REF ID:A64566

517974 1

[redacted]

~~*Confidential*~~

WAR DEPARTMENT

OFFICE OF THE CHIEF SIGNAL OFFICER

WASHINGTON

MILITARY CRYPTANALYSIS

PART II

NOTICE.—This document contains information affecting the national defense of the United States within the meaning of the Espionage Act (U. S. C., 50: 31, 32). The transmission of this document or the revelation of its contents in any manner to any unauthorized person is prohibited.

REF ID:A64566

Keep copy

clean

~~CONFIDENTIAL~~ 117974 2

~~Restricted~~

WAR DEPARTMENT
OFFICE OF THE CHIEF SIGNAL OFFICER
WASHINGTON

MILITARY CRYPTANALYSIS
Part II

SIMPLER VARIETIES OF POLYALPHABETIC SUBSTITUTION SYSTEMS

By
WILLIAM F. FRIEDMAN
Principal Cryptanalyst
Signal Intelligence Service

PREPARED UNDER THE DIRECTION OF THE
CHIEF SIGNAL OFFICER

(THIRD EDITION) - set

UNITED STATES
GOVERNMENT PRINTING OFFICE
WASHINGTON : 1941) - set

3/

~~CONFIDENTIAL~~

Over

517974 3

The Golden Guess
Is Morning-Star to the full round of Truth.
—Tennyson.

8 pt

c 130 36½ pi

(Reprint of original edition of 1938)

, with minor corrections and the addition of Appendix 3

517974 4-5

36½ pi.
8pt leaded

MILITARY CRYPTANALYSIS. PART II. SIMPLER VARIETIES OF POLYALPHABETIC SUBSTITUTION SYSTEMS

(III) – set

Italic

Appendix 3 – – – – – – 119 – ~~000~~ 126 – set

30 April 1959

This document is re-graded ~~"CONFIDENTIAL"~~ UP of DOD Directive 5200.1 dated 8 July 1957, and by authority of the Director, National Security Agency.

Paul S. Willard
Paul S. Willard
Colonel, AGC
Adjutant General

517974 6

SECTION I

INTRODUCTORY REMARKS

1. The essential difference between monoalphabetic and polyalphabetic substitution.—*a.* In the substitution methods thus far discussed it has been pointed out that their basic feature is that of monoalphabeticity. From the cryptanalytic standpoint, neither the nature of the cipher symbols, nor their method of production is an essential feature, although these may be differentiating characteristics from the cryptographic standpoint. It is true that in those cases designated as monoalphabetic substitution with variants or multiple equivalents, there is a departure, more or less considerable, from strict monoalphabeticity. In some of those cases, indeed, there may be available two or more wholly independent sets of equivalents, which, moreover, may even be arranged in the form of completely separate alphabets. Thus, while a loose terminology might permit one to designate such systems as polyalphabetic, it is better to reserve this nomenclature for those cases wherein polyalphabeticity is the essence of the method, specifically introduced with the purpose of imparting a *positional* variation in the substitutive equivalents for plain-text letters, in accordance with some rule directly or indirectly connected with the absolute *positions* the plain-text letters occupy in the message. This point calls for amplification.

b. In monoalphabetic substitution with variants the object of having different or multiple equivalents is to suppress, so far as possible by simple methods, the characteristic frequencies of the letters occurring in plain text. As has been noted, it is by means of these characteristic frequencies that the cipher equivalents can usually be identified. In these systems the varying equivalents for plain-text letters are subject to the free choice and caprice of the enciphering clerk; if he is careful and conscientious in the work, he will really make use of all the different equivalents afforded by the system; but if he is slip-shod and hurried in his work, he will use the same equivalents repeatedly rather than take pains and time to refer to the charts, tables, or diagrams to find the variants. Moreover, and this is a crucial point, even if the individual enciphering clerks are extremely careful, when many of them employ the same system it is entirely impossible to insure a complete diversity in the encipherments produced by two or more clerks working at different message centers. The result is inevitably to produce plenty of repetitions in the texts emanating from several stations, and when texts such as these are all available for study they are open to solution, by a comparison of their similarities and differences.

c. In true polyalphabetic systems, on the other hand, there is established a rather definite procedure which automatically determines the shifts or changes in equivalents or in the manner in which they are introduced, so that these changes are beyond the momentary whim or choice of the enciphering clerk. When the method of shifting or changing the equivalents is scientifically sound and sufficiently complex, the research necessary to establish the values of the cipher characters is much more prolonged and difficult than is the case even in complicated monoalphabetic substitution with variants, as will later be seen. These are the objects of true polyalphabetic substitution systems. The number of such systems is quite large, and it will be possible to

Over

describe in detail the cryptanalysis of only a few of the more common or typical examples of methods encountered in practical military communications.

d. The three methods, (1) single-equivalent monoalphabetic substitution, (2) monoalphabetic substitution with variants, and (3) true polyalphabetic substitution, show the following relationships as regards the equivalency between plain-text and cipher-text units:

A. In method (1), there is a set of 26 symbols; a plain-text letter is always represented by one and only one of these symbols; conversely, a symbol always represents the same plain-text letter. The equivalence between the plain-text and the cipher letters is constant in both encipherment and decipherment.

B. In method (2), there is a set of *n* symbols, where *n* may be any number greater than 26 and often is a multiple of that number; a plain-text letter may be represented by 1, 2, 3, . . . different symbols; conversely, a symbol always represents the same plain-text letter, the same as is the case in method (1). The equivalence between the plain-text and the cipher letters is variable in encipherment but constant in decipherment.[1]

C. In method (3) there is, as in the first method, a set of 26 symbols; a plain-text letter may be represented by 1, 2, 3, . . . 26 different symbols; conversely, a symbol may represent 1, 2, 3, . . . 26 different plain text letters, depending upon the system and the specific key. The equivalence between the plain-text and the cipher letters is variable in both encipherment and decipherment.

2. Primary classification of polyalphabetic systems.—*a.* A primary classification of polyalphabetic systems into two rather distinct types may be made: (1) periodic systems and (2) aperiodic systems. When the enciphering process involves a cryptographic treatment which is repetitive in character, and which results in the production of *cyclic phenomena* in the cryptographic text, the system is termed *periodic.* When the enciphering process is not of the type described in the foregoing general terms, the system is termed *aperiodic.* The substitution in both cases involves the use of two or more cipher alphabets.

b. The cyclic phenomena inherent in a periodic system may be exhibited externally, in which case they are said to be *patent*, or they may not be exhibited externally, and must be uncovered by a preliminary step in the analysis, in which case they are said to be *latent.* The periodicity may be quite definite in nature, and therefore determinable with mathematical exactitude allowing for no variability, in which case the periodicity is said to be *fixed.* In other instances the periodicity is more or less flexible in character and even though it may be deter-

[1] There is a monoalphabetic method in which the inverse result obtains, the correspondence being constant in encipherment but variable in decipherment; this is a method not found in the usual books on cryptography but in an essay on that subject by Edgar Allan Poe, entitled, in some editions of his works, *A few words on secret writing* and, in other editions, *Cryptography.* The method is to draw up an enciphering alphabet such as the following (using Poe's example):

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	S	U	A	V	I	T	E	R	I	N	M	O	D	O	F	O	R	T	I	T	E	R	I	N	R	E

In such an alphabet, because of repetitions in the cipher component, the plain-text equivalents are subject to a considerable degree of variability, as will be seen in the deciphering alphabet:

Cipher	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain	C			M	G	O			E				K	J	L			H	A	F	B	D				
					U				I					X	N			Q		R						
					Z				S						P			V		T						
									W									Y								

This type of variability gives rise to ambiguities in decipherment. A cipher group such as TIE, would yield such plain-text sequences as REG, FIG, TEU, REU, etc., which could be read only by *context.* No system of such a character would be practical for serious usage. For a further discussion of this type of cipher alphabet see Friedman, William F., *Edgar Allan Poe, Cryptographer*, Signal Corps Bulletins Nos. 97 (July-Sept.) and 98 (Oct.-Dec.), 1937. ~~1937-38.~~

8 pt leaded
36½ pi

minable mathematically, allowance must be made for a degree of variability subject to limits controlled by the specific system under investigation. The periodicity is in this case said to be *flexible*, or *variable within limits*.

3. Primary classification of periodic systems.—*a*. Periodic polyalphabetic substitution systems may primarily be classified into two kinds:

(1) Those in which only a few of a whole set of cipher alphabets are used in enciphering individual messages, these alphabets being employed repeatedly in a fixed sequence throughout each message. Because it is usual to employ a secret word, phrase, or number as a key to determine the number, identity, and sequence with which the cipher alphabets are employed, and this key is used over and over again in encipherment, this method is often called the *repeating-key system*, or the *repeating-alphabet system*. It is also sometimes referred to as the *multiple-alphabet system* because if the keying of the entire message be considered as a whole it is composed of multiples of a short key used repetitively.[2] In this text the designation "repeating-key system" will be used.

(2) Those in which all the cipher alphabets comprising the complete set for the system are employed one after the other successively in the encipherment of a message, and when the last alphabet of the series has been used, the encipherer begins over again with the first alphabet. This is commonly referred to as a *progressive-alphabet system* because the cipher alphabets are used in progression.

4. Sequence of study of polyalphabetic systems.—*a*. In the studies to be followed in connection with polyalphabetic systems, the order in which the work will proceed conforms very closely to the classifications made in paragraphs 2 and 3. Periodic polyalphabetic substitution ciphers will come first, because they are, as a rule, the simpler and because a thorough understanding of the principles of their analysis is prerequisite to a comprehension of how aperiodic systems are solved. But in the final analysis the solution of examples of both types rests upon the conversion or reduction of polyalphabeticity into monoalphabeticity. If this is possible, solution can always be achieved, granted there are sufficient data in the final monoalphabetic distributions to permit of solution by recourse to the ordinary principles of frequency.

b. First in the order of study of periodic systems will come the analysis of repeating-key systems. Some of the more simple varieties will be discussed in detail, with examples. Subsequently, ciphers of the progressive type will be discussed. There will then follow a more or less detailed treatment of aperiodic systems.

[2] French terminology calls this the "double-key method", but there is no logic in such nomenclature.

Over

SECTION II

CIPHER ALPHABETS FOR POLYALPHABETIC SUBSTITUTION

5. Classification of cipher alphabets upon the basis of their derivation.—*a.* The substitution processes in polyalphabetic methods involve the use of a plurality of cipher alphabets. The latter may be derived by various schemes, the exact nature of which determines the principal characteristics of the cipher alphabets and plays a very important role in the preparation and solution of polyalphabetic cryptograms. For these reasons it is advisable, before proceeding to a discussion of the principles and methods of analysis, to point out these various types of cipher alphabets, show how they are produced, and how the method of their production or derivation may be made to yield important clues and short-cuts in analysis.

b. A primary classification of cipher alphabets for polyalphabetic substitution may be made into the two following types:

(1) Independent or unrelated cipher alphabets.

(2) Derived or interrelated cipher alphabets.

c. Independent cipher alphabets may be disposed of in a very few words. They are merely separate and distinct alphabets showing no relationship to one another in any way. They may be compiled by the various methods discussed in Section IX of *Elementary Military Cryptography.* The solution of cryptograms written by means of such alphabets is rendered more difficult by reason of the absence of any relationship between the equivalents of one cipher alphabet and those of any of the other alphabets of the same cryptogram. On the other hand, from the point of view of practicability in their production and their handling in cryptographing and decryptographing, they present some difficulties which make them less favored by cryptographers than cipher alphabets of the second type.

d. Derived or interrelated alphabets, as their name indicates, are most commonly produced by the *interaction of* two primary components, which when juxtaposed at the various points of coincidence can be made to yield *secondary alphabets.*[1]

6. Primary components and secondary alphabets.—Two basic, slidable sequences or components of n characters each will yield n secondary alphabets. The components may be classified according to various schemes. For cryptanalytic purposes the following classification will be found useful:

CASE A. The primary components are both normal sequences.

(1) The sequences proceed in the same direction. (The secondary alphabets are direct standard alphabets.) (Pars. 13–15.)

(2) The sequences proceed in opposite directions. (The secondary alphabets are reversed standard alphabets; they are also reciprocal cipher alphabets.) (Par. 13*i*, 14*g*.)

CASE B. The primary components are not both normal sequences.

(1) The plain component is normal, the cipher component is a mixed sequence. (The secondary alphabets are mixed alphabets.) (Par. 16–25.)

[1] See Sec. VIII and IX, *Elementary Military Cryptography.*

(2) The plain component is a mixed sequence, the cipher component is normal. (The secondary alphabets are mixed alphabets.) (Par. 26.)

(3) Both components are mixed sequences.

(a) Components are identical mixed sequences.

I. Sequences proceed in the same direction. (The secondary alphabets are mixed alphabets.) (Par. 28.)

II. Sequences proceed in opposite directions. (The secondary alphabets are reciprocal mixed alphabets.) (Par. 38.)

(b) Components are different mixed sequences. (The secondary alphabets are mixed alphabets.) (Par. 39.)

7. Primary components, cipher disks, and square tables.—*a.* In preceding texts it has been shown that the equivalents obtainable from the use of quadricular or square tables may be duplicated by the use of revolving cipher disks or of sliding primary components. It was also stated that there are various ways of employing such tables, disks, and sliding components. Cryptographically the results may be quite diverse from different methods of using such paraphernalia, since the specific equivalents obtained from one method may be altogether different from those obtained from another method. But from the cryptanalytic point of view the diversity referred to is of little significance; only in one or two cases does the specific method of employing these cryptographic instrumentalities have an important bearing upon the procedure in cryptanalysis. However, it is advisable that the student learn something about these different methods before proceeding with further work.

b. There are, not *two*, but *four* letters involved in every case of finding equivalents by means of sliding primary components; furthermore, the determination of an equivalent for a given plain-text letter is representable by *two* equations involving *four* elements, usually letters. Three of these letters are by this time well-known to and understood by the student, viz, Θ_k, Θ_p, and Θ_c. The fourth element or letter has been passed over without much comment, but cryptographically it is just as important a factor as the other three. Its function may best be indicated by noting what happens when two primary components are juxtaposed, for the purpose of finding equivalents. Suppose these components are the following sequences:

```
(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X
```

Now suppose one is merely asked to find the equivalent of P_p when the key letter is K. Without further specification, the cipher equivalent cannot be stated; for it is necessary to know not only which K will be used as the key letter, the one in the component labeled (1) or the one in the component labeled (2), but also what letter the K_k will be set against, in order to juxtapose the two components. Most of the time, in preceding texts, these two factors have been tacitly assumed to be fixed and well understood: the K_k is sought in the mixed, or cipher component, and this K is set against A in the normal, or plain component. Thus:

```
                                   Plain      Index
                                     ↓          ↓
(1) Plain........ ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
(2) Cipher.......           FBPYRCQZIGSEHTDJUMKVALWNOX
                                     ↑          ↑
                                   Cipher      Key
```

With this setting $P_p = Z_c$.

Over

c. The letter A in this case may be termed the *index letter*, symbolized A_i. The index letter constitutes the fourth element involved in the two equations applicable to the finding of equivalents by sliding components. The four elements are therefore these:

(1) The key letter, Θ_k
(2) The index letter, Θ_i
(3) The plain-text letter, Θ_p
(4) The cipher letter, Θ_c

The index letter is commonly the initial letter of the component; but this, too, is only a convention. It *might be any letter* of the sequence constituting the component, as agreed upon by the correspondents. *However, in the subsequent discussion it will be assumed that the index letter is the initial letter of the component in which it is located, unless otherwise stated.*

d. In the foregoing case the enciphering equations are as follows:

$$\text{(I)}\ K_k=A_i;\ P_p=Z_c$$

But there is nothing about the use of sliding components which excludes other methods of finding equivalents than that shown above. For instance, despite the labeling of the two components as shown above, there is nothing to prevent one from seeking the plain-text letter in the component labeled (2), that is, the cipher component, and taking as its cipher equivalent the letter opposite it in the other component labeled (1). Thus:

```
           Cipher           Index
              ↓               ↓
(1) ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
(2)         FBPYRCQZIGSEHTDJUMKVALWNOX
              ↑               ↑
            Plain            Key
```

Thus:

$$\text{(II)}\ K_k=A_i;\ P_p=K_c$$

e. Since equations (I) and (II) yield different resultants, even with the same index, key, and plain-text letters, it is obvious that an accurate formula to cover a specific pair of enciphering equations must include data showing in what component each of the four letters comprising the equations is located. Thus, equations (I) and (II) should read:

(I) K_k in component (2)=A_i in component (1); P_p in component (1)=Z_c in component (2).
(II) K_k in component (2)=A_i in component (1); P_p in component (2)= K_c in component (1).

For the sake of brevity, the following notation will be used:

$$(1)\ K_{k/2}=A_{i/1};\ P_{p/1}=Z_{c/2}$$

$$(2)\ K_{k/2}=A_{i/1};\ P_{p/2}=K_{c/1}$$

f. Employing two sliding components and the four letters entering into an enciphering equation, there are, in all, twelve different resultants possible for the same set of components and the same set of four basic elements. These twelve differences in resultants arise from a set of twelve different enciphering conditions, as set forth below (the notation adopted in subparagraph *e* is used):

(1) $\Theta_{k/2}=\Theta_{i/1};\ \Theta_{p/1}=\Theta_{c/2}$
(2) $\Theta_{k/2}=\Theta_{i/1};\ \Theta_{p/2}=\Theta_{c/1}$
(3) $\Theta_{k/1}=\Theta_{i/2};\ \Theta_{p/1}=\Theta_{c/2}$
(4) $\Theta_{k/1}=\Theta_{i/2};\ \Theta_{p/2}=\Theta_{c/1}$
(5) $\Theta_{k/2}=\Theta_{p/1};\ \Theta_{i/1}=\Theta_{c/2}$
(6) $\Theta_{k/2}=\Theta_{c/1};\ \Theta_{i/1}=\Theta_{p/2}$
(7) $\Theta_{k/2}=\Theta_{p/1};\ \Theta_{i/2}=\Theta_{c/1}$
(8) $\Theta_{k/2}=\Theta_{c/1};\ \Theta_{i/2}=\Theta_{p/1}$
(9) $\Theta_{k/1}=\Theta_{p/2};\ \Theta_{i/1}=\Theta_{c/2}$
(10) $\Theta_{k/1}=\Theta_{c/2};\ \Theta_{i/1}=\Theta_{p/2}$
(11) $\Theta_{k/1}=\Theta_{p/2};\ \Theta_{i/2}=\Theta_{c/1}$
(12) $\Theta_{k/1}=\Theta_{c/2};\ \Theta_{i/2}=\Theta_{p/1}$

10 pt leaded
36½ pi – F.i.c

517974 - - - - 12

g. The twelve resultants obtainable from juxtaposing sliding components as indicated under the preceding subparagraph may also be obtained either from one square table, in which case twelve different methods of finding equivalents must be applied, or from twelve different square tables, in which case one standard method of finding equivalents will serve all purposes.

N

h. If but one table such as that shown below as Table ~~1~~–A is employed, the various methods of finding equivalents are difficult to keep in mind. I

TABLE I–A

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X
B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F
P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B
Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P
R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y
C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R
Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C
Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q
I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z
G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I
S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G
E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S
H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E
T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H
D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T
J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D
U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J
M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U
K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M
V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K
A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V
L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A
W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L
N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W
O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N
X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O

For example:

(1) For enciphering equations $\Theta_{k/2}=\Theta_{1/1}$; $\Theta_{p/1}=\Theta_{c/2}$:

Locate Θ_p in top sequence; locate Θ_k in first column;

Θ_c is letter within the square at intersection of the two lines thus determined.

Thus:

$$K_{k/2}=A_{1/1};\ P_{p/1}=Z_{c/2}$$

Over

(2) For enciphering equations $\Theta_{k/2}=\Theta_{1/1}$; $\Theta_{p/2}=\Theta_{c/1}$:

Locate Θ_k in first column; follow line to right to Θ_p; proceed up this column; Θ_c is letter at top.

Thus:

$$K_{k/2}=A_{1/1};\ P_{p/2}=K_{c/1}$$

(3) For enciphering equations $\Theta_{k/1}=\Theta_{1/2}$; $\Theta_{p/1}=\Theta_{c/2}$:

Locate Θ_k in top sequence and proceed down column to Θ_1;

Locate Θ_p in top sequence; Θ_c is letter at other corner of rectangle thus formed.

Thus:

$$K_{k/1}=A_{1/2};\ P_{p/1}=X_{c/2}$$

Only three different methods have been shown and the student no doubt already has encountered difficulty in keeping them segregated in his mind. It would obviously be very confusing to try to remember all twelve methods. But if one standard or fixed method of finding equivalents is followed with several different tables, then this difficulty disappears. Suppose that the following method is adopted: Arrange the square so that the plain-text letter may be sought in a separate sequence, arranged alphabetically, above the square and so that the key letter may be sought in a separate sequence, also arranged alphabetically, to the left of the square; look for the plain-text letter in the top row; locate the key letter in the 1st column to the left; find the letter standing within the square at the intersection of the vertical and horizontal lines thus determined. Then *twelve* squares, equivalent to the twelve different conditions listed in subparagraph *f*, can readily be constructed. They are all shown in Appendix 1, pp. 96–107.

i. When these square tables are examined carefully, certain interesting points are noted. In the first place, the tables may be paired so that one of a pair may serve for enciphering and the other of the pair may serve for deciphering, or vice versa. For example, tables I and II bear this reciprocal relationship to each other; III and IV, V and VI, VII and VIII, IX and X, XI and XII. In the second place, the internal dispositions of the letters, although the tables are derived from the same pair of components, are quite diverse. For example, in table I–B the horizontal sequences are identical with those of Table I-A but are merely displaced to the right and to the left different intervals set as necessary according to the successive key letters. Hence this square shows what may be termed a horizontally-displaced, direct symmetry of the cipher component. Vertically, it shows no symmetry, or if there is symmetry, it is not visible.[2] But when Table I–B is more carefully examined, an invisible, or indirect, vertical symmetry may be discerned where at first glance it is not apparent. If one takes any two *columns* of the table, it is found that the interval between the members of any pair of letters in one column is the same as the interval between the members of the homologous pair of letters in the other column, *if the distance is measured on the cipher component*. For example, consider the 2d and 15th columns (headed by L and I, respectively); take the letters P and G in the 2d column, and J and W in the 15th column. The distance between P and G on the cipher component is 7 intervals; the distance between J and W on the same component is *also* 7 intervals. This phenomenon implies a kind of hidden, or latent, or indirect symmetry within the cipher square. In fact, it may be stated that every table which sets forth in systematic fashion the various secondary alphabets derivable by sliding two primary sequences through all points of coincidence to find cipher equivalents must show some kind of symmetry, both horizontally and

[2] It is true that the first column within the table shows the plain-component sequence, but this is merely because the method of finding the equivalents in this case is such that this sequence is bound to appear in that column, since the successive key letters are A, B, C, . . . Z, and this sequence happens to be identical with the plain component in this case. The same is true of Tables V and XI; it is also applicable to the first row of Tables IX and X.

vertically. The symmetry may be termed *visible* or *direct,* if the sequences of letters in the rows (or columns) are the same throughout and are identical with that of one of the primary components; it may be termed *hidden* or *indirect* if the sequences of letters in the rows or columns are different, apparently not related to either of the components, but are in reality decimations of one of the primary components.

j. When the twelve tables of Appendix 1 are examined in the light of the foregoing remarks, the type of symmetry found in each may be summarized in the following manner:

Table	Horizontal				Vertical			
	Visible or direct		Invisible or indirect		Visible or direct		Invisible or indirect	
	Follows plain component	Follows cipher component	Follows plain component	Follows cipher component	Follows plain component	Follows cipher component	Follows plain component	Follows cipher component
I		x						x
II			x				x	
III		x				x		
IV			x		x			
V		x						x
VI			x				x	
VII	x						x	
VIII	x						x	
IX				x				x
X				x				x
XI			x		x			
XII		x				x		

Of these twelve types of cipher squares, corresponding to the twelve different ways of using a pair of sliding primary components to derive secondary alphabets, the ones best known and most often encountered in cryptographic studies are Tables I–B and II, referred to as being of the Vigenère type; Tables V and VI, referred to as being of the Beaufort type; and Tables IX and X, referred to as being of the Delastelle type. It will be noted that the tables of the Delastelle type show no direct or visible symmetry, either horizontally or vertically and because of this are supposed to yield more security than do any of the other types of tables. But it will presently be shown that the supposed increase in security is more illusory than real.

k. The foregoing facts concerning the various types of quadricular tables generated by diverse methods of using sliding primary components or their equivalent rotating cipher disks will be employed to good advantage, when the studies presently to be undertaken will bring the student to the place where he can comprehend them in the analysis of polyalphabetic systems. But in order not to confuse him with a multiplicity of details which have no direct bearing upon basic principles, one and only one standard method of finding equivalents by means of sliding components will be selected from among the twelve available, as set forth in the preceding subparagraphs. Unless otherwise stated, this method will be the one denoted by the first of the formulae listed in subpar. *f, viz:*

$$\Theta_{k/2}=\Theta_{1/1};\ \Theta_{p/1}=\Theta_{c/2}$$

Calling the plain component "1" and the cipher component "2", this will mean that the keyletter on the cipher component will be set opposite the index, which will be the first letter of the plain component; the plain-text letter to be enciphered will then be sought on the plain component and its equivalent will be the letter opposite it on the cipher component.

Over

SECTION III

THEORY OF SOLUTION OF REPEATING-KEY SYSTEMS

8. The three steps in the analysis of repeating-key systems.—*a.* The method of enciphering according to the principle of the repeating key, or repeating alphabets is adequately explained in Section XI of *Elementary Military Cryptography*, and no further reference need be made at this time. The analysis of a cryptogram of this type, regardless of the kind of cipher alphabets employed, or their method of production, resolves itself into three distinct and successive steps.

(1) Determination of the length of the repeating key, which is the same as the determination of the exact number of alphabets involved in the cryptogram;

(2) Allocation or distribution of the letters of the cipher text into the respective cipher alphabets to which they belong. This is the step which reduces the polyalphabetic text to monoalphabetic terms;

(3) Analysis of the individual monoalphabetic distributions to determine plain-text values of the cipher letters in each distribution or alphabet.

b. The foregoing steps will be treated in the order in which mentioned. The first step may be described briefly as that of *determining the period.* The second step may be described briefly as that of *reduction to monoalphabetic terms.* The third step may be designated as *identification of cipher-text values.*

9. First step: finding the length of the period.—*a.* The determination of the period, that is, the length of the key or the number of cipher alphabets involved in a cryptogram enciphered by the repeating-key method is, as a rule, a relatively simple matter. The cryptogram itself usually manifests externally certain phenomena which are the direct result of the use of a repeating key. The principles involved are, however, so fundamental in cryptanalysis that their elucidation warrants a somewhat detailed treatment. This will be done in connection with a short example of encipherment, shown in Fig. 1.

MESSAGE

THE ARTILLERY BATTALION MARCHING IN THE REAR OF THE ADVANCE GUARD KEEPS ITS COMBAT TRAIN WITH IT INSOFAR AS PRACTICABLE.

1

[Key: BLUE, using direct standard alphabets]

CIPHER ALPHABETS

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	(1)	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
	(2)	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
	(3)	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	(4)	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D

B L U E	B L U E	B L U E	B L U E
T H E A	A R D K	T H E A U S Y E	A R D K B C X O
R T I L	E E P S	R T I L S E C P	E E P S F P J W
L E R Y	I T S C	L E R Y M P L C	I T S C J E M G
B A T T	O M B A	B A T T C L N X	O M B A P X V E
A L I O	T T R A	A L I O B W C S	T T R A U E L E
N M A R	I N W I	N M A R O X U V	I N W I J Y Q M
C H I N	T H I T	C H I N D S C R	T H I T U S C X
G I N T	I N S O	G I N T H T H X	I N S O J Y M S
H E R E	F A R A	H E R E I P L I	F A R A G L L E
A R O F	S P R A	A R O F B C I J	S P R A T A L E
T H E A	C T I C	T H E A U S Y E	C T I C D E C G
D V A N	A B L E	D V A N E G U R	A B L E B M F I
C E G U		C E G U D P A Y	
a	*a*	*b*	*b*

CRYPTOGRAM

U S Y E S E C P M P L C C L N X B W C S O X U V D S C R H T

H X I P L I B C I J U S Y E E G U R D P A Y B C X O F P J W

J E M G P X V E U E L E J Y Q M U S C X J Y M S G L L E T A

L E D E C G B M F I

FIGURE 1.

Over

b. Regardless of what system is used, identical plain-text letters enciphered by the same cipher alphabet [1] must yield identical cipher letters. Referring to Fig. 1, such a condition is brought about every time that identical plain-text letters happen to be enciphered with the same key-letter, or every time identical plain-text letters fall into the same column in the encipherment.[2] Now since the number of columns or positions with respect to the key is very limited (except in the case of very long key words), and since the repetition of letters is an inevitable condition in plain text, it follows that there will be in a message of fair length many cases where identical plain-text letters *must* fall into the same column. They will thus be enciphered by the same cipher alphabet, resulting, therefore, in the production of many identical letters in the cipher text and these will represent identical letters in the plain text. When identical plain-text polygraphs fall into identical columns the result is the formation of identical cipher-text polygraphs, that is, repetitions of groups of 2, 3, 4, . . . letters are exhibited in the cryptogram. Repetitions of this type will hereafter be called *causal repetitions*, because they are produced by a definite, traceable cause, *viz*, the encipherment of identical letters by the same cipher alphabets.

c. It will also happen, however, that *different* plain-text letters falling in *different* columns will, by mere accident, produce identical cipher letters. Note, for example, in Fig. 1 that in Column 1, R_p becomes S_c and that in Column 2, H_p also becomes S_c. The production of an identical cipher text letter in these two cases (that is, a repetition where the plain-text letters are different and enciphered by different alphabets) is merely fortuitous. It is, in every day language, "a mere coincidence", or "an accident." For this reason repetitions of this type will hereafter be called *accidental repetitions*.

d. A consideration of the phenomenon pointed out in *c* makes it obvious that in polyalphabetic ciphers it is important that the cryptanalyst be able to tell whether the repetitions he finds in a specific case are causal or accidental in their origin, that is, whether they represent actual encipherments of identical plain-text letters by identical keying elements, or mere coincidences brought about purely fortuitously.

e. Now accidental repetitions will, of course, happen fairly frequently with individual letters, but less frequently with digraphs, because in this case the same kind of an "accident" must take place twice in succession. Intuitively one feels that the chances that such a purely fortuitous coincidence will happen two times in succession must be much less than that it will happen every once in a while in the case of single letters. Similarly, intuition makes one feel that the chances of such accidents happening in the case of three or more consecutive letters are still less than in the case of digraphs, decreasing very rapidly as the repetition increases in length.

f. The phenomena of cryptographic repetition may, fortunately, be dealt with statistically, thus taking the matter outside the realm of intuition and putting it on a firm mathematical or objective basis. Moreover, often the statistical analysis will tell the cryptanalyst when he has arranged or rearranged his text properly, that is, when he is approaching or has reached monoalphabeticity in his efforts to reduce polyalphabetic text to its simplest terms. However, in order to preserve continuity of thought it is deemed inadvisable to inject these statistical considerations at this place in the text proper; they have been incorporated in Appendix 2 hereof. The student is advised to study the Appendix very carefully after he has finished reading this section of the text.

g. At this point it will merely be indicated that if a cryptanalyst were to have at hand only the cryptogram of Fig. 1, with the repetitions underlined as below, a statistical study of the

[1] It is to be understood, of course, that cipher alphabets with single equivalents are meant in this case.

[2] The frequency with which this condition may be *expected* to occur can be definitely calculated. A discussion of this point falls beyond the scope of the present text.

number and length of the repetitions within the message (Par. 5 of Appendix 2) would tell him that while some of the digraphic repetitions may be accidental, the chances that they all are accidental are small. In the case of the tetragraphic repetition he would realize that the chances of its being accidental are very small indeed.

A.	USYES	ECPMP	LCCLN	XBWCS	OXUVD
B.	SCRHT	HXIPL	IBCIJ	USYEE	GURDP
C.	AYBCX	OFPJW	JEMGP	XVEUE	LEJYQ
D.	MUSCX	JYMSG	LLETA	LEDEC	GBMFI

h. A consideration of the facts therefore leads to but one conclusion, *viz,* that the repetitions exhibited by the cryptogram under investigation are *not accidental* but are *causal* in their origin; and the cause is in this case not difficult to find: repetitions in the plain text were actually enciphered by identical alphabets. In order for this to occur, it was necessary that the tetragraph USYE, for example, fall *both* times in *exactly* the same relative position with respect to the key. Note, for example, that USYE in Fig. 1 represents in both cases the plain-text polygraph THEA. The first time it occurred it fell in positions 1–2–3–4 with respect to the key; the second time it occurred it happened to fall in the very same relative positions, although it might just as well have happened to fall in any of the other three possible relative positions with respect to the key, *viz,* 2–3–4–1, 3–4–1–2, or 4–1–2–3. tr/

i. Lest the student be misled, however, a few more words are necessary on this subject. In the preceding subparagraph the word "happened" was used; this word correctly expresses the idea in mind, because the insertion or deletion of a single plain-text letter between the two occurrences would have thrown the second occurrence one letter forward or backward, respectively, and thus caused the polygraph to be enciphered by a sequence of alphabets such as can no longer produce the cipher polygraph USYE from the plain-text polygraph THEA. On the other hand, the insertion or deletion of this one letter might bring the letters of some other polygraph into similar columns so that some other repetition would be exhibited in case the USYE repetition had thus been suppressed.

j. The encipherment of similar letters by similar cipher alphabets is therefore the *cause* of the production of repetitions in the cipher text in the case of repeating-key ciphers. What principles can be derived from this fact, and how can they be employed in the solution of cryptograms of this type?

k. If a count is made of the number of letters from and including the first USYE to, but not including, the second occurrence of USYE, a total of 40 letters is found to intervene between the two occurrences. This number, 40, must, of course, be an exact multiple of the length of the key. Having the plain-text before one, it is easily seen that it is the 10th multiple; that is, the 4-letter key has repeated itself 10 times between the first and the second occurrence of USYE. It follows, therefore, that if the length of the key were not known, the number 40 could safely be taken to be an exact multiple of the length of the key; in other words, one of the *factors* of the number 40 would be equal to the length of the key. The word "safely" is used in the preceding sentence to mean that the interval 40 applies to a repetition of 4 letters and it has been shown that the chances that this repetition is accidental are small. The factors of 40 are 2, 4, 5, 8, 10, and 20. So far as this single repetition of USYE is concerned, if the length of the key were not known, all that could be said about the latter would be that it is equal to one of these factors. The repetition by itself gives no further indications. How can the exact factor be selected from among a list of several possible factors?

~~128801—3—41——2~~

Over

l. Let the intervals between all the repetitions in the cryptogram be listed. They are as follows:

Repetition	Interval	Factors
1st USYE to 2d USYE	40	2, 4, 5, 8, 10, 20.
1st BC to 2d BC	16	2, 4, 8.
1st CX to 2d CX	25	5.
1st EC to 2d EC	88	2, 4, 11, 22, 44.
1st LE to 2d LE	16	2, 4, 8.
2d LE to 3d LE	4	2, 4.
1st LE to 3d LE	20	2, 4, 5, 10.
1st JY to 2d JY	8	2, 4.
1st PL to 2d PL	24	2, 3, 4, 6, 8, 10, 12.
1st SC to 2d SC	52	2, 4, 13, 26.
(1st SY to 2d SY, already included in USYE.)		
(1st US to 2d US, already included in USYE.)		
2d US to 3d US	36	2, 3, 4, 6, 9, 18.
(1st US to 3d US, already included in USYE.)		
(1st YE to 2d YE, already included in USYE.)		

m. Are all these repetitions *causal* repetitions? It can be shown (Appendix 2, par. 4*c*) that the odds against a theory that the USYE repetition is accidental are about 99 to 1 (since the probability for its occurrence is .01). It can also be shown that the odds against a theory that the 10 digraphs which occur two or more times are accidental repetitions are over 4 to 1 (Appendix 2, par. 5*c*); the odds against a theory that the two digraphs which occur 3 times are accidental repetitions are quite large. (Probability is calculated to be about .06.) The chances are very great, therefore, that all or nearly all these repetitions are causal. Certainly the chances against the two occurrences of the tetragraph USYE and the three occurrences of the two different digraphs (LE and US) being accidental are quite high, and it is therefore not astonishing that the intervals between all the various repetitions, except in one case, contain the factors 2 and 4.

n. This means that if the cipher is written out in either 2 columns or 4 columns, all these repetitions (except the CX repetition) would fall into the same columns. From this it follows that the length of the key is either 2 or 4, the latter, on practical grounds, being more probable than the former. Doubts concerning the matter of choosing between a 2-letter and a 4-letter key will be dissolved when the cipher text is distributed into its component uniliteral frequency distributions.

o. The repeated digraph CX in the foregoing message is an accidental repetition, as will be apparent by referring to Fig. 1. Had the message been longer there would have been more such accidental repetitions, but, on the other hand, there would be a proportionately greater number of causal repetitions. This is because the phenomenon of repetition in plain text is so all-pervading.

p. Sometimes it happens that the cryptanalyst quickly notes a repetition of a polygraph of four or more letters, the interval between the first and second occurrences of which has only two factors, of which one is a relatively small number, the other a relatively high incommensurable number. He may therefore assume at once that the length of the key is equal to the smaller factor without searching for additional recurrences upon which to corroborate his assumption. Suppose, for example, that in a relatively short cryptogram the interval between the first and second occurrences of a polygraph of five letters happens to be a number such as 203, the factors of which are 7 and 29. Evidently the number of alphabets may at once be assumed to be 7, unless one is dealing with messages exchanged among correspondents known to use long keys. In the latter case one could assume the number of alphabets to be 29.

g. The foregoing method of determining the period in a polyalphabetic cipher is commonly referred to the literature as "factoring the intervals between repetitions"; or more often it is simply called "factoring." Because the latter is an apt term and is brief, it will be employed hereafter in this text to designate the process. in

10. General remarks on factoring.—*a*. The statement made in Par. 2 with respect to the cyclic phenomena said to be exhibited in cryptograms of the periodic type now becomes clear. The use of a short repeating key produces a periodicity of recurrences or repetitions collectively termed "cyclic phenomena", an analysis of which leads to a determination of the length of the period or cycle, and this gives the length of the key. Only in the case of relatively short cryptograms enciphered by a relatively long key does factoring fail to lead to the correct determination of the number of cipher alphabets in a repeating-key cipher; and of course, the fact that a cryptogram contains repetitions whose factors show constancy is in itself an indication and test of its periodic nature. It also follows that if the cryptogram is not a repeating-key cipher, then factoring will show no definite results, and conversely the fact that it does not yield definite results at once indicates that the cryptogram is not a periodic, repeating-key cipher.

b. There are two cases in which factoring leads to no definite results. One is in the case of monoalphabetic substitution ciphers. Here recurrences are very plentiful as a rule, and the intervals separating these recurrences may be factored, *but the factors will show no constancy;* there will be several factors common to many or most of the recurrences. This in itself is an indication of a monoalphabetic substitution cipher, if the very fact of the presence of many recurrences fails to impress itself upon the inexperienced cryptanalyst. The other case in which the process of factoring is nonsignificant involves certain types of nonperiodic, polyalphabetic ciphers. In certain of these ciphers recurrences of digraphs, trigraphs, and even polygraphs may be plentiful in a long message, but the intervals between such recurrences bear no definite multiple relation to the length of the key, such as in the case of the true periodic, repeating-key cipher, in which the alphabets change with successive letters and repeat themselves over and over again.

c. Factoring is not the only method of determining the length of the period of a periodic, polyalphabetic substitution cipher, although it is by far the most common and easily applied. At this point it will merely be stated that when the message under study is relatively short in comparison with the length of the key, so that there are only a few cycles of cipher text and no long repetitions affording a basis for factoring, there are several other methods available. However, it being deemed inadvisable to interject the data concerning those other methods at this point, they will be explained subsequently. It is desirable at this juncture merely to indicate that methods other than factoring do exist and are used in practical work.

d. Fundamentally, the factoring process is merely a more or less simple mathematical method of studying the phenomena of periodicity in cryptograms. It will usually enable the cryptanalyst to ascertain definitely whether or not a given cryptogram is periodic in nature, and if so, the length of the period, *stated in terms of the cryptographic unit involved.* By the latter statement is meant that the factoring process may be applied not only in analyzing the periodicity manifested by cryptograms in which the plain-text units subjected to cryptographic treatment are monographic in nature (i. e. are single letters) but also in studying the periodicity exhibited by those occasional cryptograms wherein the plain-text units are digraphic, trigraphic, or *n*-graphic in character. The student should bear this point in mind when he comes to the study of substitution systems of the latter sort. However, the present text will deal solely with cases of the former type, wherein the plain-text units subjected to cryptographic treatment are single letters.

Over

11. Second step: distributing the cipher text into the component monoalphabets.—*a.* After the number of cipher alphabets involved in the cryptogram has been ascertained, the next step is to rewrite the message in groups corresponding to the length of the key, or in columnar fashion, whichever is more convenient, and this automatically divides up the text so that the letters belonging to the same cipher alphabet occupy similar positions in the groups, or, if the columnar method is used, fall in the same column. The letters are thus allocated or distributed into the respective cipher alphabets to which they belong. This reduces the polyalphabetic text to monoalphabetic terms.

b. Then separate uniliteral frequency distributions for the thus isolated individual alphabets are compiled. For example, in the case of the cipher on page 13, having determined that four alphabets are involved, and having rewritten the message in four columns, a frequency distribution is made of the letters in Column 1, another is made of the letters in Column 2, and so on for the rest of the columns. *Each of the resulting distributions is therefore a monoalphabetic frequency distribution.* If these distributions do not give the characteristic irregular crest and trough appearance of monoalphabetic frequency distributions, then the analysis which led to the hypothesis as regards the number of alphabets involved is fallacious. In fact, the appearance of these individual distributions may be considered to be an index of the correctness of the factoring process; for theoretically, and practically, the individual distributions constructed upon the *correct* hypothesis will tend to conform more closely to the irregular crest and trough appearacne of a monoalphabetic frequency distribution than will the graphic tables constructed upon an incorrect hypothesis. These individual distributions may also be tested for monoalphabeticity by statistical methods.

12. Third step: solving the monoalphabetic distributions.—The difficulty experienced in analyzing the individual or isolated frequency distributions depends mostly upon the type of cipher alphabets that is used. It is apparent that mixed alphabets may be used just as easily as standard alphabets, and, of course, the cipher letters themselves give no indication as to which is the case. However, just as it was found that in the case of monoalphabetic substitution ciphers, a uniliteral frequency distribution gives clear indications as to whether the cipher alphabet is a standard or a mixed alphabet, by the relative positions and extensions of the crests and troughs in the table, so it is found that in the case of repeating-key ciphers, uniliteral frequency distributions for the isolated or individual alphabets will also give clear indications as to whether these alphabets are standard alphabets or mixed alphabets. Only one or two such frequency distributions are necessary for this determination; if they appear to be standard alphabets, similar distributions can be made for the rest of the alphabets; but if they appear to be mixed alphabets, then it is best to compile triliteral frequency distributions for all the alphabets. The analysis of the values of the cipher letters in each table proceeds along the same lines as in the case of monoalphabetic ciphers. The analysis is more difficult only because of the reduced size of the tables, but if the message be very long, then each frequency distribution will contain a sufficient number of elements to enable a speedy solution to be achieved.

SECTION IV

REPEATING-KEY SYSTEMS WITH STANDARD CIPHER ALPHABETS

13. Solution by applying principles of frequency.—*a.* In the light of the foregoing principles, let the following cryptogram be studied:

MESSAGE

	1	2	3	4	5
A.	A U K H Y	J A M K I	Z Y M W M	J M I G X	N F M L X
B.	E T I M I	Z H B H R	A Y M Z M	I L V M E	J K U T G
C.	D P V X K	Q U K H Q	L H V R M	J A Z N G	G Z V X E
D.	N L U F M	P Z J N V	C H U A S	H K Q G K	I P L W P
E.	A J Z X I	G U M T V	D P T E J	E C M Y S	Q Y B A V
F.	A L A H Y	P O E X W	P V N Y E	E Y X E E	U D P X R
G.	B V Z V I	Z I I V O	S P T E G	K U B B R	Q L L X P
H.	W F Q G K	N L L L E	P T I K W	D J Z X I	G O I O I
J.	Z L A M V	K F M W F	N P L Z I	O V V F M	Z K T X G
K.	N L M D F	A A E X I	J L U F M	P Z J N V	C A I G I
L.	U A W P R	N V I W E	J K Z A S	Z L A F M	H S

A search for repetitions discloses the following short list with the intervals and factors above 10 omitted (for previous experience may lead to the conclusion that it is unlikely that the cryptogram involves more than 10 alphabets, showing the number of recurrences which it does):

Repetition	Location	Interval	Factors
LUFMPZJNVC	D1, K3	160	2, 4, 5, 8, 10.
JZXIG	E1, H4	90	2, 3, 5, 6, 9, 10.
EJK	B4, L2	215	5.
PTE	E3, G3	50	2, 5, 10.
QGK	D4, H1	85	5.
UKH	A1, C2	55	5.
ZLA	J1, L4	65	5.
AS	D3, L3	175	3, 5, 7,
EJ	B4, L2	115	5.
FM	A5, D1	57	3.
FM	A5, J2	185	5.
FM	J2, J4	12	2, 3, 4, 6.
FM	J4, K3	20	2, 4, 5, 10.
FM	K3, L4	30	2, 3, 5, 6, 10.
JA	A2, C4	60	2, 3, 4, 5, 6, 10.
LA	F1, J1	75	3, 5.
LA	J1, L4	65	5.
LL	G5, H2	10	2, 5.
NL	D1, H2	105	3, 5, 7.
NL	H2, K1	45	3, 5, 9.
VX	C1, C5	20	2, 4, 5, 10.
YM	A3, B3	25	5.

Over

b. The factor 5 appears in all but two cases, each of which involves only a digraph. It seems almost certain that the number of alphabets is five. Since the text already appears in groups of five letters, it is unnecessary to rewrite the message. The next step is to make a uniliteral frequency distribution for Alphabet 1 to see if it can be determined whether or not standard alphabets are involved. It is as follows:

Alphabet 1

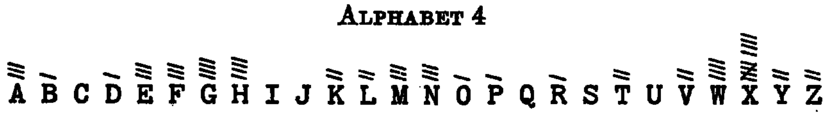
c. Although the indications are not very clear cut, yet if one takes into consideration the small amount of data the assumption of a direct standard alphabet with $W_c=A_p$, is worth further test. Accordingly a similar distribution is made for Alphabet 2.

Alphabet 2

d. There is every indication of a direct standard alphabet, with $H_c=A_p$. Let similar distributions be made for the last three alphabets. They are as follows:

Alphabet 3

Alphabet 4



Alphabet 5

e. After but little experiment it is found that the distributions can best be made to fit the normal when the following values are assumed:

Alphabet 1.......... $A_p=W_c$
Alphabet 2.......... $A_p=H_c$
Alphabet 3.......... $A_p=I_c$
Alphabet 4.......... $A_p=T_c$
Alphabet 5.......... $A_p=E_c$

f. Note the key word given by the successive equivalents of A_p: WHITE. The real proof of the correctness of the analysis is, of course, to test the values of the solved alphabets on the cryptogram. The five complete cipher alphabets are as follows:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	1	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	2	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
	3	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
	4	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	5	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D

Figure 2.

g. Applying these values to the first few groups of our message, the following is found:

	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Cipher	A U K H Y	J A M K I	Z Y M W M	J M I G X	N F M L X . . .
Plain	E N C O U	N T E R E	D R E D I	N F A N T	R Y E S T . . .

h. Intelligible text at once results, and the solution can now be completed very quickly. The complete message is as follows:

ENCOUNTERED RED INFANTRY ESTIMATED AT ONE REGIMENT AND MACHINE GUN COMPANY IN TRUCKS NEAR EMMITSBURG. AM HOLDING MIDDLE CREEK NEAR HILL 543 SOUTHWEST OF FAIRPLAY. WHEN FORCED BACK WILL CONTINUE DELAYING REDS AT MARSH CREEK. HAVE DESTROYED BRIDGES ON MIDDLE CREEK BETWEEN EMMITSBURG–TANEYTOWN ROAD AND RHODES MILL.

i. In the foregoing example (which is typical of the system erroneously attributed, in cryptographic literature, to the French cryptographer Vigenère, although to do him justice, he made no claim of having "invented" it), direct standard alphabets were used, but it is obvious that reversed standard alphabets may be used and the solution accomplished in the same manner. In fact, the now obsolete cipher disk used by the United States Army for a number of years yields exactly this type of cipher, which is also known in the literature as the Beaufort Cipher, and by other names. In fitting the isolated frequency distributions to the normal, the direction of "reading" the crests and troughs is merely reversed.

14. Solution by completing the plain-component sequence.—*a.* There is another method of solving this type of cipher, which is worthwhile explaining, because the underlying principles will be found useful in many cases. It is a modification of the method of solution by completing the plain-component sequence, already explained in *Military Cryptanalysis*, Part I.

b. After all, the individual alphabets of a cipher such as the one just solved are merely direct standard alphabets. It has been seen that monoalphabetic ciphers in which standard cipher alphabets are employed may be solved almost mechanically by completing the plain-component sequence. The plain text reappears on only one generatrix and this generatrix is the same for the whole message. It is easy to pick this generatrix out of all the other generatrices because it is the only one which yields intelligible text. Is it not apparent that if the same process is applied to the cipher letters of the *individual alphabets* of the cipher just solved that the plain-text equivalents of these letters must all reappear on one and the same generatrix? But how will the generatrix which actually contains the plain-text letters be distinguishable from the other generatrices, since these plain-text letters are not consecutive letters in the plain text but only letters separated from one another by a constant interval? The answer is simple. The plain-text generatrix should be distinguishable from the others *because it will show more and a better assortment of high-frequency letters, and can thus be selected by the eye from the whole set of generatrices.* If this is done with all the alphabets in the cryptogram, it will merely be necessary to assemble the letters of the thus selected generatrices in proper order, and the result sould be consecutive letters forming intelligible text.

c. An example will serve to make the process clear. Let the same message be used as before. Factoring showed that it involves five alphabets. Let the first ten cipher letters *in each alphabet* be set down in a horizontal line and let the normal alphabet sequences be completed. Thus:

Over

	Alphabet 1	Alphabet 2	Alphabet 3	Alphabet 4	Alphabet 5
1	AJZJNEZAIJ	UAYMFTHYLK	KMMIMIBMVU	HKWGLMHZMT	YIMXXIRMEG
2	BKAKOFABJK	VBZNGUIZML	LNNJNJCNWV	ILXHMNIANU	ZJNYYJSNFH
3	CLBLPGBCKL	WCAOHVJANM	MOOKOKDOXW	JMYINOJBOV	AKOZZKTOGI
4	DMCMQHCDLM	XDBPIWKBON	NPPLPLEPYX	KNZJOPKCPW	BLPAALUPHJ
5	ENDNRIDEMN	YECQJXLCPO	OQQMQMFQZY	LOAKPQLDQX	CMQBBMVQIK
6	FOEOSJEFNO	ZFDRKYMDQP	PRRNRNGRAZ	MPBLQRMERY	DNRCCNWRJL
7	GPFPTKFGOP	AGESLZNERQ	QSSOSOHSBA	NQCMRSNFSZ	EOSDDOXSKM
8	HQGQULGHPQ	BHFTMAOFSR	RTTPTPITCB	ORDNSTOGTA	FPTEEPYTLN
9	IRHRVMHIQR	CIGUNBPGTS	SUUQUQJUDC	PSEOTUPHUB	GQUFFQZUMO
10	JSISWNIJRS	DJHVOCQHUT	TVVRVRKVED	QTFPUVQIVC	HRVGGRAVNP
11	KTJTXOJKST	EKIWPDRIVU	UWWSWSLWFE	RUGQVWRJWD	ISWHHSBWOQ
12	LUKUYPKLTU	FLJXQESJWV	VXXTXTMXGF	SVHRWXSKXE	JTXIITCXPR
13	MVLVZQLMUV	GMKYRFTKXW	WYYUYUNYHG	TWISXYTLYF	KUYJJUDYQS
14	NWMWARMNVW	HNLZSGULYX	XZZVZVOZIH	UXJTYZUMZG	LVZKKVEZRT
15	OXNXBSNOWX	IOMATHVMZY	YAAWAWPAJI	VYKUZAVNAH	MWALLWFASU
16	PYOYCTOPXY	JPNBUIWNAZ	ZBBXBXQBKJ	WZLVABWOBI	NXBMMXGBTV
17	QZPZDUPQYZ	KQOCVJXOBA	ACCYCYRCLK	XAMWBCXPCJ	OYCNNYHCUW
18	RAQAEVQRZA	LRPDWKYPCB	BDDZDZSDML	YBNXCDYQDK	PZDOOZIDVX
19	SBRBFWRSAB	MSQEXLZQDC	CEEAEATENM	ZCOYDEZREL	QAEPPAJEWY
20	TCSCGXSTBC	NTRFYMARED	DFFBFBUFON	ADPZEFASFM	RBFQQBKFXZ
21	UDTDHYTUCD	OUSGZNBSFE	EGGCGCVGPO	BEQAFGBTGN	SCGRRCLGYA
22	VEUEIZUVDE	PVTHAOCTGF	FHHDHDWHQP	CFRCGHCUHO	TDHSSDMHZB
23	WFVFJAVWEF	QWUIBPDUHG	GIIEIEXIRQ	DGSCHIDVIP	UEITTENIAC
24	XGWGKBWXFG	RXVJCQEVIH	HJJFJFYJSR	EHTDIJEWJQ	VFJUUFOJBD
25	YHXHLCXYGH	SYWKDRFWJI	IKKGKGZKTS	FIUEJKFXKR	WGKVVGPKCE
26	ZIYIMDYZHI	TZXLESGXKJ	JLLHLHALUT	GJVFKLGYLS	XHLWWHQLDF

FIGURE 3.

d. If the high-frequency generatrices underlined in Figure 3 are selected and their letters are juxtaposed *in columns* the consecutive letters of intelligible plain text immediately present themselves. Thus:

Selected Generatrices:

For Alphabet 1, generatrix 5	....	E N D N R I D E M N
For Alphabet 2, generatrix 20	....	N T R F Y M A R E D
For Alphabet 3, generatrix 19	....	C E E A E A T E N M
For Alphabet 4, generatrix 8	....	O R D N S T O G T A
For Alphabet 5, generatrix 23	....	U E I T T E N I A C

Columnar juxtaposition of letters from selected generatrices......

1	2	3	4	5
E	N	C	O	U
N	T	E	R	E
D	R	E	D	I
N	F	A	N	T
R	Y	E	S	T
I	M	A	T	E
D	A	T	O	N
E	R	E	G	I
M	E	N	T	A
N	D	M	A	C

FIGURE 4.

Plain text: ENCOUNTERED RED INFANTRY ESTIMATED AT ONE
REGIMENT AND MAC . . . .

e. Solution by this method can thus be achieved without the compilation of any frequency tables whatever and is very quickly attained. The inexperienced cryptanalyst may have difficulty at first in selecting the generatrices which contain the most and the best assortment of high-frequency letters, but with increased practice, a high degree of proficiency is attained. After all it is only a matter of experiment, trial, and error to select and assemble the proper generatrices so as to produce intelligible text.

f. If the letters on the sliding strips were accompanied by numbers representing their relative frequencies in plain text, and these numbers were added *across* each generatrix, then that generatrix with the highest total frequency would *theoretically* always be the plain-text generatrix. Practically it will be among the generatrices which show the first three or four greatest totals. Thus, an entirely mathematical solution for this type of cipher may be applied.

g. If the cipher alphabets are reversed standard alphabets, it is only necessary to convert the cipher letters of each isolated alphabet into their normal, plain-component equivalents and then proceed as in the case of direct standard alphabets.

h. It has been seen how the key word may be discovered in this type of cryptogram. Usually the key is made up of those letters in the successive alphabets whose equivalents are A_p but other conventions are of course possible. Sometimes a key number is used, such as 8-4-7-1-12, which means merely that A_p is represented by the eighth letter from A (in the normal alphabet) in the first cipher alphabet, by the fourth letter from A in the second cipher alphabet, and so on. This modification is known in the literature as the Gronsfeld cipher. However, the method of solution as illustrated above, being independent of the nature of the key, is the same as before.

15. Solution by the "probable-word method."—*a.* The common use of key words in cryptograms such as the foregoing makes possible a method of solution that is simple and can be used where the more detailed method of analysis using frequency distributions or by completing the plain-component sequence is of no avail. In the case of a very short message which may show no recurrences and give no indications as to the number of alphabets involved, this modified method will be found most useful.

b. Briefly, the method consists in assuming the presence of a probable word in the message, and referring to the alphabets to find the key letters applicable when this hypothetical word is assumed to be present in various positions in the cipher text. If the assumed word happens to be correct, and is placed in the correct position in the message, the key letters produced by referring to the alphabets will yield the key word. In the following example it is assumed that reversed standard alphabets are known to be used by the enemy.

MESSAGE

M D S T J L Q C X C K Z A S A N Y Y K O L P

c. Extraneous circumstances lead to the assumption of the presence of the word AMMUNITION. One may assume that this word begins the message. Using sliding normal components, one reversed, the other direct, the key letters are ascertained by noting what the successive equivalents of A_p are. Thus:

Cipher	M	D	S	T	J	L	Q	C	X	C
Plain text	A	M	M	U	N	I	T	I	O	N
"Key"	M	P	E	N	W	T	J	K	L	P

Over

The key does not spell any intelligible word. One therefore shifts the assumed word one letter forward and another trial is made.

Cipher	D	S	T	J	L	Q	C	X	C	K
Plain text	A	M	M	U	N	I	T	I	O	N
"Key"	D	E	F	D	Y	Y	V	F	Q	X

This also yields no intelligible key word. One continues to shift the assumed word forward one space at a time until the following point is reached.

Cipher	L	Q	C	X	C	K	Z	A	S	A
Plain text	A	M	M	U	N	I	T	I	O	N
"Key"	L	C	O	R	P	S	S	I	G	N

The key now becomes evident. It is a cyclic permutation of SIGNAL CORPS. It should be clear that since the key word or key phrase repeats itself during the encipherment of such a message, the plain-text word upon whose assumed presence in the message this test is being based may begin to be enciphered at any point in the key, and continue over into its next repetition if it is longer than the key. When this is the case it is merely necessary to shift the latter part of the sequence of key letters to the first part, as in the case noted: LCORPSSIGN is transposed into SIGN . . . LCORPS, and thus SIGNAL CORPS.

d. It will be seen in the foregoing method of solution that the length of the key is of no particular interest or consequence in the steps taken in effecting the solution. The determination of the length and elements of the key comes after the solution rather than before it. In this case the length of the period is seen to be eleven, corresponding to the length of the key (SIGNAL CORPS).

e. The foregoing method is one of the other methods of determining the length of the key (besides factoring), referred to in Par. 10*c*.

f. If the assumption of reversed standard alphabets yields no good results, then direct standard alphabets are assumed and the test made exactly in the same manner. As will be shown subsequently, the method can also be used as a last resort when mixed alphabets are employed.

g. When the assumed word is longer than the key, the sequence of recovered key letters will show a periodicity equal to the length of the key; that is, after a certain number of letters the sequence of key letters will repeat. This phenomenon would be most useful in the case of keys that are not intelligible words but are composed of random letters or figures. Of course, if such a key is longer than the assumed word, this method is of no avail.

h. This method of solution by searching for a word is contingent upon the following circumstances:

(1) That the word whose presence is assumed actually occurs in the message, is properly spelled, and correctly enciphered.

(2) That the sliding components (or equivalent cipher disks or squares) employed in the search for the assumed word are actually the ones which were employed in the encipherment, or are such as to give identical results as the ones which were actually used.

(3) That the pair of enciphering equations used in the test is actually the pair which was employed in the encipherment; or if a cipher square is used in the test, the method of finding equivalents gives results that correspond with those actually obtained in the encipherment. (See par. 9.)

i. The foregoing appears to be quite an array of contingencies and the student may think that on this account the method will often fail. But examining these contingencies one by one, it will be seen that successful application of the method may not be at all rare—after the solution of some messages has disclosed what sort of paraphernalia and methods of employing them are favored by the enemy. From the foregoing remark it is to be inferred that the probable-word method has its greatest usefulness not in an initial solution of a system, but only after successful study of enemy communications by more difficult processes of analysis has told its story to the alert cryptanalyst. Although it is commonly attributed to Bazeries, the French cryptanalyst of 1900, the probable-word method is very old in cryptanalysis and goes back several centuries. Its usefulness in practical work may best be indicated by quoting from a competent observer [1]:

> There is another [method] which is to this first method what the geometric method is to analysis in certain sciences, and, according to the whims of individuals, certain cryptanalysts prefer one to the other. Certain others, incapable of getting the answer with one of the methods in the solution of a difficult problem, conquer it by means of the other, with a disconcerting masterly stroke. This other method is that of the probable word. We may have more or less definite opinions concerning the subject of the cryptogram. We may know something about its date, and the correspondents, who may have been indiscreet in the subject they have treated. On this basis, the hypothesis is made that a certain word probably appears in the text. . . . In certain classes of documents, military or diplomatic telegrams, banking and mining affairs, etc., it is not impossible to make very important assumptions about the presence of certain words in the text. After a cryptanalyst has worked for a long time with the writings of certain correspondents, he gets used to their expressions. He gets a whole load of words to try out; then the changes of key, and sometimes of system, no longer throw into his way the difficulties of an absolutely new study, which might require the analytical method.

2lo

[1] Givierge, M., *Cours de Cryptographie*, Paris, 1925, p. 30.

To which I am prompted to add the amusing definition of cryptanalysis attributed to a British wag: "All cryptanalysis is divided into two parts: trance-titution and supposition."

10pt

This is not part of the quotation. Do not set.

Over

Section V

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, I

16. Reason for the use of mixed alphabets.—*a.* It has been seen in the examples considered thus far that the use of several alphabets in the same message does not greatly complicate the analysis of such a cryptogram. There are three reasons why this is so. Firstly, only relatively few alphabets were employed; secondly, these alphabets were employed in a periodic or repeating manner, giving rise to cyclic phenomena in the cryptogram, by means of which the number of alphabets could be determined; and, thirdly, the cipher alphabets were *known* alphabets, by which is meant merely that the sequences of letters in both components of the cipher alphabets were known sequences.

b. In the case of monoalphabetic ciphers it was found that the use of a mixed alphabet delayed the solution to a considerable degree, and it will now be seen that the use of mixed alphabets in polyalphabetic ciphers renders the analysis much more difficult than the use of standard alphabets, but the solution is still fairly easy to achieve.

17. Interrelated mixed alphabets.—*a.* It was stated in Par. 5 that the method of producing the mixed alphabets in a polyalphabetic cipher often affords clues which are of great assistance in the analysis of the cipher alphabets. This is so, of course, only when the cipher alphabets are interrelated secondary alphabets produced by sliding components or their equivalents. Reference is now made to the classification set forth in Par. 6, in connection with the types of alphabets which may be employed in polyalphabetic substitution. It will be seen that thus far only Cases A (1) and (2) have been treated. Case B (1) will now be discussed.

b. Here one of the components, the plain component, is the normal sequence, while the cipher component is a mixed sequence, the various juxtapositions of the two components yielding mixed alphabets. The mixed component may be a systematically-mixed or a random-mixed sequence. If the 25 successive displacements of the mixed component are recorded in separate lines, a symmetrical cipher square such as that shown in Fig. 5 results therefrom. It is identical in form with the square table shown on p. 7, labeled Table I–A.

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z
	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L
	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E
	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A
	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V
	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N
	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W
	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O
	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R
	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T
	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H
	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B
Cipher	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C
	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D
	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F
	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G
	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I
	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J
	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K
	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M
	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P
	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q
	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S
	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U
	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X
	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y

FIGURE 5.

c. Such a cipher square may be used in exactly the same manner as the Vigenère square. With the key word BLUE and conforming to the normal enciphering equations ($\theta_{k/2}=\theta_{1/1}$; $\theta_{p/1}=\theta_{c/2}$), the following lines of the square would be used:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H
L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z
U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S
E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L

FIGURE 6a.

These lines would, of course, yield the following cipher alphabets:

(1)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H
(2)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z
(3)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher	U	X	Y	Z	L	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S
(4)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher	E	A	V	N	W	O	R	T	H	B	C	D	F	G	I	J	K	M	P	Q	S	U	X	Y	Z	L

FIGURE 6b.

Over

18. Principles of direct symmetry of position.—*a*. It was stated directly above that Fig. 5 is a symmetrical cipher square, by which is meant that the letters in its successive horizontal lines show a *symmetry of position* with respect to one another. They constitute, in reality, one and only one sequence or series of letters, the sequences being merely displaced successively 1, 2, 3, . . . intervals. The symmetry exhibited is obvious and is said to be visible, or *direct*. This fact can be used to good advantage, as has already been alluded to in par. 7*j*.

b. Consider, for example, the pair of letters G_c and V_c in cipher alphabet (1) of Fig. 6*b*. The letter V_c is the 15th letter to the right of G_c. In cipher alphabet (2), V_c is also the 15th letter to the right of G_c, as is the case in each of the four cipher alphabets in Fig. 6*b*, since the *relative* positions they occupy are the same in each horizontal line in Fig. 6*a*, that is, in each of the successive recordings of the cipher component as the latter is slid to the right against the plain or normal component. If, therefore, the relative positions occupied by two letters, Θ_1 and Θ_2, in such a cipher alphabet, C_1, are known, and if the position of Θ_1 in another cipher alphabet, C_2, belonging to the same series is known, then Θ_2 may at once be placed into its correct position in C_2. Suppose, for example, that as the result of an analysis based upon considerations of frequency, the following values in four cipher alphabets have been tentatively determined:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
(1)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher					G										Y					V						
(2)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher					N										G					P						
(3)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher					L										B					I						
(4)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher					W										I					Q						

FIGURE 7*a*.

c. The cipher components of these four secondary alphabets may, for convenience, be assembled into a cellular structure, hereinafter called a *sequence reconstruction skeleton*, as shown in Fig. 7*b*. Regarding the top line of the reconstruction skeleton in Fig. 7*b* as being common to all four secondary cipher alphabets listed in Fig. 7*a*, the successive lines of the reconstruction skeleton may now be termed cipher alphabets, and may be referred to by the numbers at the left.

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	1					G										Y					V						
	2					N										G					P						
	3					L										B					I						
	4					W										I					Q						

FIGURE 7*b*.

d. The letter G is common to Alphabets 1 and 2. In Alphabet 2 it is noted that N occupies the 10th position to the left of G, and the letter P occupies the 5th position to the right of G. One may therefore place these letters, N and P, in their proper positions in Alphabet 1, the letter N being placed 10 letters before G, and the letter P, 5 letters after G. Thus:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1					G					P					Y					V	N					

Thus, the values of two new letters in Alphabet 1, viz, $P_c=J_p$, and $N_c=U_p$ have been automatically determined; these values were obtained without any analysis based upon the *frequency* of P_c and N_c. Likewise, in Alphabet 2, the letters Y and V may be inserted in these positions:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2				V	N										G					P					Y	

This gives the new values $V_c=D_p$ and $Y_c=Y_p$ in Alphabet 2. Alphabets 3 and 4 have a common letter I, which permits of the placement of Q and W in Alphabet 3, and of B and L in Alphabet 4.

e. The new values thus found are of course immediately inserted throughout the cryptogram, thus leading to the assumption of further values in the cipher text. This process, *viz, the reconstruction of the primary components,* by the application of the principles of direct symmetry of position to the cells of the reconstruction skeleton, thus facilitates and hastens solution.

f. It must be clearly understood that before the principles of direct symmetry of position can be applied in cases such as the foregoing, *it is necessary that the plain component be a known sequence.* Whether it is the normal sequence or not is immaterial, so long as the sequence is known. Obviously, if the sequence is unknown, symmetry, even if present, cannot be detected by the cryptanalyst because he has no *base* upon which to try out his assumptions for symmetry. In other words, direct symmetry of position is manifested in the illustrative example because the plain component is a known sequence, and not because it is the normal alphabet. The significance of this point will become apparent later on in connection with the problem discussed in Par. 26*b*.

19. Initial steps in the solution of a typical example.—*a.* In the light of the foregoing principles let a typical message now be studied.

Message

	1	2	3	4	5
A.	Q W B R I	V W Y C A	I S P J L	R B Z E Y	Q W Y E U
B.	L W M G W	I C J C I	M T Z E I	M I B K N	Q W B R I
C.	V W Y I G	B W N B Q	Q C G Q H	I W J K A	G E G X N
D.	I D M R U	V E Z Y G	Q I G V N	C T G Y O	B P D B L
E.	V C G X G	B K Z Z G	I V X C U	N T Z A O	B W F E Q
F.	Q L F C O	M T Y Z T	C C B Y Q	O P D K A	G D G I G
G.	V P W M R	Q I I E W	I C G X G	B L G Q Q	V B G R S
H.	M Y J J Y	Q V F W Y	R W N F L	G X N F W	M C J K X
J.	I D D R U	O P J Q Q	Z R H C N	V W D Y Q	R D G D G
K.	B X D B N	P X F P U	Y X N F G	M P J E L	S A N C D
L.	S E Z Z G	I B E Y U	K D H C A	M B J J F	K I L C J
M.	M F D Z T	C T J R D	M I Y Z Q	A C J R R	S B G Z N
N.	Q Y A H Q	V E D C Q	L X N C L	L V V C S	Q W B I I
P.	I V J R N	W N B R I	V P J E L	T A G D N	I R G Q P
Q.	A T Y E W	C B Y Z T	E V G Q U	V P Y H L	L R Z N Q
R.	X I N B A	I K W J Q	R D Z Y F	K W F Z L	G W F J Q
S.	Q W J Y Q	I B W R X			

Over

b. The principal repetitions of three or more letters have been underlined in the message and the factors (up to 20 only) of the intervals between them are as follows:

QWBRIVWY........ 45=3, 5, 9, 15.
CGXGB........ 60=2, 3, 4, 5, 6, 10, 12, 15, 20.
PJEL........ 95=5, 19.
ZZGI........ 145=5.
BRIV........ 285=3, 5, 15, 19.
BRI........ 45=3, 5, 9, 15.
KAG........ 75=3, 5, 15.
QRD........ 165=3, 5, 15.
QWB........ 45=3, 5, 9, 15.
QWB........ 275=5, 11.
WIC........ 130=2, 5, 10, 13.
XNF........ 45=3, 5, 9, 15.
YZT........ 225=3, 5, 15.
ZTC........ 145=3, 5.

The factor 5 is common to all of these repetitions, and there seems to be every indication that five alphabets are involved. Since the message already appears in groups of five letters, it is unnecessary in this case to rewrite it in groups corresponding to the length of the key. The uniliteral frequency distribution for Alphabet 1 is as follows:

FIGURE 8.

c. Attempts to fit this distribution to the normal on the basis of a direct or reversed standard alphabet do not give positive results, and it is assumed that mixed alphabets are involved. Individual triliteral frequency distributions are then compiled and are shown in Fig. 9. These tables are similar to those made for single mixed alphabet ciphers, and are made in the same way except that instead of taking the letters one after the other, the letters which belong to the separate alphabets now must be assembled in separate tables. For example, in Alphabet 1, the trigraph QAC means that A occurs in Alphabet 1; Q, its prefix, occurs in Alphabet 5, and C, its suffix, occurs in Alphabet 2. All confusion may be avoided by placing numbers indicating the alphabets in which they belong above the letters, thus: $\overset{5}{\text{Q}}\overset{1}{\text{A}}\overset{2}{\text{C}}$

ALPHABET 1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
QC	GW	NF		TV		AE		AS		UD	UW	IT	UT	QP	NX	-W	LB	LA	LA		IW	NN	QI	UX	QR
PT	OP	TG				AD		WC		FI	QX	II		UP		YW	YW	DE			IW				
	GK	TT				LX		HW		FW	LV	OT				NW	QD	RB			UE				
	OW	WB				LW		ND			LR	SY				QC	QD				LC				
	GL							GV				WC				GI					GP				
	GX							WC				GP				QL					QB				
								XD				AB				RI					NW				
								GB				JF				YV					QE				
								IV				DI				NY					IP				
								NR								SW					UP				
								AK								QW									
								QB																	

Alphabet 2

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
SN	RZ	IJ	IM	GG	MD			MB		IW	QF		WB		BD		ZH	IP	MZ		IX	QB	GN	MJ	
TG	VG	QG	GG	VZ				QG		BZ	BG				OD		IG		CG		QF	VY	BD	QA	
	IE	VG	ID	SZ				QI							VW		LZ		NZ		LV	QY	PF		
	MJ	CB	RG	VD				KL							OJ				MY		IJ	LM	YN		
	SG	IG	KH					MY							MJ				CJ		EG	QB	LN		
	CY	MJ	RZ					XN							VJ				AY			VY			
	IW	AJ													VY							BN			
																						IJ			
																						BF			
																						RN			
																						VD			
																						QB			
																						KF			
																						GF			
																						QJ			

Alphabet 3

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
YH	WR		PB	BY	WE	CQ	RC	IE	CC		IC	WG	WB		SJ						VC	PM	VC	WC	BE
	IK		PK		LC	EX	DC		WK			DR	WF									KJ		WE	TE
	WR		DR		VW	IV			YJ				XF									BR		WI	EY
	CY		WY		XP	TY			CK				XF											TZ	KZ
	WI		XB		WZ	CX			PQ				AC											IZ	TA
	NR		FZ		WJ	DI			PE				XC											TE	EZ
			EC			CX			BJ				IB											BZ	RN
						LQ			TR															PH	DY
						BR			CR																
						DD			VR																
						BZ			PE																
						AD			WY																
						RQ																			
						VQ																			

Alphabet 4

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
ZO	NQ	YA	GG	ZY	NL	MW	AQ	YG	PL	BN		WR	ZQ		FU	GH	BI				GN	FY	GN	ZG	ZG
	DL	JI	GN	YU	NW		YL	GG	JY	JA						GQ	BI						GG	GO	YT
	DN	XU		ZI	NG			BI	JF	DA						JQ	MU						GG	BQ	ZG
	NA	FO		FQ					WQ	JX						GP	GS							DQ	DT
		HN		IW					FQ							GU	DU							EU	YQ
		ND		JL													JD							ZF	GN
		HA		JL													JR							JQ	YT
		LJ		YW													JN								FL
		DQ															BI								
		NL															WX								
		VS																							

~~288074 O—41——3.~~

strip fig. 3

Over

ALPHABET 5

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
CI			CS		JK	IB	QI	RV	CM		JR		KQ	YB	QA	BQ	MQ	RM	ZC	EL		GI	KI	EQ	
KG			RM		YK	YQ		CM			BV		XI	AB		EQ	RS	CQ	ZC	RV		EI	R–	JQ	
KG						XB		EM			FG		VC	CM		YO			ZE	CN		FM		WR	
CM						ZI		RV			ES		CV			QV				RO		EC			
BI						IV		II			CL		BP			QZ				PY					
						XB		RV			ET		ZQ			YR				YK					
						DB					HL		RW			ZA				QV					
						FM					ZG		DI			HV									
						ZI										CL									
																NX									
																JR									
																JQ									
																YI									

Condensed table of repetitions

1-2-3-4-5-1-2-3
Q W B R I V W Y-2

2-3-4-5-1
C G X G B-2

2-3-4-1
P J E L-2

3-4-5-1
B-R-I-V
Z-Z-G-I-2

1-2-3
Q W B-3
V W Y-2

2-3-4
C G X-2
P J E-2
W B R-2
X N F-2

3-4-5
B R I-3
G X G-2
J E L-2
Y Z T-2
Z Z G-2

4-5-1
K A G-2
X G B-2
Z G I-2
Z T C-2
R I V-3

5-1-2
I V W-2
Q R D-2
W I C-2

1-2
Q W-5
V P-3
V W-3

2-3
C G-3
C J-3
P J-3
W B-3
W F-3
W Y-3
X N-3

3-4
B R-3
G Q-4
G X-3
J R-3
N F-3
Y Z-3

4-5
R I-3
Y Q-3
Z T-3

5-1
G B-4
I V-3
Q Q-3

FIGURE 9.

d. One now proceeds to analyze each alphabet distribution, in an endeavor to establish identifications of cipher equivalents. First, of course, attempts should be made to separate the vowels from the consonants in each alphabet, using the same test as in the case of a single mixed-alphabet cipher. There seems to be no doubt about the equivalent of E_p in each alphabet:

$$E = \overset{1}{I}_c, \overset{2}{W}_c, \overset{3}{G}_c, \overset{4}{C}_c, \overset{5}{Q}_c$$

e. The letters of greatest frequency in Alphabet 1 are I, M, Q, V, B, G, L, R, S, and C. I_c has already been assumed to be E_p. If $\overset{2}{W}_c$ and $\overset{5}{Q}_c = E_p$, then one should be able to distinguish the vowels from the consonants among the letters M, Q, V, B, G, L, R, S, and C by examining the prefixes of $\overset{2}{W}_c$, and the suffixes of $\overset{5}{Q}_c$. The prefixes and suffixes of these letters, as shown by the triliteral frequency distributions, are these:

Prefixes of $\overset{2}{W}_c$ $(=\overset{2}{E}_p)$	Suffixes of $\overset{5}{Q}_c$ $(=\overset{5}{E}_p)$
Q G K V R B I L	I Q R X L V A Z O

f. Consider now the letter $\overset{1}{M}_c$; it does not occur either as a prefix of $\overset{2}{W}_c$, or as a suffix of $\overset{5}{Q}_c$. Hence it is most probably a vowel, and on account of its high frequency it may be assumed to be O_p. On the other hand, note that $\overset{1}{Q}_c$ occurs five times as a prefix of $\overset{2}{W}_c$ and three times as a suffix of $\overset{5}{Q}_c$. It is therefore a consonant, most probably R_p, for it would give the digraph ER $(=\overset{5\,1}{QQ}_c)$ as occurring three times and RE $(=\overset{1\,2}{QW}_c)$ as occurring five times.

g. The letter $\overset{1}{V}_c$ occurs three times as a prefix of $\overset{2}{W}_c$ and twice as a suffix of $\overset{5}{Q}_c$. It is therefore a consonant, and on account of its frequency, let it be assumed to be T_p. The letter $\overset{1}{B}_c$ occurs twice as a prefix of $\overset{2}{W}_c$ but not as a suffix of $\overset{5}{Q}_c$. Its frequency is only medium, and it is probably a consonant. In fact, the twice repeated digraph $\overset{1\,2}{BW}_c$ is once a part of the trigraph $\overset{5\,1\,2}{GBW}$, and $\overset{5}{G}_c$, the letter of second highest frequency in Alphabet 5, looks excellent for T_p. Might not the trigraph $\overset{5\,1\,2}{GBW}$ be THE? It will be well to keep this possibility in mind.

h. The letter $\overset{1}{G}_c$ occurs only once as a prefix of $\overset{2}{W}_c$ and does not occur as a suffix of $\overset{5}{Q}_c$. It may be a vowel, but one can not be sure. The letter $\overset{1}{L}_c$ occurs once as a prefix of $\overset{2}{W}_c$ and once as a suffix of $\overset{5}{Q}_c$. It may be considered to be a consonant. $\overset{1}{R}_c$ occurs once as a prefix of $\overset{2}{W}_c$, and twice as a suffix of $\overset{5}{Q}_c$, and is certainly a consonant. Neither the letter $\overset{1}{S}_c$ nor the letter $\overset{1}{C}_c$ occurs as a prefix of $\overset{2}{W}_c$ or as a suffix of $\overset{5}{Q}_c$; both would seem to be vowels, but a study of the prefixes and suffixes of these letters lends more weight to the assumption that $\overset{1}{C}_c$ is a vowel than that $\overset{1}{S}_c$ is a vowel. For all the prefixes of C, viz, $\overset{5}{N}$, $\overset{5}{T}$, and $\overset{5}{W}$, are in subsequent analysis of Alphabet 5 classified as consonants, as are likewise its suffixes, viz, T, C, and B in Alphabet 2. On the other hand, only one prefix, $\overset{5}{L}_c$, and one suffix, $\overset{2}{B}_c$, of $\overset{1}{S}_c$ are later classified as consonants. Since vowels are

Over

more often associated with consonants than with other vowels, it would seem that $\overset{1}{C}_\theta$ is more likely to be a vowel than $\overset{1}{S}_\theta$. At any rate $\overset{1}{C}_\theta$ is assumed to be a vowel, for the present, leaving $\overset{1}{S}_\theta$ unclassified.

i. Going through the same steps with the remaining alphabets, the following results are obtained:

Alphabet	Consonants	Vowels
1	Q, V, B, L, R, G?	I, M, C.
2	B, C, D, T.	W, P, I.
3	J, N, D, Y, F.	G, Z.
4	Y, Z, J. Q.	C, E?, R?, B?
5	G, N, A, I, W, L, T.	Q, U.

20. Application of principles of direct symmetry of position.—*a*. The next step is to try to determine a few values in each alphabet. In Alphabet 1, from the foregoing analysis, the following data are on hand:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	C?				I				C?						M			Q		V						

Let the values of E_p already assumed in the remaining alphabets, be set down in a reconstruction skeleton, as follows:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher 1	C?				I				C?						M			Q		V						
Cipher 2					W																					
Cipher 3					G																					
Cipher 4					C																					
Cipher 5					Q																					

FIGURE 10.

b. It is seen that by good fortune the letter Q is common to Alphabets 1 and 5, and the letter C is common to Alphabets 1 and 4. If it is assumed that one is dealing with a case in which a mixed component is sliding against the normal component, one can apply the principles of direct symmetry of position to these alphabets, as outlined in Par. 18. For example, one may insert the following values in Alphabet 5:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher 1	C?				I				C?						M			Q		V						
Cipher 5		M			Q		V							C?				I				C?				

FIGURE 11.

c. The process at once gives three definite values: $\overset{5}{M}_c=B_p$, $\overset{5}{V}_c=G_p$, $\overset{5}{I}_c=R_p$. Let these deduced values be substantiated by referring to the frequency distribution. Since B and G are normally low or medium frequency letters in plain text, one should find that M_c and V_c, their hypothetical equivalents in Alphabet 5, should have low frequencies. As a matter of fact, they do not appear in this alphabet, which thus far corroborates the assumption. On the other hand, since $\overset{5}{I}_c=R_p$, if the values derived from symmetry of position are correct, $\overset{5}{I}_c$ should be of high frequency, and reference to the distribution shows that I_c is of high frequency. The position of C is doubtful; it belongs either under N_p or V_p. If the former is correct, then the frequency of $\overset{5}{C}_c$ should be high, for it would equal N_p; if the latter is correct, then its frequency should be low, for it would equal V_c. As a matter of fact, $\overset{5}{C}_c$ does not occur, and it must be concluded that it belongs under V_p. This in turn settles the value of $\overset{1}{C}_c$, for it must now be placed definitely under I_p and removed from beneath A_p.

d. The definite placement of C now permits the insertion of new values in Alphabet 4, and one now has the following:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain																											
Cipher	1					I				C						M			Q		V						
	2					W																					
	3					G																					
	4	I				C						M			Q		V										
	5		M			Q		V											I				C				

FIGURE 12.

21. Subsequent steps in solution.—*a.* It is high time that the thus far deduced values, as recorded in the reconstruction skeleton, be inserted in the cipher text, for by this time it must seem that the analysis has certainly gone too far upon unproved hypotheses. The following results are obtained:

MESSAGE

```
          1          2          3          4          5
A.  Q W B R I  V W Y C A  I S P J L  R B Z E Y  Q W Y E U
    R E     R  T E     E  E                     R E

B.  L W M G W  I C J C I  M T Z E I  M I B K N  Q W B R I
      E        E     E R  O       R  O          R E     R

C.  V W Y I G  B W N B Q  Q C G Q H  I W J K A  G E G X N
    T E   A      E     E  R   E N    E E          E
```

```
D.  I D M R U   V E Z Y G   Q I G V N   C T G Y O   B P D B L
    E           T           R   E P     I   E

E.  V C G X G   B K Z Z G   I V X C U   N T Z A O   B W F E Q
    T   E                   E     E                   E     E

F.  Q L F C O   M T Y Z T   C C B Y Q   O P D K A   G D G I G
    R     E     O           I       E                   E A

G.  V P W M R   Q I I E W   I C G X G   B L G Q Q   V B G R S
    T     K     R           E   E           E N E   T   E

H.  M Y J J Y   Q V F W Y   R W N F L   G X N F W   M C J K X
    O           R             E                     O

J.  I D D R U   O P J Q Q   Z R H C N   V W D Y Q   R D G D G
    E                 N E         E     T E     E       E

K.  B X D B N   P X F P U   Y X N F G   M P J E L   S A N C D
                                        O                 E

L.  S E Z Z G   I B E Y U   K D H C A   M B J J F   K I L C J
                E                 E     O                 E

M.  M F D Z T   C T J R D   M I Y Z Q   A C J R R   S B G Z N
    O           I           O       E                   E

N.  Q Y A H Q   V E D C Q   L X N C L   L V V C S   Q W B I I
    R       E   T     E E         E           E     R E   A R

P.  I V J R N   W N B R I   V P J E L   T A G D N   I R G Q P
    E                   R   T               E       E   E N

Q.  A T Y E W   C B Y Z T   E V G Q U   V P Y H L   L R Z N Q
                I               E N     T                   E

R.  X I N B A   I K W J Q   R D Z Y F   K W F Z L   G W F J Q
                E       E                 E           E     E

S.  Q W J Y Q   I B W R X
    R E     E   E
```

b. The combinations given are excellent throughout and no inconsistencies appear. Note the trigraph $\overset{1\,2\,3}{\text{QWB}}$, which is repeated in the following polygraphs (underlined in the foregoing text):

1	2	3	4	5	1				5	1	2	3	4	5	1
Q	W	B	R	I	V	.	.	.	S	Q	W	B	I	I	I
R	E			R	T	.	.	.		R	E		A	R	E

c. The letter $\overset{3}{B}_c$ is common to both polygraphs, and a little imagination will lead to the assumption of the value $\overset{3}{B}_c=P_p$, yielding the following:

1	2	3	4	5	1				5	1	2	3	4	5	1
Q	W	B	R	I	V	.	.	.	S	Q	W	B	I	I	I
R	E	P	O	R	T	.	.	.	P	R	E	P	A	R	E

d. Note also (in F5) the polygraph $\underset{A}{\overset{4}{I}}\ \overset{5}{G}\ \underset{T}{\overset{1}{V}}\ \overset{2}{P}\ \overset{3}{W}\ \underset{K}{\overset{4}{M}}$, which looks like the word ATTACK. The frequency distributions are consulted to see whether the frequencies given for $\overset{5}{G}_c$ and $\overset{2}{P}_c$ are high enough for T_p and A_p, respectively, and also whether the frequency of $\overset{3}{W}_c$ is good enough for C_p; it is noted that they are excellent. Moreover, the digraph $\overset{5\,1}{GB}_c$, which occurs four times, looks like TH, thus making $\overset{1}{B}_c=H_p$. Does the insertion of these four new values in our diagram of alphabets bring forth any inconsistencies? The insertion of the value $\overset{2}{P}_c=A_p$ and $\overset{1}{B}_c=H_p$ gives no indications either way, since neither letter has yet been located in any of the other alphabets. The insertion of the value $\overset{5}{G}_c=T_p$ gives a value common to Alphabets 3 and 5, for the value $\overset{3}{G}_c=E_p$ was assumed long ago. Unfortunately an inconsistency is found here. The letter I has been placed two letters to the left of G in the mixed component, and has given good results in Alphabets 1 and 5; if the value $\overset{3}{W}_c=C_p$ (obtained above from the assumption of the word ATTACK) is correct, then W, and not I, should be the second letter to the left of G. Which shall be retained? There has been so far nothing to establish the value of $\overset{3}{G}_c=E_p$; this value was assumed from frequency considerations solely. Perhaps it is wrong. It certainly behaves like a vowel, and one may see what happens when one changes its value to O_p. The following placements in the reconstruction skeleton result from the analysis, when only two or three new values have been added as a result of the clues afforded by the deductions:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	1			S		I		G	B	C						M		P	Q	R	V	W					
	2	P	Q	R	V	W								S		I		G	B	C						M	
	3	R	V	W								S		I		G	B	C						M		P	Q
	4	I		G	B	C						M		P	Q	R	V	W								S	
	5		M		P	Q	R	V	W								S		I		G	B	C				

FIGURE 13c.

Over

e. Many new values are produced, and these are inserted throughout the message, yielding the following:

```
        1           2           3           4           5
A.  Q W B R I   V W Y C A   I S P J L   R B Z E Y   Q W Y E U
    R E P O R   T E   E     E M Y       S R         R E

B.  L W M G W   I C J C I   M T Z E I   M I B K N   Q W B R I
      E W C H   E S   E R   O       R   O O P       R E P O R

C.  V W Y I G   B W N B Q   Q C G Q H   I W J K A   G E G X N
    T E   A T   H E   D E   R S O N     E E         G   O

D.  I D M R U   V E Z Y G   Q I G V N   C T G Y O   B P D B L
    E   W O     T       T   R O O P     I   O       H A   D

E.  V C G X G   B K Z Z G   I V X C U   N T Z A O   B W F E Q
    T S O   T   H       T   E D   E                 H E     E

F.  Q L F C O   M T Y Z T   C C B Y Q   O P D K A   G D G I G
    R     E     O           I S P   E     A         G   O A T

G.  V P W M R   Q I I E W   I C G X G   B L G Q Q   V B G R S
    T A C K F   R O M   H   E S O   T   H   O N E   T R O O P

H.  M Y J J Y   Q V F W Y   R W N F L   G X N F W   M C J K X
    O           R D   Q     S E         G       H   O S

J.  I D D R U   O P J Q Q   Z R H C N   V W D Y Q   R D G D G
    E     O       A   N E     C   E     T E     E   S   O   T

K.  B X D B N   P X F P U   Y X N F G   M P J E L   S A N C D
    H     D     Q     M             T   O A         C     E

L.  S E Z Z G   I B E Y U   K D H C A   M B J J F   K I L C J
    C       T   E R               E     O R           O   E

M.  M F D Z T   C T J R D   M I Y Z Q   A C J R R   S B G Z N
    O           I     O     O O     E     S   O F   C R O

N.  Q Y A H Q   V E D C Q   L X N C L   L V V C S   Q W B I I
    R       E   T     E E         E       D B E P   R E P A R

P.  I V J R N   W N B R I   V P J E L   T A G D N   I R G Q P
    E D   O     U   P O R   T A             O       E C O N D

Q.  A T Y E W   C B Y Z T   E V G Q U   V P Y H L   L R Z N Q
            H   I R           D O N     T A           C     E

R.  X I N B A   I K W J Q   R D Z Y F   K W F Z L   G W F J Q
      O   D     E       E   S             E         G E     E

S.  Q W J Y Q   I B W R X
    R E     E   E R   O
```

22. Completing the solution.—*a.* Completion of solution is now a very easy matter. The mixed component is finally found to be the following sequence, based upon the word EXHAUSTING:

E X H A U S T I N G B C D F J K L M O P Q R V W Y Z

and the completely reconstructed skeleton of the cipher square is shown in Fig. 13*b*.

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	1	A	U	S	T	I	N	G	B	C	D	F	J	K	L	M	O	P	Q	R	V	W	Y	Z	E	X	H
	2	P	Q	R	V	W	Y	Z	E	X	H	A	U	S	T	I	N	G	B	C	D	F	J	K	L	M	O
	3	R	V	W	Y	Z	E	X	H	A	U	S	T	I	N	G	B	C	D	F	J	K	L	M	O	P	Q
	4	I	N	G	B	C	D	F	J	K	L	M	O	P	Q	R	V	W	Y	Z	E	X	H	A	U	S	T
	5	L	M	O	P	Q	R	V	W	Y	Z	E	X	H	A	U	S	T	I	N	G	B	C	D	F	J	K

FIGURE 13*b*.

b. Note that the successive equivalents of A_p spell the word APRIL, which is the key for the message. The plain-text message is as follows:

REPORTED ENEMY HAS RETIRED TO NEWCHESTER. ONE TROOP IS REPORTED AT HENDERSON MEETING HOUSE: TWO OTHER TROOPS IN ORCHARD AT SOUTHWEST EDGE OF NEWCHESTER. 2D SQ IS PREPARING TO ATTACK FROM THE SOUTH. ONE TROOP OF 3D SQ IS ENGAGING HOSTILE TROOP AT NEWCHESTER. REST OF 3D SQ IS MOVING TO ATTACK NEWCHESTER FROM THE NORTH. MOVE YOUR SQ INTO WOODS EAST OF CROSSROAD 539 AND BE PREPARED TO SUPPORT ATTACK OF 2D AND 3D SQ. DO NOT ADVANCE BEYOND NEWCHESTER. MESSAGES HERE.

TREER,
COL.

c. The preceding case is a good example of the value of the principles of direct symmetry of position when applied properly to a cryptogram enciphered by the sliding of a mixed component against the normal. The cryptanalyst starts off with only a very limited number of assumptions and builds up many new values as a result of the placement of the few original values in the reconstruction skeleton.

23. Solution of subsequent messages enciphered by the same cipher component.—*a. Preliminary remarks.*—Let it be supposed that the correspondents are using the same basic or primary component but with different key words for other messages. Can the knowledge of the sequence of letters in the reconstructed primary component be used to solve the subsequent messages? It has been shown that in the case of a monoalphabetic cipher in which a mixed alphabet was used, the process of completing the plain component could be applied to solve subsequent messages in which the same cipher component was used, even though the cipher component was set at a different key letter. A modification of the procedure used in that case can be used in this case, where a plurality of cipher alphabets based upon a sliding primary component is used.

Over

b. The message.—Let it be supposed that the following message passing between the same two correspondents as in the preceding message has been intercepted:

MESSAGE

SFDZR	YRRKX	MIWLL	AQRLU	RQFRT	IJQKF	XUWBS	MDJZK
MICQC	UDPTV	TYRNH	TRORV	BQLTI	QBNPR	RTUHD	PTIVE
RMGQN	LRATQ	PLUKR	KGRZF	JCMGP	IHSMR	GQRFX	BCABA
OEMTL	PCXJM	RGQSZ	VB				

c. Factoring and conversion into plain component equivalents.—The presence of a repetition of a four-letter polygraph whose interval is 21 letters suggests a key word of seven letters. There are very few other repetitions, and this is to be expected in a short message with a key of such length.

1	2	3	4	5	6	7
S	F	D	Z	R	Y	R
R	K	X	M	I	W	L
L	A	Q	R	L	U	R
Q	F	R	T	I	J	Q
K	F	X	U	W	B	S
M	D	J	Z	K	M	I
C	Q	C	U	D	P	T
V	T	Y	R	N	H	T
R	O	R	V	B	Q	L
T	I	Q	B	N	P	R
R	T	U	H	D	P	T
I	V	E	R	M	G	Q
N	L	R	A	T	Q	P
L	U	K	R	K	G	R
Z	F	J	C	M	G	P
I	H	S	M	R	G	Q
R	F	X	B	C	A	B
A	O	E	M	T	L	P
C	X	J	M	R	G	Q
S	Z	V	B			

FIGURE 14.

1	2	3	4	5	6	7
F	N	M	Z	V	Y	V
V	P	B	R	H	X	Q
Q	D	U	V	Q	E	V
U	N	V	G	H	O	U
P	N	B	E	X	K	F
R	M	O	Z	P	R	H
L	U	L	E	M	T	G
W	G	Y	V	I	C	G
V	S	V	W	K	U	Q
G	H	U	K	I	T	V
V	G	E	C	M	T	G
H	W	A	V	R	J	U
I	Q	V	D	G	U	T
Q	E	P	V	P	J	V
Z	N	O	L	R	J	T
H	C	F	R	V	J	U
V	N	B	K	L	D	K
D	S	A	R	G	Q	T
L	B	O	R	V	J	U
F	Z	W	K			

FIGURE 15.

d. Transcription into periods.—Let the message be written in groups of seven letters, in columnar fashion, as shown in Fig. 14. The letters in each column belong to a single alphabet. Let the letters in each column be converted into their plain-component equivalents by setting the reconstructed cipher component against the normal alphabet at any arbitrarily selected point, for example, that shown below:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	E	X	H	A	U	S	T	I	N	G	B	C	D	F	J	K	L	M	O	P	Q	R	V	W	Y	Z

The columns of equivalents are now as shown in Fig. 15.

e. Examination and selection of generatrices.—It has been shown that in the case of a monoalphabetic cipher it was merely necessary to complete the normal alphabet sequence beneath the plain-component equivalents and the plain text all reappeared on one generatrix. It was also found that in the case of a multiple-alphabet cipher involving standard alphabets, the plain-text equivalents of each alphabet reappeared on the same generatrix, and it was necessary only to combine the proper generatrices in order to produce the plain text of the message. In the case at hand both processes are combined: the normal alphabet sequence is continued beneath the letters of each column and then the generatrices are combined to produce the plain text. The completely developed generatrix diagrams for the first two columns are as follows (Fig. 16):

	Column 1		Column 2
	FVQUPRLWVGVHIQZHVDLF		NPDNNMUGSHGWQENCNSBZ
1	GWRVQSMXWHWIJRAIWEMG	1	OQEOONVHTIHXRFODOTCA
2	HXSWRTNYXIXJKSBJXFNH	2	PRFPPOWIUJIYSGPEPUDB
3	IYTXSUOZYJYKLTCKYGOI	3	QSGQQPXJVKJZTHQFQVEC
4	JZUYTVPAZKZLMUDLZHPJ	4	RTHRRQYKWLKAUIRGRWFD
5	KAVZUWQBALAMNVEMAIQK	5	SUISSRZLXMLBVJSHSXGE
6	LBWAVXRCBMBNOWFNBJRL	6	TVJTTSAMYNMCWKTITYHF
7	MCXBWYSDCNCOPXGOCKSM	7	UWKUUTBNZONDXLUJUZIG
8	NDYCXZTEDODPQYHPDLTN	8	VXLVVUCOAPOEYMVKVAJH
9	OEZDYAUFEPEQRZIQEMUO	9	WYMWWVDPBQPFZNWLWBKI
10	PFAEZBVGFQFRSAJRFNVP	10	XZNXXWEQCRQGAOXMXCLJ
11	QGBFACWHGRGSTBKSGOWQ	11	YAOYYXFRDSRHBPYNYDMK
12	RHCGBDXIHSHTUCLTHPXR	12	ZBPZZYGSETSICQZOZENL
13	SIDHCEYJITIUVDMUIQYS	13	ACQAAZHTFUTJDRAPAFOM
14	TJEIDFZKJUJVWENVJRZT	14	BDRBBAIUGVUKESBQBGPN
15	UKFJEGALKVKWXFOWKSAU	15	CESCCBJVHWVLFTCRCHQO
16	VLGKFHBMLWLXYGPXLTBV	16	DFTDDCKWIXWMGUDSDIRP
17	WMHLGICNMXMYZHQYMUCW	17	EGUEEDLXJYXNHVETEJSQ
18	XNIMHJDONYNZAIRZNVDX	18	FHVFFEMYKZYOIWFUFKTR
19	YOJNIKEPOZOABJSAOWEY	19	GIWGGFNZLAZPJXGVGLUS
20	ZPKOJLFQPAPBCKTBPXFZ	20	HJXHHGOAMBAQKYHWHMVT
21	AQLPKMGRQBQCDLUCQYGA	21	IKYIIHPBNCBRLZIXINWU
22	BRMQLNHSRCRDEMVDRZHB	22	JLZJJIQCODCSMAJYJOXV
23	CSNRMOITSDSEFNWESAIC	23	KMAKKJRDPEDTNBKZKPYW
24	DTOSNPJUTETFGOXFTBJD	24	LNBLLKSEQFEUOCLALQZX
25	EUPTOQKVUFUGHPYGUCKE	25	MOCMMLTFRGFVPDMBMRAY

Figure 16.

1	2
C	O
S	Q
N	E
R	O
M	O
O	N
I	V
T	H
S	T
D	I
S	H
E	X
F	R
N	F
W	O
E	D
S	O
A	T
I	C
C	A

Figure 17a.

f. Combining the selected generatrices.—After some experimenting with these generatrices the 23d generatrix of Column 1 and the 1st of Column 2, which yield the digraphs shown in Fig. 17*a*, are combined. The generatrices of the subsequent columns are examined to select those which may be added to these already selected in order to build up the plain text. The results are shown in Fig. 17*b*. This process is a very valuable aid in the solution of messages after the primary component has been recovered as a result of the longer and more detailed analysis of the frequency distributions of the first message intercepted. Very often a short message can be solved in no other way than the one shown, if the primary component is completely known.

g. Recovery of the key.—It may be of interest to find the key word for the message. Assuming that enciphering method number 1 (see Par. 7*f*, page 6) were known to be employed, all that is necessary is to set the mixed component of the cipher alphabet underneath the plain component so as to produce the cipher letter indicated as the equivalent of any given plain-text letter in each of the alphabets. For example, in the first alphabet it is noted that $C_p = S_c$. Adjust the two components under each other so as to bring S of the cipher component beneath C of the plain component, thus:

1	2	3	4	5	6	7
C	O	F	I	R	S	T
S	Q	U	A	D	R	O
N	E	N	E	M	Y	T
R	O	O	P	D	I	S
M	O	U	N	T	E	D
O	N	H	I	L	L	F
I	V	E	N	I	N	E
T	H	R	E	E	W	E
S	T	O	F	G	O	O
D	I	N	T	E	N	T
S	H	X	L	I	N	E
E	X	T	E	N	D	S
F	R	O	M	C	O	R
N	F	I	E	L	D	T
W	O	H	U	N	D	R
E	D	Y	A	R	D	S
S	O	U	T	H	X	I
A	T	T	A	C	K	R
I	C	H	A	R	D	S
C	A	P	T			

Figure 17b.

Over

Plain.......... ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher........ EXHAUSTINGBCDFJKLMOPQRVWYZ

It is noted that $A_p = A_c$. Hence, the first letter of the key word to the message is A. The 2d, 3d, 4th, . . . 7th key letters are found in exactly the same manner, and the following is obtained:

When C O F I R S T equals
S F D Z R Y R then A_p successively equals
A Z I M U T H

24. Summation of relative frequencies as an aid to the selection of the correct generatrices.—*a*. In the foregoing example, under subparagraph *f*, there occurs this phrase: "After some experimenting with these generatrices . . ." By this was meant, of course, that the selection of the correct initial pair of generatrices of plain-text equivalents is in this process a matter of trial and error. The test of "correctness" is whether, when juxtaposed, the two generatrices so selected yield "good" digraphs, that is, high-frequency digraphs such as occur in normal plain text. In his early efforts the student may have some difficulty in selecting, merely by ocular examination, the most likely generatrices to try. There may be in each diagram several generatrices which contain good assortments of high-frequency letters, and the number of trials of combinations of generatrices may be quite large. Perhaps a simple mathematical method may be of assistance in the process.

b. Suppose, in Fig. 16, that each letter were accompanied by a number which corresponds to its relative frequency in normal English telegraphic text. Then, by adding the numbers along each *horizontal* line, the totals thus obtained will serve as relative numerical measures of the frequency values of the respective generatrices. Theoretically, the generatrix with the greatest value will be the correct generatrix because its total will represent the sum of the individual values of the actual plaintext letters. In actual practice, of course, the generatrix with the greatest value may not be the correct one, but the correct one will certainly be among the three or four generatrices with the largest values. Thus, the number of trials may be greatly reduced, in the attempt to put together the correct generatrices.

c. Using the preceding message as an example, note the respective generatrix values in Fig. 18. The frequency values of the respective letters shown in the figure are based upon the normal distribution for War Department telegraphic text (see Table 3, Appendix 1, Military Cryptanalysis, Part I).

Column 1

Generatrix																					Frequency value
0	F	V	Q	U	P	R	L	W	V	G	V	H	I	Q	Z	H	V	D	L	F	57
	3	2	0	3	3	8	4	2	2	2	2	3	7	0	0	3	2	4	4	3	
1	G	W	R	V	Q	S	M	X	W	H	W	I	J	R	A	I	W	E	M	G	77
	2	2	8	2	0	6	2	0	2	3	2	7	0	8	7	7	2	13	2	2	
2	H	X	S	W	R	T	N	Y	X	I	X	J	K	S	B	J	X	F	N	H	66
	3	0	6	2	8	9	8	2	0	7	0	0	0	6	1	0	0	3	8	3	
3	I	Y	T	X	S	U	O	Z	Y	J	Y	K	L	T	C	K	Y	G	O	I	74
	7	2	9	0	6	3	8	0	2	0	2	0	4	9	3	0	2	2	8	7	
4	J	Z	U	Y	T	V	P	A	Z	K	Z	L	M	U	D	L	Z	H	P	J	49
	0	0	3	2	9	2	3	7	0	0	0	4	2	3	4	4	0	3	3	0	
5	K	A	V	Z	U	W	Q	B	A	L	A	M	N	V	E	M	A	I	Q	K	74
	0	7	2	0	3	2	0	1	7	4	7	2	8	2	13	2	7	7	0	0	
6	L	B	W	A	V	X	R	C	B	M	B	N	O	W	F	N	B	J	R	L	73
	4	1	2	7	2	0	8	3	1	2	1	8	8	2	3	8	1	0	8	4	
7	M	C	X	B	W	Y	S	D	C	N	C	O	P	X	G	O	C	K	S	M	66
	2	3	0	1	2	2	6	4	3	8	3	8	3	0	2	8	3	0	6	2	
8	N	D	Y	C	X	Z	T	E	D	O	D	P	Q	Y	H	P	D	L	T	N	91
	8	4	2	3	0	0	9	13	4	8	4	3	0	2	3	3	4	4	9	8	
9	O	E	Z	D	Y	A	U	F	E	P	E	Q	R	Z	I	Q	E	M	U	O	110
	8	13	0	4	2	7	3	3	13	3	13	0	8	0	7	0	13	2	3	8	
10	P	F	A	E	Z	B	V	G	F	Q	F	R	S	A	J	R	F	N	V	P	82
	3	3	7	13	0	1	2	2	3	0	3	8	6	7	0	8	3	8	2	3	
11	Q	G	B	F	A	C	W	H	G	R	G	S	T	B	K	S	G	O	W	Q	67
	0	2	1	3	7	3	2	3	2	8	2	6	9	1	0	6	2	8	2	0	
12	R	H	C	G	B	D	X	I	H	S	H	T	U	C	L	T	H	P	X	R	82
	8	3	3	2	1	4	0	7	3	6	3	9	3	3	4	9	3	3	0	8	
13	S	I	D	H	C	E	Y	J	I	T	I	U	V	D	M	U	I	Q	Y	S	90
	6	7	4	3	3	13	2	0	7	9	7	3	2	4	2	3	7	0	2	6	
14	T	J	E	I	D	F	Z	K	J	U	J	V	W	E	N	V	J	R	Z	T	83
	9	0	13	7	4	3	0	0	0	3	0	2	2	13	8	2	0	8	0	9	
15	U	K	F	J	E	G	A	L	K	V	K	W	X	F	O	W	K	S	A	U	65
	3	0	3	0	13	2	7	4	0	2	0	2	0	3	8	2	0	6	7	3	
16	V	L	G	K	F	H	B	M	L	W	L	X	Y	G	P	X	L	T	B	V	50
	2	4	2	0	3	3	1	2	4	2	4	0	2	2	3	0	4	9	1	2	
17	W	M	H	L	G	I	C	N	M	X	M	Y	Z	H	Q	Y	M	U	C	W	52
	2	2	3	4	2	7	3	8	2	0	2	2	0	3	0	2	2	3	3	2	
18	X	N	I	M	H	J	D	O	N	Y	N	Z	A	I	R	Z	N	V	D	X	86
	0	8	7	2	3	0	4	8	8	2	8	0	7	7	8	0	8	2	4	0	
19	Y	O	J	N	I	K	E	P	O	Z	O	A	B	J	S	A	O	W	E	Y	103
	2	8	0	8	7	0	13	3	8	0	8	7	1	0	6	7	8	2	13	2	
20	Z	P	K	O	J	L	F	Q	P	A	P	B	C	K	T	B	P	X	F	Z	51
	0	3	0	8	0	4	3	0	3	7	3	1	3	0	9	1	3	0	3	0	
21	A	Q	L	P	K	M	G	R	Q	B	Q	C	D	L	U	C	Q	Y	G	A	55
	7	0	4	3	0	2	2	8	0	1	0	3	4	4	3	3	0	2	2	7	
22	B	R	M	Q	L	N	H	S	R	C	R	D	E	M	V	D	R	Z	H	B	88
	1	8	2	0	4	8	3	6	8	3	8	4	13	2	2	4	8	0	3	1	
23	C	S	N	R	M	O	I	T	S	D	S	E	F	N	W	E	S	A	I	C	129
	3	6	8	8	2	8	7	9	6	4	6	13	3	8	2	13	6	7	7	3	
24	D	T	O	S	N	P	J	U	T	E	T	F	G	O	X	F	T	B	J	D	102
	4	9	8	6	8	3	0	3	9	13	9	3	2	8	0	3	9	1	0	4	
25	E	U	P	T	O	Q	K	V	U	F	U	G	H	P	Y	G	U	C	K	E	78
	13	3	3	9	8	0	0	2	3	3	3	2	3	3	2	2	3	3	0	13	

Over

Column 2

Generatrix																					Frequency value
0	N	P	D	N	N	M	U	G	S	H	G	W	Q	E	N	C	N	S	B	Z	
	8	3	4	8	8	2	3	2	6	3	2	2	0	13	8	3	8	6	1	0	90
1	O	Q	E	O	O	N	V	H	T	I	H	X	R	F	O	D	O	T	C	A	
	8	0	13	8	8	8	2	3	9	7	3	0	8	3	8	4	8	9	3	7	119
2	P	R	F	P	P	O	W	I	U	J	I	Y	S	G	P	E	P	U	D	B	
	3	8	3	3	3	8	2	7	3	0	7	2	6	2	3	13	3	3	4	1	84
3	Q	S	G	Q	Q	P	X	J	V	K	J	Z	T	H	Q	F	Q	V	E	C	
	0	6	2	0	0	3	0	0	2	0	0	0	9	3	0	3	0	2	13	3	46
4	R	T	H	R	R	Q	Y	K	W	L	K	A	U	I	R	G	R	W	F	D	
	8	9	3	8	8	0	2	0	2	4	0	7	3	7	8	2	8	2	3	4	88
5	S	U	I	S	S	R	Z	L	X	M	L	B	V	J	S	H	S	X	G	E	
	6	3	7	6	6	8	0	4	0	2	4	1	2	0	6	3	6	0	2	13	79
6	T	V	J	T	T	S	A	M	Y	N	M	C	W	K	T	I	T	Y	H	F	
	9	2	0	9	9	6	7	2	2	8	2	3	2	0	9	7	9	2	3	3	94
7	U	W	K	U	U	T	B	N	Z	O	N	D	X	L	U	J	U	Z	I	G	
	3	2	0	3	3	9	1	8	0	8	8	4	0	4	3	0	3	0	7	2	68
8	V	X	L	V	V	U	C	O	A	P	O	E	Y	M	V	K	V	A	J	H	
	2	0	4	2	2	3	3	8	7	3	8	13	2	2	2	0	2	7	0	3	73
9	W	Y	M	W	W	V	D	P	B	Q	P	F	Z	N	W	L	W	B	K	I	
	2	2	2	2	2	2	4	3	1	0	3	3	0	8	2	4	2	1	0	7	50
10	X	Z	N	X	X	W	E	Q	C	R	Q	G	A	O	X	M	X	C	L	J	
	0	0	8	0	0	2	13	0	3	8	0	2	7	8	0	2	0	3	4	0	60
11	Y	A	O	Y	Y	X	F	R	D	S	R	H	B	P	Y	N	Y	D	M	K	
	2	7	8	2	2	0	3	8	4	6	8	3	1	3	2	8	2	4	2	0	75
12	Z	B	P	Z	Z	Y	G	S	E	T	S	I	C	Q	Z	O	Z	E	N	L	
	0	1	3	0	0	2	2	6	13	9	6	7	3	0	0	8	0	13	8	4	85
13	A	C	Q	A	A	Z	H	T	F	U	T	J	D	R	A	P	A	F	O	M	
	7	3	0	7	7	0	3	9	3	3	9	0	4	8	7	3	7	3	8	2	93
14	B	D	R	B	B	A	I	U	G	V	U	K	E	S	B	Q	B	G	P	N	
	1	4	8	1	1	7	7	3	2	2	3	0	13	6	1	0	1	2	3	8	73
15	C	E	S	C	C	B	J	V	H	W	V	L	F	T	C	R	C	H	Q	O	
	3	13	6	3	3	1	0	2	3	2	2	4	3	9	3	8	3	3	0	8	79
16	D	F	T	D	D	C	K	W	I	X	W	M	G	U	D	S	D	I	R	P	
	4	3	9	4	4	3	0	2	7	0	2	2	2	3	4	6	4	7	8	3	77
17	E	G	U	E	E	D	L	X	J	Y	X	N	H	V	E	T	E	J	S	Q	
	13	2	3	13	13	4	4	0	0	2	0	8	3	2	13	9	13	0	6	0	108
18	F	H	V	F	F	E	M	Y	K	Z	Y	O	I	W	F	U	F	K	T	R	
	3	3	2	3	3	13	2	2	0	0	2	8	7	2	3	3	3	0	9	8	76
19	G	I	W	G	G	F	N	Z	L	A	Z	P	J	X	G	V	G	L	U	S	
	2	7	2	2	2	3	8	0	4	7	0	3	0	0	2	2	2	4	3	6	59
20	H	J	X	H	H	G	O	A	M	B	A	Q	K	Y	H	W	H	M	V	T	
	3	0	0	3	3	2	8	7	2	1	7	0	0	2	3	2	3	2	2	9	59
21	I	K	Y	I	I	H	P	B	N	C	B	R	L	Z	I	X	I	N	W	U	
	7	0	2	7	7	3	3	1	8	3	1	8	4	0	7	0	7	8	2	3	81
22	J	L	Z	J	J	I	Q	C	O	D	C	S	M	A	J	Y	J	O	X	V	
	0	4	0	0	0	7	0	3	8	4	3	6	2	7	0	2	0	8	0	2	56
23	K	M	A	K	K	J	R	D	P	E	D	T	N	B	K	Z	K	P	Y	W	
	0	2	7	0	0	0	8	4	3	13	4	9	8	1	0	0	0	3	2	2	66
24	L	N	B	L	L	K	S	E	Q	F	E	U	O	C	L	A	L	Q	Z	X	
	4	8	1	4	4	0	6	13	0	3	13	3	8	3	4	7	4	0	0	0	85
25	M	O	C	M	M	L	T	F	R	G	F	V	P	D	M	B	M	R	A	Y	
	2	8	3	2	2	4	9	3	8	2	3	2	3	4	2	1	2	8	7	2	77

Figure 18.

d. It will be noted that the frequency value of the 23d generatrix for the first column of cipher letters is the greatest; that of the first generatrix for the second column is the greatest. In both cases these are the correct generatrices. Thus the selection of the correct generatrices in such cases has been reduced to a purely mathematical basis which is at times of much assistance in effecting a quick solution. Moreover, an understanding of the principles involved will be of considerable value in subsequent work.

25. Solution by the probable-word method.—*a*. Occasionally one may encounter a cryptogram which is so short that it contains no recurrences even of digraphs, and thus gives no indications of the number of alphabets involved. If the sliding mixed component is known, one may apply the method illustrated in Par. 15, assuming the presence of a probable word, checking it against the text and the sliding components to establish a key, if the correspondents are using key words.

b. For example, suppose that the presence of the word ENEMY is assumed in the message in Par. 23*b* above. One proceeds to check it against an unknown key word, sliding the already reconstructed mixed component against the normal and starting with the first letter of the cryptogram, in this manner:

When ENEMY equals
SFDZR then A_p successively equals
XENFW

The sequence XENFW spells no intelligible word. Therefore, the location of the assumed word ENEMY is shifted one letter forward in the cipher text, and the test is made again, just as was explained in Par. 15. When the group AQRLU is tried, the key letters ZIMUT are obtained, which, taken as a part of a word, suggests the word AZIMUTH. The method must yield solution when the correct assumptions are made.

c. The danger to cryptographic security resulting from the inclusion of *cryptographed* addresses and signatures in cryptographic messages becomes quite obvious in the light of solution by the probable-word method. To illustrate, reference is made to the message employed in Pars. 19–22. It will be noted in Par. 22*b* that the message carried a signature (Treer, Col.) and that the latter was enciphered. Suppose that this were an authorized practice, and that every message could be assumed to conclude with a cryptographed signature. The signature "TREER COL" would at once afford a very good basis for the quick solution of subsequent messages emanating from the same headquarters as did the first message, because presumably this same signature would appear in other messages. It is for this reason that addresses and signatures must not be cryptographed; if they must be included they should be cryptographed in a totally different system or by a wholly different method, perhaps by means of a special address and signature code. It would be best, however, to omit all addresses and signatures, and to let the call signs of the headquarters concerned also convey these parts of the message, leaving the delivery to the addressee a matter for local action.

26. Solution when the plain component is a mixed sequence, the cipher component, the normal.—*a*. This falls under Case B (2) outlined in Par. 6. It is not the usual method of employing a single mixed component, but may be encountered occasionally in cipher devices.

b. The preliminary steps, as regards factoring to determine the length of the period, are the same as usual. The message is then transcribed into its periods. Frequency distributions are then made, as usual, and these are attacked by the principles of frequency and recurrence. An attempt is made to apply the principles of direct symmetry of position, but this attempt will be futile, for the reason that the plain component is in this case an *unknown* mixed sequence.

Over

(See Par. 18*d*.) Any attempt to find symmetry in the secondary alphabets based upon the normal sequence can therefore disclose no symmetry because the symmetry which exists is based upon a wholly different sequence.

c. However, if the principles of direct symmetry of position are of no avail in this case, there are certain other principles of symmetry which may be employed to great advantage. To explain them an actual example will be used. Let it be assumed that it is known to the cryptanalyst that the enemy is using the general system under discussion, *viz*, a mixed sequence, variable from day to day, is used as plain component; the normal sequence is used as cipher component; and a repeating key, variable from message to message, is used in the ordinary manner.

The following message has been intercepted:

	1	2	3	4	5	6
A.	Q E O V K	L R M L Z	J V G T G	N D L V K	E V N T Y	E R M U E
B.	V R Z M O	Y A A M P	D K E I J	S F M Y O	Y H M M E	G Q A M B
C.	U Q A X R	H U F B U	K Q Y M U	N E L V T	K Q I L E	K Z B U E
D.	U L I B K	N D A X B	X U D G L	L A D V K	P O A Y O	D K K Y K
E.	L A D H Y	B V N F V	U E E M E	F F M T E	G V W B Y	T V D Z L
F.	S P B H B	X V A Z C	U D Y U E	L K M M A	E U D D K	N C F S H
G.	H S A H Y	T M G U J	H Q X P P	D K O U E	X U Q V B	F V W B X
H.	N X A L B	T C D L M	I V A A A	N S Z I L	O V W V P	Y A G Z L
J.	S H M M E	G Q D H O	Y H I V P	N C R R E	X K D Q Z	G K N C G
K.	N Q G U Y	J I W Y Y	T M A H W	X R L B L	O A D L G	N Q G U Y
L.	J U U G B	J H R V X	E R F L E	G W G U O	X E D T P	D K E I Z
M.	V X N W A	F A A N E	M K G H B	S S N L O	K J C B Z	T G G L O
N.	P K M B X	H G E R Y	T M W L Z	N Q C Y Y	T M W I P	D K A T E
P.	F L N U J	N D T V X	J R Z T L	O P A H C	D F Z Y Y	D E Y C L
Q.	G P G T Y	T E C X B	H Q E B R	K V W M U	N I N G J	I Q D L P
R.	J K A T E	G U W B R	H U Q W M	V R Q B W	Y R F B F	K M W M B
S.	T M U L Z	L A A H Y	J G D V K	L K R R E	X K N A O	N D S B X
T.	X C G Z A	H D G T L	V K M B W	I S A U E	F D N W P	N L Z I J
V.	S R Q Z L	A V N H L	G V W V K	F I G H P	G E C Z U	K Q A P

d. A study of the recurrences and factoring their intervals discloses that five alphabets are involved. Uniliteral frequency distributions are made and are shown in Fig. 19*a*:

ALPHABET 1

ALPHABET 2

ALPHABET 3

ALPHABET 4

ALPHABET 5

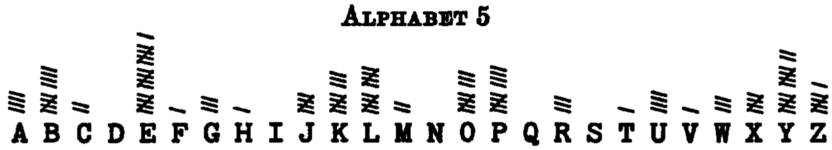
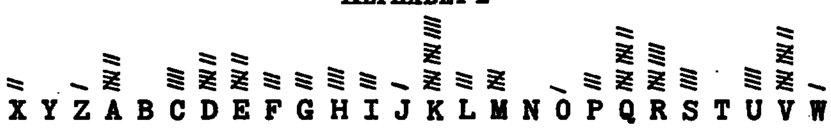


FIGURE 19a.

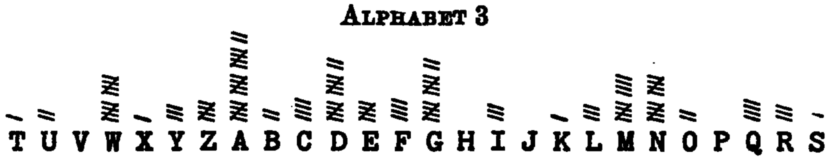
e. Since the cipher component in this case is the normal alphabet, *it follows that the five frequency distributions are based upon a sequence which is known, and therefore, the five frequency distributions should manifest a direct symmetry of distribution of crests and troughs*. By virtue of this symmetry and by shifting the five distributions relative to one another to proper superimpositions, the several distributions may be combined into a single uniliteral distribution. Note how this shifting has been done in the case of the five illustrative distributions:

ALPHABET 1

ALPHABET 2



ALPHABET 3



ALPHABET 4

ALPHABET 5

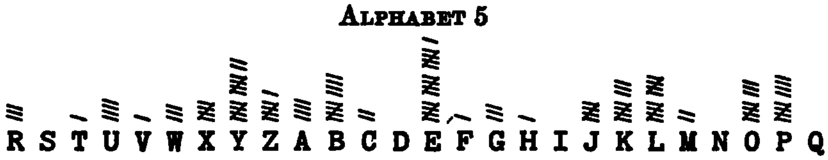


FIGURE 19b.

~~—41——4~~

skip sig. 4

Over

f. The superimposition of the respective distributions enables one to convert the cipher letters of the five alphabets into one alphabet. Suppose it is decided to convert Alphabets 2, 3, 4, and 5 into Alphabet 1. It is merely necessary to substitute for the respective letters in the four alphabets those which stand above them in Alphabet 1. For example, in Fig. 19*b*, X_c in Alphabet 2 is directly under A_c in Alphabet 1; hence, if the superimposition is correct then $\overset{2}{X_c}=\overset{1}{A_c}$. Therefore, in the cryptogram it is merely necessary to replace every X_c in the second position by A_c. Again T_c in Alphabet 3=A_c in Alphabet 1; therefore, in the cryptogram one replaces every T_c in the third position by A_c. The entire process, hereinafter designated as *conversion into monoalphabetic terms*, gives the following *converted* message:

	1	2	3	4	5	6
A.	Q H V H T	L U T X I	J Y N F P	N G S H T	E Y U F H	E U T G N
B.	V U G Y X	Y D H Y Y	D N L U S	S I T K X	Y K T Y N	G T H Y K
C.	U T H J A	H X M N D	K T F Y D	N H S H C	K T P X N	K C I G N
D.	U O P N T	N G H J K	X X K S U	L D K H T	P R H K X	D N R K T
E.	L D K T H	B Y U R E	U H L Y N	F I T F N	G Y D N H	T Y K L U
F.	S S I T K	X Y H L L	U G F G N	L N T Y J	E X K P T	N F M E Q
G.	H V H T H	T P N G S	H T E B Y	D N V G N	X X X H K	F Y D N G
H.	N A H X K	T F K X V	I Y H M J	N V G U U	O Y D H Y	Y D N L U
J.	S K T Y N	G T K T X	Y K P H Y	N F Y D N	X N K C I	G N U O P
K.	N T N G H	J L D K H	T P H T F	X U S N U	O D K X P	N T N G H
L.	J X B S K	J K Y H G	E U M X N	G Z N G X	X H K F Y	D N L U I
M.	V A U I J	F D H Z N	M N N T K	S V U X X	K M J N I	T J N X X
N.	P N T N G	H J L D H	T P D X I	N T J K H	T P D U Y	D N H F N
P.	F O U G S	N G A H G	J U G F U	O S H T L	D I G K H	D H F O U
Q.	G S N F H	T H J J K	H T L N A	K Y D Y D	N L U S S	I T K X Y
R.	J N H F N	G X D N A	H X X I V	V U X N F	Y U M N O	K P D Y K
S.	T P B X I	L D H T H	J J K H T	L N Y D N	X N U M X	N G Z N G
T.	X F N L J	H G N F U	V N T N F	I V H G N	F G U I Y	N O G U S
V.	S U X L U	A Y U T U	G Y D H T	F L N T Y	G H J L D	K T H B

The uniliteral frequency distribution for this converted text follows. Note that the frequency of each letter is the sum of the five frequencies in the corresponding columns of Fig. 19*b*.

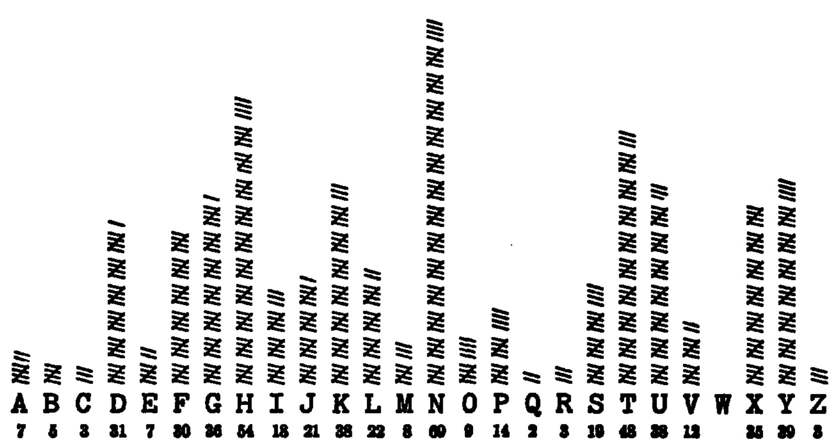


FIGURE 20.

g. The problem having been reduced to monoalphabetic terms, a triliteral frequency distribution can now be made and solution readily attained by simple principles. It yields the following:

JAPAN CONSULTED GERMANY TODAY ON REPORTS THAT THE COMMUNIST INTERNATIONAL WAS BEHIND THE AMAZING SEIZURE OF GENERALISSIMO CHIANG KAI SHEK IN CHINA. TOKYO ACTED UNDER THE ANTICOMMUNIST ACCORD RECENTLY SIGNED BY JAPAN AND GERMANY. THE PRESS SAID THERE WAS INDISPUTABLE PROOF THAT THE COMINTERN INSTIGATED THE SEIZURE OF GENERAL CHIANG AND SOME OF HIS GENERALS. MILITARY OBSERVERS SAID THE COUP WOULD HAVE BEEN IMPOSSIBLE UNLESS GENERAL CHANG HSUEN LIANG HOTHEADED FORMER WAR LORD OF MANCHURIA HAD FORMED AN ALLIANCE WITH THE COMMUNIST LEADERS HE WAS SUPPOSED TO BE FIGHTING. SUCH AN ALLIANCE THESE OBSERVERS DECLARED OPENED UP A RED ROUTE FROM MOSCOW TO NORTH AND CENTRAL CHINA.

h. The reconstruction of the plain component is now a very simple matter. It is found to be as follows:

H Y D R A U L I C B E F G J K M N O P Q S T V W X Z

Note also, in Fig. 19*b*, the keyword for the message, (HEAVY), the letters being in the columns headed by the letter H.

i. The solution of subsequent messages with different keys can now be reached directly, by a simple modification of the principles explained in Par. 18. This modification consists in using for the completion sequence the *mixed plain component* (now known) instead of the normal alphabet, after the cipher letters have been converted into their plain-component equivalents. Let the student confirm this by experiment.

j. The probable-word method of solution discussed under Paragraph 20 is also applicable here, in case of very short cryptograms. This method presupposes of course, possession of the mixed component and the procedure is essentially the same as that in Par. 20. In the example discussed in the present paragraph, the letter A on the plain component was successively set against the key letters HEAVY; but this is not the only possible procedure.

k. The student should go over carefully the principle of "conversion into monoalphabetic terms" explained in subparagraph *f* above until he thoroughly understands it. Later on he will encounter cases in which this principle is of very great assistance in the cryptanalysis of more complex problems. (Another example will be found under Par. 45.)

l. The principle illustrated in subparagraph *e*, that is, shifting two or more monoalphabetic frequency distributions relatively so as to bring them into proper alignment for amalgamation into a single monoalphabetic distribution, is called *matching*. It is a very important cryptanalytic principle. Note that its practical application consists in sliding one monoalphabetic distribution against the other so as to obtain the best coincidence between the *entire sequence* of crests and troughs of one distribution and the *entire sequence* of crests and troughs of the other distribution. When the best point of coincidence has been found, the two sequences may be amalgamated and *theoretically* the single resultant distribution will also be monoalphabetic in character. The successful application of the principle of matching depends upon several factors. First, the cryptographic situation must be such that matching is a correct cryptographic step. For example, the distributions in figure 19*b* are properly subject to matching because the cipher component in the basic sequences concerned in this problem is the normal sequence, while the plain component is a mixed sequence. But it would be futile to try to match the distributions in figure 9, for in that case the cipher component is a mixed sequence, the plain component is the normal sequence. Hence, no amount of shifting or matching can bring the distributions of

Over

figure 9 into proper superimposition for correct amalgamation. (If the occurrences in the various distributions in figure 9 had been distributed according to the sequence of letters in the mixed component, then matching would be possible; but in order to be able to distribute these occurrences according to the mixed component, the latter has to be *known*—and that is just what is unknown until the problem has been solved.) A second factor involved in successful matching is the number of elements in the two distributions forming the subject of the test. If both of them have very few tallies, there is hardly sufficient information to permit of matching with any degree of assurance that the work is not in vain. If one of them has many tallies, the other only a few, the chances for success are better than before, because the positions of the *blanks* in the two distributions can be used as a guide for their proper superimposition.

m. There are certain mathematical and statistical procedures which can be brought to bear upon the matter of cryptanalytic matching. These will be presented in a later text. However, until the student has studied these mathematical and statistical methods of matching distributions, he will have to rely upon mere ocular examination as a guide to proper superimposition. Obviously, the more data he has in each distribution, the easier is the correct superimposition ascertained by any method.

SECTION VI

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, II

27. Further cases to be considered.—*a.* Thus far Cases B (1) and (2), mentioned in Paragraph 6 have been treated. There remains Case B (3), and this case has been further subdivided as follows:

CASE B (3). Both components are mixed sequences.

(a) Components are identical mixed sequences.

(1) Sequences proceed in the same direction. (The secondary alphabet s are mixed alphabets.) set s/

(2) Sequences proceed in opposite directions. (The secondary alphabets are reciprocal mixed alphabets.)

(b) Components are different mixed sequences. (The secondary alphabets are mixed alphabets.)

b. The first of the foregoing subcases will now be examined.

28. Identical primary mixed components proceeding in the same direction.—*a.* It is often the case that the mixed components are derived from an easily remembered word or phrase, so that they can be reproduced at any time from memory. Thus, for example, given the key word QUESTIONABLY, the following mixed sequence is derived:

Q U E S T I O N A B L Y C D F G H J K M P R V W X Z

b. By using this sequence as both plain and cipher component, that is, by sliding this sequence against itself, a series of 26 secondary mixed alphabets may be produced. In enciphering a message, sliding strips may be employed with a key word to designate the particular and successive positions in which the strips are to be set, the same as was the case in previous examples of the use of sliding components. The method of designating the positions, however, requires a word or two of comment at this point. In the examples thus far shown, the key letter, as located on the cipher component, was always set opposite A, as located on the plain component; possibly an erroneous impression has been created, *viz,* that this is invariably the rule. This is decidedly not true, as has already been explained in paragraph 7*c*. If it has seemed to be the case that Θ_k always equals A_p, it is only because the text has dealt thus far principally with cases in which the plain component is the normal sequence and its ~~intital~~ letter, which usually constitutes the index for juxtaposing cipher components, is A. It must be emphasized, however, that various conventions may be adopted in this respect; but the most common of them is to employ the initial letter of the plain component as the index letter. That is, the index letter, Θ_1, will be the initial letter of the mixed sequence, in this case, Q. Furthermore, to prevent the possibility of ambiguity it will be stated again that the pair of enciphering equations employed in the ensuing discussion will be the first of the 12 set forth under Par. 7*f*, *viz,* $\Theta_k/_2=\Theta_1/_1$; $\Theta_p/_1=\Theta_c/_2$. In this case the subscript "1" means the plain component, the subscript "2", the cipher component, so that the enciphering equation is the following: $\Theta_k/_c=\Theta_1/_p$; $\Theta_p/_p=\Theta_c/_c$. set initial

Over

c. By setting the two sliding components against each other in the two positions shown below, the cipher alphabets labeled (1) and (2) given by two key letters, A and B, are seen to be different.

KEY LETTER=A

```
                                                 θ1
                                                 ↓
Plain component........  QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ
Cipher component.......                     QUESTIONABLYCDFGHJKMPRVWXZ
                                                 ↑
                                                 θk
```

Secondary alphabet (1):

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	H	J	P	R	L	V	W	X	D	Z	Q	K	U	G	F	E	A	S	Y	C	B	T	I	O	M	N

KEY LETTER=B

```
                                                 θ1
                                                 ↓
Plain component........  QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ
Cipher component.......                    QUESTIONABLYCDFGHJKMPRVWXZ
                                                 ↑
                                                 θk
```

Secondary alphabet (2):

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	J	K	R	V	Y	W	X	Z	F	Q	U	M	E	H	G	S	B	T	C	D	L	I	O	N	P	A

d. Very frequently a quadricular or square table is employed by the correspondents, instead of sliding strips, but the results are the same. The cipher square based upon the word QUESTIONABLY is shown in Fig. 21. It will be noted that it does nothing more than set forth the successive positions of the two primary sliding components; the top line of the square is the plain component, the successive horizontal lines below it, the cipher component in its various juxtapositions. The usual method of employing such a square (i. e., corresponding to the enciphering equations $\theta_{k/c}=\theta_{1/p}$; $\theta_{p/p}=\theta_{c/c}$) is to take as the cipher equivalent of a plain-text letter that letter which lies at the intersection of the vertical column headed by the plain-text letter and the horizontal row begun by the key letter. For example, the cipher equivalent of E_p with keyletter T is the letter O_c; or $E_p (T_k)=O_c$. The method given in paragraph *b*, for determining the cipher equivalents by means of the two sliding strips yields the same results as does the cipher square.

Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z
U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q
E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U
S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E
T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S
I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T
O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I
N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O
A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N
B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A
L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B
Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L
C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y
D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C
F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D
G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F
H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G
J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H
K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J
M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K
P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M
R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P
V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R
W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V
X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W
Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X

FIGURE 21.

29. Cryptographing and decryptographing by identical primary mixed components.—There is nothing of special interest to be noted in connection with the use either of identical mixed components or of an equivalent quadricular table such as that shown in Fig. 21, in enciphering or deciphering a message. The basic principles are the same as in the case of the sliding of one mixed component against the normal, the displacements of the two components being controlled by changeable key words of varying lengths. The components may be changed at will and so on. All this has been demonstrated adequately enough in *Elementary Military Cryptography*, and *Advanced Military Cryptography*.

30. Principles of solution.—*a.* Basically the principles of solution in the case of a cryptogram enciphered by two identical mixed sliding components are the same as in the preceding case. Primary recourse is had to the principles of frequency and repetition of single letters, digraphs, trigraphs, and polygraphs. Once an entering wedge has been forced into the problem, the subsequent steps may consist merely in continuing along the same lines as before, building up the solution bit by bit.

b. Doubtless the question has already arisen in the student's mind as to whether any principles of symmetery of position can be used to assist in the solution and in the reconstruction of the cipher alphabets in cases of the kind under consideration. This phase of the subject will be taken up in the next section and will be treated in a somewhat detailed manner, because the theory and principles involved are of very wide application in cryptanalytics.

Over

Section VII

THEORY OF INDIRECT SYMMETRY OF POSITION IN SECONDARY ALPHABETS[1] set

31. Reconstruction of primary components from secondary alphabets.—*a.* Note the two secondary alphabets (1) and (2) given in paragraph 28*c*. Externally they show no resemblance or symmetry despite the fact that they were produced from the same primary components. Nevertheless, when the matter is studied with care, a symmetry of position is discoverable. Because it is a hidden or latent phenomenon, it may be termed *latent symmetry of position.* However, in previous texts the phenomenon has been designated as an *indirect symmetry of position* and this terminology has grown into usage, so that a change is perhaps now inadvisable. Indirect symmetry of position is a very interesting and exceedingly useful phenomenon in cryptanalytics.

b. Consider the following secondary alphabet (the one labeled (2) in paragraph 28*c*):

(2)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher	J	K	R	V	Y	W	X	Z	F	Q	U	M	E	H	G	S	B	T	C	D	L	I	O	N	P	A

c. Assuming it to be known that this is a secondary alphabet produced by two primary identical mixed components, it is desired to reconstruct the latter. Construct a chain of alternating plain-text and cipher-text equivalents, beginning at any point and continuing until the chain has been completed. Thus, for example, beginning with $A_p=J_c$, $J_p=Q_c$, $Q_p=B_c$, . . ., and dropping out the letters common to successive pairs, there results the sequence A J Q B . . .. By completing the chain the following sequence of letters is established:

A J Q B K U L M E Y P S C R T D V I F W O G X N H Z

d. This sequence consists of 26 letters. *When slid against itself it will produce exactly the same secondary alphabets as do the primary components based upon the word* QUESTIONABLY. To demonstrate that this is the case, compare the secondary alphabets given by the two settings of the externally different components shown below:

Plain component...... QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ
Cipher component..... QUESTIONABLYCDFGHJKMPRVWXZ

Secondary alphabet (1):

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	J	K	R	V	Y	W	X	Z	F	Q	U	M	E	H	G	S	B	T	C	D	L	I	O	N	P	A

Plain component...... AJQBKULMEYPSCRTDVIFWOGXNHZAJQBKULMEYPSCRTDVIFWOGXNHZ
Cipher component..... AJQBKULMEYPSCRTDVIFWOGXNHZ

Secondary alphabet (2):

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	J	K	R	V	Y	W	X	Z	F	Q	U	M	E	H	G	S	B	T	C	D	L	I	O	N	P	A

Footnote [1] After the student has read this and the next section it would be well for him to study Appendix 3, where another and perhaps simpler method is explained.

36½ pi 8pt Single leaded F.i.c set

e. Since the sequence A J Q B K . . . gives exactly the same equivalents in the secondary alphabets as the sequence Q U E S T . . . gives, the former sequence is cryptographically equivalent to the latter sequence. For this reason the A J Q B K . . . sequence is termed an *equivalent primary component*.[1] If the real or original primary component is a key-word mixed sequence, it is hidden or *latent* within the equivalent primary sequence; but it can be made *patent* by decimation of the equivalent primary component. The procedure is as follows: Find three letters in the equivalent primary component such as are likely to have formed an unbroken sequence in the original primary component, and see if the interval between the first and second is the same as that between the second and third. Such a case is presented by the letters W, X, and Z in the equivalent primary component above. Note the sequence . . . W O G X N H Z . . . ; the distance or interval between the letters W, X, and Z is two letters. Continuing the chain by adding letters two intervals removed, the latent original primary component is made patent. Thus:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V

f. It is possible to perform the steps given in *c* and *e* in a combined single operation when the original primary component is a key-word mixed sequence. Starting with any pair of letters (in the cipher component of the secondary alphabet) likely to be sequent in the key-word mixed sequence, such as JK_c in the secondary alphabet labeled (2), the following chain of digraphs may be set up. Thus, J, K, in the plain component stand over Q, U, respectively, in the cipher component; Q, U, in the plain component stand over B, L, respectively, in the cipher component, and so on. Connecting the pairs in a series, the following results are obtained:

JK → QU → BL → KM → UE → LY → MP → ES → YC → PR → ST → CD → RV →
TI → DF → VW → IO → FG → WX → ON → GH → XZ → NA → HJ → ZQ → AB → JK . . .

These may now be united by means of their common letters:

JK → KM → MP → PR → RV → etc.=J K M P R V W X Z Q U E S T I O N A B L Y C D F G H

The original primary component is thus completely reconstructed.

g. Not all of the 26 secondary alphabets of the series yielded by two sliding primary components may be used to develop a complete equivalent primary component. If examination be made, it will be found that only 13 of these secondary alphabets will yield complete equivalent primary components when the method of reconstruction shown in subparagraph *c* above is followed. For example the following secondary alphabet, which is also derived, from the primary components based upon the word QUESTIONABLY will not yield a complete chain of 26 plain text-cipher-plain text equivalents:

Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	C	D	H	J	O	K	M	P	B	R	V	F	W	Y	L	X	T	Z	N	A	I	Q	U	E	G	S

[1] Such an equivalent component is merely a sequence which has been or can be developed or derived from the original sequence or basic primary component by applying a *decimation* process to the latter; conversely, the original or basic component can be derived from an equivalent component by applying the same sort of process to the equivalent component. By decimation is meant the selection of elements from a sequence according to some fixed interval. For example, the sequence A E I M . . . is derived, by decimation, from the normal alphabet by selecting every fourth letter.

Over

Equivalent primary component:

1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	
A	C	H	P	X	E	O	L	F	K	V	Q	T	A	C	H	. . . (The A C H sequence begins again.)

h. It is seen that only 13 letters of the chain have been established before the sequence begins to repeat itself. It is evident that exactly one-half of the chain has been established. The other half may be established by beginning with a letter not in the first half. Thus:

1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	
B	D	J	R	Z	S	N	Y	G	M	W	U	I	B	D	J	. . . (The B D J sequence begins again.)

i. It is now necessary to distribute the letters of each half-sequence within 26 spaces, to correspond with their placements in a complete alphabet. This can only be done by allowing a constant *odd* number of spaces between the letters of one of the half-sequences. Distributions are therefore made upon the basis of 3, 5, 7, 9, . . . spaces. Select that distribution which most nearly coincides with the distribution to be expected in a key-word component. Thus, for example, with the first half-sequence the distribution selected is the one made by leaving three spaces between the letters. It is as follows:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
A	–	L	–	C	–	F	–	H	–	K	–	P	–	V	–	X	–	Q	–	E	–	T	–	O	–

j. Now interpolate, by the same constant interval (three in this case), the letters of the other half-sequence. Noting that the group F – H appears in the foregoing distribution, it is apparent that G of the second half-sequence should be inserted between F and H. The letter which immediately follows G in the second half-sequence, *viz,* M, is next inserted in the position three spaces to the right of G, and so on, until the interpolation has been completed. This yields the original primary component, which is as follows:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
A	B	L	Y	C	D	F	G	H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N

k. Another method of handling cases such as the foregoing is indicated in subparagraph *f* By extending the principles set forth in that subparagraph, one may reconstruct the following chain of 13 pairs from the secondary alphabet given in subparagraph *g:*

1	2	3	4	5	6	7	8	9	10	11	12	13	1
CD →	HJ →	PR →	XZ →	ES →	ON →	LY →	FG →	KM →	VW →	QU →	TI →	AB →	CD . . .

Now find, in the foregoing chain, two pairs likely to be sequent, for example HJ and KM and count the interval between them in the chain. It is 7 (counting by pairs). If this decimation interval is now applied to the chain of pairs, the following is established:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
H	J	K	M	P	R	V	W	X	Z	Q	U	E	S	T	I	O	N	A	B	L	Y	C	D	F	G

l. The reason why a complete chain of 26 letters cannot be constructed from the secondary alphabet given under subparagraph *g* is that it represents a case in which two primary components of 26 letters were slid an *even* number of intervals apart. (This will be explained in further detail in subparagraph *r* below.) There are in all 12 such cases, none of which will admit of the construction of a complete chain of 26 letters. In addition, there is one case wherein, despite the fact that the primary components are an *odd* number of intervals apart, the secondary alphabet cannot be made to yield a complete chain of 26 letters for an equivalent primary component. This is the case in which the displacement is 13 intervals. Note the secondary alphabet based upon the primary components below (which are the same as those shown in subparagraph *d*):

PRIMARY COMPONENTS

```
Q U E S T I O N A B L Y C D F G H J K M P R V W X Z
D F G H J K M P R V W X Z Q U E S T I O N A B L Y C
```

SECONDARY ALPHABET

```
Plain........ A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Cipher....... R V Z Q G U E S K T I W O P M N D A H J F B L Y X C
```

m. If an attempt is made to construct a chain of letters from this secondary alphabet alone, no progress can be made because the alphabet is completely reciprocal. However, the cryptanalyst need not at all be baffled by this case. The attack will follow along the lines shown below in subparagraphs *n* and *o*.

n. If the original primary component is a key-word mixed sequence, the cryptanalyst may reconstruct it by attempting to "dovetail" the 13 reciprocal pairs (AR, BV, CZ, DQ, EG, FU, HS, IK, JT, LW, MO, NP, and XY) into one sequence. The members of these pairs are all 13 intervals apart. Thus:

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13
A . . . . . . . . . . . . R
B . . . . . . . . . . . . V
C . . . . . . . . . . . . Z
D . . . . . . . . . . . . Q
E . . . . . . . . . . . . G
F . . . . . . . . . . . . U
H . . . . . . . . . . . . S
I . . . . . . . . . . . . K
J . . . . . . . . . . . . T
L . . . . . . . . . . . . W
M . . . . . . . . . . . . O
N . . . . . . . . . . . . P
X . . . . . . . . . . . . Y
```

FIGURE 22.

Write out the series of numbers from 1 to 26 and insert as many pairs into position as possible. being guided by considerations of probable partial sequences in the key-word mixed sequence, Thus:

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
A B C D . . . . . . . . . R V Z Q
```

It begins to look as though the key-word commences with the letter Q, in which case it should be followed by U. This means that the next pair to be inserted is FU. Thus:

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
A B C D F . . . . . . . . R V Z Q U
```

The sequence A B C D F means that E is in the key. Perhaps the sequence is A B C D F G H. Upon trial, using the pairs EG and HS, the following placements are obtained:

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
A B C D F G H . . . . . . R V Z Q U E S
```

This suggests the word QUEST or QUESTION. The pair JT is added:

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
A B C D F G H J . . . . . R V Z Q U E S T
```

Over

The sequence G H J suggests G H J K, which places an I after T. Enough of the process has been shown to make the steps clear.

o. Another method of circumventing the difficulties introduced by the 14th secondary alphabet (displacement interval, 13) is to use it in conjunction with another secondary alphabet which is produced by an even-interval displacement. For example, suppose the following two secondary alphabets are available.[1]

Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	R	V	Z	Q	G	U	E	S	K	T	I	W	O	P	M	N	D	A	H	J	F	B	L	Y	X	C
2	X	Z	E	S	K	T	I	O	R	N	A	Q	B	W	V	L	H	Y	M	P	J	C	D	F	U	G

FIGURE 23.

The first of these secondaries is the 13-interval secondary; the second is one of the even-interval secondaries, from which only half-chain sequences can be constructed. But if the construction be based upon the two sequences, 1 and 2 in the foregoing diagram, the following is obtained:

R X U T N L D H M V Z E I A Y F J P W Q S O B C G K

This is a complete equivalent primary component. The original key-word mixed component can be recovered from it by decimation based upon the 9th interval:

R V W X Z Q U E S T I O N A B L Y C D F G H J K M P

p. (1) When the primary components are identical mixed sequences proceeding in *opposite* directions, all the secondary alphabets will be reciprocal alphabets. Reconstruction of the primary component can be accomplished by the procedure indicated under subparagraph *o* above. Note the following three reciprocal secondary alphabets:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	P	M	H	G	Q	F	D	C	W	Y	L	K	B	R	V	A	E	N	Z	X	U	O	I	T	J	S
2	W	V	M	K	S	J	H	G	Q	F	D	R	C	X	Z	Y	I	L	E	U	T	B	A	N	P	O
3	T	S	~~S~~	Z	L	X	W	V	N	R	P	E	M	I	O	K	C	J	B	A	Y	H	G	F	U	D

FIGURE 24.

rset Q/

(2) Using lines 1 and 2, the following chain can be constructed (equivalent primary component): rset

Y/

P W Q S O B C G K R X U T N L D H M V Z E I A ~~T~~ F J

[1] The method of writing down the secondaries shown in figure 23 will hereafter be followed in all cases when alphabet reconstruction skeletons are necessary. The top line will be understood to be the plain component; it is common to all the secondary alphabets, and is set off from the cipher components by the heavy black line. This top line of letters will be designated by the digit Ø, and will be referred to as "the zero line" in the diagram. The successive lines of letters, which occupy the space below the zero line and which contain the various cipher components of the several secondary alphabets, will be numbered serially. These numbers may then be used as reference numbers for designating the horizontal lines in the diagram. The numbers standing above the letters may be used as reference numbers for the vertical columns in the diagram. Hence, any letter in the reconstruction skeleton may be designated by coordinates, giving the horizontal or X coordinate first. Thus, D (2–11) means the letter D standing in line 2, Column 11.

Or, using lines 2 and 3:

W T Y K Z O D P U A G V S L J X I C M Q N F R E B H

The original key-word mixed primary component (based on the word QUESTIONABLY) can be recovered from either of the two foregoing equivalent primary components. But if lines 1 and 3 are used, only half-chains can be constructed:

P T F X A K E C V O H Q L and M S D W N J U Y R I G Z B

This is because 1 and 3 are both odd-interval secondary alphabets, whereas 2 is an even-interval secondary. It may be added that odd-interval secondaries are characterized by having two cases in which a plain-text letter is enciphered by itself; that is, θ_p is identical with θ_c. This phrase "identical with" will be represented by the symbol $\equiv$; the phrase "not identical with" will be represented by the symbol $\not\equiv$. (Note that in secondary alphabet number 1 above, $F_p \equiv F_c$ and $U_p \equiv U_c$; in secondary alphabet number 3 above, $M_p \equiv M_c$ and $O_p \equiv O_c$). This characteristic will enable the cryptanalyst to select at once the proper two secondaries to work with in case several are available; one should show two cases where $\theta_p \equiv \theta_c$; the other should show none.

q. (1) When the primary components are different mixed sequences, their reconstruction from secondary cipher alphabets follows along the same lines as set forth above, under *b* to *j*, inclusive, with the exception that the selection of letters for building up the chain of equivalents for the primary cipher component is restricted to those below the zero line in the reconstruction skeleton. Having reconstructed the primary cipher component, the plain component can be readily reconstructed. This will become clear if the student will study the following example.

Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	T	V	A	B	U	L	I	Q	X	Y	C	W	S	N	D	P	F	E	Z	G	R	H	J	K	M	O
2	Z	J	S	T	V	I	Q	R	M	O	N	K	X	E	A	G	B	W	P	L	H	Y	C	D	F	U

FIGURE 25.

(2) Using only lines 1 and 2, the following chain is constructed:

T Z P G L I Q R H Y O U V J C N E W K D A S X M F B

This is an equivalent primary cipher component. By finding the values of the successive letters of this chain in terms of the plain component of secondary alphabet number 1 (the zero line), the following is obtained:

T Z P G L I Q R H Y O U V J C N E W K D A S X M F B
A S P T F G H U V J Z E B W K N R L X O C M I Y Q D

The sequence A S P T . . . is an equivalent primary plain component. The original key-word mixed components may be recovered from each of the equivalent primary components. That for the primary plain component is based upon the key PUBLISHERS MAGAZINE; that for the primary cipher component is based upon the key QUESTIONABLY.

Over

(3) Another method of accomplishing the process indicated above can be illustrated graphically by the following two chains, based upon the two secondary alphabets set forth in subparagraph *q* (1):

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	T	V	A	B	U	L	I	Q	X	Y	C	W	S	N	D	P	F	E	Z	G	R	H	J	K	M	O
2	Z	J	S	T	V	I	Q	R	M	O	N	K	X	E	A	G	B	W	P	L	H	Y	C	D	F	U

Col. 1.		Col. 2.					
A (Ø–1)	→	T (1–1);	→	T (2–4)	→	D (Ø–4);	→
D (Ø–4)	→	B (1–4);	→	B (2–17)	→	Q (Ø–17);	→
Q (Ø–17)	→	F (1–17);	→	F (2–25)	→	Y (Ø–25);	→
Y (Ø–25)	→	M (1–25);	→	M (2–9)	→	I (Ø–9);	→
I (Ø–9)	→	X (1–9);	→	X (2–13)	→	M (Ø–13);	→
M (Ø–13)	→	S (1–13);	→	S (2–3)	→	C (Ø–3);	→
etc.		etc.					

FIGURE 26.

(4) By joining the letters in Column 1, the following chain is obtained: A D Q Y I M, etc. If this be examined, it will be found to be an equivalent primary of the sequence based upon PUBLISHERS MAGAZINE. By joining the letters in Column 2, the following chain is obtained: T B F M X S. This is an equivalent primary of the sequence based upon QUESTIONABLY.

r. A final word concerning the reconstruction of primary components in general may be added. It has been seen that in the case of a 26-element component sliding against itself (both components proceeding in the same direction), it is only the secondary alphabets resulting from odd-interval displacements of the primary components which permit of reconstructing a single 26-letter chain of equivalents. This is true except for the 13th interval displacement, which even though an odd number, still acts like an even number displacement in that no complete chain of equivalents can be established from the secondary alphabet. This exception gives the clue to the basic reason for this phenomenon: it is that the number 26 has two factors, 2 and 13, which enter into the picture. With the exception of displacement-interval 1, *any displacement interval which is a sub-multiple of, or has a factor in common with the number of letters in the primary sequence will yield a secondary alphabet from which no complete chain of 26 equivalents can be derived for the construction of a complete equivalent primary component.* This general rule is applicable only to components which progress in the same direction; if they progress in opposite directions, all the secondary alphabets are reciprocal alphabets and they behave exactly like the reciprocal secondaries resulting from the 13-interval displacement of two 26-letter identical components progressing in the same direction.

s. The foregoing remarks give rise to the following observations based upon the general rule pointed out above. Whether or not a complete equivalent primary component is derivable by decimation from an original primary component (and if not, the lengths and numbers of chains of letters, or incomplete components, that can be constructed in attempts to derive such equivalent components) will depend upon the number of letters in the original primary component and the specific decimation interval selected. For example, in a 26-letter original primary component, decimation interval 5 will yield a complete equivalent primary component of 26 letters, whereas decimation intervals 4 or 8 will yield 2 chains of 13 letters each. In a 24-letter component, decimation interval 5 will also yield a complete equivalent primary component (of 24 letters), but decimation interval 4 will yield 6 chains of 4 letters each, and decimation interval 8 will yield 3 chains of 8 letters each. It also follows that in the case of an original primary component in which the total number of characters is a prime number, *all* decimation intervals will yield complete equivalent primary components. The following table has been drawn up in the light of these observations, for original primary sequences from 16 to 32 elements. (All prime-number sequences have been omitted.) In this table, the column at the extreme left gives the various decimation intervals, omitting in each case the first interval, which merely gives the original primary sequence, and the last interval, which merely gives the original sequence reversed. The top line of the table gives the various lengths of original primary sequences from 32 down to 16. (The student should bear in mind that sequences containing characters in addition to the letters of the alphabet may be encountered; he can add to this table when he is interested in sequences of more than 32 characters.) The numbers within the table then show, for each combination of decimation interval and length of, original sequence, the lengths of the chains of characters that can be constructed. (The student may note the symmetry in each column.) The bottom line shows the total number of complete equivalent primary components which can be derived for each different length of original component.

Decimation interval	Number of characters in original primary component											
	32	30	28	27	26	25	24	22	21	20	18	16
2	16	15	14	27	13	25	12	11	21	10	9	8
3	32	10	28	9	26	25	8	22	7	20	6	16
4	8	15	7	27	13	25	6	11	21	5	9	4
5	32	6	28	27	26	5	24	22	21	4	18	16
6	16	5	14	9	13	25	4	11	7	10	3	8
7	32	30	4	27	26	25	24	22	3	20	18	16
8	4	15	7	27	13	25	6	11	21	5	9	2
9	32	10	28	9	26	25	8	22	7	20	2	16
10	16	3	14	27	13	5	12	11	21	2	9	8
11	32	30	28	27	26	25	24	2	21	20	18	16
12	8	5	7	9	13	25	2	11	7	5	3	4
13	32	30	28	27	2	25	24	22	21	20	18	16
14	16	15	2	27	13	25	12	11	3	10	9	8
15	32	2	28	9	26	5	8	22	7	4	6	
16	2	15	7	27	13	25	6	11	21	5	9	
17	32	30	28	27	26	25	24	22	21	20		
18	16	5	14	9	13	25	4	11	7	10		
19	32	30	28	27	26	25	24	22	21			
20	8	3	7	27	13	5	6	11				
21	32	10	4	9	26	25	8					
22	16	15	14	27	13	25	12					
23	32	30	28	27	26	25						
24	4	5	7	9	13							
25	32	6	28	27								
26	16	15	14									
27	32	10										
28	8	15										
29	32											
30	16											
Total number of complete sequences	14	6	10	16	10	18	16	8	10	6	4	6

Over

Section VIII

APPLICATION OF PRINCIPLES OF INDIRECT SYMMETRY OF POSITION

32. Applying the principles to a specific example.—*a.* The preceding section, with the many details covered, now forms a sufficient base for proceeding with an exposition of how the principles of indirect symmetry of position can be applied very early in the solution of a polyalphabetic substitution cipher in which sliding primary components were employed to produce the secondary cipher alphabets for the enciphering of the cryptogram.

b. The case described below will serve not only to explain the method of applying these principles but will at the same time show how their application greatly facilitates the solution of a single, rather difficult, polyalphabetic substitution cipher. It is realized, of course, that the cryptogram could be solved by the usual methods of frequency and long, patient experimentation. However, the method to be described was actually applied and very materially reduced the amount of time and labor that would otherwise have been required for solution.

33. The cryptogram employed in the exposition.—*a.* The problem that will be used in this exposition involves an actual cryptogram submitted for solution in connection with a cipher device having two concentric disks upon which the same random mixed alphabet appears, both alphabets progressing in the same direction. This was obtained from a study of the descriptive circular accompanying the cryptogram. By the usual process of factoring, it was determined that the cryptogram involved 10 alphabets. The message as arranged according to its period is shown in Figure 27, in which all repetitions of two or more letters are indicated.

b. The triliteral frequency distributions are given in Figure 28. It will be seen that on account of the brevity of the message, considering the number of alphabets involved, the frequency distributions do not yield many clues. By a very careful study of the repetitions, tentative individual determinations of values of cipher letters, as illustrated in Figures 29, 30, 31, and 32, were made. These are given in sequence and in detail in order to show that there is nothing artificial or arbitrary in the preliminary stages of analysis here set forth.

The Cryptogram

(Repetitions underlined)

	1	2	3	4	5	6	7	8	9	10
A	W	F	U	P	C	F	O	C	J	Y
B	G	B	Z	D	P	F	B	O	U	O
C	G	R	F	T	Z	M	Q	M	A	V
D	K	Z	U	G	D	Y	F	T	R	W
E	G	J	X	N	L	W	Y	O	U	X
F	I	K	W	E	P	Q	Z	O	K	Z
G	P	R	X	D	W	L	Z	I	C	W
H	G	K	Q	H	O	L	O	D	V	M
I	G	O	X	S	N	Z	H	A	S	E
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E
N	G	S	X	E	R	O	Z	J	S	E
O	G	V	Q	W	E	J	M	K	G	H

	1	2	3	4	5	6	7	8	9	10
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
S	C	F	B	S	C	V	X	C	H	Q
T	Z	T	Z	S	D	M	X	W	C	M
U	R	K	U	H	E	Q	E	D	G	X
V	F	K	V	H	P	J	J	K	J	Y
W	Y	Q	D	P	C	J	X	L	L	L
X	G	H	X	E	R	O	Q	P	S	E
Y	G	K	B	W	T	L	F	D	U	Z
Z	O	C	D	H	W	M	Z	T	U	Z
AA	K	L	B	P	C	J	O	T	X	E
BB	H	S	P	O	P	N	M	D	L	M
CC	G	C	K	W	D	V	B	L	S	E
DD	G	S	U	G	D	P	O	T	H	X

	1	2	3	4	5	6	7	8	9	10
EE	B	K	D	Z	F	M	T	G	Q	J
FF	L	F	U	Y	D	T	Z	V	H	Q
GG	Z	G	W	N	K	X	J	T	R	N
HH	Y	T	X	C	D	P	M	V	L	W
II	B	G	B	W	W	O	Q	R	G	N
JJ	H	H	V	L	A	Q	Q	V	A	V
KK	J	Q	W	O	O	T	T	N	V	Q
LL	B	K	X	D	S	O	Z	R	S	N
MM	Y	U	X	O	P	P	Y	O	X	Z
NN	H	O	Z	O	W	M	X	C	G	Q
OO	J	J	U	G	D	W	Q	R	V	M
PP	U	K	W	P	E	F	X	E	N	F
QQ	C	C	U	G	D	W	P	E	U	H
RR	Y	B	W	E	W	V	M	D	W	J
SS	R	Z	X							

Figure 27.

~~288074 O - 41 - 5~~

Skip fig. 5

Over

Triliteral Frequency Distributions

I

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	EB	FF			XK	YB	ES	XK	ZC	VZ	WQ			ZC	ZR	DC	HC	HR		MK		–F		YQ	QT
	HZ	FC				OR	NH		VQ	ZL	JF						MK							NT	QG
	XK					WJ	ZO		QJ								JZ							NU	
	WG					WK																		HB	
	QK					MO																			
						ES																			
						EV																			
						LH																			
						EK																			
						MC																			
						ES																			

II

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	GZ	QB			WU	ZW	GX		GX	IW	KB			GX		LZ	GF	GX	ZZ	YX	GQ				KU
	BJ	JQ			CB	BB	HV		JU	GQ				HZ		YD	PX	HP	YX						BZ
	YW	RV			LU					RU						JW	SQ	GU							RX
		OD								FV															
		GK								GB															
		CU								BD															
										BX															
										UW															

III

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	CZ		QP		RT				BI	CW					SO	KH				FP	CO	KE	JN		BD
	FS		CH													CR				ZG	KH	GN	RD		QA
	KW		KZ													RE				KH	HL	QO	OS		ZC
	LP															VW				SG		KP	SE		TS
	GW																			FY		BE	HE		OO
																				JG			TC		
																				CG			KD		
																							UO		
																							Z–		

IV

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
ZA		ZK	ZP	WP		UD	QO	JP			VA		XL	VP	UC		QQ	XN	FZ			QE		UD	BE
		XD	XW	QW		UD	UE						WK	PP	DC			BC				BT			DF
			XS	XR		UD	VP							WO	BC			ZD				KD			
				XR		UD	DW							XP	WE							BW			
				WW										ZW											

V

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
AA		PF	GY	ZX	ZM					CQ	NW		SZ	HL	DF	RF	EO	DO	WL			DL			TM
LQ		SV	SM	WJ						NX				OT	EQ		EO					EM			
		PJ	WV	HQ											IQ							HM			
		PJ	GP	PF											ON							WO			
			YT												HJ							OM			
			GP												ON							EV			
			GW												OP										
			GW																						

VI

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
AM					CO				EM		WZ	ZQ	PB	RZ	DO	PZ			DZ		CX	LY	EQ	DF	NH
					PB				PJ		OO	WL	PM	RQ	DM	PF			OT		DB	DQ	KJ		
					QV				CX		TF	DX		WQ	PY	KO					WM	DP			
					EX				CO			WZ		SZ		EE									
												FT				AQ									
												WX													

VII

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	FO			QD	YT		ZA		JK		MN	JK		FC	WE	MM			MG		FM		VC	WO	QO
	NL				QJ				XT			AD		LD		XT			TN				MW	PO	LI
	VL				LD							ND		QI		OP							JL		OJ
												PV		JT		OR							MC		MT
												VD		PT		QV							FE		TV
																WR									OR

VIII

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
HS		OJ	OV	XN		TQ		ZC	FH	MG	BC	QA	LA	BU	QS		QG		FR		ZH	XC			
		XH	MC	PU				OK	ZS	JJ	XL	VL	TV	YU			ZS		QX		ML				
		XG	EG								BS			ZK			QV		ZU		QA				
			FU											YX					OX						
			ML																OH						
			MY																JR						

IX

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
MV		IW				KH	JD		CY	OZ	MH		EF			GJ	TW	AE		OO	DM		TZ	DJ	
NE		LW				DX	CQ		KY	IF	LL						TN	JE		OX	NQ		TE		
VV		DH				RN	TX				DM							PE		DZ	RM		OZ		
		WM				CQ	VQ				VW							LE		TZ					
																		RN		EH					

X

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	Z	Y	Z
			HQ	SB	KC		LS		QL		LG	VG	RY	UG		HZ					AK	RG	UI	JG	KP
				AG	NC		GR		YR			CR	GH			HZ					AJ	CG	GF	JY	XJ
				SG			CB					LG	SY			VB						CL	HB		UO
				SG			UY					VU				GJ						LB			UK
				XH																					XH
				SG																					

FIGURE 28.

Over

Initial Values From Assumptions

10/ $\overset{1}{G_c}=E_p$; $\overset{2}{K_c}=E_p$; $\overset{3}{X_c}=E_p$; and $\overset{5}{D_c}=E_p$, from frequency considerations. set

$\overset{345}{\text{UGD}}$=THE; $\overset{456}{\text{PCJ}}$=THE; and $\overset{9\,10\,1}{\text{SEG}}$=THE, from study of repetitions.

	1	2	3	4	5	6	7	8	9	10
A	W	F	U	P	C	F	O	C	J	Y
			T	T	H					
B	G	B	Z	D	P	F	B	O	U	O
	E									
C	G	R	F	T	Z	M	Q	M	A	V
	E									
D	K	Z	U	G	D	Y	F	T	R	W
			T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E								
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E								
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E
										H
N	G	S	X	E	R	O	Z	J	S	E
	E		E						T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				

	1	2	3	4	5	6	7	8	9	10
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
S	C	F	B	S	C	V	X	C	H	Q
					H					
T	Z	T	Z	S	D	M	X	W	C	M
					E					
U	R	K	U	H	E	Q	E	D	G	X
		E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E				E				
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				
X	G	H	X	E	R	O	Q	P	S	E
	E		E						T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
AA	K	L	B	P	C	J	O	T	X	E
				T	H	E				H
BB	H	S	P	O	P	N	M	D	L	M
CC	G	C	K	W	D	V	B	L	S	E
	E				E				T	H
DD	G	S	U	G	D	P	O	T	H	X
	E		T	H	E					

	1	2	3	4	5	6	7	8	9	10
EE	B	K	D	Z	F	M	T	G	Q	J
		E								
FF	L	F	U	Y	D	T	Z	V	H	Q
			T		E					
GG	Z	G	W	N	K	X	J	T	R	N
HH	Y	T	X	C	D	P	M	V	L	W
			E		E					
II	B	G	B	W	W	O	Q	R	G	N
JJ	H	H	V	L	A	Q	Q	V	A	V
KK	J	Q	W	O	O	T	T	N	V	Q
LL	B	K	X	D	S	O	Z	R	S	N
		E	E						T	
MM	Y	U	X	O	P	P	Y	O	X	Z
NN	H	O	Z	O	W	M	X	C	G	Q
OO	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
PP	U	K	W	P	E	F	X	E	N	F
		E		T						
QQ	C	C	U	G	D	W	P	E	U	H
			T	H	E					
RR	Y	B	W	E	W	V	M	D	Y	J
SS	R	Z	X							
			E							

Figure 20.

ADDITIONAL VALUES FROM ASSUMPTIONS (I)

Refer to line DD in Figure 29; $\overset{2}{S}_e$ assumed to be N_p.

Refer to line M in figure 29; $\overset{9}{A}_e$ assumed to be W_p.

Then in lines C–D, $\overset{9}{A}\ \overset{10}{V}\ \overset{1}{K}\ \overset{2}{Z}\ \overset{3}{U}\ \overset{4}{G}\ \overset{5}{D}$ is assumed to be WITH THE.

	1	2	3	4	5	6	7	8	9	10
A	W	F	U	P	C	F	O	C	J	Y
			T	T	H					
B	G	B	Z	D	P	F	B	O	U	O
	E									
C	G	R	F	T	Z	M	Q	M	A	V
	E								W	I
D	K	Z	U	G	D	Y	F	T	R	W
	T	H	T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E								
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E								
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E
									W	H
N	G	S	X	E	R	O	Z	J	S	E
	E	N	E						T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				

	1	2	3	4	5	6	7	8	9	10
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
		H								
S	C	F	B	S	C	V	X	C	H	Q
					H					
T	Z	T	Z	S	D	M	X	W	C	M
					E					
U	R	K	U	H	E	Q	E	D	G	X
		E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E				E				
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				
X	G	H	X	E	R	O	Q	P	S	E
	E		E						T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
AA	K	L	B	P	C	J	O	T	X	E
	T			T	H	E				
BB	H	S	P	O	P	N	M	D	L	M
		N								
CC	G	C	K	W	D	V	B	L	S	E
	E				E				T	H
DD	G	S	U	G	D	P	O	T	H	X
	E	N	T	H	E					

	1	2	3	4	5	6	7	8	9	10
EE	B	K	D	Z	F	M	T	G	Q	J
		E								
FF	L	F	U	Y	D	T	Z	V	H	Q
			T		E					
GG	Z	G	W	N	K	X	J	T	R	N
HH	Y	T	X	C	D	P	M	V	L	W
			E		E					
II	B	G	B	W	W	O	Q	R	G	N
JJ	H	H	V	L	A	Q	Q	V	A	V
									W	I
KK	J	Q	W	O	O	T	T	N	V	Q
LL	B	K	X	D	S	O	Z	R	S	N
		E	E						T	
MM	Y	U	X	O	P	P	Y	O	X	Z
NN	H	O	Z	O	W	M	X	C	G	Q
OO	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
PP	U	K	W	P	E	F	X	E	N	F
		E		T						
QQ	C	C	U	G	D	W	P	E	U	H
			T	H	E					
RR	Y	B	W	E	W	V	M	D	Y	J
SS	R	Z	X							
		H	E							

FIGURE 30.

Over

Additional Values from Assumptions (II)

Refer to Figure 30, line A; W F U P C F O C J Y (1–10); assume to be BUT THOUGH.
— — T T H — — — — —

Refer to Figure 30, lines N and X, where repetition X E R O (3–6) occurs; assume EACH
E — — —

11 pi.

v497

	1	2	3	4	5	6	7	8	9	10
A	W	F	U	P	C	F	O	C	J	Y
	B	U	T	T	H	O	U	G	H	
B	G	B	Z	D	P	F	B	O	U	O
	E					O				
C	G	R	F	T	Z	M	Q	M	A	V
	E								W	I
D	K	Z	U	G	D	Y	F	T	R	W
	T	H	T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E		A						
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E					U			
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
						Q				
M	S	R	Q	E	W	M	L	N	A	E
				A					W	H
N	G	S	X	E	R	O	Z	J	S	E
	E	N	E	A	C	H			T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				

	1	2	3	4	5	6	7	8	9	10
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
		H					U			
S	C	F	B	S	C	V	X	C	H	Q
		U			H			G		
T	Z	T	Z	S	D	M	X	W	C	M
					E					
U	R	K	U	H	E	Q	E	D	G	X
	T	E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E					E		H	
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				
X	G	H	X	E	R	O	Q	P	S	E
	E		E	A	C	H			T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
AA	K	L	B	P	C	J	O	T	X	E
	T			T	H	E	U			H
BB	H	S	P	O	P	N	M	D	L	M
		N								
CC	G	C	K	W	D	V	B	L	S	E
	E				E				T	H
DD	G	S	U	G	D	P	O	T	H	X
	E	N	T	H	E		U			

	1	2	3	4	5	6	7	8	9	10
EE	B	K	D	Z	F	M	T	G	Q	J
		E								
FF	L	F	U	Y	D	T	Z	V	H	Q
		U	T		E					
GG	Z	G	W	N	K	X	J	T	R	N
HH	Y	T	X	C	D	P	M	V	L	W
			E			E				
II	B	G	B	W	W	O	Q	R	G	N
						H				
JJ	H	H	V	L	A	Q	Q	V	A	V
									W	I
KK	J	Q	W	O	O	T	T	N	V	Q
LL	B	K	X	D	S	O	Z	R	S	N
		E	E			H			T	
MM	Y	U	X	O	P	P	Y	O	X	Z
NN	H	O	Z	O	W	M	X	C	G	Q
								G		
OO	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
PP	U	K	W	P	E	F	X	E	N	F
		E		T		O				
QQ	C	C	U	G	D	W	P	E	U	H
			T	H	E					
RR	Y	B	W	E	W	V	M	D	Y	J
				A						
SS	R	Z	X							
		H	E							

Figure 31.

I/ set

J/

set

set

change indicator as indicated

M/ set

Change wiring between "K+L" as per all other alphabetical letters.

E/ set line

Additional Values From Assumptions (III)

456
OPN—assume ING from repetition and frequency.

901
HQZ—assume ING from repetition and frequency.

	1	2	3	4	5	6	7	8	9	10
A	W	F	U	P	C	F	O	C	J	Y
	B	U	T	T	H	O	U	G	H	
B	G	B	Z	D	P	F	B	O	U	O
	E				N	O				
C	G	R	F	T	Z	M	Q	M	A	V
	E								W	I
D	K	Z	U	G	D	Y	F	T	R	W
	T	H	T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E		A	N					
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E					U			
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
					N				I	
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	O	R	Q	F	V	M	L	H
						O				
M	S	R	Q	E	W	M	L	N	A	E
				A					W	H
N	G	S	X	E	R	O	Z	J	S	E
	E	N	E	A	C	H			T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				

	1	2	3	4	5	6	7	8	9	10
P	R	C	V	O	P	N	B	L	C	W
				I	N	G				
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
		H					U			
S	C	F	B	S	C	V	X	C	H	Q
		U			H			G	I	N
T	Z	T	Z	S	D	M	X	W	C	M
	G				E					
U	R	K	U	H	E	Q	E	D	G	X
		E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E			N	E			H	
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				
X	G	H	X	E	R	O	Q	P	S	E
	E		E	A	C	H			T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
AA	K	L	B	P	C	J	O	T	X	E
	T			T	H	E	U			H
BB	H	S	P	O	P	N	M	D	L	M
		N		I	N	G				
CC	G	C	K	W	D	V	B	L	S	E
	E				E				T	H
DD	G	S	U	G	D	P	O	T	H	X
	E	H	T	H	E		U		I	

	1	2	3	4	5	6	7	8	9	10
EE	B	K	D	Z	F	M	T	G	Q	J
		E								
FF	L	F	U	Y	D	T	Z	V	H	Q
		U	T		E				I	N
GG	Z	G	W	N	K	X	J	T	R	N
	G									
HH	Y	T	X	C	D	P	M	V	L	W
			E		E					
II	B	G	B	W	W	O	Q	R	G	N
						H				
JJ	H	H	V	L	A	Q	Q	V	A	V
									W	I
KK	J	Q	W	O	O	T	T	N	V	Q
				I						N
LL	B	K	X	D	S	O	Z	R	S	N
		E	E			H			T	
MM	Y	U	X	O	P	P	Y	O	X	Z
				I	N					
NN	H	O	Z	O	W	M	X	C	G	Q
				I				G		N
OO	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
PP	U	K	W	P	E	F	X	E	N	F
		E		T		O				
QQ	C	C	U	G	D	W	P	E	U	H
			T	H	E					
RR	Y	B	W	E	W	V	M	D	Y	J
				A						
SS	R	Z	X							
		H	E							

Figure 32.

not set

set

10/

set

tr/

VHPJJK

E NEE H

58

Over

c. From the initial and subsequent tentative identifications shown in Figures 29, 30, 31, and 32, the values obtained were arranged in the form of the secondary alphabets in a reconstruction skeleton, shown in Figure 33.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1		W			G		Z													K						
2					K			Z						S							F					
3					X															U						
4	E							G	O											P						
5			R		D			C						P												
6					J		N	O							F											
7																					O					
8							C																			
9								J	H											S			A			
10								E	V																	

FIGURE 33.

C447

34. Fundamental theory.—*a*. In paragraph 31, methods of reconstructing primary components from secondary alphabets were given in detail. It is necessary that those methods be fully understood before the following steps be studied. It was there shown that the primary component can be one of a series of equivalent primary sequences, all of which will give exactly similar results so far as the secondary alphabets and the cryptographic text are concerned. It is not necessary that the identical or original primary component employed in the cryptographing be reconstructed; any equivalent primary sequence will serve. The whole question is one of establishing a sequence of letters the interval between which is either identical with that in the original primary component or else is an exact constant multiple of the interval separating the letters in the original primary component. For example, suppose K P X N Q forms a sequence in the original primary component. Here the interval between K and P, and P and X, X and N, N and Q is one; in an equivalent primary component, say the sequence K . . P . . X . . N . . Q, the interval between K and P is three, that between P and X also three, and so on; and the two sequences will yield the same secondary alphabets. So long as the interval between K and P, P and X, X and N, N and Q, . . ., is a constant one, the sequence will be cryptographically equivalent to the original primary sequence and will yield the same secondary alphabets as do those of the original primary sequence. However, in the case of a 26-letter component, it is necessary that this interval be an odd number other than 13, as these are the only cases which will yield one unbroken sequence of 26 letters. Suppose a secondary alphabet to be as follows:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
(1)	Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	Cipher														X		K	N							P		

It can be said that the primary component contains the following sequences:

XN KP NQ PX

These, when united by means of their common letters, yield K P X N Q.

Suppose also the following secondary alphabet is at hand:

(2)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plain	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher														P			X							K		N

Here the sequences PN, XQ, KX, and NZ can be obtained, which when united yield the two sequences KXQ and PNZ.

By a comparison of the sequences K P X N Q, K X Q, and P N Z, one can establish the following:

```
K P X N Q
K . X . Q
  P . N . Z
```

It follows that one can now add the letter Z to the sequence, making it K P X N Q Z.

b. The reconstruction of a primary component from one of the secondary alphabets by the process given in paragraph 31 requires a complete or nearly complete secondary alphabet. This is at hand only *after* a cryptogram has been completely solved. But if one could employ several very scant or skeletonized secondary alphabets simultaneously with the analysis of the cryptogram, one could then possibly build up a primary component from fewer data and thus solve the cryptogram much more rapidly than would otherwise be possible.

c. Suppose only the cipher components of the two secondary alphabets (1) and (2) given above be placed into juxtaposition. Thus:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	X	.	K	N	.	.	.	.	.	.	P	.	.
(2)	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	.	X	.	.	.	.	.	.	K	.	N

The sequences PX, XN, and KP are given by juxtaposition. These, when united, yield KPXN as part of the primary sequence. It follows, therefore, that *one can employ the cipher components of secondary alphabets as sources of independent data* to assist in building up the primary sequences. The usefulness of this point will become clearer subsequently.

35. Application of principles.—*a.* Refer now to the reconstruction skeleton shown in Figure 33. Hereafter, in order to avoid all ambiguity and for ease in reference, the position of a letter in Figure 33 will be indicated as stated in footnote 1, page 56. Thus, N (6–7) refers to the letter N in line 6 and in column 7 of Figure 33.

b. (1) Now, consider the following pairs of letters:

```
 E (Ø–5)     J (6–5)
 G (Ø–7)     N (6–7)
{H (Ø–8)     O (6–8) }
{O (Ø–15)    F (6–15)} HO, OF=HOF
```

(One is able to use the line marked zero in Figure 33 since this is a mixed sequence sliding against *itself*.)

Over

(2) The immediate results of this set of values will now be given. Having HOF as a sequence, with EJ as belonging to the same displacement interval, suppose HOF and EJ are placed into juxtaposition as portions of sliding components. Thus:

Plain......	. . . H O F . . .
Cipher....	. . . E J . . . .

When $H_p=E_c$, then $O_p=J_c$.

(3) Refer now to alphabet 10, Figure 33, where it is seen that $H_p=E_c$. *The derived value,* $O_p=J_c$, *can immediately be inserted in the same alphabet* and substituted in the cryptogram.

(4) The student may possibly get a clearer idea of the principles involved if he will regard the matter as though he were dealing with arithmetical proportion. For instance, given any three terms in the proportion 2:8=4:16, the 4th term can easily be found. Furthermore, given the pair of values on the left-hand side of the equation, one may find numerous pairs of values which may be inserted in the right-hand side, or vice versa. For instance, 2:8=4:16 is the same as 2:8=5:20, or 9:36=4:16, and so on. An illustration of each of these principles will now be given, reference being made to Figure 33. As an example of the first principle, note that E (Ø–5):H (Ø–8)=J (6–5):O (6–8). Now find E (10–8):H (Ø–8)=? (10–15):O (Ø–15). It is clear that J may be inserted as the 3d term in this proportion, thus giving the important new value, $\overset{10}{O}_p=J_c$, which is exactly what was obtained directly above, by means of the partial sliding components. As an example of the second principle, note the following pairs:

E (Ø–5)	H (Ø–8)
K (2–5)	Z (2–8)
D (5–5)	C (5–8)
J (6–5)	O (6–8)

These additional pairs are also noted:

K (1–20)	Z (1–7)
T (Ø–20)	G (Ø–7)

Therefore, E:H=K:Z=D:C=J:O=T:G, and T may be inserted in position (4–5).

c. (1) Again, GN belongs to the same set of displacement-interval values as do EJ and HOF. Hence, by superimposition:

Plain......	. . . H O F . . .
Cipher....	. . . G N . . . .

(2) Referring to alphabet 4, when $H_p=G_c$, then $O_p=N_c$. Therefore, the letter N can be inserted in position (4–15) in Figure 33, and the value $\overset{4}{N}_c=O_p$ can be substituted in the cryptogram.

(3) Furthermore, note the corroboration found from this particular superimposition:

H (Ø–8)	G (Ø–7)
O (6–8)	N (6–7)

This checks up the value in alphabet 6, $G_p=N_c$.

d. (1) Again superimpose HOF and GN:

```
. . . H O F . . .
. . . . G N . . .
```

(2) Note this corroboration:

O (6–8)	G (4–8)
F (6–15)	N (4–15)

which has just been inserted in Figure 33, as stated above.

e. (1) Again using HOF and EJ, but in a different superimposition:

```
. . . H O F . .
. . E J . . . .
```

(2) Refer now to H (9–9), J (9–8). Directly under these letters is found V (10–9), E (10–8). Therefore, the V can be added immediately before H O F, making the sequence V H O F.

f. (1) Now take V H O F and juxtapose it with E J, thus:

```
. . . V H O F . . .
. . . E J . . .
```

(2) Refer now to Figure 33, and find the following:

V (10–9)	E (10–8)
H (9–9)	J (9–8)
O (4–9)	G (4–8)
I (Ø–9)	H (Ø–8)

(3) From the value O G it follows that G can be set next to J in E J. Thus:

```
. . . V H O F . . .
. . . E J G . . .
```

(4) But G N already is known to belong to the same set of displacement-interval values as E J. Therefore, it is now possible to combine E J, J G, and G N into one sequence, E J G N, yielding:

```
. . . V H O F . . .
. . . E J G N . . .
```

g. (1) Refer now to Figure 33.

V (Ø–22)	E (Ø–5)
? (1–22)	G (1–5)
? (2–22)	K (2–5)
? (3–22)	X (3–5)
? (5–22)	D (5–5)
? (6–22)	J (6–5)

(2) The only values which can be inserted are:

O (1–22)	G (1–5)
H (6–22)	J (6–5)

(3) This means that $V_p=O_c$ in alphabet 1 and that $V_p=H_c$ in alphabet 6. There is one O_c in the frequency distribution for alphabet 1, and no H_c in that for alphabet 6. The frequency distribution is, therefore, corroborative insofar as these values are concerned.

(*h*) (1) Further, taking E J G N and V H O F, superimpose them thus:

```
  . . . E J G N . . .
. . . V H O F . . .
```

(2) Refer now to Figure 33.

E (Ø–5)	H (Ø–8)
G (1–5)	? (1–8)

Over

(3) From the diagram of superimposition the value G (1–5) F (1–8) can be inserted, which gives $H_p=F_e$ in alphabet 1.

i. (1) Again, V H O F and E J G N are juxtaposed:

. . . V H O F . . .

. . . E J G N . . .

(2) Refer to Figure 33 and find the following:

H (Ø–8)	G (4–8)
A (Ø–1)	E (4–1)

This means that it is possible to add A, thus:

. . . A V H O F . . .

. . . E J G N . . .

(3) In the set there are also:

E (Ø–5)	G (1–5)
G (Ø–7)	Z (1–7)

Then in the superimposition

. . . E J G N . . .

. . . E J G N . . .

It is possible to add Z under G, making the sequence E J G N Z.

(4) Then taking

. . . A V H O F . . .

. . . E J G N Z . . .

and referring to Figure 33:

H (Ø–8)	N (Ø–14)
O (6–8)	? (6–14)

It will be seen that O=Z from superimposition, and hence in alphabet 6 $N_p=Z_e$, an important new value, but occurring only once in the cryptogram. Has an error been made? The work so far seems too corroborative in interlocking details to think so.

j. (1) The possibilities of the superimposition and sliding of the AVHOF and the EJGNZ sequences have by no means been exhausted as yet, but a little different trail this time may be advisable.

497

x

set

E (Ø–5)	T (Ø–20)
G (1–5)	K (1–20)
~~K~~ (3–5)	U (3–20)

(2) Then:

. . . E J G N Z . . .

. . . T . K . . .

(3) Now refer to the following:

E (Ø–5)	K (2–5)
N (Ø–14)	S (2–14)

whereupon the value S can be inserted:

```
   . . . E J G N Z . . .
. . . T . K . . S . . .
```

k. (1) Consider all the values based upon the displacement interval corresponding to JG:

```
J (6-5)  G (1-5)→|J (9- 8)  G (4- 8)
N (6-7)  Z (1-7) |H (9- 9)  O (4- 9)
                 |S (9-20)  P (4-20)→|S (2-14)  P (5-14)
                                     |Z (2- 8)  C (5- 8)
                                     |K (2- 5)  D (5- 5)
```

(2) Since J and G are sequent in the E J G N Z sequence, it can be said that all the letters of the foregoing pairs are also sequent. Hence Z C, S P, and K D are available as new data. These give E J G N Z C and T . K D . S P.

(3) Now consider:

```
T (Ø-20)   P (4-20)
A (Ø- 1)   E (4- 1)
H (Ø- 8)   G (4- 8)
I (Ø- 9)   O (4- 9)
```

Now in the T . K D . S P sequence the interval between T and P is T . . . . . P (123456). Hence the interval between A and E is 6 also. It follows therefore that the sequences A V H O F and E J G N Z C should be united, thus:

```
       1 2 3 4 5 6
. . . A V H O F . E J G N Z C . . .
```

(4) Corroboration is found in the interval between H and G, which is also six. The letter I can be placed into position, from the relation I (Ø-9) O (4-9), thus:

```
       1 2 3 4 5 6
. . . I . . A V H O F . E J G N Z C . . .
```

l. (1) From Figure 33:

```
H (Ø- 8)   Z (2- 8)
E (Ø- 5)   K (2- 5)
N (Ø-14)   S (2-14)
U (Ø-21)   F (2-21)
```

(2) Since in the I . . A V H O F . E J G N Z C sequence the letters H and Z are separated by 8 intervals one can write:

```
        1 2 3 4 5 6 7 8
. . . H . . . . . . . . Z . . .
. . . E . . . . . . . . K . . .
. . . N . . . . . . . . S . . .
. . . U . . . . . . . . F . . .
```

Over

(3) Hence one can make the sequence

```
                         1 2 3 4 5 6 7 8
      . . . I . . A V H O F . E J G N Z C . . K . . .
Then  . . . I . . A V H O F . E J G N Z C T . K D . S P . . .
and   . . U I . . A V H O F . E J G N Z C T . K D . S P . . .
          1 2 3 4 5 6 7 8           1 2 3 4 5 6 7 8
```

m. (1) Subsequent derivations can be indicated very briefly as follows:

E (∅–5)   C (∅–3)
D (5–5)   R (5–3)

```
              1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
From          U I . . A V H O F .  E  J  G  N  Z  C  T  .  K  D  .  S  P  .  .  .
one can write                . . . E  .  .  .  .  C  .  .  .
                                      1  2  3  4  5
and                                                  .  .  .  D  .  .  .  .  R  .
                                                                 1  2  3  4  5
making the sequence
              1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
              U I . . A V H O F .  E  J  G  N  Z  C  T  .  K  D  .  S  P  .  R  .
```

(2) Another derivation:

U (3–20)   T (∅–20)
X (3– 5)   E (∅– 5)

```
              1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
From          U I . . A V H O F .  E  J  G  N  Z  C  T  .  K  D  .  S  P  .  R  .
one can write
              U I . . . . . . . .  .  .  .  .  .  .  T  .  .  .
and           . . . . . . . . . .  E  .  .  .  .  .  .  .  .  .  X  .  .  .  .  .
making the sequence
              1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
              U I . . A V H O F .  E  J  G  N  Z  C  T  .  K  D  X  S  P  .  R  .
```

(3) Another derivation:

E (∅–5)   G (1–5)
B (∅–2)   W (1–2)

```
From           . . . E J G . . .
one can write  . . . E . G . . .
and then       . . . B . W . . .
```

There is only one place where B . W can fit, *viz*, at the end:

```
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
U I . . A V H O F .  E  J  G  N  Z  C  T  .  K  D  X  S  P  B  R  W
```

n. Only four letters remain to be placed into the sequence, *viz*, L, M, Q, and Y. Their positions are easily found by application of the primary component to the message. The complete sequence is as follows:

```
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
U I M Y A V H O F L  E  J  G  N  Z  C  T  Q  K  D  X  S  P  B  R  W
```

Having the primary component fully constructed, decipherment of the cryptogram can be completed with speed and precision. The text is as follows:

0497

W F U P C F O C J Y	R C V O P N B L C W	B K D Z F M T G Q J
B U T T H O U G H W	P O S I N G T H E S	S E L F W I L L G O
G B Z D P F B O U O	L Q Z A A A M D C H	L F U Y D T Z V H Q
E C A N N O T A S Y	O L A R S Y S T E M	O U T B E C O M I N
I G R F ~~X~~ Z M Q M A V	B Z Z C K Q O I K F	Z G W N K X J T R N
E T R E V I E W W I	S H A L L T U R N A	G A C O L D A N D L
K Z U G D Y F T R W	C F B S C V X C H Q	Y T X C D P M V L W
T H T H E M I N D S	N U N C H A N G I N	I F E L E S S M A S
G J X N L W Y O U X	Z T Z S D M X W C M	B G B W W O Q R G N
E Y E O U R P A S T	G F A C E I N P E R	S A N D T H E S O L
I T W E P Q Z O K Z	R K U H E Q E D G X	H H V L A Q Q V A V
W E C A N T O A N E	P E T U I T Y T O T	A R S Y S T E M W I
P R X C W L Z I C W	F K V H P J J K J Y	J Q W O O T T N V Q
X T E N T F O R E S	H E S U N E A C H W	L L C I R C L E U N
G K Q H O L O D V M	Y Q D P C J X L L L	B K X D S O Z R S N
E E O U R F U T U R	I L L T H E N H A V	S E E N G H O S T L
G O X S N Z H A S E	G H X E R O Q P S E	Y U X O P P Y O X Z
E W E C A N W I T H	E R E A C H E D T H	I K E I N S P A C E
B B J I P Q F J H D	G K B W T L F D U Z	H O Z O W M X C G Q
S C I E N T I F I C	E E N D O F I T S E	A W A I T I N G O N
Q C B Z E X Q T X Z	O C D H W M Z T U Z	J J U G J W Q R V M
C O N F I D E N C E	V O L U T I O N S E	L Y T H E R E S U R
M J C Q R Q F V M L ~~H~~	K L B P C J O T X E	U K W P E F X E N F
L O O K F O R W A R	T I N T H E U N C H	R E C T I O N O F A
S R Q E W M L N A E	H S P O P N M D L M	C C U G D W P E U H
D T O A T I M E W H	A N G I N G S T A R	N O T H E R C O S M
G S X E R O Z J S E	G C K W D V B L S E	Y B W E W V M D Y J
E N E A C H O F T H	E O F D E A T H T H	I C C A T A S T R O
G V Q W E J M K G H	G S U G D P O T H X	R Z X
E B O D I E S C O M	E N T H E S U N I T	P H E

FIGURE 34.

o. The primary component appears to be a random-mixed sequence; no key word is to be found, at least none reappears on experimentation with various hypotheses as to enciphering equations. Nevertheless, the random construction of the primary component did not complicate or retard the solution.

Over

p. Some students may prefer to work exclusively with the reconstruction skeleton, rather than with sliding strips. One method is as good as the other and personal preferences will dictate which will be used by the individual student. If the reconstruction skeleton is used, the original letters should be inserted in ink, so as to differentiate them from derived letters.

36. General remarks.—*a*. It is to be stated that the sequence of steps described in the preceding paragraphs corresponds quite closely with that actually followed in solving the problem. It is also to be pointed out that this method can be used as a control in the early stages of analysis because it will allow the cryptanalyst to check assumptions for values. For example, the very first value derived in applying the principles of indirect symmetry to the problem herein described was $H_c=A_p$ in alphabet 1. As a matter of fact the writer had been inclined toward this value, from a study of the frequency and combinations which H_c showed; when the indirect-symmetry method actually substantiated his tentative hypothesis he immediately proceeded to substitute the value given. If he had assigned a different value to H_c, or if he had assumed a letter other than H_c for A_p in that alphabet, the conclusion would immediately follow that either the assumed value for H_c was erroneous, or that one of the values which led to the derivation of $H_c=A_p$ by indirect symmetry was wrong. Thus, these principles aid not only in the systematic and nearly automatic derivation of new values (with only occasional, or incidental references to the actual frequencies of letters), but they also assist very materially in serving as corroborative checks upon the validity of the assumptions already made.

b. Furthermore, while the writer has set forth, in the reconstruction skeleton in Figure 33, a set of 30 values apparently obtained before he began to reconstruct the primary component, this was done for purposes of clarity and brevity in exposition of the principles herein described. As a matter of fact, what he did was to watch very carefully, when inserting values in the reconstruction skeleton to find the very first chance to employ the principles of indirect symmetry; and just as soon as a value could be derived, he substituted the value in the cryptographic text. This is good procedure for two reasons. Not only will it disclose impossible combinations but also it gives opportunity for making further assumptions for values by the addition of the derived values to those previously assumed. Thus, the processes of reconstructing the primary component and finding additional data for the reconstruction proceed simultaneously in an ever-widening circle.

c. It is worth noting that the careful analysis of only 30 cipher equivalents in the reconstruction skeleton shown in Figure 33 results in the derivation of the entire table of secondary alphabets, 676 values in all. And while the elucidation of the method seems long and tedious, in its actual application the results are speedy, accurate, and gratifying in their corroborative effect upon the mental activity of the cryptanalyst.

d. (1) The problem here used as an illustrative case is by no means one that most favorably presents the application and the value of the method, for it has been applied in other cases with much speedier success. For example, suppose that in a cryptogram of 6 alphabets the equivalents of only THE in all 6 alphabets are fairly certain. As in the previous case, it is supposed that the secondary alphabets are obtained by sliding a mixed alphabet against itself. Suppose the secondary alphabets to be as follows:

Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1					B			Q												E						
2					C			L												X						
3					I			V												C						
4					N			P												B						
5					X			O												P						
6					T			Z												V						

FIGURE 35.

(2) Consider the following chain of derivatives arranged diagrammatically:

```
H (Ø- 8)  O (5- 8)
T (Ø-20)  P (5-20)
E (Ø- 5)  X (5- 5)→E (1-20)  X (2-20)
                    Q (1- 8)  L (2- 8)
                    B (1- 5)  C (2- 5)→B (4-20)  C (3-20)
                                        N (4- 5)  I (3- 5)
                                        P (4- 8)  V (3- 8)→

→P (5-20)  V (6-20)
 O (5- 8)  Z (6- 8)
 X (5- 5)  T (6- 5)→X (2-20)  T (Ø-20)
                    L (2- 8)  H (Ø- 8)
                    C (2- 5)  E (Ø- 5)→C (3-20)  E (1-20)
                                        V (3- 8)  Q (1- 8)
                                        I (3- 5)  B (1- 5)
```

FIGURE 36.

(3) These pairs manifestly all belong to the same displacement interval, and therefore unions can be made immediately. The complete list is as follows:

E X, Q L, N I, L H, H O, B C, O Z, C E, T P, P V, X T, V Q, I B

(4) Joining pairs by their common letters, the following sequence is obtained:

. . . N I B C E X T P V Q L H O Z . . .

e. With this as a nucleus the cryptogram can be solved speedily and accurately. When it is realized that the cryptanalyst can assume THE's rather readily in some cases, the value of this principle becomes apparent. When it is further realized that if a cryptogram has sufficient text to enable the THE's to be found easily, it is usually also not at all difficult to make correct assumptions of values for two or three other high-frequency letters, it is clear that the principles of indirect symmetry of position may often be used with gratifyingly quick success to reconstruct the complete primary component.

f. When the probable-word method is combined with the principles of indirect symmetry the solution of a difficult case is often accomplished with astonishing ease and rapidity.

strip Fig. 6

Over

SECTION IX

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, III

37. Solution of subsequent messages enciphered by the same primary components.—*a.* In the discussion of the methods of solving repeating-key ciphers using secondary alphabets derived from the sliding of a mixed component against the normal component (Section V), it was shown how subsequent messages enciphered by the same pair of primary components but with different keys could be solved by application of principles involving the completion of the plain-component sequence (paragraphs 23, 24). The present paragraph deals with the application of these same principles to the case where the primary components are identical mixed sequences.

b. Suppose that the following primary component has been reconstructed from the analysis of a lengthy cryptogram:

Q U E S T I O N A B L Y C D F G H J K M P R V W X Z

A new message exchanged between the same correspondents is intercepted and is suspected of having been enciphered by the same primary components but with a different key. The message is as follows:

```
N F W W P   N O M K I   W P I D S   C A A E T   Q V Z S E
Y O J S C   A A A F G   R V N H D   W D S C A   E G N F P
F O E M T   H X L J W   P N O M K   I Q D B J   I V N H L
T F N C S   B G C R P
```

c. Factoring discloses that the period is 7 letters. The text is transcribed accordingly, and is as follows:

```
N F W W P N O
M K I W P I D
S C A A E T Q
V Z S E Y O J
S C A A A F G
R V N H D W D
S C A E G N F
P F O E M T H
X L J W P N O
M K I Q D B J
I V N H L T F
N C S B G C R
P
```

FIGURE 37.

d. The letters belonging to the same alphabet are then employed as the initial letters of completion sequences, in the manner shown in paragraph 23*e*, using the already reconstructed primary component. The completion diagrams for the first five letters of the first three alphabets are as follows:

ALPHABET 1

N	M	S	V	S
A	P	T	W	T
B	R	I	X	I
L	V	O	Z	O
Y	W	N	Q	N
C	X	A	U	A
D	Z	B	E	B
F	Q	L	S	L
G	U	Y	T	Y
*H	E	C	I	C
J	S	D	O	D
K	T	F	N	F
M	I	G	A	G
P	O	H	B	H
R	N	J	L	J
V	A	K	Y	K
W	B	M	C	M
X	L	P	D	P
Z	Y	R	F	R
Q	C	V	G	V
U	D	W	H	W
E	F	X	J	X
S	G	Z	K	Z
T	H	Q	M	Q
I	J	U	P	U
O	K	E	R	E

ALPHABET 2

F	K	C	Z	C
G	M	D	Q	D
H	P	F	U	F
J	R	G	E	G
K	V	H	S	H
M	W	J	T	J
P	X	K	I	K
R	Z	M	O	M
V	Q	P	N	P
W	U	R	A	R
X	E	V	B	V
Z	S	W	L	W
Q	T	X	Y	X
U	I	Z	C	Z
E	O	Q	D	Q
S	N	U	F	U
T	A	E	G	E
I	B	S	H	S
O	L	T	J	T
N	Y	I	K	I
*A	C	O	M	O
B	D	N	P	N
L	F	A	R	A
Y	G	B	V	B
C	H	L	W	L
D	J	Y	X	Y

ALPHABET 3

W	I	A	S	A
X	O	B	T	B
Z	N	L	I	L
Q	A	Y	O	Y
U	B	C	N	C
E	L	D	A	D
S	Y	F	B	F
T	C	G	L	G
I	D	H	Y	H
O	F	J	C	J
N	G	K	D	K
A	H	M	F	M
B	J	P	G	P
L	K	R	H	R
Y	M	V	S	V
C	P	W	K	W
D	R	X	M	X
F	V	Z	P	Z
G	W	Q	R	Q
H	X	U	V	U
J	Z	E	W	E
K	Q	S	X	X
M	U	T	Z	T
P	E	I	Q	I
R	S	O	U	O
*V	T	N	E	N

FIGURE 38.

set

e. Examining the successive generat~~ives~~ to select the ones showing the best assortment of high-frequency letters, those marked in Figure 38 by asterisks are chosen. These are then assembled in columnar fashion and yield the following plain text: rices

1	2	3	4	5	6	7
H	A	V				
E	C	T				
C	O	N				
I	M	E				
C	O	N				

FIGURE 39.

Over

f. The corresponding key-letters are sought, using enciphering equations $\Theta_{k/c}=\Theta_{1/p}$; $\Theta_{p/p}=\Theta_{c/c}$, and are found to be JOU, which suggests the keyword JOURNEY. Testing the key-letters RNEY for alphabets 4, 5, 6, and 7, the following results are obtained:

1	2	3	4	5	6	7
J	O	U	R	N	E	Y
N	F	W	W	P	N	O
H	A	V	E	D	I	R
M	K	I	W	P	I	D
E	C	T	E	D	S	E

FIGURE 40.

The message may now be completed with ease. It is as follows:

J	O	U	R	N	E	Y		J	O	U	R	N	E	Y
H	A	V	E	D	I	R		S	A	I	N	C	E	I
N	F	W	W	P	N	O		P	F	O	E	M	T	H
E	C	T	E	D	S	E		N	T	H	E	D	I	R
M	K	I	W	P	I	D		X	L	J	W	P	N	O
C	O	N	D	R	E	G		E	C	T	I	O	N	O
S	C	A	A	E	T	Q		M	K	I	Q	D	B	J
I	M	E	N	T	T	O		F	H	O	R	S	E	S
V	Z	S	E	Y	O	J		I	V	N	H	L	T	F
C	O	N	D	U	C	T		H	O	E	F	A	L	L
S	C	A	A	A	F	G		N	C	S	B	G	C	R
T	H	O	R	O	R	E		S						
R	V	N	H	D	W	D		P						
C	O	N	N	A	I	S								
S	C	A	E	G	N	F								

FIGURE 41.

38. Solution of repeating-key ciphers in which the identical mixed components proceed in opposite directions.—The secondary alphabets in this case (paragraph 6, Case B (3) (a) (II) are reciprocal. The steps in solution are essentially the same as in the preceding case (paragraph 28); the principles of indirect symmetry of position can also be applied with the necessary modifications introduced by virtue of the reciprocity existing within the respective secondary alphabets (paragraph 31*p*).

39. Solution of repeating-key ciphers in which the primary components are different mixed sequences.—This is Case B (3) (b) of paragraph 6. The steps in solution are essentially the same as in paragraphs 28 and 31, except that in applying the principles of indirect symmetry of position it is necessary to take cognizance of the fact that the primary components are different mixed sequences (paragraph 31*q*).

40. Solution of subsequent messages after the primary components have been recovered.—*a*. In the case in which the primary components are identical mixed sequences proceeding in opposite directions, as well as in that in which the primary components are different mixed sequences, the solution of subsequent messages[1] is a relatively easy matter. In both cases, however, the student must remember that before the method illustrated in paragraph 37 can be applied it is necessary to convert the cipher letters into their plain-component equivalents before completing the plain-component sequence. From there on, the process of selecting and assembling the proper generatrices is the same as usual.

b. Perhaps an example may be advisable. Suppose the enemy has been found to be using primary components based upon the keyword QUESTIONABLY, the plain component running from left to right, the cipher component in the reverse direction. The following new message has arrived from the intercept station:

M V X O X B Z I Y Z N L W Z H O X I E O O O E P Z
F X S R X E J B S H B O N A U R A P Z I N R A M V
X O X A I J Y X W F K N D O W J E R C U R A L V B
Z A Q U W J W X Y I D G R K D Q B D R M Q E C Y V
Q W

c. Factoring discloses that the period is 6 and the message is accordingly transcribed into 6 columns, Fig. 42. The letters of these columns are then converted into their plain component equivalents by juxtaposing the two primary components at any point of coincidence, for example $Q_p = Z_c$. The converted letters are shown in Fig. 43. The letters of the individual columns are then used as the initial letters of completion sequences, using the QUESTIONABLY primary sequence. The final step is the selection and assembling of the selected generatrices. The results for the first ten letters of the first three columns are shown below:

1	2	3	4	5	6
M	V	X	O	X	B
Z	I	Y	Z	N	L
W	Z	H	O	X	I
E	O	O	O	E	P
Z	F	X	S	R	X
E	J	B	S	H	B
O	N	A	U	R	A
P	Z	I	N	R	A
M	V	X	O	X	A
I	J	Y	X	W	F
K	N	D	O	W	J
E	R	C	U	R	A
L	V	B	Z	A	Q
U	W	J	W	X	Y
I	D	G	R	K	D
Q	B	D	R	M	Q
E	C	Y	V	Q	W

FIGURE 42.

1	2	3	4	5	6
O	S	U	M	U	H
Q	P	F	Q	K	G
E	Q	B	M	U	P
W	M	M	M	W	I
Q	Y	U	V	T	U
W	A	H	V	B	H
M	K	J	X	T	J
I	Q	P	K	T	J
O	S	U	M	U	J
P	A	F	U	E	Y
N	K	C	M	E	A
W	T	D	X	T	J
G	S	H	Q	J	Z
X	E	A	E	U	F
P	C	L	T	N	C
Z	H	C	T	O	Z
W	D	F	S	Z	E

FIGURE 43.

[1] That is, messages intercepted after the primary components have been reconstructed and enciphered by keys different from those used in the messages upon which the reconstruction of the primary components was accomplished.

Over

COLUMN 1

O	Q	E	W	Q	W	M	I	O	P
N	U	S	X	U	X	P	O	N	R
A	E	T	Z	E	Z	R	N	A	V
B	S	I	Q	S	Q	V	A	B	W
L	T	O	U	T	U	W	B	L	X
Y	I	N	E	I	E	X	L	Y	Z
C	O	A	S	O	S	Z	Y	C	Q
D	N	B	T	N	T	Q	C	D	U
*F	A	L	I	A	I	U	D	F	E
G	B	Y	O	B	O	E	F	G	S
H	L	C	N	L	N	S	G	H	T
J	Y	D	A	Y	A	T	H	J	I
K	C	F	B	C	B	I	J	K	O
M	D	G	L	D	L	O	K	M	N
P	F	H	Y	F	Y	N	M	P	A
R	G	J	C	G	C	A	P	R	B
V	H	K	D	H	D	B	R	V	L
W	J	M	F	J	F	L	V	W	Y
X	K	P	G	K	G	Y	W	X	C
Z	M	R	H	M	H	C	X	Z	D
Q	P	V	J	P	J	D	Z	Q	F
U	R	W	K	R	K	F	Q	U	G
E	V	X	M	V	M	G	U	E	H
S	W	Z	P	W	P	H	E	S	J
T	X	Q	R	S	R	J	S	T	K
I	Z	U	V	Z	V	K	T	I	M

COLUMN 2

S	P	Q	M	Y	A	K	Q	S	A
T	R	U	P	C	B	M	U	T	B
*I	V	E	R	D	L	P	E	I	L
O	W	S	V	F	Y	R	S	O	Y
N	X	T	W	G	C	V	T	N	C
A	Z	I	X	H	D	W	I	A	D
B	Q	O	Z	J	F	X	O	B	F
L	U	N	Q	K	G	Z	N	L	G
Y	E	A	U	M	H	Q	A	Y	H
C	S	B	E	P	J	U	B	C	J
D	T	L	S	R	K	E	L	D	K
F	I	Y	T	V	M	S	Y	F	M
G	O	C	I	W	P	T	C	G	P
H	N	D	O	X	R	I	D	H	R
J	A	F	N	Z	V	O	F	J	V
K	B	G	A	Q	W	N	G	K	W
M	L	H	B	U	X	A	H	M	X
P	Y	J	L	E	Z	B	J	P	Z
R	C	K	Y	S	Q	L	K	R	Q
V	D	M	C	T	U	Y	M	V	U
W	F	P	D	I	E	C	P	W	E
X	G	R	F	O	S	D	R	X	S
Z	H	V	G	N	T	F	V	Z	T
Q	J	W	H	A	I	G	W	O	I
U	K	X	J	B	O	H	X	U	O
E	M	Z	K	L	N	J	Z	E	N

COLUMN 3

U	F	B	M	U	H	J	P	U	F
E	G	L	P	E	J	K	R	E	G
S	H	Y	R	S	K	M	V	S	H
T	J	C	V	T	M	P	W	T	J
I	K	D	W	I	P	R	X	I	K
O	M	F	X	O	R	V	Z	O	M
N	P	G	Z	N	V	W	Q	N	P
A	R	H	Q	A	W	X	U	A	R
B	V	J	U	B	X	Z	E	B	V
L	W	K	E	L	Z	Q	S	L	W
Y	X	M	S	Y	Q	U	T	Y	X
C	Z	P	T	C	U	E	I	C	Z
D	Q	R	I	D	E	S	O	D	Q
F	U	V	O	F	S	T	N	F	U
G	E	W	N	G	T	I	A	G	E
H	S	X	A	H	I	O	B	H	S
J	T	Z	B	J	O	N	L	J	T
K	I	Q	L	K	N	A	Y	K	I
M	O	U	Y	M	A	B	C	M	O
P	N	E	C	P	B	L	D	P	N
*R	A	S	D	R	L	Y	F	R	A
V	B	T	F	V	Y	C	G	V	B
W	L	I	G	W	C	D	H	W	L
X	Y	O	H	X	D	F	J	X	Y
Z	C	N	J	Z	F	G	K	Z	C
Q	D	A	K	Q	G	H	N	Q	D

FIGURE 44.

Columnar assembling of selected generatrices gives what is shown in Fig. 45.

1	2	3	4	5	6
F	I	R	.	.	.
A	V	A	.	.	.
L	E	S	.	.	.
I	R	D	.	.	.
A	D	R	.	.	.
I	L	L	.	.	.
U	P	Y	.	.	.
D	E	F	.	.	.
F	I	R	.	.	.
E	L	A	.	.	.

FIGURE 45.

d. The key letters are sought, and found to be NUM, which suggests NUMBER. The entire message may now be read with ease. It is as follows:

N	U	M	B	E	R
F	I	R	S	T	C
M	V	X	O	X	B
A	V	A	L	R	Y
Z	I	Y	Z	N	L
L	E	S	S	T	H
W	Z	H	O	X	I
I	R	D	S	Q	U
E	O	O	O	E	P
A	D	R	O	N	W
Z	F	X	S	R	X
I	L	L	O	C	C
E	J	B	S	H	B
U	P	Y	A	N	D
O	N	A	U	R	A
D	E	F	E	N	D
P	Z	I	N	R	A
F	I	R	S	T	D
M	V	X	O	X	A

N	U	M	B	E	R
E	L	A	Y	I	N
I	J	Y	X	W	F
G	P	O	S	I	T
K	N	D	O	W	J
I	O	N	A	N	D
E	R	C	U	R	A
W	I	L	L	P	R
L	V	B	Z	A	Q
O	T	E	C	T	L
U	W	J	W	X	Y
E	F	T	F	L	A
I	D	G	R	K	D
N	K	O	F	B	R
Q	B	D	R	M	Q
I	G	A	D	E	X
E	C	Y	V	Q	W

FIGURE 46.

e. If the primary components are different mixed sequences, the procedure is identical with that just indicated. The important point to note is that one must not fail to convert the letters into their plain-component equivalents before the completion-sequence method is applied.

Over

Section X

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, IV

41. General remarks.—The preceding three sections have been devoted to an elucidation of the general principles and procedure in the solution of typical cases of repeating-key ciphers. This section will be devoted to a consideration of the variations in cryptanalytic procedure arising from special circumstances. It may be well to add that by the designation "special circumstances" it is not meant to imply that the latter are necessarily *unusual* circumstances. *The student should always be on the alert to seize upon any opportunities that may appear in which he may apply the methods to be described.* In practical work such opportunities are by no means rare and are seldom overlooked by competent cryptanalysts.

42. Deriving the secondary alphabets, the primary components, and the key, given a cryptogram with its plain text.—*a.* It may happen that a cryptogram and its equivalent plain text are at hand, as the result of capture, pilferage, compromise, etc. This, as a general rule, affords a very easy attack upon the whole system.

b. Taking first the case where the plain component is the normal alphabet, the cipher component a mixed sequence, the first thing to do is to write out the cipher text with its letter-for-letter decipherment. From this, by a slight modification of the principles of "factoring", one discovers the length of the key. It is obvious that when a word of three or four letters is enciphered by the same cipher text, the interval between the two occurrences is almost certainly a multiple of the length of the key. By noting a few recurrences of plain text and cipher letters, one can quickly determine the length of the key (assuming of course that the message is long enough to afford sufficient data). Having determined the length of the key, the message is rewritten according to its periods, with the plain text likewise in periods under the cipher letters. From this arrangement one can now reconstruct complete or partial secondary alphabets. If the secondary alphabets are complete, they will show direct symmetry of position; if they are but fragmentary in several alphabets, then the primary component can be reconstructed by the application of the principles of direct symmetry of position.

c. If the plain component is a mixed sequence, and the cipher component the normal (direct or reversed sequence), the secondary alphabets will show no direct symmetry unless they are arranged in the form of deciphering alphabets (that is, $A_c \ldots Z_c$ above the zero line, with their equivalents below). The student should be on the lookout for such cases.

d. (1) If the plain and cipher primary components are identical mixed sequences proceeding in the same direction, the secondary alphabets will show indirect symmetry of position, and they can be used for the speedy reconstruction of the primary components (Paragraph 31*a* to *o*).

(2) If the plain and the cipher primary components are identical mixed sequences proceeding in opposite directions, the secondary alphabets will be completely reciprocal secondary alphabets and the primary component may be reconstructed by applying the principles outlined in paragraph 31*p*.

(3) If the plain and the cipher primary components are different mixed sequences, the secondary alphabets will show indirect symmetry of position and the primary components may be reconstructed by applying the principles outlined in paragraph 31*q*.

e. In all the foregoing cases, after the primary components have been reconstructed, the keys can be readily recovered.

43. Deriving the secondary alphabets, the primary components, and the keywords for messages, given two or more cryptograms in different keys and suspected to contain identical plain text.—*a*. The simplest case of this kind is that involving two monoalphabetic substitution ciphers with mixed alphabets derived from the same pair of sliding components. An understanding of this case is necessary to that of the case involving repeating-key ciphers.

b. (1) A message is transmitted from station A to station B. B then sends A some operating signals which indicate that B cannot decipher the message, and soon thereafter A sends a second message, identical in length with the first. This leads to the suspicion that the plain text of both messages is the same. The intercepted messages are superimposed. Thus:

1. NXGRV MPUOF ZQVCP VWERX QDZVX WXZQE TBDSP VVXJK RFZWH ZUWLU IYVZQ FXOAR
2. EMLHJ FGVUB PRJNG JKWHM RAPJM KMPRW ZTAXG JJMCD HBPKY PVKIV QOJPR BMUSH

(2) Initiating a chain of cipher-text equivalents from message 1 to message 2, the following complete sequence is obtained:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
N	E	W	K	D	A	S	X	M	F	B	T	Z	P	G	L	I	Q	R	H	Y	O	U	V	J	C

(3) Experimentation along already-indicated lines soon discloses the fact that the foregoing component is an equivalent primary component of the original primary based upon the keyword QUESTIONABLY, decimated on the 21st interval. Let the student decipher the cryptogram.

(4) The foregoing example is somewhat artificial in that the plain text was consciously selected with a view to making it contain every letter of the alphabet. The purpose in doing this was to permit the construction of a complete chain of equivalents from only two short messages, in order to give a simple illustration of the principles involved. If the plain-text message does not contain every letter of the alphabet, then only partial chains of equivalents can be constructed. These may be united, if circumstances will permit, by recourse to the various principles elucidated in paragraph 31.

(5) The student should carefully study the foregoing example in order to obtain a thorough comprehension of the *reason* why it was possible to reconstruct the primary component from the two cipher messages without having any plain text to begin with at all. Since the plain text of both messages is the same, the relative displacement of the primary components in the case of message 1 differs from the relative displacement of the same primary components in the case of message 2 by a *fixed* interval. Therefore, the distance between N and E (the first letters of the two messages), on the primary component, regardless of what plain-text letter these two cipher letters represent, is the same as the distance between E and W (the 18th letters), W and K (the 17th letters), and so on. Thus, this fixed interval permits of establishing a complete chain of letters separated by constant intervals and this chain becomes an equivalent primary component.

Over

44. The case of repeating-key systems.—*a.* With the foregoing basic principles in mind the student is ready to note the procedure in the case of two repeating-key ciphers having identical plain texts. First, the case in which both messages have keywords of identical length but different compositions will be studied.

b. (1) Given the following two cryptograms suspected to contain the same plain text:

Message 1

```
YHYEX  UBUKA  PVLLT  ABUVV  DYSAB
PCQTU  NGKFA  ZEFIZ  BDJEZ  ALVID
TROQS  UHAFK
```

Message 2

```
CGSLZ  QUBMN  CTYBV  HLQFT  FLRHL
MTAIQ  ZWMDQ  NSDWN  LCBLQ  NETOC
VSNZR  BJNOQ
```

(2) The first step is to try to determine the length of the period. The usual method of factoring cannot be employed because there are no long repetitions and not enough repetitions even of digraphs to give any convincing indications. However, a subterfuge will be employed, based upon the theory of factoring.

c. (1) Let the two messages be superimposed.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1.	Y	H	Y	E	X	U	B	U	K	A	P	V	L	L	T	A	B	U	V	V
2.	C	G	S	L	Z	Q	U	B	M	N	C	T	Y	B	V	H	L	Q	F	T

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1.	D	Y	S	A	B	P	C	Q	T	U	N	G	K	F	A	Z	E	F	I	Z
2.	F	L	R	H	L	M	T	A	I	Q	Z	W	M	D	Q	N	S	D	W	N

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1.	B	D	J	E	Z	A	L	V	I	D	T	R	O	Q	S	U	H	A	F	K
2.	L	C	B	L	Q	N	E	T	O	C	V	S	N	Z	R	B	J	N	O	Q

(2) Now let a search be made of cases of identical superimposition. For example, $\overset{4}{\overset{E}{L}}$ and $\overset{44}{\overset{E}{L}}$ are separated by 40 letters, $\overset{6}{\overset{U}{Q}}$, $\overset{18}{\overset{U}{Q}}$, and $\overset{30}{\overset{U}{Q}}$ are separated by 12 letters. Let these intervals between identical superimpositions be factored, just as though they were ordinary repetitions. That factor which is the most frequent should correspond with the length of the period for the following reason. If the period is the same and the plain text is the same in both messages, then the condition of identity of superimposition can only be the result of identity of encipherments by identical cipher alphabets. This is only another way of saying that the same relative position in the keying cycle has been reached in both cases of identity. Therefore, the distance between identical superimpositions must be either equal to or else a multiple of the length of the period. Hence, factoring the intervals must yield the length of the period. The complete list of intervals

10 pt leaded 36½ pi. - F.i.c.

reset

and factors applicable to cases of identical superimposed pairs is as follows (factors above 12 are omitted):

Repetition	Interval	Factors	Repetition	Interval	Factors
1st EL to 2d EL	40	2, 4, 5, 8, 10.	1st TV to 2d TV	36	2, 3, 4, 6, 9, 12.
1st UQ to 2d UQ	12	2, 3, 4, 6, 12.	1st AH to 2d AH	8	2, 4, 8.
2d UQ to 3d UQ	12	2, 3, 4, 6, 12.	1st BL to 2d BL	8	2, 4, 8.
1st UB to 2d UB	48	2, 3, 4, 6, 8, 12.	2d BL to 3d BL	16	2, 4, 8.
1st KM to 2d KM	24	2, 3, 4, 6, 8, 12.	1st SR to 2d SR	32	2, 4, 8.
1st AN to 2d AN	36	2, 3, 4, 6, 9, 12.	1st FD to 2d FD	4	2, 4.
2d AN to 3d AN	12	2, 3, 4, 6, 12.	1st ZN to 2d ZN	4	2, 4.
1st VT to 2d VT	8	2, 4, 8.	1st DC to 2d DC	8	2, 4, 8.
2d VT to 3d VT	28	2, 4, 7.			

set

(3) The ~~factor 4 is the only one~~ factors 4 and 2 are the only ones common to every one of these intervals and it may be taken as beyond question that the length of the period is 4. Since a period of 2 is not very probable!

d. Let the messages now be superimposed according to their periods:

```
     1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4
1.   Y H Y E   X U B U   K A P V   L L T A   B U V V   D Y S A   B P C Q
2.   C G S L   Z Q U B   M N C T   Y B V H   L Q F T   F L R H   L M T A

1.   T U N G   K F A Z   E F I Z   B D J E   Z A L V   I D T R   O Q S U
2.   I Q Z W   M D Q N   S D W N   L C B L   Q N E T   O C V S   N Z R B

1.   H A F K
2.   J N O Q
```

e. (1) Now distribute the superimposed letters into a reconstruction skeleton of "secondary alphabets."
Thus:

ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1		L		F	S			J	O		M	Y			N					I				Z	C	Q
2	N			C		D		G				B				M	Z				Q				L	
3	Q	U	T			O			W	B		E		Z		C			R	V		F			S	
4	H				L		W				Q						A	S			B	T				N

(2) By the usual methods, construct the primary or an equivalent primary component. Taking lines ø and 1, the following sequences are noted:

BL, DF, ES, HJ, IO, KM, LY, ON, TI, XZ, YC, ZQ.

which, when united by means of common letters and study of other sequences, yield the complete original primary component based upon the keyword QUESTIONABLY:

Q U E S T I O N A B L Y C D F G H J K M P R V W X Z

set

(3) The fact that the pair of lines with which the process was commenced yield the original primary sequence is purely accidental; it might have just as well yielded an equivalent primary sequence.

Over

f. (1) Having the primary component, the solution of the messages is now a relatively simple matter. An application of the method elucidated in paragraph 37 is made, involving the completion of the plain-component sequence for each alphabet and selecting those generatrices which contain the best assortments of high-frequency letters. Thus, using Message 1:

FIRST ALPHABET	SECOND ALPHABET	THIRD ALPHABET	FOURTH ALPHABET
Y X K L B	H U A L U	Y B P T V	E U V A V
C Z M Y L	J E B Y E	C L R I W	S E W B W
D Q P C Y	K S L C S	D Y V O X	T S X L X
F U R D C	M T Y D T	F C W N Z	I T Z Y Z
G E V F D	P I C F I	G D X A Q	O I Q C Q
H S W G F	R O D G O	H F Z B U	N O U D U
J T X H G	V N F H N	J G Q L E	*A N E F E
K I Z J H	W A G J A	K H U Y S	B A S G S
M O Q K J	X B H K B	M J E C T	L B T H T
P N U M K	Z L J M L	P K S D I	Y L I J I
R A E P M	Q Y K P Y	R M T F O	C Y O K O
V B S R P	U C M R C	V P I G N	D C N M N
W L T V R	E D P V D	W R O H A	F D A P A
X Y I W V	S F R W F	X V N J B	G F B R B
Z C O X W	T G V X G	Z W A K L	H G L V L
Q D N Z X	I H W Z H	Q X B M Y	J H Y W Y
U F A Q Z	O J X Q J	U Z L P C	K J C X C
E G B U Q	N K Z U K	E Q Y R D	M K D Z D
S H L E U	A M Q E M	S U C V F	P M F Q F
T J Y S E	B P U S P	T E D W G	R P G U G
I K C T S	*L R E T R	I S F X H	V R H E H
O M D I T	Y V S I V	O T G Z J	W V J S J
N P F O I	C W T O W	N I H Q K	X W K T K
*A R G N O	D X I N X	A O J U M	Z X M I M
B V H A N	F Z O A Z	B N K E P	Q Z P O P
L W J B A	G Q N B Q	*L A M S R	U Q R N R

FIGURE 48.

(2) The selected generatrices (those marked by asterisks in Fig. 48) are assembled in columnar manner:

```
A L L A
R R A N
G E M E
N T S F
O R R E
```

FIGURE 49.

(3) The key letters are sought and give the keyword SOUP. The plain text for the second message is now known, and by reference to the cipher text and the primary components, the keyword for this message is found to be TIME. The complete texts are as follows:

S	O	U	P		T	I	M	E
A	L	L	A		A	L	L	A
Y	H	Y	E		C	G	S	L
R	R	A	N		R	R	A	N
X	U	B	U		Z	Q	U	B
G	E	M	E		G	E	M	E
K	A	P	V		M	N	C	T
N	T	S	F		N	T	S	F
L	L	T	A		Y	B	V	H
O	R	R	E		O	R	R	E
B	U	V	V		L	Q	F	T
L	I	E	F		L	I	E	F
D	Y	S	A		F	L	R	H
O	F	Y	O		O	F	Y	O
B	P	C	Q		L	M	T	A
U	R	O	R		U	R	O	R
T	U	N	G		I	Q	Z	W
G	A	N	I		G	A	N	I
K	F	A	Z		M	D	Q	N
Z	A	T	I		Z	A	T	I
E	F	I	Z		S	D	W	N
O	N	H	A		O	N	H	A
B	D	J	E		L	C	B	L
V	E	B	E		V	E	B	E
Z	A	L	V		Q	N	E	T
E	N	S	U		E	N	S	U
I	D	T	R		O	C	V	S
S	P	E	N		S	P	E	N
O	Q	S	U		N	Z	R	B
D	E	D	X		D	E	D	X
H	A	F	K		J	N	O	Q

FIGURE 50.

45. The case of identical messages enciphered by keywords of different lengths.—*a.* In the foregoing case the keywords for the two messages, although different, were identical in length. When this is not true and the keywords are of different lengths, the procedure need be only slightly modified.

Over

b. Given the following two cryptograms suspected of containing the same plain-text enciphered by the same primary components but with different keywords of different lengths, solve the messages.

MESSAGE No. 1

```
V M Y Z G   E A U N T   P K F A Y   J I Z M B   U M Y K B   V F I V V
S E O A F   S K X K R   Y W C A C   Z O R D O   Z R D E F   B L K F E
S M K S F   A F E K V   Q U R C M   Y Z V O X   V A B T A   Y Y U O A
Y T D K F   E N W N T   D B Q K U   L A J L Z   I O U M A   B O A F S
K X Q P U   Y M J P W   Q T D B T   O S I Y S   M I Y K U   R O G M W
C T M Z Z   V M V A J
```

MESSAGE No. 2

```
Z G A M W   I O M O A   C O D H A   C L R L P   M O Q O J   E M O Q U
D H X B Y   U Q M G A   U V G L Q   D B S P U   O A B I R   P W X Y M
O G G F T   M R H V F   G W K N I   V A U P F   A B R V I   L A Q E M
Z D J X Y   M E D D Y   B O S V M   P N L G X   X D Y D O   P X B Y U
Q M N K Y   F L U Y Y   G V P V R   D N C Z E   K J Q O R   W J X R V
G D K D S   X C E E C
```

c. The messages are long enough to show a few short repetitions which permit factoring. The latter discloses that Message 1 has a period of 4 and Message 2, a period of 6 letters. The messages are superimposed, with numbers marking the position of each letter in the corresponding period, as shown below:

```
        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  V M Y Z G E A U N T P K F A Y J I Z M B U M Y K
No. 2.  Z G A M W I O M O A C O D H A C L R L P M O Q O
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  B V F I V V S E O A F S K X K R Y W C A C Z O R
No. 2.  J E M O Q U D H X B Y U Q M G A U V G L Q D B S
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  D O Z R D E F B L K F E S M K S F A F E K V Q U
No. 2.  P U O A B I R P W X Y M O G G F T M R H V F G W
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  R C M Y Z V O X V A B T A Y Y U O A Y T D K F E
No. 2.  K N I V A U P F A B R V I L A Q E M Z D J X Y M
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  N W N T D B Q K U L A J L Z I O U M A B O A F S
No. 2.  E D D Y B O S V M P N L G X X D Y D O P X B Y U
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  K X Q P U Y M J P W Q T D B T O S I Y S M I Y K
No. 2.  Q M N K Y F L U Y Y G V P V R D N C Z E K J Q O
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

        1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No. 1.  U R O G M W C T M Z Z V M V A J
No. 2.  R W J X R V G D K D S X C E E C
        1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4
```

d. A reconstruction skeleton of "secondary alphabets" is now made by distributing the letters in respective lines corresponding to the 12 different superimposed pairs of numbers. For example, all pairs corresponding to the superimposition of position 1 of Message 1 with position 1 of Message 2 are distributed in lines Ø and 1 of the skeleton. Thus, the very first superimposed pair is $\left\{\begin{matrix}1\\V\\Z\\1\end{matrix}\right.$; the letter Z is inserted in line 1 under the letter V. The next $\left\{\begin{matrix}1\\1\end{matrix}\right.$ pair is the 13th superimposition, with $\left\{\begin{matrix}F\\D\end{matrix}\right.$; the letter D is inserted in line 1 under the letter F, and so on. The skeleton is then as follows:

Superimposed pairs. Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1–1	I	J		P		D					Q	G	C	E				K	O		R	Z				
2–2	H	V	N										G		U			W				E	D	M	L	X
3–3	E					M			X		G		I	D	J		N			R					A	O
4–4							X		O	C					D	K		A	F	Y	Q				V	M
1–5				B		T	W		L				R		E				N		Y	Q			U	A
2–6	M	O			I				C				D									U	V		F	R
3–1	O		G			R							L		P		S		D						Z	
4–2	L	P			H					U	V								E	D	M			F		
1–3			Q	J							V	W	K	O	X	Y					M	A				
2–4	B								J		X	P	O							A		F	Y			D
3–5	N	R				Y									B	C	G								Q	S
4–6					M					L	O							S	U	V	W	X				

FIGURE 51.

e. There are more than sufficient data here to permit of the reconstruction of a complete equivalent primary component, for example, the following:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
I	T	K	N	P	Z	H	M	W	B	Q	E	U	L	F	C	S	J	A	X	R	G	D	V	O	Y

f. The subsequent steps in the actual decipherment of the text of either of the two messages are of considerable interest. Thus far the cryptanalyst has only the cipher component of the primary sliding components. The plain component may be identical with the cipher component and may progress in the same direction, or in the reverse direction; or, the two components may be different. If different, the plain component may be the normal sequence, direct or reversed. Tests must be made to ascertain which of these various possibilities is true.

g. (1) It will first be assumed that the primary plain component is the normal direct sequence. Applying the procedure outlined in Par. 23 to the message with the shorter key (Message No. 1, to give the most data per secondary alphabet), an attempt is made to solve the message. It is unnecessary here to go further into detail in this procedure; suffice it to indicate that the attempt is unsuccessful and it follows that the plain component is not the normal direct sequence. A normal reversed sequence is then assumed for the plain component and the proper procedure applied. Again the attempt is found useless. Next, it is assumed that the plain component is identical with the cipher component, and the procedure outlined in Par. 37 is tried. This also is unsuccessful. Another attempt, assuming the plain component runs in the reverse direction, is likewise unsuccessful. There remains one last hypothesis, viz, that the two primary components are different mixed sequences.

Over

(2) Here is Message No. 1 transcribed in periods of four letters. Uniliteral frequency distributions for the four secondary alphabets are shown below in Fig. 52, labeled 1*a*, 2*a*, 3*a*, and 4*a*. These distributions are based upon the normal sequence A to Z. But since the reconstructed cipher component is at hand these distributions can be rearranged according to the sequence of the cipher component, as shown in distributions labeled 1*b*, 2*b*, 3*b*, and 4*b* in Fig. 52. *The latter distributions may be combined by shifting distributions 2b, 3b, and 4b to proper superimpositions with respect to 1b so as to yield a single monoalphabetic distribution for the entire message. In other words, the polyalphabetic message can be converted into monoalphabetic terms, thus very considerably simplifying the solution.*

MESSAGE NO. 1

V M Y Z	V A B T
G E A U	A Y Y U
N T P K	O A Y T
F A Y J	D K F E
I Z M B	N W N T
U M Y K	D B Q K
B V F I	U L A J
V V S E	L Z I O
O A F S	U M A B
K X K R	O A F S
Y W C A	K X Q P
C Z O R	U Y M J
D O Z R	P W Q T
D E F B	D B T O
L K F E	S I Y S
S M K S	M I Y K
F A F E	U R O G
K V Q U	M W C T
R C M Y	M Z Z V
Z V O X	M V A J

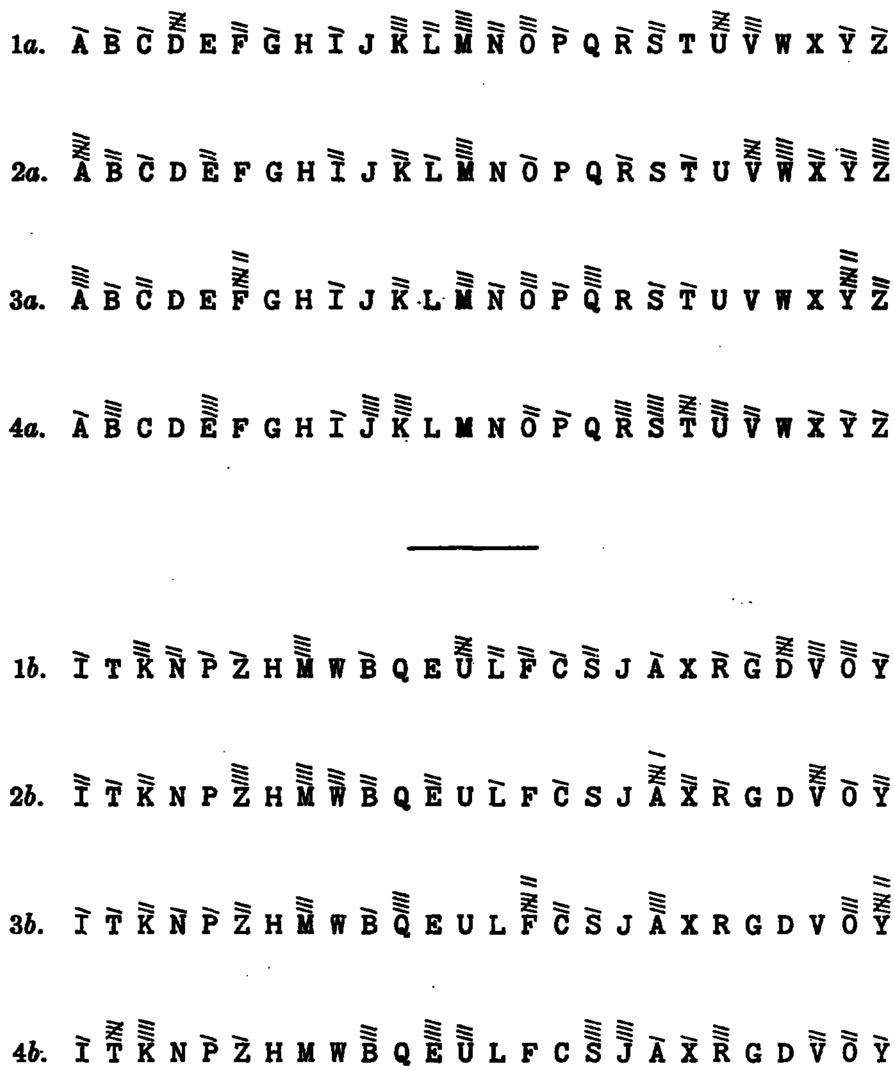


FIGURE 52.

(3) Note in Fig. 53 how the four distributions are shifted for superimposition and how the combined distribution presents the characteristics of a typical monoalphabetic distribution.

1*b*.	I T K N P Z H M W B Q E U L F C S J A X R G D V O Y
2*b*.	E U L F C S J A X R G D V O Y I T K N P Z H M W B Q
3*b*.	K N P Z H M W B Q E U L F C S J A X R G D V O Y I T
4*b*.	P Z H M W B Q E U L F C S J A X R G D V O Y I T K N
1*b*.–4*b*. combined	I T K N P Z H M W B Q E U L F C S J A X R G D V O Y

FIGURE 53.

(4) The letters belonging to alphabets 2, 3, and 4 of Fig. 52 may now be transcribed in terms of alphabet 1. That is, the two E's of alphabet 2 become I's; the L of alphabet 2 becomes a K; the C becomes a P, and so on. Likewise, the two K's of alphabet 3 become I's, the N becomes a T, and so on. The entire message is then a monoalphabet and can readily be solved. It is as follows:

V D V T G	I S W N S	K O F M V	L I R Z Z	U D V O B	U U D V U
E N E M Y	H A S C A	P T U R E	D H I L L	O N E T W	O O N E O
F M O M U	U K W I S	Y V L F C	R D S D L	N S D I U	Z L J U M
U R T R O	O P S H A	V E D U G	I N A N D	C A N H O	L D F O R
S D I U F	M U M K U	W W R P Z	G Z U D C	V M M V A	F V W O M
A N H O U	R O R P O	S S I B L	Y L O N G	E R R E Q	U E S T R
V V D J U	M N V T V	D O W O U	K S L L R	O R U D S	Z O M U U
E I N F O	R C E M E	N T S T O	P A D D I	T I O N A	L T R O O
K W W I U	F Z L P V	W V D O Y	R S C V U	M C V O U	B D J M V
P S S H O	U L D B E	S E N T V	I A G E O	R G E T O	W N F R E
L V M R N	X M U S L				
D E R I C	K R O A D				

(5) Having the plain text, the derivation of the ~~cipher~~ plain component (an equivalent) is an easy matter. It is merely necessary to base the reconstruction upon any of the secondary alphabets, since the plain text—cipher relationship is now known directly, and the primary cipher component is at hand. The primary plain component is found to be as follows:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
H	M	P	C	B	L	.	R	S	W	.	.	O	D	U	G	A	F	Q	K	I	Y	N	E	T	V

(6) The keywords for both messages can now be found, if desirable, by finding the equivalent of A_p in each of the secondary alphabets of the original polyalphabetic messages. The keyword for No. 1 is STAR; that for No. 2 is OCEANS.

~~288074 O—41——7~~

strip fig. 7

Over

(7) The student may, if he wishes, try to find out whether the primary components reconstructed above are the original components or are equivalent components, by examining all the possible decimations of the two components for evidences of derivation from keywords.

h. As already stated in Par. 26*l*, there are certain statistical and mathematical tests that can be employed in the process of "matching" distributions to ascertain proper superimpositions for monoalphabeticity. In the case just considered there were sufficient data in the distributions to permit the process to be applied successfully by eye, without necessitating statistical tests.

i. This case is an excellent illustration of the application of the process of *converting a polyalphabetic cipher into monoalphabetic terms*. Because it is a very valuable and important cryptanalytic "trick," the student should study it most carefully in order to gain a good understanding of the principle upon which it is based and its significance in cryptanalysis. The conversion in the case under discussion was possible because the sequence of letters forming the cipher component had been reconstructed and was known, and therefore the uniliteral distributions for the respective secondary cipher alphabets could theoretically be shifted to correct superimpositions for monoalphabeticity. It also happened that there were sufficient data in the distributions to give proper indications for their relative displacements. Therefore, the theoretical possibility in this case became an actuality. Without these two necessary conditions the superimposition and conversion cannot be accomplished. The student should always be on the lookout for situations in which this is possible.

46. Concluding remarks.—*a*. The observant student will have noted that a large part of this text is devoted to the elucidation and application of a very few basic principles. These principles are, however, extremely important and their proper usage in the hands of a skilled cryptanalyst makes them practically indispensable tools of his art. The student should therefore drill himself in the application of these tools by having someone make up problem after problem for him to practice upon, until he acquires facility in their use and feels competent to apply them in practice whenever the least opportunity presents itself. This will save him much time and effort in the solution of bona fide messages.

b. Continuing the analytical key introduced in Military Cryptanalysis Part I, the outline for the studies covered by Part II follows herewith.

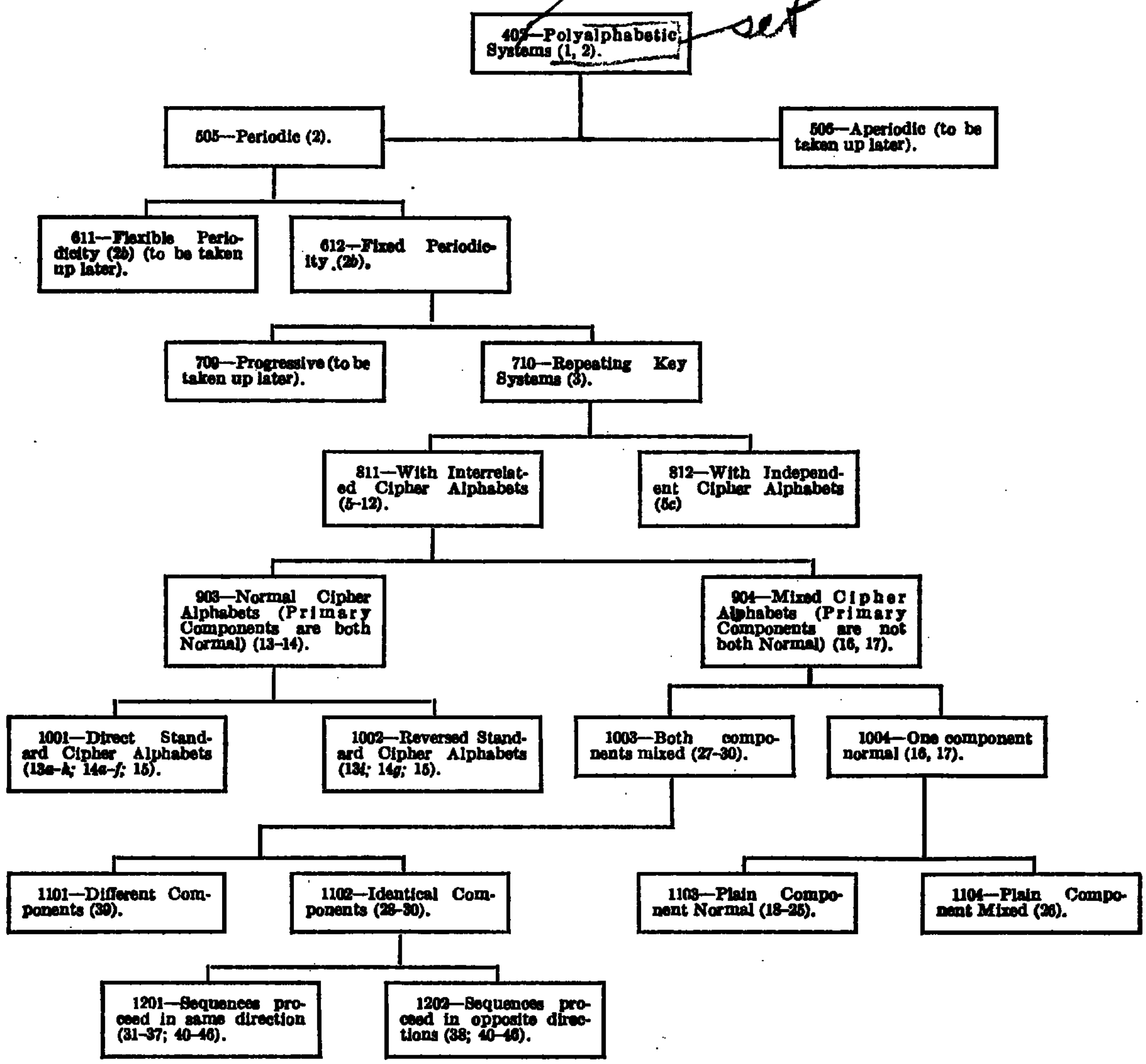
517974 100

4/6pt

Analytical Key for Military Cryptanalysis, Part II *

(Numbers in parentheses refer to Paragraph Numbers in this text)

set



*For explanation of the use of this chart see Par. 50 of Military Cryptanalysis, Part I.

Over

APPENDIX 1

The 12 Types of Cipher Squares

(See Paragraph 7)

Table I–B.[1]

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/2}=\theta_{1/1}$; $\theta_{p/1}=\theta_{c/2}$ ($\theta_{1/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V
B	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F
C	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R
D	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T
E	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S
F	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X
G	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I
H	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E
I	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z
J	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D
K	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M
L	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A
M	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U
N	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W
O	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N
P	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B
Q	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C
R	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y
S	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G
T	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H
U	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J
V	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K
W	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L
X	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O
Y	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P
Z	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q

[1] This table is labeled "Table 1–B" because it is the same as Table 1–A on page 7, except that the horizontal lines of the latter have been shifted so as to begin the successive alphabets with the successive letters of the normal sequence.

Table II

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\Theta_{k/2}=\Theta_{1/1}$; $\Theta_{p/2}=\Theta_{c/1}$ ($\Theta_{1/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	H	L	U	R	G	P	S	O	V	Y	B	X	D	E	I	M	K	Q	T	W	Z	C	F	J	N
B	T	A	E	N	K	Z	I	L	H	O	R	U	Q	W	X	B	F	D	J	M	P	S	V	Y	C	G
C	P	W	A	J	G	V	E	H	D	K	N	Q	M	S	T	X	B	Z	F	I	L	O	R	U	Y	C
D	G	N	R	A	X	M	V	Y	U	B	E	H	D	J	K	O	S	Q	W	Z	C	F	I	L	P	T
E	J	Q	U	D	A	P	Y	B	X	E	H	K	G	M	N	R	V	T	Z	C	F	I	L	O	S	W
F	U	B	F	O	L	A	J	M	I	P	S	V	R	X	Y	C	G	E	K	N	Q	T	W	Z	D	H
G	L	S	W	F	C	R	A	D	Z	G	J	M	I	O	P	T	X	V	B	E	H	K	N	Q	U	Y
H	I	P	T	C	Z	O	X	A	W	D	G	J	F	L	M	Q	U	S	Y	B	E	H	K	N	R	V
I	M	T	X	G	D	S	B	E	A	H	K	N	J	P	Q	U	Y	W	C	F	I	L	O	R	V	Z
J	F	M	Q	Z	W	L	U	X	T	A	D	G	C	I	J	N	R	P	V	Y	B	E	H	K	O	S
K	C	J	N	W	T	I	R	U	Q	X	A	D	Z	F	G	K	O	M	S	V	Y	B	E	H	L	P
L	Z	G	K	T	Q	F	O	R	N	U	X	A	W	C	D	H	L	J	P	S	V	Y	B	E	I	M
M	D	K	O	X	U	J	S	V	R	Y	B	E	A	G	H	L	P	N	T	W	Z	C	F	I	M	Q
N	X	E	I	R	O	D	M	P	L	S	V	Y	U	A	B	F	J	H	N	Q	T	W	Z	C	G	K
O	W	D	H	Q	N	C	L	O	K	R	U	X	T	Z	A	E	I	G	M	P	S	V	Y	B	F	J
P	S	Z	D	M	J	Y	H	K	G	N	Q	T	P	V	W	A	E	C	I	L	O	R	U	X	B	F
Q	O	V	Z	I	F	U	D	G	C	J	M	P	L	R	S	W	A	Y	E	H	K	N	Q	T	X	B
R	Q	X	B	K	H	W	F	I	E	L	O	R	N	T	U	Y	C	A	G	J	M	P	S	V	Z	D
S	K	R	V	E	B	Q	Z	C	Y	F	I	L	H	N	O	S	W	U	A	D	G	J	M	P	T	X
T	H	O	S	B	Y	N	W	Z	V	C	F	I	E	K	L	P	T	R	X	A	D	G	J	M	Q	U
U	E	L	P	Y	V	K	T	W	S	Z	C	F	B	H	I	M	Q	O	U	X	A	D	G	J	N	R
V	B	I	M	V	S	H	Q	T	P	W	Z	C	Y	E	F	J	N	L	R	U	X	A	D	G	K	O
W	Y	F	J	S	P	E	N	Q	M	T	W	Z	V	B	C	G	K	I	O	R	U	X	A	D	H	L
X	V	C	G	P	M	B	K	N	J	Q	T	W	S	Y	Z	D	H	F	L	O	R	U	X	A	E	I
Y	R	Y	C	L	I	X	G	J	F	M	P	S	O	U	V	Z	D	B	H	K	N	Q	T	W	A	E
Z	N	U	Y	H	E	T	C	F	B	I	L	O	K	Q	R	V	Z	X	D	G	J	M	P	S	W	A

Over

TABLE III

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\Theta_{k/1}=\Theta_{1/2}$; $\Theta_{p/1}=\Theta_{c/2}$ ($\Theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X
B	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O
C	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N
D	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W
E	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L
F	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A
G	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V
H	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K
I	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M
J	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U
K	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J
L	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D
M	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H	T
N	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E	H
O	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S	E
P	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G	S
Q	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I	G
R	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z	I
S	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q	Z
T	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C	Q
U	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R	C
V	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y	R
W	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P	Y
X	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B	P
Y	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F	B
Z	B	P	Y	R	C	Q	Z	I	G	S	E	H	T	D	J	U	M	K	V	A	L	W	N	O	X	F

TABLE IV

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/1}=\theta_{1/2}$; $\theta_{p/2}=\theta_{c/1}$ ($\theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	U	B	F	O	L	A	J	M	I	P	S	V	R	X	Y	C	G	E	K	N	Q	T	W	Z	D	H
B	V	C	G	P	M	B	K	N	J	Q	T	W	S	Y	Z	D	H	F	L	O	R	U	X	A	E	I
C	W	D	H	Q	N	C	L	O	K	R	U	X	T	Z	A	E	I	G	M	P	S	V	Y	B	F	J
D	X	E	I	R	O	D	M	P	L	S	V	Y	U	A	B	F	J	H	N	Q	T	W	Z	C	G	K
E	Y	F	J	S	P	E	N	Q	M	T	W	Z	V	B	C	G	K	I	O	R	U	X	A	D	H	L
F	Z	G	K	T	Q	F	O	R	N	U	X	A	W	C	D	H	L	J	P	S	V	Y	B	E	I	M
G	A	H	L	U	R	G	P	S	O	V	Y	B	X	D	E	I	M	K	Q	T	W	Z	C	F	J	N
H	B	I	M	V	S	H	Q	T	P	W	Z	C	Y	E	F	J	N	L	R	U	X	A	D	G	K	O
I	C	J	N	W	T	I	R	U	Q	X	A	D	Z	F	G	K	O	M	S	V	Y	B	E	H	L	P
J	D	K	O	X	U	J	S	V	R	Y	B	E	A	G	H	L	P	N	T	W	Z	C	F	I	M	Q
K	E	L	P	Y	V	K	T	W	S	Z	C	F	B	H	I	M	Q	O	U	X	A	D	G	J	N	R
L	F	M	Q	Z	W	L	U	X	T	A	D	G	C	I	J	N	R	P	V	Y	B	E	H	K	O	S
M	G	N	R	A	X	M	V	Y	U	B	E	H	D	J	K	O	S	Q	W	Z	C	F	I	L	P	T
N	H	O	S	B	Y	N	W	Z	V	C	F	I	E	K	L	P	T	R	X	A	D	G	J	M	Q	U
O	I	P	T	C	Z	O	X	A	W	D	G	J	F	L	M	Q	U	S	Y	B	E	H	K	N	R	V
P	J	Q	U	D	A	P	Y	B	X	E	H	K	G	M	N	R	V	T	Z	C	F	I	L	O	S	W
Q	K	R	V	E	B	Q	Z	C	Y	F	I	L	H	N	O	S	W	U	A	D	G	J	M	P	T	X
R	L	S	W	F	C	R	A	D	Z	G	J	M	I	O	P	T	X	V	B	E	H	K	N	Q	U	Y
S	M	T	X	G	D	S	B	E	A	H	K	N	J	P	Q	U	Y	W	C	F	I	L	O	R	V	Z
T	N	U	Y	H	E	T	C	F	B	I	L	O	K	Q	R	V	Z	X	D	G	J	M	P	S	W	A
U	O	V	Z	I	F	U	D	G	C	J	M	P	L	R	S	W	A	Y	E	H	K	N	Q	T	X	B
V	P	W	A	J	G	V	E	H	D	K	N	Q	M	S	T	X	B	Z	F	I	L	O	R	U	Y	C
W	Q	X	B	K	H	W	F	I	E	L	O	R	N	T	U	Y	C	A	G	J	M	P	S	V	Z	D
X	R	Y	C	L	I	X	G	J	F	M	P	S	O	U	V	Z	D	B	H	K	N	Q	T	W	A	E
Y	S	Z	D	M	J	Y	H	K	G	N	Q	T	P	V	W	A	E	C	I	L	O	R	U	X	B	F
Z	T	A	E	N	K	Z	I	L	H	O	R	U	Q	W	X	B	F	D	J	M	P	S	V	Y	C	G

Over

TABLE V

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/2}=\theta_{p/1}$; $\theta_{i/1}=\theta_{c/2}$ ($\theta_{i/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L
B	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P
C	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q
D	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J
E	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H
F	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B
G	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S
H	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T
I	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G
J	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U
K	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V
L	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W
M	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K
N	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O
O	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X
P	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y
Q	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z
R	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C
S	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E
T	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D
U	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M
V	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A
W	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N
X	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F
Y	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R
Z	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I

Table VI

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/2}=\theta_{c/1}$; $\theta_{1/1}=\theta_{p/2}$ ($\theta_{1/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	T	P	G	J	U	L	I	M	F	C	Z	D	X	W	S	O	Q	K	H	E	B	Y	V	R	N
B	H	A	W	N	Q	B	S	P	T	M	J	G	K	E	D	Z	V	X	R	O	L	I	F	C	Y	U
C	L	E	A	R	U	F	W	T	X	Q	N	K	O	I	H	D	Z	B	V	S	P	M	J	G	C	Y
D	U	N	J	A	D	O	F	C	G	Z	W	T	X	R	Q	M	I	K	E	B	Y	V	S	P	L	H
E	R	K	G	X	A	L	C	Z	D	W	T	Q	U	O	N	J	F	H	B	Y	V	S	P	M	I	E
F	G	Z	V	M	P	A	R	O	S	L	I	F	J	D	C	Y	U	W	Q	N	K	H	E	B	X	T
G	P	I	E	V	Y	J	A	X	B	U	R	O	S	M	L	H	D	F	Z	W	T	Q	N	K	G	C
H	S	L	H	Y	B	M	D	A	E	X	U	R	V	P	O	K	G	I	C	Z	W	T	Q	N	J	F
I	O	H	D	U	X	I	Z	W	A	T	Q	N	R	L	K	G	C	E	Y	V	S	P	M	J	F	B
J	V	O	K	B	E	P	G	D	H	A	X	U	Y	S	R	N	J	L	F	C	Z	W	T	Q	M	I
K	Y	R	N	E	H	S	J	G	K	D	A	X	B	V	U	Q	M	O	I	F	C	Z	W	T	P	L
L	B	U	Q	H	K	V	M	J	N	G	D	A	E	Y	X	T	P	R	L	I	F	C	Z	W	S	O
M	X	Q	M	D	G	R	I	F	J	C	Z	W	A	U	T	P	L	N	H	E	B	Y	V	S	O	K
N	D	W	S	J	M	X	O	L	P	I	F	C	G	A	Z	V	R	T	N	K	H	E	B	Y	U	Q
O	E	X	T	K	N	Y	P	M	Q	J	G	D	H	B	A	W	S	U	O	L	I	F	C	Z	V	R
P	I	B	X	O	R	C	T	Q	U	N	K	H	L	F	E	A	W	Y	S	P	M	J	G	D	Z	V
Q	M	F	B	S	V	G	X	U	Y	R	O	L	P	J	I	E	A	C	W	T	Q	N	K	H	D	Z
R	K	D	Z	Q	T	E	V	S	W	P	M	J	N	H	G	C	Y	A	U	R	O	L	I	F	B	X
S	Q	J	F	W	Z	K	B	Y	C	V	S	P	T	N	M	I	E	G	A	X	U	R	O	L	H	D
T	T	M	I	Z	C	N	E	B	F	Y	V	S	W	Q	P	L	H	J	D	A	X	U	R	O	K	G
U	W	P	L	C	F	Q	H	E	I	B	Y	V	Z	T	S	O	K	M	G	D	A	X	U	R	N	J
V	Z	S	O	F	I	T	K	H	L	E	B	Y	C	W	V	R	N	P	J	G	D	A	X	U	Q	M
W	C	V	R	I	L	W	N	K	O	H	E	W	F	U	T	U	Q	S	M	J	G	D	V	X	T	P
X	F	Y	U	L	O	Z	Q	N	R	K	H	A	I	C	B	X	T	V	P	M	J	G	D	A	W	S
Y	J	C	Y	P	S	D	U	R	V	O	L	I	M	G	F	B	X	Z	T	Q	N	K	H	E	A	W
Z	N	G	C	T	W	H	Y	V	Z	S	P	M	Q	K	J	F	B	D	X	U	R	O	L	I	E	A

Over

Table VII

Components:

(1)—A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

(2)—F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\Theta_{k/2}=\Theta_{p/1}$; $\Theta_{1/2}=\Theta_{c/1}$ ($\Theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
B	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
C	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
D	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
E	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
F	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
G	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
H	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
I	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
J	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
K	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
L	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
M	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
N	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
O	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
P	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Q	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
R	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
S	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
T	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
U	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
V	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
W	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
X	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
Y	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Z	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S

Table VIII

per N

Componets:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/2}=\theta_{e/1}$; $\theta_{1/2}=\theta_{p/1}$ ($\theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
D	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
E	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
F	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
G	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
H	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
K	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
L	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
M	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
N	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
O	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
P	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
Q	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
R	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
S	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
T	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
U	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
V	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Y	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
Z	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G

Over

TABLE IX [2]

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/1}=\theta_{p/2}$; $\theta_{i/1}=\theta_{c/2}$ ($\theta_{1/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	V	F	R	T	S	X	I	E	Z	D	M	A	U	W	N	B	C	Y	G	H	J	K	L	O	P	Q
C	K	X	Y	H	G	O	Z	S	Q	T	U	V	J	L	W	F	R	P	I	E	D	M	A	N	B	C
D	M	O	P	E	I	N	Q	G	C	H	J	K	D	A	L	X	Y	B	Z	S	T	U	V	W	F	R
E	U	N	B	S	Z	W	C	I	R	E	D	M	T	V	A	O	P	F	Q	G	H	J	K	L	X	Y
F	J	W	F	G	Q	L	R	Z	Y	S	T	U	H	K	V	N	B	X	C	I	E	D	M	A	O	P
G	D	L	X	I	C	A	Y	Q	P	G	H	J	E	M	K	W	F	O	R	Z	S	T	U	V	N	B
H	T	A	O	Z	R	V	P	C	B	I	E	D	S	U	M	L	X	N	Y	Q	G	H	J	K	W	F
I	H	V	N	Q	Y	K	B	R	F	Z	S	T	G	J	U	A	O	W	P	C	I	E	D	M	L	X
J	E	K	W	C	P	M	F	Y	X	Q	G	H	I	D	J	V	N	L	B	R	Z	S	T	U	A	O
K	S	M	L	R	B	U	X	P	O	C	I	E	Z	T	D	K	W	A	F	Y	Q	G	H	J	V	N
L	G	U	A	Y	F	J	O	B	N	R	Z	S	Q	H	T	M	L	V	X	P	C	I	E	D	K	W
M	I	J	V	P	X	D	N	F	W	Y	Q	G	C	E	H	U	A	K	O	B	R	Z	S	T	M	L
N	Z	D	K	B	O	T	W	X	L	P	C	I	R	S	E	J	V	M	N	F	Y	Q	G	H	U	A
O	Q	T	M	F	N	H	L	O	A	B	R	Z	Y	G	S	D	K	U	W	X	P	C	I	E	J	V
P	C	H	U	X	W	E	A	N	V	F	Y	Q	P	I	G	T	M	J	L	O	B	R	Z	S	D	K
Q	R	E	J	O	L	S	V	W	K	X	P	C	B	Z	I	H	U	D	A	N	F	Y	Q	G	T	M
R	Y	S	D	N	A	G	K	L	M	O	B	R	F	Q	Z	E	J	T	V	W	X	P	C	I	H	U
S	P	G	T	W	V	I	M	A	U	N	F	Y	X	C	Q	S	D	H	K	L	O	B	R	Z	E	J
T	B	I	H	L	K	Z	U	V	J	W	X	P	O	R	C	G	T	E	M	A	N	F	Y	Q	S	D
U	F	Z	E	A	M	Q	J	K	D	L	O	B	N	Y	R	I	H	S	U	V	W	X	P	C	G	T
V	X	Q	S	V	U	C	D	M	T	A	N	F	W	P	Y	Z	E	G	J	K	L	O	B	R	I	H
W	O	C	G	K	J	R	T	U	H	V	W	X	L	B	P	Q	S	I	D	M	A	N	F	Y	Z	E
X	N	R	I	M	D	Y	H	J	E	K	L	O	A	F	B	C	G	Z	T	U	V	W	X	P	Q	S
Y	W	Y	Z	U	T	P	E	D	S	M	A	N	V	X	F	R	I	Q	H	J	K	L	O	B	C	G
Z	L	P	Q	J	H	B	S	T	G	U	V	W	K	O	X	Y	Z	C	E	D	M	A	N	F	R	I

[2] An interesting fact about this case is that if the plain component is made identical with the cipher component (both being the sequence FBPY ...), and if the enciphering equations are the same as for Table 1-B, then the resultant cipher square is identical with Table IX, except that the key letters at the left are in the order of the reversed mixed component, FXON .... In other words, the secondary cipher alphabets produced by the interaction of two identical mixed components are the same as those given by the interaction of a mixed component and the normal component.

Table X[3]

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\Theta_{k/1}=\Theta_{e/2}$; $\Theta_{i/1}=\Theta_{p/2}$ ($\Theta_{1/1}$ is A).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	L	P	Q	J	H	B	S	T	G	U	V	W	K	O	X	Y	Z	C	E	D	M	A	N	F	R	I
C	W	Y	Z	U	T	P	E	D	S	M	A	N	V	X	F	R	I	Q	H	J	K	L	O	B	C	G
D	N	R	I	M	D	Y	H	J	E	K	L	O	A	F	B	C	G	Z	T	U	V	W	X	P	Q	S
E	O	C	G	K	J	R	T	U	H	V	W	X	L	B	P	Q	S	I	D	M	A	N	F	Y	Z	E
F	X	Q	S	V	U	C	D	M	T	A	N	F	W	P	Y	Z	E	G	J	K	L	O	B	R	I	H
G	F	Z	E	A	M	Q	J	K	D	L	O	B	N	Y	R	I	H	S	U	V	W	X	P	C	G	T
H	B	I	H	L	K	Z	U	V	J	W	X	P	O	R	C	G	T	E	M	A	N	F	Y	Q	S	D
I	P	G	T	W	V	I	M	A	U	N	F	Y	X	C	Q	S	D	H	K	L	O	B	R	Z	E	J
J	Y	S	D	N	A	G	K	L	M	O	B	R	F	Q	Z	E	J	T	V	W	X	P	C	I	H	U
K	R	E	J	O	L	S	V	W	K	X	P	C	B	Z	I	H	U	D	A	N	F	Y	Q	G	T	M
L	C	H	U	X	W	E	A	N	V	F	Y	Q	P	I	G	T	M	J	L	O	B	R	Z	S	D	K
M	Q	T	M	F	N	H	L	O	A	B	R	Z	Y	G	S	D	K	U	W	X	P	C	I	E	J	V
N	Z	D	K	B	O	T	W	X	L	P	C	I	R	S	E	J	V	M	N	F	Y	Q	G	H	U	A
O	I	J	V	P	X	D	N	F	W	Y	Q	G	C	E	H	U	A	K	O	B	R	Z	S	T	M	L
P	G	U	A	Y	F	J	O	B	N	R	Z	S	Q	H	T	M	L	V	X	P	C	I	E	D	K	W
Q	S	M	L	R	B	U	X	P	O	C	I	E	Z	T	D	K	W	A	F	Y	Q	G	H	J	V	N
R	E	K	W	C	P	M	F	Y	X	Q	G	H	I	D	J	V	N	L	B	R	Z	S	T	U	A	O
S	H	V	N	Q	Y	K	B	R	F	Z	S	T	G	J	U	A	O	W	P	C	I	E	D	M	L	X
T	T	A	O	Z	R	V	P	C	B	I	E	D	S	U	M	L	X	N	Y	Q	G	H	J	K	W	F
U	D	L	X	I	C	A	Y	Q	P	G	H	J	E	M	K	W	F	O	R	Z	S	T	U	V	N	B
V	J	W	F	G	Q	L	R	Z	Y	S	T	U	H	K	V	N	B	X	C	I	E	D	M	A	O	P
W	U	N	B	S	Z	W	C	I	R	E	D	M	T	V	A	O	P	F	Q	G	H	J	K	L	X	Y
X	M	O	P	E	I	N	Q	G	C	H	J	K	D	A	L	X	Y	B	Z	S	T	U	V	W	F	R
Y	K	X	Y	H	G	O	Z	S	Q	T	U	V	J	L	W	F	R	P	I	E	D	M	A	N	B	C
Z	V	F	R	T	S	X	I	E	Z	D	M	A	U	W	N	B	C	Y	G	H	J	K	L	O	P	Q

[3] Footnote 2 to Table IX, page 104, also applies to this table, except that the key letters at the left will follow the order of the direct mixed component.

Over

Table XI

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/1}=\theta_{p/2}$; $\theta_{1/2}=\theta_{c/1}$ ($\theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	G	Z	V	M	P	A	R	O	S	L	I	F	J	D	C	Y	U	W	Q	N	K	H	E	B	X	T
B	H	A	W	N	Q	B	S	P	T	M	J	G	K	E	D	Z	V	X	R	O	L	I	F	C	Y	U
C	I	B	X	O	R	C	T	Q	U	N	K	H	L	F	E	A	W	Y	S	P	M	J	G	D	Z	V
D	J	C	Y	P	S	D	U	R	V	O	L	I	M	G	F	B	X	Z	T	Q	N	K	H	E	A	W
E	K	D	Z	Q	T	E	V	S	W	P	M	J	N	H	G	C	Y	A	U	R	O	L	I	F	B	X
F	L	E	A	R	U	F	W	T	X	Q	N	K	O	I	H	D	Z	B	V	S	P	M	J	G	C	Y
G	M	F	B	S	V	G	X	U	Y	R	O	L	P	J	I	E	A	C	W	T	Q	N	K	H	D	Z
H	N	G	C	T	W	H	Y	V	Z	S	P	M	Q	K	J	F	B	D	X	U	R	O	L	I	E	A
I	O	H	D	U	X	I	Z	W	A	T	Q	N	R	L	K	G	C	E	Y	V	S	P	M	J	F	B
J	P	I	E	V	Y	J	A	X	B	U	R	O	S	M	L	H	D	F	Z	W	T	Q	N	K	G	C
K	Q	J	F	W	Z	K	B	Y	C	V	S	P	T	N	M	I	E	G	A	X	U	R	O	L	H	D
L	R	K	G	X	A	L	C	Z	D	W	T	Q	U	O	N	J	F	H	B	Y	V	S	P	M	I	E
M	S	L	H	Y	B	M	D	A	E	X	U	R	V	P	O	K	G	I	C	Z	W	T	Q	N	J	F
N	T	M	I	Z	C	N	E	B	F	Y	V	S	W	Q	P	L	H	J	D	A	X	U	R	O	K	G
O	U	N	J	A	D	O	F	C	G	Z	W	T	X	R	Q	M	I	K	E	B	Y	V	S	P	L	H
P	V	O	K	B	E	P	G	D	H	A	X	U	Y	S	R	N	J	L	F	C	Z	W	T	Q	M	I
Q	W	P	L	C	F	Q	H	E	I	B	Y	V	Z	T	S	O	K	M	G	D	A	X	U	R	N	J
R	X	Q	M	D	G	R	I	F	J	C	Z	W	A	U	T	P	L	N	H	E	B	Y	V	S	O	K
S	Y	R	N	E	H	S	J	G	K	D	A	X	B	V	U	Q	M	O	I	F	C	Z	W	T	P	L
T	Z	S	O	F	I	T	K	H	L	E	B	Y	C	W	V	R	N	P	J	G	D	A	X	U	Q	M
U	A	T	P	G	J	U	L	I	M	F	C	Z	D	X	W	S	O	Q	K	H	E	B	Y	V	R	N
V	B	U	Q	H	K	V	M	J	N	G	D	A	E	Y	X	T	P	R	L	I	F	C	Z	W	S	O
W	C	V	R	I	L	W	N	K	O	H	E	B	F	Z	Y	U	Q	S	M	J	G	D	A	X	T	P
X	D	W	S	J	M	X	O	L	P	I	F	C	G	A	Z	V	R	T	N	K	H	E	B	Y	U	Q
Y	E	X	T	K	N	Y	P	M	Q	J	G	D	H	B	A	W	S	U	O	L	I	F	C	Z	V	R
Z	F	Y	U	L	O	Z	Q	N	R	K	H	E	I	C	B	X	T	V	P	M	J	G	D	A	W	S

TABLE XII

Components:

(1) A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(2) F B P Y R C Q Z I G S E H T D J U M K V A L W N O X

Enciphering equations: $\theta_{k/1}=\theta_{c/2}$; $\theta_{1/2}=\theta_{p/1}$ ($\theta_{1/2}$ is F).

PLAIN TEXT

KEY	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B
B	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P
C	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y
D	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R
E	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C
F	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q
G	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z
H	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I
I	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G
J	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S
K	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E
L	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T	H
M	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D	T
N	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J	D
O	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U	J
P	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M	U
Q	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K	M
R	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V	K
S	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A	V
T	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L	A
U	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W	L
V	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N	W
W	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O	N
X	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X	O
Y	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F	X
Z	X	O	N	W	L	A	V	K	M	U	J	D	T	H	E	S	G	I	Z	Q	C	R	Y	P	B	F

APPENDIX 2[1]

Elementary Statistical Theory Applicable to the Phenomena of Repetition in Cryptanalysis

1. **Introductory.**—*a.* In Par. 9*c* it was stated that the phenomena of repetition in cryptanalytics may be removed from the realm of intuition and dealt with statistically. The discussion of the matter will here be confined to relatively simple phases of the theory of probability, a definition of which implies philosophical questions of no practical interest to the student of cryptanalysis. For his purposes, the following definition of *a priori* probability will be sufficient:

> The probability that an event will occur is the ratio of the number of "favorable cases" to the number of total possible cases, all cases being equally likely to occur. By a "favorable case" is meant one which will produce the event in question.

b. In what follows, reference will be made to *random assortments* of letters and especially to *random text.* By the latter will be meant merely that the text under consideration has been assumed to have been enciphered by some more or less complex cryptographic system so that for all practical purposes the sequence of letters constituting this text is a random assortment; that is, the sequence is just about what would have been obtained if the letters had been drawn at random out of a box containing a large number of the 26 letters of the alphabet, all in equal proportions, so that there are exactly the same numbers of A's, B's, C's, . . . Z's. It is assumed that each time in making a drawing from such a box, the latter is thoroughly shaken so that the letters are thoroughly mixed and then a single letter is selected at random, recorded, and replaced in the same box. In what follows, the word "box" will refer to the box as described.

c. A uniliteral frequency distribution of a large volume of random text will be "flat," i. e., lacking crests and troughs.

d. For purposes of statistical analysis, the text of a monoalphabetic substitution cipher is equivalent to plain text. As a corollary, when a polyalphabetic substitution cipher has been reduced to the simple terms of a set of monoalphabets, i. e., when the letters constituting the cipher text have been allocated into their proper uniliteral distributions, the letters falling into the respective distributions are statistically equivalent to plain text.

2. **Data pertaining to single letters.**—*a.* (1) A single letter will be drawn at random from the box. What is the probability that it will be an A? According to the foregoing definition of probability, since the total number of possible cases is 26 and the number of favorable cases is here only 1, the probability is $1:26=\frac{1}{26}=.0385$. This is the probability of drawing an A from the box. The probability that the letter drawn will be a B, a C, a D, . . ., a Z is the same as for A. In other words, the probability of drawing *any specified single letter* is $p=.0385$.

(2) The value $p=.0385$, as found above, may also be termed the probability constant for single letters in random text of a 26-letter alphabet. For any language this constant is merely the reciprocal of the total number of different characters which may be employed in writing the text in question.

[1] In the preparation of this appendix, the author has had the benefit of the very helpful suggestions of Capt. H. G. Miller, Signal Corps, Mr. F. B. Rowlett, Dr. S. Kullback, and Dr. A. Sinkov, Assistant Cryptanalysts, O. C. Sig. O. Certain parts of Dr. Kullback's important paper "Statistical Methods in Cryptanalysis" form the basis of the discussion.

(3) Another way of interpreting the notation $p=.0385$ is to say that in a large volume of random text, for example in 100,000 letters, any letter that one may choose to specify may be expected to occur about 3,850 times; in 10,000 letters it may be expected to occur about 385 times; in 1,000 letters, about 38.5 times, and so on. In every-day language it would be said that "in the long run" or "on the average" in 1,000 letters of random text there will be about 38.5 occurrences of each of the 26 letters of the alphabet.

(4) But unfortunately, in cryptanalysis it is not often the case that one has such a large number of letters available for study in any single cipher alphabet. More often the cryptanalyst has a relatively small number of letters and these must be distributed over several cipher alphabets. Hence it is necessary to be able to deal with smaller numbers of letters. Consider a specific piece of random text of only 100 letters. It has been seen that "in the long run" each letter may be expected to occur about 3.85 times in this amount of random text; that is, the 26 letters will have an *average* frequency of 3.85. But in reaching this average of 3.85 occurrences in 100 letters, it is obvious that some letter or letters may not appear at all, some may appear once, some twice, and so on. How many will not appear at all; how many will appear 1, 2, 3, . . . times? In other words, how will the different categories of letters (different in respect to frequency of occurrence) be distributed, or what will the *distribution* be like? Will it follow any kind of law or pattern? The cryptanalyst also wants to know the answer to questions such as these: What is the probability that a specified letter will not appear at all in a given piece of text? That it will appear *exactly* 1, 2, 3, . . . times? That it will appear *at least* 1, 2, 3, . . . times? The same sort of questions may be asked with respect to digraphs, trigraphs, and so on.

b. (1) It may be stated at once that questions of this nature are not easily answered, and a complete discussion falls quite outside the scope of this text. However, it will be sufficient for the present purposes if the student is provided with a more or less simple and practical means of finding the answers. With this in view certain curves have been prepared from data based upon Poisson's exponential expansion, or the "law of small probabilities" and their use will now be explained. Students without a knowledge of the mathematical theory of probability and statistics will have to take the curves "on faith" Those interested in their derivation are referred to the following texts:

Fisher, R. A., *Statistical Methods for Research Workers*, London, 1937.

Fry, T. C., *Probability and Its Engineering Uses*, New York, 1928.

(2) By means of these *probability curves*, it is possible to find, in a relatively easy manner, the probability for 0, 1, 2, . . . 11 occurrences of an event in n cases, if the *mean* (expected, average, probable) number of occurrences in these n cases is known. For example, given a cryptogram equivalent to 100 letters of random text, what is the probability that any specified single letter, whatever will not appear at all in the cryptogram? Since the probability of the occurrence of a specified single letter is $\frac{1}{26}=.0385$, and there are 100 letters in the cryptogram, the average or expected or mean number of occurrences of an A, a B, a C, . . ., is $.0385\times100=3.85$. Refer now to that probability curve which is marked "f_0", meaning "frequency zero", or "zero occurrences." On the horizontal or x axis of that curve find the point corresponding to the value 3.85 and follow the vertical coordinate determined by this value up to the point of intersection with the curve itself; then follow the horizontal coordinate determined by this intersection point over to the left and read the value on the vertical axis of the curve. It is approximately .021. This means that the probability that a specified single letter (an A, a B, a C, . . .) will not appear at all in the cryptogram, if it really were a perfectly random assortment of 100 letters, is .021.

strip Sig. 8

Over

That is, according to the theory of probability, in 1,000 cases of random-text messages of 100 letters each, one may expect to find about 21 messages in which a specified single letter will not appear at all. Another way of saying the same thing is: If 1,000 sets of 100 letters of random text are examined, in about 21 out of the 1,000 such sets any letter that one may choose to name will be absent. This, of course, is merely a theoretical expectancy; it indicates only what probably will happen in the long run.

(3) What is the probability that a specified single letter will appear *exactly* once in 100 letters of random text? To answer this question, find on the curve marked f_1, the point of intersection of the vertical coordinate corresponding to the mean or average value 3.85 with the curve; follow the horizontal coordinate thus determined over to the vertical scale at the left; read the value on this scale. It is .082, which means that in 1,000 cases of random-text messages of 100 letters each, one may expect to find about 82 messages in which any letter one chooses to specify will occur exactly once, no more and no less.

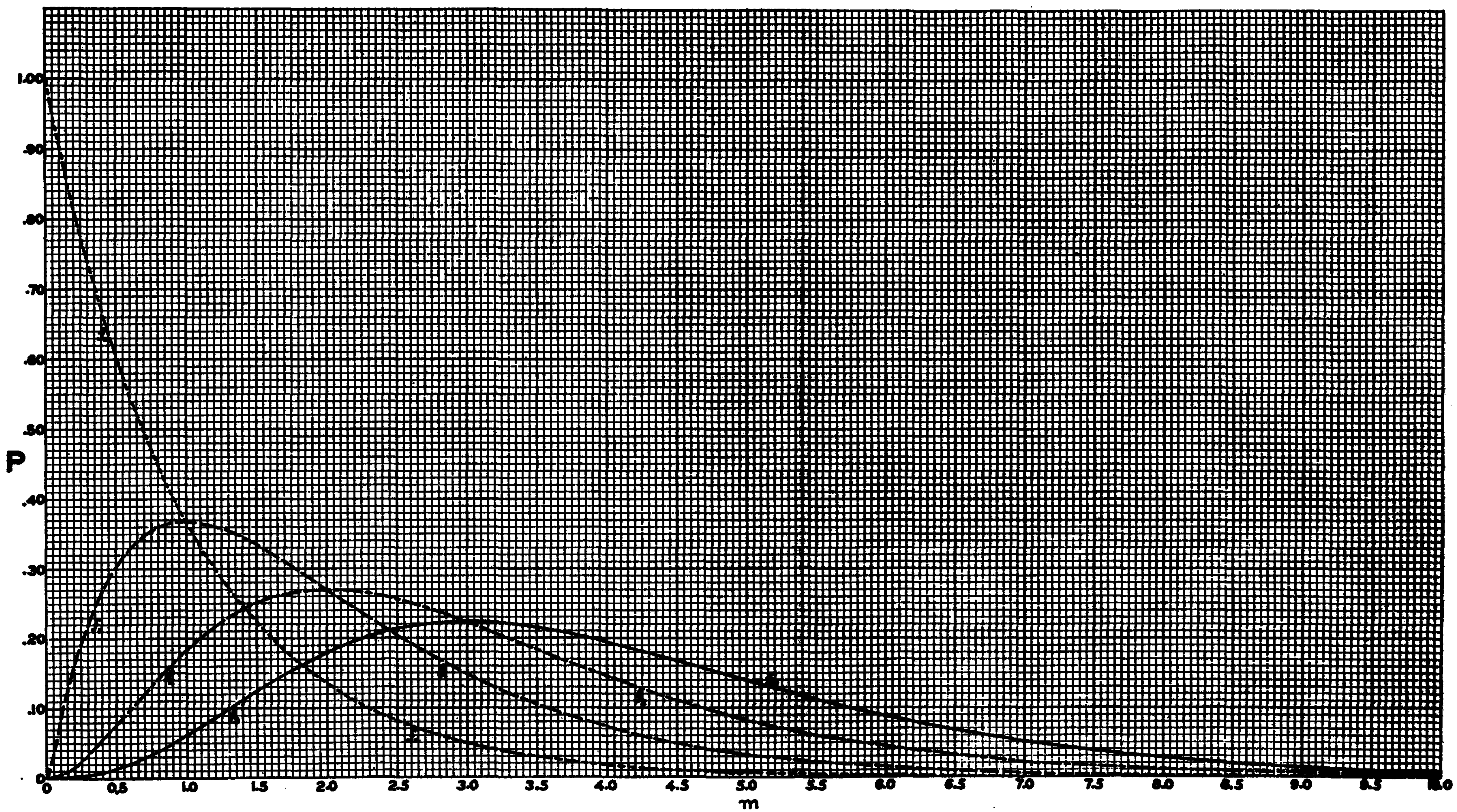
(4) In the same way, the probability that a specified single letter will appear *exactly* twice is found to be .158; exactly 3 times, .202; and so on, as shown in the table below:

100 letters of random text

Frequency (x)	Probability that a specified single letter will occur exactly x times
0	0.021
1	.082
2	.158
3	.202
4	.195
5	.150
6	.096
7	.053
8	.026
9	.011
10	.004
11	.001

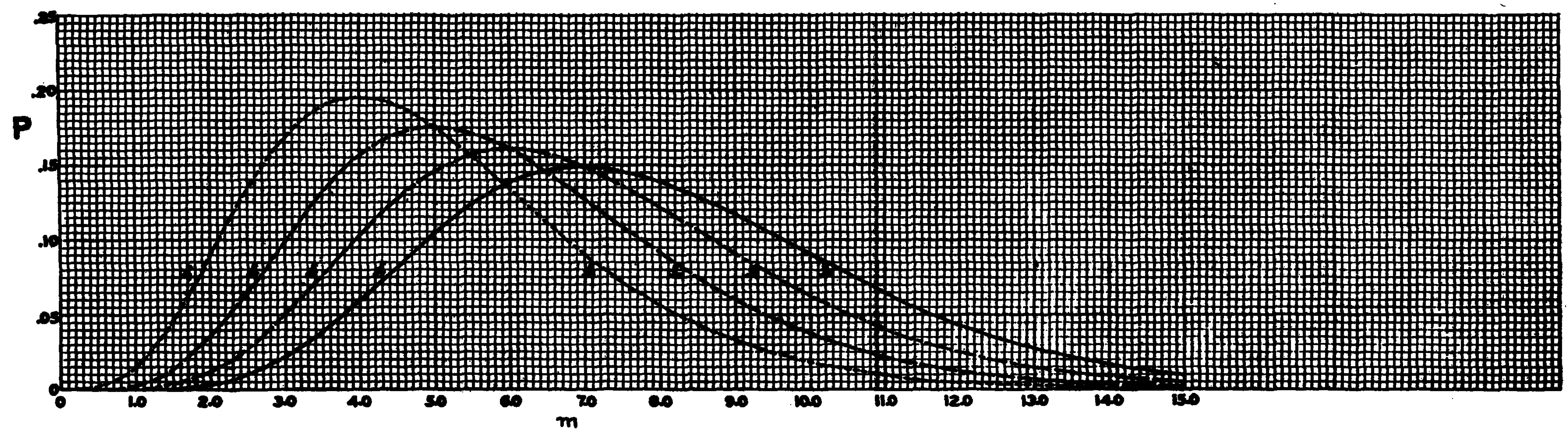
(5) To find the probability that a specified single letter will occur *at least* 1, 2, 3, . . . times in a series of letters constituting random text, one reasons as follows: Since the concept "at least 1" implies that the number specified is to be considered only as the minimum, with no limit indicated as to maximum, occurrences of 2, 3, 4, . . . are also "favorable" cases; the probabilities for *exactly* 1, 2, 3, 4, . . . occurrences should therefore be added and this will give the probability for "at least 1." Thus, in the case of 100 letters, the sum of the probabilities for exactly 1 to 11 occurrences, as set forth in the table directly above, is .978, and the latter value approximates the probability for at least 1 occurrence.

(6) A more accurate result will be obtained by the following reasoning. The probability for zero occurrences is .021. Since it is certain that a specified letter will occur either zero times or 1, 2, 3, . . . times, to find the probability for *at least* one time it is merely necessary to subtract the probability for zero occurrences from unity. That is, $1-.021=.979$, which is .001 greater than the result obtained by the other method. The reason it is greater is that the value .979 includes occurrences beyond 11, which were excluded from the previous calculation. Of course, the probabilities for these occurrences beyond 11 are very small, but taken all together they

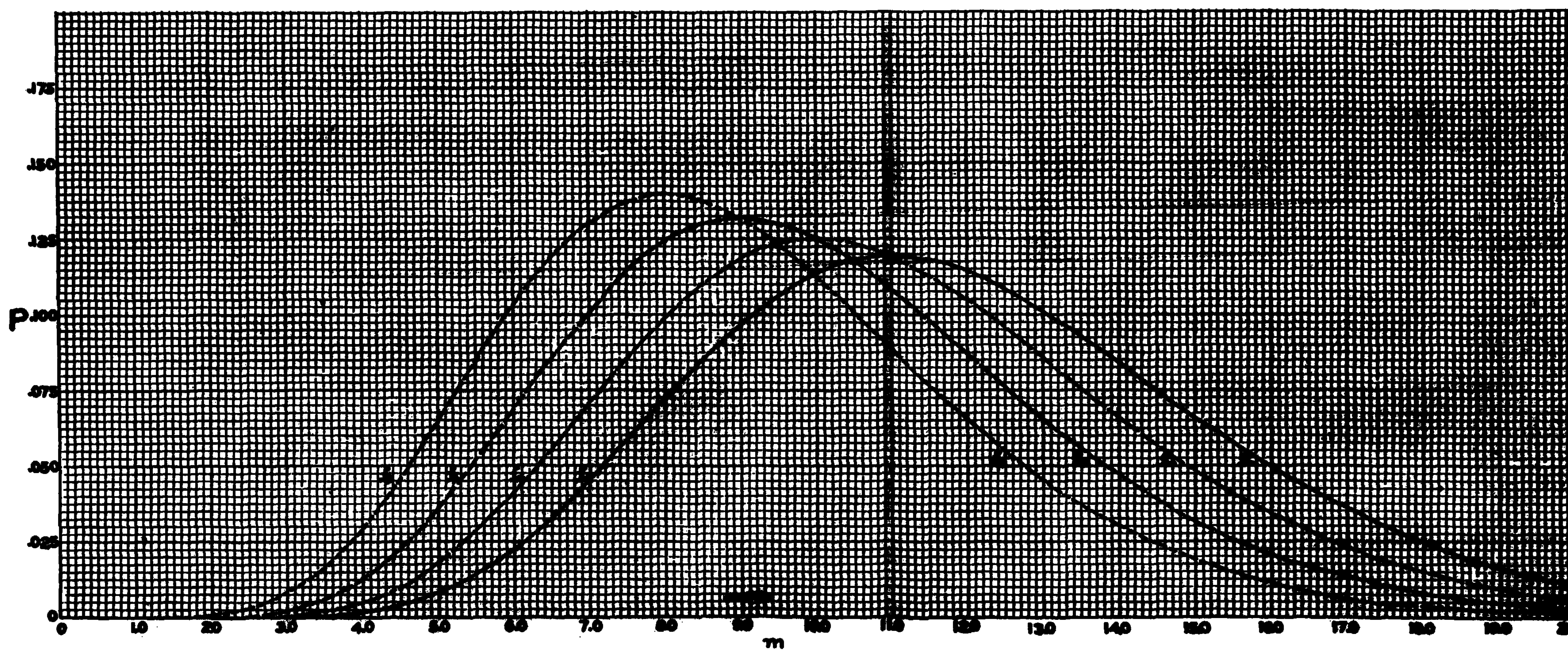


Curves showing probability for 0, 1, 2, and 3 occurrences of an event in *n* cases, given the mean number of occurrences. ~~[illegible]~~ (Face p. 110) No. 1

strip big



Curves showing probability for 4, 5, 6, and 7 occurrences of an event in n cases, given the mean number of occurrences.



Curves showing probability for 8, 9, 10, and 11 occurrences of an event in n cases, given the mean number of occurrences.

(Face p. 118) No. 2

mend sig

add up to .001, the difference between the results obtained by the two methods. The probability for at least 2 occurrences is the difference between unity and the sum of the probability for zero and exactly 1 occurrences; that is, $1-(P_0+P_1)=1-(.021+.082)=1-.103=.897$. The respective probabilities for various numbers of occurrences of a specified single letter (from 0 to 11) are given in the following table:

100 letters of random text

Frequency (x)	Probability that a specified single letter will occur exactly x times	Probability that a specified single letter will occur at least x times
0	0.021	1.000
1	.082	.979
2	.158	.897
3	.202	.739
4	.195	.537
5	.150	.342
6	.096	.192
7	.053	.096
8	.026	.043
9	.011	.017
10	.004	.006
11	.001	.002

(7) The foregoing calculations refer to random text composed of 100 letters. For other numbers of letters, it is merely necessary to find the mean (multiply the probability for drawing a specified single letter out of the box, which is $\frac{1}{26}$ or .0385, by the number of letters in the assortment) and refer to the various curves, as before. For example, for a random assortment of 200 letters, the mean is $200 \times .0385$, or 7.7, and this is the value of the point to be sought along the horizontal or x axes of the curves; the intersections of the respective vertical lines corresponding to this mean with the various curves for 0, 1, 2, 3, . . . occurrences give the probabilities for these occurrences, the reading being taken on the vertical or y axes of the curves.

(8) The discussion thus far has dealt with the probabilities for 0, 1, 2, 3, . . . occurrences of specified single letters. It may be of more practical advantage to the student if he could be shown how to find the answer to these questions: Given a random assortment of 100 letters *how many* letters may be expected to occur *exactly* 0, 1, 2, 3, . . . times? How many may be expected to occur *at least* 1, 2, 3, . . . times? The curves may here again be used to answer these questions, by a very simple calculation: multiply the probability value as obtained above for a specified single letter by the number of different elements being considered. For example, the probability that a specified single letter will occur exactly twice in a perfectly random assortment of 100 letters is .158; since the number of different letters is 26, the absolute number of single letters that may be expected to occur exactly 2 times in this assortment is $.158\times26=4.108$. That is, in 100 letters of random text there should be about four letters which occur exactly 2 times. The following table gives the data for various numbers of occurrences.

Over

100 letters of random text

Frequency (x)	Probability that a specified single letter will occur exactly x times	Probability that a specified single letter will occur at least x times	Probable number of letters appearing exactly x times	Probable number of letters appearing at least x times
0	0.021	1.000	0.546	26.000
1	.082	.979	2.132	25.454
2	.158	.897	4.108	23.322
3	.202	.739	5.252	19.214
4	.195	.537	5.070	13.962
5	.150	.342	3.900	8.892
6	.096	.192	2.496	4.992
7	.053	.096	1.378	2.496
8	.026	.043	.676	1.118
9	.011	.017	.286	.442
10	.004	.006	.104	.156
11	.001	.002	.026	.052

(9) Referring again to the curves, and specifically to the tabulated results set forth directly above, it will be seen that the probability that there will be exactly two occurrences of a specified single letter in 100 letters of random text (.158), is less than the probability that there will be exactly three occurrences (.202); in other words, the chances that a specified single letter will occur exactly three times are better, by about 25 percent, than that it will occur only two times. Furthermore, there will be about five letters which will occur exactly 3 times, and about five which will occur exactly 4 times, whereas there will be only about two letters which will occur exactly 1 time. Other facts of a similar import may be deduced from the foregoing table.

c. The discussion thus far has dealt with random assortments of letters. What about other types of texts, for example, normal plain text? What is the probability that E will occur 0, 1, 2, 3, . . . times in 50 letters of normal English? The relative frequency value or probability that a letter selected at random from a large volume of normal English text will be E is .12604. (In 100,000 letters E occurred 12,604 times.) For 50 letters this value must be multiplied by 50, giving 6.3 as the mean or point to be found along the x axes of the curves. The probabilities for 0, 1, 2, 3, . . . occurrences are tabulated below:

50 letters of normal English plain text

Frequency (x)	Probability that an E will be drawn exactly x times	Probability that an E will be drawn at least x times
0	0.002	1.000
1	.011	.998
2	.036	.987
3	.076	.951
4	.120	.875
5	.151	.755
6	.159	.604
7	.143	.445
8	.113	.302
9	.079	.223
10	.050	.173
11	.029	.123

(3

d. (1) It has been seen that the probability of occurrence of a specified single letter in random text employing a 26-letter alphabet is $p=\frac{1}{26}=.0385$. If a considerable volume of such text is written on a large sheet of paper and a pencil is directed at random toward this text, the probability that the pencil point will hit the letter A, *or any other letter which may be specified in advance*, is .0385. Now suppose two pencils are directed simultaneously toward the sheet of paper. The probability that both pencil points will hit two A's is $\frac{1}{26}\times\frac{1}{26}=\frac{1}{26^2}=.00148$, since in this case one is dealing with the probability of the simultaneous occurrence of two events which are independent. The probability of hitting two B's, two C's, . . ., two Z's is likewise $\frac{1}{26^2}$. Hence, if no particular letter is specified, and merely this question is asked: "What is the probability that both pencil points will hit the same letter?" the answer must be the sum of the separate probabilities for simultaneously hitting two A's, two B's, and so on, for the whole alphabet, which is $26\times\frac{1}{26^2}=\frac{1}{26}=.0385$. This, then, is the probability that *any* two letters selected *at random* in random text of a 26-letter alphabet will be identical or will *coincide*. Since this value remains the same so long as the number of alphabetic elements remains fixed, it may be said that *the probability of monographic coincidence in random text of a 26-element alphabet is* .0385. The foregoing italicized expression [2] is important enough to warrant assigning a special symbol to it, *viz*, κ_r (read "kappa sub-*r*"). For a 26-element alphabet, then, $\kappa_r=.0385$.

(2) Now if one asks: "Given a random assortment of 10 letters, what are the respective probabilities of occurrence of 0, 1, 2, . . . single-letter coincidences?" one proceeds as follows. As before, it is first necessary to find the mean or expected number of coincidences and then refer to the various probability curves. To find the mean, one reasons as follows. Given a sequence of 10 letters, one may begin with the 1st letter and compare it with the 2d, 3d, . . . 10th letter to see if any two letters coincide; 9 such comparisons may be made, or in other words there are, beginning with the 1st letter, 9 opportunities for the occurrence of a coincidence. But one may also start with the 2nd letter and compare it with the 3d, 4th . . . 10th letter, thus yielding 8 more opportunities for the occurrence of a coincidence, and so on. This process may continue until one reaches the 9th letter and compares it with the 10th, yielding but one opportunity for the occurrence in question. The total number of comparisons that can be made is therefore the sum of the series of numbers 9, 8, 7, . . . 1, which is 45 comparisons.[3] Since in the 10 letters there are 45 opportunities for coincidence of single letters, and since the probability

---

[2] The expression itself may be termed a *parameter*, which in mathematics is often used to designate a constant that characterizes by each of its particular values some particular member of a system of values, functions, etc. The word is applicable in the case under discussion because the value obtained for κ_r is .0385; for a 25-element alphabet, $\kappa_r=.0400$; for a 27-element alphabet, $\kappa_r=.0370$, etc.

[3] The number of comparisons may readily be found by the formula $\frac{n(n-1)}{2}$, where n is the total number of letters involved. This formula is merely a special case under the general formula for ascertaining the number of combinations that may be made of n different things taken r at a time, which is ${}_nC_r=\frac{n!}{r!(n-r)!}$. In the present case, since only two letters are compared at a time, r is always 2, and hence the expression $\frac{n!}{r!(n-r)!}$, which is the same as $\frac{n(n-1)(n-2)!}{2(n-2)!}$, becomes by cancellation of the term $(n-2)!$ reduced to $\frac{n(n-1)}{2}$.

Over

for monographic coincidence in random text is .0385 the expected number of coincidences is $.0385 \times 45 = 1.7325$. With $m=1.7$ one consults the various probability curves and an approximate distribution for exactly and for at least 0, 1, 2, . . . coincidences may readily be ascertained.[4]

e. (1) Now consider the matter of monographic coincidence in English plain text.[5] Following the same reasoning outlined in subpar. *d* (1), the probability of coincidence of two A's in plain text is the square of the probability of occurrence of the single letter A in such text. The probability of coincidence of two B's is the square of the probability of occurrence of the single letter B, and so on. The sum of these squares for all the letters of the alphabet, as shown in the following table, is found to be .0667.

Letter	Frequency [1] in 1,000 letters	Probability of separate occurrence of the letter	Square of probability of separate occurrence
A	73. 66	0. 0737	0. 0054
B	9. 74	. 0097	. 0001
C	30. 68	. 0307	. 0009
D	42. 44	. 0424	. 0018
E	129. 96	. 1300	. 0169
F	28. 32	. 0283	. 0008
G	16. 38	. 0164	. 0003
H	33. 88	. 0339	. 0012
I	73. 52	. 0735	. 0054
J	1. 64	. 0016	. 0000
K	2. 96	. 0030	. 0000
L	36. 42	. 0364	. 0013
M	24. 74	. 0247	. 0006
N	79. 50	. 0795	. 0063
O	75. 28	. 0753	. 0057
P	26. 70	. 0267	. 0007
Q	3. 50	. 0035	. 0000
R	75. 76	. 0758	. 0057
S	61. 16	. 0612	. 0037
T	91. 90	. 0919	. 0084
U	26. 00	. 0260	. 0007
V	15. 32	. 0153	. 0002
W	15. 60	. 0156	. 0002
X	4. 62	. 0046	. 0000
Y	19. 34	. 0193	. 0004
Z	. 98	. 0010	. 0000
Total	1,000. 00	1. 0000	. 0667

[1] The data given are taken from Table 3, Appendix 1, Military Cryptanalysis, Part I.

This then is the probability that any two letters selected at random in a large volume of normal English telegraphic plain text will coincide. Since this value remains the same so long as the character of the language does not change radically, it may be said that *the probability of monographic coincidence in English telegraphic plain text is* .0667, or $\kappa_p=.0667$.

[4] The approximation given by the Poisson distribution in the case of single letters is not as good as that in the case of digraphs, trigraphs, etc., discussed in paragraphs 3, 4, below.

[5] The theory of monographic coincidence in plain text was originally developed and applied by the author in a technical paper written in 1925 dealing with his solution of messages enciphered by a cryptograph known as the "Hebern Electric Super-Code." The paper was printed in 1934.

(2) Given 10 letters of English plain text, what is the probability that there will be 0, 1, 2, . . . single-letter coincidences? Following the line of reasoning in subparagraph *d* (2), the expected number of coincidences is $.0667 \times 45 = 3.00$, or $m=3$. The distribution for exactly and for at least 0, 1, 2, . . . coincidences may readily be found by reference to the various probability curves. (See footnote 4.)

f. The fact that κ_p (for English) is almost twice as great as κ_r is of considerable importance in cryptanalysis. It will be dealt with in detail in a subsequent text. At this point it will merely be said that κ_p and κ_r for other languages and alphabets have been calculated and show considerable variation, as will be noted in the table shown in paragraph 3*d*.

3. Data pertaining to digraphs.—*a.* (1) The foregoing discussion has been restricted to questions concerning single letters, but by slight modification it can be applied to questions concerning digraphs, trigraphs, and longer polygraphs.

(2) In the preceding cases it was necessary, before referring to the various probability curves, to find the mean or expected number of occurrences of the event in question in the total number of cases or trials being considered. Given a piece of random text totalling 100 letters, for example, what is the mean (average, probable, expected) number of occurrences of digraphs in this text? Since there are 676 different digraphs, the probability of occurrence of any specified digraph is $\frac{1}{676} = .00148$; since in 100 letters there are 99 digraphs (if the letters are taken consecutively in pairs) the mean or average number of occurrences in this case is $.00148 \times 99 = .147$. Having the mean number of occurrences of the event under consideration, one may now find the answers to these questions: What is the probability that any specified digraph, say XY, will not occur? What is the probability that it will occur *exactly* 1, 2, 3, . . . times? *At least* 1, 2, 3, . . . times?

(3) Again the probability curves may be used as before, for the type of distribution is the same. The following values are obtainable by reference to the various curves, using the mean value $.00148 \times 99 = .147$.

100 letters of random text

Frequency (x)	Probability that a specified digraph will occur exactly x times	Probability that a specified digraph will occur at least x times	Probable number of digraphs appearing exactly x times	Probable number of digraphs appearing at least x times
0	0. 86	1. 00	581. 36	676. 00
1	. 13	. 14	87. 88	94. 64
2	. 01	. 01	6. 76	6. 76
3	. 00	. 00	0. 00	0. 00

(4) Thus it is seen that in 100 letters of random text the probability that a specified digraph will occur exactly once, for example, is .13; at least once, .14; at least twice, .01. The probability that a specified digraph will occur at least 3 times is negligible. (By calculation, it is found to to be .0005.)

b. (1) The probability of digraphic coincidence in random text based upon a 26-element alphabet is of course quite simply obtained: since there are 26^2 different digraphs, the probability of selecting any specified digraph in random text is $\frac{1}{26^2}$. The probability of selecting two identical digraphs in such text, *when the digraphs are specified*, is $\frac{1}{26^2} \times \frac{1}{26^2} = \frac{1}{26^4}$. Since there are 26^2 different digraphs, the probability of digraphic coincidence in random text, κ_r, is $26^2 \times \frac{1}{26^4} = \frac{1}{26^2} =$ .00148.

Over

[redacted]

(2) Given a random assortment of 100 letters, what is the probability of occurrence of 0, 1, 2, . . . digraphic coincidences? Following the line of reasoning in paragraph 2*d* (2), in 100 letters the total number of comparisons that may be made to see if two digraphs coincide is 4,851. This number is obtained as follows: Consider the 1st and 2d letters in the series of 100 letters; they may be combined to form a digraph to be compared with the digraphs formed by combining the 2d and 3d, the 3d and 4th, the 4th and 5th letters, and so on, giving a total of 98 comparisons. Consider the digraph formed by combining the 2d and 3d letters; it may be compared with the digraphs formed by combining the 3d and 4th, 4th and 5th letters, and so on, giving a total of 97 comparisons. This process may be continued down to the digraph formed by combining the 98th and 99th letters, which yields only one comparison, since it may be compared only with the digraph resulting from combining the 99th and 100th letters. The total number of comparisons is the sum of the sequence of numbers 98, 97, 96, 95, . . . 1, which is 4,851.[6]

(3) Since in the 100 letters there are 4,851 opportunities for the occurrence of a digraphic coincidence, and since $\kappa_r{}^2=.00148$, the expected number of coincidences is $.00148\times4851=7.17948=7.2$. The various probability curves may now be referred to and the following results are obtained:

Distribution for 100 letters of random text

Frequency (x)	Probability for exactly x digraphic coincidences	Probability for at least x digraphic coincidences
0	0.001	1.000
1	.005	.999
2	.019	.994
3	.046	.975
4	.083	.929
5	.120	.846
6	.144	.726
7	.148	.582
8	.134	.434
9	.107	.300
10	.077	.193
11	.050	.116

c. In this table it will be noted that it is almost certain that in 100 letters of random text there will be at least one digraphic coincidence, despite the fact that there are 676 possible digraphs and only 99 of them have appeared in 100 letters. When one thinks of a total of 676 different digraphs from which the 99 digraphs may be selected it may appear rather incredible that the chances are better than even (.582) that one will find at least 7 digraphic coincidences in 100 letters of random text, yet that is what the statistical analysis of the problem shows to be the case. *These are, of course, purely accidental repetitions.* It is important that the student should fully realize that more coincidences or accidental repetitions than he feels intuitively should occur in random text will actually occur in the cryptograms he will study. He must therefore be on guard against putting too much reliance upon the surface appearances of the phenomena of repetition; he must calculate what may be expected from pure chance, to make sure that the number and length of the repetitions he does see in a cryptogram are really better than what may be expected in random text. In studying cryptograms composed of figures this

[6] The ~~formula for finding the number of comparisons that can be made is as follows, where n=the total number of letters in the sequence and t is the length of the polygraph: No. of comparisons=$\frac{(n-t)(n-t+1)}{2}$.~~

Substitute accompanying footnote.

equal sign

¶ 6 The formula for finding the number of comparisons that can be made is as follows, where *n* = the total number of letters in the sequence and *t* is the length of the polygraph: Since the number of polygraphs possible is *n* − *t* + 1, the number of comparisons is

minus

$$\frac{(n-t+1)\ (n-t)}{2}$$

ctr

ital ital ital ital ital ital

because any one of the *n*-*t*+1 polygraphs may be compared with any one of the remaining *n*-*t* but as a comparison of A with B is the same as a comparison of B with A, the product must be halved.

set

is very important, for as the number of different symbols decreases the probability for purely chance coincidences increases.

d. (1) For convenience the following values of the reciprocals of various numbers from 20 to 36, and of the reciprocals of the squares, cubes, and 4th powers of these numbers are listed:

x	$1/x$	$1/x^2$	$1/x^3$	$1/x^4$
20	0.0500	0.002500	0.000125	0.00000625
21	.0476	.002266	.000108	.00000514
22	.0455	.002070	.000094	.00000429
23	.0435	.001892	.000082	.00000358
24	.0417	.001739	.000073	.00000302
25	.0400	.001600	.000064	.00000256
26	.0385	.001482	.000057	.00000220
27	.0370	.001369	.000051	.00000187
28	.0357	.001274	.000046	.00000162
29	.0345	.001190	.000041	.00000142
30	.0333	.001109	.000037	.00000123
31	.0323	.001043	.000034	.00000109
32	.0313	.000980	.000031	.00000096
33	.0303	.000918	.000028	.00000084
34	.0294	.000864	.000025	.00000075
35	.0286	.000818	.000023	.00000067
36	.0278	.000773	.000021	.00000060

(2) The following table gives the probabilities for monographic and digraphic coincidence for plain-text in several languages.

Language	κ_p	$\kappa_p{}^2$
English	0.0667	0.0069
French	.0778	.0093
German	.0762	.0112
Italian	.0738	.0081
Spanish	.0775	.0093

4. **Data pertaining to trigraphs, etc.**—*a.* Enough has been shown to make clear to the student how to calculate probability data concerning trigraphs, tetragraphs, and longer polygraphs.

b. (1) For example, in 100 letters of random text the value of m (the mean) for trigraphs is $.00005689\times100=.005689$. With so small a value, the probability curves are hardly usable, but at any rate they show that the probability of occurrence of a specified trigraph in so small a volume of text is so small as to be practically negligible. The probability of a specified trigraph occurring twice in that text is an even smaller quantity.

(2) The calculation for finding the probability of at least one trigraphic coincidence in 100 letters of random text is as follows:

$$m=\left(\frac{97\times98}{2}\right)\left(\frac{1}{26^3}\right)=4{,}753\times.0000568912=.2704=.27$$

Referring to curve f_0, with $m=.27$ the probability of finding no trigraphic coincidence is .76. The probability of finding at least one trigraphic coincidence is therefore $1-.76=.24$.

c. The calculation for a tetragraphic coincidence is as follows:

$$m=\left(\frac{96\times97}{2}\right)\left(\frac{1}{26^4}\right)=4{,}656\times.0000021883=.0101=.01$$

Referring to curve f_0, with $m=.01$ the probability of finding no tetragraphic coincidence is so high as to amount almost to certainty. Consequently, the probability of finding at least

strip fig. 9

Over

one tetragraphic coincidence is practically nil. (It is calculated to be .0094 = approximately .01. This means that in a hundred cases of 100-letter random-text cryptograms, one might expect to find but one cryptogram in which a 4-letter repetition is brought about purely by chance; it is, in common parlance, a "hundred to one shot.") Consequently, if a tetragraphic repetition is found in a cryptogram of 100 letters, the probability that it is an accidental repetition is extremely small. If not accidental, then it must be causal, and the cause should be ascertained.

5. An example.—*a.* The message of Par. 9*a* of the text proper will be employed. First, let the repetitions be sought and underlined; then the repetitions are listed for convenience.

```
A.  USYES  ECPMP  LCCLN  XBWCS  OXUVD
B.  SCRHT  HXIPL  IBCIJ  USYEE  GURDP
C.  AYBCX  OFPJW  JEMGP  XVEUE  LEJYQ
D.  MUSCX  JYMSG  LLETA  LEDEC  GBMFI
```

Group	Number of occurrences
BC	2
CX	2
EC	2
LE	3
JY	2
PL	2
SC	2
SY	2
US	3
YE	2
SYE	2
USY	2
USYE	2

b. Referring to the table in Par. 3*a* (3) above, it will be seen that in 100 letters of random text one might expect to find about 7 digraphs appearing at least twice and no digraph appearing 3 times. The list of repetitions shows 8 digraphs occurring twice and 2 occurring 3 times.

c. Again, the list of repetitions shows 10 digraphs each repeated at least twice; the table in Par. 3*b* (3) above shows that in 100 letters of random text the probability of finding at least that many digraphic coincidences is only .193. That is, the chances of this being an accident are but 176 in a thousand; or another way of expressing the same thing is to say that the odds against this phenomenon being an accident are as 807 is to 193 or roughly 4 to 1.

d. The probability of finding at least one trigraphic coincidence in 100 letters of random text is very small, as noted in Par. 4*b*; the probability of finding at least one tetragraphic coincidence is still smaller (Par. 4*c*). Yet this cipher message of but 100 letters contains a repetition of this length.

e. A consideration of the foregoing leads to the conclusion that the number and length of the repetitions manifested by the cryptogram are not accidental, such as might be expected to occur in random text of the same length; hence they must be causal in their origin. The cause in this case is not difficult to find: repeated isolated letters and repeated sequences of letters (digraphs, trigraphs) in the plain text were actually enciphered by identical alphabets, resulting in producing repeated letters and sequences in the cipher text.