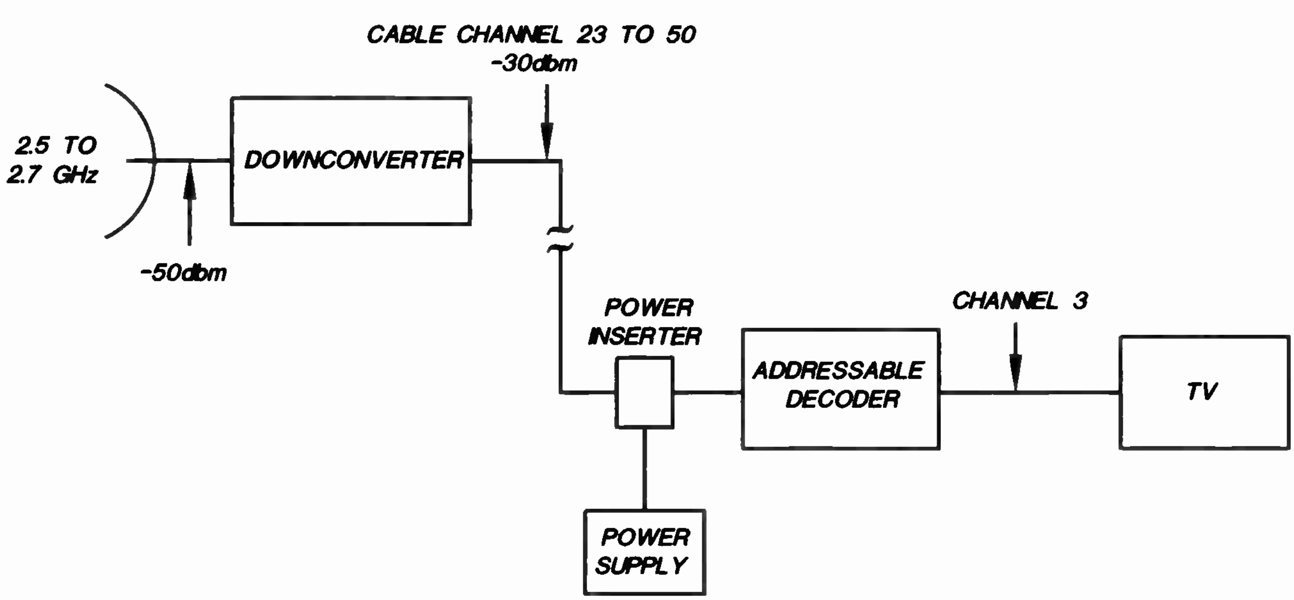


Figure 17. S-band receiver.

jumper usually connects the antenna feed directly to the downconverter. In some designs, the receive feedhorn is directly mounted to the downconverter. A typical antenna is shown in Fig. 18.

Downconverters

A typical downconverter is shown in Fig. 19. The primary important characteristics in downconverters are: sensitivity, signal handling capability, gain, spurious products, phase stability, and reliability. Sensitivity is achieved through the use of low noise transistor stages preceding the mixer. Noise figures of 4 dB are typical. Signal handling capability is a limitation in a multichannel environment. The front end mixer and the VHF output amplifier are both sources of distortion. This distortion takes the form of composite triple-beat, a common cable television distortion characteristic. This distortion can be minimized through improved mixer design and increased VHF amplifier output capability. It is common to test this characteristic with a two-tone input test signal. A reasonable objective to achieve is −27 dBm per tone at the input of the converter without significant output distortion.

The gain of a downconverter is generally not critical unless there are long transmission line feeders after the converter. A normal range for converters is 17 dB to 30 dB. Downconverters with as much as 40 dB of gain are available, but these units may suffer from output signal distortion because of high signal levels in the output amplifier.

Spurious products are a significant problem in many cases. Downconverters use various local oscillator techniques including phase-locked loops, dielectric resonators, free running VCOs, and SAW oscillators to achieve clean local oscillator signals. Spurious

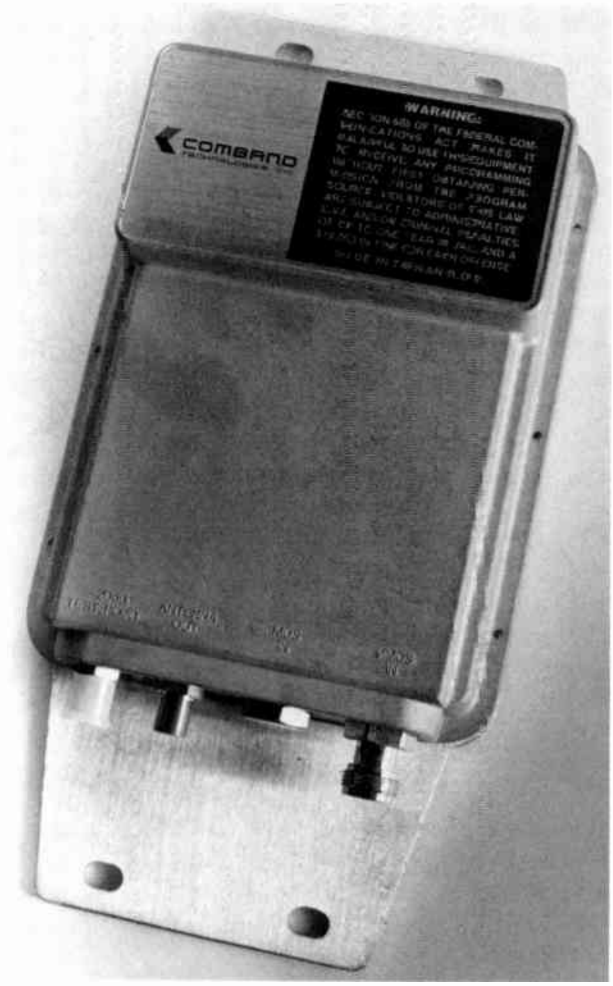


Figure 19. ITFS/MMDS downconverter. (Courtesy Comband Technologies, Inc.)

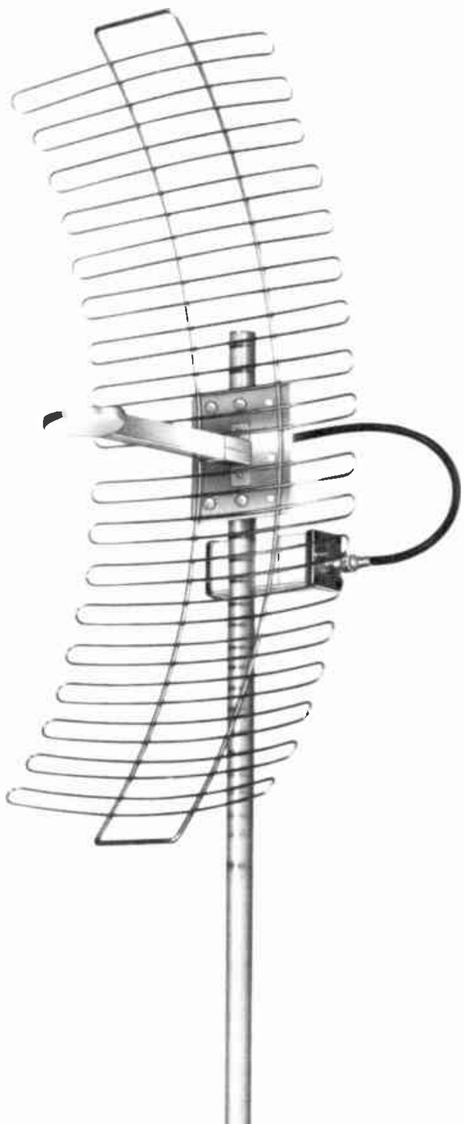


Figure 18. 21 dB gain ITFS/MMDS receiving antenna. (Courtesy Lance Industries.)

products in the output band should be suppressed at least 50 dB below normal output levels.

Phase jitter is a result of local oscillator electrical noise and mechanical vibration sensitivity. This characteristic leads to horizontal streaking in the picture and difficulty in phase-lock demodulating television signals. This characteristic can be observed on a spectrum analyzer tuned to an output unmodulated carrier.

Decoders

Decoders typically have synthesized front ends and specialized IF decoding circuitry. Addressability is a desirable characteristic of decoders but adds to the cost of the unit. Decoders designed for cable television service will occasionally have difficulty with the widely varying levels in an MMDS system and may be unable to handle the normal adjacent channel variations that it experiences. The time constant of an automatic gain control of the decoder may also be inappropriate for an MMDS signal environment. As MMDS systems grow, it is expected that new decoders developed specifically for this service will become available.

Figure 20. Set-top converter. (Courtesy Comband Technologies, Inc.)

Since decoders heterodyne convert or remodulate the signal onto a low band output VHF carrier, additional video and audio signal degradation may be incurred. A photograph of a typical television decoder is shown in Fig. 20.

INDEX

ABB. *See* Asea Brown Boveri (ABB)
ABES. *See* Association for Broadcast Engineering Standards (ABES)
ABU Review, 69
AC power systems
  assessing the lightning hazard, 1286–1287
  choosing a generator, 1293
  classifications of transient disturbances, 1284–1285
  critical system bus protection, 1294–1295
  discrete transient protection devices, 1290–1292
  high-performance isolation transformer, 1288, 1290
  motor-generator unit (MGU), 1288
  power factor, 1281–1283
  protection alternatives, 1285–1286
  protection using the systems approach, 1287–1290
  selective or secondary spot networks, 1290
  solid state line voltage filter, 1290
  standards of measurement of transient disturbances, 1285
  standby power systems, 1292–1295
  transient disturbances, 1283–1292
  uninterruptible power system (UPS), 1288, 1293–1294
  utility company interfacing, 1283
  utility power system, 1281–1283
ACGIH. *See* American Conference of Governmental Industrial Hygienists (ACGIH)
Acoustic gain devices, 760
Acoustics
  absorption characteristics of people, 1022
  absorption of sound, 1017–1023
  basic characteristics of sound propagation, 1013–1015
  cavity resonators, 1022–1023
  definition of, 1013
  definitions and sound measurements, 1014–1015
  diffusion, 1024
  directionality of sound sources, 1015
  echo control, 1018
  humidity's effect on sound level, 1015
  limitations of RT_{60}, 1026
  need for acoustical design in broadcasting, 1013
  noise control techniques, 1026–1031
  noise reduction, 1018
  nonparallel walls, 1017
  panel absorbers, 1022
  porous absorbers, 1018, 1022
  reverberation control, 1018
  room modes, 1015–1017
  sound decay in room, 1025–1026
  sound in enclosed spaces, 1015–1026
  sound level as function of distance in enclosed room, 1024–1025
  speed of sound, 1014
  studio acoustics, 999–1003
  types of absorbers, 1018–1023
Acronyms, 71–77
Addressability, 744–745
Administrative Procedures Act (APA), 7
Advanced Television Systems Committee (ATSC), 65
Advisory Notice (FCC Form 790), 16–22
Aeolian vibrations, 92
AES. *See* Audio Engineering Society (AES)
AFC. *See* Automatic frequency control (AFC)
AFCCE. *See* Association of Federal Communications Consulting Engineers (AFCCE)
Air contaminants, 50
Air-dielectric cable, 118
Aircraft marking, for towers, 83
ALJ. *See* Office of Administrative Law Judges (ALJ)
Allocations. *See* Frequency allocations
Allocations Branch, of the FCC, 5
AM Branch, of the FCC, 3
AM broadcast antennas
  for AM stereo broadcasting, 409
  antenna monitor system approval for AM directional antenna systems, 330
  authorizations for AM directional antenna systems, 330–332
  conditions for proper array performance, 191
  directional antenna system, 327–332
  electrical maintenance of, 187–188
  environmental changes, 190
  FCC form for license for AM directional antenna systems, 332
  FCC policy statement on AM directional antenna systems, 332
  graphical analysis of, 329–330
  ground system, 190
  license for AM directional antenna systems, 332
  lightning protection for, 188–190
  maintenance of, 187–193
  measurements of, 327–332
  mechanical maintenance of, 187
  monitoring points, 327
  partial proof, 328–329
  proof-of-performance, 327–328
  radiation hazards, 190–191
  readjustment data and analysis, 191–193
  readjustment of an array, 191
  reference proof-of-performance, 328
  routine maintenance of, 187–188
  soil conductivity and, 191
  special temporary authority for AM directional antenna systems, 330
  sudden component failures, 188–190
  towers for, 188
  transmission lines 188
AM broadcasting
  auxiliary and translator services, 7
  channel utilization in the U.S. prior to 1990, 37
  characteristics of, compared with FM broadcasting, 429–431
  clear channels, 36–37
  expanded band, 38
  FCC Broadcast Service Checklist, 18–22
  frequency allocations for, 35–38
  international and domestic classifications of, 37
  licensing of, 5–6
  local channels, 36–37
  power limitations by class of station, 37
  regional channels, 36–37
  RF radiation, 316–317
  studio facilities, 1005–1007
  switching for, 163
AM monitors, 611, 613
AM parallel transmitters, 172–176
AM power sampling, 577–578
AM signal-to-noise ratios, 467
AM stereo broadcasting
  antenna and phasor alignment, 409

audio chain, 405
audio interfacing, 411
equalization, 411–415
equalization path selection, 413
equalization procedure, 413
group delay equalization, 411–412
high-frequency equalization, 412–413
ISB AM stereo, 417–421
low-frequency equalization, 412
maintenance for, 415
monaural proof-of-performance, 410
Motorola C-QUAM AM stereo system, 422–427
phone lines, 407–408
processing, 405–407
proof-of-performance, 415
radio links, 408
RF interfacing, 410
station preparation for, 405–409
stereo exciter installation, 410–411
stereo synthesizers, 408
studio-to-transmitter link, 407–408
transmission audio processing for, 977–979
transmitter preparation, 408–409
X-Y oscilloscope, 413
AM stereo transmitters, 379
AM towers, and installation of semi-flexible cable, 133
AM transmitters
AM-stereo transmitter characteristics, 379
amplitude modulation theory, 353–354
audio frequency response, 374
audio phase linearity, 375–376
audio preemphasis, 380–381
basic AM systems, 356–361
carrier amplitude regulation (carrier shift), 375
carrier frequency generator/exciter, 370
carrier output power, 377–378
Chireix "outphasing" modulation, 356–357
circuitry for, 369
Class-B high level anode modulation, 357–358
control and monitoring of, 373
Doherty high-efficiency linear amplifier for high-power facilities, 358–359
factory tests for, 374–379
harmonic and spurious output, 376–377
high level anode modulation, 356
high-efficiency screen/impedance modulation, 361
high-voltage power supplies, 373–374
incidental phase modulation (IPM), 375, 379
international shortwave broadcast transmitters, 379–380
occupied bandwidth, 376
operating efficiency and input power, 378–379
performance requirements for, 374
phase noise, 379
power amplifiers, 354–356
preparation for AM stereo, 405–410
pulse-width high-level anode modulation, 359–361
radio frequency, 354–356
residual AM hum and noise, 374–375
residual PM hum and noise, 375
RF output networks, 370–373
RF power amplifiers, 370
solid-state broadcast transmitters, 361–363
solid-state equipment and circuitry, 363–369
stereophonic phase/gain equalization, 379
American Conference of Governmental Industrial Hygienists (ACGIH), 315
American National Standards Institute (ANSI)
description of, 68
radiation protection guide, 312–314
revision of RF radiation standard, 316
American Radio Relay League (ARRL), 68
American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE), 50
AMPFET transmitters, 365
Amplifiers
aural exciter, 510
cathode-driven triode amplifiers, 450
double-tuned tube power amplifiers for TV transmitter applications, 524
for FM transmitters, 445–449
grid-driven tetrode and pentode amplifiers, 451
intermediate power amplifiers (IPA), 446–449
power amplifier output circuits, 452–455
pulse and subcarrier distribution amplifiers, 812–813
solid-state television transmitters, 517
for studio audio, 959–961
tuner-amplifiers, 1143–1144
vacuum tube power amplifier circuits, 449–452
video distribution amplifiers, 811–813
Amplitude measurement, 1121–1123
Amplitude modulated (AM). *See* AM
Amplitude modulation theory, 353–354
Anchors, inspection and maintenance procedures, 99
Anchors, for towers, 94–96
ANSI. *See* American National Standards Institute (ANSI)
Antenna gain, 273–274
Antenna structures. *See* Towers
Antennas. *See* AM broadcast antennas; FM broadcast antennas; Satellite earth station antennas; Television broadcast antennas
APA. *See* Administrative Procedures Act (APA)
Appeals
of FCC rulings to federal courts, 9
of FOB rulings to FCC, 18–22
ARRL. *See* American Radio Relay League (ARRL)
ART (Additional Reference Transmission) color system, 1203
Asbestos, 51
Asea Brown Boveri (ABB) pulse-step modulator, 364
ASHRAE. *See* American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE)
Association for Broadcast Engineering Standards (ABES), 69
Association for Maximum Service Television (MSTV), 65
Association of Federal Communications Consulting Engineers (AFCCE), 60, 65
Associations, 65–69
Asynchronous AM, 467–468
ATSC. *See* Advanced Television Systems Committee (ATSC)
Audio, 69
Audio. *See* Studio audio
Audio cart machines, 963
Audio consoles, 956
Audio digitization, 863–866
Audio distribution and routing, 956–959
Audio Engineering Society (AES), 69, 874
Audio measurement
amplitude measurement, 1121–1123
audio test signals, 1134–1136
distortion measurement, 1126–1132
frequency analysis, 1123–1126
metering, 1122–1123
noise measurement, 1132–1133
phase monitoring and measurement, 1133–1134
real time analyzer, 1125–1126
swept frequency, 1124–1125
Audio processing in transmission. *See* Transmission audio processing
Audio recording
amplitude quantizing, 863–866
audio digitization, 863–866
compact disc format, 866–872
digital audio tape format, 872–874
digital audio technology, 863–876

digital signal processing (DSP), 875–876
discrete time sampling, 863
equalization, 849–850
erasable CD-E format, 870–872
future of analog recording, 853
historic perspective on, 845
mechanical alignment, 850–852
modulation noise, 853
need for bias, 847–848
noise reduction, 852–853
optimum bias, 848
playback head contour effects, 850
playback process, 848–849
recordable CD-R format, 870
recording system performance, 847
reference levels, 856–857
self erasure, 848
serial interfacing, 874–875
system performance, 852–853
tape and tape cartridges, 860
tape characteristics and types, 845–847
tape machine electronics, 853–856
tape machine maintenance, 857–860
tape machine performance comparison, 853
test tapes and alignment tools, 860–861
Audio Services Division, 3
Audio sources, 961–964
Audio standards for television signal transmission. *See* Television signal transmission standards
Audio systems, digital, 964
Audio tape machine
bias, 853
control, 854
electronics, 853–856
equalization, 853–854
head and tape path demagnetization, 859
head cleaning, 858–859
interfacing, 854–856
maintenance for, 857–860
performance comparison, 856
remote control, 855–856
tape path cleaning, 859
Aural broadcast auxiliary links
active repeaters, 677–678
antenna gain, 676–677
backup aural broadcast links, 678
cost of, 674–675
fading and fade margin, 676
frequency availability and coordination, 674
heterodyne repeater, 677–678
intercity relay (ICR), 674
licensing, 674
microwave booster, 678
passive repeater, 677
path analysis, 675–678
path attenuation between antennas, 675
path data calculation sheet, 677
preliminary system planning considerations, 674–675
problem path, 677
receiver sensitivity, 676
remodulating repeater, 677
studio-to-transmitter link (STL), 671–673
system requirements and equipment, 671–674
terrain considerations, 674
transmitter power (gain), 676
transmitter-studio links (TSL), 673–674
Aural broadcast auxiliary stations, 43
Aural exciter
audio processing circuits, 509–510
IF modulation, 510
RF amplifiers and diplexers, 510
Aural modulation monitoring
calibration, 614
coverage and, 609
essential features of aural modulation monitor, 611–613
FCC requirements for, 609–610
maintenance, 614
proof-of-performance, 609
reasons for, 609–610
without a "monitor" frequency, 614–615
Automatic frequency control (AFC), 441
Automatic power control (APC), 457
Auxiliary broadcast services
aural broadcast auxiliary stations, 43
frequency allocation, 42–44
licensing of, 6
Part 94 Operational Fixed Services, 43
remote pickup broadcast stations (RPUs), 42–43
RF radiation, 317
television broadcast auxiliary stations, 43–44
Auxiliary Services Branch, of the FCC, 3
Azimuth calculations, 1261
Azimuth pattern, 273, 275

Baffle effect, 755
Band-pass and band-stop filters, 165–69
BAS/CARS. *See* Broadcast Auxiliary Spectrum/Cable TV Relay Service (BAS/CARS)
Base insulators, 82, 99
BEA. *See* Broadcast Education Association (BEA)
Bias, 847–848, 853
Blumlein miking, 772
BOCA. *See* Building Officials and Code Administrators International (BOCA)
Boom boxes, 1143
Boost transformers, 88
Booster stations
FM broadcasting, 39
microwave booster stations, 43
television broadcasting, 40
Boosters
FM boosters, 445
low power television transmission systems (LPTV), 1309–1310
synchronous FM boosters, 443–444
Brewster angle, 240
Broadcast Auxiliary Spectrum/Cable TV Relay Service (BAS/CARS), 24–25
Broadcast coordination. *See* Frequency coordination
Broadcast Education Association (BEA), 69
Broadcast Engineering, 69
Broadcast facilities. *See also* headings beginning with Studio
air conditioning, 997
air distribution, 997–998
AM and FM studio facilities, 1005–1007
budget development checklist, 990–992
building codes, 989–990
building construction, 992–993
building planning, 990
control room/studio design, 1007–1008
cost estimation for, 1003–1005
electrical systems, 994–997
emergency systems, 996
engineering considerations, 993–994
fire alarm system, 995
grounding systems, 995, 997
lighting, 996–997
mechanical systems, 997–998
noise control, 997–998
planning, 989–1011
radio studios on low budget, 1008–1011
security, 996

site selection, 990
stand-by operation, 998
studio acoustics, 999–1003
studio production lighting system, 996
telephone system, 995
zoning regulations, 990
Broadcast Remote Pickup, 7
Broadcast Technology Society, 67
Broadcast Television Systems Committee (BTSC), 553–554
Broadcasting, 70
BTSC. *See* Broadcast Television Systems Committee (BTSC)
BTSC system
background of, 353
companding system, 559–561
compatibility with CATV systems, 565
compatibility with monophonic receivers, 564–565
compatibility with television translators, 565
incidental carrier phase modulation, 563–564
measuring and correcting ICPM, 564
modulation monitoring, 565–566
modulation summary, 561
monophonic channel, 556
performance factors in, 554
pilot carrier, 558
professional use channel, 561
separate audio program channel, 559
separation and modulation accuracy of, 612–613
stereo subcarrier, 557–558
synchronous amplitude modulation, 564
system requirements, 554–555
transmission format, 555–556
transmission of, 565
transmission requirements, 561–563
visual over-modulation, 563
Building codes
for broadcast facilities, 989–990
for towers, 89
Building Officials and Code Administrators International (BOCA), 89
Bureau of Land Management, 64

C-band satellites, 684–685
C-QUAM. *See* Motorola C-QUAM AM stereo system
Cable television
AML and FML microwave links, 739
amplifier distortions, 743
amplitude-modulated coaxial supertrunk, 739–740
baseband scrambling systems, 745
channel carriage capacity, 736–738
channelization, 738–739
characteristics of coaxial cable, 731
compatibility of BTSC system with, 565
composite triple-beat (CTB) interference in, 735
configurations and trends in, 736–739
consumer electronics compatibility, 746
distortion in, 741–744
fiber in, 746–748
fiber interconnect, 740
flat loss in, 735
frequency band usage regulation, 740
frequency bands affected by radio frequency interference, 743–744
frequency modulated coaxial trunk, 739
group delay through cable plant, 742
increasing channel capacity, 740
industry trends, 738
maintaining amplitude-versus-frequency response, 742
network design, 731–733
off-premises systems, 745
phase noise, 743
RF synchronization suppression systems, 745
scrambling and addressability, 744–745
signal and customer's equipment, 745–746
signal quality, 733–734
signal security systems, 744–745
signal splitting at customer's premises, 745–746
signal transportation systems, 739–740
system reflections, 742–749
trade-offs in, 734–735
upgrades (retrofitting) and rebuilds, 740
Cable Television Branch, of the FCC, 3
Cable Television Laboratories Inc. (CableLabs), 65–66
Camera lens
angle of view, 799
characteristics of, 798–803
color temperature, 803
depth of field and depth of focus, 801
11-step gray scale, 804
field of view, 800
focal length, 799
geometric distortion, 802–803
longitudinal chromatic aberrations and lateral chromatic aberrations, 802
minimum object distance (MOD), 803
modulation transfer function (MTF), 801
pick-up tube cameras, 803–804
ramping, 801
relative aperture (f-number or f-stop), 800
uniformity of field (vignetting) 801–802
zoom lens, 798–799
Cameras, television
anti-aliasing filter, 778
aperture correction, 1184
basic camera system, 776–777
CCD solid-state imager, 780–786
color camera alignment, 1183–1186
color correction and neutral density filters, 777–778
color matching techniques, 1184–1185
color separation system, 778–779, 797–798
color TV film chains, 1186
digital camera, 793–794
electronic field production (EFP), 796–797
electronic news gathering (ENG), 797
flare in pickup tubes, 1185
gamma correction, 1184
image diagonal, 779
imaging devices, 779–786
infrared filter, 778
introduction to, 775–776
lens characteristics, 777, 798–803
lighting for color cameras, 1185
low-light pickup, 1186
optical filters, 777–778
optical system, 777–779
outdoor broadcast pickup, 1185
performance verification, 794–795
pick-up tube, 780
quarter-wavelength filter, 778
robotics for, 981–988
selection considerations, 795–797
signal processing, 786–794
signal-to-noise and sensitivity, 1184
specular highlights, 1184
specular reflections, 1185–1186
studio remote applications, 795–796
three-tube concept, 1183–1184
Canadian method, for distance calculations, 1260
Candelabras, 94
Captioning systems. *See* Closed captioning systems
Car radios, 1144–1145
Carbon microphones, 749
Cardioid microphones, 756–757

Cascading studios, 819
Case effect, 755
Cathode-driven triode amplifiers, 450
CATV. *See* Cable television
CAV. *See* Component analog video (CAV)
CCB. *See* Common Carrier Bureau (CCB)
CCD line array, 938–939
CCD solid-state imager
  basics of, 781–782
  compared with pick-up tubes, 780–781
  on-chip lens array (OCL), 785–786
  spatial offsetting, 785
  transfer mechanism, 782–785
CCIR. *See* International Radio Consultative Committee (CCIR)
CCIR Aural Scale, 339
CCIR Comparison Scale, 339
CD. *See* Compact disc (CD)
CD-R format, 5.5–26
CDRB. *See* Committee for Digital Radio Broadcasting (CDRB)
CEC. *See* Continental Electronics Corporation (CEC)
Cellular telephones, 1089–1090
Centrex, 1089
CERCLA. *See* Comprehensive Environmental Response, Compensation and Liability Act (CERCLA)
CFCs. *See* Chloroflorocarbons (CFCs)
CFR. *See* Code of Federal Regulations (CFR)
Charge Coupled Device. *See* CCD
Chilton's ECN Electronic Component NEWS, 70
Chireix "outphasing" modulation, 356–357
Chloroflorocarbons (CFCs), 55
Choke networks, 1087–1088
Circularly-polarized "V" (CPV) antennas, 290
Clarke spheroid versus WGS ellipsoid, 1261–1262
Clear channels, for AM broadcasting, 36–37
Closed captioning systems
  automatic live encoding, 1245
  caption data encoders, 1240–1241
  cautionary notes, 1248–1249
  data transmission format, 1246–1247
  definition of, 1237
  display format, 1240
  emergency messages, 1249
  FCC standard, 1238–1239
  history of, 1237–1238
  line 21 technical advisory, 1247–1249
  live captioning, 1242–1245
  live-display captions, 1244
  methods of creating caption data, 1241–1246
  newsroom computer captions, 1244–1245, 1248–1249
  prerecorded captioning, 1241–1242
  real-time captions, 1242–1244
  simple encoder, 1240–1241
  smart encoder, 1240
  text service, 1245–1246
Coaxial cable
  attenuation, 119
  cut-off frequency, 119–120
  design considerations of, 118–120
  impedance of, 118–119
  power handling, 118–119
Coaxial patch panels, 161–162
Coaxial switches, 162
Code of ethics, for consultants, 60
Code of Federal Regulations (CFR), 50, 64–65
Color encoder, 1180–1183
Color fidelity
  burst amplitude error, 1170–1171
  burst phase error, 1170
  carrier unbalance, 1171
  characteristics of the eye, 1161–1164
  color adaptation, 1163–1164
  color system analysis, 1161
  distortions in transmission system, 1172–1179
  encoding and decoding distortions, 1169–1172
  envelope delay, 1173–1175
  gain characteristic, 1172
  gain stability of M, I, and Q transmission path, 1170
  matrixing process errors, 1169–1170
  modulation and demodulation, 1170
  modulation errors, 1170–1172
  nonlinearities of practical transmission system, 1179
  perfectly linear transmission system, 1172
  phase characteristic, 1172–1173
  possible errors in transducers, 1164–1166
  primary colors, 1164
  reference white, 1164
  state of the art, 1179–1180
  subcarrier frequency error, 1171–1172
  transfer characteristics, 1166–1168
  two-phase modulation errors, 1171
  video unbalance, 1171
Color picture tubes, 1229–1231
Color separation, 778–779, 797–798
Color television
  ART (Additional Reference Transmission) color system, 1203
  background on, 1193–1194
  band shaping, 1155–1156
  camera alignment, 1183–1186
  color encoder, 1180–1183
  color fidelity, 1161–1180
  color receiving system, 1160–1161
  color test equipment, 1187–1190
  color-frequency standards, 1158
  comparison of systems standards and specifications, 1204–1206
  compatibility, 1149
  displaying RGB signals, 1154
  electronic aspects of, 1155–1161
  FCC time-base specifications, 1191
  frequency interface, 1157–1158
  fundamentals, 1149–1154
  generating RGB signals, 1152–1154
  international exchange of images, 1206–1208
  matrixing, 1155
  monochrome compatible color TV systems, 1194
  monochrome television system, 1151–1152
  NIR color system, 1203
  NTSC color system, 1195–1199
  overall color system, 1158–1159
  PAL color system, 1194–1195, 1199–1201
  primary colors, 1152
  radiated picture signal, 1151–1152
  receiver, 1152
  resolution and bandwidth, 1150
  scanning, 1150
  SECAM color system, 1194–1195, 1202–1203
  sine-squared pulse and bar, 1188
  stair-step generator, 1187–1188
  standards, 1193–1208
  synchronizing, 1151
  test charts, 1190
  three variables of color, 1152–1154
  two-phase modulation and generation of color subcarrier, 1156–1157
  vectorscope, 1188–1189
  vertical interval reference and test signals, 1189
Color television receivers
  advanced audio systems for "full feature" receivers, 1219–1221
  audio demodulation and processing, 1218–1221

basic monaural systems, 1218–1219
block diagram, 1210–1211
color picture tubes, 1229–1231
comb filter applications, 1225–1229
deflection control, 1232–1234
dynamic noise reduction (DNR), 1221–1222
general design considerations, 1209–1210
IF (intermediate-frequency) section, 1215–1218
introduction, 1209–1211
luminance and chrominance processing, 1222–1225
power supply and protection systems, 1234–1236
sync separation, 1232
television tuners, 1211–1215
video drivers, 1231–1232
Color temperature, 803, 788–797
Comb filter, 1225–1229
Committee for Digital Radio Broadcasting (CDRB), 66
Common Carrier Bureau (CCB), 3
Communications Act of 1934, 1, 35
Community and multi-user antennas, 258–260
Community Right to Know Law, 52
Compact disc (CD)
data decoding, 868–870
data readout, 867–868
EFM encoding and frame assembly, 867
encoding, 866
erasable CD-E format, 870–872
physical design, 866
recordable CD-R format, 870
subcode data, 866–867
Compact disc (CD) players, 961–962
Complaints and Investigations Branch, 5
Component analog video (CAV), 819–821
Composite triple-beat (CTB) interference, 735
Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), 52
Compression insulators, 82
Computers, *See* headings beginning with Microprocessor
Condenser microphones, 751–752
Conductor sizes, 88
Conduit sizing, 88
Consulting services
code of ethics, 60
data rights and, 61
for FCC applications, 61
qualifications, registration, and certification, 59–60
range of, 59, 61–62
selecting a consulting engineer, 60
working with a consulting engineer, 60–61
Continental Electronics Corporation (CEC) pulse-step modulator, 364–365
Conversion algorithms between datums, 1263
Corrosion, 82, 97–98
CPV antennas, *See* Circularly-polarized "V" (CPV) antennas
CTB, *See* Composite triple-beat (CTB) interference

D-1 component digital VTR, 891–898
D-2 composite digital VTR, 898–900
DAT, *See* Digital audio tape (DAT)
db Magazine, 70
DBS, licensing of, 5, 7
DDS, *See* Direct digital synthesis (DDS)
Deicing, *See* Icing protection
Design standards, for electrical systems for towers, 87–89
Digital 950 MHz STLs, 672
Digital audio systems, 964
Digital audio tape (DAT), 872–874
Digital audio tape (DAT) recorder, 873–874
Digital camera, 793–794
Digital recording, 828
Digital signal processing (DSP), 875–876
Digital telephone systems, 1088–1089
Digital video systems, 821–823
Diplexers, 168-170, 172–173
aural exciter, 510
FM broadcast antennas, 260–263
Direct digital synthesis (DDS), 442
Directional microphones, 755–757
Dispatch and aural remote pickup, 25–26
Distance and bearing calculations
azimuth calculations, 1261
Canadian method for, 1260
Clarke spheroid versus WGS ellipsoid, 1261–1262
conversion algorithms between datums, 1263
FCC flat-earth method, 1259–1261
Mexican method for, 1260–1261
NAD27/NAD83 datums, 1262
rounding practices, 1262
spherical-earth method, 1261
Distortion
amplifier distortions, 743
audio harmonic distortion measurement, 466–467
cable television, 741–744
chrominance nonlinear distortion, 633, 637–638
color fidelity, 21–31
field-time waveform distortion, 627
gain-frequency distortion, 633–634
geometric distortion in camera lens, 802–803
harmonic distortion, 1126–1127
intermodulation distortion (IMD), 245, 1127–1132
line-time waveform distortion, 627–629
luminance nonlinear distortion, 633, 635–636
measurement of, 466, 1126–1132
short-time waveform distortion, 629–631
total harmonic distortion, 466–467
transient intermodulation distortion, 975
Distribution amplifiers
studio audio, 959–962
video, 811–813
Distribution Service Branch, of the FCC, 3
DNR, *See* Dynamic noise reduction (DNR)
DOL, *See* U.S. Department of Labor (DOL)
Doppler radar, 1252–1253
Drinking water coolers, 51
DS-1 based multiplexers, 673
DSP, *See* Digital signal processing (DSP)
Dummy loads, 271
Dynamic noise reduction (DNR), 1221–1222

EAN, *See* Emergency Action Notification (EAN)
Earth stations, *See* Satellite earth stations
EBS, *See* Emergency Broadcast System (EBS)
EBS Advisory Committee (EBSAC), 1265
EBSAC, *See* EBS Advisory Committee (EBSAC)
EBU, *See* European Broadcasting Union (EBU)
EBU Review—Technical, 70
Editing, *See* Videotape editing
EE Evaluation Engineering, 70
EEPROM, 590–591
Effective radiated power (ERP)
definition of, 273
of FM broadcast antennas, 244
EIA, *See* Electronic Industries Association (EIA)
18 GHz STLs, 672–673
"Electrical Safety-Related Work Practices" (ESWP) rule, 53
Electrical systems for broadcast facilities, 994–997
Electrical systems for studio lighting, 1045–1048
Electrical systems for towers
AC utility circuits, 88
antenna deicing circuits, 87–88
boost transformers, 88

circuits for isolated antenna structures, 88
conductor sizes, 88
conduit sizing, 88
design criteria for, 88
elevator control circuits, 88
and future antennas or height extensions, 94
junction and pull box sizing, 88
lighting, 84
power at platforms for communications systems, 88
sound-powered telephone, 88
tower erection and, 96
Electromagnetic radiation, 52. *See also* Radiation
Electronic Buyers' News, 70
Electronic Engineering Times, 70
Electronic field production (EFP), 1073
Electronic Industries Association (EIA), 66–67, 88–89, 380
Electronic Media, 69–70
Electronic news gathering (ENG). *See also* Radio electronic news gathering and field production
camera selection for, 797
frequency coordination and, 24, 28–29
guidelines for, 28–29
7 GHz band, 29
for television, 1073
13 GHz band, 29
Electronic telephone systems, 1090–1093
Electronics Design, 70
Electronics Industry Association (EIA), 553–554
Electroplating, of towers, 83
Elevation pattern, 273, 274–275
Elevator control circuits, 88
Elevators, for towers, 83-84, 99
Eleven-step gray scale, 804
Emergency Action Notification (EAN), 1265–1267
Emergency Broadcast System (EBS)
checklist for nonparticipating stations, 1273–1277
checklist for participating stations, 1270–1273
national advisory committees, 1277–1279
national level activation and termination, 1266–1268
operational (local) area operation, 1269–1270
state level operation, 1268–1269
system operation, 1266–1270
system organization, 1265–1266
Enforcement Division, of the FCC, 3, 5, 11
ENG. *See* Electronic news gathering (ENG)
Engineering Division, of the FCC, 11
Engineering Policy Branch, of the FCC, 5
Engineers
professional engineers, 59–60
selecting a consulting engineer, 60
working with a consulting engineer, 60–61
Ennes Educational Foundation, 68
Environmental concerns
air contaminants, 50–51
asbestos, 51
drinking water coolers, 51
electromagnetic radiation, 52
federal agencies for, 49
hazard communications standard, 52–53
information sources on, 49–50
lockout/tagout, 53
occupational noise exposure, 53
polychlorinated biphenyls (PCBs), 53–54
protection of stratospheric ozone, 54–55
radon, 55–56
underground storage tank (UST) regulations, 56–57
Environmental Protection Agency (EPA), 49-57, 64, 315
EPA. *See* Environmental Protection Agency (EPA)
EPROM, 589-590
Equal Employment Opportunity Branch, of the FCC, 5
ERP. *See* Effective radiated power (ERP)
ESWP rule. *See* "Electrical Safety-Related Work Practices" (ESWP) rule
Ethics code, for consultants, 60
European Broadcasting Union (EBU), 381, 875
European Independent Broadcasting Authority (IBA), 941
Ex parte considerations, 9
Expanded band, 38
Eyes, color perception of, 1161–1164

FAA. *See* Federal Aviation Administration (FAA)
Facilities. *See* Broadcast facilities
Fairness/Political Programming Branch, of the FCC, 5
FCC. *See* Federal Communications Commission (FCC)
FCC flat-earth method, 1259–1261
Federal Aviation Administration (FAA), 26, 49, 64, 83, 84
Federal Communications Commission (FCC)
Audio Services Division
AM Branch, 3
Auxiliary Services Branch, 3
FM Branch, 3
Common Carrier Bureau (CCB), 3
current copy of broadcast rules, 64–65
Enforcement Division
Complaints and Investigations Branch, 5
Equal Employment Opportunity Branch, 5
Fairness/Political Programming Branch, 5
Hearing Branch, 5
Field Operations Bureau (FOB), 3, 11–17
appeals, 13
Broadcast Service Checklist, 18–22
criminal penalties, 13
Enforcement Division, 11
Engineering Division, 11
field offices, 11
FM Broadcast Signal Analysis Report (FCC), 14
headquarters, 11
inspections, 12
local offices, 11–12
monetary forfeitures, 13
off-the-air monitoring, 12
Public Service Division, 11
regional directors, 11
Television Signal Analysis Report (FCC), 15
violations, 12–13
frequency allocations and, 35–36
licensing process
application for the station license, 6
auxiliary and translator services, 7
contested applications, 6
DPS, 7
first step in, 6
LPTV, 7
renewal of license, 6
sales and transfer, 6–7
satellite earth stations, 687
station construction and equipment testing, 6
types of facilities, 5–6
Mass Media Bureau (MMB)
Audio Services Division, 3
Enforcement Division, 3
organization of, 3–5
Policy and Rules Division, 5
Video Services Division, 3
membership of, 1
Office of Administrative Law Judges (ALJ), 3, 6
Office of General Counsel (OGC), 1
Office of Legislative Affairs (OLA), 1
Office of Plans and Policy (OPP), 1
Office of Public Affairs (OPA), 1
Office of Science and Technology (OST), 3

Office of the Managing Director (OMD), 1
organization of, 5
organizational chart, 2
Policy and Rules Division
Allocations Branch, 5
Engineering Policy Branch, 5
International Branch, 5
Legal Branch, 5
Policy Branch, 5
Private Radio Bureau (PRB), 3
Review Board (RB), 3
rule making process
appeal to the courts, 9
chart of, 8
early options, 7, 9
ex parte considerations, 9
initiation of actions, 7
Notice of Proposed Rule Making, 9
petition for reconsideration, 9
Report and Order, 9
rules and regulations, 7
statutory authority for, 1, 35
transmitter control systems, 604–607
Video Services Division
Cable Television Branch, 3
Distribution Service Branch, 3
Low Power Television Branch, 3
Television Branch, 3
Federal Emergency Management Agency (FEMA), 1265
Federal government. *See* names of specific federal departments and agencies
Federal Labor Standards Act, 64
Federal Radio Commission, 35
Federal Register, 50
FEMA. *See* Federal Emergency Management Agency (FEMA)
Fiber news gathering, 727–728
Fiber optic transmission systems
advantages of, 717–718
applications of, 727–730
attenuation, 724
broadband digital transmission services, 721–722
chronology of development of, 719
components, 724–725
configurations of, 718–719
design considerations, 726–727
dispersion, 723
fiber cable specifications, 727
fiber optic circuit designs, 728–729
high performance system, 729–730
light and fiber characteristics, 722–724
light detectors (receivers), 721–724
light sources (transmitters), 724
lightwave transmission theory, 722
modulation techniques, 725–726
multiple analog channel (MAC), 720
multiple digital channel, 721
multiplexing techniques, 726
performance requirements, 726–727
principles of, 717–722
remote studio interconnection, 728
repeaters, 724–725
simple low-cost system, 729
single analog channel (SAC), 719–720
single digital channel, 720–721
switched fiber systems for news gathering, 727–728
switching, 725
synchronous optical network (SONET), 722
T and DS series hierarchy, 721–722
technology principles, 722–727
testing optical circuits, 727
transmission channel options, 719–721
types of fibers, 723–724
video and audio transmission techniques, 725–726
Field Operations Bureau (FOB)
appeals, 13
Broadcast Service Checklist, 18–22
criminal penalties, 13
Enforcement Division, 11
Engineering Division, 11
field offices, 11
FM Broadcast Signal Analysis Report (FCC), 14
headquarters, 11
inspections, 12
local offices, 11–12
monetary forfeitures, 13
off-the-air monitoring, 12
Public Service Division, 11
regional directors, 11
responsibilities of, 11–13
Television Signal Analysis Report (FCC), 15
violations, 12–13
Field production. *See* Radio electronic news gathering and field production; Television field production
Field strength. *See* FM broadcast field strengths; Television broadcast field strengths
Film for television
advantages of, 935
broadcast telecine, 935
CCD line array, 938–939
cement splices, 944
characteristics of, 941–942
cleaning, 943–944
CRT flying spot scanner, 937–938
film-to-tape transfer, 939–940
film-video interface definitions, 945
future of, 944–945
handling and storage, 943–944
introduction to, 933
monitor setup, 941
photoconductive telecine, 936–937
production telecine, 935
programmable color correction, 942
sample films, 946
sound for, 942
splicing, 944
standards, 945
tape splices, 944
telecine design, 935–939
telecine set up, 940–941
transfer telecine, 935
Filterplexing, 463–465
Filters
band-pass and band-stop filters, 165–168
comb filter, 1225–1229
diplexers, 168–170, 172–173
group delay correctors, 168
harmonic filters, 164–165
harmonic notch filters, 459
low power television transmission systems (LPTV), 1311
multiplexers, 168–171
optical filters, 777–778
rejection filters, 250
requirements for, 164
RF output low-pass filters, 458–459
SAW filters, 1215–1216
special filtering for UHF, 183
Fire alarm system, for broadcast facilities, 995
Flame spraying, of towers, 81
Flashlamp, 86–87
Flat loss, in cable television, 735
Flat-earth method, 1259–1261
FM booster stations, 40

FM boosters
  leased phone lines with, 445
  radio link used with, 445
  synchronous, 443–444
FM Branch, of the FCC, 3
FM broadcast antennas
  accessory equipment for, 271
  air pressurization, 267–268
  antenna gain, 243
  antenna supporting mast, 347
  beam tilt, 244
  blanketing, 243
  Brewster angle, 240
  cavity backed panel antennas, 254
  characteristics of, 243–246
  checking system VSWR, 245
  circular polarization, 246
  Class A station coverage, 246–247
  Class B and C-2 station coverage, 247–248
  commercially available antennas, 250–255
  coverage standards, 242–243
  crossed dipole theory, 254–255
  directional antennas, 245–246
  dummy loads, 271
  effective radiated power (ERP), 244
  element spacing and downward radiation, 255
  evaluation of system performance, 341–342
  FCC service contours, 242
  flat panel antennas, 253–254
  FM SCA multiplexing, 270
  FM signal measurements, 242–243
  Fresnel zone clearance, 240–241
  general features of, 237
  ground reflections, 239
  harmonic filters, 270–271
  high-gain antenna contradictions, 249
  horizontal and vertical polarization, 246
  icing protection, 268–269
  importance of low VSWR, 245
  installation on AM towers, 263–264
  installation procedures, 264–266
  intermodulation distortion, 245
  lightning protection, 269–270
  linear height gain effect, 242
  matching coverage and antennas, 246–248
  matching transmitter power and antennas, 248–249
  multipath problems, 238–239
  null fill, 244
  pattern optimization, 256–258
  performance of, 253
  polarization, 246
  pole mounting, 266
  problems with sidemounted antennas, 255
  propagation, 237–242
  propagation loss, 238
  radomes on, 269
  receiving antennas, 345–347
  recommended minimum antenna heights, 241
  reflectometers, 271
  required signal strengths, 243
  RF switches, 271
  ring radiators, 251
  ring-stub antennas, 251–252
  series fed antennas, 253
  series-fed V-dipole antennas, 251
  short helix antennas, 252
  short helix—multi-arm antennas, 253
  shunt fed slanted dipole antennas, 252
  site selection, 249
  soil conductivity, 241
  spurious frequencies, 270
  standard sidemount antennas, 251
  station diplexing, 260–263
  structural considerations, 266
  synchronous AM (SAM) noise, 245
  transmission line systems, 266–267
  TV channel 6/FM antenna problem, 249–250
  twisted ring antennas, 252
  VSWR bandwidth, 244–245
  wideband community and multi-user antennas, 258–260
FM broadcast field strengths
  analysis of measurements to depict coverage, 339
  antenna supporting mast, 347
  antenna system performance evaluation, 341–342
  equipment principles, 343–347
  FCC standard method for collection of propagation data, 348
  FCC standard method for measurement of coverage of communities, 348–349
  measurement of, 335–351
  measurement of coverage, 336–339
  measurements versus predictions, 335–336
  measuring procedures and techniques, 347–349
  planning a measurement program, 336–341
  power supplies, 347
  practical field strength meters, 343–345
  practical problems in measurement of, 349
  propagation, 341, 343
  receiving antennas, 345–347
  recorder, 347
  safety considerations in measurement of, 349–351
  spurious emission measurement, 341, 343
  time variation of field strength, 349
  transmitting system evaluation, 340–341
FM Broadcast Signal Analysis Report (FCC), 14
FM broadcast transmitters
  angular modulation, 431
  bandwidth limitation, effects of, 433–434
  Bessel nulls, 432–433
  FM exciters, 438–443
  FM sideband structure, 431–432
  frequency modulation theory, 431–436
  group delay response symmetry versus amplitude response symmetry, 434
  improvement of RF path, 434
  limiting factors within, 434
  matching transmitter power and FM broadcast antennas, 248–249
  occupied bandwidth, 433
  preemphasis, 436–437
  purpose of, 438
  subsystems of, 438
  transmitter power output (TPO) requirements, 437–438
FM broadcasting
  auxiliary and translator services, 7
  booster stations, 40
  characteristics of, compared with AM broadcasting, 429–431
  classifications of stations, 38–39
  commercial, 5–6
  FCC Broadcast Service Checklist, 18–22
  FCC FM Broadcast Signal Analysis Report, 14
  frequency allocations, 38–40
  history of, 429
  licensing of, 5–6
  minimum distance separation requirements, 40
  noncommercial, 5–6
  RF radiation, 317
  satellite earth stations and, 702
  studio facilities, 1005–1007
  translator stations, 40
  transmission audio processing for, 977
FM exciters
  automatic frequency control, 441

direct digital synthesis, 442
direct FM, 439
FM modulator linearity, 439–441
general discussion of, 438–439
hybrid splitting of exciter power, 462–463
metering, 442
output stage, 443
packaging, 442–443
phase-locked loop automatic frequency control, 441–442
FM monitors, 611–613
FM required signal strengths, 243
FM signal measurements, 242–243
FM signal-to-noise ratio, 467
FM squared satellite, 603
FM stereo broadcasting
crosstalk from main channel program into SCA subchannels, 496–497
digital stereo generation, 486–487
FMX stereo noise reduction, 486
frequency division multiplexing, 477–484
frequency spectrum and modulation limits, 476–477
generating the stereophonic baseband signal, 477–487
glossary of terms for, 475–476
performance of, 492–497
SCA transmission, 488–491
SCA-to-stereo interference considerations, 494–496
stereo decoder circuits, 487
time division multiplex, 484–486
FM translator stations, 40
FM transmitters
AM signal-to-noise ratios, 467–468
asynchronous AM measurement, 467–468
audio frequency response measurement, 466
audio harmonic distortion measurement, 466–467
automatic power control (APC), 457
automatic power control headroom, 470
booster systems, 445
built-in AM noise measurement capability, 468
carrier frequency measurement, 465
centering the passband, 470
combined transmitters, 461–463
control systems, 457–458
controller back-up systems, 458
cooling of, 472
deviation calibration, 445
effect of tuning on FM sidebands, 470–471
effects of circuit topology and tuning on FM modulation performance, 455
efficiency versus optimum tuning, 471
filterplexing, 463–465
FM modulation measurement, 465
FM signal-to-noise ratio, 467
folded-half-wavelength cavity, 453–454
hybrid combiners, 461–462
hybrid couplers, 461
hybrid splitting of exciter power, 462–463
initial tuning, 469
installation of, 471–472
interference correction, 445
intermediate power amplifiers (IPA), 446–448
logging of transmitter operating parameters, 466
maintenance for, 472–473
microprocessor control systems, 457–458
modulation effects, 444–445
operational measurements, 465–469
output coupling, 454–455
output loading, 470
output power sampling, 577
output tuning, 469–470
plant maintenance for, 472–473
power amplifier cavity, 453
power amplifier output circuits, 452–455
power amplifier source impedance, 455
power output requirements, 445–446
power supplies, 455–457
preventive maintenance for, 473
proof-of-performance measurements, 466
quarter-wavelength cavity, 453
remote control interfacing, 458
RF intermodulation, 463–464
RF output low-pass filters, 458–459
RF power amplifiers, 445–446
RF power output measurement, 465–466
solid-state FM broadcast transmitters, 449
standing wave ratio (SWR), 459–460
step-start, 457
synchronous AM measurement, 468–469
synchronous FM boosters, 443–444
time delay equalization, 444–445
transmission line power and SWR measurements, 459–461
tuning for best modulation performance, 469–471
turn-around-loss, 464
two-tone intermodulation distortion measurement, 467
vacuum tube power amplifier circuits, 449–452
VSWR measurement, 461
wiring the transmitter plant, 471–472
FMX stereo noise reduction, 486
FNPRM. *See Further Notice of Proposed Rule Making* (FNPRM)
Foam-dielectric cable, 118
FOB. *See* Field Operations Bureau (FOB)
Foreign exchange (FX) loops, 1087
Forest Service, 63
Foundations, for towers, 94–96
Free space propagation, 299–300
Frequency allocations
AM broadcasting, 35–38
aural broadcast auxiliary stations, 43
auxiliary broadcast services, 42–44
coordination in the allocation and use of spectrum, 33–35
definition of allocation, 33
domestic process of, 35
FM broadcasting, 38–40
international allocation, 33
Part 94 operational Fixed Services, 43
process of, 34–35
remote pickup broadcast stations (RPUs), 42–43
technical submissions, 33–34
television broadcast auxiliary stations, 43–44
television broadcasting, 40–42
Frequency analysis, 1123–1126
Frequency coordination
aural studio-to-transmitter link (STL) bank, 27
dispatch and aural remote pickup, 25–26
high band VHF and UHF coordination, 26–27
history of, 23–24
interference and the Part 74 licensee, 31–32
for major events, 30
950 band channel stacking, 27
950 band recent developments, 27
942–944 band segment, 27
non-RF alternatives, 27
operational guidelines for television ENG operations, 28–29
process of, 24–25
propagation characteristic, 24
rationale for, 23
7 GHz Band, 29
television microwave considerations, 27–28
television STL coordination, 29–31
13 GHz Band, 29
various broadcast uses of Part 74 spectrum, 25–27
wireless microphones, 27

Frequency modulation recording, 828–830
Frequency modulation theory, 431–435
Frequency monitoring
calibration, 614
coverage and, 609
essential features of frequency monitors or meters, 610–611
FCC requirements for, 609–610
maintenance, 614
off-site monitoring techniques, 614
on-site monitoring techniques, 613
peak modulation duration, 611
proof-of-performance, 609
reasons for, 609–610
without a "monitor" frequency, 614–615
Fresnel zone clearance, 240–241
Fresnels, 1037
Further Notice of Proposed Rule Making (FNPRM), 9

Galloping, 92
Genlock, 816
Government Printing Office (GPO), 50
Government Printing Office (GPO) Bookstores, 50
GPO. *See* Government Printing Office (GPO)
Gray scale, 804
Grounding system
for AM broadcast antennas, 188
bonding and grounding recommendations, 113
for broadcast facilities, 995, 997
external ground system, 106–109
formulas for calculation of resistance to ground, 107
ground electrode testing, 106–108
grounding resistance, 108–109
isolated ground receptacle, 113
low power television transmission systems (LPTV), 1304–1305
for semi-flexible cable, 132–133
single point grounding, 110–111
tower-building system, 109–111
Group delay correctors, 168
Guy anchors, 95, 99
Guy cables 263–264
Guy installation, for towers, 97
Guy insulators, 82, 98–99
Guyed towers, 80–81, 95–97
Guys
clips for, 81–82
connections for, 81–82, 98
corrosion of, 98
dead end grips for, 81
inspection and maintenance procedures, 98–99
materials for, 81–82
nonconductive, 264–265
serving for, 82,
sockets for, 81
swaged sleeves for, 82
tensions of, 98–99
wedge type sockets for, 82

Halons, 55
Harmonic distortion, 1126–1127
Harmonic filters, 164–165, 270–271
Harris DX system, 367–369
Hazard communications standard, 52–53
HCFC-123, 55
HDTV. *See* High definition television system (HDTV)
Headset microphones, 759
Hearing Branch, of the FCC, 5
Heaters, electrical, as icing protection, 268
Heising constant current modulation, 356
Helical scan recorders
capstan and scanner servos, 882
definitions concerning, 878–879
digital recording systems, 880
error protection strategy, 881
frequency modulation, 879–880
helical recorders, 881–882
recording and playback systems, 880–881
signal system, 879–880
tape tension servos, 882
tape transports, 881–882
video head position servo, 882
Heterodyne conversion, 1211–1212
High band VHF and UHF coordination, 26–27
High definition television (HDTV), 41, 830, 944–945
Hilbert transform clipper, 968
Home and Studio Recording, 70
Home tuners and tuner-amplifiers, 1143–1144
Horns, 760
Hot-dip galvanizing, of towers, 82–83
Hybrid combiners, 461–462
Hybrid couplers, 461

IBA. *See* European Independent Broadcasting Authority (IBA)
Ice load, 91
Icing protection
deicing circuits for antennas, 87–88
electrical heaters, 268
radomes, 269
for towers, 83
ICR. *See* Intercity relay (ICR)
IEEE. *See* Institute of Electrical and Electronic Engineers (IEEE)
IES. *See* Illumination Engineering Society of North America (IES)
IF (intermediate-frequency) system
applications of SAW filters, 1215–1216
objectives, 1215
quasi-parallel IF systems, 1218
typical integrated electronics for, 1216–1218
IFRB. *See* International Frequency Registration Board (IFRB)
IHVT. *See* Integrated high voltage transformer (IHVT)
Illumination Engineering Society of North America (IES), 69
IMD. *See* Intermodulation distortion (IMD)
Impedance, definition of, 274
Incidental phase modulation (IPM), 375, 379
Independent sideband (ISB) AM stereo, 416–420
Institute of Electrical and Electronic Engineers (IEEE), 67
Instructional Television Fixed Service (ITFS), 5, 1298–1299
Insulation resistance test
for rigid line, 149
for semi-flexible cable, 135
Insulators, porcelain, 264
Insulators for towers
base insulators, 82, 99
compression insulators, 82
guy insulators, 82, 98
sectionalizing insulators, 82
tension insulators, 82
Integrated high voltage transformer (IHVT), 1233, 1235–1236
Integrated services digital network (ISDN), 1105–1106, 1116–1117
Intercity relay (ICR), 43, 674
Interference
acoustic phase interference in microphones, 766–769
cable frequency bands affected by radio frequency interference, 743–744
composite triple-beat (CTB) interference, 735
cures for, 31–32
between FM SCA subcarriers, 489
FM transmitters, 445
between ground waves and sky waves, 302
and Part 74 licensee, 31–32

radio receivers, 1140–1141
satellite earth stations, 698–702
SCA-to-stereo interference considerations, 494–496
Intermediate power amplifiers (IPA), 446–449
Intermodulation distortion (IMD), 245, 1127–1132
International Branch, of the FCC, 5
International Conference of Building Officials (ICBO), 89
International exchange of images, 1206–1208
International frequency allocations, 33
International Frequency Registration Board (IFRB), 34
International Radiation Protection Association (IRPA), 314–315
International Radio Consultative Committee (CCIR), 33–34, 44–45, 380
International shortwave broadcast transmitters, 379–380
International Telecommunication Union (ITU), 33–34, 380
Inverse square law, 758–759
IPA. *See* Intermediate power amplifiers (IPA)
IPM. *See* Incidental phase modulation (IPM)
IRPA. *See* International Radiation Protection Association (IRPA)
ISB. *See* Independent sideband (ISB) AM stereo
ISDN. *See* Integrated services digital network (ISDN)
ITFS. *See* Instructional Television Fixed Service (ITFS)
ITU. *See* International Telecommunication Union (ITU)

Journal of the Audio Engineering Society, 70
Journals. *See* Periodicals
Junction and pull box sizing, 88

K factor, 654–656, 659
Kahn/Hazeltine AM stereo system, 416–420
Kahn power-side system, 369
Key systems, 1090–1094
Kirchoff's law, 109
KLYSTRODE, 536–537
Klystrons
basic klystron theory and practice, 527–533
beam current pulsing, 532
cavity tuning, 529–530
cavity/transmission line coupling, 530
effect of pulsing on transmitter precorrection, 532–533
effect on RF drive power on RF output power, 530
efficiency, 530–531
KLYSTRODE, 536–537
magnetic field, 529
multi-stage depressed collector (MSDC) klystrons, 533–534
perveance and the modulating anode, 529
transmitter design using multiple depressed collector klystrons, 534–536
tuning, 531
variable coupler for integral cavity tubes, 531–532
Ku-band satellites, 685–686

Ladders, for towers, 83, 99
Lead Contamination Control Act, 51
Legal Branch, of the FCC, 5
Licensing process
application for station license, 6
auxiliary and translator services, 7
contested applications, 6
DBS, 7
first step in, 6
LPTV, 7
microwave licensing considerations, 29–30
renewal of license, 6
sales and transfer, 6–7
station construction and equipment testing, 6
types of facilities, 5–6
Lighting. *See also* Studio lighting
for broadcast facilities, 996–997
for color television, 1185
for television, 1033–1050
Lighting for towers
energy storage and, 86
FAA/FCC requirements, 84–85
flashlamp characteristics, 86–88
high intensity lights, 87
inspection and maintenance procedures for, 84–85, 99
medium intensity lights, 87
monitoring of, 576–577
optical design and, 87
physical installation requirements, 87
power conversion and, 86
red lights, 86
white xenon lights, 86
Lightning protection
AC power systems, 1286–1287
for AM broadcast antennas, 188–190
bonding and grounding recommendations, 113
equipment damage mechanisms, 105
equipment protection, 103, 105–113
external ground system, 106–109
for FM broadcast antennas, 269–271
formulas for calculation of resistance to ground, 107
ground electrode testing, 106–108
grounding resistance, 108–109
isolated ground receptacle, 113
nature of lightning, 101–103
objectives of, 103
personnel protection, 103–105
signal line suppressors, 112–113
single point grounding, 110–111
structural protection, 103–104
surge suppression, 112
tower-building system, 109–111
Lightwave transmission theory, 722
Line-level microphones, 761–762
Local channels, for AM broadcasting, 36, 36–37
Lockout/tagout, 53
Low Power Television Branch, of the FCC, 3
Low power television transmission systems (LPTV)
AC power requirements, 1304
antennas, 1302–1303
boosters, 1309–1310
decoders, 1315
design considerations, 1299–1305
downconverters, 1314–1315
encoding and decoding, 1305
environmental considerations, 1302, 1304
filters, 1311
frequency allocations for, 40
grounding and transient protection, 1304–1304
hardware analysis, 1305–1315
Instructional Television Fixed Service (ITFS), 1298–1299
introduction to, 1297–1299
licensing of, 5, 7
Multichannel Multipoint Distribution Service (MMDS), 1298–1299
propagation characteristics, 1299–1300
receive antennas, 1311–1314
receiver sensitivity, 1300–1302
S-Band receiving equipment, 1311, 1313
site requirements, 1302–1305
translators, 1309
transmission lines, 1311
transmitters, 1305–1308
transmitting antennas, 1311
LPTV. *See* Low power television transmission systems (LPTV)

Magazines. *See* Periodicals
Magnetic recording media. *See also* Audio recording; Videotape recording
accidental erasure, 843–844
basic principles of, 825–830
cassette formats, 843
digital recording, 828
features of, 825
frequency modulation recording, 828–830
head-to-tape spacing, 835
high definition television, 830
linearity and nonlinear systems, 827–830
manufacturing of, 830–833
operating practice recommendations for all formats, 840–842
particle orientation and size, 835
progress toward higher recording densities, 833–835
properties of, 826–827
recording and operations area environment, 839–840
reel-to-reel formats, 842–843
storage environment, 843
track density, 834–835
types of material in general use, 831–832
Maintenance
AM broadcast antenna systems, 187–193
AM stereo broadcasting, 415
audio tape machine, 857–860
rigid coaxial transmission line system, 149
robotics for television cameras, 987–988
semi-flexible cable, 135
television transmitters, 545–547
towers, 97–100
videotape recording, 900–905
Major event frequency coordination, 30
Mass Media Bureau (MMB)
Audio Services Division, 3
Enforcement Division, 3, 5
organizational chart, 4
Policy and Rules Division, 5
Video Services Division, 3
Master International Frequency Register (MIFR), 34
MDS. *See* Multipoint Distribution Service (MDS)
Medium frequency propagation, 301–302
Memorandum Opinion and Order (MO&O), 7, 9
Mexican method, for distance calculations, 1260–1261
MGU. *See* Motor-generator unit (MGU)
Microphones
AC hum rejection, 765–766
acoustic gain devices, 760
acoustic phase interference, 766–769
Blumlein miking, 772
carbon microphones, 749
cardioid microphones, 755–757
care of, 772–774
coincident microphone techniques, 770–772
condenser microphones, 751
directional microphones, 755–757
distant-miking applications, 759
durability and reliability, 763
dynamic moving coil microphones, 750–751
dynamic range, 762–763
dynamic ribbon microphones, 749–750
EIA sensitivity rating, 761
frequency response, 752–753
frequency response and distant miking, 760
handling or mechanical noise, 764–765
headset microphones, 759
inverse square law, 758–759
line-level microphones, 761–762
mid-side coincident microphones, 771–772
noise problems, 763–770
omnidirectional microphones, 754–757
open circuit output voltage, 760–761
output impedance, 762
polar patterns, 753–757
polar scaling and dynamic range, 754
polarity reversal, 769–770
power level, 761
pressure microphones, 755
sensitivity ratings, 760–762
shotgun microphones, 759–760
spaced-pair microphones, 770
stereo microphone techniques for broadcasting, 770–772
studio audio, 765
transducer types, 749–752
wireless microphones, 1069–1070
working distance, 757–760
X-Y microphones, 771–772
Microprocessor-based transmitter remote control
A/D converter and analog multiplexor, 592
EEPROM, 590–591
EPROM, 589–590
memory I/O map decode, 589
parallel output ports, 591–592
processor in, 587–589
RAM, 591
scanned keyboard port, 592
watchdog timer, 589
Microprocessor control systems, 457–458
Microstrip Wilkinson combiner, 519–520
Microwave booster stations, 43
Microwave engineering
antenna systems, 657, 660
clearance requirements, 656
FCC licensing for, 668
free space loss, 653
frequency planning, 667–668
frequency re-use, 668
horn antennas, 660
isotropic antenna, 657, 660
K factor, 654–656, 654–659
multipath outages, 654
omnidirectional antennas, 660
parabolic antennas, 660
passive reflectors, 660–661
passive repeaters, 661–662
path loss, 653–654
path performance, 653–657
path profiles, 665–658
power plants, 667
power supplies, 666
radio considerations, 662–667
rainfall outages, 654, 658
receivers, 665–666
repeaters, 662–663
subcarriers, 666–667
system path loss, 653–654
terrestrial microwave, 1079
transmission lines, 661
transmitters (heterodyne), 664–665
transmitters (remodulating), 663–664
Microwave receivers, 665–666
Microwave transmission. *See* Television microwave transmission
Mid-side coincident microphones, 771–772
MIFR. *See* Master International Frequency Register (MIFR)
MIX, 7
Mix-minus, 1100–1101
Mixing loss, 464
MMDS. *See* Multichannel Multipoint Distribution Service (MMDS)
MO&O. *See Memorandum Opinion and Order* (MO&O)
Modulation monitoring. *See* Aural modulation monitoring

Monitoring. *See* Aural modulation monitoring; Frequency monitoring
Monochrome compatible color TV systems, 1194
Monochrome television system, 1151–1152
Motor-generator unit (MGU), 1288
Motorola C-QUAM AM stereo system
decoding/receiving techniques, 423–424
encoding system, 421–423
equations for, 420–421
performance considerations, 424–426
Moving coil microphones, 750–751
MSTV. *See* Association for Maximum Service Television (MSTV)
MTS. *See* Multichannel Television Sound (MTS)
Multi-transmitter different frequencies coaxial system, 183–185
Multichannel Multipoint Distribution Service (MMDS), 1298–1299
Multichannel Television Sound (MTS)
background on, 553–554
BTSC system, 405–418
general discussion of, 553
operating practices, 566–568
Multiplexers
DS-1 based multiplexers, 673
FM SCA multiplexing, 270
telephone network interfacing, 1104
Multipoint Distribution Service (MDS), 5

NAB. *See* National Association of Broadcasters (NAB)
NAD27/NAD83 datums, 1262
National Association of Broadcasters (NAB), 49, 67, 380
National Cable Television Association (NCTA), 67
National Captioning Institute (NCI), 1237
National Contingency Plan (NCP), 57
National Council on Radiation Protection and Measurements (NCRP), 314
National Emissions Standards for Hazardous Air Pollutants (NESHAPS), 51
National Environmental Policy Act, 316, 324
National Environmental Protection Act (NEPA), 64
National Institute of Standards and Technology (NIST), 63
National Labor Relations Board (NLRB), 65
National Oil and Hazardous Substances Pollution Plan, 57
National Political Conventions Frequency Coordinating Committee, 23
National Radio Systems Committee, 380–381
National Society of Professional Engineers (NSPE), 60
National Telecommunications and Information Administration (NTIA), 34–35, 63
National Weather Service (NWS), 1255
Nautel Corporation—AMPFET transmitters, 365–366
NCI. *See* National Captioning Institute (NCI)
NCP. *See* National Contingency Plan (NCP)
NCRP. *See* National Council on Radiation Protection and Measurements (NCRP)
NCTA. *See* National Cable Television Association (NCTA)
NEPA. *See* National Environmental Protection Act (NEPA)
NESHAPS. *See* National Emissions Standards for Hazardous Air Pollutants (NESHAPS).
Network analyzers
for rigid line, 147–148
for semi-flexible cable, 134
News gathering. *See* Electronic news gathering (ENG); Fiber news gathering
NEXRAD, 1254
NFCC. *See* SBE National Frequency Coordinating Committee (NFCC)
950 band, 27
950 MHz STLs, 671–672
942–944 band segment, 27
NIR color system, 1203
NIST. *See* National Institute of Standards and Technology
NLRB. *See* National Labor Relations Board (NLRB)
NOI. *See Notice of Inquiry* (NOI)
Noise, occupational, 53
Noise control
acoustics of, 1018
airborne noise reduction, 1028
assessing noise levels, 1027–1028
audio recording, 852–853
broadcasting facilities, 997–998
checklist for, 1026–1027
discontinuous construction, 1028–1029
doors, 1030
dynamic noise reduction (DNR), 1221–1222
floating construction, 1030
FMX stereo noise reduction, 486
HVAC considerations, 1031
impact noise reduction, 1030–1031
noise reduction between rooms, 1028
STC rating, 1028
studio audio, 965
techniques for, 1026–1031
transmission loss of solid materials, 1028
vibration isolation, 1030
windows, 1028, 1030
Noise measurement, 468, 1132–1133
Non-RF alternatives, 27
North American Regional Broadcasting Agreement (NARBA), 36
Notice of Inquiry (NOI), 7, 9
Notice of Proposed Rule Making (NPRM), 7, 9, 35
NPRM. *See Notice of Proposed Rule Making* (NPRM)
NRSC, 975–976
NRSC-1 audio standard, 382–392
NRSC-2, 393–404
NSPE. *See* National Society of Professional Engineers (NSPE)
NTIA. *See* National Telecommunications and Information Administration (NTIA)
NTSC color system, 1194–1199

Occupational noise exposure, 53
Occupational Safety and Health Act, 64
Occupational Safety and Health Administration (OSHA), 49–53, 64, 315
Off-the-air monitoring, 12
Office of Administrative Law Judges (ALJ), 3, 6
Office of General Counsel (OGC), 1
Office of Legislative Affairs (OLA), 1
Office of Plans and Policy (OPP), 1
Office of Public Affairs (OPA), 1
Office of Science and Technology (OST), 3
Office of the Managing Director (OMD), 1
OGC. *See* Office of General Counsel (OGC)
OLA. *See* Office of Legislative Affairs (OLA)
OMD. *See* Office of the Managing Director (OMD)
Omnidirectional microphones, 754–757
1/2-inch digital VTR (D-3), 900
1/2-inch video format analog composite S-VHS system, 890–891
1/2-inch video format analog systems, 887–890
One-inch video format analog system, 882–885
OPA. *See* Office of Public Affairs (OPA)
Operational Area Emergency Communications Committee, 1265
Operational Fixed Services, 43
OPP. *See* Office of Plans and Policy (OPP)
OST. *See* Office of Science and Technology (OST)
Ozone, 54–55

P channel, 603, 673
Paint, for towers, 83, 97
PAL color system, 1199–1202

Par unit, 1038
Parabolas, 760
Parabolic reflector,760
Part 74 licensee, and interference, 31–32
Part 74 spectrum, 25–27
Part 94 Operational Fixed Services, 43
Patch panels, 161
PBX and key systems, 1090–1094, 1101
PCBs. *See* Polychorinated biphenyls
PE. *See* Professional engineers (PE)
Peak clipping, 968
Peak power, definition of, 279–280
Periodicals, 69–71
Petition for Rule Making, 5, 7
Phase-locked loop (PLL) technology, 441–442
PLL. *See* Phase-locked loop (PLL) technology
Pocket radios, 1142–1143
Polarization, definition of, 273
Policy and Rules Division, of the FCC, 5
Policy Branch, of the FCC, 5
Polychlorinated biphenyls (PCBs), 53–54
Polyphase pulse-width modulation, 365
Popular Communications, 70
Porcelain insulators, 264
Port patch panels, 161
Post Sunset Authorizations (PSSA), 38
PRB. *See* Private Radio Bureau (PRB)
Pressure microphones, 755
Presunrise Authorizations (PSRAs), 37–38
Preventive maintenance. *See* Maintenance
Private Radio Bureau (PRB), 3
PRO Sound News, 70
Professional engineers (PE), 59–61
Propagation. *See also* Radio wave propagation
  basic characteristics, 1013–1015
  FCC standard method for collection of propagation data, 348
  FM and television broadcast field strengths, 341, 343
  FM broadcast antennas, 237–242
  free space propagation, 299–300
  frequency coordination and, 24
  low power television transmission systems (LPTV), 329–330
  medium frequency propagation, 301–302
  TEM mode, 135
PSRAs. *See* Presunrise Authorizations (PSRAs)
PSSA. *See* Post Sunset Authorizations (PSSA)
Public Notice, filing of, 6, 7
Public Service Division, of the FCC, 21
Pulse-width modulation (PWM), 359–361
PWM. *See* Pulse-width modulation (PWM)
Pylon antenna, 292–293

QKT, 1094
QST, 70
Quadruplex recorders, 878

R&O. *See Report and Order* (R&O)
RADAC C-Band weather radar network map, 1257
RADAC S-Band weather radar network map, 1256
Radar. *See* Weather radar
Radiation
  electromagnetic, 52
  from FM broadcast antennas, 255
  human exposure to RF radiation, 323
  unwanted television antenna radiation, 280–281
Radiation hazards, from AM broadcast antennas, 190–191
Radio consoles, 951–953
Radio control rooms, 1007–1008
Radio electronic news gathering and field production
  emerging technologies, 1070–1071
  planning the RENG network, 1051
  remote site equipment, 1054–1056
  set-up and testing procedures, 1056–1058
  setting up a wired remote, 1053–1054
  wired systems, 1052–1058
  wired vs. wireless, 1051–1052
  wireless microphones, 1069–1070
  wireless systems, 1058–1069
Radio-Electronic-Television Manufacturers Association (RETMA), 88–89
Radio Electronics, 70
Radio frequency, 354–356
Radio frequency radiation. *See* RF radiation
Radio receivers
  AM and FM band frequencies, 1147–1148
  car radios, 1144–1145
  double-conversion receivers, 1141–1142
  filter design considerations, 1142
  group delay and frequency response, 1142
  home tuners and tuner-amplifiers, 1143–1144
  interference sources, 1140–1141
  overview of current receiver technology, 1139–1142
  pocket radios, 1142–1143
  SCA receivers, 1145–1146
  table radios and combination units (boom boxes), 1143
Radio studios. *See also* AM broadcasting; FM broadcasting; Studio audio
equipment, 1009–1011
  layouts of, 947–948
  on low budget, 1008–1011
  physical layout, 1008–1009
  planning, 1005–1007
  radio control room as, 1007–1008
Radio wave propagation
  atmosphere, 308
  atmospheric absorption, 309
  beyond line of sight conditions, 303
  buildings, 307
  computer databases, 310
  coverage areas, 309–310
  definition of, 297
  ducting, 308–309
  effects of obstacles on, 303
  field strength prediction, 309–310
  free space propagation, 299–300
  ground waves, 301–307
  hills, 303–307
  interference between ground waves and sky waves, 302
  medium frequency propagation, 301–302
  presence of earth, 300–301
  propagation above 30 MHz, 302–309
  quantifying propagation, 297–299
  rainfall, 309
  sky waves, 302
  smooth earth conditions, 302–303
  solar activity effects on, 302
  vegetation, 303
Radio World, 70
Radomes, 269
Radon, 55–56
Rainfall
  effects on satellite earth stations, 693–696
  microwave engineering and, 654, 658
RAM, 521
RARCs. *See* Regional Administrative Radio Conferences (RARCs)
RB. *See* Review Board (RB), of the FCC
Receivers. *See* Color television receivers; Radio receivers
Recording. *See* Audio recording; Magnetic tape media; Videotape recording
Recording Engineer/Producer, 70
Reel-to-reel audio recorders, 963–964
Reel-to-reel magnetic formats, 842–843

Reference white, 1160
Reflectometers, 271
Regional Administrative Radio Conferences (RARCs), 33–34
Regional channels, for AM broadcasting, 36–37
Rejection filters, 250
Remote control. *See* Transmitter remote control systems
Remote pickup broadcast stations (RPUs), 43–44
Remote production vehicle (RPV), 1074–1077
Renewal of license, 6
RENG. *See* Radio electronic news gathering and field production
Repeaters, for microwave system, 662–663
Report and Order (R&O), 7, 9
RETMA. *See* Radio-Electronic-Television Manufacturers Association (RETMA)
Review Board (RB), of the FCC, 1–3
RF clipping, 968
RF output networks, 370–373
RF power amplifiers, 370–371
RF pulse test
  for rigid line, 147
  for semi-flexible cable, 134
  for waveguide, 158–160
RF radiation
  ACGIH RF threshold limit values, 315
  AM stations, 316–317
  ANSI radiation protection guide, 312–314, 316
  auxiliary services, 317
  broadband instruments, 318–319
  calculation methods, 320–323
  EPA guidelines, 316
  exposure standards, 311–316
  FM stations, 317
  foreign standards, 315
  guidelines on, 324–325
  human exposure to, 311–323
  IRPA guidelines, 314–315
  local standards, 315–316
  measurement procedures, 318–320
  measurement techniques, 319–320
  mitigation of hazard, 321–323
  multi-user sites, 317–318
  narrowband instruments, 319
  National Council on Radiation Protection and Measurements, 314
  OSHA standard, 315
  physical parameters, 318
  special adaptations or applications of instruments, 320
  specific absorption rate (SAR), 312
  television stations, 317
RF switches, 271
Ribbon microphones, 749–750
Rigid line
  applications of, 137
  attenuation, 139, 143
  average power, 139
  characteristics of, 138
  coaxial components, 144
  cut-off frequency, 136
  derating, 139, 142
  efficiency, 139
  expansion/contraction compensation devices, 137
  general considerations, 135
  horizontal run, 145
  installation of, 144–145
  mechanical interfaces, 144
  operation and maintenance of, 149
  peak power, 138–140
  pressurization, 145–146
  propagation factors, 135–137
  purging, 146
  section lengths, 138–139
  selection issues, 137–142
  systems, 142, 144
  TEM mode of propagation, 135
  testing of, 146–149
  thermal analysis, 137
  velocity of propagation, 135–136
  vertical run, 140–141
  VSWR, return loss, reflection coefficient, 136–137, 141
Rio Agreement, 36–37
Robotics for television cameras
  bar-coded signs, 985
  considerations for, 981–983
  installation of, 983–984
  interface with other systems, 987
  labor costs, 981–982
  maintenance, 987–988
  manual control, 987
  mechanical track system, 985
  navigation schemes, 984–985
  operating control systems, 986–987
  operation of cameras located in inaccessible places, 983
  operational considerations, 987–988
  operator training, 987
  personnel considerations, 983
  reliability and durability, 983
  repeatability of shots, 982
  safety, 985–986
  scheduling, 982
  smooth execution of complex camera work, 982
  tape track system, 985
  target system, 985
Rounding practices, in distance and bearing calculations, 1262
RPUs. *See* Remote pickup broadcast stations (RPUs)
RPV. *See* Remote production vehicle (RPV)
RT_{60} formula, 1026

S-VHS video system, 890–891
Sales, and licensing process, 6–7
SAM noise. *See* Synchronous AM (SAM) noise
SARA. *See* Superfund Amendment and Authorization Act (SARA)
Satellite backhaul, 1079
Satellite earth station antennas
  axisymmetric dual-reflector antennas, 704–705
  beam waveguide antenna, 706
  electrical performance, 704
  importance of, 703–704
  multiple beam antenna, 706
  offset-fed reflector antenna, 706
  prime-focus-fed paraboloidal (PFFP) antenna, 705–706
  reflector feed configurations, 706–707
  single-beam antennas, 704
  types of, 704–706
Satellite earth stations
  antennas, 707
  audio service, 712–713
  automatic transmit identification system, 712
  baseband performance link analysis, 697–698
  block diagram, 698–699
  C-band satellites, 684–686
  characteristics of satellite system, 681–683
  digital transmission, 684
  equipment characteristics, 703–715
  example of system link calculation, 696–697
  FCC licensing, 687
  frequency, 680
  geosynchronous orbit of, 679–680
  interference analysis, 698–702
  interference analysis for FM/TV service, 702
  introduction to, 679–687

Ku-band satellites, 685–686
low-noise amplifiers/low-noise converters, 713
mechanical performance, 707–711
monitor and control systems, 714–715
performance analysis techniques, 687–703
phase shift keying (PSK), 698
pointing and tracking, 710–711
polarization, 681
positioning systems, 709–710
protection switching, 714
rain effects on system performance, 693–696
receive electronics, 713–714
receive figure-of-merit G/T, 688–690
regulatory issues, 685–687
satellite footprint, 681
satellite link analysis, 691–693
satellite stationkeeping, 680–681
satellite transponder, 691
sun transits and eclipses, 702–703
transmission modes for television, 683–684
transmit electronics, 711–713
video receivers, 713–714
video services, 711–712
SAW. *See* Surface acoustic wave (SAW)
SBE. *See* Society of Broadcast Engineers (SBE)
SBE National Frequency Coordinating Committee (NFCC), 24–25, 68
SCFCC. *See* Southern California Frequency Coordinating Committee (SCFCC)
SCPC satellite. *See* Single Channel Per Carrier (SCPC) satellite
Scrambling, 744–745
SECAM color system, 1202–1203
SECC. *See* State Emergency Communications Committee (SECC)
Sectionalizing insulators, 82
Seismic loads, 91
Self-supporting towers, 83
Semi-flexible cable
advantages of, 118
air-dielectric cable, 118
AM tower installations, 133
attenuation, 119
attenuation, 123, 124
characteristics of, 126, 127
coaxial cable design considerations, 118–120
coaxial components, 120–121
comparison chart for, 116
connector loss, 123
connectors, splices, and adaptors, 126–129
cut-off frequency, 119–120
derating, 123, 125–126
efficiency, 123
foam-dielectric cable, 118
general considerations, 116, 118
grounding of, 132–133
hanger installation of, 130–131
horizontal runs, 131
impedance of, 118–119
installation of, 129–133
maintenance of, 135
minimum bending radius of, 118
moisture purging, 132
power handling, 119
power ratings, 120, 122–123
pressurization, 131–132
sampling lines, 133
selection considerations, 120–126
testing of, 133–135
tower installation of, 130
types of, 118
voltage standing wave ratio (VSWR), 123, 125
7 GHz band, 29
Shortwave broadcast transmitters, international, 379–380
Shortwave broadcasting, switching for, 163
Shortwave parallel transmitters, 172–175
Signal line suppressors, 112–113
Sine-squared pulse and bar, 1188
Single Channel Per Carrier (SCPC) satellite, 603
Single sideband AM subcarriers, 494
SMPTE. *See* Society of Motion Picture and Television Engineers (SMPTE)
SMPTE Journal, 70
SMPTE Time Code, 906–910, 1127–1129
Society of Broadcast Engineers (SBE), 24, 27, 30, 32, 67–68
Society of Motion Picture and Television Engineers (SMPTE), 68, 877–878, 934, 940–941
Soil conductivity
AM broadcast antennas, 191
FM broadcast antennas, 242
Soil investigation, for towers, 95
Solid-state broadcast transmitters, 361–363
Solid-state equipment and circuitry
advantages of, 364
Asea Brown Boveri (ABB) pulse-step modulator, 364
Continental Electronics Corporation (CEC) pulse-step modulator, 364–365
general discussion of, 363–364
Harris DX system, 365, 367–369
Kahn power-side system, 369
Nautel Corporation—AMPFET transmitters, 365–366
polyphase pulse-width modulation, 365
Solid-state FM transmitters, 449, 465
Solid-state television transmitters
AC distribution, 518
automatic control gain (ACG), 517
combiners and dividers, 519–520
combining multiple amplifier cabinets, 517
control systems, 518
cooling system, 517–518
design of, 514
Microstrip Wilkinson combiner, 519–520
power supplies, 518
RF amplifiers, 517
ring combiner, 519
solid-state devices, 517
system on-air availability, 514, 516–517
Sound. *See* Acoustics; Multichannel Television Sound (MTS); Studio audio; and headings beginning with Audio
Sound-powered telephone, 88
Sound Transmission Class (STC) rating, 1028
Southern Building Code Congress, 89
Southern California Frequency Coordinating Committee (SCFCC), 23–24
Spherical-earth method, 1261
Stair-step generator, 1187–1188
Stairways, for towers, 84
Standing wave ratio (SWR), 459–460
State Emergency Communications Committee (SECC), 1265
STC. *See* Sound Transmission Class (STC) rating, 1028
Stereo broadcasting. *See* AM stereo broadcasting; FM stereo broadcasting
Stereo microphone techniques for broadcasting, 770–772
Stereo television, 977–978
STL. *See* Studio-to-transmitter link (STL)
Stratospheric ozone, protection of, 54–55
Studio acoustics, 999–1003
Studio audio
audio cart machines, 963
audio consoles, 956
audio distribution and routing, 956–959
audio enhancements, 964–965

audio routing switchers, 959
audio sources, 961–964
compact disc (CD) players, 961–962
digital audio systems, 964
distribution amplifiers, 959–960
effects generators, 965
equalization, 964–965
microphones, 963
noise reduction systems, 965
patch panel wiring and termination, 957–959
patch panels, 956–957
reel-to-reel recorders, 963–964
studio monitors, 959, 961
television studio layout, 948–951
turntables, 962–963
Studio layouts, for radio studios, 948–949
Studio lighting
accessories, 1039
air conditioning and, 1046
architectural criteria, 1045–1046
basic principles of, 1033–1034
battens, 1041–1042
broad area lighting, 1037
catwalks, 1041
color temperature, 1034–1035
control consoles, 1048–1050
control system, 1048–1050
cyc pit, 1044–1045
cyclorama, 1042–1045
dimmer bank, 1046–1048
electrical criteria, 1045–1046
electrical distribution, 1046–1048
ellipsoidal reflector spotlight, 1038–1039
fixed grid, 1040–1041
fresnels, 1037
hard cyc, 1044
lighting fixtures, 1035–1039
lighting levels, 1034
location lighting, 1035–1036
modular grids—self-hoisting, 1042
Par unit, 1038
production system, 996
softlights, 1037–1038
spot lighting (fresnels), 1036
studio design and, 1039–1040
studio lighting, 1036–1039
suspension systems, 1040–1042
three-point system, 1033–1034
winches, 1042
Studio Sound and Broadcast Engineering, 71
Studio-to-transmitter link (STL)
AM stereo broadcasting, 407–408
aural broadcast auxiliary links, 671–673
aural studio-to-transmitter link (STL) band, 27
frequency allocations, 43
television STL coordination, 29–31
television STL stations, 43–44
transmission audio processing, 973
Studios. *See* Radio studios; Television studios
Subcarrier, AM, 494
Subcarrier, FM
composite baseband, 475–477
crosstalk from main channel program into SCA subchannels, 496–497
examples of analog SCA operation, 490–491
FM SCA transmission, 488–491
frequency spectrum and modulation limits, 476–477
glossary of terms for, 475–476
interference between FM SCA subcarriers, 489
performance of, 492–497
SCA data systems, 491–492
SCA-to-stereo interference considerations, 494–496
types of analog subcarrier modulation, 489–490
Subcarriers, and microwave systems, 488–489
Subjective listening tests, 978–979
Subscriber loops, 1083–1087
Subsidiary Communications Authorization (SCA) services, 475–477, 488–497
Superfund Amendment and Authorization Act (SARA), 52, 56
Surface acoustic wave (SAW), 505
Surge suppression, 112
Sweep test
for rigid line, 146–147
for semi-flexible cable, 134
for waveguide, 158
Switching
coaxial patch panels, 161–162
manual coaxial switches, 162
motorized coaxial switches, 162
patch panels, 161
port patch panels, 161
RF switches, 271
for shortwave and AM applications, 163
waveguide patch panels, 162
waveguide switches, 163–164
Switchless combiners, 540–542
SWR. *See* Standing wave ratio (SWR)
Synchronous AM, 468–469
Synchronous AM (SAM) noise, 245

Table radios and combination units (boom boxes), 1143
TAF. *See* Telecine analysis film (TAF)
Talk-show screening software, 1102
TASO Picture Quality Scale, 338–339
TCA. *See* Toxic Substances Control Act (TCA)
TCB. *See* Tri-chlorobenzene
TDM antennas. *See* Transmission dual mode (TDM) antennas
TDR test. *See* Time domain reflectometry (TDR) test
Telco audio program service
analog circuits, 1111–1114
communication lines, 1117
cost of services, 1117–1118
data compression, 1114
data rates, 1114
dealing with telco, 1111
digital audio circuits, 1114–1117
DS1 or T-1 service, 1115
fractional T-1, 1115–1116
glossary of telco terms, 1118–1120
Integrated Services Digital Network (ISDN), 1116–1117
interfacing procedures, 1113–1114
obtaining and testing telco lines, 1111–1113
switched, 1116
Telecine
basic principles, 935–936
broadcast telecine, 935
CCD line array, 938–939
CRT flying spot scanner, 937–938
design, 935–939
photoconductive telecine, 936–937
production telecine, 935
setup for, 940–941
transfer telecine, 935
Telecine analysis film (TAF), 940
Telephone, sound-powered, 88
Telephone network interfacing
broadcast audio interfacing, 1094–1098
cellular telephones, 1089–1090
Centrex, 1089
channel cards, 1104
channel service unit (CSU), 1104
choke networks, 1087–1088

dial-up network basics, 1083–1090
dial-up remotes, 1106–1108
digital service unit (DSU) and multiplexer, 1104
digital telephone systems, 1088–1089
electronic telephone systems, 1090–1094
FCC regulations concerning, 1090
filtering, 1098
foreign exchange (FX) loops, 1087
four-wire intercom systems, 1108–1110
frequency shifting using one line, 1106–1107
frequency shifting using two or more lines, 1107
half-speed transmission, 1107–1108
hybrids, 1095–1098
improving phone audio quality, 1098
integrated services digital network (ISDN), 1005–1006
integration of on-air systems with PBXs, 1101
interfacing production intercom systems to telco lines, 1108–1110
mix-minus, 1100–1101
on-air systems for broadcast interfacing, 1098–1102
1A2 key systems, 1090
PBX and key systems, 1090–1094
QKT, 1094
recording phone calls, 1101
remotes with phones, 1101–1102
subscriber loops, 1083–1087
switching technique, 1094–1095
T-1 digital service, 1102–1106
talk-show screening software, 1102
two-way interfacing, 1094–1098
two-wire systems, 1110
Telephone system, for broadcast facilities, 995
Television audio consoles, 951
Television booster stations, 40
Television Branch, of the FCC, 3
Television broadcast antennas
antenna gain, 273–274
antenna input impedance, 281–282
antenna supporting mast, 347
azimuth pattern, 273, 275
beam tilt, 279
branching feed systems, 285
characteristics of, 281–286
deflection and wind load, 284
effective radiated power (ERP), 273, 277
evaluation of system performance, 341–342
feed systems of, 285–286
field strength, 277–278
impedance, 274
null fill, 279
panel antenna, 276
polarization, 273, 278–279
power capability, 279–281
radiating characteristics, 274–275
radiating structures, 275–278
receiving antennas, 345–347
slot antenna, 275–276
standards for TV STL systems, 29
standing wave feed systems, 285
technical terms concerning, 273
tests for, 295
transmission line and components, 282–283
traveling wave antenna, 276–277
traveling wave feed systems, 285
types of, 286–287
UHF antennas, 292–295
unwanted antenna radiation, 280–281
vertical pattern (elevation plane), 273–275
VHF antennas, 287–292
Television broadcast auxiliary stations, 43–44
Television broadcast field strengths
analysis of measurements to depict coverage, 339
antenna supporting mast, 347
antenna system performance evaluation, 341–342
equipment principles, 343–347
FCC standard method for collection of propagation data, 348
FCC standard method for measurement of coverage of communities, 348–349
measurement of, 335–351
measurement of coverage, 336–339
measurements versus predictions, 335–336
measuring procedures and techniques, 347–349
planning a measurement program, 336–341
power supplies, 347
practical field strength meters, 343–345
practical problems in measurement of, 349
propagation, 341, 343
receiving antennas, 345–347
recorder, 347
safety considerations in measurement of, 349–351
spurious emission measurement, 341, 343
time variation of field strength, 349
transmitting system evaluation, 340–341
Television broadcasting
auxiliary and translator services, 7
booster stations, 40
FCC Broadcast Service Checklist, 18–22
FCC Television Signal Analysis Report, 15
frequency allocations, 40–42
high definition television system (HDTV), 40–41
licensing of, 5–6
low-power television stations (LPTV), 40
RF radiation, 317
satellite earth stations and, 702
translator stations, 40
transmission audio processing for, 977–978
Zone I, 40, 46
Zone II, 40
Zone III, 40, 47
Television cameras. *See* Cameras, television
Television channel 6/FM antenna problem, 249–258
Television Decoder Circuitry Act, 1238
Television field production
assessing options, 1074
camera selection for, 796–797
communications, 1079–1080
contact with headquarters, 1079
definition of, 1073–1074
electronic field production, 1073
electronic news gathering, 1073
equipment for, 1077–1079
guidelines for, 28–29
leasing vs. buying a remote production vehicle, 1076–1077
outside broadcasting, 1073–1074
planning for, 1074
remote production vehicle (RPV), 1074–1077
satellite backhaul, 1079
terrestrial microwave, 1079
Television film. *See* Film for television
Television lighting. *See* Studio lighting
Television microwave booster stations, 44
Television microwave transmission
benefits of coordination, 31
frequency coordination considerations, 27–28
guidelines for, 28–29
licensing for, 29–30
7 GHz band, 29
temporary operation for TV microwave RPU, 30
13 GHz band, 29
Television monitors, 613
Television pickup stations, 43–44
Television relay stations, 43–44

Television Signal Analysis Report (FCC), 15
Television signal transmission standards
amplitude vs. frequency response, 618
audio performance standards, 618, 622–624
audio practices, 625
audio-to-video time differential, 623–624
chart, 619–621
chrominance nonlinear gain distortion, 633, 638
chrominance nonlinear phase distortion, 633, 637–638
chrominance-luminance delay inequality, 629, 632
chrominance-luminance gain inequality, 629, 632
chrominance-to-luminance intermodulation, 640, 644
continuity of audio service, 623
crosstalk noise, 647
definitions concerning, 617–618
determining rates of accumulated degradation of performance, 624
differential gain, 637, 640
differential phase, 640–642
dynamic gain distortions, 637, 639–640
example of weekly VITS log, 680
field-time waveform distortion, 627
FM terrestrial and satellite systems, 624–625
gain/frequency distortion, 633–634
general and waveform technology, 625–626
harmonic distortion, 618, 622
impulsive noise, 644–646, 648
insertion gain, 626, 617–623
introduction to, 617–618
line-time waveform distortion, 627–628
luminance nonlinear distortion, 633, 635–636
periodic noise, 646–647
random noise-weighted, 644–645
short-time waveform distortion, 629–631
signal-to-noise ratio, 622–623
technique for approximating weighted video random noise, 651
Television sound. *See* Multichannel Television Sound (MTS); Studio audio
Television STL coordination
antenna standards, 29
benefits of, 31
major event frequency coordination for RF equipment, 30
microwave licensing considerations, 29–30
minimum path length, 29
penalties of not coordinating, 31
power limitations, 29
risks of not coordination, 31
temporary operation for TV microwave RPU, 30
Television STL stations, 43–44
Television studios
design of, 1039–1040
equipment purchasing, 949–951
layout of, 948–951
lighting for, 766
planning, 949
Television Technology, 71
Television translator relay stations, 44
Television translator stations, 40
Television translators, 565
Television transmitters
air conditioning, 550
air systems for, 547–550
aural exciter, 509–510
circuit design, 513–515
combination of circuits and RF systems, 538–544
compensation of aural passband for optimum stereo performance, 510–514
exhaust air, 547–548
IF group delay correction, 511–513
intake air, 548–549
offset frequency control, 508–509
output power sampling, 577
performance measurements for, 544–545
preventative maintenance for, 545–547
solid-state transmitters, 514–528
switchless combiners, 540–542
UHF magic tee FR systems, 542–544
UHF transmitters, 527–537
VHF tube transmitters, 520–527
visual exciter, 499–508
Television tuners
band partitioning, 1213
block diagram, 1212–1213
channel selection and fine tuning, 1213–1214
dual-conversion option, 1214–1215
heterodyne conversion fundamentals, 1211–1212
implementation requirements, 1211
TEM (Transverse Electromagnetic) mode of propagation, 135
Temperature changes, and towers, 91
Temporary operation for TV microwave RPU, 30
Tension insulators, 82
THD. *See* Total harmonic distortion (THD)
13 GHz band, 29
3/4-inch video format analog systems, 885–887
Three-tube switchless combiner, 183
Tier-one violation (Form 793), 12–13, 17
Tier-two violation (Form 790), 12, 16
Time and Frequency Division, 63
Time code, for videotape editing, 906–910, 920–921
Time-domain reflectometry (TDR) test
for rigid line, 149
for semi-flexible cable, 135
TIS. *See* Travelers Information Stations (TIS)
Topographic maps, 64
Total harmonic distortion (THD), 466–467
Tower base, inspection and maintenance procedures for, 99
Towers
aircraft marking for, 83
alignment of, 97–98
allowable stresses and safety factors, 92–93
for AM broadcast antennas, 188
AM tower installation of semi-flexible cable, 133
base of, 99
configurations of, 80–81
construction of foundation, 95–96
corrosion, 97
corrosion protection for, 82–83
design standards for, 88–89
dynamic considerations, 92
effects of antennas and transmission lines on, 93–94
electrical systems for, 84–88
electroplating of, 83
elevators for, 83–84, 99
erection of, 96–97
finishes for, 82–83
flame spraying of, 83
foundations and anchors for, 94–96
guy anchors, 95, 99
guy installation for, 97
guyed tower base foundations, 95
guyed towers, 81
guys, 81–82, 98
hot-dip galvanizing of, 82–83
ice load, 89, 91
ice prevention for, 83
industry standards for, 88–89
initial and future considerations for, 94
inspection and maintenance procedures, 97–100
insulators for, 82, 98–99
ladders for, 83, 99

lighting for, 84–88, 99
lighting of, 576–577
lightning protection for, 104
loads, 89–91, 93
materials for, 81–82
reinforcement and modification of existing towers, 97
replacement, relocation, or additions to existing towers, 94
seismic loads, 91
self-supporting tower foundations, 95
self-supporting towers, 80
semi-flexible cable installation on, 130
soil investigation for, 131
stairways for, 120
statutory standards for, 125
structural models and analysis, 127–128
temperature changes and, 127
transmission line bridges for, 120
types of, 79–80
width restrictions for, 93–94
wind load, 89–90
zinc rich paint for, 83
Toxic Substances Control Act (TCA), 54
Transfers, and licensing process, 6–7
Transient Voltage Surge Suppression (TVSS) devices, 110–111
Translator stations
FM broadcasting, 40
licensing of, 7
television broadcasting, 40
Translators, for low power television transmission systems (LPTV), 1309
Transmission audio processing
for AM stereo, 976–977
bounce, 975
building a system, 972
compression, 967–968
definition of linearity, 978
equalization, 969–970
expansion, 968
fundamentals of, 967–970
gating, 969
location of system components, 972–973
loudness, 970
metering, 973
multiband compression and frequency-selective limiting, 969
NRSC-1 audio standard, 975–976
peak limiting and clipping, 968–969
peak modulation control, 970–971
performance requirements, 970–971
processing for stereo, 971
requirements for FM (VHF) broadcast transmission, 977
requirements for MW and HF broadcast transmission, 974–977
requirements for studio-to-transmitter links (STLs), 973
sinewave measurements and nonlinearity, 978
slew rate limiting (transient intermodulation distortion), 975
for stereo television, 977–978
studio levels, 973–974
subjective audio processing concepts, 971
subjective listening tests, 978–979
system considerations, 971–974 technical evaluation of, 978–979
for television broadcast transmission, 977–978
transmission levels, 973–974
transmission-link limiting, 974
transmitter equalization, 974–975
Transmission dual mode (TDM) antennas, 288, 290
Transmission line analyzers, 135
Transmission line bridges, for towers, 84
Transmission lines
AM broadcast antennas, 188
attenuation, 116
comparison chart for, 116
effects on towers, 93–94
examples of, 117
FM broadcast antennas, 266–267
FM transmitters, 459–461
frequency branch of operation, 115–116
low power television transmission systems (LPTV), 1311
microwave equipment, 661
power handling, 116
rigid line, 135–149
selection criteria, 115–116
semi-flexible cable, 116–135
television broadcast antennas, 282–283
Transmission standards for television. *See* Television signal transmission standards
Transmission system, for BTSC system, 561–566
Transmitter control systems
AC voltage sampling circuits, 574–575
alternate control points, 606–607
AM pattern change sequencing, 582–583
AM power sampling, 577–578
AM subaudible, 602–603
automatic transmission systems, 570
broadcast carrier limitations, 603
calculated differential voltage, 572–573
clamp diode for safety, 570–571
common mode attenuator, 571
common mode rejection, 571
common mode voltages, 571
communications circuits, 601–604
continuous telemetry transmission, 599
control interface panel, 582
control outputs, 581–583
data analysis, 600
data reporting, 600–601
DC current samples, 573
DC voltage sampling circuits, 570–573
dedicated radio link (P channel), 603
differential amplifier, 572
differential input circuitry, 571–572
directional antenna monitor, 579–580
division of responsibility, 598–601
exception reporting, 600
extended voltmeter, 599
extension metering, 569–570
fail safe, 605
FCC regulation of, 604–607
filament voltage sampling, 575
flying capacitor isolator, 572
FM output power sampling, 577
FM squared satellite, 603
FM subcarrier, 602
introduction to, 569
isolation amplifier, 572
parameter sampling circuitry, 570–577
problems in analog sampling interfacing, 580–581
remote control, 570
remote control authorization, 604
remote control hardware design, 583–598
remote diagnostics, 601
required control capability, 605
required levels of transmitter control, 569–570
required metering, 604
required metering accuracy, 605
resistor stability, 574
SCPC satellite, 603
site polling, 599
status sampling, 581
STL subcarrier, 602
telco dial-up voice grade circuit, 601–602
telco digital circuits, 602

telco metallic pair, 601
telco voice grade pair, 601
television output power sampling, 577
three-phase line voltage sampling, 575
tower light sampling, 576–577
TV aural subcarrier, 602
vertical blanking interval, 602
VSAT satellite, 603–604
Transmitter remote control systems
A/D with FSK transmission, 585
alphanumeric display, 592–598
analog FM encoding of telemetry, 584–585
analog multiplexing, 584–585
control encoding, 585–586
frequency shift keyed control, 586
general discussion of, 583–584
microprocessor-based remote control, 587–592
nonprocessor-based designs, 586
nonprocessor digital remote control, 586–587
telemetry encoding, 584
Transmitter-studio links (TSL), 673–674
Transmitter systems
AM and short wave parallel transmitters, 172–175
low power television transmission systems (LPTV), 1305–1308
methods of connecting, 170–171
microwave system, 663–665
multi-transmitter different frequencies coaxial system, 183–185
parallel transmitters, 171–172, 174
single and alternate main transmitter systems, 171, 173
special filtering for UHF, 183
three-tube switchless combiner, 183
UHF parallel transmitters (waveguide), 180
UHF switchless system (waveguide), 180–183
VHF parallel transmitters (coaxial), 174–176
VHF switchless system (coaxial), 177–180
Transmitters. *See* AM transmitters; FM broadcast transmitters; Television transmitters; UHF transmitters
Travelers Information Stations (TIS), 35
Tree-and-branch architecture, 732
Tri-chlorobenzene (TCB), 54
Triangular top platforms, 94
TSL. *See* Transmitter-studio links (TSL)
Tuners and tuner-amplifiers, 1143–1144
Turn-around-loss, 464
Turntables, 962–963
TV. *See* Television
TV Broadcasters All-Industry Committee, 23
TVSS devices. *See* Transient Voltage Surge Suppression (TVSS) devices
23 GHz STLs, 672–673

UHF, and high band coordination, 26–27
UHF antennas
circular polarized, 294–295
doublet array, 293–294
linear polarized, 292–294
slotted coax, 294–295
slotted coaxial, 292–293
slotted director, 293
UHF magic tee FR systems, 542–544
UHF parallel transmitters (waveguide), 180
UHF switchless system (waveguide), 180–183
UHF transmitters
basic klystron theory and practice, 527–533
KLYSTRODE, 536–537
multi-stage depressed collector (MSDC) klystrons, 533–534
transmitter design using multiple depressed collector klystrons, 534–536
UL. *See* Underwriters Laboratories (UL)
Underground storage tanks (USTs), 56–57
Underwriters Laboratories (UL), 69
Uninterruptible power system (UPS), 1288, 1293–1294
U.S. Department of Agriculture (USDA), 63
U.S. Department of Commerce, 35, 63
U.S. Department of Interior, 64
U.S. Department of Labor (DOL), 64
U.S. Department of State (DOS), 34
U.S. Department of Transportation, 64, 1076
UPS. *See* Uninterruptible power system (UPS)
USDA. *See* U.S. Department of Agriculture

Vacuum tube power amplifier circuits
cathode-driven triode amplifiers, 450
general discussion of, 449–450
grid-driven tetrode and pentode amplifiers, 451
grounded-grid versus grid-driven tetrode operation, 450–451
impedance matching into the grid, 451–452
neutralization, 452
VCO. *See* Voltage controlled oscillator (VCO)
Vectorscope, 1188–1189
Vertical interval reference and test signals, 1189
Very Small Aperture Terminal (VSAT) satellite, 603–604
VHF, and high band coordination, 26–27
VHF antennas
circular polarized, 288–292
circularly-polarized "V" (CPV), 290
linear polarized, 287–288
panels, 288, 289, 290–292
slotted coax, 292
spiral, 292
superturnstile/batwing, 287
transmission dual mode (TDM), 288, 290
traveling wave, 287–288
VHF parallel transmitters (coaxial), 174–176
VHF switchless system (coaxial), 177–180
VHF tube transmitters
cathode in, 521
coaxial construction of a tetrode RF power amplifier tube, 522–523
double-tuned tube power amplifiers for TV transmitter applications, 524
general discussion of, 520–521
grids in, 521
transmission line cavity, 521–522
tuning for power, 524–527
Video drivers, 1231–1232
Video recording
cassette formats, 844
film-to-tape transfer, 939–940
film-video interface definitions, 945
reel-to-reel formats, 843–844
Video Services Division, 3
Video signals
cascading studios, 819
component analog video (CAV), 819–821
description of, 807–808
digital systems, 821–823
distribution amplifiers, 813
genlock, 816
interconnection conventions, 808
introduction to, 807
multiple-bus switchers, 814
non-NTSC systems, 819
problems, 808
pulse and subcarrier distribution amplifiers, 812–813
selection of distribution amplifiers, 813
single-bus switchers, 813–814
switching, 815
switching system control, 814–815

system design, 808–811
timing, 815–819
Video standards for television signal transmission. *See* Television signal transmission standards
Video Systems, 71
Videotape editing
audio mixer interfaces, 923
audio tape interfaces, 921
component analog video, 926–927
control systems and interfaces, 917–924
digital video, 927–928
digital video effects and character generator interfaces, 922–923
edit suite physical considerations, 930–932
edit suite signal configurations, 924–930
editing controllers, 917–919
equipment centralization, 932
general purpose interfaces (GPI), 923–924
historical perspective, 913–914
interfacing VTRs, 919–920
multichannel VTR audio, 930
off-line edit suite, 915–916
operational types, 916
SMPTE Time Code, 906–910, 1127–1129
system functional blocks, 916–917
system types, 914–916
time code, 920–921
video switcher interfaces, 922
workstation editing, 928–930
Videotape recording
D-1 component digital VTR, 891–898
D-2 composite digital VTR, 898–900
helical scan recorders, 878–882
history, 877–878
maintenance, 900–905
1/2-inch digital VTR (D-3), 900
1/2-inch format analog composite S-VHS system, 890–891
1/2-inch format analog systems, 887–890
one-inch format analog system, 882–885
quadruplex recorders, 878
SMPTE Time Code, 906–910
system features, 905–906
3/4-inch format analog systems, 885–887
troubleshooting, 904–905
Violations
Advisory Notice (FCC Form 790), 12, 16
tier-one violation (Form 793), 12–13, 17
tier-two violation (Form 790), 12, 16
VIR. *See* Vertical interval reference and test signals
Visual exciter
active group delay equalizer, 882–885
group delay compensation, 501
IF and video delay compensators, 501–503
IF linearity precorrection, 506
incidental carrier phase modulation (ICPM), 506–507
modulator, 501
passive group delay equalizer, 503–504
RF generator, 507–508
vestigial sideband filter, 504–506
video processing, 499–501
VIT. *See* Vertical interval reference and test signals
Voltage controlled oscillator (VCO), 439
Voltage standing wave ratio (VSWR), 123, 136–137, 141, 244–245, 461
Voltage tuned oscillator (VTO), 439
VSAT. *See* Very Small Aperture Terminal (VSAT) satellite
VSWR. *See* Voltage standing wave ratio (VSWR)
VTO. *See* Voltage tuned oscillator (VTO)

Wage and Hours Division, 64
WARCs. *See* World Administrative Radio Conferences (WARCs)
Washington Executive Broadcast Engineers (WEBE), 23
Waveguide
attenuation, 149, 151–152
examples of, 150
general considerations for, 149
installation of, 155–157
operating costs of, 151
power handling capability, 151
pressurization, 151, 157–158
selection considerations, 149–153
system components, 156
system layout considerations, 153–155
testing of, 158–160
wind load, 151
Waveguide patch panels, 162
Waveguide switches, 163–164
Weather radar
development of, 1251
Doppler radar, 1252–1253
maintenance of the radar, 1255
maps, 1256–1257
NEXRAD, 1254
operational weather radar theory, 1251–1252
RADAC C-Band weather radar network map, 1257
RADAC S-Band weather radar network map, 1256
radar equipment, 1254
radar remoting, 1255
range effects, 1252
siting a radar, 1254–1255
space loss, 1252
typical radar applications, 1253–1254
WEBE. *See* Washington Executive Broadcast Engineers (WEBE)
WHCA. *See* White House Communication Agency (WHCA)
White House Communication Agency (WHCA), 1265
White, reference, 1164
Wideband community and multi-user antennas, 258–260
Wind load, 89–90, 93, 95, 151, 284
Wired systems
compared with wireless systems, 1051–1052
description of, 1052–1058
remote site equipment, 1054–1056
set-up and testing procedures, 1056–1058
setting up a wired remote, 1053–1054
Wireless microphones, 27, 1069–1070
Wireless systems
antenna considerations, 1062–1063
building an RPU system, 1059
compared with wired systems, 1051–1052
fade margin, determining, 1060–1061
licensing procedures, 1058
path engineering for fixed stations, 1059–1060
planning considerations, 1058–1069
system configuration, 1065–1069
transmitter-receiver considerations, 1063–1065
World Administrative Radio Conferences (WARCs), 33–34

X-Y microphones, 770–771
X-Y oscilloscope, 413

Zoning regulations, 990

AUTHOR LISTING

Anthony, Edward J. (3.2 AM Stereo Broadcasting, 3.3 FM Transmitters)
Broadcast Electronics, Inc.
4100 N. 24th St., P.O. Box 3606
Quincy, Illinois 62305-3606
217 224-9600

Baron, Stanley N. (5.6 Videotape Recording)
National Broadcasting Company, Inc.
30 Rockefeller Plaza
New York, New York 10112
212 664-7557

Barringer, Martin H. (2.8 Radio Wave Propagation)
Frederick G. Griffin, P.C.
3229 Waterlick Rd.
Lynchburg, Virginia 24502
804 237-2044

Bauer, Richard W. (5.8 Film for Television)
Eastman Kodak Company
Motion Picture and Television Imaging Division
Kodak Park
Third Floor Building 69
Rochester, New York 14650-1910
716 477-7717

Bentz, Carl (6.2 TV Field Production)
Broadcast Engineering Magazine
P. O. Box 12901
Overland Park, Kansas 66282-2901
913 888-4664

Best, Gregory L. (3.5 TV Transmitters)
Harris Corporation
Broadcast Division
P.O. Box 4290
Quincy, Illinois 62305-4290
217 222-8200

Bialik, David (1.7 Source Guide: Broadcast Standards and Information)
WKDM Radio
570 Seventh Avenue
New City, New York 10018
212 594-1380

Bisset, John P. (7.1 Broadcast Audio Measurement)
Multiphase Consulting
5827 Columbia Pike
Suite 310-A
Falls Church, Virginia 22041
703 379-1665

Block, James (7.6 Weather Radar Systems)
Kavouras, Inc.
6301 34th Ave. South
Minneapolis, Minnesota 55450
612 726-9515

Breeze, Frank C. (2.2 Lightning Protection for Broadcast Facilities)
Tilden, Lobnitz and Cooper, Inc.
1717 South Orange Avenue
Orlando, Florida 32806
407 841-9050

Bruene, Warren B. (3.3 FM Transmitters)
Broadcast Electronics, Inc.
4100 N. 24th Street, P.O. Box 3606
Quincy, Illinois 62305-3606
217 224-9600

Buchwald, Greg (3.2 AM Stereo Broadcasting)
Modulation Systems Laboratory
Corporate Research
Motorola Center, SH-1
1301 E. Algonquin Road
Schaumburg, Illinois 60196
708 576-4893

Carlson, Bruce J. (2.3 Transmission Lines)
Andrew Corporation
10500 West 153rd Street
Orland Park, Illinois 60462
708 349-3300

Carpenter, J. R. (5.5 Audio Magnetic Recording)
Broadcast Electronics, Inc.
P.O. Box 3606
Quincy, Illinois 62305-3606
217 224-9600

Church, Steve (6.3 Telephone Network Interfacing)
Telos Systems
1729 Superior Avenue
Cleveland, Ohio 44114
216 241-7225

Ciciora, Walter S., Ph.D.(4.6 Cable Television Systems)
American Television and Communications Corporation
300 First Stamford Place
Stamford, Connecticut 06902-6732
203 328-0622

Claudy, Lynn (5.13 Principles of Acoustics for Broadcast Applications)
National Association of Broadcasters
1771 N Street, NW
Washington, D.C. 20036
202 429-5340

Cohen, Jules (2.9 RF Radiation Hazards)
Jules Cohen & Associates
1725 DeSales Street, NW
Washington, D.C. 20036
202 659-3707

Cook, James H., Jr.(4.4 Satellite Earth Stations)
Scientific Atlanta, Inc.
Electronic Systems
3845 Pleasantdale Road
Atlanta, Georgia 30340
404 903-2468

DeCormier, William (2.4 Diplexers, Combiners, and Filters)
Dielectric Communications
Tower Hill Road
Raymond, Maine 04071
800 341-9678

Dye, Eric (2.6 FM Broadcast Antennas)
Jampro Antennas, Inc.
6939 Power Inn Road
Sacramento, California 95828
916 383-1177

Edison, Edward, P.E.(2.5 AM Broadcast Antennas)
Hammett & Edison, Inc.
Box 280068
San Francisco, California 94128-0068
415 342-5200

Ericksen, Dane E., P.E.
(7.7 Distance and Bearing Calculations, 7.8 The Emergency Broadcast System)
Hammett & Edison, Inc.
P.O. Box 280068
San Francisco, California 94128-0068
415 342-5200

Everist, Donald G. (2.10 AM Field Strength Measurements and Proof of Performance)
Cohen, Dippell and Everist, P.C.
1300 L Street, NW
Suite 1100
Washington, D.C. 20005
202 898-0111

FCC (1.1 FCC Organization and Administrative Procedures)
For more information on organization and procedures, contact:
William Hassinger
Mass Media Bureau, Room 314
Federal Communications Commission
1919 M Street, NW
Washington, DC 20554
202 632-6460

FCC (1.2 FCC Field Operations Bureau)
For more information on field operations, contact:
Richard M. Smith
Field Operations Bureau, Room 734
Federal Communications Commission
1919 M Street, NW
Washington, DC 20554
202 632-6980

Freda, Gerald (7.5 Closed Captioning)
National Captioning Institute
5203 Leesburg Pike, Suite 1500
Falls Church, Virginia 22041
703 998-2451

Gloeggler, Peter (5.2 The Broadcast Television Camera System)
Sony Business and Professional Group
4 Paragon Dr
Montvale, New Jersey 07645
201 358-4133

GrosJean, Jon P. (7.2 Radio Receivers)
Woodstock Engineering
310 Pulpit Rock Road
Woodstock, Connecticut 06281
203 974-2035

Hallikainen, Harold (3.7 Transmitter Cont Systems)
Hallikainen & Friends
141 Suburban Road
San Luis Obispo, California 93401-7590
805 541-0200

Hammett, William F. ,P.E. (2.9 RF Radiation Hazards)
Hammett & Edison, Inc.
Box 280068
San Francisco, California 94128-0068
415 342-5200

Happel, Warren P. (1.5 Environmental Hazards)
6791 Westwood Drive
Brecksville, Ohio 44141-2614
216 526-7147

Hickin, Ernest M. (4.2 Microwave Engineering for the Broadcaster)
Myton Associates Inc.
21 Sawmill Brook Road
Winchester, Massachusetts 01890-1056
617 721-0564

Kahn, Leonard (3.2 AM Stereo Broadcasting)
Kahn Communications, Inc.
222 Westbury Ave
Carle Place, New York 11514
516 222-2221

Kean, John (3.4 FM Stereo Broadcasting and SCA)
Moffet, Larson & Johnson
5203 Leesburg Pike, Suite 800
Falls Church, Virginia 22041
703 824-5660

Keller, Michael T. (5.7 Videotape Editing)
Keller Systems Technology
12 Prentiss St.
Southborough, Massachusetts 01772
508 485-3603

Kuca, Jay (5.3 Video Signal Switching, Timing, and Distribution)
Grass Valley Group
P.O. Box 1114 MS 8E
Grass Valley, California 95945
916 478-3000

Leonard, Robert D. (2.3 Transmission Lines)
Andrew Corporation
10500 West 153rd Street
Orland Park, Illinois 60462
708 349-3300

Lowery, Walt (5.9 Studio Audio)
Broadcast Supply West
7012 27th Street West
Tacoma, Washington 98466
206 565-2301

Lucia, Frank M. (7.8 The Emergency Broadcast System)
Office of Managing Director
FCC
Washington, D.C. 20554
202 632-3906

Marshall, Bill (5.14 Planning and Design of Studio Lighting)
Imero Fiorentino and Associates
33 West 60th Street
New York, New York 10023
212 246-0600

McClanahan, M. E. (4.3 Aural STL and Intercity Relay Service)
Marti Electronics Inc.
Box 661
Cleburne, Texas 76031
817 645-9163

McKee, Harry D. (2.3 Transmission Lines)
Andrew Corporation
10500 West 153rd Street
Orland Park, Illinois 60462
708 349-3300

Meintel, William (1.4 Frequency Allocations for Broadcasting and the Broadcast Auxiliary Services)
DATEL
1515 North Court House Road
Suite 301
Arlington, Virginia 22201
703 276-9007

Mendenhall, Geoffrey N. (3.3 FM Transmitters)
Broadcast Electronics, Inc.
4100 N. 24th Street, P.O. Box 3606
Quincy, Illinois 62305-3606
217 224-9600

Midkiff, John A. (2.3 Transmission Lines)
Andrew Corporation
10500 West 153rd Street
Orland Park, Illinois 60462
708 349-3300

Murch, Robert S. (5.11 Television Camera Robotics)
WPIX
220 E. 42nd Street
New York, New York 10017
212 210-2551

Orban, Robert (5.10 Transmission Audio Processing)
Orban, a division of AKG Acoustics Inc.
1525 Alvarado Street
San Leandro, California 94577
510 351-0500

Paulson, C. Robert (4.5 Fiber Optic Transmission Systems)
AVP Communication
15 Kay Street
Box 1356
Westborough, Massachusetts 01581
508-366-4694

Pizzi, Skip
(6.1 Radio Field Production, 6.4 Telco Audio Program Service)
Broadcast Engineering Magazine
P.O. Box 12901
Overland Park, Kansas 66282-2901
913 341-1300

Plonka, Robert (3.5 TV Transmitters)
Harris Corporation
Broadcast Division
P.O. Box 4290
Quincy, Illinois 62305-4290
217 222-8200

Pohlmann, Ken C. (5.5 Audio Magnetic Recording)
School of Music
University of Miami
P.O. Box 248165
Coral Gables, Florida 33124-7610
305 284-2439

Powis, Warren (1.6 Consultant Services: When and How to Use Them)
Cohen, Dippell and Everist, P.C.
1300 L Street, NW, Suite 1100
Washington, D.C. 20005
202 898-0111

Reiser, John (1.7 Source Guide: Broadcast Standards and Information)
Federal Communications Commission
Mass Media Bureau, Room 8330
2025 M Street, N.W.
Washington, DC 20554
202 254-3394

Reph, Ronald C. (5.4 Magnetic Recording Media)
3M Company
St. Paul, Minnesota 55144-1000
612 733-1110

Ritter, Norman C. (5.4 Magnetic Recording Media)
3M Company
St. Paul, Minnesota 55144-1000
612 733-1110

Rudman, Richard (1.3 Frequency Coordination)
Radio Station KFWB
6230 Yucca St.
Los Angeles, California 90028
213 871-4680

Salek, Stanley (7.1 Broadcast Audio Measurement)
Hammett & Edison, Inc.
Box 280068
San Francisco, California 94128-0068
415 342-5200

Salomon, Amnon (7.5 Closed Captioning)
National Captioning Institute
5203 Leesburg Pike, Suite 1500
Falls Church, Virginia 22041
703 998-2451

Sharrock, Michael P. (5.4 Magnetic Recording Media)
3M Company
St. Paul, Minnesota 55144-1000
612 733-1110

Shrestha, Mukunda B. (3.3 FM Transmitters)
Broadcast Electronics, Inc.
4100 N. 24th Street, P.O. Box 3606
Quincy, Illinois 62305-3606
217 224-9600

Silsby, Greg (5.1 Microphones)
Greg Silsby & Associates
29 Maplewood Ave.
Akron, Ohio 44313
216 869-0408

Slenker, Richard D. (5.11 Television Camera Robotics)
WTTG
5151 Wisconsin Avenue, N.W.
Washington, D.C. 20016-4124
202-244-5151

Smith, Carl E. (2.5 AM Broadcast Antennas)
Carl E. Smith Electronics
8200 Snowville Road
Cleveland, Ohio 44141
216 526-4386

Springer, Gary (4.4 Satellite Earth Stations)
Scientific Atlanta, Inc.
Network Systems Group
4356 Communications Drive
Norcross, Georgia 30093
404 925-5000

Stielper, Joseph W. (2.11 FM and TV Field Strength Measurements)
du Treil, Lundin & Rackley, Inc.
1019 19th Street, NW, 3rd Floor
Washington, D.C. 20036
202 223-6700

Tell, Richard A. (2.9 RF Radiation Hazards)
Richard Tell Associates, Inc.
6141 West Racel Street
Las Vegas, Nevada 89131-1912
702 645-3338

Unetich, Robert M. (7.10 TV Translators and LPTV; ITFS and MDS)
ITS Corporation
375 Valley Brook Rd.
McMurray, Pennsylvania 15317
412 941-1500

Van Buhler, Robert (1.3 Frequency Coordination)
Radio Station KNIX-FM/KCWW-AM/Real Country Network
600 E. Gilbert Dr.
Tempe, Arizona 85281
602 921-6311

Vaughan, Thomas J. (2.7 TV Broadcast Antennas)
Micro Communications, Inc.
Grenier Field
P. O. Box 4365
Manchester, New Hampshire 03108
603 624-4351

Vespoli, Jorge B. (4.4 Satellite Earth Stations)
Scientific Atlanta, Inc.
Network Systems Group
4356 Communications Drive
Norcross, Georgia 30093
404 925-5000

Weirather, Robert R. (3.5 TV Transmitters)
Harris Corporation
Broadcast Division
P.O. Box 4290
Quincy, Illinois 62305-4290
217 222-8200

Wentworth, John, P.E. (7.4 Color Television Receivers)
37 State Route 5
P.O. Box 259
Lovell, Maine 04051-0259
207 925-3058

Wetzel, Lew (2.1 Design, Erection, and Maintenance of Antenna Structures)
E G & G
35 Congress Street
Salem, Massachusetts 01970
508 745-3200

Whitaker, Jerry
(6.1 Radio Field Production, 7.9 AC Power Systems)
3830 NW 172 Place
Beaverton, Oregon 97006
503 690-8798

Williams, Edmund A. (3.6 Multichannel Television Sound)
Public Broadcasting Service
1320 Braddock Place
Alexandria, Virginia 22314
703 739-5172

Windle, John (2.1 Design, Erection, and Maintenance of Antenna Structures)
Stainless, Inc.
210 South Third Street
North Wales, Pennsylvania 19454
215 699-4871

Woodard, George, P.E. (3.1 AM Transmitters)
Radio Free Europe/Radio Liberty, Inc.
Oettingenstrasse 67
W-8000 Munich 22
Federal Republic of Germany
49-89-2102-3952

Wu, Joe (3.8 Frequency and Aural Modulation Monitoring)
TFT Inc.
3090 Oakmead Village Dr.
Santa Clara, California 95051
408 727-7272