PROGRESS IN ATOMIC PHYSICS

NEUTRINOS AND GRAVITATION

THIS BOOK IS FOR PERSONAL USE ONLY
NOT TO BE DISPLAYED IN
LIBRARY OR PUBLIC PLACE

XIIth Moriond Workshop
Les Arcs, Savoie, France - January 25- February 1, 1992

PROGRESS IN ATOMIC PHYSICS
NEUTRINOS AND GRAVITATION
Series : Moriond Workshops

ISBN 2-86332-117-X
Copyright 1992 by Editions Frontières

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

EDITIONS FRONTIERES
B. P. 33
91192 Gif-sur-Yvette Cedex - France

Printed in Singapore by Fong and Sons Printers Pte. Ltd.

Proceedings of the XXVIIth RENCONTRE DE MORIOND
Series : Moriond Workshops
Les Arcs, Savoie, France January 25 - February 1, 1992

PROGRESS IN ATOMIC PHYSICS
NEUTRINOS AND GRAVITATION

edited by

G. Chardin
O. Fackler
J. Trân Thanh Vân

THIS BOOK IS FOR PERSONAL USE ONLY
NOT TO BE DISPLAYED IN
LIBRARY OR PUBLIC PLACE

S 7637

EDITIONS
FRONTIERES

M 72 1992

XIIth Moriond Workshop on :
Progress in Atomic Physics, Neutrinos and Gravitation

was organized by

Trân Thanh Vân J. *(Orsay)*

with the active collaboration of :

Boehm F.	*(Caltech)*
Chardin G.	*(Saclay)*
Damour Th.	*(Bures sur Yvette)*
Fackler O.	*(Livermore)*
Faller J.	*(Boulder)*
Fischbach E.	*(Purdue)*
Fontaine G.	*(Paris)*
Gerbier G.	*(Saclay)*
Giacobino E.	*(Paris)*
Greene G.	*(NBS)*
Kayser B.	*(Washington)*
Mugge M.	*(Livermore)*
Pain R.	*(Paris)*
Petcov S.	*(Sofia/Trieste)*
Spiro M.	*(Saclay)*
Wilkerson J.	*(LANL)*

Sponsored by :

. Centre National de la Recherche Scientifique (IN2P3, MPB)
. Commissariat à l'Energie Atomique (DAPNIA)
. National Research Foundation

Foreword

The XIIth Moriond Workshop was held at Les Arcs, Savoie (France) from January 25 to February 1, 1992. It was attended by 106 participants from 15 countries.

The main purpose of the Rencontres de Moriond is to discuss recent developments in contemporary physics and also to promote effective collaboration between experimentalists and theorists in similar fields. By bringing together a relatively small number of participants we hope to develop personal contacts as well as a more thorough and detailed discussion of the scientific contributions in an informal and friendly atmosphere.

The scientific program of this workshop focused in 3 main subjects : atomic physics, neutrino physics and gravitation. Special emphasis was given to solar neutrinos, neutrino mass and test of fundamental principles in atomic physics and gravitation. The participation of physicists from various domains : astrophysics, atomic physics, particle physics and solid state physics allowed exchange of knowledge and technology between different fields and confered to this workshop a frontier physics character.

The success of this workshop was due to the active participation of all participants, to the lively discussions following lectures, and to the efforts of the speakers in providing pedagogical and enlightening lectures.

I would like to thank all participants and programm committee members : F. Boehm, G. Chardin, Th. Damour, O. Fackler, J. Faller, E. Fischbach, G. Fontaine, G. Gerbier, E. Giacobino, G. Greene, B. Kayser, M. Mugge, R. Pain, S. Petcov, M. Spiro and J. Wilkerson for preparing and organizing this workshop.

I am also grateful to Mr and Mrs J. M. Milan and Ms R. Chenal for their hospitality at les Arcs, and to the Conference secretaries : L. Besson, J. Boratav, L. Norry, J. Raguideau, D. Soudry, D. Vernet and A. C. Tran Thanh Van who have devoted much of their time and energy to the success of this meeting.

The 1992 Moriond workshop is sponsored by the Centre National de la Recherche Scientifique, the Commissariat à l'Energie Atomique and the National Science Foundation. We are deeply grateful for their financial support.

J. Trân Thanh Vân

Contents

III. Physics with Neutrons and Atoms

IV. Physics and Computers

V. Tests of Special and General Relativity

VI. Tests of Newtonian Gravity

Solar Neutrinos

THE CLASSICAL SOLAR MODEL AND BEYOND THIS FRAMEWORK

S. Turck-Chièze
DAPNIA, Service d'Astrophysique
Centre d'Etudes de Saclay
91191 Gif sur Yvette Cedex 01, FRANCE

Abstract

From the comparison of recent updated solar model predictions, the extreme sensitivity of the ^8B neutrino flux to opacity coefficients, to some specific reactions and to the screening effect is once more illustrated. Errors and too simple prescriptions have been identified in any calculation with only serious consequences on the pp III chain. Considering that these differences are clarified and are useful for a more correct prediction on neutrino fluxes, I show how the specific ^8B neutrino flux is still fluctuating with the recent microscopic improvements: reestimate of the pp reaction, improved opacity calculations, better determination of the solar abundances, new chlorine absorption cross section. This is the reason why we prefer to express our result as a range of possibility to emphasize the still large uncertainties: chlorine prediction between 5.4 - 7.6 SNU, for Kamiokande $\Phi^8\mathrm{B} = 3.7 - 5.5\ 10^6 cm^{-2}s^{-1}$, gallium: 118 - 128 SNU.

In the second part, I briefly recall that, even if the framework of classical stellar evolution is not complete, it seems from the properties of the sound speed, the detection of acoustic modes, and the present studies of extra phenomena that it is difficult to imagine today large variations of the solar structure (several %), this point justifies to consider with interest, results coming from the simplified scheme of classical stellar evolution.

First, I would like to thank, for all of us, the organisers of this workshop who give us the occasion to really discuss between different fields of physics in this so exciting atmosphere which is Moriond. It is true that this beautiful place permits extensive exchanges together with appreciable insights in so different parts of physics. It seems to me that it is mainly due to the effort of organisers to largely let express the different aspects of each field.

I think, we have really a lot of chance to discuss this year, once more, about solar neutrinos in Moriond. In fact, since the results of SAGE experiment last year [1] and before the result of GALLEX experiment (see talk of L. Zanotti), which both attend to measure for the first time the pp neutrinos, this subject is really open: do we approach the properties of the Sun by measuring solar neutrinos on earth or do we approach the properties of the neutrinos? Nobody knows the answer to this question, for the moment, and it seems that everything has been prepared for an answer to this question rather soon.

In this review, I shall take the astrophysical point of view, and after some general considerations on solar modelling and neutrino predictions, I would like to recall what is a solar model in the classical framework of stellar evolution (section 1) . Then I confront independent and updated calculations of solar modelling to examine the astrophysical situation, the helioseismological one (this part will be very brief because it will be extensively reviewed by D. Gough and I. Lopes) and the variety of neutrino predictions (section 2). More and more neutrino experiments are running or funded nowadays, so, it is extremely useful to understand the present neutrino prediction differences to interpret as clearly as possible the experimental results (section 3). Even, in the classical framework of stellar evolution, the solar modelling calculation requires a lot of ingredients and as an active field of investigation, real improvements have been performed in the last three years. I shall comment on them and study the consequences for neutrino predictions (section 4). We are convinced by stellar observations that one needs to go beyond the classical framework to explain a lot of measurements on stars and in particular on the Sun, so I shall examine if one can extract a quantitative estimate of the validity of what one calls the "Standard solar modelling" (section 5). Finally, as a conclusion is not yet possible at the present stage, I shall discuss what could be or not, evident today and how one may perhaps interpret present and near future neutrino measurements on earth...

Introduction

I am not completely convinced that the conventional *"Standard solar model"* would be a really appropriate expression. When it is applied, people usually thinks that it represents, first, consensus and, second, a finality for comparisons with solar neutrino experiments. In fact, it is not completely true. One traditionally, considers a classical stellar evolution of the Sun, with specific but reasonable hypotheses of hydrostatic equilibrium, assuming the absence of rotation and magnetic field in the Sun. This represents the framework astrophysists have already checked for a lot of other stars and other problems with a reasonable amount of success. In this classical approach, one deduces a solar *reference* model which translates not only the quality of theoretical

hypotheses and structure equations but also which absorbs the state of sophistication of different fields as nuclear physics, atomic physics and plasma physics: it is why this *reference* solar model is regularly updated. At the end of such a complex calculation, the first step is reached, but any astrophysicist knows that the second step, which consists in introducing all the forgotten phenomena that we know to exist in the Sun as differential rotation, varying surface magnetic field, mass loss, internal motions, diffusion...must be considered if one would like to give interesting and precise predictions. This is an extremely difficult task to examine the consequences on the solar structure of such clearly identified processes: turbulence, mixing or ..?, and secondly to follow such induced effects in time and space.

Normally, confrontations with observables would begin to be interesting only after this second step, which is not yet reached. The question we need effectively to answer, is the following: are we convinced that, even at the end of the first stage, when the classical picture of stellar evolution is used, the neutrino predictions are reasonably well representative of what we hope to detect on earth? The parallel development of helioseismology seems to help us to answer to this question, and in this review, I would like to remind the conclusions of this community.

1. The solar model in the classical framework of stellar evolution

To define the solar structure at the present age (4.5- 4.55 Gyr) with the present luminosity ($3.826\ 10^{33}$erg s^{-1}) and the present radius ($6.9599\ 10^{10}$ cm) one follows the spatial and temporal evolution of a star of $1M_{\odot}$ from the onset of hydrogen burning where the star is considered as totally convective. To do so, one solves the four stellar structure equations which govern the hydrostatic equilibrium, the mass conservation, the energy transport, and the temperature gradient [2]. To quantify the results, a precise determination of the initial composition (which is assumed to be determined by the present photospheric one) is an essential ingredient. This is not only true for the determination of the nuclear energy (knowledge of all the species up to oxygen to calculate pp chain and CNO cycles), but also for the determination of the opacity coefficients where species up to iron have important contributions (due to bound-bound and bound-free processes). In addition to this crucial ingredient, all the nuclear reaction rates must be estimated from laboratory experiments and the opacity coefficients calculated for precise range of temperature density and composition.

In the classical solar calculation, only two parameters are adjusted: the initial helium content which is unmeasurable in the photosphere and the mixing length parameter which is the only parameter which plays an important role when convection and radiation are competing processes (mainly near the photosphere). These two parameters can be determined due to the exceptional knowledge of the present luminosity and radius of the Sun coupled with the knowledge of the age of this star.

2. Confrontation of independent and updated calculations of solar modelling

A lot of solar models have been calculated in the last years, let discuss the most updated ones which have been recently published [3],[4], [5], [6]. Table 1 summarizes

the main astrophysical results which are the central thermodynamical quantities, the temperature at the basis of the convective zone (T_{BCZ}), the initial helium content Y and the mixing length parameter α. One notices that the structure of the different models is very similar (within less than 1% in the central part). This is not really true for the case of the Cox et al calculation, where explicit adjustment to the heliosismological data has been applied to reproduce the real behaviour of the sound speed at the basis of the convective zone and where absence of coulomb correction in the equation of state has important consequences on density, temperature and neutrino fluxes. The agreement between calculations is mainly due to the fact that these independent models use practically the same updated values of nuclear reaction rates, composition and opacity coefficients. The initial helium content obtained is also rather consistent with other helium measurements in different astrophysical sites and with chemical evolution of galaxies (see more detailed comments [7],[8]).

	Bahcall and Ulrich 1988 [3]	Turck-Chièze et al 1988 [4]	Cox et al 1989 [5]	Sackmann et al 1990 [6]
T_C	15.62 10^6 K	15.52 10^6 K	15.68 10^6 K	15.42 10^6 K
ρ_C	148.	147.2	162.4	146.6
T_{BCZ}	2.11 10^6 K	2.04 10^6 K	2.29 10^6	1.96 10^6 K
Y	0.271	0.275	0.291	0.278
α		2.11 *	1.89	2.07

Table 1 : Rough comparison of recent reference solar models. The model of Cox et al, particularly adapted for seismological calculations has been fine tuned at the bottom of the convective zone with an opacity increase of 15 to 20 %. Their equation of state leads to a pressure slightly too large, thus a higher helium abundance (overestimation of 0.01 by the authors) and higher predicted neutrino fluxes due to the absence of coulomb correction. *: the published value was 1.6 due to an error in the hydrogen equation of state.

The rapid development of the helioseismological investigation gives us an idea of the real sound speed behaviour (see talk of D. Gough). Even if the central errors are still large, it seems that the agreement with the prediction is within 2%, except in the central part where deviations of more than 2 or 3% could be possible. Between theory and observations of acoustic modes, deviations of the order of 10 μHz on 3000 μHz are noticed (see for more details the talk of I. Lopes).

The solar model predicts the radial profile of the production rate of each species of neutrinos (see table 3 for the reactions involved). A comparison with detection of neutrinos on earth needs a complementary estimate of the energy dependent interaction of neutrinos with the different detectors and the corresponding absorption cross sections. Table 2 shows the comparison between experiments and theoretical predictions of the solar calculations discussed above.

experimental results	Theoretical	Predictions	
	Bahcall-Ulrich [3]	Turck-Chièze et al [4]	Sackman et al [6]
Chlorine experiment			
2.33 ± 0.25 (SNU)	7.9 ± 2.5 (3σ)	$5.8 \pm 1.$ (1σ)	7.7
ratio $\frac{exp}{th}(R_H)$	$(29.4 \pm 11 \pm 10)\%$	$(40 \pm 11\ ^{+7}_{-7})\%$	$(30.3 \pm 11)\%$
Kamiokande experiment			
0.28 ± 0.03 ev/d ($E \geq 7.5$ MeV)	0.62 ev/d	0.41 ev/d	
ratio $\frac{exp}{th}$ (R_K)KII	$(46 \pm 5 \pm 6 \pm 17)\%$	$(70 \pm 8 \pm 9\ ^{+15}_{-13})\%$	
	(th. err: 3σ)	(th. err:1σ)	
ratio $\frac{exp}{th}$ (R_K)KIII(preliminary)	$(65 \pm 9\ ^{+18}_{-15})\%$	$(88 \pm 14^{+19}_{-16})\%$	
Gallium experiments			
SAGE : <55 SNU (68%)	132^{+20}_{-16}	125 ± 5	125
< 79 SNU (95%)			
Gallex : (in 1992)	$(3\ \sigma)$	$(1\ \sigma)$	

Table 2 : Comparison of the experimental neutrino results [9], [11], [1] with the predicted calculations. The ratio experiment/theory includes the experimental errors (statistical or statistical and systematic) and the theoretical ones. The Kamiokande results are partially deduced from [8], and theoretical estimates include a correction for detector efficiency [10]. The comparison with the Kamiokande experiment (where only statistical error appears in left column) is obtained from [11] (1 SNU= 10^{-36} captures/atom/s).

Looking at table 2, two important remarks may be done.

- Very often, people says that Kamiokande experiment confirms the disagreement noticed between the chlorine experiment and the prediction. In fact, if the experiment seems to present a deficit, its order of magnitude seems to be smaller than in the case of the chlorine detector. If it is confirmed (the errors are still large!), this is *incompatible* with astrophysical interpretation, because in this context, the water experiment (only ^8B neutrinos detected) is a subset of the chlorine one ($^8B\ \nu + ^7Be\ \nu$ + small other contributions). Any modification of the solar internal structure will lead to a greater effect on ^8B, which varies as T^{18}, than on any other neutrino sources. So the possible solutions, if confirmed, could be the experimental precision, or some uncertainty on absorption cross section, (the neutrino electron scattering appearing much easy to estimate) or neutrino properties.

- Very similar models, from the astrophysical point of view, lead to significantly different predictions. To analyze this point in details, one needs first to come back to the respective neutrino fluxes (table 3) and to the corresponding errors (table 4).

Neutrino sources	Bahcall and Ulrich [3]	Neutrino fluxes Turck-Chièze et al [4]	$(10^{10} cm^{-2} s^{-1})$ Sackman et al [6]
pp	6.0 (1 ± 0.02)	5.98 $(1. \pm 0.03)$	6.00
pep	$1.4\ 10^{-2}$ (1 ± 0.05)	$1.30\ 10^{-2}$	$1.29\ 10^{-2}$
^{7}Be	$4.7\ 10^{-1}$ $(1. \pm 0.15)$	$4.18\ 10^{-1}$	$4.23\ 10^{-1}$
^{8}B	$5.8\ 10^{-4}(1. \pm 0.37)$	$3.83\ 10^{-4}(1._{-0.21}^{+0.18})$	$5.8\ 10^{-4}$
^{13}N	$6.1\ 10^{-2}(1. \pm 0.5)$	$6.27\ 10^{-2}$	$3.99\ 10^{-2}$
^{15}O	$5.2\ 10^{-2}(1 \pm 0.58)$	$5.60\ 10^{-2}$	$3.09\ 10^{-2}$
^{17}F	$5.2\ 10^{-4}(1. \pm 0.46)$		$4.23\ 10^{-4}$

Table 3 : Theoretical predictions for the neutrino fluxes coming from the different sources .

uncertainties sources	uncertainty	gallium detector	uncertainties chlorine detector	water detector
pp reaction	3%	2.7 %	6%	8 %
^{3}He ^{3}He reaction	5 %	< 0.1 %	1.6 %	2 %
^{3}He ^{4}He reaction	4 %	1 %	2.7 %	3.3 %
^{7}Be p reaction	15 %	1 %	13 %	15 %
Luminosity	0.5 %	0.3 %	3.6 %	3.6 %
Age	1 %	< 1 %	1 %	1.5 %
Opacity	$^{+6}_{-11}$ %	3%	$^{+4}_{-8}$ %	$^{+5}_{-10}$%
Absorption cross section		4%	3%	< 1%
Total		6%	$^{+16}_{-17}$ %	$^{+18}_{-21}$ %

Table 4 : Estimated uncertainties for the different present or near future experiments, the errors correspond to 1σ estimate.

One notices that pp neutrino flux is in excellent agreement, mainly due to the constraint on the luminosity and the same estimated cross section used. A small difference is observed on pep flux, without consequence on the global prediction for the different detectors because of the small contribution of the reaction. The 10% difference on ^{7}Be neutrino flux is compatible with the difference noticed on the temperature ($\Phi^{7}Be$ proportional to T^{8}). Small differences are also noticed on neutrinos coming from the CNO cycle but, once more the global effect on the different detectors is small: this difference is easily explained by different reaction rates used (see [3], [4], [6]).

The largest difference is noticed on ^{8}B neutrino prediction where difference of the order of 30% is observed between Turck-Chièze et al calculation and the two others while the apparent agreement between Bahcall and Ulrich calculation and Sackman et al one is in contradiction with the difference of 1.5% on central temperature. These points have been attentively studied first by Turck-Chièze [2] [13], then very recently by A. Boothroyd ([14] and by Lopes and Turck-Chièze [7]).

3. Understanding the neutrino prediction differences

The possibility of large differences in the predictions on ^8B neutrino flux could be understood by the fact that:

- the $^7Be(p,\gamma)^8B$ cross section is not well established;
- the ^8B neutrino flux is extremely dependent of the temperature (see above),
- there is no constraint coming from the present solar luminosity, on the ppIII chain;

- As it appeared in table 4, the uncertainty on opacity coefficients has a large impact on the ^8B neutrino flux. So, we have noticed that subtle differences in the different calculations have large influence on high energy neutrino fluxes without subtantial ones on the structure of the Sun or on the helioseismological variables. The conclusions are the following: a small difference in composition may have noticeable impact on the opacity coefficients calculations even if the different authors use the Los Alamos library (in the central part, nearly 40% of this coefficient comes from heavy elements and 20% from iron). Added to this point, small errors were still existing: small spurious enhancement due to the treatment of CNO modification of the opacity coefficients in the central part in Bahcall and Ulrich estimate, small artificial underestimate of the electron collective effect in the Turck-Chièze et al one (following the too crude estimate [15]) without *double counting* [7]. All these differences may explain the 0.7 % difference in temperature and the consequent 13 % difference in the ^8B neutrino flux.

- a small difference in the (^3He,^3He) reaction rate due a correction of electron screening in the laboratory experiment may explain less than 1 % underestimate of the Sackman et al and Turck-Chièze et al ^8B neutrino flux.

- Turck-Chièze et al do not use the same prescription for the $^7Be(p,\gamma)$ cross section. In view of the crude theoretical extrapolation of Tombrello, they use a more complete calculation of Barker which leads to a decrease of 15% on ^8B neutrino flux (see detailed discussion in [8]).

- finally A. Boothroyd discovered a more subtle difference which is first the different treatment of screening in the solar plasma (weak treatment in Bahcall and Ulrich and Sackman et al calculations following Salpeter [16] and intermediate treatment in Turck-Chièze et al following Graboske et al [17]). This leads to a 12% enhancement of the ^8B neutrino flux in these two first calculations (see table 5). This kind of correction was also used by error in the case of electron capture rate on ^7Be in the Sackman et al calculation leading to an artificial increase of 15% on Φ^8B and consequently explaining the artificial agreement between Bahcall and Ulrich and Sackman et al predictions.

These differences, when explained, show clearly the amplification of some rather subtle effects and point out the improvements needed in the classical framework of stellar evolution for performing precise neutrino predictions: precise determination of the initial composition, the importance of the determination of the $^7Be(p,\gamma)^8B$ for solar high energy neutrinos detection, importance of a correct treatment of screening at the nuclear reaction level. The importance of such effect was previously pointed out also for the opacity [18], and for the equation of state [19].

The errors discussed in table 4 and included in table 2 for Turck-Chièze et al solar calculation are supposed to point out the main sources of uncertainties to encourage

improvements. In the case of gallium prediction there is not an identified source of uncertainty which would dominate the others. For the other detectors, the main actors seem: pp reaction rate (see next section), the $^7Be(p,\gamma)$ experimental cross: 15 % unaccurate due to the fact we add experiments performed in very different conditions and for which the theoretical extrapolation and the screening effect is not still so clear, the opacity coefficients for which the error is partly due to the calculation, partly due to the variation of the composition still possible. We do not discuss here the error on screening factor which may be not negligeable for nuclei with Z greater or equal to 4.

4. The implications of recent improvements on solar neutrino predictions

Recently, noticeable improvements appear in the litterature. They concern, first, the most abundant elements: spatial infrared measurements have improved the determination of the CNO abundances: C/H= 3.98 10^{-4} (increase of 10%), N/H= 1. 10^{-4} (decrease of 10 %) , O/H= 7.24 10^{-4} (decrease of 15 %) in fraction number. Reduction of the order of 10 % of the opacity coefficients may be estimated. The consequent influence on neutrino flux is not yet studied and must be analyzed taking into account the microscopic diffusion.We have studied another composition effect which is the reestimation of the photospheric iron. In 1989, Anders and Grevesse had pointed out the discrepancy of 30% between the photospheric value and the meteoritic one obtained with a detection of neutral iron (5% of the total iron). The consequences on neutrino fluxes had been examined by Courtaud et al [12]. This measurement has been reexamined and completed with a determination coming from ionized iron. The conclusion seems to lead to an agreement with the lower meteoritic one: Fe/X = 3.24 10^{-5} [21], [22]. Following table 5, this leads to a decrease of 15 % of the ^8B flux.

The pp reaction has been reestimated. Gould and Guessoum propose an astrophysical factor S(0) = 4.21 10^{-15}keV - b corresponding to an increase of 2.2 % [23].

A more precise (^3He,^3He) astrophysical factor of 5.24 MeV.b is proposed due to a proper laboratory screening correction [20].

Concerning $^7Be(p,\gamma)^8B$ cross section the theoretical extrapolation has been recalculated by Johnson et al [24] who confirm the estimate of Barker but recommend to use only low energy cross section measurements. Due to systematic error between the two experiments concerned, we consider that the error is still large (15%) and largely contribute to the error on ^8B neutrino flux. We have adopted the new astrophysical factor of 0.0225 keV.b.

The absorption chlorine cross section has been reestimated, following a beta experiment, which is an improvement in comparison with the p,n estimate. This leads to an increase of 3% of the absorption chlorine cross section with a reduced error which is now only 3% [25].

With all these new improvements, we have recalculated new solar reference models which are summarized in table 5. Even if the intermediate screening is justified for the central solar conditions, some surprising behaviour of the prescription [17] may in fact favour the Salpeter prescription [16] (see [7] for details on this point and the other microscopic behaviours).

As was shown by Pinsonneault (see corresponding talk), some other estimate of the

pp reaction, differing from Gould and Guessoum by the calculation of the matrix element (5%), has been done by Carlson et al [26]. In fact in these two recent calculations, they mainly concentrate on meson exchange currents and radiative correction but such unconsistency on the central part of the calculation let a confusion at a level of 0.5 SNU in chlorine experiment (see table 5). Moreover, the introduction of Livermore opacity coefficients [27] also increases the estimates by no far from 1 SNU for the chlorine experiment and 8% for the Kamiokande experiment. This effect could yet be compensated by the decrease of the opacity due to the CNO reestimate. It is why we have not yet introduced such opacity coefficients variations.

	IS, GG pp	WS GG pp	WS Cpp
	Fe_{low}	Fe_{low}	Fe_{low}
T_C (10^6K)	15.33	15.33	15.41
ρ_C (g/ cm^3)	142.8	142.8	144.3
$Y_{initial}$	0.2664	0.2664	0.2661
$\Phi\ pp(10^{10}cm^{-2}s{-}1)$	6.08	6.05	6.02
$\Phi\ pep(10^8cm^{-2}s{-}1)$	1.30	1.30	1.37
$\Phi\ ^7$Be ($10^9cm^{-2}s^{-1}$)	3.91	4.08	4.37
$\Phi\ ^8B(10^6cm^{-2}s^{-1})$	3.51	4.28	4.59
$\Phi\ ^{13}N$ ($10^8cm^{-2}s^{-1}$)	2.98	3.40	3.73
$\Phi\ ^{15}O$ ($10^8cm^{-2}s^{-1}$)	2.37	2.82	3.17
chlorine detector	5.22 SNU	5.87 SNU	6.60 SNU
water detector ($10^6cm^{-2}s^{-1}$)	3.51	4.28	4.59
gallium detector	116 SNU	119 SNU	123 SNU

Table 5: Evolution of the different predictions for low value of iron and Los Alamos opacity in case of intermediate (Graboske et al: IS) or weak (Salpeter: WS) screening and for the two reaction rates of pp (Gould and Guessoum: GGpp) and Carlson et al (Cpp). In the last column we have also considered a low value S(0) = 5 MeV.b for (^3He,^3He).

So, the predictions for chlorine and Kamiokande are still fluctuating by 15- 20 % for very different and independent reasons; and a reasonable range of possibility would be 5.4- 7.6 SNU for chlorine and 3.7 to 5.5 $10^6cm^{-2}s^{-1}$ for ^8B neutrino flux, including Livermore opacity improvements and large uncertainty of $^7Be(p,\gamma)$ cross section.The gallium prediction is really more stable at a level of 5% (118- 128 SNU). Let us remark that the lower gallium prediction in comparison with the previous values ([4]) comes mainly from the modification of the nuclear reaction rates of the CNO cycle.

5. Beyond the classical framework

Normally, as was expressed in the introduction, there is no reason to stop the calculation at this point. In fact, more and more precise measurements on the Sun or on other stars demonstrate that several processes are not yet included in the classical

framework of stellar evolution. Their consequences on central conditions are extremely difficult to estimate. It is not the place here to do a review on this very active field. I would like only to summarize the present ideas and illustrate the constraints given by the helioseismology. In this aim, figure 1 proposes a comparison between the radial profile of the sound speed deduced from the observation of acoustic modes with that calculated from a reference model or from a more sophisticated model of the Sun.

If one would like to quantify other phenomena, one sees the necessity to express as correctly as possible the first step which is the result of the classical stellar evolution framework. Figure 1a illustrates the squared difference between a typical reference model [7] and the "real" solar sound speed deduced by Christensen Dalsgaard. The difference is surprisingly small and partly understood in the vicinity of the basis of the convection zone. Nevertheless, it seems (the error in this part is still rather large), that the central part is not yet well under control. So, let us summarize the different idea already explored.

From the observation of photospheric lithium, one is convinced that present lithium is about 1% the value normally estimated in the classical framework. Lithium burns at rather low temperature : 2.4 10^6 K but without diffusion processes one cannot modify the photospheric value from the beginning of the solar life because the convective external part is too thin. Of course it could be possible to introduce an ad hoc parameter to reconciliate the observation but what is interesting is to understand the origin of such observation to be able to follow the time dependence of such phenomena in order to reproduce the observations for different stellar ages and masses. From an extensive analysis of a large amount of data, no definite conclusion is still found. People thinks that microscopic diffusion is part of the explanation: in the study of such process, one tries to determine the gravitational settling of the different species. Such process has been estimated for the Sun but to reconciliate the solar observation and the observation of the stars classified Am, one should not forgotten to consider the fact that such a process may be inhibited by turbulence. This turbulence would be due to differential rotation in the stellar interior. This phenomenon has been largely studied in the case of the Sun and helioseismology provides us a determination of the differential rotation from the surface down to 0.4 $R_\odot$. During a long time, one has thought that instabilities generated by the differential rotation could explain together mixing near the basis of the convective zone and in the center, leading to a small observed neutrino flux. In reality, one has shown that turbulence would be certainly inhibited in the central region due to the presence of a composition gradient and one has been convinced by the sound speed behaviour that only small mixing could be perhaps compatible with the helioseismological data (figure 1b).

One begins also to study the influence of the formation of the Sun or the effect of mass loss, but one is pratically convinced that this early phase does not largely influenced the present time. Large initial mass loss is ruled out [29] (figure 1b).

The effect of cosmions has been studied to find a solution to the neutrino problem, but definitely this hypothetical particle does not seem finally to explain all the observations [31](figure 1c).

Any study of such extra phenomena has shown us that all the plausible effects

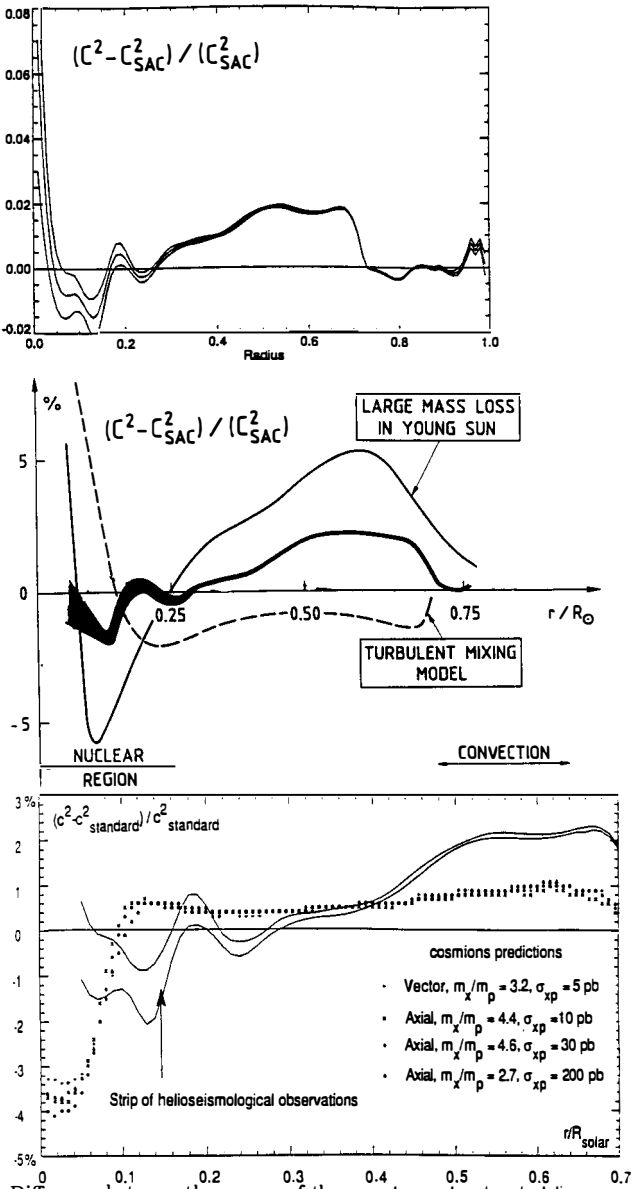


Figure 1: Difference between the square of the sound speed extracted from acoustic modes and the sound speed obtained in a solar model. a) corresponds to the reference model [7], b) to the effect of 1 M $\odot$ mass loss [29] or turbulent mixing [30], c) to different prescriptions for cosmions [31].

must have few consequence on the solar structure, and very often less than we thought previously.

Now the present orientation would be to take into account the presence of acoustic and gravity waves and their consequences on the different species or on the solar structure. This is a very difficult task which couples dynamical time scale and evolutionary time scale. We are not in a situation to deduce from such studies a quantitative modification of the corresponding neutrino fluxes. The only possibility nowadays is to build different ad hoc models consistent with the sound speed behaviour (as shown in figures 1b,c) and examine the consequences on the structure. A more interesting way is to estimate directly from the helioseismological data, information on the internal structure. There is not yet a consensus on this point but presently, people believes that only a modification smaller than 30 % in the high energy neutrino fluxes would be compatible with helioseismology [28].

It is only in this context that we may consider that a comparison between observed solar neutrino fluxes and classical theoretical predictions may be reasonably done as a first step with a precision of may be 30 %. In reality, the sensitivity to the extreme interior of the acoustic modes already detected, is not sufficiently large to be more precise. If a 10 % variation of the temperature seems practically excluded, 1 or 2 % variation is extremely difficult to check [32].

6. What is perhaps established?

I have tried to show that solar modelling is always a very active field due to a lot of improvements in the observations, experiments and calculations and due to the stress of incoming experiments. Present uncertainties, in the classical framework, have been identified and improvements have been already obtained in the carbon, nitrogen, oxygen, iron abundances, the theoretical extrapolation of $^7Be(p,\gamma)^8B$ cross section, the opacity calculations, the determination of the absorption cross section on chlorine. The situation of the pp reaction is surprising and slightly confused, as it seems that the corrections are better understood than the spin matrix element, this point could certainly be clarified very soon. One hopes that the $^7Be(p,\gamma)^8B$ cross section will be remeasured to reduce the error on the classical ^8B prediction neutrino flux. The importance of introducing correctly the effect of screening at different levels of the calculation is now evident for all the astrophysical community but of course a comparable plasma is not yet easy to interpret in the laboratory with laser experiments and definitive prescription is not yet achieved.

From the recent improvements, new solar predictions have been proposed in section 4. Several effects: pp reaction rate, $^7Be(p,\gamma)^8B$ astrophysical factor, opacity coefficients could yet individually modify the ^8B prediction by 15-20 %. Instead large efforts have been done to clearly identify the most important sources of uncertainties, the accuracy of the predictions is not sensibly improved yet and the predictions not significantly modified, but due the well identified ambiguous points, the accuracy could progress rather soon. Statistically, as independent effects, important changes may cancel each other but this is not always the case and consequently an astrophysical explanation to the discrepancy between experiment and predictions for the detection of high energy

neutrino may not be ruled out. However the situation is rather different for pp neutrinos, and it is extremely difficult to reconciliate the solar prediction on gallium with the SAGE results and one waits for an increase of statistics and GALLEX result before definitive conclusion. If the gallium experiments lead to a capture rate smaller than 70-80 SNU, one would orientate the efforts on the neutrino properties, if not neutrino detections will help in the understanding of the central properties of the stars with an extreme accuracy. In any case, these active researches are justified by the improvements in the understanding of such important laboratory which is the Sun and its consequences on particle physics and astrophysics.

This work has been done with Ilidio LOPES and will be largely comment in [8] and detailed in [7]. We would like to thank A. Boothroyd, J. Christensen - Dalsgaard, J. Provost and C. Lagrange for fruitful discussions.

References

[1] V.N.Gavrin, Nucl. Phys. B (Proc. Suppl.) 19 (1991) 84.

[2] S. Turck-Chièze, 1990, New and Exotic phenomena 90, ed. by O. Fackler and J. Tran Thanh Vân, Editions Frontieres, p. 571.

S. Turck-Chièze, in 5TH Force and Neutrino Physics, ed O. Fackler and J. Tran Thanh Van, Ed Fontières, (Gif sur Yvette), p 165.

[3] J.N.Bahcall and R.K.Ulrich, Rev. of Mod. Phys. 60 (1988) 297

[4] S.Turck-Chièze, S. Cahen, M. Cassé and C. Doom, Ap. J. 335 (1988) 415.

[5] A. N. Cox, J. A. Guzik and R. B. Kidman, Ap. J., (1989), 342, 1187.

[6] I. J. Sackmann, A. I. Boothroyd and W. A. Fowler, Ap. J. 360 (1990) 727.

[7] I. Lopes and S. Turck-Chièze, Astrophys. J., (1992), to be published.

[8] S. Turck-Chièze, W. Däppen, E. Fossat, J. Provost, E. Schatzman, D. Vignaud, (1992), Physics Report, submitted.

[9] R.Davis, in Proceedings of the IAU Colloquium 121 "Inside the Sun", Versailles (1989), ed. by G.Berthomieu and M.Cribier, p. 171

[10] K. S. Hirata et al., Phys. Rev. Lett. 66 (1991) 9.

[11] Y. Totsuka, Nucl. Phys. B (Proc. Suppl.) 19 (1991) 69.

[12] D. Courtaud, G. Damamme, E. Genot, M. Vuillemin and S. Turck-Chièze, Solar Phys. 128 (1990) 49.

[13] S. Turck-Chièze, in Proceedings of the IAU Colloquium 121 "Inside the Sun", Versailles (1989), ed. by G.Berthomieu and M.Cribier, p. 125.

[14] A. I. Boothroyd, private communication

[15] J. N. Bahcall, W. F. Huebner, S. H. Lubow, P. D. Parker and R. K. Ulrich, Rev. Mod. Phys. 54 (1982) 767.

[16] E. E. Salpeter, 1954, Australian J. Phys., 7, 353.

[17] H. E. DeWitt, H. C. Graboske and M. C. Cooper, Astrop. J., (1973), 439.

H. C. Graboske, H. E. DeWitt, A.S. Grossman and M. S. Cooper, Ap. J. 181 (1973) 457.

[18] D. B. Boercker, Ap. J. Letters, 316, (1987), L95.

[19] J. Christensen-Dalsgaard and W. Däppen, Astronomyand Astrophysics Review, (1992), submitted.

[20] H. J. Assembaum, K. Langanke and C. Rolfs, Z. Phys. A 327 (1987) 461.

[21] E. Biémont, A. Hibbert, M. Godefroid, N. Vaeck and B. C. Fawcett, Ap. J. 375 (1991) 818.

[22] H. Holweger, C. Heise, M. Kock, Astron. Astrophys. 232 (1990) 510.

[23] R. J. Gould and N. Guessoum, Astrophys. J. 359 (1990) L67.

[24] C.W. Johnson,E. Kolbe, S. E. Koonin, and K. Langanke, 1992, Astrophys. J. (submitted).

[25] A. Garcia, E. G. Adelberger, P. V. Magnus, H. E. Swanson, O. Tengblad and the Isolde collaboration, D. M. Motz, 1991, Phys. Rev. Lett.

[26] J. Carlson, D. O. Riska, R. Schiavilla, and R. B. Wiringa, Phys. Rev. C,(1991), 44, 619.

[27] C. A. Iglesias and F. J. Rogers, Astrophys. J. , (1991), 371, 408.

[28] D.O. Gough and A.G. Kosovichev, in *Seismology of the Sun & Sun-like Stars*, ed. E. Rolfe, ESA SP-286, p.195, 1988.

[29] S. Turck-Chièze, W. Dappen and M. Cassé, in *Seismology of the Sun & Sun-like Stars*' ed. E.J. Rolfe, ESA SP-286, pp. 629-635 (1988).

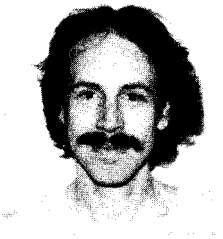
[30] E. Schatzman and A. Maeder, F. Angrand and R. Glowinski, Astron. Astrophys 96 (1981) 1.

[31] J. Kaplan, F. Martin de Volnay, C. Tao and S. Turck-Chièze, Astrophys. J. 378 (1991) 315. and references therein.

[32] S. Turck-Chièze, Proceedings of the Second International Workshop on Theoretical and Phenomenological Aspects of Undergroud Physics, Toledo, 1991, to be published in Nuclear Physics B.

SOLAR MODELS

Marc Pinsonneault
Yale University
Department of Astronomy
P.O. Box 6666
New Haven, CT 06511
USA



Abstract

The current status of the predicted neutrino fluxes from solar models is examined. We first examine the consistency of the predicted predicted neutrino fluxes from different stellar evolution codes with the same input physics; excellent agreement is found between 4 different codes. We then discuss the impact of recent changes in astrophysical opacities and the mixture of heavy elements in the Sun; the increase in the neutrino fluxes caused by increased opacity is largely balanced by the decrease caused by a lowering of the solar iron abundance. The effect of helium diffusion on solar models is investigated, and a deepening of the solar surface convection zone and a modest increase in the predicted neutrino fluxes are found. The current best solar model is found to be in very good agreement with solar properties inferred from helioseismology, including the depth of the solar surface convection zone; any effects not included in standard solar models are therefore unlikely to affect the properties of the outer layers of the Sun. The errors in the neutrino fluxes for the solar model are discussed; the best solar model predicts neutrino fluxes of 8.0±3.0 SNU for the chlorine experiment and 131.5^{+21}_{-17} SNU for the gallium experiment.

1. Introduction

The observed deficit of solar neutrinos relative to the predictions of solar models is an important clue that either the models are in significant error or our understanding of neutrino physics needs to be modified. I will concentrate on the robustness of the solar models in this review. There are two basic sources of uncertainty in the predicted solar neutrino fluxes. Either the ingredients of the standard solar model itself could be incorrect, or some of the physical effects not usually included in the models could be important. I will begin by listing the ingredients of a standard solar model and comparing the agreement between solar models constructed by different investigators using the same input physics (§2). I then discuss both recent changes in the standard solar model itself (§3) and the impact of a physical effect, namely helium diffusion, which is neglected in most solar model calculations (§4). The errors are discussed in §5, and the results are summarized in §6. The work reported here is discussed in more detail in a review paper by Bahcall and Pinsonneault[1], and is based on a joint research effort.

2. Standard Solar Models : Ingredients and Consistency.

Before addressing changes in the solar models, it is necessary to define a "standard" solar model. The initial mass and composition of the Sun must be specified, and the model is then evolved from an initially homogeneous state by solving the equations of stellar structure. A calibrated solar model must reproduce the solar luminosity and radius at the solar age; in practice, this calibration is achieved by adjusting the initial (unknown) helium abundance to reproduce the solar luminosity and the efficiency of convection in the outer layers to reproduce the solar radius. Solving the equations of stellar structure requires knowledge of an equation of state, the efficiency of convective energy transport in the outer layers, and the correct nuclear reaction rates; in addition, the results depend sensitively on the initial solar composition and the opacity of matter to radiation. The latter two are

particularly important for the neutrino fluxes, and have been recently changed; I will discuss the impact of these changes in the next section. For a discussion of the status of the nuclear reaction rates, see the paper by Turck-Chièze[2] in these proceedings.

Given the range of published neutrino fluxes, it is worth investigating how well different authors can reproduce the same neutrino fluxes for the same input physics in the context of a standard solar model. We have therefore systematically changed the input physics for the Yale stellar evolution code, to compare with the neutrino fluxes of Bahcall and Ulrich [3]. We find that the neutrino fluxes of the two codes, for the same input physics, agree to within 0.1 SNU for the chlorine experiment; a similar level of agreement is present for the solar model of Sienkiewicz *et al.* [4]. Although the fluxes reported by Turck-Chièze *et al.* appear significantly lower[5], these lower fluxes are caused by differences in input physics; applying the published partial derivatives of the fluxes with respect to the relative changes in input physics yields a total neutrino flux within 0.1-0.2 SNU of the other results. This excellent agreement, between independently constructed evolution codes built for different purposes, is encouraging and indicates that relative changes in fluxes from carefully calibrated models can be reliably extrapolated to other solar models.

3. Recent Improvements in the Standard Solar Model : Abundances and Opacities.

It is important to ensure that any change in the solar models is consistent with other astrophysical constraints, given the possibility that neutrino physics, rather than the models, is responsible for the solar neutrino problem. The observed non-radial oscillations of the Sun[6]-[7] provide a particularly strong constraint on non-standard solar models. The frequencies of these "p-mode" oscillations provide a measure of the speed of sound in the solar interior, and even provide a measure of the depth of the solar surface convection

zone[8]. Although the p-modes are primarily sensitive to the properties of the outer layers, discrepancies between the observed and theoretical frequencies may nonetheless provide evidence for changes in the solar models which may affect the properties of deeper layers - and thus by extension the observed neutrino fluxes. It is therefore significant that recent improvements in the input physics for solar models provide good agreement between the theoretical and observed p-mode frequencies[9].

The net effect of all the changes in the standard solar models other than those in the initial solar composition, opacities are relatively modest, and can be gauged by comparing the fluxes in the first line of Table 1 below with those of Bahcall and Ulrich. There are two independent sources for the initial solar composition : the relative abundances of heavy elements in meteorites and the solar photosphere. Agreement between the two techniques is in general good[10)-11], with the notable (and important!) exception of iron, which is a significant opacity source in the Sun. The best estimate for the photospheric iron abundance has, until recently, been higher than the meteoritic value; however, this discrepancy has recently be resolved and the best current estimate for the photospheric iron has been reduced to the meteoritic value[12]. A lower iron abundance reduces the opacity at a given temperature and density, and thus reduces the predicted neutrino fluxes (see table below).

However, the opacity for a given composition, temperature, and density has recently been revised, in the sense of a small increase. Until recently, astrophysical opacities have been based on calculations performed at Los Alamos[13]. A group at Lawrence Livermore National Laboratory, led by Rogers and Iglesias, have produced an independent set of calculations which consider more absorption lines and use a more sophisticated quantum mechanical treatment [14]. These opacities increase the predicted depth of the surface convection zone to near the value inferred from helioseismology[9], and

also resolve astrophysical problems in other contexts. As a result, they should be adopted in preference to the older calculations; using these "OPAL" opacities produces an increase in the expected neutrino fluxes (Table 1).

Table 1

The first column is the opacity source (LANL for Los Alamos, OPAL for Livermore). The second is the assumed mixture of heavy elements (Grevesse 1984 for high iron, Anders-Grevesse 1989 for low iron). The third, fourth, and fifth are the predicted B^8 flux, chlorine SNU rate, and gallium SNU rate respectively.

Opacity Source	Mixture	$\phi(B^8)$	Cl(SNU)	Ga(SNU)
LANL	Grevesse(1984)	5.10	7.3	128
LANL	Anders and Grevesse(1989)	4.47	6.4	124
OPAL	Grevesse(1984)	5.97	8.4	134
OPAL	Anders and Grevesse(1989)	5.06	7.2	127.5

Although it is possible to obtain relatively low neutrino fluxes by lowering the iron, the net effect on the neutrino fluxes is minimal because of the compensating effect of the increased opacity.

4. Helium Diffusion in Solar Models.

Heavy elements sink relative to hydrogen by gravitational settling and thermal diffusion[15]. Among major elements, this process is most rapid for helium[16]; because helium is also the most abundant heavy element, its diffusion is also the most important. We model helium diffusion using the method of Bahcall and Loeb[17]. Although helium diffusion is a long timescale process ($\tau \sim 10^{13}$ yr for the solar core), the diffusion of helium into the core increases the central mean molecular weight and temperature enough to cause a modest rise in the expected neutrino fluxes. In addition, the timescale for helium diffusion is shorter in the envelope than in the core[16], which leads to an increase in the hydrogen abundance at the surface. The smaller surface helium abundance causes a decrease in the radius of the

base of the solar model surface convection zone from 0.721 to 0.709 of the total radius. These values for the depth of the convection zone bracket the best estimate of 0.713±0.003 obtained from helioseismology[8]. The change in the surface hydrogen abundance from its initial value also affects our estimate for the total metal content of the Sun because the abundances of elements heavier than helium are estimated relative to hydrogen, and it is usually assumed in stellar models that the surface hydrogen abundance does not change during the stellar lifetime. Therefore, in models with helium diffusion the "true" initial heavy element abundances are higher than the amount expected if the surface hydrogen abundance of the present-day Sun equalled the initial hydrogen abundance. A higher metal abundance increases the total neutrino flux; in fact, this effect roughly doubles the change in the neutrino flux obtained solely from the direct effect of helium diffusion in the solar core. The net effect is a 12% increase in the predicted B^8 neutrino flux and an 11% increase in the SNU rate for the chlorine experiment. Because inhibiting effects, such as rotational mixing, could reduce this effect the solar neutrino fluxes should lie somewhere between the solar models with and without helium diffusion.

5. Errors.

The impact of changes in solar model ingredients on the neutrino fluxes can be gauged by taking the partial derivatives of the different fluxes with respect to the input physics (opacities, mixture, nuclear reaction rates, etc.) and multiplying by the estimated errors in these quantities. The hard part, of course, is determining what the errors are. For the nuclear reaction cross-sections, the experimental errors can be used. However, the opacities and total solar metal abundance (Z) have a substantial impact on the neutrino fluxes and their errors are more difficult to quantify.

For the total solar Z, the published error estimate[11] is ±5% (1σ); this is comparable to the 19% error (3σ) adopted by Bahcall and Ulrich, which we retain. We note that both

the meteoritic and photospheric abundances would have to be in error (in the same sense) to cause a substantial change in the total Z. The opacities used in stellar models are the result of complex quantum mechanical computer calculations, whose accuracy depends on the assumptions taken in creating them. Although large differences have been obtained for low temperature opacities, the new OPAL calculations differ from the earlier LANL opacities by ~2% for solar core conditions; we adopt this as an estimate for the total error. Because it is easier to neglect an opacity source than to overestimate one, a substantial reduction in the opacity (which would reduce the neutrino fluxes) is unlikely. Further small increases, however, are conceivable. The sum of all the individual error contributors in quadrature yields a total 3σ error of 3.0 SNU for the chlorine experiment, 21 SNU for the gallium experiment, and 43% of the B^8 flux

6. Summary.

The challenge posed by solar observations has stimulated considerable progress on the input physics for solar and stellar models. In the case of helioseismology, the new physics has produced significant improvement in the agreement with observations. For the solar neutrino problem, however, all of the various effects have left us almost in the same place where we began; our current best solar model has a predicted chlorine rate of 8.0±3.0 SNU, compared with 7.9±2.6 SNU for Bahcall and Ulrich in 1988. This indicates that there is no systematic trend towards small predicted fluxes as the input physics improves; furthermore, we have verified that different stellar evolution codes give the same neutrino fluxes for the same input physics, indicating that the relative effects of the different changes we have looked at are generally applicable.

The upcoming generation of neutrino experiments will tell us definitively whether the solar neutrino problem is in the solar model or in the neutrino physics. If the solar model is the source of the solar neutrino problem the answer probably lies not in the

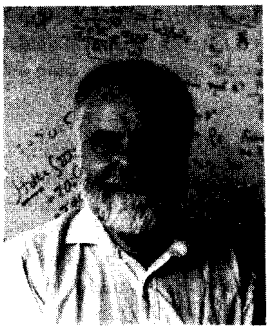
ingredients within the standard solar model but in the effects it neglects. However, any acceptable additional input physics must not destroy the current agreement between other solar and stellar observations and the current generation of models. Work on quantifying the impact of such "non-standard" effects is important, and may ultimately lead to an expansion in the definition of what constitutes a standard stellar model. We have found that helium diffusion, which is a physical effect not usually included in solar models, raises the predicted neutrino fluxes by 11-12% for the Kamiokande and chlorine experiments. It also deepens the solar surface convection zone, in better agreement with that inferred from helioseismology.

References.

1. Bahcall, J.N., and M.H. Pinsonneault, 1992, Rev. Mod. Phys., in press.
2. Turck-Chièze, S., 1992, these proceedings.
3. Bahcall, J.N. and R.K. Ulrich, 1992, Rev. Mod. Phys. **60**, 297.
4. Sienkiewicz, R., B. Paczyński, and S.J. Ratcliff, 1988, Astrophys. J. **326**, 392.
5. Turck-Chièze, S., S. Cahen, M. Cassé, C. Doom, 1988, Astrophys. J. **335**, 415.
6. Libbrecht, K.G., M.F. Woodard, and J.M. Kaufman, 1990, Astrophys. J. Supp. **74**, 1129.
7. Gough, D.O., 1992, these proceedings.
8. Christensen-Dalsgaard, J., D.O. Gough, and M.J. Thompson, 1991, Astrophys. J. **378**, 413.
9. Guenther, D.B., P. Demarque, Y.-C. Kim, and M.H. Pinsonneault, 1992, Astrophys. J. **387**, 372.
10. Grevesse, N., 1984, Physica Scripta **T8**, 49.
11. Anders, E., and N. Grevesse, 1989, Geochim. Cosmochim. Acta **53**, 197.
12. Holweiger, H., A. Bard, A. Kock, and M. Kock, 1991, Astron. Astrophys. **249**, 545.
13. Huebner, W.F., A.L. Merts, N.H. Magee Jr., and M.F. Argo, 1977, in *Astrophysical Opacity Library* (Los Alamos Scientific Laboratory Report LA-6760-M).
14. Iglesias, C.A., and F.J. Rogers, 1991, Astrophys. J. **371**, 408.
15. Michaud, G., Y. Charland, S. Vauclair, and G. Vauclair, 1976, Astrophys. J. **210**, 447.
16. Michaud, G., G. Fontaine, and G. Beaudet, 1984, Astrophys. J. **282**, 206.
17. Bahcall, J.N., and A. Loeb, 1990, Astrophys. J. **360**, 267.

THE ROLE OF HELIOSEISMOLOGY IN NEUTRINO PHYSICS

D O Gough
Institute of Astronomy & Department of Applied Mathematics
and Theoretical Physics, University of Cambridge
Madingley Road, Cambridge, CB3 OHA, UK



Abstract

For the sun to provide a useful laboratory for neutrino physics, we must be sure of the limits on its possible structure that are set by observations. We must also be aware of the assumptions upon which those limits depend. Thus, traditional, so-called standard solar models, which depend on many *ad hoc* assumptions and much uncertain physics, provide only a theoretical guide, albeit an extremely useful one. Helioseismic inversions, which seek constraints on the basic structure subject solely to assuming hydrostatics, are therefore more reliable. However, at present they are less robust, which is as one might expect if the set of possibilities have fewer *ad hoc* limitations. Indeed, unless the tentative negative results from SAGE are accepted as showing that the pp neutrino flux is significantly below theoretical expectation, one cannot yet deduce that the sun provides compelling evidence that nuclear or particle physics, without neutrino transitions, is truly inconsistent with observation.

Introduction

The only direct apparent observational evidence that neutrino transitions take place is offered by the deficiency of observed solar neutrinos with respect to the theoretical production rate. Is that real evidence? From the astrophysical point of view, it can be so only if we are really sure that the theoretical solar models faithfully represent the structure of the sun. If that were the case, the astrophysicist could then conclude that the resolution of the inconsistency must lie in nuclear or particle physics: either the thermonuclear reactions generating the thermal energy and producing the neutrinos have been misrepresented, or neutrinos are not massless, and transitions, such as Mikheyev-Smirnov-Wolfenstein transitions, modify the flux of electron neutrinos during their flight to Earth. Neutrino transitions are currently very fashionable, and are providing amusement to many particle physicists in their quest for a 'solution' to the so-called solar neutrino problem. Connoisseurs who take seriously the suggestion that the solar neutrino flux measured at Homestake from capture by ^{37}Cl is varying with time, possibly in association with the solar cycle, are having even more fun modulating the flux of left-handed neutrinos with helicity flipping by a varying magnetic field, the interactions being contrived in such a way as to have also permitted neutrinos to have escaped from the core of Supernova 1987A.

It is my task here to discuss what we know about the internal structure of the sun, based on a combination of observation and plausible physical reasoning. The purpose is to provide an assessment of the degree to which a solar neutrino problem really exists. To what extent are we really forced into contemplating neutrino oscillations or a modification to the physics of thermonuclear reactions? How likely is it that the misunderstanding lies in the astrophysics? We have had two discussions of the so-called standard theoretical models of the sun. Dr Turck-Chièze has quite rightly pointed out that the models are certainly not standard, and nor is the theory by which they are calculated, though the name seems to have stuck; her sobriquet 'classical' is perhaps more apt, the principal connotation presumably relating to the simplicity of the theory and its acceptance by much of the establishment. But how reliable are the many assumptions that are adopted by the classical theory, and how reliably do the theoretical models incorporate those assumptions? Could a physically more plausible model be constructed that is actually in better agreement with observation? Indeed, what are the constraints set by observation? I shall spend most of my time addressing the last of these three questions, but before doing so I shall first briefly remark here on the other two, and return to them, also quite briefly, later.

Many of the assumptions of classical solar theory have been mentioned by the previous two speakers. Broadly speaking, the idea is to adopt the simplest assumptions from a macroscopic point of view, and then to compute models as accurately as possible using the best microscopic physics. Thus one assumes the sun to have evolved spherically symmetrically from a chemically homogeneous initial state for a time $t_\odot$ whose value, about 4.5×10^9y, we believe we know. We assume that, except in the outer 30 per cent by radius where the stratification is unstable to convection, no motion takes place, and that there is neither segregation nor, in the chemically inhomogeneous core, homogenization of chemical elements. What we have learned from the previous two speakers is that even with such apparently simple assumptions there is considerable diversity amongst the values of the neutrino fluxes that are predicted. Dr Turck-Chièze and her collaborators, for example, using somewhat different microphysics, such as the treatment of electron screening in the thermonuclear reaction rates, and apparently taking more care over the precision of the numerical solution of the governing differential equations, obtains a rather lower value of the ^{8}B neutrino flux than the widely quoted standard value of about 6 snu

obtained by Bahcall and Ulrich[1]. Indeed, the total neutrino flux from the latest Saclay model is hardly significantly at variance with the Kamiokande-II measurements. So is there really a neutrino problem?

Dr Pinsonneault has mentioned the effect on theoretical neutrino fluxes of considering additional physical processes not normally taken into account in the classical models. Rotation, for example, modifies the balance of forces and induces a slow circulatory flow which advects the products of the nuclear reactions; both modify the reaction rates. Dr Pinsonneault has ignored the latter, taking into account only the influence on hydrostatic balance of the centrifugal forces. He finds that, for an assumed uniform angular velocity equal to the photospheric value at the equator, namely $2.86 \times 10^{-6} s^{-1}$, the neutrino flux is reduced by 2 per cent. This is most surprising, because the value of the ratio of centrifugal to gravitational acceleration in the equatorial plane is only $\lambda \simeq 2.1 \times 10^{-5}$ at the photosphere, decreasing monotonically inwards to about 1.8×10^{-7} at the centre of the sun. Scaling arguments suggest that, despite the high temperature sensitivity of particularly the ^{8}B neutrino production rate: $\epsilon_{\nu 8} \propto ((1-X)/(1+X))T^{20.5}$ per unit mass at fixed energy-generation rate $\epsilon \propto X^2 \rho T^4$, where X is hydrogen abundance, T temperature and ρ density, the relative modification $\delta F_\nu / F_\nu$ to the neutrino flux (as measured by capture by ^{37}Cl) is only about $+0.8\lambda$ at fixed heavy-element abundance Z. The magnitude of the modification is only 10^{-3} of Pinsonneault's result. Moreover, the argument predicts an *increase* rather than a *decrease* in F_ν. That this should not really be surprising can be appreciated by recognizing that because centrifugal force opposes gravity, the influence of rotation is roughly similar to decreasing the gravitational constant G, which, according to homology scaling[2] increases F_ν. Owing to the sensitive dependence of the photon luminosity L on G, any decrease in G must be offset by a correspondingly larger decrease in X (at fixed Z) to maintain the radiative flux beneath the convection zone; ρ and T are therefore increased in the core to maintain the rate of the nuclear energy generation, and consequently the ^{7}Be and ^{8}B neutrino fluxes rise. (At fixed Z/X, the increase is about 70 per cent of that at fixed Z.) Centrifugal force acts preferentially in the outer layers of the star, and produces an even greater increase in F_ν. It is only when the core rotates very rapidly compared with the rest of the star, so that the direct relief of pressure locally permits a decrease in central temperature, that F_ν is decreased. It is evident that if Dr Pinsonneault were to be correct, something would be severely amiss with the scaling arguments; and that would certainly need to be understood. If, on the other hand, he were not, this issue would provide an illustration of a complicated numerical computer programme to calculate theoretical solar models yielding misleading results. In either case it upholds the general statement that reports of the properties of even neoclassical models cannot unreservedly be trusted. With regard to the specific issue in question, stellar physicists must welcome Dr Pinsonneault's intention to check his calculations. Even though neither of us think it likely that centrifugal force alone will directly influence F_ν significantly, to uncover an error in the scaling arguments would reveal a serious flaw in our conception of stellar structure.

Another point that neither of the two previous speakers mentioned is that classical solar models are unstable. Therefore, they cannot accurately represent reality. The instability cannot render the models wholly unrepresentative, however, because the theory has very successfully reproduced a very large body of astronomical observations; it has provided us with a model of much of the life-history of a wide range of stars. However, by just how much the nonlinear development of the instability vitiates the predictions of relatively subtle properties such as the solar neutrino flux is a matter of some dispute. There are suggestions in the literature that the instability could lead to motion in the energy-generating core of such a nature as to reduce the ^{7}Be and ^{8}B neutrino fluxes to levels consistent with the ^{37}Cl and Kamiokande-II detection rates. Thus one can certainly entertain the idea that models more plausible than the classical models

might be in better agreement with observation. I should point out immediately, however, that the major relevant difference between such models and the classical models is in the relatively unimportant ppII and ppIII branches of the proton-proton chain of nuclear reactions. The main reactions are hardly modified. Therefore, if the suggestion from the Soviet-American Gallium Experiment (SAGE) of a very low pp neutrino flux is verified, by GALLEX, for example, the model I have just described would no longer agree with observation. Indeed, there would then certainly be no plausible solar model that is both consistent with observation and which relies on generally accepted nuclear physics with no neutrino transitions.

Let me return to the classical theory for another brief comment. It was emphasized by Dr Pinsonneault that the strength of the classical theory lies not only in its ability to explain observations of other stars, but also in its robustness. This is demonstrated particularly with calibrations against the wealth of seismic data that are now available. Moreover, as the microphysics, such as the equation of state, is improved, so too is the propensity of the models to fit the data. This is certainly an indication that in some respects the theory is on the right track. Please note that , strictly speaking, the procedure should not be thought of merely as one of introducing additional parameters that can be adjusted to fit one's preconceptions (or real data) more tightly. The strength of the argument lies in the fact that *first* one decides what is an improvement, and only *afterwards* does one discover that the fit to observations can be more successfully carried out. The fit does not become tighter with every improvement, however. We have just been told that the latest improvement to the Yale model, that of taking gravitational settling of helium into account, has led to a deterioration of the fit, which indicates that the model has not yet been adequately honed. But the trend is clear. Of course, a cynic might retort that the general trend does little more than exhibit the deficiencies of the earlier models, and logically perhaps that is correct, though I regard is as hardly fair. The classical models actually fit much of the seismic data very well indeed. But what does the residual misfit tell us? Is it a product of some minor property of the model that is insignificant at least so far as neutrino physics is concerned? Or is it a symptom of one of the more serious *ad hoc* tenets of the classical theorists that unduly constrains the models not to reproduce conditions in the solar core, yet, at the same time, assures their robustness?

Broadly speaking, solar physicists working on these problems can be divided into two camps. In one, theoretical models are massaged in an attempt to bring their properties closer to observation. In the other, an attempt is made to characterize the class of structures that actually do reproduce the observations within the suspected errors. Evidently, the failure of classical models to reproduce the observations satisfactorily is an indication that the products of the two camps do not yet overlap. In days before the seismic constraints, the objective was merely to find a theoretical model that reproduces the low neutrino flux. But now that neutrino transitions are more widely entertained, that end is less commonly accepted as an objective. It is evident that if the sun is to provide a testbed for neutrino physics, it is necessary for us really to know what constraints the observations set on conditions where the neutrino flux is likely to be influenced. To this end a very careful assessment from the second camp is absolutely essential. That is a task for helioseismology.

What is helioseismology

The basic objective of helioseismology is to determine the structure of the sun by analysing its spectrum of oscillation eigenfrequencies. The procedure is not unlike identifying the instruments in a symphony orchestra. It is from the quality of the tones, which is determined principally by the transient amplitudes and the deviations of the frequencies of the overtones from harmonics of

their fundamentals, that one can recognize the difference between, say, an oboe and a clarinet: the different constructions of their mouthpieces and their reeds and the differences in the shapes of their bores make them quite easy to tell apart. In all cases, identification relies on a combination of recognition and presumption. One recognizes a tone as being similar to one that one has heard and identified before, and presumes this time that that tone is unlikely to have been produced otherwise. Note that the recognition stage is essential: I am quite sure that any physicist who had never seen a 'cello before, should such a person exist, would be unable to draw a picture of the instrument after having experienced no more than the sound of its being played, no matter how good his ear and his artistic talents. However, had he seen and heard a violin before, he might have deduced that the 'cello is simply a larger version, and of course he would be essentially correct. (Indeed, this latter example illustrates one of the scaling techniques of the subject of asteroseismology, which is the seismological investigation of the natures of stars other than the sun.)

It takes a connoisseur to hear the difference between a Stradivarius and a Guarnerius del Gesù. The degree of sophistication required to do so is similar to that at which we are aiming for the sun. The connoisseur's technique must be formalized in mathematical terms, and applied in essentially the same way. As such it is called an 'inverse calculation'. When so translated explicitly into our language, one then realizes that it is simply based on what one might call the Principle of Scientific Arrogance, namely, that our prejudice cannot be very far from the truth. That prejudice is typically based on what was the subject of the first two lectures: a classical solar model or some slight deviant from it. One therefore looks in the vicinity of such a model for a slightly adjusted representation whose eigenfrequencies correspond, within observational error, to those of the sun. (We have not yet been able to use oscillation amplitudes as a diagnostic because the processes that determine them are not understood well enough.)

An important point to realize now is that there are infinitely many places in even a finite domain of function space in which to search for an acceptable representation of the sun. Therefore, even disregarding the leeway provided by uncertainty in the observations, with only a finite amount of data one can never find the unique spot. The problem in hand is said to be ill-posed. Headway is often made by one, or both, of two approaches. One of them is to constrain the task yet further, by strengthening the guiding principle: the Strong Principle of Scientific Arrogance is that one's prejudice is as near to the truth as it possibly could be. Thus, one seeks that representation of the structure of the sun that is as close as possible to a reference (possibly classical) model and whose eigenfrequencies are acceptably close to the observations. This does not yet specify the task uniquely, for it still remains for the investigator to define a measure on the space in order to quantify proximity. This is usually represented as some combination of a direct measure in the space of models and a measure in the space of eigenfrequencies. Thus the helioseismologist is still afforded a substantial degree of flexibility in the manner by which he can practise his art. The other procedure is to specify the region in function space by some set of functionals on it, such as localized averages of the structure of the sun. I shall show some later. These provide a smoothed representation of the structure of the sun which is relatively easy to comprehend. (The raw frequencies themselves can be represented as averages of the structure, but the weight functions are not simple enough for those averages to be readily interpreted.) In practice, the latter procedure provides averages of the differences between the structure of the sun and that of a reference theoretical model, obtained by linearization (with possible iteration to take nonlinearities into account), so it too relies to some degree on the Principle of Arrogance, but only in the weak form.

I should point out that there is another method based on an asymptotic representation of the

modes which does not rely explicitly on a particular reference model. However, like the other methods I have mentioned, it does restrict the function space within which a representation of the solar structure is sought, by assuming the sun to satisfy the conditions under which the asymptotic approximation is valid. Broadly speaking, asymptotic methods, when carried out carefully enough, yield results that are consistent with the linearized perturbation methods.

Helioseismic inversion

Oscillations of the sun fall into basically two classes, called p modes and g modes. Modes of the first kind are standing sound waves: they reside mainly where acoustic propagation is slow, which is in the cooler outer regions of the sun. Thus, they are relatively easy to observe. The g modes are internal gravity waves, the restoring force for which is buoyancy. (They are not to be confused with gravitational waves, which propagate the curvature of space-time.) The g modes of greatest interest are almost completely confined deep in the solar interior, and have relatively small amplitudes at the surface. Thus they are very difficult to observe. Indeed, I think it is true to say that no g mode has yet been convincingly observed and identified in the sun.

The frequencies of the modes depend mainly on the structure of the sun in the regions where their amplitudes are greatest. Some weighting functions are illustrated in Figure 1. Our interest here is principally in the energy-generating core. Therefore, in principle, we would be most interested in g modes. However, since we have no reliable data, we must work with acoustic modes alone. My analogy with musical instruments is therefore perhaps closer than one might have expected.

The amplitudes of all the observed solar oscillations are small. Therefore linear theory provides an excellent description of the motion. Because at least the large-scale structure of the sun does not change substantially over the period of an oscillation, the oscillation frequency ω is a well defined quantity, and the equation governing the spatial structure of the mode can be written

$$\mathcal{L}_\omega \Psi = 0, \tag{1}$$

where $\mathcal{L}_\omega$ is a linear elliptic differential operator in spatial coordinates $\mathbf{x}$, depending on both ω and the structure $\mathbf{S}(\mathbf{x})$ of the sun, and $\Psi(\mathbf{x})$ is the amplitude of some wave-disturbance quantity. Equation (1) must be solved subject to boundary conditions $\mathcal{B}$ on the surface $\mathcal{S}$ of the sun, which determines possible values of ω as eigenvalues. Of course, equation (1) is only an idealization of reality. It is based on certain physical approximations, such as assuming adiabatic motion and neglecting Reynolds stresses in the turbulent convection zone that occupies the outer 30 per cent, by radius, of the sun. However, we believe that for most purposes it is adequate, or that if it is not it provides a firm basis from which one can subsequently estimate small corrections.

With a knowledge of $\mathbf{S}$ one can solve equation (1) for eigenfunctions Ψ_i satisfying $\mathcal{B}$ to yield the eigenfrequencies ω_i. The eigenfrequencies satisfy a variational principle of the form

$$\omega^2 = \frac{\mathcal{K}(\mathbf{S}, \Psi)}{\mathcal{I}(\mathbf{S}, \Psi)}: \tag{2}$$

at fixed $\mathbf{S}$, amongst all functions Ψ satisfying $\mathcal{B}$, the stationary values of $\mathcal{K}/\mathcal{I}$ occur at the eigenfunctions Ψ_i, at which also ω^2 takes the values ω_i^2. This is the so-called forward problem.

The inverse problem is to find $\mathbf{S}$ given ω_i. At this point the Principle of Arrogance is invoked to justify a perturbation approach. The variational principle (2) provides a basis for linearization

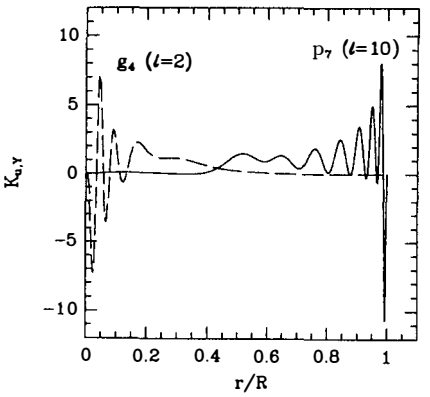


Figure 1. Components K_{ui} of kernels $\mathbf{K}_i$ for structure perturbations $\delta \ln u$ where $u = p/\rho$, for a g mode of order n = 4 and degree l = 2, (dashed line) and a p mode of order 7 and degree 10 (continuous line).

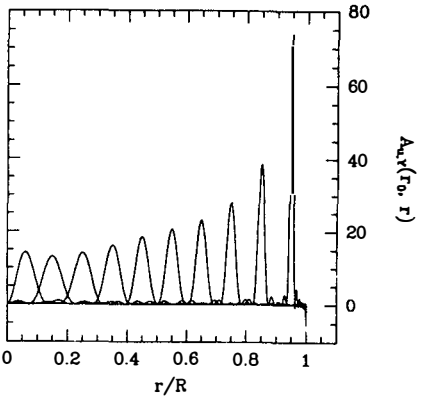


Figure 2. A selection of components A_u of averaging kernels $\mathbf{A}$ defined by equation (6), constructed to be localized about various radii in such a way that the other component A_γ is everywhere small. Only kernels $\mathbf{K}_i$ for p modes that have been observed in the sun were used.

about a reference model $\mathbf{S}_0$ with eigenfunctions Ψ_{0i} and eigenfrequencies ω_{0i} for the presumed small differences $\delta\mathbf{S} := \mathbf{S} - \mathbf{S}_0$, $\delta\omega_i^2 := \omega_i^2 - \omega_{0i}^2$, where $\mathbf{S}$ and ω_i are the actual structure and oscillation frequencies of the sun. Thus

$$\delta\omega_i^2 = (\delta\mathcal{K}_i - \omega_{0i}^2 \delta\mathcal{I}_i)/\mathcal{I}_{0i}, \tag{3}$$

where $\delta\mathcal{K}_i$ and $\delta\mathcal{I}_i$ are linearized variations of $\mathcal{K}$ and $\mathcal{I}$ for the mode (Ψ_{0i}, ω_{0i}) resulting from variations $\delta\mathbf{S}$, and $\mathcal{I}_{0i} := \mathcal{I}(\mathbf{S}_0, \Psi_{0i})$. In view of the stationary property of the ratio $\mathcal{K}/\mathcal{I}$, variations in Ψ_i do not contribute linearly to $\delta\omega_i^2$, and therefore the expression on the right-hand side of equation (3) does not depend explicitly on the actual eigenfunctions of the sun; only the eigenfunctions of the reference need be evaluated. Equation (3) may be usefully rewritten in the form

$$\frac{\delta\omega_i^2}{\omega_{0i}^2} = \int_{\mathcal{V}} \mathbf{K}_i . \delta\mathbf{S} dV, \tag{4}$$

the integral being over the volume $\mathcal{V}$ of the sun. The kernel $\mathbf{K}_i$ depends on $\mathbf{S}_0$ and on the eigenfunction Ψ_{0i}.

The objective now is to estimate, or to provide easily comprehensible constraints on the possible structures, $\mathbf{S}(\mathbf{x})$. Whatever the techniques employed, the details of which I shall not discuss here, the outcome can be expressed as a set of averages

$$\overline{\delta\mathbf{S}}(\mathbf{x}_0) := \int_{\mathcal{V}} \mathbf{A}(\mathbf{x}; \mathbf{x}_0) . \delta\mathbf{S} dV = \sum_i \alpha_i(\mathbf{x}_0) \frac{\delta\omega_i^2}{\omega_{0i}^2}, \tag{5}$$

where it is hoped that the averaging kernel

$$\mathbf{A}(\mathbf{x}; \mathbf{x}_0) = \sum_i \alpha_i(\mathbf{x}_0) \mathbf{K}_i(\mathbf{x}; \mathbf{S}_0, \Psi_{0i}) \tag{6}$$

is highly localized about a chosen point $\mathbf{x}_0$. In some cases, in addition, an estimate of an example of a function $\delta\mathbf{S}$ whose averages are $\overline{\delta\mathbf{S}}$ is also provided. In Figure 2 I illustrate some averaging

kernels $\mathbf{A}$. That demonstrates how well the data can resolve structure. In this example, only modes that have already been observed in the sun have been employed. Moreover, the present uncertainty in the measured eigenfrequencies has also been assumed. Therefore, the resolution is not as high as might be possible in the future. The kernels are for spherical averages of the sun, so the independent variable is simply the normalized radius $x := r/R$, where R is the radius of $\mathcal{S}$. Magnetic field, rotation and other kinds of essentially steady internal motion are believed to be small perturbations and can be treated separately. Here, the structure is represented simply by the function $u = p/\rho$, p being pressure and ρ density, which is proportional to the square of the sound speed. This is but one component of a two-component vector $\mathbf{S}$ necessary to specify the hydrostatic stratification of the spherical averages. The averages have been constructed such that the other component A_γ of $\mathbf{A} = (A_u, A_\gamma)$ is everywhere small.

Before presenting results, I should make a further, very important point. The adiabatic eigenvalue equation (1) depends on $\mathbf{S}$ only via $p(\mathbf{x}), \rho(\mathbf{x})$ and the adiabatic compressibility factor $\gamma(\mathbf{x}) = (\partial \ln p/\partial \ln \rho)_s$, the partial derivative being taken at constant specific entropy s and constant chemical composition. The dynamics of the oscillations are described basically in terms of material with inertia density ρ being accelerated by pressure gradients, the perturbation p' to p being determined in terms of the perturbation ρ' to ρ in terms of γ. Therefore, so far as the hydrostatic stratification is concerned, by seismological reasoning alone one can hope to do no more than determine $p(\mathbf{x})$, $\rho(\mathbf{x})$ and $\gamma(\mathbf{x})$, and any other thermodynamic quantity that depends on those three quantities only.

A second point worthy of mentioning is that in reducing $\mathbf{S}$ from a three-dimensional vector function such as (p, ρ, γ) to a vector in two dimensions, I have related p to ρ via the equation of hydrostatic support. Since this equation is used also in the derivation of the linear eigenvalue equation (1), this reduction introduces no new approximation. It does, however, emphasize that hydrostatic balance has been explicitly used.

The purely seismic description of the solar interior is inadequate for studying nuclear reactions. The hydrostatic stratification must be augmented by a knowledge of the temperature $T(\mathbf{x})$ and the composition $\mathbf{X}(\mathbf{x})$, where X_i is the abundance $[i]$ by mass of a chemical element identified by a label i. If $\mathbf{X}$ were known, then T can be determined from p and ρ from the equation of state. One might estimate $\mathbf{X}$ with the minimal addition of new assumptions by taking from the classical model just the assumptions: (i) that the values of the abundances of the heavy elements (i.e. elements heavier than helium) take their usually adopted values, except that the abundances of C, N and O have been redistributed amongst each other, at constant [C]+[N]+[O], by the CNO cycle, and (ii) that the sun is in radiative energy balance beneath the convection zone, and satisfies the usual energy-transport relations in the convection zone. Together with the seismic determination of p, ρ and γ, this completely determines the structure. However, it is evident that the additional nonseismic variables are less reliably estimated, though, in principle, the model is a more reliable representation of the sun than the classical model, because it depends on fewer untested assumptions.

Seismic determination of conditions in the solar interior

Because classical models appear to provide a good first-approximation to the structure of the sun, I present the seismic inferences as relative deviations from a classical model. The reference model I adopt is the central model of Gough and Novotny [3], chosen partly because its structure has been computed accurately enough for it to be possible to obtain precise eigenfrequencies. Its total heavy-element abundance Z is 0.02, and it was evolved to an age of $t_\odot = 4.60 \times 10^9$y.

The initial helium abundance is $Y_0 = 0.2762$; the present central values of u and ρ are 0.1536 $Mm^2 s^{-2}$ and 148.2 g cm^{-3} respectively.

The eigenfrequencies used for the inversion are a combination of modes observed in 1986 by Libbrecht, Woodard and Kaufman[4], which are of degree $l \geq 5$. Basically they were obtained from spatially resolved Doppler measurements, projected onto spherical harmonics of degree l and order m, and then Fourier analysed in time. The m dependence of the frequency represents degeneracy splitting by rotation and asphericity in the solar structure, which I shall not discuss here. These data were supplemented by spatially unresolved measurements of modes with $l \leq 3$ by Elsworth *et al.*[5]; they are time-averaged frequencies that have been scaled to 1986 for observed solar-cycle variation[6], assuming that all the variation in the solar structure arises in the photospheric layers, as indicated by Libbrecht and Woodard[7].

The results of the inversion are illustrated in Figure 3. They are averages of $\delta u/u$ and $\delta\rho/\rho$ weighted by kernels A such as those illustrated in Figure 2, the characteristic widths of which are represented by the horizontal bars. The vertical bars represent $\pm$ 1 standard error based on the published estimates of the measurement error.

Discussion of the inversion

The most prominent feature of $\delta u/u$ illustrated in Figure 3 is the positive hump between $r \simeq 0.25R$ and $r \simeq 0.7R$, indicating that the sound speed in the sun is greater than that

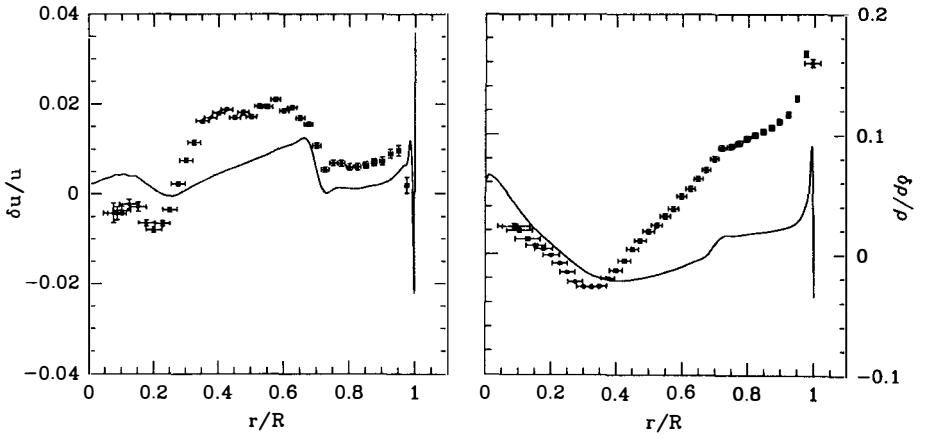


Figure 3. Averages $\overline{\delta \ln u}$ and $\overline{\delta \ln \rho}$ computed by D.O. Gough and A.G. Kosovichev representing relative differences between the structure of the sun and that of the central model of Gough and Novotny[3]. The continuous lines represent the corresponding differences between the standard solar model and a model computed with the Livermore[9] opacity corrections.

of the model, by nearly 1 per cent in places. (Recall that $c^2 = \gamma u$; matter is essentially fully ionized for $r \lesssim 0.97R$, and therefore $\gamma \simeq 5/3$ and $\delta c/c \simeq 0.5\delta u/u$ in the region under discussion.) This has been known since the first sound-speed inversion that extended beneath the convection zone[8]. Its cause was attributed to a 20 per cent error in the opacity in the radiative interior at temperatures less than about 4×10^6K, which, broadly speaking, subsequent opacity computations at Livermore[9] have confirmed. Thus we see an example of a helioseismic measurement being used to make a microphysical inference. It is interesting to notice also in Figure 3 that readjusting the opacity produces a superficially similar correction to c. Moreover, it produces a density correction very roughly of the right shape. However, the computed opacity corrections do not explain the entire discrepancy between the sun and the solar model.

What is of greater interest to this workshop is any deduction that can be made about conditions in the reacting core. Recall that, according to the perfect gas law, $u \propto T/\mu$, where μ is the mean molecular mass of the material. The perfect gas law is not satisfied exactly – electrons are partially degenerate – but it does serve as an adequate guide to our thinking. Thus, the slight reduction in c in the core and the augmentation of ρ throughout much of the core, coupled with the requirement that the density-weighted integral of the total thermal energy generation rate $\epsilon \propto X^2\rho T^4$ is unaltered, suggests that the minimal adjustment to X is quite small and that therefore the ^{8}B neutrino flux, in particular, which arises predominantly from the inner 5 per cent or so by radius, is perhaps just slightly overestimated by the theoretical model. However, I must emphasize that this conclusion is based on the assumption that the inner regions of the sun are and always have been (since significant nuclear processing began) in a state of no motion and in thermal balance.

Concluding remarks

The helioseismic evidence confirms that the distribution of density and sound-speed in the reacting core of the sun is similar to that in classical solar models, though the inversions illustrated in Figure 3 do indicate that a discrepancy remains even after the new opacity calculations. Thus, it is not unlikely that the story told by Dr Turck-Chièze is basically correct. Classical predictions of the solar neutrino luminosity L_ν are somewhat greater than observation, though the significance of the discrepancy I leave for you to judge. However, we do not actually have measurements of temperature T and hydrogen abundance X, so any inference about nuclear reaction rates is necessarily model dependent. Let me conclude briefly by questioning the models in two different ways.

The first concerns $X(r)$. I mentioned a minimal adjustment to reconcile the inversions, by which I meant modifying the initial uniform abundance and assuming the profile to have evolved in the classical way. But what if the initial abundance were not uniform? It has been suggested, for example, that the very centre of the sun is devoid of hydrogen, being composed either of heavy elements that formed an initially solid seed to attract gravitationally the rest of the solar material at the time of the formation of the sun, or that at early times a phase change caused helium to rain into the central core, displacing the hydrogen. Either hypothesis might lead to a severe reduction in the ^{8}B neutrino production rate, and a lesser reduction in the production of ^{7}Be neutrinos, bringing the total luminosity L_ν more-or-less into agreement with the Kamiokande-II and Homestake values.

The second concerns the nonlinear development of the instability I mentioned earlier in this talk. That negates the classical assumption of no motion, and leads to spatial and temporal perturbations which, as Roxburgh[10] and Ghosal and Spiegel[11] have pointed out, might also

reduce L_ν to within the observed limits; a simple pilot calculation has confirmed that view[12].

Finally, let me reiterate that macroscopic phenomena of this kind have no influence on pp neutrino production, provided the current thermal energy production rate is equal to the observed photon luminosity. That equality too has been questioned[13], in a proposal that the core is currently, and atypically, out of thermal balance. Nevertheless, the balance is upset by only a few per cent. Therefore, if it were to be confirmed that the low-energy pp neutrino flux is substantially lower than classical theoretical predictions, as the preliminary SAGE results suggest, then one would be forced to conclude that classical nuclear or particle physics with massless neutrinos must be revised. If, on the other hand, it is eventually found that the ^{71}Ga capture rate is about 110 snu, which is typical of models with motion in the core that reproduce the neutrino flux measured by Kamiokande-II, then no serious revision will be demanded.

References

1. Bahcall, J.N. and Ulrich, R.K., 1988, *Rev. Mod. Phys.*, **60**, 297-372
2. Gough, D.O., 1990, *Kong. Danske Vid. Selsk., Mat-fys. Medd.*, **42**:4, 13-50
3. Gough, D.O. and Novotny, E., 1990, *Solar Phys.*, **128**, 143-160
4. Libbrecht, K.G., Woodard, M.F. and Kaufman, J.M., 1990, *Astrophys. J. Suppl.*, **74**, 1129-1149
5. Elsworth, Y., Howe, R., Isaak, G.R., McLeod, C.P. and New, R., 1991, *Mon. Not. R. astr. Soc.*, **251**, 7P-9P
6. Elsworth, Y., Howe, R., Isaak, G.R., McLeod, C.P. and New, R., 1990, *Nature*, **345**, 322-324
7. Libbrecht, K.G. and Woodard, M.F., 1990, *Nature*, **345**, 779-782
8. Christensen-Dalsgaard, J., Duvall Jr, T.L., Gough, D.O., Harvey, J.W. and Rhodes Jr, E.J., 1985, *Nature*, **315**, 378-382
9. Iglesias, C.A. and Rogers, F.J., 1990, *Inside the sun* (ed. G. Berthomieu and M. Cribier, Kluwer, Dordrecht) 81-90
10. Roxburgh, I.W., 1985, *Solar Phys.*, **100**, 21-51
11. Ghosal, S. and Spiegel, E.A., 1992, preprint
12. Gough, D.O., 1992, *Ann. N.Y. Acad. Sci.*, in press
13. Dilke, F.W.W. and Gough, D.O., 1972, *Nature*, **240**, 262-264, 293-294

HELIOSEISMOLOGY: Solar p-modes from Standard Model Probing the Solar interior

*I.Lopes**
DAPNIA, Service d'Astrophysique
Centre d'Etudes de Saclay
91191 Gif-sur-Yvette Cedex 01, FRANCE

Abstract

This is an introduction to the helioseismological investigation of the solar interior. The equations and the methods are briefly introduced. Acoustic $p-$ *mode* frequencies obtained from the Saclay Standard Model will be presented and compared with observation data.

* *Also Centro de Astrofísica, Universidade do Porto*
Rua do Campo Alegre, 823, 4100 Porto - PORTUGAL

Introduction

The Observation of the five-minute oscillations in 1960 by Leighton [1], predicted by Ledoux and Walraven in 1958 [2], has opened a new field of research in the study of the Solar Interior. The object of Helioseismology is to make some deductions about the chemical composition and physical mechanism acting in the interior of the Sun.

This new discipline enables us to study with more details one of the basic difficulties of the standard model, the microphysics: the equation of state, the opacity and the nuclear energy generation rate. This motivates also the study of non standard mechanisms in the Solar model. Among those we have some macrophysical processes like energy tranport, dynamics of convection, convective overshoot, molecular diffusion, core mixing and magnetic fields.

In the first section we present the physical mechanisms of solar oscillations and the different methods used in Helioseismology to study the structure of the Sun. In the last section, we will discuss some results of the Standard Solar Model from the point of view of Helioseismology.

1 Adiabatic equations of solar Oscillations

The evolution of the Sun in the Main Sequence is determined by the nuclear network ($p-p$ chain, CNO cycle) with a typical time scale of about 10^{10} years, which is larger than the dynamical time scale ($0.5h$) associated to the non-radial oscillations observed in the Sun. This condition allows to the linearisation of the hydrodynamic equations around an equilibrium state (the Sun is spherical), where all the equilibrium quantities are functions of the radius only.

Oscillations of the Sun were interpreted as a superposition of eigenmodes. For a star without magnetic field and without rotation a eigenmode of oscillation is described in terms of spherical harmonics of degree l. For each l the modes are distinguished by the radial eigenfunction, which can be identified by the number of nodes in the radial direction defined as the radial order n. The local vibrational property of the structure is characterized by the characteristic frequencies of the equilibrium model, the Lamb frequency F_l, the *Brünt – Väisälä* frequency N and the cut-off frequency ω_C (*Figure* 1). For high-frequency oscillations ($\omega > F_l, \omega_c, N$), the restoring force is due mainly to the pressure and locally the oscillations present characteristics of acoustic waves($P-modes$). For low-frequency oscillations ($\omega < N, \omega_c, F_l$), the restoring force is due mainly to buoyancy, and the oscillations show characteristics of a gravity wave ($G-modes$).

A great adiabaticity is achieved locally to a high degree of accuracy, from the center to the basis of the convective zone. For the more superficial layers and in the atmosphere as the thermal time scale is of the order of the dynamical time scale, the adiabatic approximation is not entirely correct to calculate the theoretical frequencies. The correction induced by the nonadiabaticity in the adiabatic eigenfrequencies was estimated by Cox [6], who found that this effect reduced the difference between observed frequencies and eigenfrequencies at the order of $3\mu Hz$ for the modes with $3 < l < 10$ and $3mHz < \nu < 4mHz$, and is negligeable for the modes with $\nu < 3mHz$.

In the calculation of solar oscillation frequencies we also ignored magnetic field, detailed atmospheric behaviours, dynamical effects of convection and excitation mechanisms which may change the adiabatic frequencies of several μHz.

1.1 Asymptotic methods

Different asymptotic theories have been used to study the solution of the linear adiabatic wave equations, and to understand the main properties of the waves and their physical mechanisms.

To put in evidence the physical properties of the p-mode oscillations we have reduced the wave-equation of linear adiabatic oscillations. If we consider the Cowling approximation [5], which is valid for the high degree or high radial order modes, the system of linear adiabatic oscillations is reduced to a second-order differential equation.

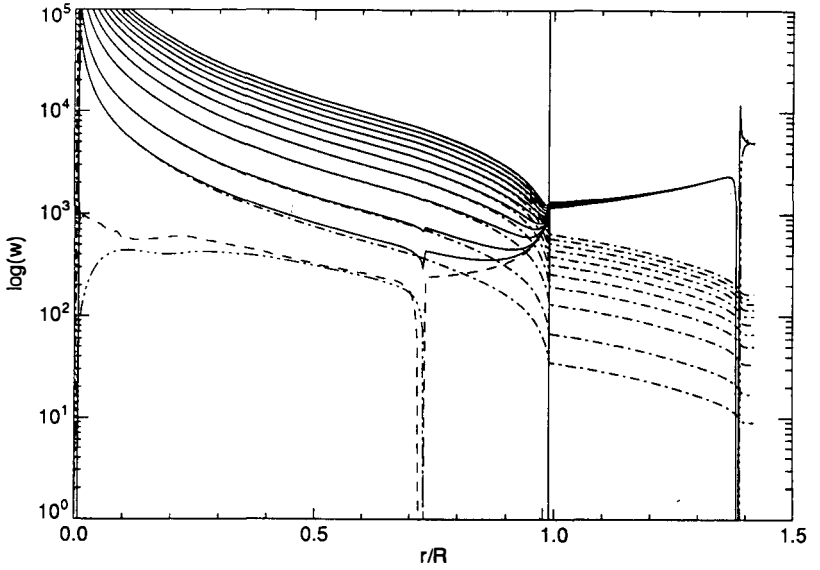


Figure 1 : Propagation diagram for the Present Sun ($l = 5$ *and* $l = 10,(10),100$): the logarithms of the characteristic frequencies and Acoustic Potential are plotted as a function of the radius.The *Brünt – Väissälä* frequency $\frac{N}{2\pi}$ by a *3 doted line*, the Lamb frequency $\frac{F_l}{2\pi}$ by a *1 doted line* , the Cut-off frequency $\frac{w_c}{2\pi}$ by a *dashed line* and Acoustic Potential by a *continous line.* To put in evidence the reflection zone the scale between 0.9 and 1.0 has been multiplied by 40.

If we are only interested in the high order $p-modes$ we can reduce the system to a *Schrödinger* equation in quantum mechanics. If we use the approximation of Deubner and Gough [3] a mode is defined as :

$$\rho^{\frac{1}{2}} c^2 \nabla \vec{\xi} \equiv \xi(r,\theta,\phi) = \xi_{nl}(r) Y_l^m(\theta,\phi) \exp(i w_{nl} t)$$

A mode of degree l and radial order n, has a radial component $\xi_{nl}(r)$ which satisfies the equation:

$$\frac{d^2\xi_{nl}(r)}{dr^2} + \frac{1}{c^2(r)}[\omega_{nl}^2 - \Phi_{nl}(r)]\xi_{nl}(r) = 0$$

where ω_{nl} is the eigenfrequency of the model and the $\Phi_{nl}(r)$ is the Acoustic Potential cavity in which the mode oscillates :

$$\Phi_{nl}(r) = \omega_c^2 + F_l^2(1 - \frac{N^2}{\omega_{nl}^2}) = c^2(r)\left[\frac{d}{dr}(\frac{1}{2H_\rho}) - (\frac{1}{2H_\rho})^2\right] + F_l^2\left(1 - \frac{g}{\omega_{nl}^2}\left[\frac{1}{\Gamma_1}\frac{dlnP}{dr} - \frac{dln\rho}{dr}\right]\right)$$

The observed $five-minute$ oscillations correspond to standing acoustic waves ($p-modes$). For any particular wave, of degree l and frequency ω_{nl}, the guide is a spherical shell(*Acoustic cavity*) whose boundaries depend on the Lamb frequency $F_l = c(r)\sqrt{l(l+1)}/r$ and the cut-off frequency ω_c. At the surface, waves are reflected beneath the photosphere, where the density

scale height $H_\rho = -(dln\rho/dr)^{-1}$ is comparable with the wavelength. In the outer layers the wave reflection is pratically independent of the degree and depends only on the frequency (*Figure 1*). The refraction of the wave initially propagating downward deflects it back towards the surface, at the point where the local horizontal component of the speed of the wave is equal to the sound speed ($c(r_t) = r_t\omega_{nl}/\sqrt{l(l+1)}$).

The modes of low degree ($l \leq 4$) are differently treated, and the last approach developed is not valid in this case. For these modes, a second-order asymptotic analysis has been made in the Cowling approximation by Tassoul [8] for modes with a sufficiently high number of radial order. This method overcomes the difficulties related to the plane wave approach by carefully treating the center as a singular point. It leads to two expressions for the $p-mode$ frequencies, the first-order and second-order separations. Second-order separation is approximately equal to:

$$\Delta\nu_{l,l+2} \equiv \nu_{n,l} - \nu_{n-1,l+2} \approx -\frac{(4l+6)}{4\pi^2\nu_{nl}}\nu_0 \int_0^{R_\odot} \frac{dc}{dr}\frac{dr}{r} \qquad \nu_0 = \left[2\int_0^{R_\odot}\frac{dr}{c}\right]^{-1}$$

The first-order separation $\delta\nu_l \equiv \nu_{n,l} - \nu_{n-1,l} \approx \nu_0$ is expected to be approximately constant (*high radial order*), and depends on the sound speed near the surface where the integrand provides its large contribution. This frequency spacing has been used to identify the different modes of oscillation in the Sun and Solar type stars, and its value is $135.1 \pm 0.1\mu Hz$. The second-order spacing is most sensitive to the interior because its integrand $r^{-1}(dc/dr)$ increases with depth (*see Figure 3*) [11].

1.2 Fonctional Analyses

Principles of Helioseismology can be obtained by regarding the linear adiabatic equations of non-radial oscillations as an eigenvalue problem in a Hilbert Space [2]. The perturbated variables, p', ρ', and ϕ' can be obtained by a linear operation of ξ_r and ξ_h [4] [5] which are the solutions of the adiabatic equations with the boundary conditions, at the center and surface. The eigenfrequencies can be seen as the eigenvalues of a Linear Operator in a Hilbert Space defined with the scalar product:

$$\langle\xi,\eta\rangle = \int_0^{R_\odot} (\xi_r^*(r)\eta_r(r) + l(l+1)\xi_h(r)\eta_h^*(r))dr \qquad \langle\xi,\xi\rangle = \| \xi \|^{\frac{1}{2}} \qquad \forall\xi,\eta \in H$$

where $\xi = (\xi_r, \xi_h)$ is the solution of the linear adiabatic wave equation. The eigenvalue problem can formally be written as $L_{ad}\xi = \omega^2\xi$, where L_{ad} is a *Linear* Operator (*function of the equilibrium variables*) in H with the domain $D(L_{ad}) = \{\xi = (\xi_r,\xi_h) : \xi_r - l\xi_h \to 0 \quad for \quad r \to 0 \wedge \delta p(R_\odot) = 0\}$. We can prove that the Operator is Hermitian [5] in $D(L_{ad})$. Then follow immediately the two useful properties of L_{ad}:

$$\omega^2 = \frac{\langle\xi, L_{ad}\xi\rangle}{\langle\xi,\xi\rangle} \qquad \delta\omega^2 = \frac{\langle\xi, \delta L_{ad}\xi\rangle}{\langle\xi,\xi\rangle}$$

These two results establish the three complementary approaches for the study of the internal structure of the Sun:

-The inverse problem in the asymptotic method, which tries to approach the thermodynamical quantities from the data (*asymptotic inversion of the sound speed* [5]).

-The inverse problem in the Geophysical method, which tries to approach the thermodynamical quantities from the data using a reference model (*determination of the density* [5]).

- The forward problem, which directly compares the measured and calculated quantities and consequently examines the effect of a physical modification, even if it does not guarantee the uniqueness of the proposed solution. We will illustrate this approach in the next section.

2 Discussion of Results

In view of understanding the physics of the Solar interior, it is obvious the interest to compare observed frequencies of the five-minute oscillations with theoretical frequencies of representative models of the present Sun. We have calculated adiabatic $p-$ *mode* frequencies from our reference solar model. The observations were reported by Libbrecht, Woodword and Kaufman (1990) [14].

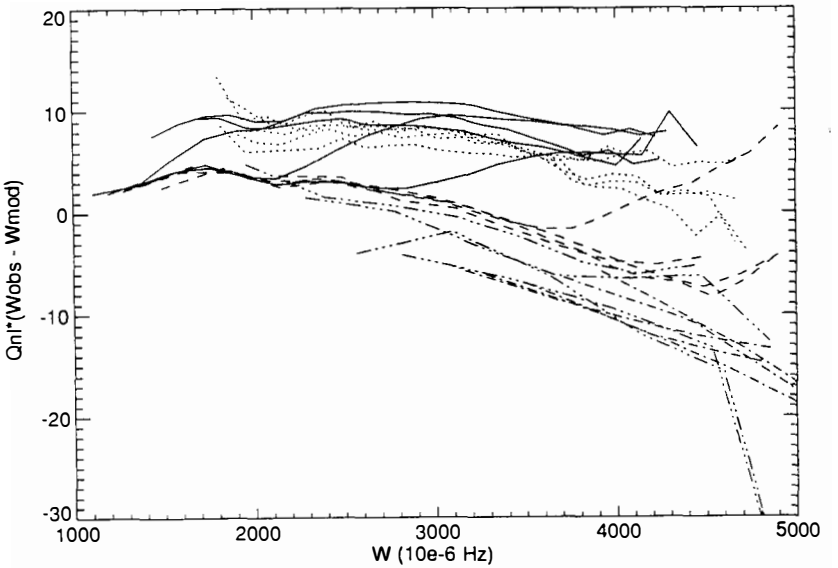


Figure 2 : Scaled Frequency differences between observed frequencies of solar oscillation from the compilation by Libbrecht, Woodard and Kaufman(1990) and theoretical frequencies of Solar Standard Model for selected values of l. Points correponding to a given value of l have been connected, according to the following convention: $l = 0-3$(*doted line*),$l = 4, 5, 10, 20, 30$(*continous line*), $l = 40, 50, 70, 100$(*dashed line*), $l = 150, 200, 300, 400$(*1 doted line*) and $l = 500, 600, 700, 800, 900, 1000$ (*3 doted line*).

Our standard model [13] was computed with approximately 300 spatial meshpoints, 26 or 27 time steps between the Zero Age Main Sequency and the present Sun. Updated model presented here used the Los Alamos Opacity tables, the Vardya Equation of State, the low recent measurement of iron and the intermediate screening for the reaction rates (*see [11] [12] for discussion of the influence of the microphysics*). The model, is calibrated to have at the Solar age ($4.54\ 10^9 years$) the Solar radius (assumed $6.9599\ 10^{10} cm$) and the Solar luminosity ($3.846\ 10^{33} ergs\ sec^{-1}$) by adjusting the mixing length parameter $\alpha(= 2.1456)$ and the initial helium abundance $Y_0(= 0.2664)$. In the calibration a relative error in surface radius and luminosity of less than 10^{-6} was required.

We calculate the theoretical frequencies (*variational method*) [7] in the adiabatic aproximation using a relaxation method to solve the system of adiabatic non-radial oscillations with the condition of regularity at the center and an isothermal atmosphere at the surface [9] [5]. The

errors in the computed frequencies arise partly from errors in the equilibrium model, partly from the errors in the calculation of the frequencies for a given model. What is important in an helioseismic purpose is to avoid any discontinuity which may produce an enhancement in the derivatives used. This could be the case at the bottom of the convective zone, in the treatment of the equation of state and in the treatment of the opacity coefficients. Peculiar care and small modifications have been introduced to avoid such effects.

Figure 2 shows the comparison of the theoretical predictions of acoustic modes with the compilation of Libbrecht et al (1990)[14] for l = 0 to 1000, scaled by Q_{nl}. The modes of high degree penetrate less deeply than modes of low degree so they present a smaller inertia, and hence their frequencies are more easier perturbated. To compensate this effect, those are scaled by the ratio $Q_{nl} = E_{nl}/E_0(\omega_{nl})$, where E_{nl} is a dimensionless measure of kinetic energy of the pulsation (*dimensionless inertia*) and $E_0(\omega_{nl})$ is the extrapolation of dimensionless kinetic energy for radial modes with the same frequency [7]. Then the scaled frequency $Q_{nl}\delta\omega$, measures the effect of the perturbation δL_{ad} (*real Sun – Solar model*) which is localized in the region where the mode is trapped.

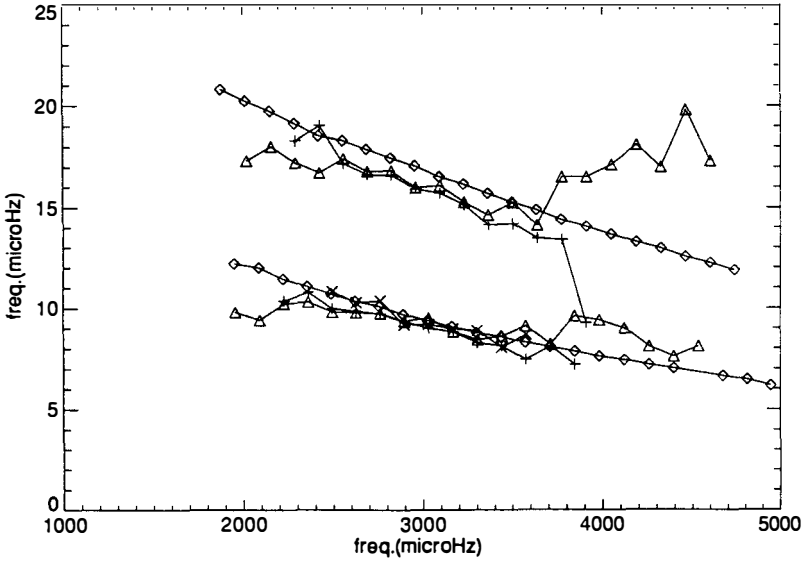


Figure 3 :Variation of the $p-modes$ frequency splitting for consecutive radial order and pair of degrees $\Delta\nu_{2,0}$ and $\Delta\nu_{3,1}$ as a function of the frequency. The observations from the Libbrecht et al. compilation[14] (◁),the observations of Toutain et al.[15](×), the observations of Birmingham group[16](+) and the theoretical calculations of the Standard Model(◇) are shown.

We notice a difference between theoretical frequencies and observations smaller than ± 10 μHz which is largely greater than the experimental error (0.1 to 0.5 μ Hz). This difference translates part of our knowledge of the external layers of the Sun, part of what is not yet understood in the Solar interior. The general trend is very similar to other results [10], the small slope, independent of the degree of the modes, has been remarquely reduced by the treatment of opacities from Los Alamos at the surface. One observes two groups of modes in the scaled frequency differences (*see [12]*) which translate the imperfect behaviour of the sound

speed at the basis of the convective zone. One group of modes do not penetrate the critical region ($r \geq 0.7 R_{\odot}, l \geq 40$), the other one is affected by this region ($0 \leq l \leq 40$). This difference could be due to the treatment of opacities in the solar model. Nevertheless, a degradation of this slope with the degree l shows that the modelling of the atmosphere is not sufficiently well treated. We consider that the uncertainty of our complete calculation is still of the order of $1 - 1.5\ \mu Hz$, and mainly due to the accuracy of our solar modelling.

To better investigate the internal part, we consider the last expression of the section 1.1, $\Delta\nu_{0,2}$ and $\Delta\nu_{1,3}$ [8], which is the best representation to express the sensitivity to the very center. In this case, the influence of the external layers is reduced to about 30 % of the integral and consequently the influence of the most internal part is enhanced. *Figure* 3 shows the theoretical predictions together with all the experimental results already obtained. Differences are of the order of 1- 2 μHz. The rather good agreement between predictions and observations seems to show that, even if the classical frame work of stellar evolution is certainly not the finality of the calculation, the impact of extra phenomena must be relatively small [15].

Acknowledgments

I am grateful to Jørgen Christensen-Dalsgaard who supplied us the pulsation code, which was a fundamental tool for this work, and Werner Däppen who contributed to this work. Also, I would like to thank Sylvaine Turck-Chièze with whom this work has been done, for the useful discussions and helpful comments on this manuscript. This research was partly financed by grant BD/749/90-RM from JNICT, Portugal.

References

[1] R. B. Leighton, 1960, Proc. IAU Symposium n 12, p. 321.
R. B. Leighton,R.W. Noyes and G. W. Simon, Ap. J. 135 (1962) 474.

[2] Ledoux P., Walraven T.: Handbuch der Physik 51, 353 1958.

[3] Deubner F.L., Gough D.O.: Ann. Rev. Astron. Astrophys. 22, 593-619(1984)

[4] Unno W., Osaki Y., Ando H.,Shibahashi H.: *Non-radial oscillations of stars*, University of Tokyo Press (1979).

[5] Unno W., Osaki Y., Ando H.,Saio H., Shibahashi H.: *Non-radial oscillations of stars*, Second Edition, University of Tokyo Press (1989).

[6] A. N. Cox, J. A. Guzik and R. B. Kidman, Ap. J., (1989), 342, 1187.

[7] Christensen-Dalsgaard J., Berthomieu G.: 1990, to appear in *Solar interior and atmosphere*, Chapter 11, eds. Cox A.N., Livingston W. C. and Matthews M., Space Science series, University of Arizona Press.

[8] Tassoul M.: Astrophys. J. Suppl. 43, 467-490 (1980).

[9] Christensen-Dalsgaard J.: Mon. Not. R. astr. Soc. 199, 735-761 (1982).

[10] J. Christensen-Dalsgaard :1991,in ITP Proccedings: Challenges to theories of the structure of moderate-mass stars (eds D. O. Gough and J. Toomre) 11.

[11] I. Lopes and S. Turck-Chièze, 1992, Astrophys. J., to be published.

[12] S. Turck-Chièze, these proceedings.

[13] S.Turck-Chièze, S. Cahen, M. Cassé and C. Doom, Ap. J. 335 (1988) 415.

[14] K.G. Libbrecht, M. F. Woodard and J. M. Kaufman, 1990, Astrophys. J. Supplement, 74, 1129.

[15] Th. Toutain and C. Fröhlich, 1992 Astron. Astrophys., 257(1992)287.

[16] Y. Elsworth, R. Howe, G. R. Isaak, C. P. McLeod, and R. New, MNRAS 251(1988)6.

NEUTRINO OSCILLATIONS IN VACUUM AND THE SOLAR NEUTRINO PROBLEM

P.I. Krastev [1,3] and S.T. Petcov [2,3] *

[1] Theory Division, CERN, CH-1211 Geneva 23, Switzerland.
[2] Scuola Internazionale Superiore di Studi Avanzati, and Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, I-34014 Trieste, Italy.
[3] Permanent address: Institute of Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria.

Abstract

We study the possibility that vacuum oscillations of the solar neutrinos are at the origin of the solar neutrino problem. The available data from the Homestake and Kamiokande II experiments are analyzed allowing for seasonal variations of the solar neutrino flux due to the oscillations. We find that two-neutrino oscillations of the solar ν_e in vacuum allow to reconcile the results of these experiments for 0.5×10^{-10} eV$^2 \leq \Delta m^2 \leq 1.1 \times 10^{-10}$ eV^2 and $\sin^2 2\theta \geq 0.75$.

* Presented by S.T. Petcov.

If the existence of unconventional neutrino properties and/or interactions is at the origin of the solar neutrino problem [1–5], as the current status of the problem suggests [6], several physically rather different solutions are possible. The solar neutrinos, ν_e, can take part (together with the other flavour neutrinos ν_μ and ν_τ, for instance) in oscillations in vacuum [7, 8], while propagating from the surface of the Sun to the Earth *. They can undergo matter-enhanced transitions (oscillations) into neutrinos of a different type (or types) while propagating from the central region of the Sun, where the solar neutrinos are supposedly produced, to the surface of the Sun [9–12]. Furthermore, spin or spin-flavour precession of the solar neutrinos ν_e can take place in the magnetic field of the Sun [13, 14] provided the electron neutrinos possess a sufficiently large magnetic dipole moment or transition moment of the magnetic dipole type. The spin-flavour precession and the spin precession of the solar ν_e can be resonantly enhanced by the solar matter [15, 16]. Finally, the observed deficit of solar neutrinos can be caused by ν_e instability [17].

The vacuum oscillation and the matter-enhanced transition solutions of the solar neutrino problem have been especially well studied [18–20]. As is well known, neutrino oscillations can take place in vacuum if neutrinos possess non-zero and different masses and non-trivial lepton mixing exists in vacuum [21]. As a rule, the solar neutrino data are analyzed under the simplest assumption that solar neutrinos take part in two-neutrino oscillations, which for concreteness we shall suppose to be of the type $\nu_e \leftrightarrow \nu_\mu$ or $\nu_e \leftrightarrow \nu_s$, ν_s being a sterile neutrino [21]. The two-neutrino vacuum oscillations are characterized by two parameters: $\Delta m^2 = m_2^2 - m_1^2 > 0$, $m_{1,2}$ being the masses of two neutrinos $\nu_{1,2}$ with definite mass in vacuum, and θ, the lepton mixing angle in vacuum **. The oscillations of interest ($\nu_e \leftrightarrow \nu_\mu$ or $\nu_e \leftrightarrow \nu_s$) are possible in vacuum provided $\Delta m^2 \neq 0$ and $\sin^2 2\theta \neq 0$.

A rather thorough study of the possibility to explain the solar-neutrino observations on the basis of the two-neutrino vacuum oscillation hypothesis was performed in ref. [18]. The data of the Cl–Ar experiment of Davis and collaborators [1] and those on the total flux and spectrum of the solar ^{8}B neutrinos obtained by the Kamiokande II collaboration [2] were used in the analysis done in ref. [18]. It was found, in particular, that the two-neutrino oscillations in vacuum of the type $\nu_e \leftrightarrow \nu_\mu$ provide a solution of the solar neutrino problem approximately for

$$0.5 \times 10^{-10}\ \mathrm{eV}^2 \leq \Delta m^2 \leq 2.5 \times 10^{-10}\ \mathrm{eV}^2 \qquad (1)$$

and

$$\sin^2 2\theta \geq 0.7 \qquad (95\%\mathrm{C.L.}). \qquad (2)$$

For the values of Δm^2 and $\sin^2 2\theta$ satisfying (1) and (2), the probability of oscillations of the solar pp, ^{7}Be and lower-energy fraction of the ^{8}B neutrinos exhibit a strong periodic

* The fact that ν_e oscillations can lead to a "deficit" of solar electron neutrinos detected at the Earth's surface was pointed out in ref. [8] before the first data on the solar neutrino flux were obtained, i.e., before the solar neutrino problem existed.

** For further details concerning the theory of neutrino oscillations in vacuum see refs. [21].

dependence on the Sun–Earth distance [21]. As a consequence, seasonal variations of the indicated components of the solar ν_e flux should take place.

In the present note the data of the Cl–Ar and Kamiokande II solar neutrino experiments are analyzed in terms of the two-neutrino vacuum oscillation hypothesis by taking into account the possibility of seasonal variations of the different components of the solar ν_e flux. The recent results of the Kamiokande II collaboration on the seasonal variations of the solar neutrino flux [23] are used in our analysis. The indicated Kamiokande I-I data were not available at the time when the previous analogous studies [18,19] were performed, and the possible effects of seasonal variations were not taken into account in obtaining constraints on the values of the parameters Δm^2 and $\sin^2 2\theta$ in these studies.

In our analysis we have paid special attention to the change of the distance between the Earth and the Sun, which is due to the ellipticity of the Earth's orbit. Besides the change of the overall neutrino flux with the distance R, like R^{-2}, this leads to a change of the solar neutrino survival probability:

$$P(t) = 1 - \sin^2 2\theta \sin^2\Big[\pi\frac{R(t)}{L_v}\Big], \tag{3}$$

Here $L_v = 4\pi E/\ \Delta m^2$ is the oscillation length in vacuum, and $R(t)$ is the Sun–Earth distance,

$$R(t) = R_0\big[1 - \epsilon\cos\big(2\pi\frac{t}{T}\big)\big], \tag{4}$$

where $R_0 = 1.49 \times 10^{13}$ cm is the mean Sun–Earth distance, $\epsilon = 0.0167$ is the ellipticity of the Earth's orbit, t is the corresponding time of the year at which the solar neutrino flux is being measured, and $T = 365$ days. In order to evaluate the predicted signal (i.e., event rate) in a solar neutrino detector for a given time interval $[t_1, t_2]$ one has to compute the following integral:

$$Q_{th} = \int \sigma(E_\nu)F(E_\nu)\bar{P}(E_\nu, \Delta m^2, \sin^2 2\theta)dE_\nu, \tag{5}$$

where $\sigma(E_\nu)$ is the cross-section for the neutrino reaction in the relevant solar neutrino detector, $F(E_\nu)$ is the solar neutrino flux spectrum, and

$$\bar{P}(E_\nu, \Delta m^2, \sin^2 2\theta) = \frac{1}{(t_2 - t_1)}\int_{t_1}^{t_2} P(t)dt. \tag{6}$$

The theoretically predicted signal Q_{th} is normalized to reproduce the event rates derived on the basis of solar model calculations in ref. [4] or ref. [5]. For fixed Δm^2 and

$\sin^2 2\theta$ we evaluate the corresponding value of Q_{th} for each run and then compute the χ^2 function

$$\chi^2 = \sum_{i=1}^{N} \frac{(Q^i_{th} - Q^i_{exp})^2}{\sigma_i^2}, \tag{7}$$

where σ_i is the standard deviation of the experimental result for the i-th run. We then compare the χ^2 value computed in this way with the corresponding percentage point for the χ^2 distribution function for N degrees of freedom and determine whether, at the chosen confidence level, the pair Δm^2–$\sin^2 2\theta$ can be rejected or not.

The results of the Homestake experiment are available for 84 runs (from run No. 18 to run No. 106, except runs No. 23, 25, 26, 34, 93). The mean value of the ^{37}Ar production rate and the beginning and end date of each run determining t_1 and t_2, correspondingly, are used as input in formula (6). The error bars are asymmetrical and do not allow negative values of the ^{37}Ar production rate because of the chosen method for data handling, which has been a matter of extensive discussion in the literature. In our analysis we use symmetrical error bars, taking the larger of the two values for each particular run. The predictions for the fluxes of the solar ^{8}B, ^{7}Be, pep, ^{13}N and ^{15}O neutrinos and the corresponding ^{37}Ar production rates are taken from ref. [4] and ref. [5]. The allowed regions of values of the parameters Δm^2 and $\sin^2 2\theta$ thus obtained at 90% C.L. are given in fig. 1. The irregularities at $\sin^2 2\theta \geq 0.8$ and $\Delta m^2 \approx (0.5\text{–}1.0) \times 10^{-10}$ eV^2 are due to the ^{7}Be-neutrinos, which have a line spectrum and give a strongly time-dependent contribution to the signal in the Cl–Ar experiment.

For the analysis of the Kamiokande II results we proceed in a similar way. First, for the case of oscillations into active neutrinos, e.g. ν_μ, we evaluate the theoretically predicted signal as

$$Q_{\nu e} = \int \sigma_{\nu_e e}(E_\nu) F(E_\nu) \bar{P}_{\nu_e \to \nu_e}(E_\nu) dE_\nu + \int \sigma_{\nu_\mu e}(E_\nu) F(E_\nu) \bar{P}_{\nu_e \to \nu_\mu}(E_\nu) dE_\nu, \tag{8}$$

where σ_{ν_e} (σ_{ν_μ}) is the cross-section for the ν_e (ν_μ) scattering off electrons, $F(E_\nu)$ is the solar neutrino flux spectrum and $\bar{P}_{\nu_e \to \nu_e}$ ($\bar{P}_{\nu_e \to \nu_\mu}$) is the survival (transition) probability determined by eqs. (3) and (6). For the case of sterile neutrinos, i.e., neutrinos that do not contribute at all to the signal in ν_e–e^- scattering experiments, the second term in (8) is dropped. Next, for each pair of values of the parameters Δm^2 and $\sin^2 2\theta$ we compute the corresponding χ^2 values and determine the allowed regions of these parameters in the standard way described above. In our analysis we have used slightly more detailed data in the form of three-months averages of the data sample reported by the Kamiokande II

collaboration [22] *. There are 13 points altogether, and the last point is a four-month average from January to April 1990. The error bars are slightly asymmetric and we have made our analysis with and without symmetrizing them. The results in both cases do not differ substantially. Only ^{8}B neutrinos, which have a continuous spectrum, contribute to the signal in the Kamiokande II detector. Therefore the variations of the Earth–Sun distance cannot produce drastic changes of the average signal rate in this detector. The resulting allowed regions are given in fig. 1 and fig. 2 for the cases of $\nu_e \leftrightarrow \nu_\mu$ and $\nu_e \leftrightarrow \nu_s$ oscillations, respectively. We have not used in our analysis the results of the Kamiokande II collaboration on the spectrum of electrons scattered by the solar ^{8}B neutrinos. The procedure of extraction of the relevant signal from the raw data is described in ref. [22]. It includes, in particular, a maximum likelihood analysis which relies on the prediction for the spectrum of electrons resulting from the elastic scattering of the ^{8}B neutrinos off electrons.

The regions of values of the parameters Δm^2 and $\sin^2 2\theta$ allowed by both the Homestake and Kamiokande II data are given as shaded areas in fig. 1 and fig. 2 for $\nu_e \leftrightarrow \nu_\mu$ and $\nu_e \leftrightarrow \nu_s$ oscillations, respectively. Our results differ somewhat from those obtained in ref. [18]. Whereas our allowed by the Cl–Ar data region of values of Δm^2 and $\sin^2 2\theta$ almost coincides with the corresponding one in [18], the relevant regions allowed by the Kamiokande II data are different. For large $\sin^2 2\theta$ our excluded regions around $\Delta m^2 \approx 7 \times 10^{-11}$ eV^2 are substantially smaller than the corresponding ones in ref. [18]. On the other hand, with our approach we are able to exclude values of $\sin^2 2\theta \leq 0.54$ for $\nu_e \leftrightarrow \nu_\mu$ oscillations and $\sin^2 2\theta \leq 0.47$ for $\nu_e \leftrightarrow \nu_s$ oscillations at the 95% C.L., whereas the corresponding limits from ref. [18] are 0.45 and 0.40, respectively. Note also that taking into account the change of the neutrino oscillation probabilities with the change of the Sun–Earth distance during the experimental runs in which the solar neutrino flux was measured leads to the exclusion by the Cl–Ar and Kamiokande II data of the region of values of Δm^2 and $\sin^2 2\theta$ around $\Delta m^2 \simeq 2.5 \times 10^{-10}$ eV^2 and $\sin^2 2\theta \geq 0.95$ found to be allowed in ref. [18]. We believe that these differences are due to the different treatment of errors as well as to the different procedures for obtaining the limits on the neutrino oscillation parameters. We do not use information about the spectrum of electrons scattered by the solar neutrinos in the Kamiokande II detector and do not take into account the theoretical uncertainty as was done in ref. [18]. In our view a more adequate way to estimate the uncertainty of the obtained limits is to compare the restrictions resulting from different solar model predictions for the solar neutrino fluxes.

The comparison of the allowed regions obtained on the basis of the predictions of the event rates in the solar neutrino detectors made in refs. [4] and [5] shows that using the results of ref. [5] leads to somewhat weaker restrictions on $\sin^2 2\theta$ and Δm^2 than using the results of ref. [4]. This is because the authors of ref. [5] predict signals in the Cl-Ar and Kamiokande II detectors that are lower than those predicted in ref. [4] and therefore the suppression of the solar neutrino flux needed to explain the data is smaller in the former

* We thank Prof. Y. Totsuka for providing us with these data. We do not use the published in ref. [22] five point data because these represent averages over periods of about 200 days and therefore are less sensitive to the possible effects of seasonal variations of the solar ^{8}B neutrino flux.

case.

From the results of the analysis described above we conclude that the neutrino oscillations in vacuum are still a possible solution to the solar neutrino problem. Two-neutrino oscillations of electron neutrinos into active as well as into sterile neutrinos can account for the observed deficit of neutrinos from the Sun in the Cl-Ar and Kamiokande II experiments. * Future experiments with H_2O, D_2O, ^{37}Cl, ^{11}B, ^{40}Ar, ^{7}Li, and ^{19}F with different thresholds and counting rates, can restrict even more the allowed regions of values of the parameters Δm^2 and $\sin^2 2\theta$.

The measurement with a sufficient accuracy of the ^{7}Be neutrino flux over a long period of time will provide, perhaps, the best possibility to test the vacuum oscillation solution of the solar neutrino problem [21, 23].

* For the restrictions on the values of Δm^2 and $\sin^2 2\theta$ following from the data of the SAGE collaboration [3], see ref. [24], where the implications of our results for the Ga-Ge experiments are also analized.

References

1. R. Davis et al., XXI Int. Cosmic Ray Conf., 6-19 January, Adelaide, Australia, Conference papers, vol. 7, p. 155, 1990.
2. K.S. Hirata et al., Phys.Rev.Lett. 63 (1989) 16; 65 (1990) 1297 and 1301.
3. A.I. Abazov et al. (SAGE collaboration), Phys.Rev.Lett. 67 (1991) 332.
4. J.N. Bahcall and R.K. Ulrich, Rev.Mod.Phys. 60 (1988) 298.
5. S. Turk-Chieze et al., Astrophys.J. 335 (1988) 425.
6. J.N. Bahcall, Talk given at the Int. Workshop on Electroweak Physics Beyond the Standard Model, Valencia, Spain, 1991.
7. B. Pontecorvo, Zh.Eksp.Teor.Fiz. 33 (1957) 549.
8. B. Pontecorvo, Zh.Eksp.Teor.Fiz. 34 (1958) 247; 53 1967) 1717.
9. S.P. Mikheyev and A.Yu. Smirnov, Yad.Fiz. 42 (1985) 1441; Nuovo Cimento 9C (1986) 17.
10. L. Wolfenstein, Phys.Rev. D 17 (1978) 2369.
11. V. Barger, S. Pakvasa, R.J.N. Phillips and K. Whisnant, Phys.Rev. D 22 (1980) 2718.
12. M.M. Guzzo, A. Masiero and S.T. Petcov, Phys.Lett. B 260 (1991) 154; M.M. Guzzo and S.T. Petcov, Phys.Lett. B 271 (1991) 172; E.Roulet, Phys.Rev. D 44 (1991) R935.
13. A. Cisneros, Astrophys.Space Sci. 10 (1970) 87.
14. L.B. Okun, M.B. Voloshin and M.I. Vysotsky, Yad.Fiz. 44 (1986) 677.
15. C.-S. Lim and W.J. Marciano, Phys.Rev. D 37 (1988) 1368; E.Kh. Akhmedov, Phys.Lett. B 213 (1988) 64.
16. A.Yu. Smirnov, Phys.Lett. B 260 (1991) 161; E.Kh. Akhmedov, P.I. Krastev and A.Yu. Smirnov, Z. Phys. C 52 (1991) 701.
17. J.N. Bahcall, N. Cabibbo and A. Yahill, Phys.Rev.Lett. 28 (1972) 316; J.N. Bahcall et al., Phys.Lett. B 181 (1986) 369; Z.G. Berezhiani et al., Univ. of Ferrara preprint INFN-FE-05-91, Nov. 1991.
18. V. Barger, R.J.N. Phillips and K. Whisnant, Phys.Rev. D 43 (1991) 1110.
19. A. Acker, S. Pakvasa and J. Pantaleone, Phys.Rev. D 43 (1991) 1754.
20. P.I. Krastev, S.P. Mikheyev and A.Yu. Smirnov, INR preprint P-695, Moscow, 1991; H.A. Bethe and J.N. Bahcall, Phys.Rev. D 44 (1991) 2962; see also: S.P. Rosen and J.M. Gelb, Phys.Rev. D 39 (1989) 3190; J.N. Bahcall and W.C. Haxton, Phys.Rev. D 40 (1989) 931.
21. S.M. Bilenky and B. Pontecorvo, Phys.Rep. 41 (1978) 225; S.M. Bilenky and S.T. Petcov, Rev.Mod.Phys. 59 (1987) 671.
22. K.S. Hirata et al., Phys.Rev.Lett. 66
23. S.T. Petcov and J. Rich, Phys.Lett. B 224 (1989) 426.
24. P.I. Krastev and S.T. Petcov, preprint CERN-TH.6401/92 (to be published in Phys. Lett. B).

Figure captions

Fig.1. Regions of values of the parameters Δm^2 and $\sin^2 2\theta$ allowed at 90% C.L. by the Homestake data (dashed line) and by the Kamiokande II data (solid line) in the case of solar neutrino oscillations into active neutrinos: $\nu_e \leftrightarrow \nu_a$, a = μ or τ. Fig. 1a and fig. 1b were obtained by using the predictions of ref. [4] and ref. [5], respectively, for the fluxes of solar neutrinos and the corresponding event rates in the two detectors in the absence of oscillations.

Fig.2. The same as fig. 1 but for solar neutrino oscillations into sterile neutrinos: $\nu_e \leftrightarrow \nu_s$.

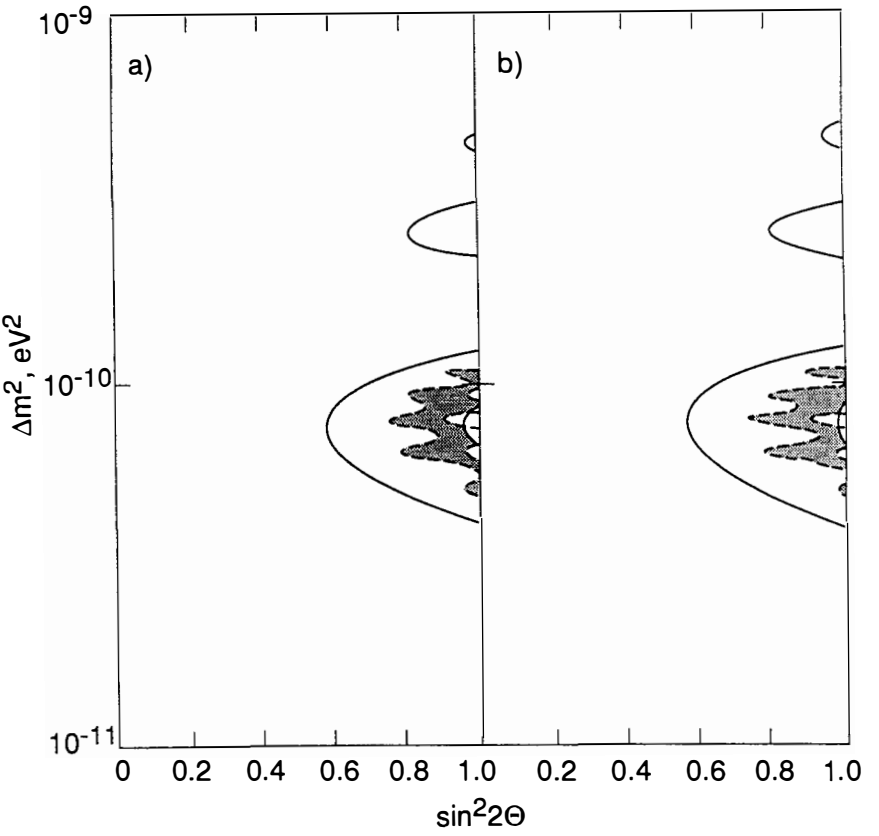


Fig. 1

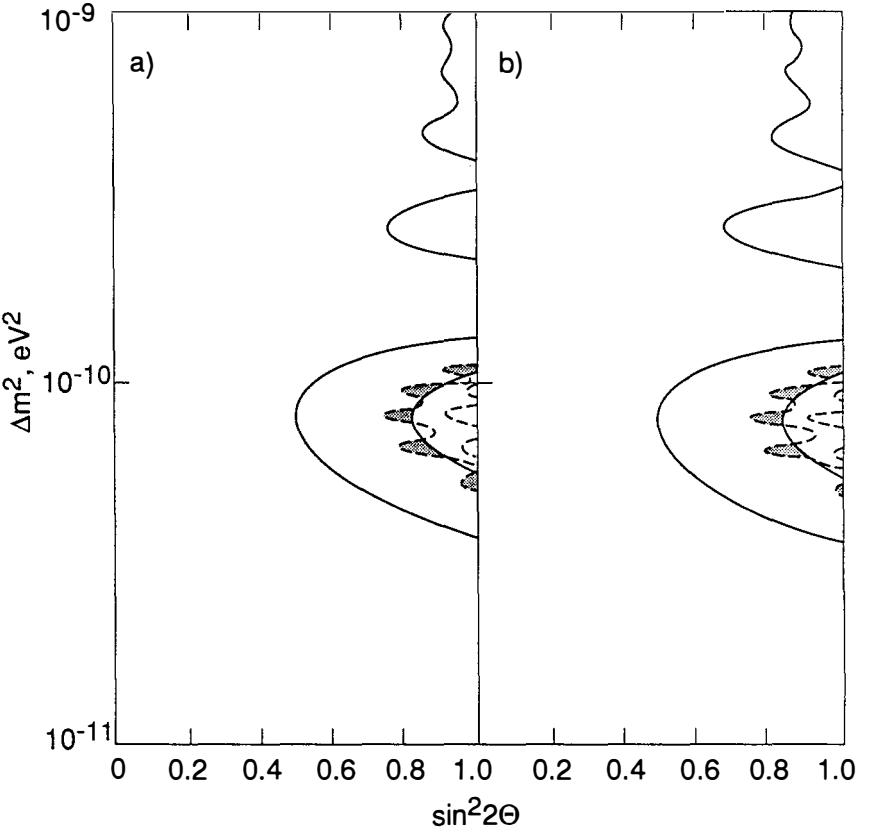


Fig. 2

Status Report on the GALLEX Experiment

presented by L. Zanotti

L.N.G.S.: M. Balata, E. Bellotti, N. Ferrari, H. Lalla, T. Stolarczyk - **M.P.I. Heidelberg:** P. Anselmann, W. Hampel, G. Heusser, J. Kiko, T. Kirsten, E. Pernicka, R. Plaga, B. Povh, U. Roenn, M. Sann, C. Schlosser, H. Völk, R. Wink, M. Wojcik - **KfK Karlsruhe:** R. von Ammon, K. Ebert, T. Fritsch, K. Hellriegel, E. Henrich, L. Stieglitz, F. Weyrich - **Milano (Univ. & I.N.F.N.):** C. Cattadori, O. Cremonesi, E. Fiorini, S. Ragazzi, L.Zanotti - **T.U. München:** F. von Feilitzsch, R. L. Mössbauer, U. Schanda - **Nice:** G. Berthomieu, E. Schatzmann - **Saclay:** S. Charbit, L. Gosset, M. Cribier, G. Dupont, B. Pichard, J. Rich, M. Spiro, C. Tao, D. Vignaud - **Rehovot:** I. Carmi, I. Dostrovsky - **Roma (Univ. & I.N.F.N.):** S. d'Angelo, C. Bacci, P. Belli, R. Bernabei, L. Paoluzi - **Brookhaven:** R. L. Hahn, F. X. Hartmann, J. K. Rowley, R. W. Stoenner, J. Weneser

Abstract

The GALLEX collaboration is performing an experiment for the detection of neutrinos coming principally from the p-p fusion reaction in the Sun, via the reaction $\nu_e + {}^{71}Ga \rightarrow {}^{71}Ge + e^-$. The experiment is actually running in the Gran Sasso Laboratory of I.N.F.N. since more than 1 year, using as a target a solution of $GaCl_3$ containing 30 tons of natural Gallium.

A report is given of the status of GALLEX after the end of the operations devoted to the removal from the solution of the cosmogenically formed ^{68}Ge, which represented a big challenge during the last year, since ^{68}Ge decay has the same signature as ^{71}Ge decay (electron capture). The experiment is now collecting data on solar neutrinos since the middle of 1991, and tha data analysis is in progress; results are expected during 1992.

1 - Introduction

It is well known (and it has been discussed elsewhere in this conference[1,2]) that the solar neutrino problem is principally based on the deficit in the neutrino flux measured by R. Davis with the Chlorine experiment and by the Kamioka collaboration with a water Cerenkov detector; both experiments are sensitive to neutrinos produced by the ^8B decay in the Sun. Since the decay rate is extremely sensitive to parameters of the solar model, the question of existence of neutrino oscillations or other exotic neutrino properties is still controversial.

Two experiments (SAGE and GALLEX) are now trying to cast a new light on the problem by studying the flux of neutrinos from the p-p fusion reaction in the Sun, which is strictly related to the solar luminosity, which is well known ($L_S = 3.9 \times 10^{33}$ erg/s); in case of deficit of neutrinos in this channel a solution of the puzzle can be given by attributing to neutrinos yet unknown properties (oscillations, magnetic moment, etc.). Both the experiments are planning to detect solar neutrinos of the p-p chain by the reaction $\nu_e + {}^{71}Ga \rightarrow {}^{71}Ge + e^-$, first proposed by A. Kuzmin in 1965[3], whose feasibility was checked by a pilot experiment of M.P.I. Heidelberg and B.N.L. in the past years.

The threshold for the reaction is E = 233 keV, well below the endpoint for the p-p neutrinos (E = 420 keV), and the interaction can be detected by looking at the decay of ^{71}Ge by electron capture ($\tau_{1/2}$ = 11.43 d), with suitable counters. The predicted results, according to various solar models, range from 115 to 135 SNU (1 SNU = 10^{-36} ν captures per second and per target nucleus) for all neutrino sources; about 71 SNU come from p-p reaction.

SAGE experiment has already published results that seem to indicate a real deficit in solar neutrino flux[4], while GALLEX started taking data since the middle of 1991, and the analysis is in progress.

2 - Experimental Setup

The target is a solution of 30 tons of natural Gallium, in the form of $GaCl_3$, whose composition is reported in Table 1. The solution is stored since June 1990 in tank B (each tank is 70 m^3; the other tank, named A, is equipped with the thimble for the insertion of a calibration n source) located in a building inside hall A of the Gran Sasso Underground Laboratory (Laboratori Nazionali del Gran Sasso - I.N.F.N., near the town of L'Aquila in the Central Italy). The rock overburden of about 3600 m.w.e. provides a natural shielding against cosmic rays, whose flux is reduced to ~ 1 μ/(m^2 h).

$GaCl_3$	H_2O	HCl	Density
76 ton	20 ton	5 ton	1.9 g/cm^3

Table 1 : Composition of the solution

In the same building all the chemical equipment for the extraction of Germanium is operating, while the counting facility is installed in a nearby building. A Faraday cage is located in it, which contains a large tank of Iron, selected Lead and Copper (to reduce radioactive background) designed to host the proportional counters of the Davis type for the

measurement. Part of the counters (16 maximum) can be put in the purely passive shield, while 8 counters can be put in a NaI counter of the well type, also inside the large shielding tank, in order to check for coincidences. Such a configuration has proved to be useful to identify the chain decay:

^{68}Ge ($\tau_{1/2}$ = 288.6 d, E.C.) --> ^{68}Ga ($\tau_{1/2}$ = 68.3 m, β^+) --> ^{68}Zn

where ^{68}Ge is cosmogenically produced in Ga when stored in presence of cosmic rays.

3 - Exposition and Extraction

The procedure runs as follows. At the beginning of each exposition about 1 mg of Ge isotopes are added to the 53 m^3 solution; it will be used as a check of extraction efficiency and as filling gas for counters in the form of GeH_4. After three or four weeks, corresponding to a fraction of about 75% and 87% of saturation (equilibrium between production and decay of ^{71}Ge) extraction starts by flowing about 2000 m^3 of N_2 in the tank. Since $GeCl_4$ is highly volatile in a high acidity solution, it is swept out of the tank by such flow, and is immediately sent to a series of absorption columns, where it is trapped in pure water (low acidity). It is afterwards concentrated in about 0.5 l of pure water and chemically treated to reduce it to GeH_4 with tritium free water and high purity chemicals; GeH_4 is further purified by gaschromatography, and used to fill the counters with Xe, in the ratio 70% Xe and 30% GeH_4. Extraction and filling efficiency is ~ 90%.

4 - Counting

A 132 SNU ν flux would yield, after 20 exposure days, 11 ^{71}Ge atoms in the counters; it is therefore extremely important to reduce backgrounds. Proportional counters of 1 cm^3 volume, of which 90% is effective, have been built with special care for background reduction[5]; they can detect both the K-capture and L-capture lines at 10.4 keV and 1.3 keV respectively with efficiencies of 36% and 30%. Backgrounds of .01 and .15 counts/day have been measured in the regions of the above quoted peaks after pulse shape analysis to identify pulses due to a localized energy deposition, as expected in the case of electron capture events. Calibration of counters is performed by irradiating them with a Ce source[6], whose effect is to produce a complex series of lines clustering around the energies of 1 keV, 5 keV, 10keV, therefore allowing to check resolution and linearity of the counters. After filling, counters are placed inside appropriate copper boxes which are inserted in the shielding described in the previous section.

5 - ^{68}Ge reduction

^{68}Ge is produced by exposition of the Gallium target to cosmic rays via (p,2n) and (p,4n) reactions on ^{69}Ga and ^{71}Ga. Since Gallium for GALLEX has been exposed for many months to cosmic rays at ground level, a production of ~ 20 million atoms was extimated, but a few extraction cycles should have reduced them to an amount compatible

with the expected backgrounds. It was found instead that, after an initial strong decrease, the amount of ^{68}Ge atoms extracted was proportional to the exposure time, at the level of ~ 100 per day (fig. 1), indicating the presence of a reservoir of Ge trapped in non-volatile chemical species inside the solution. In order to increase the Ge release by the reservoir the solution was heated up to 40 C, starting from November 1990. The effect is shown in fig. 1, from run 14 to run 27; it can be seen that after run 21 a new equilibrium has been reached, in which Ge release is comparable to the preceding cold runs. It was therefore decided to cool down the solution to room temperature, therefore strongly reducing the Ge release; after such operation ^{68}Ge release decreased well below the dotted line in the plot, and the background counting rate reduced to a level consistent with the counters background. Therefore, starting from run 29 (May 1991), the experiment is collecting data for the evaluation of the solar ν flux.

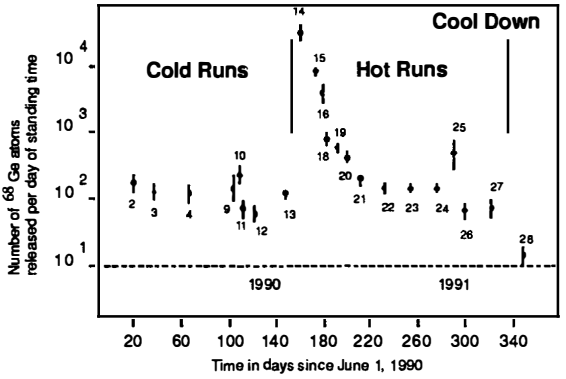


Fig. 1 - Number of ^{68}Ge atoms released per day of standing time as a function of solar time; the dotted line is the level below which a solar neutrino experiment is feasible.

6 - Conclusion

After completion of the installation procedures, and reduction to an acceptable level of the residual ^{68}Ge contamination, the experiment is fully operational since May 1991. Extraction and filling efficiency is ~ 90% and background counting rates are at the expected level of ~ .01 and ~ .15 counts/day in the K and L peak regions. Ten expositions of 21 days each have been performed and counting is proceeding; analysis is in progress and results will be communicated when they will be of statistical relevance.

References

[1] S. Turck-Chieze, this conference
[2] Kamioka collaboration, this conference
[3] A. Kuzmin, Sov. Phys. JETP **49** (1965) 1532
[4] J. Wilkerson, this conference
[5] R. Wink, Ph. D. Thesis, Heidelberg, 1988
[6] A. Urban, Ph. D. Thesis, München, 1989

RECENT RESELTS FROM KAMIOKANDE

Yuichi Oyama
National Laboratory for High Energy Physics (KEK)
Tsukuba, Ibaraki 305
Japan
(for the Kamiokande collaboration*)

Abstract

Recent results on solar neutrinos and atmospheric neutrinos from Kamiokande Experiment are reported.

1 : Introduction

Kamiokande-II, which is the upgraded phase of Kamiokande-I, is an imaging water Cherenkov detector located at 2700 m.w.e. underground in the Kamioka mine, about 300km west of Tokyo. 2140 tons of water in a cylindrical steel tank are viewed by 948 20-inch photomultiplier tubes(PMT's) covering 20% of the tank surface. The inner counter is surrounded by a 4π anticounter layer which contains on average 1.5-m-thick water viewed by 123 20-inch PMT's.

In 1990, we stopped the data-taking for 8 months. During the long shutdown, more than 100 of dead PMTs were replaced and reflective mirrors were added to PMTs. A new electronics system, which is a prototype of Superkamiokande project was also installed. New phase of the experiment, namely Kamiokande-III, was started in December 1990.

2 : Solar neutrinos

In Kamiokande, ^{8}B solar neutrinos can be observed by neutrino-electron scattering, $\nu_e e^- \rightarrow \nu_e e^-$. From event samples of energy larger than 7.5MeV, external γ-rays are removed by limiting the fiducial volume to the central 680 tons. Events due to unstable spallation products by through going muons are removed by examining timing and spatial correlation between preceding muons and low energy events. More detailed analysis procedure are presented in Ref.1. Figure 1 shows the angular correlation of the momentum vector of the selected events with the direction from the Sun. An excess of events in the broad peak at $\cos\theta_{sun} = 1$ is clearly seen. The solar neutrino flux is finally obtained to be

$$\text{Data/SSM} = 0.46 \pm 0.05(\text{stat}) \pm 0.06(\text{syst})$$

at an energy threshold of 7.5MeV, where SSM is a standard solar model prediction.[2]

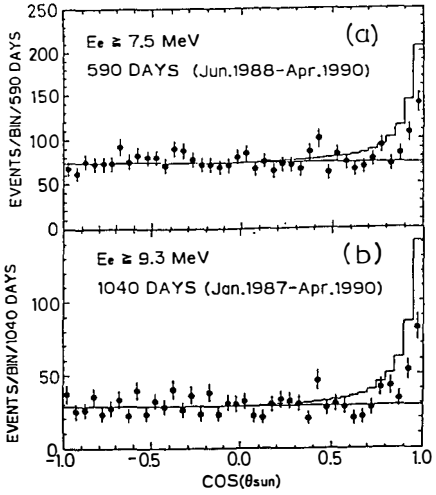


Fig.1. Distribution in $\cos\theta_{sun}$ of; (a) the 590-days data for $E_e \geq 7.5$MeV and (b) the 1040-days data for $E_e \geq 9.3$MeV. θ_{sun} is the angle between the momentum vector of the observed electron and the direction from the Sun. The histogram shows the calculated signal based on SSM value.

The energy distribution of the scattered electrons from the sample is shown in Fig.2. Above $E_e = 7.5$MeV, the shape of the distribution seems to be consistent with the shape of Monte Carlo(MC) simulation of the interaction of ^{8}B solar neutrinos. Energy independence of solar neutrino deficit gives some constraints on the neutrino oscillation parameters.[3]

Time variations of solar neutrino flux with various time scale is also interesting aspect of solar neutrino observation.

Time variation[4] correlated with 11-years cycle of solar activity which was claimed by ^{37}Cl experiment[5] is one of interesting topics. To study this time variation, total data are divided into 6 time interval. Each data has about 200 days observation. The flux of each period is plotted together with sunspot numbers in Fig.3. Clearly, there is no correlation or anti-correlation between them.

A semiannual variation[6] of the solar neutrino flux is possible because of the opening angle between the solar equatorial plane and the ecliptic plane. When a line from the earth to the center of the Sun crosses equatorial plane, the detector views the core of the Sun through the solar equator, where the magnetic field is expected to be weaker than at a higher solar latitude. The strength of the solar magnetic field would be less at those times and consequently a maximum modulation of the solar neutrino flux might occur.

To search for this effect the year was divided into two periods: period I, April 22 $\rightarrow$ July 21 and October 21 $\rightarrow$ January 20, and period II, January 21 $\rightarrow$ April 21 and July 22 $\rightarrow$ October 20. Period I (II) corresponds to the time interval in which the Earth is near (far from) the intersection of the ecliptic plane with the solar equatorial were obtained as described above. The flux of each period are plotted in Fig.4.

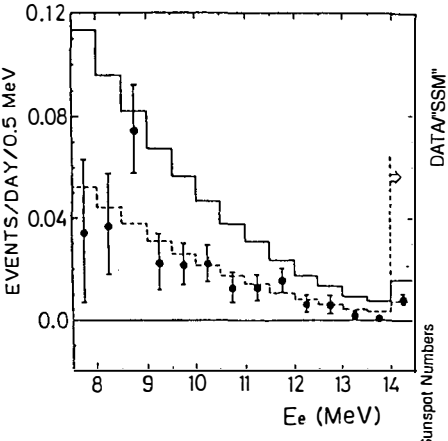


Fig.2. Differential energy distribution of the recoil electrons from $\nu_e e \rightarrow \nu_e e$. The histograms show the best fit to the data (dashed line, 0.46 × SSM) and predicted one by the SSM (solid line).

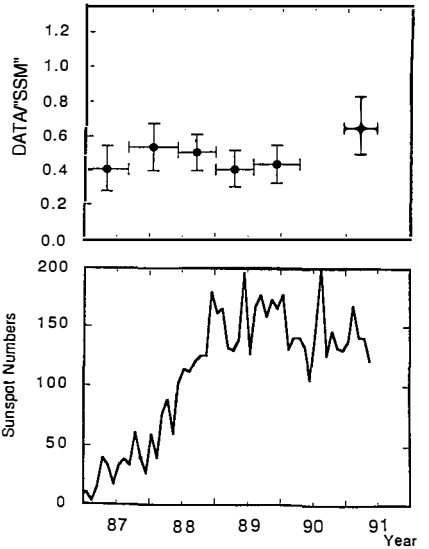


Fig.3. (a) Time variation of solar neutrino signal. (b)sunspot numbers of the same period.

The relative flux difference is expressed by

$$\frac{\text{Period I} - \text{Period II}}{\text{Period I} + \text{Period II}} = -0.06 \pm 0.11(\text{stat.}) \pm 0.02(\text{syst.}).$$

These results do not indicate a significant anticorrelation with the strength of the magnetic field, and consequently contain no evidence for a magnetic interaction of ν_e in the Sun.

The MSW effect[7] in the earth gives rise to a possible difference between the fluxes of day-time and night-time,[6] if the neutrino oscillation with MSW effect in the earth is assumed. The obtained data were divided into several bins as a function of zenith angle. No zenith angle dependence of the flux was found. The difference between the daytime and nighttime fluxes is expressed by;

$$\frac{\text{Day} - \text{Night}}{\text{Day} + \text{Night}} = -0.08 \pm 0.11(\text{stat.}) \pm 0.03(\text{syst.}).$$

This result rejects neutrino oscillation parameters around a region of $\Delta m^2 = 10^{-5} \sim 10^{-6}\text{eV}^2$ and $\sin^2 2\theta \gtrsim 10^{-2}$.

3 : Study of neutrino oscillation using contained events[8]

Atmospheric neutrinos are decay products of pions, kaons and muons which are produced by interaction of primary cosmic rays in atmosphere. In the energy range below 1 GeV, ν_μ/ν_e is essentially $\sim$2 because one charged pion produces two muon neutrinos and one electron neutrino. A study of neutrino oscillation is possible by examining μ-type neutrino events and e-type neutrino events observed in the detector.

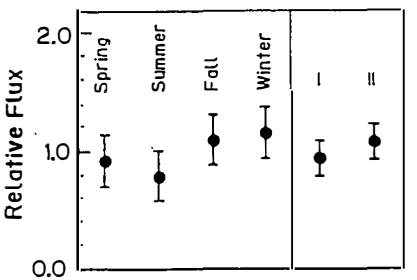


Fig.4. Seasonal and semiannual solar neutrino fluxes relative to the average value.

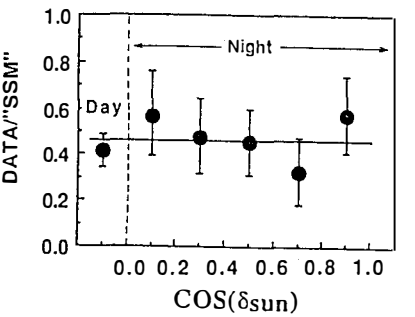


Fig.5. Measured solar neutrino fluxes of daytime and subdivided nighttime relative to the SSM prediction. δ_{sun} is the zenith angle of the Sun relative to the z-axis of the detector ($\delta_{sun} = 0$ corresponds to the direction in which the Sun is just below the horizon.) The average value is also shown by a solid line.

We have observed 457 fully contained events during 4.92 kton·yr of the detector exposure. A selection of contained events is detailed in Ref.8. Of 457 fully contained events, 310 events have single Cherenkov rings and 147 have two or more Cherenkov rings. The 310 single-ring events are then classified as either e- or μ-type.

The particle identification program makes use of the spatial distribution of photoelectrons. Cherenkov rings from electromagnetic showers exhibit more multiple scattering than those from muons or charged pions, thus producing more diffuse rings. The reconstructed opening angle of the Cherenkov cone and the timing information is additionally used. The probability of misidentifying the particle species of single ring events was estimated to be less than 5%, using Monte Carlo simulated neutrino events. Finally, 310 single-ring events are classified into 159 e-type events and 151 μ-type events.

Table 1 summarizes the data and compares them with 3 different Monte Carlo estimations. Figure 4 shows momentum distributions of (a) e-type and (b) μ-type events. The observed number of the e-like single ring events is consistent with the expectation. The number of the μ-like single ring events is, on the other hand, less than expectation.

	Data	M.C.(1)	M.C.(2)	M.C.(3)
Single ring	310	425.5	380.2	332.9
electron-like	159	164.9	146.0	127.7
muon-like	151	260.6	234.2	205.2
Multi ring	147	169.0	155.4	133.0
Total	457	594.5	535.6	465.9

Table.1. Summary of the fully-contained events from 4.92 Kton·year observation. Results from Monte Carlo simulation based on three independent atmospheric neutrino fluxes are also shown. Atmospheric neutrino fluxes are M.C.(1) form Ref.9, M.C.(2) from Ref.10 and M.C.(3) from Ref.11.

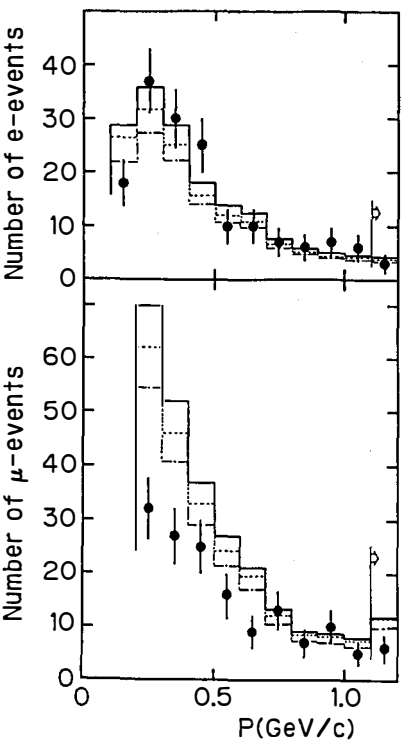


Fig.6. The momentum spectra of e- and μ-events. The histograms show the M.C. predictions based on M.C.(1)(solid), M.C.(2)(dot) and M.C.(3)(dash-dot).

Because of a large uncertainty in the absolute value of the expectation (±30%), we compared the data and the MC estimation using a ratio, $R(\mu/e)$, defined by

$$R(\mu/e) \equiv \frac{(\text{number of observed } \mu \text{ event})/(\text{number of expected } \mu \text{ event})}{(\text{number of observed e event})/(\text{number of expected e event})}.$$

R ratio with MC(1) is calculated to be

$$R(\mu/\text{e}) = 0.60^{+0.07}_{-0.06}(\text{syst.}) \pm 0.05(\text{stat.}).$$

We are unable to explain the data as the result of systematic detector effects or uncertainties in the atmospheric neutrino fluxes.

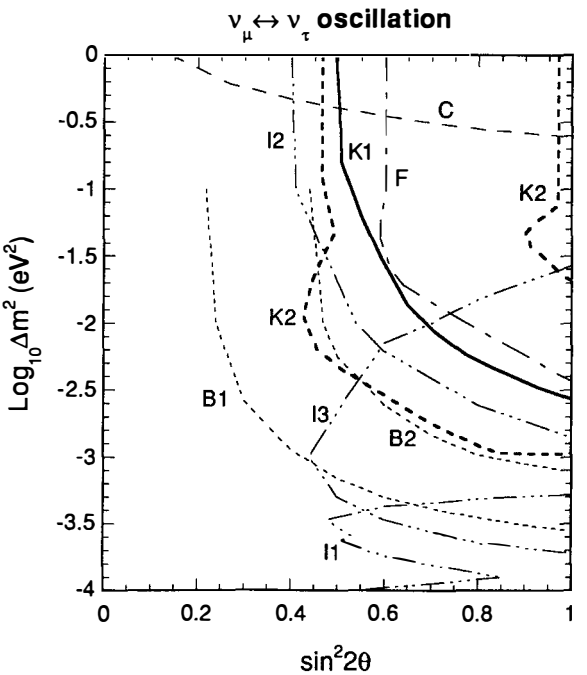


Fig.7. Excluded regions in $\sin^2 2\theta - \Delta m^2$ plane of $\nu_\mu \leftrightarrow \nu_\tau$ oscillation. Notations are; K1; Kamiokande upward-going muons, K2; Kamiokande contained events,[8] B1,B2; Baksan[12] (for two different neutrino flux calculations) C; CDHSW,[13] F; Frejus,[14] I1,I2,I3; IMB[15] (for different data sets). The right-hand sides of the curves are excluded except for K2. K2 curves surround the regions to be compatible with the ν_μ deficit.

A possible explanation of the data may be sought in neutrino oscillation in vacuum and/or in matter. Constraints on oscillation parameters are carefully studied using minimum χ^2 method shown in Ref.8. Results of the calculation for $\nu_\mu \leftrightarrow \nu_\tau$ oscillation is shown in Fig.7 together with other experiments. A region $\sin^2 2\theta \sim 1$ and $\Delta m^2 \gtrsim 10^{-3}\mathrm{eV}^2$ is favored as neutrino oscillation parameters.

4 : Study of neutrino oscillation using upward-going muons[16]

Atmospheric muon neutrinos of higher ($\gtrsim$10GeV) energy range can be observed as upward-going muon events, which are produced in the surrounding rock and penetrate the detector upward. Comparison of the upward-going muon flux with the theoretical expectation provides another information about the neutrino oscillation hypothesis.

We have observed 252 upward-going muon events which penetrate the detector with path length larger than 7 m from 1920 days data. Average flux of upward-going muons is calculated to be

$$(2.04 \pm 0.13) \times 10^{-13}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}.$$

On the other hand, expected upward-going flux can be calculated from atmospheric neutrino flux, weak interaction in the surrounding rock and muon energy loss in the rock. The result of the calculation is found to be

$$1.92 \times 10^{-13}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}.$$

Zenith angle distribution of the experimental and the theoretical flux are shown in Fig.8. They are in good agreement.

Agreement between both fluxes leads to a constraints on the neutrino oscillation parameters. Excluded region in $\sin^2 2\theta - \Delta m^2$ plane for $\nu_\mu \leftrightarrow \nu_\tau$ oscillation is shown in Fig.7. The 90% C.L. limits are $\Delta m^2 = 2.7 \times 10^{-3}\mathrm{eV}^2$, for $\sin^2 2\theta = 1$ and $\sin^2 2\theta = 0.51$ for large Δm^2.

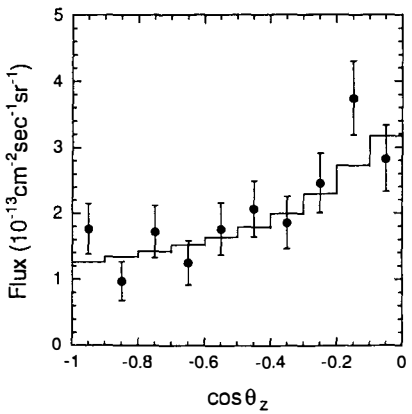


Fig.8. The zenith angle distribution of 252 upward-going muon fluxes in Kamiokande. The expected upward-going muon fluxes calculated from the atmospheric neutrino flux of Volkova,[17] the charged current cross section with the quark distributions given by Eichten et al.,[18] and the muon energy loss formula of Lohmann et al.[19] are shown by a solid line histogram.

*The Kamiokande collaboration includes E.W.Beier, L.R.Feldscher, E.D.Frank, W.Frati, Y.Fukuda, K.S.Hirata, K.Inoue, T.Ishida, T.Kajita, K.Kaneyuki, K.Kihara, S.B.Kim, T.Kajimura, E.Kodera, M.Koshiba, A.K.Mann, K.Miyano, H.Miyata, M.Mori, Y.Nagashima, M.Nakahata, K.Nakamura, F.M.Newcomer, Y.Oyama, A.Sakai, N.Sato, T.Suda, A.Suzuki, K.Takahashi, H.Takei, Y.Takeuchi, M.Takita, T.Tanimori, Y.Totsuka, R.Van.Berg, Y.Yaginuma, M.Yamada, H.Yokoyama and W.Zhang.

1 K.S.Hirata et al., Phys.Rev.D 44, 2241(1991)
2 J.N.Bahcall and R.K.Ulrich, Rev.Mod.Phys. 60, 297(1988)
3 K.S.Hirata et al., Phys.Rev.Lett 65, 1301(1990)
4 K.S.Hirata et al., Phys.Rev.Lett 65, 1297(1990)
5 R. Davis, Jr., et al., Phys.Rev.Lett. 20, 1205(1968)
K.Lande et al., in Proceedings of the 25th International Conference of High Energy Physics, Singapore, 1990 (World Scientific, Singapore, 1991)
6 K.S.Hirata et al., Phys.Rev.Lett 66, 9(1991)
7 L.Wolfenstein, Phys.Rev.D 17, 2369(1978)
S.P.Mikheyev and A.Yu.Smirnov, Nuovo Cimento 9C, 17(1986)
H.A.Bethe, Phys.Rev.Lett 56, 1305(1986)
8 K.S.Hirata et al., ICRR-Report-263-92-1, Phys.Lett.B in press
9 G.Barr, T.K.Gaisser and T.Stanev, Phys.Rev.D 39, 3532(1989)
10 M.Honda, K.Kasahara, K.Hidaka and S.Midorikawa, Phys.Lett.B 248, 193(1990)
11 H.Lee and Y.S.Koh, Nuovo Cimento 105B, 883(1990)
12 M.M.Boliev et al., Third International Workshop on Neutrino Telescopes p.235(Feb. 1991)
13 F.Dydak et al., Phys.Lett.B 207, 79(1988)
14 Ch.Berger et al., Phys.Lett.B 245, 305(1990)
15 R.M.Bionta et al., Phys.Rev.D 38, 768(1988)
R.Becker-Szendy et al., Preprint LSU-HEPA-3-92(1992)
16 Results of this section are update of Y.Oyama et al., Phys.Rev.D 39, 1481(1989), which is based on 146 upward-going events.
17 L.V.Volkova, Yad.Fiz. 31, 784(1980)[Sov.J.Nucl.Phys. 31, 1510(1980)]
18 E.Eichten et al., Rev.Mod.Phys. 56, 579(1984); ibid. 58, 1065(1986)(E)
19 W.Lohmann, R.Kopp and R.Voss, CERN Report 85-03(1985)

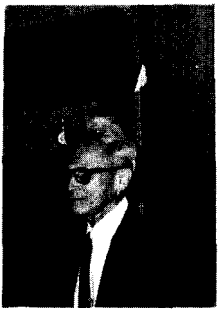
SOLAR NEUTRINO OBSERVATIONS

J. Weber

University of Maryland, College Park, Maryland 20742

and

University of California, Irvine, California 92717



Abstract

Single crystals with high Debye temperatures have large scattering cross sections for neutrinos and antineutrinos. Elastic scattering observations give a solar neutrino flux

$$\Phi = (6 \pm 2) \times 10^{10} \text{ solar neutrinos cms}^{-2} \text{ sec}^{-1} \quad ,$$

in good agreement with predictions. The neutral current interaction process has the same cross section for all types of neutrinos. Solar neutrino oscillations do not affect this result.

Introduction

When light falls upon a screen with two slits, a well defined diffraction pattern is formed on a distant screen. This diffraction pattern is destroyed if apparatus is provided to determine through which slit the photon passes. Suppose the slit separation is L, and the wavelength is λ. The photon momentum $p = 2\pi\hbar / \lambda$. Let the width of each slit be W. The first diffraction minimum occurs at an angle θ_D given by

$$\theta_D = \frac{\lambda}{2L} = \frac{\pi\hbar}{pL} \tag{1}$$

Measurement to determine through which slit a photon passes requires a precision in position at least $\pm W$. This leads to uncertainty Δp in a direction normal to the photon propagation direction,

$$\Delta p \sim \frac{\hbar}{W} \tag{2}$$

The spread due to the momentum uncertainty is θ_u given by

$$\theta_u \sim \frac{\Delta p}{p} = \frac{\hbar}{Wp} >> \theta_D \tag{3}$$

because L >> W. Identification of the slit through which the photon passes therefore destroys the diffraction pattern by uncertainty spread of the beam.

Large Cross Sections for Neutrinos and Antineutrinos

If a neutrino is scattered by quarks in a single nucleus in a liquid, momentum is exchanged with the nucleus. Measurements made before and after scattering can reveal the scattering site. If a solid such as lead is employed, scattering of a neutrino by a single nucleus results in momentary displacement and return of the nucleus to its original position by elastic forces. Phonons are emitted and these permit identification of the scattering site. Since the scatterer site can be identified, no large coherence effects are possible with different nuclei for the liquid, and the lead scatterers.

The situation is different if a single nearly perfect crystal is employed, with high Debye temperature. Such a crystal exhibits the Mossbauer effect. If a nucleus emits a gamma ray, momentum is exchanged. The crystal stiffness is so great that phonons are not excited. The exchanged momentum causes recoil of the entire crystal as a single entity.

If a neutrino is scattered by a single nucleus in such a crystal, momentum is exchanged, no phonons may be excited. The scattering site cannot be identified, large coherent effects are possible.

For single unidentifiable scatterer sites, perfect crystals with infinitely high Debye temperatures, the total cross section for elastic scattering is given by

$$\sigma_{NI} = \frac{G^2 E_\nu^2}{4\pi\hbar^4 c^4}\left[N_T - Z_T\left(1 - 4\sin^2\theta_W\right)\right]^2 \qquad (4)$$

In (4), G is the Fermi weak interaction coupling constant, N_T is the total number of neutrons, Z_T is the total number of protons, θ_W is the Weinberg angle. Important corrections are required, for finite Debye temperatures, for crystal defects, and for other processes in which groups of unidentifiable scatterers contribute.

Solar neutrino experiments were carried out employing a torsion balance. A 26 gram crystal was on one side and an equal mass of lead was on the other side.

As the sun apparently rotates around the balance, periodic torques are produced. The balance room is temperature controlled to a tolerance $\pm\ .001C$. The lead and crystal masses are exchanged periodically. Mr Gregory Wilmot analyzed magnetic tape recorded data, performing averages over approximately 100 days. As already noted, important corrections are required and these are included when a nuclear reactor is employed for calibration. With a known antineutrino source and momentum distribution, the crystal macroscopic force is measured. (4) is integrated over the momentum distributions for the solar neutrino and reactor data. The computer analyses give for the solar neutrino flux

$$\Phi = \left(6 \pm 2 \times 10^{10}\ \text{neutrinos cms}^{-2}\ \text{sec}^{-1}\right) \qquad (5)$$

Neutral current interactions leading to (4) give the same cross section for electron, muon, and tau neutrinos. Therefore (5) is not affected by neutrino oscillations.

Conclusion

The expected number of solar neutrinos is observed.

References

1. J. Weber Phys. Rev. C, 31, 1468, April 15, 1985.
2. J. Weber Phys. Rev. D, 38, 1, 32-39, July 1, 1988.

RESONANT SPIN-FLAVOR PRECESSION OF NEUTRINOS AS A POSSIBLE SOLUTION TO THE SOLAR NEUTRINO PROBLEM

Eugueni Kh. Akhmedov *
Scuola Internazionale Superiore di Studi Avanzati
Strada Costiera 11, I-34014 Trieste, Italy

ABSTRACT
Recent developments of the resonant neutrino spin-flavor precession scenario and its applications to the solar neutrino problem are reviewed. We discuss in particular the possibilities of reconciliation of strong time variations of the solar neutrino flux observed in the Homestake ^{37}Cl experiment with little or no time variation seen in the Kamiokande II experiment.

*on leave from Kurchatov Institute of Atomic Energy, Moscow 123182, Russia

1. Introduction

There are two issues in the solar neutrino problem:

(1) the deficiency of solar neutrinos observed in the Homestake[1], Kamiokande II[2] and, most recently, SAGE [3] experiments;

(2) time variation of the solar neutrino flux in anticorrelation with solar activity (11-yr variations) for which there is a strong indication in the chlorine experiment of Davis and his collaborators but which is not seen in the Kamiokande data.

In this talk I will discuss mainly the second issue with the emphasis on various possibilities of conciliation of the strong time variations in the Homestake experiment with no or little variation in Kamiokande II.

The most natural explanation of the time variation of the solar neutrino flux is related to the possible existence of a large magnetic or electric dipole moments of neutrinos, $\mu \sim 10^{-11}\mu_B$. As was pointed out by Vysotsky, Voloshin and Okun (VVO)[4,5], strong toroidal magnetic field in the convective zone of the sun $B_\perp$ could then rotate left-handed electron neutrinos ν_{eL} into right-handed ν_{eR} which escape the detection. In the periods of quiet sun the solar magnetic field is much weaker and the neutrino spin precession is less efficient which explains the 11-yr variation of the neutrino flux.

Subsequently, it was noted[5,6] that the matter effects can suppress the neutrino spin precession. The reason for this is that ν_{eL} and ν_{eR} are not degenerate in matter since ν_{eL} interact with medium whereas ν_{eR} are sterile, and their energy splitting reduces the precession probability. It was also shown [7] that, unlike in the MSW effect case, the adiabaticity may play a bad role for the VVO effect resulting in a reflip of neutrino spin and thus reducing the probability of $\nu_{eL} \to \nu_{eR}$ transition. In order to break the adiabaticity, the precession length should be large as compared to the characteristic lengths over which matter density and magnetic field vary significantly, which gives an upper bound on $\mu B_\perp$. This parameter should be also bounded from below in order for the precession phase not to be too small. Therefore one gets a rather narrow range of allowed values of $\mu B_\perp$ [7].

Another interesting possibility is the neutrino spin-flavor precession (SFP) due to the interaction of flavor-off-diagonal (transition) magnetic or electric dipole moments of neutrinos μ_{ij} with transverse magnetic fields[8,5]. The SFP is the rotation of neutrino spin with its flavor being simultaneously changed. Such a process can occur even for Majorana neu-

trinos since the CPT invariance does not preclude the transition magnetic dipole moments of Majorana particles. Until recently, the neutrino SFP has not attracted much attention because it was expected to be suppressed by the energy splitting of the neutrinos of different species. If the "Zeeman energy" $\mu_{ij}B_\perp$ is small as compared to the kinetic energy difference $\Delta m^2_{ij}/2E$, the SFP probability is heavily suppressed. However, in 1988 it was noted independently by the present author[9,7] and by Lim and Marciano[10] that in matter the situation can change drastically. Since ν_{eL} and right-handed neutrinos or antineutrinos of another flavor interact with matter differently, the difference of their potential energies can cancel their kinetic energy difference resulting in a resonant amplification of the SFP. Therefore in matter the SFP of neutrinos can be enhanced, unlike the VVO neutrino spin precession [1]. The resonant spin-flavor precession (RSFP) of neutrinos has also some more advantages as compared to the VVO mechanism:

- the adiabaticity plays a good role for the RSFP increasing the conversion probability, and therefore the $\mu_{ij}B_\perp$ should be bounded only from below; the required magnitude of this parameter is a factor of $2-3$ smaller than that for the VVO effect;

- some energy dependence of the neutrino conversion seems to be necessary to reconcile the Homestake and Kamiokande II data (see below). The RSFP probability has the desired energy dependence whereas the VVO neutrino spin precession is energy independent.

Although the above arguments disfavor the VVO effect as a solution of the solar neutrino problem, they do not rule it out, given the uncertainties of the experimental data.

2. General features of RSFP of neutrinos

The RSFP of neutrinos is analogous to the resonant neutrino oscillations [13,14], but differs from the latter in a number of important respects. The main features of this effect have been discussed in detail in my talk at the last Moriond meeting[15], and so I will just briefly mention them here.

The magnetic-field induced mixing of ν_{eL} and $\nu_{\mu R}(\bar{\nu}_{\mu R})$ can be described by the mixing angle θ,

$$\tan 2\theta = \frac{2\mu_{e\mu}B_\perp}{\sqrt{2}G_F(N_e - \alpha N_n) - \frac{\Delta m^2_{e\mu}}{2E}\cos 2\theta_0} \qquad (1)$$

Here N_e and N_n are the electron and neutron number densities, $\alpha = 1/2$ for Dirac neutrinos

[1]The VVO neutrino spin rotation can also be resonantly enhanced provided the magnetic field twists along the neutrino trajectory, see [11,12] and below.

and 1 for Majorana neutrinos, G_F is the Fermi constant, and θ_0 is the ordinary neutrino mixing angle in vacuum. The resonant density is defined as a density at which the mixing angle θ becomes $\pi/4$:

$$\sqrt{2}G_F(N_e - \alpha N_n)|_r = \frac{\Delta m^2_{e\mu}}{2E}\cos 2\theta_0 \qquad (2)$$

The efficiency of the $\nu_{eL} \to \nu_{\mu R}(\bar{\nu}_{\mu R})$ transition is defined by the degree of the adiabaticity which depends on both the neutrino energy and magnetic field strength at the resonance:

$$\lambda \equiv \pi\frac{\Delta r}{l_r} = 8\frac{E}{\Delta m^2_{e\mu}}(\mu_{e\mu}B_{\perp r})^2 L_\rho \qquad (3)$$

Here

$$\Delta r = \frac{8E\mu_{e\mu}B_{\perp r}}{\Delta m^2_{e\mu}}L_\rho \qquad (4)$$

is the resonance width, $l_r = \pi/\mu_{e\mu}B_\perp$ is the precession length at the resonance and L_ρ is the characteristic length over which matter density varies significantly in the sun. For the RSFP to be efficient, λ should be > 1. In non-uniform magnetic field the field strength at resonance $B_{\perp r}$ depends on the resonance coordinate and so, through eq. (2), on neutrino energy. Therefore the energy dependence of the adiabaticity parameter λ in eq. (3) is, in general, more complicated than just $\lambda \sim E$, and is defined by the magnetic field profile inside the sun [2]. The main difficulty in the analyses of the RSFP as a possible solution of the solar neutrino problem is that this profile is essentially unknown, so that one is forced to use various more or less plausible magnetic field configurations.

If the ν_{eL} are produced at a density which is much higher than the resonant one, then in the adiabatic regime ($\lambda >> 1$) the ν_{eL} survival probability is

$$P(\nu_{eL} \to \nu_{eL}) \approx \cos^2\theta_f \qquad (5)$$

where θ_f is the mixing angles (1) on the surface of the sun. Since the magnetic field becomes very weak at the sun's surface, the mixing angle $\theta_f \approx \pi/2$, and so the ν_{eL} survival probability is very small. The adiabaticity parameter λ in eq. (3) depends drastically on the magnetic field strength at resonance, which gives a natural explanation of time variations of the solar magnetic flux in anticorrelation with solar activity.

The RSFP requires non-vanishing flavor-off-diagonal magnetic dipole moments of neutrinos and so is only possible if the neutrino flavor is not conserved. Therefore neutrino

[2] Note that for the MSW effect the adiabaticity parameter is inversely proportional to E [13].

oscillations must also take place, and in general one should consider the SFP and oscillations of neutrinos jointly. This have been done in a number of papers both analytically[16,17] and numerically[10,16,18,19,17] . It was shown that a subtle interplay between the RSFP and the MSW resonant neutrino oscillations can occur. In particular, although the resonant neutrino oscillations cannot give rise to the time variations of the solar neutrino flux, they can assist the RSFP to do so by improving the adiabaticity of the latter[17].

3. Neutrino spin precession in twisting magnetic fields

If the magnetic field changes its direction along the neutrino trajectory, this can result in new interesting phenomena. In particular, new kinds of resonant neutrino conversions become possible, the energy dependence of the conversion probability can be significantly distorted and the lower limit on the value of $\mu B_\perp$ required to account for the solar neutrino problem can be slightly relaxed[11,12]. Moreover, if the neutrino oscillations are also taken into account, the transitions $\nu_e \to \bar{\nu}_e$ can become resonant, and the order of the RSFP and MSW resonances can be interchanged[20].

Since the main features of the resonant neutrino spin-flip transitions in twisting magnetic fields are discussed in some detail in the contributions of Krastev and Toshev in this volume, I will confine myself to a new development which was not covered in their talks.

A few years ago, Vidal and Wudka[21] claimed that the field rotation effects can greatly enhance the neutrino spin-flip probability and reduce the needed value of $\mu B_\perp$ by a few orders of magnitudes. In [11,12] it was shown that this result is incorrect and typically the required value of $\mu B_\perp$ can only be reduced by a factor 2–3 (see also [22,23] in which the process without matter effects was considered). However, in these papers it was not proved that there cannot exist a rotating field configuration giving stronger enhancement of the spin-flip probability and larger gain in the $\mu B_\perp$ parameter. Recently, Moretti [24] has found a severe constraint on the transition probability which eliminates even this possibility. The transition probability $P(\nu_{eL} \to \nu_R)$ turns out to have the following upper bound which does not depend on the magnetic field rotation angle $\phi(t)$ [24]:

$$P(\nu_{eL} \to \nu_R; t) \leq \mu \int_0^t B_\perp(t')\, dt' \tag{6}$$

The analogous result can also be obtained for the neutrino oscillations in matter as well as for the evolution of any other two-level system.

4. RSFP and antineutrinos from the sun

If both the SFP and oscillations of neutrinos can occur, this will result in the conversion of a fraction of solar ν_e into $\bar{\nu}_e$ [10,16,25,26]. For Majorana neutrinos, the direct $\nu_e \to \bar{\nu}_e$ conversions are forbidden since the CPT invariance precludes the diagonal magnetic moment μ_{ee}. However, this conversion can proceed as a two-step process in either of two ways:

$$\nu_{eL} \xrightarrow{\text{oscill.}} \nu_{\mu L} \xrightarrow{\text{SFP}} \bar{\nu}_{eR} \tag{7}$$

$$\nu_{eL} \xrightarrow{\text{SFP}} \bar{\nu}_{\mu R} \xrightarrow{\text{oscill.}} \bar{\nu}_{eR} \tag{8}$$

One can then consider two possibilities:

(1) both oscillations and SFP take place inside sun[10,16,25]. The amplitudes of the processes (7) and (8) have opposite signs since the matrix of the magnetic moments of Majorana neutrinos is antisymmetric. Therefore there is a large cancellation between these two amplitudes (the cancellation is exact in the limit of vanishing neutron density N_n), and the probability of the $\nu_e \to \bar{\nu}_e$ conversion inside the sun turns out to be about 3–5% even for large mixing angles θ_0[16,25].

(2) Only the RSFP transition $\nu_{eL} \to \bar{\nu}_{\mu R}$ occurs in the sun with an appreciable probability whereas the oscillations of neutrinos proceed mainly in vacuum on their way between the sun and the earth [eq. (8)]. For not too small neutrino mixing angles the probability of the $\nu_e \to \bar{\nu}_e$ conversion can then be quite sizable[26].

In[27] the background events in the Kamiokande II experiment were analysed and a stringent bound on the flux of $\bar{\nu}_e$ from the sun was obtained: $\Phi(\bar{\nu}_e) \leq (0.05 - 0.07)\Phi(\nu_e)$. This poses a limit on the models in which both the RSFP and neutrino oscillations occur: the mixing angle θ_0 should be less than $6 - 8°$. This rules out the models with the large magnetic moments of pseudo Dirac neutrinos including those with only one neutrino generation for which θ_0 is the mixing between ν_{eL} and sterile $\bar{\nu}_{eL}$[28,29]. However, the models with a conserved lepton charges $L_e \pm (L_\mu - L_\tau)$ are not excluded even though the mixing angle is $\pi/4$, since the $\nu_e \to \bar{\nu}_e$ conversion probability vanishes identically in this case[30].

The $\bar{\nu}_e$ production due to the combined effect of the RSFP and oscillations of neutrinos can be easily distinguished from the other mechanisms of $\bar{\nu}_e$ generation (like $\nu \to \bar{\nu}$+Majoron decay) since (i) the neutrino flux should vary in time in *direct* correlation with solar activity, and (ii) the neutrino energy is not degraded in this case[16,25]. The $\bar{\nu}_e$ flux from the sun of the order of a few per cent of the expected ν_e flux should be detectable in the forthcoming solar neutrino experiments like BOREXINO, SNO and Super-Kamiokande[25,26,31].

5. Reconciling the Homestake and Kamiokande II data

It has been mentioned above that while there is a strong indication in favor of time variation of the neutrino detection rate in the Homestake data, the Kamiokande experiment does not see such a time variation. It still cannot rule out a small ($\leq 30\%$) time variation. Therefore a question naturally arises as to whether it is possible to reconcile large time variations in the Homestake ^{37}Cl experiment with small time variation in the water Čerenkov experiment. There are two major differences between these two experiments which could in principle give rise to different time variations of their detection rates:

(1) Homestake experiment utilizes the $\nu_e - {}^{37}$Cl charged current reaction, while in the Kamiokande detector $\nu - e$ scattering is used which is mediated by both charged and neutral currents;

(2) the energy threshold in the Homestake experiment is 0.814 MeV so that it is sensitive to high energy ^{8}B, intermediate energy ^{7}Be and partly to low energy *pep* neutrinos; at the same time the energy threshold in the Kamiokande II experiment is 7.5 MeV and so it is only sensitive to the high-energy fraction of the ^{8}B neutrinos.

In[32,17] it was noted that if the lower-energy neutrino contributions to the chlorine detection rate are suppressed stronger than that of high-energy neutrinos, the latter can vary in time with smaller amplitude and still fit the Homestake data. In that case one can expect weaker time variations in the Kamiokande II experiment. The desired suppression of the low-energy neutrino flux can be easily explained in the framework of the RSFP scenario as a consequence of flavor-changing spin-flip conversion due to a strong inner magnetic field, the existence of which seems quite plausible[33]. The alternative possibility is the suppression of low energy neutrinos by the MSW effect when RSFP and the resonant neutrino oscillations operate jointly. Another important point is that due to the RSFP solar ν_e are converted into $\bar{\nu}_{\mu R}$ or $\bar{\nu}_{\tau R}$ which are sterile for the chlorine detector but can be detected (though with a smaller cross section) by water Čerenkov detectors. This also reduces the amplitude of the time variation in the Kamiokande II detector. If both these factors are taken into account, becomes possible to reconcile the Homestake and Kamiokande data; one can expect a low signal in the gallium experiments in this case since they are primarily sensitive to low energy neutrinos whose flux is supposed to be heavily suppressed[32,17].

A similar possibility has been recently considered by Babu, Mohapatra and Rothstein[31]

and by Ono and Suematsu[35]. They pointed out that due to the energy dependence of the RSFP neutrino conversion probability, lower-energy neutrinos can exhibit stronger time variations (i.e. stronger magnetic field dependence) than the higher-energy ones. In fact, this is very natural in the RSFP scenario: with increasing neutrino energy the width of the resonance increases [see eq. (4)] and at sufficiently high energies it can be a significant fraction of the solar radius. The neutrino production point can then happen to be inside the resonant region, which reduces the conversion efficiency. The different magnetic field dependence of the Homestake and Kamiokande II detection rates is illustrated by the figures which we borrowed from ref.[34].

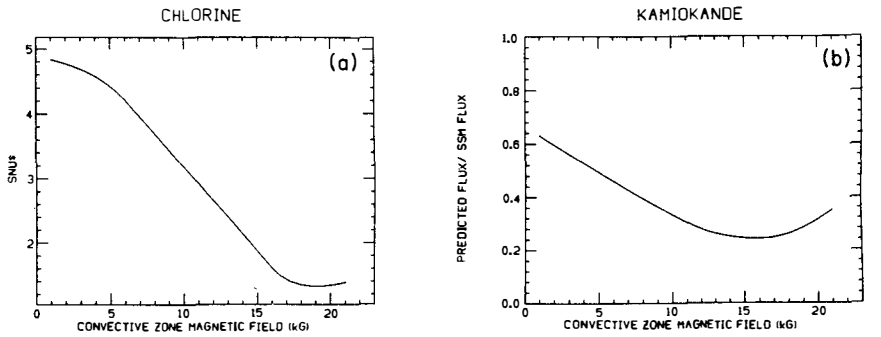


Fig. 1. (a) Expected event rate in chlorine as a function of the convective zone magnetic field. Here $\Delta m^2 = 7.8 \times 10^{-9}\,\mathrm{eV}^2$, the maximal value of the magnetic field in the core $B_i = 10^7\,G$ and $\mu = 2 \times 10^{-11}\mu_B$. (b) The same as (a) but for the Kamiokande event rate.

It should be noted that the ordinary VVO neutrino spin precession lacks energy dependence which is required to get smaller time variation in the Kamiokande II experiment. Moreover, it converts ν_{eL} into sterile ν_{eR} (unless the neutrinos are Zeldovich-Konopinski-Mahmoud particles) which do not contribute to the $\nu - e$ cross section. However, for the VVO scenario yet another possibility of reconciliation of the Homestake and Kamiokande data exists. In order to get sizable magnetic moments of neutrinos, $\mu \approx 10^{-11}\mu_B$, one has to go beyond the Standard Model. Most of the models producing large neutrino magnetic moments are based on various extensions of the Standard Model containing new charged scalars. In these models right-handed sterile neutrinos can interact with electrons via scalar exchange and therefore can contribute to the $\nu - e$ reaction which increases the signal in the Kamiokande II detector and reduces the amplitude of its time variation[36]. Note that the

models giving large transition neutrino magnetic moments usually also contain new scalars and therefore the same mechanism can be operative in case of the RSFP as well.

6. Conclusion

We conclude that the resonant neutrino spin-flavor precession mechanism provides a viable explanation of the solar-neutrino problem which complies with all the existing experimental data and yields a number of interesting predictions for the forthcoming experiments.

REFERENCES

1. R. Davis, Jr., In *Neutrino '88*, Proc. XIXth Int. Conf. on Neutrino Physics and Astrophysics, Boston, USA, 1988, ed. by J. Schneps *et al.* (World Scientific, Singapore, 1989), p. 518.
2. Kamiokande II Collaboration, K.S. Hirata *et al.*, Phys. Rev. Lett. 66 (1991) 9.
3. SAGE Collaboration, A.I. Abazov et al., Phys. Rev. Lett. 67 (1991) 3332.
4. M.B. Voloshin, M.I. Vysotsky, Sov. J. Nucl. Phys. 44 (1986) 845.
5. M.B. Voloshin, M.I. Vysotsky, L.B. Okun, Sov. Phys. JETP 64 (1986) 446.
6. R. Barbieri, G. Fiorentini, Nucl. Phys. B304 (1988) 909.
7. E.Kh. Akhmedov, Phys. Lett. B213 (1988) 64.
8. J. Schechter, J.W.F. Valle, Phys. Rev. D24 (1981) 1883.
9. E.Kh. Akhmedov, Sov. J. Nucl. Phys. 48 (1988) 382.
10. C.-S. Lim, W.J. Marciano, Phys. Rev. D37 (1988) 1368.
11. A.Yu. Smirnov, Phys. Lett. B260 (1991) 161.
12. E.Kh. Akhmedov, P.I. Krastev, A.Yu. Smirnov, Z. Phys. C52 (1991) 701.
13. S.P. Mikheyev, A.Yu. Smirnov, Sov. J. Nucl. Phys. 42 (1985) 913.
14. L. Wolfenstein: Phys. Rev. D17 (1978) 2369.
15. E.Kh. Akhmedov, In *Massive Neutrinos. Tests of Fundamental Symmetries.* Proc. XI Moriond Workshop, Les Arcs, France, 1991, ed. by O. Fackler, G. Fontaine and J. Tran Thanh Van (Editions Frontières, Gif-sur-Ivette, France, 1991), p. 53.
16. E.Kh. Akhmedov, Sov. Phys. JETP 68 (1989) 690.
17. E.Kh. Akhmedov, Phys. Lett. B255 (1991) 84.
18. H. Minakata, H. Nunokawa, Phys. Rev. Lett. 63 (1989) 121.
19. A.B. Balantekin, P.J. Hatchell, F. Loreti, Phys. Rev. D41 (1990) 3583.
20. E.Kh. Akhmedov, P.I. Krastev, M. Moretti, A.Yu. Smirnov, S.T. Petcov, in preparation.
21. J. Vidal, J. Wudka, Phys. Lett. B249 (1990) 473.
22. C. Aneziris, J. Schechter, Int. J. Mod. Phys. A6 (1991) 2375.
23. C. Aneziris, J. Schechter, Phys. Rev. D45 (1992) 1053.
24. M. Moretti, to be published.
25. E.Kh. Akhmedov, Phys. Lett. B257 (1991) 163.
26. R.S. Raghavan, X.-G. He, S. Pakvasa, Phys. Rev. D44 (1991) 3786.
27. R. Barbieri, G. Fiorentini, G. Mezzorani, M. Moretti, Phys. Lett. B259 (1991) 119.
28. M. Kobayashi, C.S. Lim, M.M. Nojiri, Phys. Rev. Lett. 67 (1991) 1685.
29. H. Minakata, H. Nunokawa, preprint TMUP-HEL-9111 (1991).
30. Z.G. Berezhiani, G. Fiorentini, M. Moretti, S.T. Petcov, Phys. Lett. B264 (1991) 381.
31. A.B. Balantekin, F. Loreti, Phys. Rev. D45 (1992) 1059.
32. E.Kh. Akhmedov, Nucl. Phys. A527 (1991) 679c. For more detailed discussion, see E.Kh. Akhmedov, preprint IAE-5017/1, 1990.
33. A. Cisneros, Astrophis. Space Sci. 10 (1970) 87.
34. K.S. Babu, R.N. Mohapatra, I.Z. Rothstein, Phys. Rev. D44 (1991) 2265.
35. Y. Ono, D. Suematsu, Phys. Lett. B271 (1991) 165.
36. M. Fukugita, T. Yanagida, Phys. Rev. Lett. 65 (1990) 1975.

SOLAR NEUTRINOS IN TWISTING MAGNETIC FIELDS †

P.I. Krastev*
Theory Division
CERN
CH-1211 Geneva 23
Switzerland

Abstract

The main characteristics of the resonant neutrino conversion in twisting magnetic fields are summarized. Possible consequencies for solar neutrino experiments are discussed.

---

† This talk is partially based on work done in collaboration with E.Kh. Akhmedov and A.Yu. Smirnov.

* Permanent address: Institute of Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, BG-1784 Sofia, Bulgaria.

1. Introduction

The spin precession [1] or spin-flavour conversion [2] of neutrinos with large magnetic moments may be at the origin of the apparent anticorrelation [3] of the signal in the Cl–Ar solar neutrino detector [4] with the solar activity cycle. It seems possible to explain the non–observation of such anticorrelation by the Kamiokande II collaboration in their water Cherenkov detector [5] by taking into account that: a) these detectors measure different parts of the solar neutrino energy spectrum, b) in the case of spin-flavour conversion of ν_{eL} in, e.g. $\bar{\nu}_{\mu R}$, there is an additional flavour independent neutral current contribution to the signal in Kamiokande II, which diminishes the amplitude of the expected signal variations in this detector and c) for $\mu_\nu \geq 10^{-10}\ \mu_B$ the magnetic moment interaction of the neutrinos with the electrons will supply an additional contribution to the signal in the Kamiokande II detector as compared with the Cl–Ar experiment, in which magnetic moment interactions cannot convert ^{37}Cl atoms into ^{37}Ar ones [6]. However, even if the neutrino has a large magnetic moment, $\mu_\nu \geq 10^{-13}\ \mu_B$, and the magnetic field at the base of the convective zone is of the order of 10^6 G [7] the spatial distribution of the magnetic field has to be tuned also [8] in order to obtain the apparently indicated variation pattern of the signal in the ^{37}Cl detector. Due to the different polarities of the solar magnetic field in the northern and in the southern hemispheres of the sun, there exists a slit in the magnetic field distribution in the solar equatorial plane. As the Earth crosses this plane twice per year, semiannual variations of the signals in solar neutrino detectors should occur [1], which on the average will tend to give a mean signal for an 11–year solar activity cycle close to one half of the one predicted in the standard solar model [9]. At the same time the ratio between the mean signal for about 20 years of observation with the Homestake detector and the one predicted in the standard solar model is 0.27 ± 0.03.

2. Propagation of neutrinos in twisting magnetic fields

An effect which might help to resolve the above problem is the spin precession or resonant spin-flavour conversion in a magnetic field rotating in a plane perpendicular to the neutrino momentum along the trajectory of the neutrino. In ref. [10] the propagation of a neutrino with magnetic moment in a rotating magnetic field was considered and it was concluded that the rotation of the magnetic field may relax the constraint on the magnetic moment of the neutrino necessary for the transitions $\nu_L \leftrightarrow \nu_R$ to take place in the sun. Strong conversion even for magnetic moments as small as $\mu_\nu \approx 10^{-13}\ \mu_B$ and magnetic fields of the order of 1 kG were predicted. The neutrino spin-precession in twisting magnetic fields was considered first in ref. [11]. However, the authors neglected the matter effect as well as the possible splitting between the masses of the right–handed and left–handed neutrinos. These were taken into account in ref. [12] where it was also shown that a resonant conversion takes place in this case in analogy with the MSW effect [13] in neutrino oscillations in matter. The conclusion in ref. [10] about the possibility of strong effects for solar neutrinos with small magnetic moments and weak magnetic fields as a result of the topological phase contribution to the neutrino wave function was shown to be erroneous.

The evolution equation for neutrinos propagating through matter in the presence of a magnetic field can be written in the form:

$$i\frac{d}{dx}\begin{pmatrix}\nu_L \\ \nu_R\end{pmatrix} = \begin{pmatrix} a & b \\ b^* & -a\end{pmatrix}\begin{pmatrix}\nu_L \\ \nu_R\end{pmatrix}. \tag{1}$$

Here x is the coordinate along the neutrino path in units of one solar radius, $x = \mathrm{r}/\mathrm{R}_\odot$, ν_L and ν_R are the left– and right–handed components of a Dirac neutrino, or a left–handed neutrino, e.g. ν_{eL}, and a right–handed antineutrino either of the same, e.g. $\bar{\nu}_{eR}$, or of a different flavour, e.g. $\bar{\nu}_{\mu R}$. The coefficients a and b in (1) are correspondingly

$$a = \frac{1}{2}(\sqrt{2}G_F N_{eff} - \Delta m^2/2E)$$

$$b = \mu_\nu B(x)e^{i\phi(x)}, \tag{2}$$

where G_F is the Fermi constant, N_{eff} is the effective density defined as:

$$N_{eff} = \begin{cases} N_e - N_n & \text{for } \nu_{eL} \leftrightarrow \bar{\nu}_{\alpha R} \text{ transitions;} \\ N_e - N_n/2 & \text{for } \nu_{eL} \leftrightarrow \nu_{\alpha R} \text{ transitions,} \end{cases} \tag{3}$$

where N_e and N_n are the electron and neutron number densities respectively, μ_ν is the neutrino magnetic moment, B is the transverse magnetic field, $\phi(x)$ is the rotation phase of B, depending on the distance travelled, E is the energy of the neutrino, $\Delta m^2 = \mathrm{m}_2^2 - \mathrm{m}_1^2$, and m_1 and m_2 are the masses of the massive neutrino components*. For the case of constant matter density and magnetic field and for a uniformly rotating magnetic field, $\dot{\phi} = \text{const}$, eq. (1) can be solved analytically [12]:

$$P(\nu_L \rightarrow \nu_R)(x) = \frac{(\mu_\nu B)^2}{(a + \frac{\dot{\phi}}{2})^2 + (\mu_\nu B)^2} \sin^2\left[x.[(a + \frac{\dot{\phi}}{2})^2 + (\mu_\nu B)^2]^{\frac{1}{2}}\right] \tag{4}$$

As shown in ref. [12] from the above expression it follows that taking into account the topological phase ϕ cannot relax the lower limit on $\mu_\nu B$ imposed by the condition that a sufficiently strong transition takes place in the convective zone of the sun. In the case of non–constant matter density a resonance appears [12]. If the resonance condition

$$\sqrt{2}G_F N_{eff} - \Delta m^2/2E + \dot{\phi} = 0 \tag{5}$$

is fulfilled, the mixing angle

$$\theta = -\frac{1}{2}\arctan\left(\frac{\mu_\nu B}{a + \frac{\dot{\phi}}{2}}\right) \tag{6}$$

becomes maximal. The adiabaticity condition

$$\frac{(2\mu_\nu B)^2}{|\dot{a} + \frac{\ddot{\phi}}{2}|} \gg 1. \tag{7}$$

* See ref. [14] for further details.

has to be fulfilled also in order for the $\nu_L \to \nu_R$ transitions to be most effective. It means that the density should change slowly, the field should rotate smoothly and the magnetic field should be strong enough at the resonant point. Different realizations of the resonant condition have been considered in ref. [14] *. It was shown, in particular, that a resonance can occur in the absence of mass–splitting, i.e. when $\Delta m^2 = 0$, in vacuum ($N_{eff} = 0$) and even when both these conditions are simultaneously fulfilled.

3. Implications for solar neutrinos

The twist of the magnetic fields in the sun can lead to drastic changes of the signals in the solar neutrino detectors. Unfortunately, little is known about the structure and even about the magnitude of the magnetic field inside the sun. Near the base of the convection zone fields as strong as 10^6 G have been inferred [7] from helioseismological data. However the spatial distribution is known only approximately. The possible magnitude and direction of an eventual twist of these magnetic fields seems not to have been studied at all. Perhaps with the advances in helioseismology [16] it will become possible to set usefull limits on these parameters. Some semi–realistic configurations of twisting magnetic fields in the sun and the resulting neutrino conversion have been studied in ref. [14]. Here we study numerically two more examples. In the first one, possibly realized in the sun, the neutrino propagates through a region with a non–twisting magnetic field except for a small region where the direction of the field changes abruptly. The method outlined in ref. [14] for solving eq. (1) by transforming it to a coordinate system rotating with the field is useless here. A solution can be found by matching three solutions of eq.(1) for the amplitudes in the case of non–rotating ($\dot{\phi} \equiv 0$) and uniformly ($\dot{\phi} = \text{const} \neq 0$) rotating magnetic field and taking the limit where the width of the region with $\dot{\phi} \neq 0$ goes to zero. Here only the results of numerical computations are given. An advantage of the numerical approach is that one can include a variation not only of the phase ϕ but also of the magnitude of the magnetic field itself. The calculations were done assuming a neutrino magnetic moment $\mu_\nu = 10^{-11}\ \mu_B$. The magnetic field distribution used in the calculations is shown in fig.1. It takes a maximum value of 35 kG at a distance from the centrum $\approx 0.84\ R_\odot$ The phase of the magnetic field changes at the point $0.85\ R_\odot$ by a finite value. Such large values of the phase difference as used in the numerical computations are certainly unrealistic and are used here only as an illustration**. In fig.2 the corresponding survival probabilities are given. They depend crucially on the direction of rotation of the magnetic field. For $\Delta\phi > 0$, where $\Delta\phi$ is the total change of the phase ϕ, the suppression pits become narrower, the minima shift to smaller values of $E/\Delta m^2$ and the asymptotic values of $P(\nu_L \to \nu_L)$ for $E/\Delta m^2 \to \infty$ rise and become close to unity. For $\Delta\phi < 0$ the minima of $P(\nu_L \to \nu_L)$ are shifted analogously to the case of an uniformly rotating magnetic field and the asymptotic values of this function decrease for not too strong changes of $\Delta\phi$.

---

* Analytic solution of eq.(1) has been found in ref. [15] for the case of a constant, non–uniformly rotating magnetic field and an exponential density distribution (eq.(33) in ref. [14]).

** A more realistic case may be the one in which the direction of the magnetic field changes in small steps randomly along the neutrino path.

As a second example, consider a magnetic field rotating in a finite region:

$$\phi(x) = \begin{cases} 0 & x \leq x_L; \\ \Omega.(x - x_L) & x_L \leq x \leq x_R; \\ \Omega.\Delta x & x_R \leq x; \end{cases} \tag{8}$$

where Ω is the angular velocity of the magnetic field and $\Delta x = x_R - x_L$ is the length of the interval $[x_L, x_R]$. The magnetic field profile for the non–rotating field is the one given in fig.1 with dashed line. The length of the region is correlated with the strength of the magnetic field in such a way that the total change of the dynamical phase

$$\Phi_{dyn} = \int_0^1 \mu_\nu B(x) dx, \tag{9}$$

remains constant. Let also the total change of the topological phase,

$$\Phi_{top} = \int_0^1 \dot{\phi}(x) dx, \tag{9}$$

remain constant. Thus the rate of change of the phase will vary in correlation with the length of the region in which the magnetic field rotates: for $\Delta x \rightarrow \Delta x' = \lambda \Delta x$, $\Omega \rightarrow \Omega' = \Omega/\lambda$ and $B \rightarrow B' = B/\lambda$.

As the adiabaticity condition depends quadratically on B, one expects that for narrower regions of rotation of B and, correspondingly, for stronger magnetic fields, the minima of the suppression pits will be lower. The numerical results confirm this conclusion as can be seen from fig.3. This is true only for not too strong magnetic fields which cannot rotate the spin of the neutrino by more than $\approx \pi$. Stronger magnetic fields can lead to both lower and higher suppression factors.

From the above examples and from the more extensive analysis of the spin-flavour conversion in twisting magnetic fields pursued in ref. [14] it follows that the predictions for the signals in solar neutrino experiments can be strongly affected by taking into account this effect. Assuming that either the spin–precession or the resonant spin-flavour conversion of neutrinos in matter are at the origin of the solar neutrino problem, the predictions for the signals in the solar neutrino detectors will depend on the presently unknown twist of the magnetic field. If the solar magnetic field exhibits significant twists in large enough regions through which the solar neutrinos pass on their way to the Earth then the suppression factors will change in comparison with the ones for the non–twisting fields. For the resonant spin–flavour conversion case the spectrum of the solar neutrinos will be distorted. As it is distorted already in the case of a non–twisting magnetic field, it will hardly be possible to distinguish between both cases by measuring the solar neutrino spectrum only. By changing the magnitude and/or the spatial distribution of the magnetic field inside the sun, it should in principle be possible to obtain similar suppression factors. However, in the case of spin-precession the transition becomes resonant. The solar neutrino spectra will be undistorted. With the additional parameter, the twist of the magnetic field, at hand, one can account for semiannual variations of the detection rates which will not smear out the 11–year ones. It becomes possible also to find a natural explanation of the apparent

discrepancy between indicated strong variations of the signal in the Cl–Ar experiment and vanishingly small ones in the Kamiokande–II experiment. In particular [17], there is no need of a strong magnetic field in the radiation zone as assumed in ref. [6].

Depending on the direction of rotation of the field, which need not be the same in the northern and southern solar hemispheres, the semiannual variations can be quite different as compared with the corresponding ones in the case of a non–twisting field. This is shown in fig.4, where four examples of possible variation patterns are shown. Sufficiently sensitive and accurate solar neutrino detectors might be able to observe such variations. The observation of the latter will be strong evidence for the twisting structure of the solar magnetic field.

The twist of the magnetic field may change also with the solar activity cycle. This will lead to correlations between the event rates in solar neutrino detectors and parameters characteristic of the solar activity cycle.

It was first pointed out in ref. [18] that the neutrino spin precession in the solar magnetic field will lead to an azimuthal asymmetry of the recoil electrons scattered by solar neutrinos in $\nu_e e$ detectors. For toroidal magnetic fields with fixed direction in space, the spin of the neutrino precesses in a plane perpendicular to the solar equatorial plane. Therefore, the azimuthal asymmetry of the scattered electrons will be maximal with respect to the solar equatorial plane. The twist of the magnetic field will lead to a rotation of the reference plane with respect to which this asymmetry is maximal and the latter will no longer coincide with the solar equatorial plane. If the twist of the magnetic field varies quickly with time and/or has a stochastic character, the azimuthal distribution of the scattered electrons will be more isotropic. On the other hand, if the twist remains constant for longer periods, then the measurement of the azimuthal asymmetry might become possible. A careful analysis of this asymmetry might help in studying the solar magnetic field.

4. Conclusions

If the neutrino has sufficiently large magnetic moments and the magnetic field inside the sun turns out to have a twisting structure, then the rotation of the magnetic field along the neutrino trajectory will affect the predictions for the event rates in solar neutrino detectors. Rotation of the field by 25 – 30 degrees over a distance of $\approx 0.05\ R_\odot$ can change the ν_{eL} survival probabilities by a factor of ≈ 2. Measurements of the neutrino spectra and of the time variations of the signals from solar neutrinos, as well as of the azimuthal asymmetry of the recoil electrons in νe scattering experiments can give information about both the neutrino magnetic moments and the structure of the solar magnetic field.

References

1 M.B.Voloshin and M.I.Vysotsky, Sov. J. Nucl. Phys. 44 (1986) 845; L.B. Okun, M.B. Voloshin and M.I. Vysotsky, Sov. Phys. JETP 64 (1986) 446.

2 C.-S. Lim and W.J. Marciano. Phys. Rev. D37 (1988) 1368; E.Kh. Akhmedov, Phys.Lett. B213 (1988) 64.

3 G.A. Bazilevskaya, Yu.I. Stozhkov and T.N. Charakhch'yan, Pis'ma Zh.ETF, 35

(1982) 273 (Soviet Phys.–JETP Letters, 35 (1982) 341), J.N. Bahcall and H. Press, Ap. J. 370 (1991) 730; B.W. Filippone and P. Vogel, Phys. Lett. B246 (1990) 546; J.W. Bieber, D. Seckel, T. Stanev and G. Steigman, Nature 348 (1990) 407; H. Nunokawa and H. Minakata, Proc. of the 25th Int. Conf. on High Energy Physics, Singapore, 1990, p.681.

4 R. Davis and K. Lande, Proc. of the 25th Int. Conf. on High High Energy Physics, Singapore, 1990, p. 667.

5 K.S. Hirata et al., Phys, Rev. D44 (1991) 2241.

6 E. Akhmedov, Nucl. Phys. A527 (1991) 679c; preprint IAE–5017/1, 1990 (unpublished); Phys. Lett. B257 (1991) 163; K.S. Babu, R.N. Mohapatra and I.Z. Rothstein, Phys. Rev. D44 (1991) 2265; Y. Ono and D. Suematsu, Phys. Lett. B271 (1991) 165.

7 W.A. Dziembowski and P.R. Goode, Ap.J. 347 (1989) 540.

8 P.I. Krastev and A.Yu. Smirnov, Z. Phys. C – Particles and Fields, 49 (1991) 675; R. Barbieri and G. Fiorentini, in Second International Workshop on Neutrino Telescopes, Venezia, 1990, ed. by Milla Baldo Ceolin, p.47.

9 J.N. Bahcall and R.K. Ulrich, Rev. Mod. Phys., 60 (1988) 297.

10 J. Vidal and J. Wudka, Phys. Lett. B249 (1990) 473.

11 C. Aneziris and J. Schechter, J. of Mod. Phys. A, 6 (1991) 2375.

12 A.Yu. Smirnov, Pis'ma ZhETF 53 (1991) 280; Phys. Lett. B260 (1991) 161.

13 S.P. Mikheyev and A.Yu. Smirnov, Nuovo Cim. C9 (1986) 17; L. Wolfenstein, Phys.Rev. D17 (1978) 2369.

14 E.Kh. Akhmedov, P.I. Krastev and A.Yu. Smirnov, Z. Phys. – C, Particles and Fields, 52 (1991) 701.

15 S.T. Petcov, to be published.

16 D. Gough, talk in these Proceedings.

17 P. Krastev, in preparation.

18 R. Barbieri and G. Fiorentini, Nucl. Phys. B304 (1988) 909.

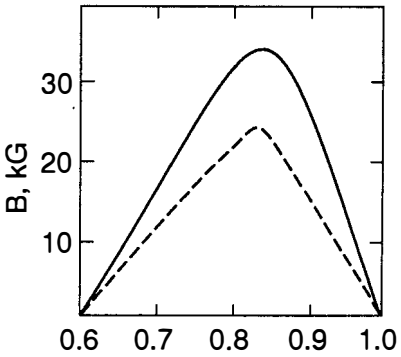


Figure 1: Standard magnetic field profile, used in the calculations (solid line) and the same profile homogeneously reduced by a factor of 0.7 (dashed line).

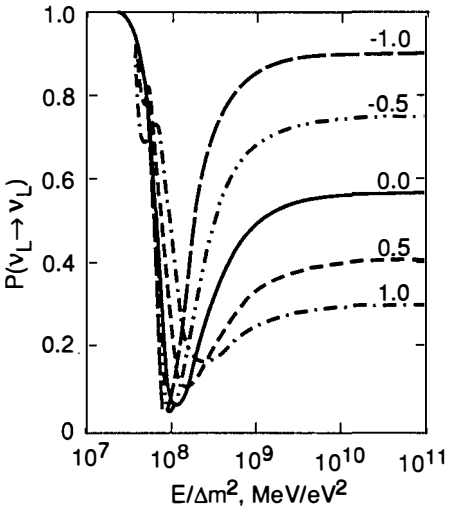


Figure 2: The probabilities P ($\nu_L \rightarrow \nu_L$) as functions of $E/\Delta m^2$ in the case of a step-like change of the phase by a finite value $\Delta\varphi$ (numbers at the curves) at x = 0.84.

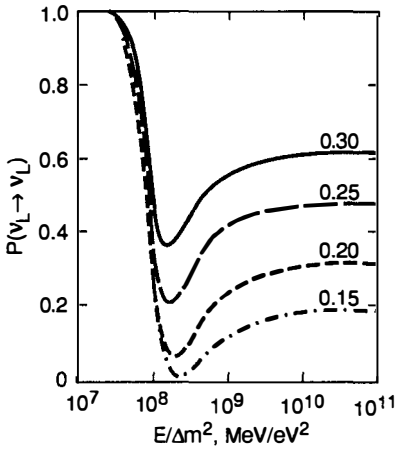


Figure 3: The probabilities P ($\nu_L \rightarrow \nu_L$) as functions of $E/\Delta m^2$ for a fixed total change of the phase $\Delta\varphi = 1$ rad and for different rotation velocities: $\Omega = 3.33$ (solid curve), 4.0 (long dashed curve), 5.0 (short dashed curve) and 6.66 (dot-dashed curve). All intervals are symmetric around $x_c = 0.85$. Numbers at the curves are the corresponding widths of the intervals.

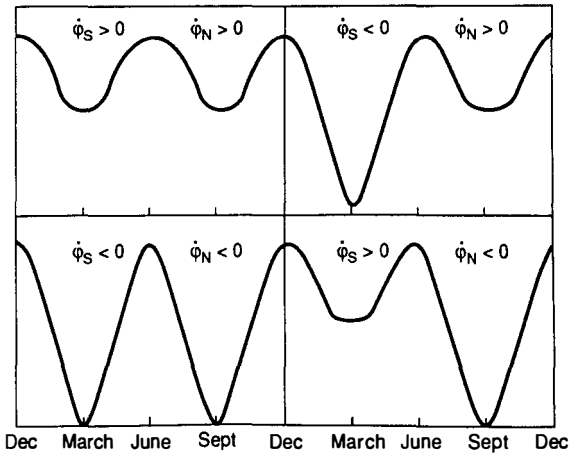


Figure 4: Annual variations of the event rates in solar neutrino experiments for different helicities of the twists in the northern ($\dot{\varphi}_N$) and the southern ($\dot{\varphi}_S$) hemispheres.

ANALYTICAL FORMULAE FOR RESONANT NEUTRINO SPIN ROTATION IN TWISTING MAGNETIC FIELD

Stoyan Toshev
Laboratoire de Physique Théorique,
Université de Bordeaux I, Bordeaux, France*

Abstract:

Analytical formulae for the probability of *resonant neutrino spin rotation* in twisting magnetic field with varying twist frequency are presented and briefly discussed.

*Permanent adress: Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria

In 1984 M.Berry has published an important result [1] showing that during the adiabatic evolution of a given quantum system the energy eigenfunctions can obtain, in addition to the usual dynamical phase, a new "geometrical" phase (now known as "Berry's phase"). This happens when the Hamiltonian depends on time through some multiple parameters and the system under consideration is transported slowly enough around a closed path in the parameter space. Later on various generalizations of this result have been considered [2].

Recently, Wudka and Vidal [3] have noticed that these conditions can be realized during the neutrino propagation in the twisting magnetic field of the sun [4]. They have argued that possible large adiabatic phases can lead to a sensible reduction in the ν_e flux even for neutrino magnetic moments as small as $10^{-13}\mu_B$. Aneziris and Schechter [5] have shown, however, in the context of a simple solvable model, that the probability of a spin rotation is large when the adiabatic (Berry's) phase is small compared to the dynamical one, otherwise this probability is small. The controversy has been solved by Smirnov [6] which have shown that the wrong conclusion in Ref. [3] results from the fact that the adiabaticity condition is not fulfilled under the circumstances considered.

It turns out that the twisting magnetic field may lead, nevertheless, to a interesting *new* phenomenon: *resonant amplification of the neutrino "spin–rotation"* [6]. In the same paper the main characteristics of this process in the adiabatic case have been briefly discussed.

Subsequently the phenomenon has been investigated using exact analytical [7] and numerical [8] methods, thus providing results valid both in the adiabatic and in the non-adiabatic regime. It has been shown that the effect of the resonant amplification of the spin rotation in twisting magnetic field is very similar to the Mikheyev-Smirnov-Wolfenstein effect [9]. Nevertheless, some important differences exist.

Mathematically, both these problems are equivalent, with the derivative of the phase of the twisting magnetic field $d(\alpha t)/dt$ in the first case corresponding to the electron number density N_e in the *MSW* case. The detailed investigations of the *MSW* effect have shown that exact analytical solutions exist only in several very specific cases when N_e is linear [10,11], exponential [12,13], $1/x$ [14], *etc.* function of the distance x to the center of the sun. According to the standard solar model [15] the density distribution in the solar interior is exponential apart from small regions in the center and close to the surface. Therefore the exponential approximation [10,11,16] has provided the best analytical results.

Unfortunately, the corresponding dependance of the derivative of the phase of the twisting magnetic field $d(\alpha t)/dt$ is not known. In such a situation the argument that one should use again the exponential approximation because, at least, N_e depends exponentially on x is a dangerous one because this means that one assumes *a priori* that the matter (*MSW*) effects dominate. We consider this to be a rather trivial case. Since our main goal is to investigate in detail the characteristics of the *new* phenomenon (*the resonant amplification of the neutrino "spin–rotation" in twisting magnetic field*) and not the characteristics of the *MSW* effect (which

are well known already) it is obvious that the best choice is the linear one because, due to the resonant character of the fenomenon and under the most general assumption that $d(\alpha t)/dt$ is arbitrary *smooth* function we can approximate it in the vicinity of the resonant point with a linear function. Essentially this means that we take from the Taylor series expansion of this unknown function only the first two terms because, sufficiently close to the resonant point, the higher order terms provide only small corrections to the final results[1].

The exact analytical solution in the case of linearly varying twist frequency α have been obtained in Ref.[7]. Here we will present the results of a more general investigation where all the parameters of the model, both the twist frequency and the electron number density, are assumed to be arbitrary smooth functions. In accordance with the arguments presented earlier sufficiently close to the resonance they are approximated with linear functions thus providing analytical solutions which are useful for a large class of magnetic field configurations.

The evolution equations for the amplitudes $A_{L,R}$ of the probabilities for transitions between the left-handed and the right-handed neutrino states in a transverse magnetic field $\vec{B}(t) = (B_\perp \cos \alpha t, B_\perp \sin \alpha t, 0)$, which winds around the direction z of the neutrino momentum $\vec{p}$ with twist frequency α, are [3,5,6]:

$$i\frac{dA_R(t)}{dt} = B_0 A_R(t) + \mu B_\perp (\cos \alpha t - i \sin \alpha t) A_L(t) \tag{1}$$

$$i\frac{dA_L(t)}{dt} = \mu B_\perp (\cos \alpha t + i \sin \alpha t) A_R(t) - B_0 A_L(t). \tag{2}$$

μ is the neutrino magnetic moment and the quantity B_0 depends on the nature of the neutrinos, *i.e.* whether they are Dirac ($B_0 = -G_F(N_e - \frac{1}{2}N_n)/\sqrt{2}$) or Majorana ($B_0 = -G_F(N_e - N_n)/\sqrt{2} + \Delta m^2/4p$) particles [2].

We proceed by replacing $A_L(t)$ in eq.(2) with the following expression obtained from eq.(1):

$$A_L(t) = \frac{\cos \alpha t + i \sin \alpha t}{\mu B_\perp} \left(i\frac{dA_R(t)}{dt} - B_0 A_R(t) \right). \tag{3}$$

The result is one second order linear differential equation for $A_R(t)$:

$$\frac{d^2 A_R(t)}{dt^2} + i\left(\frac{d\alpha}{dt}t + \alpha\right)\frac{dA_R(t)}{dt} + \left(B_0^2 - \alpha B_0 - \frac{d\alpha}{dt}tB_0 + i\frac{dB_0}{dt} + \mu^2 B_\perp^2\right) A_R(t) = 0. \tag{4}$$

The substitution:

$$A_R(t) = \exp\left(-\frac{i}{2}\alpha t\right) U_R(t) \tag{5}$$

[1] That the linear approximation is a very precise one even in the *MSW* case, especially when the neutrino mixing angles are small, has been shown by Haxton in Ref. [10]

[2] G_F is the Fermi coupling constant, N_e and N_n are the electron and neutron number densities and $\Delta m^2 = m_2^2 - m_1^2$. Since the neutrinos are relativistic we will use t instead of x.

brings it to the canonical form:

$$\frac{d^2U_R(t)}{dt^2} + \left\{ \left[B_0 - \frac{1}{2}\frac{d(\alpha t)}{dt} \right]^2 + \mu^2 B_\perp^2 + i \left[\frac{dB_0}{dt} - \frac{1}{2}\frac{d^2(\alpha t)}{dt^2} \right] \right\} U_R(t) = 0. \tag{6}$$

Obviously, when $B_0 - \frac{1}{2}\frac{d(\alpha t)}{dt}$ is linear function of t the polynomial in the brackets in the LH side of eq.(6) is of the second order with respect to t. In this case eq.(6) will be equivalent to Weber's equation [3]. The solutions of the later are the *parabolic cylinder functions* $D_n(\pm z)$ and $D_{-n-1}(\pm iz)$ [17]. In accordance with the arguments presented earlier, we take from the Taylor's series expansions (in the vicinity of the resonant point) of the two arbitrary, but smoot, functions $B_0(t)$ and $\frac{d(\alpha(t)t)}{dt}$ only the first two terms, the resonant character of the phenomenon under consideration guaranting that the higher orther terms lead only to small corrections to the final results.

Without loss of generality we can assume that:

$$\left[B_0 - \frac{1}{2}\frac{d(\alpha t)}{dt} \right]_{t=0} = 0, \tag{7}$$

which, as we will show below, corresponds to the choice the resonance to take place at $t = 0$. The transformations

$$z = \exp\left(\frac{i}{4}\pi\right) \left[2\frac{dB_0}{dt} - \frac{d^2(\alpha t)}{dt^2} \right]^{1/2} t, \qquad n = -i\frac{\mu^2 B_\perp^2}{2\frac{dB_0}{dt} - \frac{d^2(\alpha t)}{dt^2}} \tag{8}$$

give us Weber's equation:

$$\frac{d^2U_R(z)}{dt^2} + \left(n + \frac{1}{2} - \frac{z^2}{4} \right) U_R(z) = 0. \tag{9}$$

The exact solution which satisfies the initial conditions [4]

$$A_L(t_i) = 1, \qquad A_R(t_i) = 0, \qquad t_i < 0 \tag{10}$$

is superposition of parabolic cylinder functions:

$$\begin{aligned} A_R(t_i) \;=\; & \frac{\mu B_\perp}{\left[2\frac{dB_0}{dt} - \frac{d^2(\alpha t)}{dt^2} \right]^{1/2}} \exp\left[-\frac{i}{2}\pi\left(n - \frac{3}{2} \right) \right] \exp\left[-\frac{i}{2}\alpha(t_i + t_f) \right] \\ & \times \left[D_n(-z_i) D_{-n-1}(iz_f) - D_n(-z_f) D_{-n-1}(iz_i) \right]. \end{aligned} \tag{11}$$

[3] Similar problem has been investigated in Refs. [10,11] in the context of the *MSW* effect.

[4] These conditions correspond to only left-handed neutrinos in the initial moment $t = t_i$ which is the case relevant for the solar neutrinos. Due to the well-known simple relations between the amplitudes $A_{R/L}(t_f)$ in this case and the amplitudes in the case $A_L(t_i) = 0, A_R(t_i) = 1$, namely $A_{R/L}(t_f) \to -/+ A^*_{L/R}(t_f)$, the "general" solution corresponding to mixed initial state is obvious.

Assuming that the initial and final points are sufficiently far from the resonant point one can obtain very simple *approximate* solution for the averaged transition probability

$$\overline{P_{\nu_L \to \nu_R}(t_i, t_f)} = \overline{|A_R(z(t_f))|^2} \tag{12}$$

by substituting in eq.(11) the asymptotic expansions of $D_n(-z_{i(f)})$ and $D_{-n-1}(iz_{i(f)})$ for large values of $|z_{i(f)}|$. This procedure has been suggested and outlined in detail in Ref.[11] where the interested reader can find all the relevant formulae. The final result is:

$$\overline{P_{\nu_L \to \nu_{R/L}}(t_i, t_f)} = \frac{1}{2} - / + (\frac{1}{2} - P_\mu^{lin}) \cos 2\theta_\mu^i \cos 2\theta_\mu^f \tag{13}$$

where

$$P_\mu^{lin} = \exp\left(-2\pi \frac{\mu^2 B_\perp^2}{\left|2\frac{dB_0}{dt} - \frac{d^2(\alpha t)}{dt^2}\right|_R}\right) \tag{14}$$

is the relevant Landau-Zener type probability [18] in the case under consideration and $\theta_\mu^{i,f}$ are the corresponding "mixing angles" defined by:

$$\tan^2 2\theta_\mu^{i(f)} = \tan^2 2\theta_\mu(t_{i(f)}) = \frac{\mu^2 B_\perp^2}{\left[B_0 - \frac{1}{2}\frac{d(\alpha t)}{dt}\right]^2_{t=t_{i(f)}}}. \tag{15}$$

The subscript R in eq.(14) means that all the derivatives have been taken at the resonant point which, as can be veryfied easely from eq.(15), is defined by eq.(7).

The averaging procedure used to derive eqs.(13,1415) is different from the corresponding procedure in the *MSW* case [11] where we average over the neutrino production regions in the center of the sun and over the distance to the detector. Here we have time averaging over the different magnetic field configurations. This is possible because z depends explicitly on $\frac{d^2(\alpha t)}{dt^2}$.

In particular, in its adiabatic limit defined by $P_\mu^{lin} \longrightarrow 0$ eq.(13) reproduces Smirnov's results [6].

In summary, analytical formulae for the probabilites of $\nu_L \leftrightarrow \nu_R$ transitions in the twisting solar magnetic field have been derived which are applicable for a large class of magnetic field configurations.

The author is gratefull to the Organizing Committee of the Moriond Workshop for the kind hospitality at Les Arcs.

References

[1] M.Berry, Proc.R.Soc.London A 392 (1984) 45.

[2] See e.g. A.Shapere and F.Wilczek, eds., *Geometrical Phases in Physics* (World Scientific, Singapore, 1989).

[3] J.Wudka, J.Vidal, University of California Report UCD-88-40 (1988); Phys.Lett. B 249 (1990) 473.

[4] See, for example: Y.Nakagawa, in *Solar Magnetic Fields*, ed. R.Howard (Springer-Verlag, NY, 1971).

[5] C.Aneziris, J.Schechter, Int.J.Mod.Phys. A 6 (1991) 2375.

[6] A.Yu.Smirnov, Pis'ma JETP 53 (1991) 280; Phys.Lett. B 260 (1991) 161.

[7] S.Toshev, Phys.Lett. B 271 (1991) 179.

[8] E.Kh.Akhmedov, P.I.Krastev, A.Yu.Smirnov, Z.Phys. C 52 (1991) 701.

[9] S.P.Mikheyev, A.Yu.Smirnov, Yad.Fiz. 42 (1985) 1441;
L.Wolfenstein, Phys.Rev. D 17 (1978) 2369; *ibid.* 20 (1979) 2634.
For a comprehensive review see: S.P.Mikheyev, A.Yu.Smirnov, Prog.Part.Nucl.Phys. 23 (1989) 41.

[10] W.C.Haxton, Phys.Rev. D 35 (1987) 2352.

[11] S.Toshev,Phys.Lett. B 198 (1987) 551.

[12] S.Toshev, Phys.Lett. B 196 (1987) 170.

[13] T.Kaneko, Prog.Theor.Phys. 78 (1987) 532.

[14] T.K.Kuo, J.Pantaleone, Phys.Rev. D 39 (1989) 1930.

[15] J.N.Bahcall, *Neutrino Astrophysics* (Cambridge Univ. Press, 1989);
S.Turck-Chieze, S.Cahen, M.Casse, C.Doom, Ap.J. 335 (1988) 415.

[16] See also: S.Toshev, Mod.Phys.Lett. A 3 (1988) 79;
P.Pizzochero, Phys.Rev. D 36 (1987) 2293.

[17] See, *e.g.*, E.T.Whittaker, G.N.Watson, *A Course of Modern Analysis*, (Cambridge Univ. Press, 1927), p.347.

[18] L.D.Landau, Phys.Z.U.S.S.R. 1 (1932) 88;
C.Zener, Proc.R.Soc. A 137 (1932) 696.

The Sudbury Neutrino Observatory

J.R. Leslie on behalf of the SNO Collaboration †
Physics Department
Queen's University
Kingston
Ontario K7L 3N6
Canada

Abstract

The Sudbury Neutrino Observatory (SNO) detector is a 1000 tonne heavy water Čerenkov detector under construction near to Sudbury in Canada. The aim of the experiment is to detect neutrinos from the sun and other astrophysical sources. The use of heavy water allows both electron neutrinos and all types of neutrinos to be observed by three complementary reactions. The detector will be sensitive to the electron neutrino flux and energy spectrum shape and to the total neutrino flux irrespective of neutrino type. These measurements will provide information which should clarify the role of neutrino oscillations, neutrino magnetic moments and solar model predictions in explanations of the observed deficit in solar neutrino flux.

1 Introduction

The Sudbury Neutrino Observatory (SNO) is based on a 1000 tonne heavy water (D_2O) detector located in a very low background laboratory 2070 m underground. The 1000 tonnes of pure D_2O used in the experiment will be lent by AECL and a suitable site in the Creighton Mine near Sudbury, Ontario, has been provided by INCO Ltd. The objective of the SNO project is to measure the intensity, energy and direction of neutrinos from the sun and supernovae. A detailed description of the proposal[1] is given in an earlier report. Improvements (some of which are described in this report) to that reference design have since been made but the same basic principles are used. The detector and laboratory are now under construction.

The proposed[1] SNO detector makes use of three complementary neutrino reactions.

The charged current (CC) reaction

$$\text{(CC)} \qquad \nu_e + d \rightarrow p + p + e^- \qquad (Q = -1.44 \text{ MeV})$$

of the electron neutrino on the deuteron which is unique to the SNO detector. It offers excellent spectral information, thereby providing a sensitivity to the MSW effect. The large cross section will result in more than ten events per day and make the detector some fifty times more sensitive than existing experiments. The reaction also offers some directional information and can identify the sun as a source of electron neutrinos.

The neutrino-electron elastic scattering (ES) reaction,

$$\text{(ES)} \qquad \nu_x + e^- \rightarrow \nu_x + e^-$$

the primary detection mechanism for light water detectors, is sensitive to all neutrino types, but is dominated by the electron neutrino. This reaction offers excellent directional information, but provides poor information about the energy spectrum of electron neutrinos.

The third neutrino induced reaction occurring in the SNO detector is the neutral current (NC) disintegration of the deuteron,

$$\text{(NC)} \qquad \nu + d \rightarrow \nu + p + n \qquad (Q = -2.23 \text{ MeV})$$

observed by the detection of the gamma rays resulting from the subsequent neutron capture. This reaction is sensitive to all neutrinos equally and would be used to measure the total flux of neutrinos with a counting rate of about 10 per day above the threshold of 2.23 MeV.

In addition to the above reactions there are the anti-neutrino induced inverse β decays of both heavy and light hydrogen which will provide information on the possible anti-neutrino flux from the Sun. These reactions are:

$$\bar{\nu}_e + d \rightarrow n + n + e^+ \qquad (Q = -4.03 \text{ MeV})$$

observed by the detection of a relativistic β^+ followed by the detection of two neutrons as in the NC reaction and

$$\overline{\nu}_e + p \rightarrow n + e^+ \qquad (Q = -1.80 \text{ MeV})$$

produced in the region of the light water shield between the D_2O region and the PMTs. The neutron will not yield a detectable signal as the 2.23 MeV γ ray produced by neutron capture on light water will not be above the background level.

In this paper we discuss briefly the SNO detector and its role in the detection of neutrinos and anti-neutrinos from the sun.

2 The SNO Detector

The present design of the SNO detector is shown in Figure 1. It consists of 1000 tonnes of 99.85% enriched D_2O contained in a spherical thin-walled (5 cm thick) transparent acrylic vessel which itself is immersed in 7300 tonnes of ultrapure H_2O shield. The host cavity, which is barrel-shaped and of 22 m diameter and 30 m height, is under excavation at a depth of 2070 m below surface. A waterproof plastic liner is placed inside the cavity enclosing the light water shield. Some 9600 20-cm diameter photomultiplier tubes (PMTs) are uniformly arranged in the H_2O shield on a support structure at 2.5 m distance from the acrylic vessel. Each PMT is surrounded by a reflector to increase the light collection and the effective photocathode coverage is approximately 60%. The PMT array is sensitive to Čerenkov radiation produced by relativistic electrons and other relativistic particles in the central regions of the detector.

Neutrino interactions in the detector produce either relativistic electrons or free neutrons. The neutrons thermalize in the water and are subsequently captured generating γ-rays which produce relativistic electrons primarily through Compton scattering. The electrons from neutrino interactions or neutron captures will produce Čerenkov photons which pass through the D_2O, acrylic, and H_2O to be detected by the PMTs. The signals from the PMTs are interpreted to give the location, energy and direction of the electron based on the time of arrival of the photons at the PMTs and the locations of the particular PMTs hit. At E_e =10.0 MeV detector simulations predict a spatial reconstruction accuracy of 26 cm, an angular uncertainty 27° and an energy sensitivity of 9.0 PMT hits per MeV with a 12% statistical uncertainty.

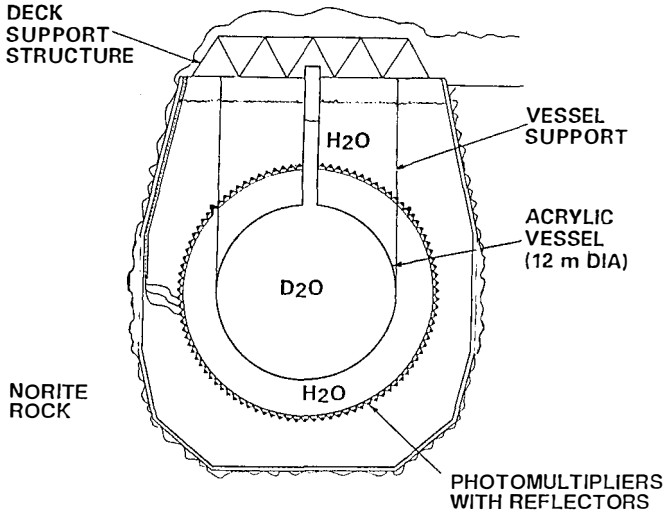


Figure 1 Line drawing of SNO detector

The expected solar neutrino signal rates of about ten events per day require an extremely low background environment for the D_2O. There are two sources of background that must be guarded against. These sources are cosmic rays and radiation produced by naturally occurring radionuclides in the rock surrounding the detector and in the materials used in building the detector. The location of the detector is so deep that the background rate from cosmic-ray muons is negligible. The purity of materials used in the detector has to be very high. The most critical components are the D_2O, H_2O, acrylic vessel and PMTs. In the D_2O and H_2O levels of U and Th in the range of 10^{-14} gram per gram of water are required. In the acrylic vessel which has less mass the tolerable level is a few times 10^{-12} gram per gram of acrylic. The PMTs are 2.5 m from the heavy water and γ-rays originating in these are attenuated in the light water shield. To keep contributions to background to a low level the PMTs are being built using glass envelopes from a special melt of low radioactivity glass by Schott Glaswerke. The U and Th content of this glass is $\sim$5–10 times lower than normal glass.

The expected performance of the SNO detector has been calculated using an extensive series of Monte Carlo simulations of the neutrino signal processes and radioactive background processes. The effective electron threshold energy for the detector, due to background β-γ cascades from U and Th contaminants, is predicted to be approximately 5 MeV. The events for the NC reaction are determined by detecting the γ-rays following the capture of the neutron released. About 2.5 tonnes of $NaC\ell$ will be added to the heavy water. Then 83% of the neutrons will be captured by $^{35}C\ell$ producing γ-rays with energies up to 8 MeV which will be observable above the background. An additional background must be considered for this reaction, corresponding to neutrons produced by

the photodisintegration of deuterium by γ-rays with energies greater than 2.23 MeV. This background will arise predominantly from radioactive decay of members of the ^{232}Th and ^{238}U decay chains contained in the D_2O. This radioactive contamination will be sampled on-line to determine the background rate accurately.

The signal for the NC reaction is the production of a free neutron and alternative schemes for the detection of this neutron, such as ^{3}He or ^{6}Li counters, are being considered. One objective is to allow real-time discrimination between events produced by neutrons and other events. Extreme purity of materials is needed for such counters as they would be located in the D_2O itself.

3 Solar Neutrinos

For some detectors the deduction of neutrino flavour oscillation parameters is dependent on the assumption that the total neutrino emission from the sun is given by the standard solar model - the unique ability of SNO to measure the total neutrino flux will make experimental interpretation independent of solar physics and of course give valuable information on the processes in the interior of the sun. The high count rate and the real time data taking of SNO will enable a study with good statistical accuracy of possible time variation of the solar neutrino flux of all types. The multiple sensitivity of SNO to reactions which are induced by CC events to detect ν_e left only, NC events to detect all type of neutrino and the ES reaction which has sensitivity to possible neutrino EM interactions will provide an invaluable tool for the study of possible neutrino magnetic moment effects. Neutrino EM effects and combined EM/matter induced oscillations have been studied extensively by Akhmedov [2], Voloshin [3] and co-workers. The prediction of these calculations include the precession of ν_e left to sterile ν_e right in the case of Dirac particles or to active $\overline{\nu}$ right in the case of Majorana particles. The variety of sensitivities of the SNO detector can give valuable data to disentangle these effects. Recent calculations [4] of neutrino induced interactions with deuterium will be valuable the interpretation of the data which will be produced by the SNO detector: There is increased confidence in the reliability of the calculation of the relative CC and NC reactions thus making the deduction of the ratio of ν_e and total ν fluxes more certain; the EM induced ν induced deuteron disintegration [5] has been shown to be a small effect relative to the weak interaction effect; possible $\overline{\nu}$ charged current interactions will occur with a cross-section of 43% of the corresponding ν_e interactions when averaged over the ^{8}B ν_e solar spectrum; the $\overline{\nu}$ and ν deuteron disintegration occur with approximately the same cross section.

4 Project Status and Schedule

The SNO detector received funding in January 1990. The order for photo-multiplier tubes have been placed with Hamamatsu Photonics and delivery and testing of the first PMTs has commenced. A final choice of reflector material and design has been made and

the PMT support structure method has been finalised. A section of the water treatment plant is in operation. Major items in the construction schedule are roughly as follows. The excavation of the cavity and the preparation of the underground laboratory will take until Nov, 1992, at which time the waterproof liner and deck will be installed. Installation of the acrylic vessel, PMTs and their support structure will take place in 1993. The initial water fill is scheduled to begin in the Nov, 1994, with a view to completing commissioning tests by April 1995.

The detector facilities at the time of initial installation will have full support for D_2O use so that the experimental procedure will be to run for a period with pure D_2O to measure the CC and ES signals and then to add neutral current detectors. The neutral current detection will be done either by adding $NaC\ell$ to the D_2O or if the attempt to produce discrete detectors of low enough radioactivity is successful to deploy these in the detector.

† The Sudbury Neutrino Observatory Collaboration: H.C. Evans, G.T. Ewan, A.L. Hallin, H.W. Lee, J.R.Leslie, J.D.MacArthur, H.B. Mak, A.B. McDonald, W. McLatchie, B.C. Robertson, B. Sur, P. Skensved. (Queen's University): I. Blevis, C.K. Hargrove, H. Mes, W.F. Davidson, M. Shatkay, D. Sinclair. (Centre for Research in Particle Physics): E.D. Earle, G.M. Milton, D. Hepburn, E. Bonvin. (Chalk River Laboratories): P. Jagam, J. Law, J.-X. Wang, J.J. Simpson. (University of Guelph): J. Bigu, E.D. Hallman, R.U. Haq. (Laurentian University): A.L. Carter, D. Kessler, B.R. Hollebone. (Carleton University): S. Gil, C.E. Waltham. (University of British Columbia): M.M. Lowry. (Princeton University): E.W. Beier, W. Frati, M. Newcomer, R. Van Berg. (University of Pennsylvania): T.J. Bowles, P.J. Doe, M.M. Fowler, A. Hime, F. McGirt, R.G.H. Robertson, T.C. Spenser, D.J. Vieira, J.B. Wilhelmy, J.F. Wilkerson, J.M. Wouters. (Los Alamos National Laboratory): Y.D. Chan, A.Garcia, E. Norman, K. Lesko, A. Smith, R. Stokstad, I. Zlimen. (Lawrence Berkeley Laboratory): N.W. Tanner, N.A. Jelley, J.C. Barton, G. Doucas, E.W. Hooper, A.B. Knox, M.E. Moorhead, M. Omori, P.T.Trent, D.L.Wark. (University of Oxford).

5 References

1. G.T. Ewan et al., Sudbury Neutrino Observatory Proposal, SNO-87-12 (1987).

2. E.K. Akhmedov, Phys Lett B213 (1988) 909.
   E.K. Akhmedov, Phys Lett B255 (1991) 84.

3. M.B. Voloshin et al., Sov. Phys. JETP 64 (1986) 446.

4. M. Doi and K. Kubodera, Phys. Rev. C - In press (April 1992 issue).
   S. Ying et al., Phys. Rev. C - In Press (April 1992 issue).

5. S. Nozawa, Private communication.
   E.K. Akhmedov and V.V. Berezin. Submitted to Z. Phys. C (Dec. 1991)

Status of the Borexino Experiment

Presented by Marco G. Giammarchi
Istituto Nazionale di Fisica Nucleare
Via Celoria 16, 20133 - Milano (Italy)

On behalf of the Borex Collaboration:

G. Alimonti[g], B. Alpat[j], C. Arpesella[a], G. Bellini[g], P. Benetti[i], B. Bertotti[i], S. Bonetti[g], S. Brambilla[g], F. P. Calaprice[l], M. Campanella[g], G. Cecchet[i], A. de Bari[i], M. Deutsch[b], F. von Feilitzsch[d], D. Franciotti[a], M. G. Giammarchi[g], D. Giugni[g], A. Golubchikov[c], V. Gracco[e], C. Hagner[d], P. Inzani[g], I. Iori[g], T. Kovacs[h], V. Lia[b], I. Manno[g], E. Meroni[g], J. Mitchell[h], L. Oberauer[d], G. Mantovani[j], G. Manuzio[e], S. Pakvasa[f], L. Perasso[g], A. Perotti[i], S. Prakhov[c], A. Preda[g], P. Raghavan[h], R.S. Raghavan[h], F. Ragusa[g], G. Ranucci[g], R. Scardaoni[g], S. Schönert[d], D. Steinberg[k], R. Tartaglia[a], G. Testera[e], P. Ullucci[g], S. Vitale[e], O. Zaimidoroga[c]

a) Laboratori Nazionali del Gran Sasso - Assergi (Aq) - Italy
b) Massachusetts Inst. of Technology - Cambridge MA - USA
c) Joint Inst. for Nuclear Research - Dubna - Russia
d) Technical University Munich - Garching - Germany
e) Physics Dep. of the University and INFN - Genova - Italy
f) University of Hawaii - Honolulu HI - USA
g) Physics Dep. of the University and INFN - Milano - Italy
h) AT&T Bell Laboratories - Murray Hill NJ - USA
i) Physics Dep. of the University and INFN - Pavia - Italy
j) Physics Dep. of the University and INFN - Perugia - Italy
k) Drexel University - Philadelphia PA - USA
l) Physics Dep. Princeton University - Princeton NJ - USA

Abstract

The main features of the Borexino project at Gran Sasso for the study of the Solar Neutrino Problem are discussed. The current situation is reviewed, with a description of the Counting Test Facility being set up at the LNGS laboratory.

1. Introduction

Borexino is a real time low-background scintillation detector for low energy neutrinos currently proposed for the LNGS laboratory. It is based on an ultrapure mixture of scintillation liquid viewed by phototubes and shielded by high purity water.

The intrinsic high purity of the scintillator and the adequate amount of shielding make it possible to lower to unprecedented levels the operational threshold of the detector, thereby achieving sensitivity down to below 250 KeV, of great importance for the study of ^{7}Be neutrinos from the Sun. In addition, Borexino can detect solar neutrino events by various other modes, allowing a complete study of the energy spectrum and a comprehensive investigation of the Solar Neutrino Problem.

2. The Detector

The Borexino detector consists of a few spherical shells of increasing radiopurity (fig. 1). The inner shell consists of the detecting medium itself, a 8.5 m diameter sphere filled with the scintillating mixture; the walls of the container will be made of tedlar-laminated acrylic. The sphere is contained in a 16 m diameter water filled steel tank. About 1600 low radioactivity PMT's coupled to light guides are positioned in a stainless steel support frame in the water at 2 m distance from the acrylic vessel.

The fiducial volume (FV) of the scintillator, set by the analysis software, is a sphere of 6 m diameter, corresponding to 100 t of target mass, which realize a *wall-less* container with a total shielding of about 5 *mwe*. For certain types of investigation, such as the interactions of ν_e from ^{8}B, a large fiducial volume (LVF) of 8 m diameter can be used; it corresponds to 240 tons target mass.

The liquid scintillator is made of trimethylborate (TMB) as a major component, and pseudocumene (PC). This results in a mixture with ultrahigh radiopurity, good transparency and high scintillation efficiency.

Radiopurity, a key issue in Borexino, has been (and currently is) studied for all the materials involved in the construction of the detector.

The intrinsic purity of TMB has been measured to be in the range of $10^{-15}/10^{-16}$ g/g for U,Th. Moreover, estimates from measured decontamination factors in the preparative steps of TMB have shown that the K content will be $< 10^{-14}$ g/g; most other impurities (e.g. ^{210}Pb and ^{60}Co) are also removed with big separation factors. Upper limits for the ^{14}C content are set by the composition of petrochemicals: in CO derived from natural gas one has $^{14}C/^{12}C < 0.75 \times 10^{-18}$. Such limits, if valid in the scintillator, will allow an event threshold of $\simeq$ 250 KeV.

The results for PC commercial samples already are in the range of several 10^{-15} g/g for U,Th due to the refinement process to extract PC from petrochemicals.

The final intrinsic limitation may be due to cosmogenic low energy untaggable activities, such as ^{14}C and Be isotopes; investigation of these activities is one of the main objectives in the Counting Test Facility (see §5).

In the acrylic, a purity of 10^{-12} g/g U,Th is certainly within reach. Moreover K measurements with neutron activation techniques have shown a level of 10^{-9} g/g.

For the water, which is the main shielding component, a purity of below 10^{-13} g/g U,Th has already been reached and a purity of 10^{-10} g/g for K is estimated.

Additional problems, such as Rn emission from contaminated surfaces and leaching from materials are currently under investigation.

Table 1 summarizes the design purity for the various components of the Borexino detector.

Table 1. Purity design goals for Borexino

Material	U (g/g)	Th (g/g)	K (g/g)
Scintillator	$10^{-15}/10^{-16}$	$10^{-15}/10^{-16}$	10^{-14}
Acrylic	10^{-12}	10^{-12}	10^{-9}
Water	10^{-13}	10^{-13}	10^{-10}
PMT's	10^{-7}	10^{-7}	10^{-4}
Steel	10^{-8}	10^{-8}	10^{-5}

With this design numbers the external background (discussed in §3) will become smaller than the internal background at a radius of r = 3 m (see fig. 2). Therefore the inner 6 m diameter sphere of scintillator will be dominated by the internal background.

The dimensions of the fiducial volume are such that low energy events will be measured calorimetrically, as a sum of all the shower processes. The signal threshold for Borexino is $\simeq$75 KeV for a 15 PMT's coincidence trigger and the event threshold is 250 KeV, possibly set by the ^{14}C contamination. PMT's will mostly operate in single photon counting. The measured quantities are:

-) The energy of the event, derived from the number of triggered PMT's.

-) The spatial position of the event, derived from the optical times of flight.

Light collection efficiency (for a scintillation output of 33% of Anthracene), energy resolution and space resolution for events randomly located in FV are shown in table 2.

Table 2. Energy and space resolution in Borexino

Event Energy (MeV)	Photoelectrons	$\sigma(E)$ KeV	$\sigma(r)$ cm
0.1	19 ± 4	24	52
0.2	37 ± 6	32	35
0.5	94 ± 9	50	18
1.0	188 ± 14	76	13

3. Backgrounds

In addition to the internal background, generated by the intrinsic impurity of the detecting medium, there is an external contribution due to various physical phenomena which take place in the material surrounding the FV. The total background can be divided in distinct physical sources, namely:

- Radioactive decay of ^{238}U,^{232}Th chains and ^{40}K in all the materials of the detector
- Beta decays of ^{14}C which can be present in the scintillator
- Background from the walls of the cave: γ and neutron induced interactions
- Cosmic rays and their induced activities

All these components have been evaluated[1] by numerical methods using the measured fluxes in the Hall C for the rock contribution and the purities in table 1 for the detector material contribution. The resulting spectrum is shown in fig. 2.

Various strategies have been devised in order to minimize (and exploit) the residual background contribution in the FV, mostly due to the internal component. It is important to realize that the activity in the scintillator also provides the means to tag a large part of it, thanks to the use of the correlation between events in the radioactive chains, the pulse shape discrimination, and the statistical background subtractions in segments of the chains. The extent to which this can be carried out in Borexino was analysed using the complete Montecarlo simulation of the detector[1]; fig. 3 presents the final reduction of the internal background. Moreover, thanks to the ability of the detector to identify specific decays † the U,Th radiation chains can be used as a live monitors, randomly sampling the active volume and calibrating the energy and space resolution of the detector.

4. The Physics

The Solar Neutrino Problem is the discrepancy between the observed and the predicted flux of neutrinos from the Sun[2]. This discrepancy, firstly observed in the Homestake chlorine experiment for the detection of ν_e from ^{8}B[3], has been recently confirmed by Kamioka[4]. At lower energy, measurements from SAGE[5] tend to indicate a deficit also in the lower-energy (and much less model dependent) neutrino flux.

A complete investigation can be made in Borexino in connection to this problem. For the currently favored *new neutrino physics* scenario for the Solar Neutrino Problem, all relevant solutions can be put to test.

One crucial aspect is the study of ^{7}Be neutrinos, which are monoenergetic and cause elastic scattering on electrons in the detector; the spectrum, with the background contribution, is depicted in fig. 4 where all the effects of finite resolution and inefficiencies have been taken into account. The observed rate, which is $\simeq$ 50 events/day in the FV ‡, will allow to apply the $1/R^2$ signature to tag the solar origin of the effect. The feasibility of this signature is summarized in fig. 5. For a *high purity* scintillator (meaning 10^{-16} g/g in U,Th) and for the SSM flux, one has a statistical significance of more than 4σ in three years of operation.

---

† This will be done mainly through pulse-shape discrimination between α and electron signals, see ref. 1.

‡ This rate is two order of magnitude higher than those observed typically so far.

Another important investigation related to ^{7}Be neutrinos which can be carried out in Borexino is the study of the day-night effect. With a $N_{night} - N_{day}$ ranging up to 15 events and with a definite dependence on the angle of incidence of the neutrino on the earth, it is possible to completely explore all the $\Delta m^2 < 10^{-6}$ eV^2 area in the $(\Delta m^2, sin^2 2\theta)$ space which is allowed for the MSW solution of the Solar Neutrino Problem. Diurnal variations can therefore directly demonstrate MSW flavor conversion in the Earth.

Electron scattering at low energy in Borexino can be made with good sensitivity to the neutrino magnetic moment. If there is a significative difference between the time dependences of Homestake and Kamioka signals caused by magnetic ν_e scattering, the effect will be magnified at low energies and the Borexino signal will be severely incompatible with the Ga signal. In this case an absolute measurement (to a limit of $\simeq 2 \times 10^{-11}\ \mu_B$) can be made with a ^{90}Sr source.

While Borexino is mainly devoted to low energy, a spectral measurement of the ^{8}B neutrino flux in the LFV at a rate about three times that at Kamioka is possible, thus allowing evidence for new neutrino physics (MSW effect, ν decay, *just so* vacuum oscillations) in the $\Delta m^2 < 10^{-5}$ region.

Various non standard scenario can be tested in Borexino. One example is the possibility of Majorana neutrinos with transition magnetic moment, which can be detected thanks to the $\overline{\nu_e}$ sensitivity of the detector.

A detailed description of the above mentioned studies (as well as others) and of their feasibility in Borexino can be found in reference 1.

5. Current Status: the CTF

After several years of work in the radiopurity of the scintillator and the other materials, in the study of the backgrounds and in the development of all the aspects of the problem, a global test of the technique is of great interest[6]. We are therefore setting up a detector based on the above mentioned principle in hall C of the LNGS; it is called Counting Test Facility (CTF).

The Counting Test Facility is the next logical step towards the construction of Borexino; it will assess our capability of reaching a very low background level, aiming at less than 10^2 counts/ton/day above 250 KeV in the scintillator †, and it will give important information on the low energy untaggable activities and on the surface effects which may determine the final sensitivity limit to the investigations in Borexino. Finally it will allow to test all the technical solutions proposed for Borexino.

A conceptual design of the CTF is sketched in fig. 6. It consists of a steel tank (10 m diameter, 9 meters height), filled with deionized water. The sensitive volume is a parallelepiped of base $1.25 \times 1.25\ m^2$ and length of 3 m, for a total active volume of 4 ton of liquid scintillator, viewed by two PMT banks of 25 PMT's each. The module, made of tedlar-laminated acrylic, is surrounded by total-internal-reflection acrylic panels The system is basically a one-dimensional light collector with an efficiency of about 150 photoelectrons per MeV of deposited energy.

† This is two orders of magnitude higher than the goal of Borexino. This goal cannot be reached in the CTF because the shielding is much less than in Borexino and because one cannot apply all the background subtractions.

The overall U,Th chain and K contribution predicted in this configuration[6] is about 3000 events/day with energy greater than 50 KeV.

Among the most important studies in CTF there is the investigation of the level of ^{14}C contamination. Since, for an isotopic abundance of 10^{-18} g/g of ^{12}C, we will have about 1500 counts/day above 100 KeV, this effect can surely be seen in CTF.

The study of ^{7}Be and ^{10}Be will answer some important question such as the contamination that one has in colemanite ores, from which TMB production process begins.

Moreover, a study of three delayed coincidence, listed in table 3, will make it possible to set a limit (at the design level of Borexino, 5×10^{-16} g/g) on the U,Th content of the scintillator. This is possible because the dark counting rate of the PMT system is such that, triggering on 6 PMT's (3 per each side, for an equivalent energy of about 60 KeV), the random rate per day will be negligible.

Table 3. Delayed coincidences to be observed in CTF

Chain	Source	Start event	Daughter	Stop Event	$T_{1/2}$
^{238}U	^{214}Bi	$\beta+\gamma$ 0-3.3	^{214}Po	α 7.8	165 μs
^{232}Th	^{220}Rn	α 6.4	^{216}Po	α 6.9	0.15 s
^{232}Th	^{212}Bi	$\beta+\gamma$ 0-2.2	^{212}Po	α 8.9	305 ns

Eventually the CTF will become a facility to test materials and mixtures to be used in Borexino. Design sensitivity for this operation mode is in the range of 10^{-12} g/g equivalent of U,Th for a 300 Kg acrylic slab positioned close to the counter.

The main auxiliary systems needed to operate the CTF will be:

-) Deionizer for water. It is a continuous automatic system which involves reverse osmosys, ultrafiltration, and an electrodialysis/ion exchange process.

-) Rn stripping system for the water. This additional system is needed in order to reduce the Rn content in LNGS water, which is about 100 PCi/l. This will be made by the use of a buffer tank, to allow Rn to decay and to bubble a suitable flux of He through the water.

-) Fluid-handling system to fill the scintillator vessel. This system will allow to mix TMB and PC and to fill the acrylic vessel; it will be made of vessels for TMB and PC containment, a vessel to mix the components, and an exhaust vessel with alla the relative plumbings. All the system will be lined with a high purity material, such as Teflon PFA.

-) Nitrogen system. This will be used to compensate pressures in various places. Firstly a nitrogen seal will be used at the top of the steel tank to prevent Rn in air from coming in. Secondly nitrogen will be used also as a compensating buffer for the fluid-handling system.

The CTF will be a field test for all this equipment. The construction of the Borexino detector is planned to begin during the operation of the CTF.

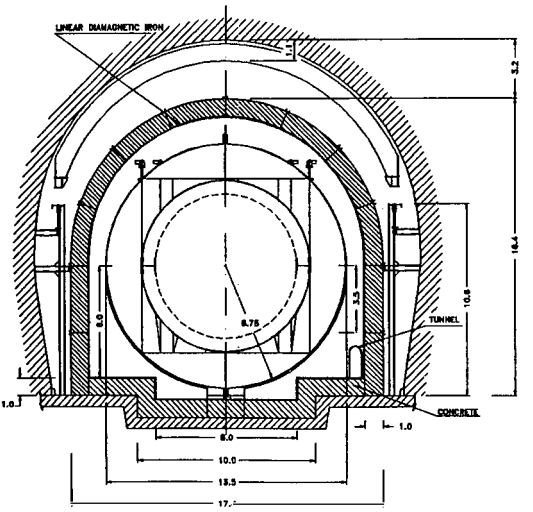


Fig. 1. Schematics of the Borexino detector configuration

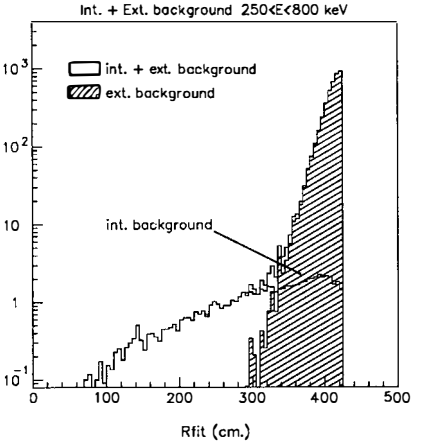


Fig. 2. Radial dependence of the background in Borexino

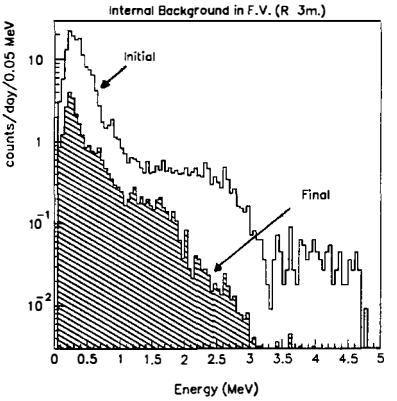


Fig. 3. Global action of the reduction cuts on the internal background in Borexino.

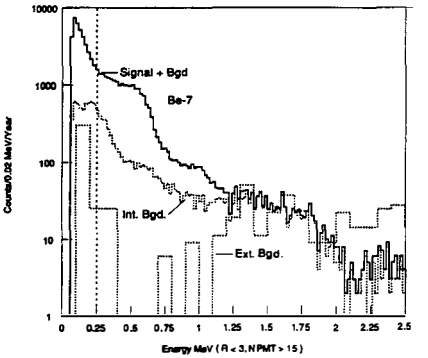


Fig. 4. SSM ^{7}Be signal and total background in Borexino FV

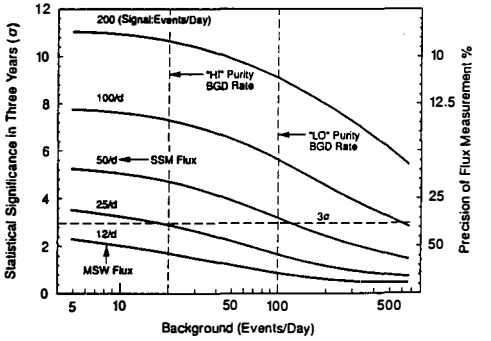


Fig. 5. Applicability of the $1/R^2$ signature to ^{7}Be ν_e flux measurement.

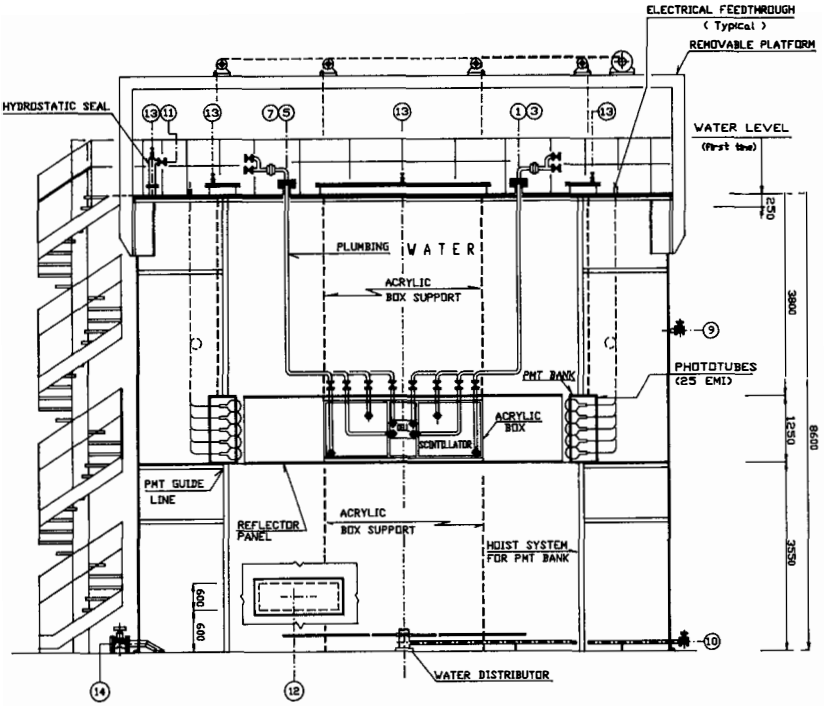


Fig. 6. Conceptual design of the Counting Test Facility.

1,3,5,7) Plumbing system to feed the acrylic vessel

9,10) Water recirculation valves

11) Nitrogen venting

12) Manhole

13) Vents

14) Drainage valve

References

1) Borexino at Gran Sasso: proposal for a real time detector for low energy solar neutrinos. Vol. I

2) J.N. Bahcall, Neutrino Astrophysics, Cambridge University Press

3) Davis R. jr, Phys. Rev. Lett. 12 (1964) 303

4) K. S. Hirata et al., Phys. Rev. D 44 (1991) 2241

5) A. I. Abazov et al., Phys. Rev. Lett. 67 (1991) 3332

6) Counting Test Facility for Borexino. Internal Report.

NEW WAYS FOR REAL TIME DETECTION OF LOW ENERGY SOLAR NEUTRINOS AND OTHER CRUCIAL EXPERIMENTS IN NUCLEAR AND PARTICLE PHYSICS

L. GONZALEZ-MESTRES

L.A.P.P. Annecy
and
L.P.C. Collège de France, Paris

Abstract

Simultaneous detection of light and phonons or heat (the *luminescent bolometer*), was proposed in 1988 in view of dark matter detection [1] and considered later [2] for double beta experiments. Fast single crystal scintillators made of indium compounds have been developed [3-6] for neutrino physics, and in most cases can in principle operate at low temperature. Combining both ideas, simultaneous detection of light and phonons in a scintillating single crystal made of an indium compound and cooled to very low temperature, may reach three basic performances: a) better effective segmentation through digital analysis of the phonon pulse read on each crystal face; b) fast timing from the light strobe; c) good energy resolution, from combined analysis of light and phonon pulses. Such performances are crucial for background rejection in any experiment (solar neutrinos or neutrino-antineutrino oscillations) involving an indium target.

Dark matter and double beta remain the main applications of the luminescent bolometer, but other uses were proposed in [6]: spectroscopy with particle identification, thermal neutron detection with a ^{6}Li target, neutrino experiments based on nucleus recoil... A new possibility is application to heavy ion physics [7], where energy resolution can be combined with fast timing from scintillation and space resolution from the phonon pulses. Simultaneous detection of light and phonons provides equally a new way for basic studies of relaxation phenomena of excitations in solids.

Photosensitive devices based on superconducting tunnel junctions appear as a suitable read out, able to collect the light pulse followed by the delayed front of phonons. They can be operated at ^{3}He temperatures ($T > 300$ mK) with excellent sensitivity [8].

Simultaneous detection of ionization and phonons may be an alternative to the luminescent bolometer for experiments with an indium target (InP, InSb...). Superheated superconducting granules with the avalanche effect are not ruled out for this purpose.

1. PHYSICS POTENTIALITIES WITH CRYOGENIC DETECTORS

Real time solar neutrino detectors are an urgent need, and must be sensitive to the pp part of the spectrum. Dark matter experiments based on nucleus recoil should reach event rates as low as 10^{-2} kg^{-1} day^{-1} . Double beta experimental programs are pursued with success, but target diversification is required to obviate nuclear physics uncertainties [9]. Neutrino oscillations remain an unsolved theoretical and experimental problem in elementary particle physics. Heavy ion physics in the energy range 1 to 40 MeV per nucleon, will require exceptional detector performances in experiments at the next generation of heavy ion facilities [10], [11], [7]. Spectroscopy is searching for ways to improve energy resolution, as compared to semiconductor devices. In any underground experiment, searching for rare events, a crucial issue is the reduction and rejection of radioactive background. Low temperature detectors provide technical solutions to such challenges (high sensitivity, low thermal noise). Working at low T allows to detect several components of the signal which can be used for particle identification: ionization + phonons [12], light + phonons [6]; in both cases, fast timing can be preserved. Cryogenic devices allow for exceptional energy resolution from thermal signals [9], and good space resolution inside a crystal using ballistic or scattered phonons [13].

Solid state physics and chemistry will also benefit from such a progress in instrumentation. By simultaneous detection of ionization or light and phonons at low temperature, with high sensitivity and low excitation energy, it will be possible to uncover new features of crystal structure and electronic energy levels, usually masked by thermal noise, nonlinearity and lack of complete information on relaxation processes. As compared to ionization, where carrier collection produces extra phonons, luminescence presents the advantage that the sensor does not disturb the crystal under study. Simultaneous detection of quasiparticles and phonons with superconducting tunnel junctions, allows to investigate the properties of nonequilibrium superconductors [14]. The measurement of the relative strength of radiative and nonradiative processes at T = 1.5 K was used [15] to study energy transfer in GaAs p-n junctions. Low temperature fluorescence is a classical tool [16] to explore electronic energy levels and electron-lattice coupling, as well as the structure and behaviour of molecular complexes. Bolometers allow to measure time-dependent specific heats, which are often missing in solid state physics litterature.

The appearance of powerful computers and digital signal processors allows for an approach based on a detailed and exhaustive study, event by event, of energy degradation in crystal detectors. In some cases, it is even possible to perform on-line digital analysis. Cryogenic particle detectors, but also solid state experiments, should serioulsy benefit of this impressive development of real-time and computing facilities.

The original proposal [1] to use simultaneous detection of light and heat for particle identification (the *luminescent bolometer*) has been presented at previous Moriond Workshops [17], [2]. The development of scintillating single crystal detectors made of indium compounds was also dealt with in these contributions. We report on new results and ideas on both topics, and suggest to combine them in a neutrino detector.

2. THE INDIUM SOLAR NEUTRINO DETECTOR

Some time ago [3], it was proposed to use *single crystal scintillators* made of indium compounds to detect solar neutrinos through Raghavan's reaction [18]. With 128 keV threshold, an inverse beta reaction on ^{115}In creates an excited state of tin which decays with a lifetime of 3.3 μs emitting two γ rays (116 keV and 497 keV). The delayed coincidence would allow for a clean signature, if the detector is fast and if the two delayed γ's are absorbed in two different cells. A planar array of $\approx 10^4$ long crystals (20 cm $\times$ 3 cm $\times$3 cm) read by photomultipliers (PM), or a cube read by avalanche photodiodes (APD) were foreseen [1]. The second solution allows for a compact detector without dead volume and radioactivity from the read-out.

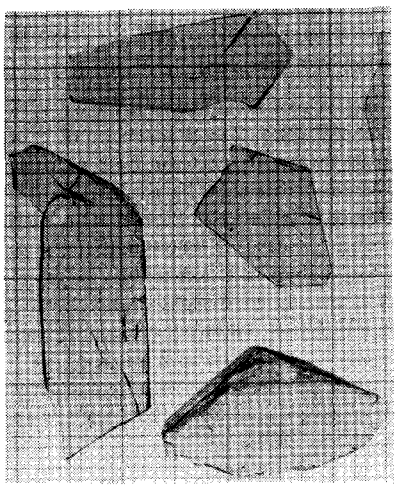
Transparent single crystals of terbium-doped $InBO_3$ (Fig. 1) were grown [4] . Tb^{3+} fluorescence is too slow, and doping with cerium was attempted. However, it is known [19], [20] that very often Ce^{3+} does not stabilize in trivalent indium sites. Attempts to make scintillating $InBO_3$:Ce^{3+} powder failed. A chemical way out was found [5]: given an indium compound InR , where R is some radical (e.g. BO_3 $^{3-}$), find another compound, XR , such that XR and InR have the same crystal structure and the ion X^{3+} allows for cerium doping. Then, it is possible to obtain compounds of the form $In_xX_{1-x}R$, and dope with Ce^{3+} the X^{3+} sites. The feasibility of this procedure was checked with powders, in compounds incorporating up to 50% of indium in volume. Scintillating single crystals of In/Sc BO_3:Ce^{3+} were grown (Fig. 2). Current goals are: a) compounds allowing for growth of large crystals at a reasonable cost (e.g. fluorides [21]); b) the best possible light yield (the most performant cerium scintillator seems to be $YAlO_3$:Ce^{3+} , with $\approx$ 5% of the energy converted into light [22]).

Background turned out to be very difficult to handle, as 1 event/day is expected with 4 tons of indium and ^{115}In is radioactive (0.2 event g^{-1} s^{-1}). The β spectrum of ^{115}In decay goes up to 490 keV in energy. It overlays the inverse β spectrum from pp solar neutrinos, and its tail lies close to the 497 keV γ . Main sources of background are: a) 3 β coincidences; b) delayed coincidences between a β from ^{115}In decay and a γ from ambient radioactivity which would fake the two γ's of ^{115}Sn** decay. It seems unrealistic to expect radioactivity rates better than 10^{-6} Bq g^{-1} in the region E $>$ 500 keV . To reject background, three basic conditions arise [18]: a) extremely good segmentation ($\approx 10^6$ elementary cells); good energy resolution ($\approx$ 10% FWHM at 100 keV); reasonably fast timing ($\approx$ 100 ns). We recently proposed [6] *cryogenic detection* as a solution. Ce^{3+} doping seems appropriate for this purpose, as trivalent cerium fluorescence remains fast ($\tau <$ 60 ns) and can improve in light yield as temperature goes down [23]. Eu^{2+} is an alternative, with slower luminescence signal ($\tau \approx$ 900 ns [24]).

Work on superconducting tunnel junctions (STJ) by the Oxford group [8] was presented at the previous January Moriond Workshop [25]. The results obtained with a 432 STJ array on a 1 cm^2 thin absorber, where the possibility to reach a threshold below 1 keV clearly emerges, are particularly relevant to the indium solar neutrino detector. At low temperature, *a STJ-based device* on the surface of a scintillating crystal can

read the fast light strobe followed by the phonon pulse. Arrays of STJ in series provide excellent results at ^{3}He temperatures and are well suited for large detectors. With 40% of In in weight, and a density of 5 g/cc , $\approx$ 30000 $4 \times 4 \times 4$ cm^3 crystals could be used, with a STJ device per crystal face ($\approx 2 \times 10^5$ electronic channels). After a trigger based on a 10 μs gate, digital analysis of the full event would extract good energy resolution and position information ($\approx$ 1 cm FWHM). Effective segmentation would be $\approx 2 \times 10^6$ elementary cells. Read-out design requires a compromise between the requirements of light collection, energy resolution and timing. Blackened STJ or a thin black absorber can improve light collection. Photosensitive STJ are an important research subject [26]. $\bar{\nu}$ - ν oscillations at reactors ($\approx$ 100 kg detector, $\approx$ 300 crystals, $\approx$ 1800 electronic channels) may be a first experiment where to use and test a cryogenic indium detector.

An alternative to low temperature scintillators is simultaneous detection of ionization and phonons in a cold semiconductor (InP, InSb, InAs...). However, the operating temperature should then be lower ($\approx$ 20 mK), and microphonics and stability problems far more delicate. High quality crystal growth and semiconductor properties are most likely the main challenges. The status of superheated superconducting granules (SSG) development was reviewed at the 1990 January Workshop [27], and recent improvements were reported at the Oxford Workshop on cryogenic detectors [28]. With the mechanism of "amplification by thermal micro-avalanche" [29], SSG remain an interesting possibility for the development of the indium solar neutrino detector.



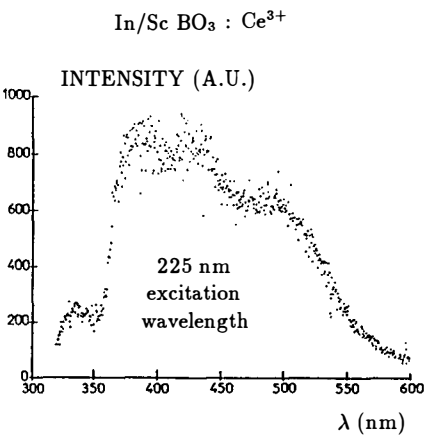


Fig. 1 (left) - Transparent single crystals of $InBO_3$:Tb^{3+} used for spectroscopic analysis in [4] (J.P. Chaminade, CNRS Talence).

Fig. 2 (right) - Fluorescence spectrum of Ce^{3+} in a single crystal of In/Sc BO_3:Ce^{3+} cooled to 7 K , from [5] (B. Jacquier, CNRS Lyon).

3. APPLICATIONS OF THE LUMINESCENT BOLOMETER

To study heavy ion reactions in the region 1 MeV < E (per nucleon) < 40 MeV , two contributions may emerge from cryogenic detectors. First, to detect the produced nuclei: very good energy resolution is required, with fast timing and position information in a $\approx$ 1mm^3 crystal. The luminescent bolometer appears as an excellent candidate. Thermal detectors for heavy ion physics are already being studied at GSI [10], [30]. More ambitious would be to use the luminescent bolometer for associated γ spectroscopy, to identify decays with a 4π massive detector. The device shoud reach better performances than the semiconductor germanium used in EUROGAM [11]. Around an array of small detectors devoted to the identification of the produced nuclei, larger bolometers devoted to gamma spectroscopy would identify the produced γ rays from the decay of excited states. Arrays of STJ operating at 300 mK would be an excellent read-out, compatible with the performance expected from a large refrigerator in a heavy ion beam. Such a development would also provide an excellent workbench in view of neutrino physics.

Dark matter detection [1] will need the luminescent bolometer to unambiguoulsy search for WIMPs with axial coupling, using a target with an odd number of protons. This is required to prevent uncertainties related to our poor knowledge of the nucleon spin structure. Simultaneous detection of light and phonons will allow for the identification of events due to recoiling nuclei and rejection of β's or γ's. If $\approx$ 10% of the energy is converted into light for a β or γ at very low temperature, it seems possible to operate with a threshold lower than 10 keV, after a suitable technical development. With a STJ read-out [6], [8], energy resolution and discrimination ability would be at least as good as reported from simultaneous detection of ionization and heat [12]. The low gap for quasiparticle excitation in superconductors plays a crucial role in the performance of the photosensitive read-out. Targets such as ^{7}Li, ^{19}F, ^{27}Al, ^{127}I, ^{183}W... can be incorporated in the cold scintillator approach. Operating at ^{3}He temperatures appears as a decisive advantage of the luminescent bolometer in view of neutralino searches, where very low event rates must be explored and large detectors (100 kg to 1 ton) will eventually be required. For such a large detector, self-active-shielding would be an excellent way to reject neutron background, where a neutron will undergo interactions in several crystals. A similar technique may apply to neutrino physics using nucleus recoil.

Double beta experiments benefit of the same technique [2], as long as the main background is given by α particles, which usually happens at the highest values of Q . Using a STJ read-out, the overall energy resolution can be extremely good ($\approx$ 0.1%). Experimental results on a luminescent bolometer with a photodiode read-out for light are presented at this Workshop by the Milano group [9]. Energy resolution remains crucial to reject the ultimate $\beta\beta 2\nu$ background. Thermal neutron detection at low counting rate may use a luminescent bolometer incorporating a ^{6}Li target [6]. The capture reaction, producing α + ^{3}H with Q = 4.78 MeV , would acquire a much better signature. Spectroscopy involving particle identification may benefit of this technique, without significantly spoiling energy resolution as compared to thermal bolometers. Working at ^{3}He temperatures with a STJ read-out would again be a substantial advantage.

References

[1] L. Gonzalez-Mestres and D. Perret-Gallix, Nucl. Instr. Meth. A279, 382 (1989); Proceedings of the XXIV International Conference on High Energy Physics, Munich August 1988 (Ed. Springer-Verlag).

[2] L. Gonzalez-Mestres and D. Perret-Gallix, Moriond Astrophysics Workshop March 1989 (Ed. Frontières).

[3] L. Gonzalez-Mestres and D. Perret-Gallix, Proceedings of the Rencontre sur la Masse Cachée, Annecy July 1997 (Annales de Physique).

[4] J.P. Chaminade et al., Journal of Crystal Growth 99, 799 (1990).

[5] J. P. Chaminade, B. Jacquier, L. Gonzalez-Mestres and D. Perret-Gallix, unpublished and ULTIMATECH project, CNRS (France) 1991.

[6] L. Gonzalez-Mestres, Proceedings of TAUP-91, Toledo September 1991 (Ed. Nuclear Physics); Proceedings of LTD-4 , Oxford September 1991 (Ed. Frontières).

[7] I. Berkes, B. Chambon, D. Drain, L. Gonzalez-Mestres and B. Jacquier, "Bolométrie et luminescence à très basse température", Int. Rep. Univ. C. Bernard, Lyon 1992.

[8] N. Booth et al., in LTD-4 .

[9] L. Zanotti et al., these Proceedings.

[10] P. Egelhof et al., Proceedings of LTD-3 , Gran Sasso September 1989 (Ed. Frontières). Low Temperature Detectors for Neutrinos and Dark Matter - III .

[11] See MM. Aleonard, preprint CENBG 9115, on the EUROGAM project.

[12] B. Sadoulet et al., in LTD-4 .

[13] See, for instance, Th. Peterreins et al., in LTD-3 .

[14] See, for instance, P. Hu et al., Phys. Rev. Lett. 38, 361 (1977).

[15] V. Narayanamurti, R.A. Logan and M.A. Chin, Phys. Rev. Lett. 40, 63 (1978).

[16] See B. Jacquier in LTD-2 , Annecy May 1988 (Ed. Frontières), and references therein. Low Temperature Detectors for Neutrinos and Dark Matter - II .

[17] J.P. Chaminade et al., Moriond January Workshop 1989 (Ed. Frontières).

L. Gonzalez-Mestres and D. Perret-Gallix, Moriond January Workshop 1988.

[18] R.S. Raghavan, Phys. Rev. Lett. 37, 259 (1976). M. Deutsch, discussions.

[19] T. Hoshina and S. Kuboniwa, J. Phys. Soc. Japan 32, 771 (1972).

[20] Y. Tsujimoto et al., Journal of Luminescence 9, 475 (1975).

[21] J.P. Chaminade et al., Acta Crystallographica, to appear.

[22] S.I. Ziegler et al., TRIUMF preprint TRI-PP-91-79 , November 1991.

[23] J. Mares et al., Revue Phys. Appl. 22 , 145-152 (1987), and references therein.

[24] B. Jacquier et al. and ULTIBOL program, CNRS (France) 1991.

[25] R.J. Gaitskell et al., Moriond January Workshop 1991 (Ed. Frontières).

[26] A. Barone and M. Russo, in Advances in Superconductivity, Plenum 1983. B. Cabrera, R. Cristiano, J.P. Maneval and G. Waysand, private discussions.

[27] L. Gonzalez-Mestres, Moriond January Workshop 1990 (Ed. Frontières).

[28] LTD-4 , IVth International Workshop on Low Temperature Detectors, Oxford September 1991, edited by N.E. Booth and G.L. Salmon (Ed. Frontières).

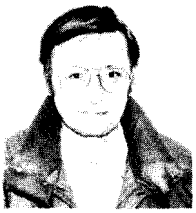
[29] L. Gonzalez-Mestres and D. Perret-Gallix in LTD-2 .

[30] P. Egelhof et al., in LTD-4 .

Neutrino Properties / Neutrino Masses

Result of the Zürich experiment to measure the mass of $\overline{\nu}_e$ from tritium β-decay

E. Holzschuh, M. Fritschi, and W. Kündig
Physik-Institut der Universität Zürich, CH-8001 Zürich



Abstract

Measurements of the tritium β-spectrum are reported. Thin sources consisting of a monolayer of tritiated hydrocarbon molecules were used. No indication of a nonzero mass m_ν of the electron antineutrino was found. The result is $m_\nu^2 = -24 \pm 48 \pm 61\ \mathrm{eV}^2$ (1σ). An upper limit $m_\nu < 11\,\mathrm{eV}$ (95% confidence level) is derived.

1. Introduction

Measuring the mass of the electron antineutrino (here denoted by m_ν) from the shape of the tritium β-spectrum is the most sensitive, direct, i.e. kinematic, method for m_ν and has been used for many decades. At the 6th Moriond Workshop 1986, we have shown [1,2] that the only experiment [3] suggesting a nonzero value of m_ν of approximately 35 eV is likely to be incorrect and an upper limit $m_\nu < 18\,\mathrm{eV}$ (95% confidence level, CL) was reported. In the mean time others have also reported results without any indication for a nonzero m_ν [4,5]. For a review of recent tritium experiments we refer to ref. [6].

Our result from 1986 was dominated by systematic uncertainties. In the mean time substantial improvements have been made in all parts of of the experiment and some preliminary results have been reported [7].

2. Measurements

The instrument employed is a toroidal field, magnetic spectrometer of the Tret'yakov type modified with a radial, electrostatic retarding field around the source [8]. Spectra are recorded by setting the magnet to a constant analyzing energy E_{mag} and stepping the source potential V_s. Analyzed electrons are detected with a position sensitive proportional counter and assigned an energy

$$E = E_{\mathrm{mag}} + eV_s + z/D \tag{1}$$

where z is the axial position relative to the centre of the detector ($|z| \leq 50\,\mathrm{mm}$). This equation was was used to construct one spectrum for the whole detector. The setting for the tritium measurements was $E_{\mathrm{mag}} = 2200\,\mathrm{eV}$ and $D = 0.97\,\mathrm{mm/eV}$ (dispersion). The spectrometer resolution function was measured with a conversion line source [7]. The width was found to be 17 eV (FWHM), in good agreement with a Monte Carlo simulation.

Two tritium sources were produced by the spontaneous formation of a selfassembling monolayer. This method gives very thin and uniform sources and has been described previously [7].

The tritium spectrum was scanned in 4 eV steps and the energy range was divided into three intervals (17650 - 18000 - 18400 - 18750 eV). The counting times in the intervals (2, 6, 20 s/step) were chosen to improve statistics around the endpoint. A total of 1709 up-down sweeps of the spectrum was accumulated where each sweep took 105 min. The raw results of the measurements were recorded as event data on magnetic tape and spectra were built off-line using eq. (1).

3. Data analysis and results

Four auxiliary distributions were required to analyse the tritium spectra. The spectrometer resolution function was taken from the measurement [7]. The backscattering distribution was minimized by choosing a source substrate with low Z (mainly carbon) and was measured in an auxiliary experiment [7]. The average thickness, $\overline{d}$, of the tritium sources was 9 Å. Hence energy loss was relatively uncritical and its distribution was constructed using data for polyethylene [9] which, for the sake of energy loss, has the same chemical structure as our sources.

The distribution of the electronic final states (see fig. 1) was taken from ref. [10]. These extensive calculations confirmed the earlier conclusion [11] that the final states are mainly

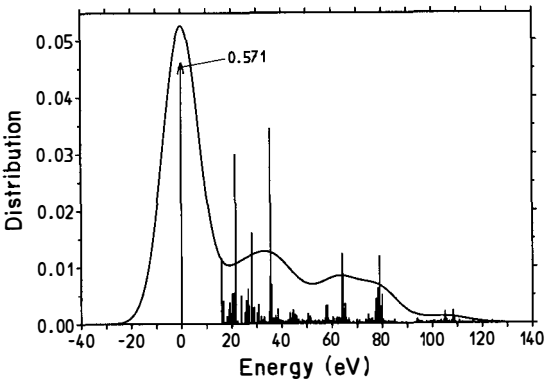


Figure 1: Distribution of electronic final states for $CH_3 - CHT - CH_3$ [10] used for the analysis. The solid line is the same distribution convoluted with a Gaussian with 17 eV FWHM (approximately the spectrometer resolution) indicating that the data analysis is not sensitive to the details of the computed distribution.

determined locally, i.e. by the fact that there is a C-T bond and the overall structure of a molecule is of relatively minor importance. The mean excitation energy, $\overline{\Delta E^*}$, can be calculated accurately from groundstate energies using a sum rule [11] ($\overline{\Delta E^*} = 19.1 \pm 0.4$ eV). Similarly, the transition probability to the groundstate, W_{00}, can be obtained from groundstate wavefunctions which can be calculated more accurately than the wavefunctions of excited states ($W_{00} = 0.571 \pm 0.015$). The mean excitation energy computed from the distribution was larger by 1.8 eV than the value from the sum rule. We corrected the distribution by shifting the excitation energy such that the resulting distribution agreed with the sum rule and took the 1.8 eV discrepancy as an indication for a 10% error in the computed excitation energies.

The theoretically expected shape of the β-spectrum was assumed to be of the form

$$\frac{dN}{dE} = f_{\text{theo}}(E) \sim F(Z,W)pW \sum_n W_{0n}\varepsilon_n \sqrt{\varepsilon_n^2 - m_\nu^2}\,\Theta(\varepsilon_n - m_\nu) \tag{2}$$

with

$$\varepsilon_n = E_{00} - \Delta E_n^* - E \tag{3}$$

and where p, W, E are momentum, total and kinetic energy. The sum runs over all electronic final states. The extrapolated endpoint energy for transitions to the electronic groundstate is denoted by E_{00} and the Fermi function by $F(Z,W)$.

The model function for the measured spectra was taken to be

$$f_{\text{exp}}(E) = AS(E)(1 + \alpha\varepsilon_0)R \otimes E_L \otimes f_{\text{theo}}(E) + B \tag{4}$$

where R is the spectrometer resolution function, E_L the energy loss distribution and $\otimes$ denotes convolution. The function $S(E)$ represents the energy-dependent solid angle accepted at the source and was calculated analytically. It varies nearly linearly over the measured range by 6.4%. Backscattering was taken into account by the parameter α, i.e. the measured backscattering probability was only used for a consistency check. The free

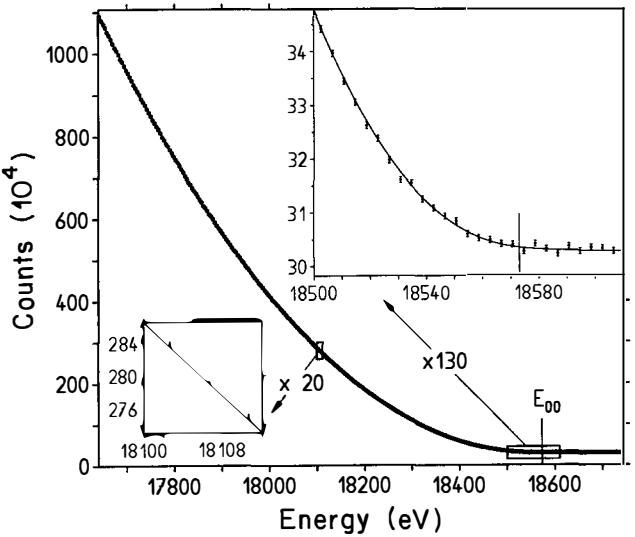


Figure 2: Measured tritium spectrum (points) from all data and best fit (solid line). For $E \geq 18400$ eV actual counts are plotted. For lower energies the data were scaled by the counting periods. Insets show parts with the vertical scale expanded by the indicated factor.

parameters in the analysis were A, B, α, E_{00}, and m_ν^2. The later was allowed to take on negative values [6].

The results of the fits are given in table 1. As can be seen the results are mutually compatible at the statistical 1σ level. All the available data were also added into a single spectrum (all) which is plotted with the best fit in fig. 2 and with which all further analysis was performed.

The fitted value of m_ν^2 is compatible with zero and gives no indication for a nonzero neutrino mass. The value for the endpoint energy E_{00} is close to what has been found recently by others [4,5]. The uncertainties for E_{00} are dominated by systematics (calibration,

Table 1: Fit results for six tritium spectra (1–6) and for their sum (all). The errors given are 1σ statistical. The number of degrees of freedom are 271.

Spectrum	χ^2	m_ν^2 (eV^2)	$E_{00} - 18570$ (eV)	$\alpha \times 10^5$ (eV^{-1})
1	308.2	-136 ± 149	3.80 ± 0.55	1.39 ± 0.36
2	274.7	$+57 \pm 122$	3.93 ± 0.46	1.91 ± 0.30
3	289.2	-95 ± 106	2.95 ± 0.40	1.64 ± 0.26
4	270.4	$+25 \pm 115$	3.17 ± 0.43	1.83 ± 0.29
5	268.9	-23 ± 108	2.84 ± 0.41	1.50 ± 0.28
6	278.0	$+8 \pm 113$	3.51 ± 0.42	1.80 ± 0.28
all	304.9	-24 ± 48	3.32 ± 0.18	1.69 ± 0.12

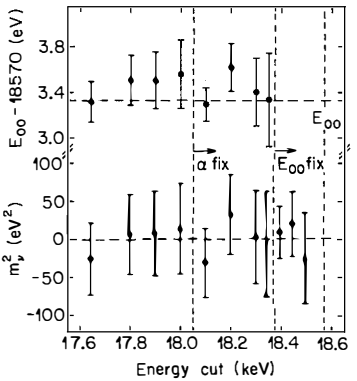


Figure 3: Fitted parameters E_{00} and m_ν^2 when the spectrum points below the energy $E_{\rm cut}$ are excluded from the fit.

etc., having no effect on m_ν^2), the analysis of which will be reported later.

To test for the presence of unaccounted spectrum distortions, the data points below an energy $E_{\rm cut}$ were excluded from the fit. The results for m_ν^2 and E_{00} are plotted in fig. 3 as a function of $E_{\rm cut}$. For $E_{\rm cut} \leq 18000$ eV all parameters were varied freely. Above this value, α was kept fixed at its value from all data, as it becomes indeterminate over a short energy range. Similarly E_{00} was kept fixed above 18400 eV. Clearly there is no systematic variation beyond statistics either for m_ν^2 or for E_{00}, indicating that the fitted model correctly describes the data over the entire range. We are thus taking the parameter values obtained with all data as the best estimate.

Systematic uncertainties for m_ν^2 were determined by changing the various distributions assumed for the analysis within estimated limits and fitting the data again with all parameters varied freely. The results, at the 1σ level, are given in table 2. As our sources were very thin, uncertainties due to energy loss are fairly small. Errors are given for the average energy loss per interaction, $\overline{\delta E}$, and for the ratio $\overline{d}/\lambda$ as the source thickness and

Table 2: Systematic uncertainties δm_ν^2 (1σ) when the most critical quantities of the various distributions, used in the data fit, are changed by the indicated amount.

Quantity	Uncertainty	δm_ν^2 (eV^2)
Energy loss distribution:		
$\overline{\delta E}$	10 %	6
$\overline{d}/\lambda$	20 %	11
Spectrometer resolution:		
Width	3 %	7
Tail intensity	15 %	30
Electronic final states:		
W_{00}	0.015	26
$\overline{\Delta E^*}$	0.4 eV	33
Distribution	–	30
Total systematic uncertainty	–	61

the mean free path ($\lambda = 410$ Å) enter the analysis only in this combination. Uncertainties due to the spectrometer resolution function are dominated by the presents of extended tails whereas the central part, containing 97% of the integrated intensity, makes only a small contribution.

The three quantities, $\overline{\Delta E^*}$, W_{00} and shape of the distribution, were calculated separately in [10] to the highest respective accuracy possible. Hence we assigned three independent systematic uncertainties to the final states. By adding all systematic uncertainties in quadrature we obtain

$$m_\nu^2 = -24 \pm 48 \pm 61\,\mathrm{eV}^2 \qquad (1\sigma) \tag{5}$$

which we consider as our final result.

There are several (reasonable) methods for computing upper limits in the presents of a physical boundary ($m_\nu \geq 0$) and upper limits are thus not unique. For our result the differences are small and we obtain

$$m_\nu < 11\,\mathrm{eV} \qquad (95\%\,\mathrm{CL}). \tag{6}$$

To conclude no indication for a nonzero neutrino mass was found in this investigation. This is in agreement with other recent experiments.

This work was supported by the Paul Scherrer Institut and by the Swiss National Science Foundation.

References

[1] M. Fritschi, E. Holzschuh, W. Kündig, J. W. Petersen, R. E. Pixley, and H. Stüssi, Proc. 6th Moriond Workshop (Tignes, 1986), ed. O. Fackler and J. Tran Thanh Van (Gif sur Yvette, France: Editions Frontieres) p 469.

[2] M. Fritschi, E. Holzschuh, W. Kündig, J. W. Petersen, R. E. Pixley and H. Stüssi, Phys. Lett. B 173 (1986) 485.

[3] V. A. Lubimov *et al.*, Phys. Lett. B 94 (1980) 266, S. Boris *et al.*, Phys. Lett. B 159 (1985) 217, and Phys. Rev. Lett. 58 (1987) 2019.

[4] R. G. H. Robertson, T. J. Bowles, G. J. Stephenson, D. J. Wark, J. F. Wilkerson, and D. A. Knapp, Phys. Rev. Lett. 67 (1991) 957.

[5] H. Kawakami *et al.*, Phys. Lett. B 256 (1991) 105.

[6] E. Holzschuh, Rep. Prog. Phys. (1992) in print.

[7] M. Fritschi, E. Holzschuh, W. Kündig and H. Stüssi, Nucl. Phys. B (Proc. Suppl.) 19 (1991) 205.

[8] W. Kündig, J. W. Petersen, R. E. Pixley, H. Stüssi and M. Warden, Proc. 4th Moriond Workshop (La Plagne, 1984), ed. J. Tran Thanh Van (Gif sur Yvette, France: Editions Frontieres) p 261.

[9] L. R. Painter, E. T. Arakawa, M. W. Williams, and J. C. Ashley, Rad. Res. 83 (1980) 1; M. J. Berger and S. M. Seltzer, ICRU Report 37 (1984); K. O. Al-Ahmad and D. E. Watt, J. Phys. D 17 (1984) 1899.

[10] T. A. Claxton, S. Schafroth, and P. F. Meier, Phys. Rev. A, in print; S. Schafroth, Thesis, University of Zürich (1992).

[11] I. G. Kaplan and V. N. Smutny, Adv. Quantum Chem. 19 (1988) 289.

FIRST TRITIUM DECAY SPECTRA FROM THE MAINZ NEUTRINO MASS EXPERIMENT

H. Backe, H. Barth, J. Bonn, B. Degen, L. Fleischmann,
J.U. Grooβ, G. Kube, A. Osipowicz, E.W. Otten, A. Picard,
M. Przyrembel, Ch. Weinheimer
Institut für Physik, Universität Mainz, Germany

P. Leiderer
Fakultät für Physik, Universität Konstanz, Germany

Presented by E.W. Otten

Abstract

After having tested the performance of the new solenoid retarding spectrometer by precision measurements on the electron conversion spectrum of ^{83m}Kr, first tritium decay spectra have been taken from a source of frozen T_2. The signal exceeds the background rate of 0.03 Hz for energies up to 30 eV below the endpoint. A preliminary analysis yields a Q-value of the tritium decay of (18589±3) eV. We expect from the final analysis of the present data set an upper limit for the electron antineutrino rest mass below 10 eV/c^2.

In preceding contributions to this workshop [1,2], we described the principle and the function of a new solenoid electron retarding spectrometer. It provides simultaneously high resolution and high transmission which are essential features for investigating the endpoint region of the β-decay of tritium in search for a finite electron antineutrino rest mass. The main components of the spectrometers are as follows: A film of 10 - 30 monolayers of molecular tritium is frozen onto a 1 cm^2 aluminum backing placed near the centre of a solenoidal field $B_0 \simeq 2.4$ T. The decay electrons emerging from the source expand adiabatically into a low field region of $B_1 \simeq 8 \times 10^{-4}$ T where they have to pass an electrostatic analyzing potential which is scanned in the vicinity of the endpoint of the β-spectrum $E_0 \simeq 18.6$ keV. Due to the adiabatic transformation of transverse cyclotron motion into longitudinal motion along the guiding magnetic field lines the whole forward solid angle of emission is analyzed within a filter width of

$$\Delta E = E\,(B_1/B_0) \qquad (1)$$

which amounts to $\Delta E \simeq 6$ eV for the settings given above. Electrons passing the filter are refocussed by two following solenoids onto a silicon detector which is separated into five concentric segments each of them having an area of 1 cm^2.

More detailed information on the source, the spectrometer and the detector may be found in references [3] - [5]. An instrument based on the same adiabatic principle has been built by a group at the Institute of Nuclear Research in Troitzk [6].

The performance of the instrument may be illustrated by the spectrum of the K-32 conversion line of ^{83m}Kr frozen onto an aluminum backing (Fig. 1). The instrument was set at a field ratio $B_0/B_1 = 3000$. In the region of steep transmission change data points were taken in steps of 2 V. The insert in Fig. 1 shows an expanded view of the spectrum close to the K-line. In this region the spectrum has been fitted with an elastic K-line component of Lorentzian shape. It is followed by a satellite spectrum of shake up and shake off events which is approximated by Gaussians convoluted with an exponential tail (see the dotted lines in the insert of Fig. 1). The fit uses the analytical transmission function given in ref. [7]. Below the shake off region the integral spectrum is rising further due to electrons backscattered

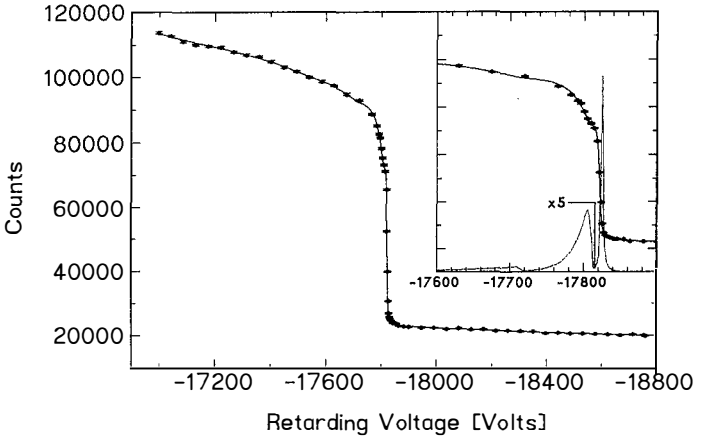


Fig. 1: Integral transmission spectrum measured in the region of the K-32 conversion line of ^{83m}Kr. Also shown is the fitted spectrum and its differential components in the insert.

from the aluminum substrate. Shape and height of the backscattered spectrum extracted from this measurement are in good agreement with a Monte-Carlo simulation. The K-32 line is sitting on a background which stems mainly from the backscattered tail of the L-32 components (compare also ref. [7]).

First tritium spectra were taken in two test runs and one production run in 1991. Fig. 2 shows the measured raw data obtained. The critical endpoint region is blown up in the insert. Plotted is the measured count rate as a function of the effective retarding potential which is the difference between the filter potential fixed to U_0 = 18877 V and the source potential U_S scanned in the respective interval. The plot represents 40 % of the data collected in the four weeks of running in November/December 1991. The signal rate equals the background rate of 0.030 Hz already 30 eV below the endpoint (the number of background counts collected in a single point above E_0 is about 550). The full line is a fit to the data with the neutrino mass m_ν fixed to zero. The preliminary fit yields an endpoint of the β-spectrum of T_2 imbedded in a matrix of frozen T_2 of

$$E_0 = 18572 \text{ eV} \qquad (2)$$

From this result we calculate a preliminary value for the mass difference

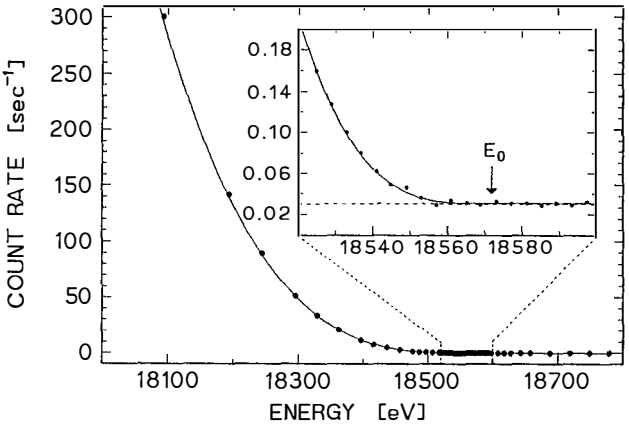


Fig. 2: Measured integral tritium decay spectrum in the endpoint region together with a fitted curve with m_ν fixed to 0. The insert shows an expanded view close to the endpoint E_0.

$$Q = m(^3H) - m(^3He) = 18589\ (\pm 3)\ eV \qquad (3)$$

The error bar on this result should be regarded as preliminary and conservative.

The source was prepared by evaporating 20 - 30 monolayers of T_2 onto the aluminum substrate under ellipsometric control. Although the source was kept at a temperature of 2.6 K, it diminished by a factor of 2 within one week of running with a slope somewhat inbetween linear and exponential. The source was refreshed after 7 to 10 days of running, depending on its actual thickness. Most of the evaporated T_2 was withheld by a tube mounted onto the source which was also cooled to 2.6 K. It limited the solid angle of free evaporation into the spectrometer to about $\Delta\Omega/4\pi \simeq 10^{-3}$. Although part of the tritium certainly evaporated into the spectrometer, it did not apparently contaminate the spectrometer: after removing the source the background rate dropped back to the value it had before the run. When the source was exposed to the spectrometer the background rate rose by about 50 % at maximum.

Fig. 3 shows the data points of Fig. 2 again but in a linearized plot. This is achieved to first approximation by taking the third root of the signal rate. However, the data bend up at about 50 eV below the endpoint. The fit

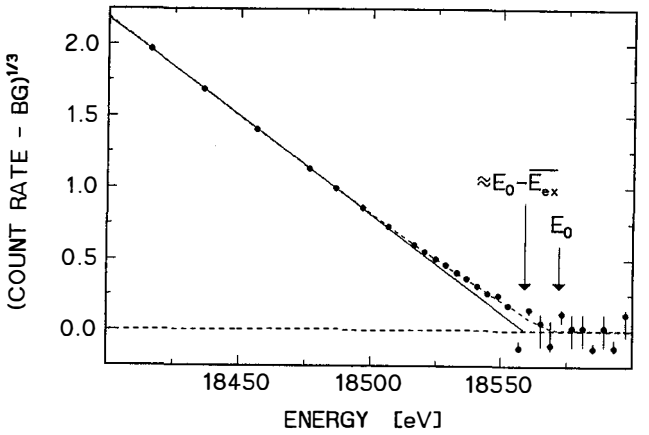


Fig. 3: Same data as in Fig. 2 but linearized by taking the third root of the signal rate. The linear extrapolation from data far below E_0 points to an endpoint shifted by the average excitation energy $\overline{E_{ex}}$ of the daughter molecule.

with $m_\nu = 0$ (broken line) indeed follows this bent and meets the abszissa at the endpoint quoted above. These events close to the endpoint correspond to an elastic emission of the electron where the daughter molecule $(^3He^3H)^+$ is left in its electronic ground state. The data at lower energies, on the other hand, point to an endpoint about 13 eV below that maximum decay energy. This shift agrees - as it should - reasonably well with the average excitation energy of the daughter molecule. As to our knowledge this is the first time that such fine details have ever been resolved in a β-spectrum.

A first impression of the sensitivity of the present data set to the neutrino mass may be gained from Fig. 4 which displays the linearized, measured spectrum again. The line indicated with $m_\nu = 0$ is a preliminary fit to the data with the neutrino mass fixed to zero. Taking the results of this fit as input parameters of a theoretical integral spectrum, but fixing the neutrino mass to 10 and 20 eV/c^2, instead, we calculate the other spectra shown in the plot. These plots show prima vista that a neutrino mass in the order of 20 eV/c^2 or more is certainly incompatible with the data. From a final evaluation of the present data set including detailed corrections for energy loss etc. we expect to be able to set a limit below 10 eV/c^2. A first result with respect to m_ν we hope to present at the Neutrino '92 conference at Granada.

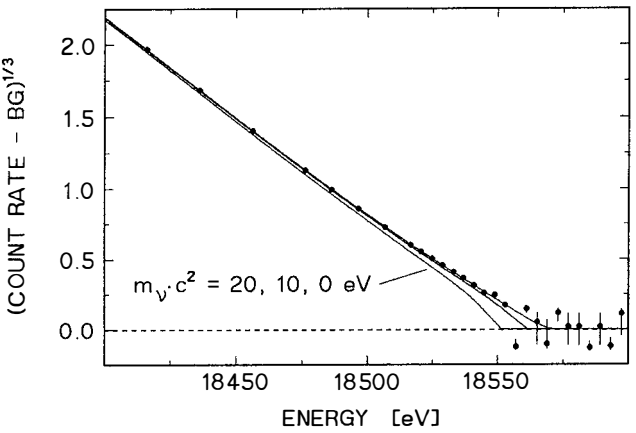


Fig. 4: Same data as in Fig. 3. The solid lines are fitted and calculated spectra, respectively (see text).

References

[1] H. Backe, J. Bonn, Th. Edling, H. Fischer, A. Hermanni, P. Leiderer, Th. Loeken, R.B. Moore, A. Osipowicz, E.W. Otten, A. Picard, Proc. of the VIIth Moriond Workshop on "New and Exotic Phenomena", Les Arcs, France, 1987 (O. Fackler, J. Tran Thanh Van, eds.), Editions Frontières, Gif-sur-Yvette, France, 1987, p. 297

[2] H. Backe, H. Barth, J. Bonn, R. Haid, A. Hermanni, A. Molz, E.W. Otten, A. Picard, M. Przyrembel, M. Schrader, M. Steininger, C. Weinheimer, R.B. Moore, P. Leiderer, A. Osipowicz, Proc. of the XIth Moriond Workshop on "Massive Neutrinos - Tests of Fundamental Symmetries", Les Arcs, France, 1991 (O. Fackler, G. Fontaine, J. Tran Thanh Van, eds.), Editions Frontières, Gif-sur-Yvette, France, 1991, p. 137

[3] M. Przyrembel, H. Fischer, A. Hermanni, E.W. Otten, P. Leiderer, Phys. Lett. A147 (1990) 517

[4] A. Picard, H. Backe, H. Barth, J. Bonn, B. Degen, Th. Edling, R. Haid, A. Hermanni, P. Leiderer, Th. Loeken, A. Molz, R.B. Moore, A. Osipowicz, E.W. Otten, M. Przyrembel, M. Schrader, M. Steininger, Ch. Weinheimer, Nucl. Instr. Meth. B63 (1992) 345

[5] Ch. Weinheimer, M. Schrader, J. Bonn, Th. Loeken, H. Backe, Nucl. Instr. Meth. A311 (1992) 273

[6] S.N. Balashov, A.I. Belesev, A.I. Bleule, E.V. Geraskin, A.A. Golubev, N.A. Golubev, V.V. Ishkin, O.V. Kazachenko, Yu.E. Kuznetsov, V.M. Lobashev, V.I. Parfenov, B.M. Ovchinnikov, I.P. Sekachev, A.P. Solodukhin, P.E. Spivak, N.A. Titov, I.E. Yarikin, Proc. of the Int. Symposium on "Weak and Electromagnetic Interactions in Nuclei (WEIN-89)", Montreal, Que., Canada, 1989, Editions Frontières, Gif-sur-Yvette, France, 1989, p. 295

[7] A. Picard, H. Backe, J. Bonn, B. Degen, R. Haid, A. Hermanni, P. Leiderer, A. Osipowicz, E.W. Otten, M. Przyrembel, M. Schrader, M. Steininger, Ch. Weinheimer, Z. Phys. A (1992), in print

Results from the KARMEN Neutrino Experiment

KARMEN Collaboration represented by

Reinhard Maschuw
Kernforschungszentrum Karlsruhe, IK 1,
D – 7500 Karlsruhe, Germany

Abstract

KARMEN denotes an experimental program of neutrino physics using a pulsed source of neutrinos ν_μ, ν_e and $\bar{\nu}_\mu$ with energies up to 52.8 MeV and a 56 t scintillation calorimeter. Major physics aims are the measurement of charged current (CC) as well as neutral current (NC) neutrino nuclear interactions on ^{12}C with their implications for specific weak couplings, nuclear formfactors and μ-e universality and the search for neutrino oscillations $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$. We present the results of the KARMEN experiment from its first two years of data taking.

KARMEN Collaboration :
G. Drexlin, V. Eberhard, K. Eitel, H. Gemmeke, W. Grandegger, M. Kleifges, J. Kleinfeller, R. Maschuw, P. Plischke, J. Rapp, J. Wochele, J. Wolf, S. Wölfle, B. Zeitnitz;
KfK Karlsruhe and Universität Karlsruhe;
B. Bodmann, F. Burtak, E. Finckh, A. Glombik, T. Hanika, J. Hößl, W. Kretschmer, R. Meyer, F. Schilling; Universität Erlangen - Nürnberg;
J.A. Edgington, T. Gorringe, A. Malik, B. Seligmann; QMWC London
A.C. Dodd; Rutherford Appleton Laboratory
N.E. Booth; Oxford University

1. Introduction

The KARMEN experiment is performed at the neutron spallation facility ISIS of the Rutherford Appleton Laboratory. From the decay of stopped pions produced in the UD_2O 'beam dump' of the pulsed 800 MeV proton beam of ISIS equal numbers of ν_μ, ν_e and $\bar{\nu}_\mu$ are emitted isotropically with energies up to 52.8 MeV according to the decay sequence $\pi^+ \rightarrow \mu^+ + \nu_\mu$; $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$. Because of the different lifetime of π^+ (26 ns) and μ^+ (2.2 μs) a prompt ν_μ - burst within the first 0.5 μs after proton beam-on-target is followed by a $(\nu_e, \bar{\nu}_\mu)$ - pulse in the later time window of 0.5 - 9 μs, where ν_μ are no longer present. The ISIS proton extraction frequency of 50 Hz defines a duty factor of the order of 10^{-4} which allows effective suppression of cosmic background. Neutrinos are detected in a high resolution 56 t liquid scintillation calorimeter [1]. Consisting entirely of hydrocarbons the calorimeter serves as a massive live target of ^{12}C - and ^{1}H - nuclei for the investigation of various ν - induced reactions [2].

2. Neutrino – Nuclear Interactions

The observation of ν - induced transitions between discrete nuclear states is an ideal tool to study in detail the structure of the weak hadronic current as well as to investigate fundamental properties of the neutrino itself. Due to the well defined change of quantum numbers in the nuclear transitions the target nucleus acts as a spin - isospin filter selecting specific parts of the weak hadronic current. In the case of the ν - induced transitions from $^{12}C_{g.s.}$ $(0^+\,0)$ to the excited state $^{12}C^*$ $(1^+\,1; 15.1$ MeV$)$ and the analogue state $^{12}N_{g.s.}$ $(1^+\,1)$, both of which are observed in the KARMEN experiment, only the isovector - axialvector weak current contributes significantly. The inverse β - decay reaction $^{12}C(\nu_e, e^-)\,^{12}N_{g.s.}$ allows to study the q^2 - dependence of the dominant isovector - axialvector formfactor of the weak hadronic current; investigation of the NC transition $^{12}C(\nu, \nu')\,^{12}C^*$ $(1^+\,1; 15.1$ MeV$)$ induced by neutrinos of different flavours, i.e. ν_μ and ν_e or $\bar{\nu}_\mu$, provides a test of the flavour universality of the ν - Z^0 coupling.

3. Exclusive CC reaction $^{12}C(\nu_e, e^-)\,^{12}N_{g.s.}$

Detection of the exclusive charged current reaction $^{12}C(\nu_e, e^-)\,^{12}N_{g.s.}$ is based on a spatially correlated delayed coincidence between an electron from the inverse β - decay on ^{12}C during the ν_e - time window and a positron from the subsequent ^{12}N - decay. This decay, which uniquely identifies ν - induced transitions to the ground state of ^{12}N, is characterized by its lifetime of $\tau = 15.9$ ms with an end point energy of $E_0 = 16.3$ MeV. In addition, electron and positron are produced at the same position in the detector thus greatly reducing random background. The data sample used for the present analysis was taken from Dec. 1989 to Dec. 1991 corresponding to 2267 C protons on target. Software cuts on the time, energy and spatial correlation of the prompt and delayed signal (for details see [3]) selected 112 coincidence events from the data sample. The proof that these ν - candidates are indeed due to exclusive CC reactions is given by fig. 1 a - d): the measured time- and energy distributions of the prompt and delayed signal are in very good agreement with what one expects from the reaction sequence

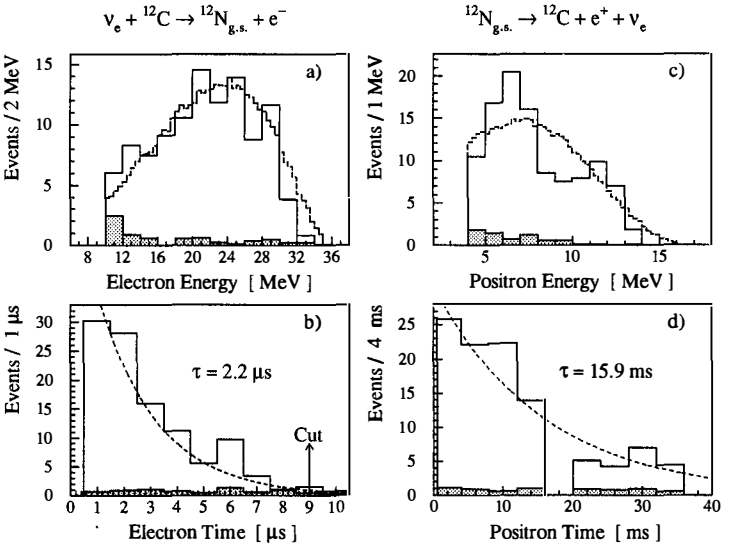


Fig. 1: Energy- and time distributions of the prompt and delayed signals of exclusive $^{12}C(\nu_e, e^-)^{12}N_{g.s.}$ reactions. Energy spectra are compared to MC simulations (broken lines), time distributions are shown with the decay curves of μ^+ ($\tau = 2.2$ µs) and ^{12}N ($\tau = 15.9$ ms) superimposed. The normalized 'beam off' background is shown as shaded area.

$^{12}C(\nu_e, e^-)^{12}N_{g.s.} \rightarrow {}^{12}C + e^+ + \nu_e$. In particular the time distribution of the prompt signal (fig. 1b) clearly demonstrates the origin of these events to be induced by ν_e from μ^+-decay. From 105.2 ± 10.6 of these CC events left after background subtraction (signal-background ratio ≈ 15) the cross section for this exclusive reaction averaged over the Michel shape ν_e energy distribution (0-52.8 MeV) was deduced to be

$$\langle \sigma_{\text{excl. CC}} \rangle^{\text{exp}} = [8.0 \pm 0.8\,(\text{stat.}) \pm 0.75\,(\text{syst.})] \times 10^{-42}\ \text{cm}^2$$

This result is in good agreement with recent theoretical calculations [4, 5, 6] yielding values between $(8.0\text{-}9.4) \times 10^{-42}$ cm^2 with a 10 % uncertainty, as well as with an earlier experimental result from an experiment at LAMPF [7].

The good calorimetric properties of the KARMEN detector also allowed for the first time a measurement of the energy dependence of the $^{12}C(\nu_e, e^-)^{12}N_{g.s.}$ cross section. An unfolding procedure transformed the electron energy spectrum into a ν_e-energy distribution which then was subdivided into four energy bins. The resulting mean cross section in each energy interval is shown in fig. 2 in comparison with theoretical calculations without any further normalization. Where currently the measurement reflects the threshold behaviour of the reaction, improvement of statistics will allow also to investigate the q^2-dependence of the dominant axialvector formfactor.

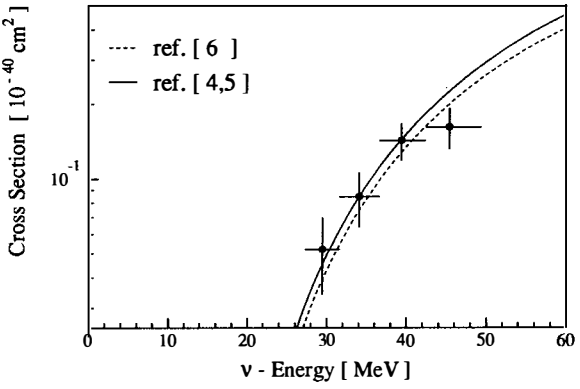


Fig. 2: Energy dependence of the $^{12}C(\nu_e, e^-)^{12}N_{g.s.}$ cross section.

4. NC nuclear excitation $^{12}C(\nu,\nu')^{12}C^*$ $(1^+1; 15.1\ MeV)$

Improved trigger conditions since summer 1990 allowed to identify for the first time the inelastic ν-scattering off ^{12}C-nuclei, i.e. NC events of the type $\nu + {}^{12}C \rightarrow \nu' + {}^{12}C^*$ $(1^+1; 15.1 MeV)$ [8]. The signal for this process is the detection of a localized scintillation event of 15 MeV visible energy from photons emitted as the (1^+1) analogue state of ^{12}C decays back to the ground state with a 94% γ-decay branching ratio. In order to optimize the signal to background ratio evaluation was restricted to the $(\nu_e, \bar{\nu}_\mu)$-time window of 0.5-3.5 μs after beam-on-target. The energy spectrum of events satisfying these criteria with residual background from off beam analysis subtracted is shown in fig. 3a). Above 17 MeV there is a broad distribution of events corresponding to inclusive charged current reactions, with no evidence for events above the kinematic limit. Between about 11 and 16 MeV lies a clearly recognizable peak which is ascribed to the $^{12}C(\nu,\nu')^{12}C^*$ $(1^+1; 15.1\ MeV)$ reaction. The time distribution of all events between 11-16 MeV shows an exponential time slope of 2.2 μs (fig. 3b) above a flat cosmic background, indicating that indeed these events are induced by ν_e and $\bar{\nu}_\mu$ from μ^+-decay. From 73.4 ± 18.6 events in the peak area between 11 and 16 MeV the flux averaged cross section for the sum of ν_e and $\bar{\nu}_\mu$ induced NC transitions was deduced to be

$$\langle \sigma_{NC}(\nu_e + \bar{\nu}_\mu) \rangle^{exp} = [\,9.1 \pm 2.3\,(stat.) \pm 1.3\,(syst.)\,] \times 10^{-42}\ cm^2$$

This result again is in good agreement with theoretical calculations for this reaction [4, 9, 10] giving values ranging from $(9.9\text{-}10.3) \times 10^{-42}$ cm^2.

5. Flavour universality of ν-NC coupling

Except for an isospin factor $1/\sqrt{2}$ the matrixelements of the dominant isovector axial-vector hadronic weak currents for the NC transition $^{12}C(\nu_e, \nu_e')^{12}C^*(1^+1)$ and the CC reaction $^{12}C(\nu_e, e^-)^{12}N_{g.s.}$ are the same [4, 6]. As the NC transition is simultaneously induced by two neutrino flavours of the same intensity, i.e. ν_e and $\bar{\nu}_\mu$, the ratio R of

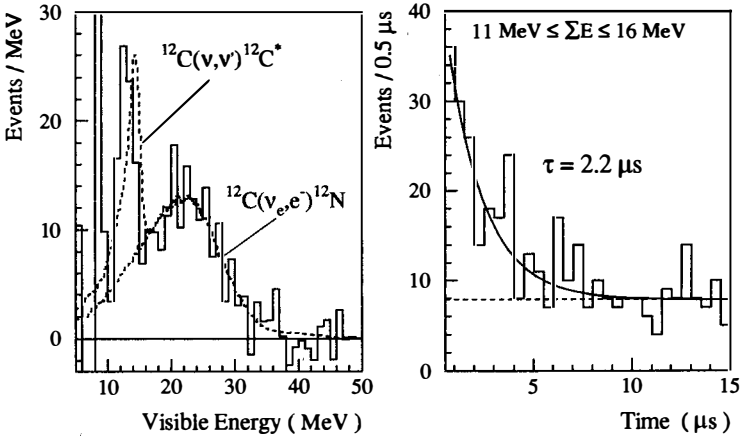


Fig. 3: a) Visible energies of single prong events in the (ν_e , $\bar{\nu}_\mu$) time window with background subtracted. The dotted line corresponds to MC simulations for inclusive CC reactions and the NC excitation of ^{12}C, respectively; b) event time with respect to beam-on-target in the neutral current energy window from 11 - 16 MeV, background not subtracted.

$\sigma_{NC}(\nu_e + \bar{\nu}_\mu)$ / $\sigma_{CC}(\nu_e)$ is about 1, provided the ν_e couples in the same way to the Z^0 as $\bar{\nu}_\mu$. Because of the slightly different energy spectra of ν_e and $\bar{\nu}_\mu$ and also accounting for the small ν - $\bar{\nu}$ difference in the NC cross section the theoretical expectation for this ratio is $R = 1.08$ [4], where from our measurement we get $R = 1.14 \pm 0.34$. This is a flux independent implicit test of the flavour universality in the neutrino NC coupling, which is going to be significantly improved with increasing statistics.

6. CC reaction $^{12}C(\nu_e, e^-)^{12}N^*$

The broad structure in the energy spectrum of single prong events of fig. 3 a) has been assigned to inclusive CC reactions. First of all, the time distribution of these events above 17 MeV again shows the typical 2.2 μs decay constant indicating these events being induced by neutrinos from μ^+ - decay. The kinematic distribution is well reproduced by a MC simulation including small contributions from ν - e scattering and from the $^{13}C(\nu_e, e^-)^{13}N$ reaction, where most of the events are inverse β - decay reactions on ^{12}C with only the electron being detected. The number of transitions to the ^{12}N ground state is well known from our measurement of the exclusive channel with subsequent ^{12}N - decay. The remaining events are thus ascribed to CC transitions to excited states of ^{12}N. From 99 ± 20 events above 17 MeV visible energy, assigned to $^{12}C(\nu_e, e^-)^{12}N^*$ reactions, the flux averaged cross section is determined to be

$$\langle \sigma_{CC\,excit.N} \rangle^{exp} = [\,13.5 \pm 2.8\,(stat.) \pm 1.5\,(syst.)\,] \times 10^{-42}\,cm^2$$

This result contradicts theoretical values of 3.7×10^{-42} cm^2 calculated by Donnelly [5] and of 6.4×10^{-42} cm^2 by Kolbe et al. [12] as well as another experimental but, as we believe, less conclusive result [11] being in agreement with [5]. Currently there is no conclusive

explanation for this deviation, but more elaborate theoretical work seems to be necessary to reliably calculate ν-induced transitions to excited nuclear states also with respect to solar neutrino detectors using nuclear inverse β- decay signatures.

7. Neutrino Oscillations

Having three types of neutrinos ν_μ, ν_e and $\bar{\nu}_\mu$, KARMEN is looking for appearance oscillations $\nu_\mu \dashrightarrow \nu_e$ and $\bar{\nu}_\mu \dashrightarrow \bar{\nu}_e$ simultaneously, whereas the actually present electron neutrinos are used for calibration and normalization purposes. Although the statistics after only two years of data taking is still very low a first preliminary analysis with respect to ν-oscillations was performed. With ν_μ from the decay $\pi^+ \dashrightarrow \mu^+ + \nu_\mu$ being monoenergetic, a search for $\nu_\mu \dashrightarrow \nu_e$ means to look for a 30 MeV ν_e during the short time window of ν_μ using the delayed coincidence signature of the exclusive CC reaction $^{12}C(\nu_e, e^-)^{12}N_{g.s.}$ described above. No event was found where one could have expected 56 events for full oscillation. From ν_e-contamination during ν_μ-time and off beam background one would have expected 0.4 and 0.3 events, respectively. Following the particle data group recommendation this means that more than 2.3 events from $\nu_\mu \dashrightarrow \nu_e$ oscillations can be excluded in the 90% CL. The oscillation probability is thus $P_{\nu_\mu \dashrightarrow \nu_e} < 4.1 \times 10^{-2}$.

Evidence for an oscillation $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ would be the detection of $\bar{\nu}_e$ with energies up to 53 MeV during the $\bar{\nu}_\mu$ time. $\bar{\nu}_e$ not being present in the primary beam will be identified by another delayed coincidence signature, i.e. inverse β-decay on protons followed by (n,γ) capture of the thermalized neutron by a small amount of Gd in the module boundaries of the detector: $^1H(\bar{\nu}_e, e^+)n \rightarrow Gd(n,\gamma)$ with $\Sigma E = 8$ MeV. Four of such events have been found, where 206 would have been expected for full oscillation. With an expected background of 2.2 events this sets an oscillation limit of $P_{\bar{\nu}_\mu \dashrightarrow \bar{\nu}_e} < 2.8 \times 10^{-2}$ at 90% CL. These limits are about an order of magnitude away from the currently published oscillation limits on this channel. After four more years of data taking KARMEN will have proven these limits with great reliability because of its good resolution figures, clear ν-signatures and well understood background conditions.

References

[1] G. Drexlin et al. (KARMEN Coll.), Nucl. Instrum. Methods **A 289** (1990) 490.

[2] G. Drexlin et al. (KARMEN Coll.), in: Neutrino Physics, Proc. of an Intern. Workshop (Heidelberg, Oct. 1987), eds. H.V. Klapdor and B. Povh (Springer, 1988) p. 147.

[3] G. Drexlin et al. (KARMEN Coll.), Phys. Lett. **B**, acc. for publ.

[4] M. Fukugita, Y. Kohyama and K. Kubodera, Phys. Lett. **B 212** (1988) 139.

[5] T.W. Donnelly, Phys. Lett. **B 43** (1973) 93 and private communication (1989).

[6] S.L. Mintz and M. Pourkaviani, Phys. Rev. **C 40** (1989) 2458.

[7] R.C. Allen et al., Phys. Rev. Letters **64** (1990) 1871.

[8] B. Bodmann et al. (KARMEN Coll.), Phys. Lett. **B 267** (1991) 321.

[9] M. Pourkaviani and S.L. Mintz, J. Phys. G: Nucl. Part. Phys. **16** (1990) 569.

[10] J. Bernabéu and P. Pascual, Nucl. Phys. **A 324** (1979) 365.

[11] D.A. Krakauer et al., Los Alamos Preprint LA-UR-91-3909 (1991).

[12] E. Kolbe et al., KFA - Report KFA-IKP(TH)-1991-48 (1991).

NEW EXPERIMENTAL RESULTS ON NEUTRINO MIXING AND DECAY

L. Oberauer, F.v. Feilitzsch, C. Hagner,
G. Kempf and R.L. Mößbauer
Physik Department E15, Technische Universität München
D-8046 Garching b. München
Y. Declais, LAPP Annecy, and E. Kajfasz, CPPM Marseille

The search for neutrino decay is a sensitive method to look for very small neutrino mixing parameter. We report about the status of an decay experiment performed at a reactor in Bugey and present preliminary new experimental limits on the coupling of a heavy neutrino to the electron state. Additionally new experimental lifetime bounds on the radiative decay mode are given. Rigid laboratory limits on this decay mode for the hypothetical 17 keV neutrino are presented. Limits on the radiative decay of a 17 keV neutrino obtained from the supernova SN1987a are discussed.

1. INTRODUCTION

Neutrino decay is, as well as the phenomenon of neutrino oscillations, an inevitable consequence of non zero neutrino masses and neutrino mixing. Experiments looking for neutrino oscillations are sensitive to small neutrino masses, because neutrino oscillations would be a quantum mechanical interference effect. Experiments looking for neutrino decay investigate the region of heavy neutrinos, because the decay widths of all modes increase strongly with the neutrino mass. Therefore the experimental search for neutrino decay is complementary to oscillation experiments.

For the mass region below ≈ 45 GeV the existance of three neutrino flavours has been proved by the recent LEP-experiments at CERN. These neutrino flavours ν_ℓ ($\ell = e, \mu, \tau$) can be expressed in the most general form as a superposition of orthogonal mass eigenstates ν_i, namely as $\nu_\ell = \sum U_{\ell i}\nu_i$. This hypothetical approach is in close analogy to the experimentally verified mixing in the quark sector.

From direct neutrino mass experiments [1] the following limits have been obtained: $m_{\nu(e)} < 9.4$ eV, $m_{\nu(\mu)} < 250$ keV, and $m_{\nu(\tau)} < 35$ MeV. These constraints can be considered as mass limits for the eigenstates ν_i coupled dominantly to the corresponding neutrino flavour. From this experimental facts one can conclude that only two modes for neutrino decay into known particles remain, which are experimentally observable under the preconditions of existing neutrino masses and mixing: $\nu_j \rightarrow \nu_i + photon(s)$ and $\nu_j \rightarrow \nu_i + e^+ + e^-$.

The former is called radiative decay, the latter corresponds to the decay of a heavy ($m_\nu > 2m_e$) neutrino into a positron electron pair.

1.1. NEUTRINO DECAY MODE $\nu_j \rightarrow \nu_i + e^+ + e^-$

We are free to arrange the mass eigenstates according to their mass values. Then we know that only $\nu_{j=3}$ is a candidate for this decay. This mass eigenstate with $m_\nu > 1MeV$ could be dominantly coupled to the tau neutrino ν_τ. Several theoretical models exist showing the possibility of a tau neutrino in the MeV range (e.g. [2]).

If such a heavy neutrino mass eigenstate exists, the decay $\nu_3 \rightarrow \nu_i e^+ e^-$ *must* occur. The transition probability is determined by the strength of the coupling to the electron: namely the mixing parameter U^2_{e3} and U^2_{ei}. This decay occurs at tree level and can be calculated in close analogy to the muon-decay. One obtains $\Gamma_i = (G_F^2 m^5_{\nu_3}/192\pi^3) U^2_{e3} U^2_{ei}$ where m_{ν_3} is the mass of the decaying neutrino and G_F is the Fermi constant. There are two possible decay branches for three neutrino flavours, which have to be added incoherently. Finally we achieve $\Gamma_{total} = \frac{G_F^2 m^5_{\nu_3}}{192\pi^3} U^2_{e3}$, because $U^2_{e1} + U^2_{e2} = 1 - U^2_{e3} \approx 1$.

1.2. RADIATIVE NEUTRINO DECAY

This decay mode occurs at loop level and one obtains in vacuum within a CP - conserving theory for Dirac neutrinos $\Gamma \simeq (m_{\nu_j}/30eV)^5 |\, U_{j\tau} U^*_{i\tau} \,|^2 \cdot 10^{-29}/y$. Here ($m_{\nu_j} >> m_{\nu_i}$). As one can see, the V-A standard model delivers extremely long lifetimes, which can further be suppressed by a weak neutrino-coupling with the tau lepton. However, several models beyond the standard theory predict considerably shorter lifetimes [3]. Additionally radiative decay in dense matter is tremendously enhanced [4]. This may be interesting for supernova models and it also set up some doubts on the exclusion of massive neutrinos with $m_\nu > MeV$ based on cosmological arguments, which did not include this fact.

2. EXPERIMENTAL SET UP AND FIRST RESULTS

We are performing an experiment searching for both decay modes. Our detector is placed in a laboratory near by a nuclear power reactor at Bugey (France). The reactor provides a high flux $\approx 10^{13} cm^{-2} s^{-1}$ of $\bar{\nu_e}$ at our detectors site. Origin of these neutrinos are nuclear beta decays of the fission products. If massive neutrinos ν_j exist which are coupled to these betas, they will be emitted with a branching ratio of U^2_{ej}. Since Q-values in nuclear beta decays up to 10 MeV are obtained,

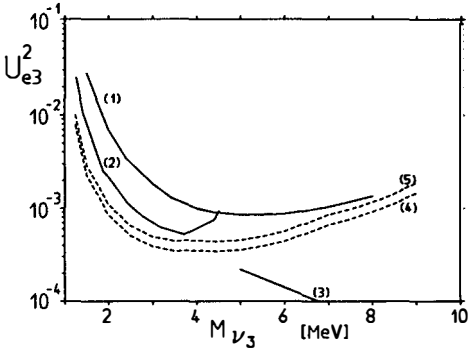


Fig.1: Limits on the neutrino mixing for heavy neutrinos. 1) search for ν decay at the reactor in Gösgen (1987) [5]; 2) at the reactor in Rovno (Ukraine) (1990) [6]; 3) limit from π decay [7]; 4) and 5) (dashed lines) are our new results with 68% and 90% CL.

neutrino masses in this range can be studied in our experiment.

The detector consists of two position sensitive multiwire XY-chambers with dimensions of $1.9m \cdot 1.8m$. Both chambers (36μ entrance window) are looking for possible neutrino decay products, namely electrons and low energy photons, emerging from a $8m^3$ big He-filled volume. This system is surrounded completely by six $2m \cdot 2m$ proportional counters, acting as an active veto shield. Copper as well as aluminium sheets are used as passive shielding material. The whole detector can be rotated by 360^0 in order to measure in and out of the neutrino direction with respect to the reactor.

2.1. RESULTS FOR THE $\nu_3 \rightarrow \nu_i + e^+ + e^-$ MODE

The measurements for this decay mode started in June 1991 during the reactor off time. This time ($\approx$ 10 days) is used as a background measurement. During the following reactor on period we took $\approx$ 20 days of measurement. A fast trigger was set when both XY-chambers fired simultaneously and in anticoincidence to the whole veto system. But only if two clusters of adjacent channels in each chamber were identified, the pattern of the channels, the energies and the pulse shape of the event were written on tape. The average trigger rate for the reactor off (on) period was 1.88 ± 0.03 (1.88 ± 0.02) counts/sec (68% CL). Here statistical as well as systematic fluctuations are included. So, no significant increase in the trigger rate during the reactor on period was observed. Taking just the trigger rate difference between reactor on minus reactor off we obtain *preliminary* limits on the mixing parameter U^2_{e3} as function of the neutrino mass. Effects due to the limited efficiency and the deadtime of the detector are included. Our new limits, derived from the trigger rate only, are depicted in fig.1 together with results from other experiments.

In the next time we will analyse the data written on tape. Only events with energies within a proper window will survive as good candidates. Additionally the pattern of the channels which fired, together with the pulse shape information will allow to reconstruct the vertex of the two fold coincidence. During spring and summer 1992 a new measurement with a reactor off period of three months is planned.

2.2 RESULTS OF THE $\nu_j \rightarrow \nu_i + \gamma$ MODE

Recently the existance of a 17 keV neutrino ν_j, weakly coupled to the electron ($U^2_{ej} = 8 \cdot 10^{-3}$) was claimed [8]. If this massive neutrino exists, it couples of course also to reactor neutrinos and we will have a flux $\Phi(\nu_{17keV}) \approx 8 \cdot 10^{10}/cm^2 sec$. Due to their mixing these neutrinos *must* decay

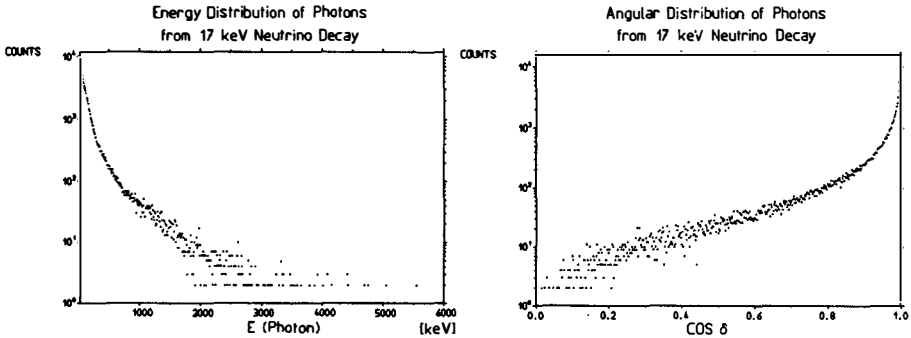


Fig.2: Energy and angular distribution of photons following the radiative decay of 17 keV neutrinos emitted by a reactor.

radiatively.

We were looking for this decay during the reactor on time. In a Monte-Carlo calculation we estimated the photon energy distribution one would expect. This spectrum is increasing for low energies as depicted in fig.2. The neutrino luminosity of a reactor is determined experimentally [9], only for the low energy part ($E < 1MeV$) we rely on the calculations of Vogel and Engel [10]. The integral over the energy window yields the counting rate one expects in the detector. It is important to notice that this number can depend very strongly on the helicity of the neutrino ν_i, which follows the radiative decay.

We set a energy window between 2 and 10 keV. For the decay photons with an energy inside this window the Helium decay volume is transparent, whereas the quanta are absorbed in the XY-chambers completely. Therefore we obtained nearly 100% detection efficiency. Due to the fact that the most of the 17 keV neutrinos are highly relativistic, the direction of the photon follows almost exactly the original direction of the decaying particle (fig.2). Therefore it is possible to determine the background just by rotating the whole detector.

No effect was observed and after one day of measurement we were able to set a 90% CL limit of 0.58 counts/sec.

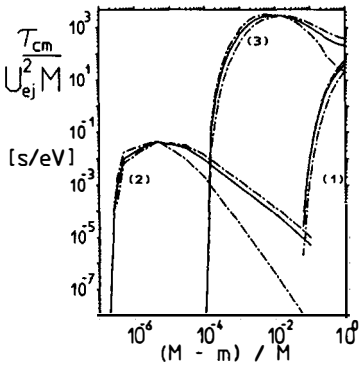


Fig.3: Limits on the neutrino radiative lifetime obtained in laboratory experiments: (1) corresponds to ref.[5], (2) to ref.[11], (3) are our new results.

Comparing the expected counting rate with the experimental bound we obtain the following 90% CL limits on the radiative lifetime (in the cm-system) of the 17 keV neutrino:

- no neutrino helicity flip: $\tau_{cm} > 3.1 \cdot 10^3\ sec$
- neutrino helicity flip: $\tau_{cm} > 5.3 \cdot 10^4\ sec$.

Note, that standard weak current interaction predicts the appearance of a neutrino helicity flip in radiative decay. To our knowledge our limits are the most stringent bounds obtained up to now in laboratory experiments.

One can also calculate the most general case of arbitrary neutrino masses. Then we get limits on the quantity $\tau_{cm}/U_{ej}^2 \cdot M$, with M the mass of the decaying neutrino as a function of $(M-m)/M$, where m describes the mass of the emitted neutrino state. This limits are depicted in fig.3 together with bounds obtained in other laboratory experiments. The dashed lines indicate the limits valid for neutrino helicity flip (upper curve at the right side) and for helicity conservation.

3. LIMITS ON THE RADIATIVE DECAY OF A 17 KEV NEUTRINO FROM SN1987a

The energy density of massive neutrinos and their decay products must not exceed the present energy density of the universe. This requirement excludes the possibility for stable neutrinos with masses over ca. 100 eV. For neutrinos with mass values above this limit one obtains bounds on the lifetime as a function of the neutrino mass [12]. For a 17 keV neutrino one estimates a maximum lifetime $\tau_{\nu,max} = 10^{12} - 10^{13}\ sec$. In principle radiative decay can solve this problem, however from standard weak interaction one obtains a lifetime which exceeds $\tau_{\nu,max}$ by many orders of magnitude. For neutrinos above 1 MeV this problem does not occure, because then the e^+e^- decay mode is open and is fast enough to meet the conditions from big-bang.

Although radiative lifetime is shortened in dense matter of the early universe, this effect is to small for a 17 keV neutrino to fulfill the limit given above. Other possible decay modes in known particles, like the decay into two photons and into three light neutrinos, yield even longer lifetimes. So, from a theoretical point of view only the introduction of unusual large magnetic transition moments leading to fast radiative decay or the invention of new light particles (Goldstone boson etc.) in which neutrinos may decay can "rescue" the 17 keV neutrino.

How is the experimental sight on radiative neutrino decay in context with cosmology? Laboratory experiments like ours can not explore the region up to 10^{13} sec. The most stringent limits [13] come from the non-observation of gamma rays in the detector of the Solar-Maximum-Mission satellite during the neutrino burst in IMB and Kamiokande. We reviseted now also the period (254 sec) of measurement following the neutrino burst in the 4.1-6.4 MeV energy band. Comparing with the background measured shortly before we obtained a limit on the gamma fluence coming from radiative neutrino decay and therefore also for the neutrino lifetime (fig.4).

How reliable are these limits? We remind, that IMB and Kamiokande detected electron antineutrinos. The lifetime limits for the electron antineutrino given in [13] are not model dependent, because they only base on the measured gamma and neutrino fluence and on the known distance from the large magellanic cloud to earth. However, the fluence of muon - and tau neutrinos reaching earth has to be estimated in models describing supernova mechanism (we rely on calculations of Wilson [14]). Therefore our limits on muon - and tau neutrinos, which are the only interesting when discussing the 17 keV neutrino, are indeed model dependent.

Additionally, the region one excludes is still below these values which are calculated within standard theory. This means, that one can describe the decay by large transition magnetic moments, leading to other may be difficult to describe phenomena inside the neutrino sphere of SN1987a. In a paper of Hatsuda et al. [15] the effect of radiationless conversion of massive neutrinos was studied. Within a certain range of the transition magnetic moment this effect will occure without reconversion, leading to a strong depletion of massive neutrinos. Then no decay of this neutrinos could be observed.

In fig.4 the exclusion plot of our work, including the arguments of [15] are shown. For a 17 keV neutrino the regions

$$10^6\ sec < \tau_{17keV} < 7 \cdot 10^7\ sec$$

$$\tau_{17keV} > 2 \cdot 10^{14}\ sec$$

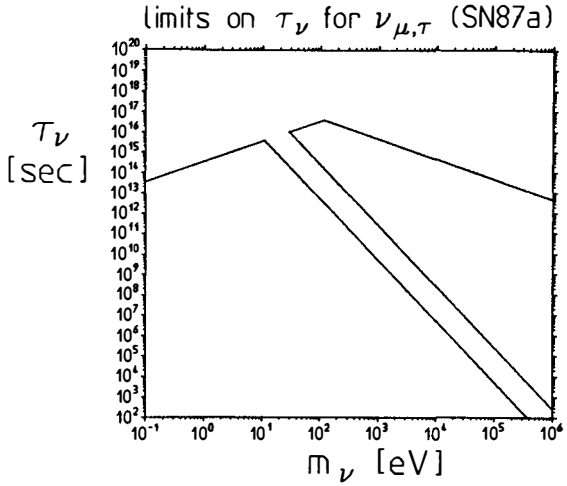


Fig.4: Limits on the neutrino radiative lifetime for massive ν_μ and ν_τ obtained from SMM data during SN1987a. Radiationless conversion of massive neutrinos inside the supernova can lead to a depletion of heavy neutrinos and no decay can be observed within a certain range of parameter.

are allowed. We would like to thank Mr. Rieger from the Max Planck Institut für Extraterrestische Physik in Garching for the access of the SMM data. This experiment is supported by the BMFT of the Federal Republic of Germany.

[1] J.F. Wilkerson et al., Nuclear Physics B (Proc. Suppl.) (1991) 19; R. Abela et al., Phys. Lett. 146B, (1984) 431; H. Albrecht et al. (ARGUS), Phys. Lett. 202B, (1988) 149
[2] C.T. Hill and E.A. Paschos, Phys. Lett. 241B, (1990) 96
[3] M. Roos, Proc. of an Int. Workshop "Neutrino Physics", Heidelberg Germany, ed. by H.V. Klapdor and B. Povh, (1987) 57
[4] T.K. Huo and J. Pantaleone, Phys. Lett. 246B, (1990) 144; J.C. D'Olivo et al., Phys. Rev. Lett. 64, (1990) 1088; C. Giunti et al., Phys. Rev. 43D, (1991) 164
[5] L. Oberauer, F. von Feilitzsch and R.L. Mössbauer, Phys. Lett. 198B, (1987) 113
[6] S.A. Fayans, V.I. Kopeykin and L.A. Mikaelyan, in L. Moscoso: "Experimental review on neutrino oscillations and decay", Nuclear Physics B (Proc. Suppl.) (1991) 19
[7] D.A. Bryman et al., Phys. Rev. Lett. 50, (1982) 1546
[8] A. Hime and N.A. Jelley, Phys. Lett. 257B, (1991) 441; J.J. Simpson and A. Hime, Phys. Rev. 39D, (1989) 1825
[9] K. Schreckenbach et al., Phys. Lett. 160B, (1985) 325; F. von Feilitzsch, A.A. Hahn and K. Schreckenbach, Phys. Lett. 118B, (1982) 162
[10] P. Vogel and J. Engel, Phys. Rev. 39D, (1989) 3378
[11] J. Bouchez et al., Phys. Lett. 207B, (1988) 217
[12] H. Harari and Y. Nir, Nucl. Phys. 292B, (1987) 251
[13] F. von Feilitzsch, L. Oberauer, Phys. Lett. 200B, (1988) 580; E.W. Kolb and M.S. Turner, Phys. Rev. Lett. 62, (1989) 509
[14] J.R. Wilson et al., Ann. N.Y. Acad. Sci. 470, (1986) 267
[15] T. Hatsuda, C.S. Lim and M. Yoshimura, Phys. Lett. 203B, (1988) 462

BUGEY III neutrino oscillation experiment
Some preliminary results

E. KAJFASZ

CPPM
163 Avenue de Luminy
Case 907
13288 MARSEILLE Cedex 9
FRANCE

Abstract

A preliminary analysis of data taken with one of our detectors at 15 and 95 m of a 2800 MW nuclear reactor core is discussed. The corresponding exclusion contour is presented. Schedule and limits we intend to reach for a 40 m measurement are also presented.

1. INTRODUCTION.

The Bugey neutrino oscillation experiment is performed near the four reactors of the Bugey Nuclear Power Plant, 30 km east from Lyon (France). Each reactor has a thermal power of 2800 MW and yields a flux of about 5.4 10^{20} antineutrinos/s. The reactor cores are 4 meter high and 3 meters in diameter. Three detectors are used as sketched in figure 1. Each one is a stainless steel tank filled with 600 liters of liquid scintillator. This pseudocumene based scintillator (NE 320) has a H/C ratio of 1.4 and is loaded with ^{6}Li (0.15%) [ref 1]. Optical tunnels define 98 cells of 8.2x8.2x85 cm^3 read by a photomultiplier (PM) at each end (figure 2). The detector(s) is (are) surrounded by a shield which, from outside inwards, is made of a 18 cm layer of lead (to reduce ambiant γ background), a 25 cm layer of water followed by a layer of B_4C (to reduce neutron background) and vessels filled with liquid scintillator read by a PM (to keep track of cosmic charged particles) used as veto counters [ref 2].

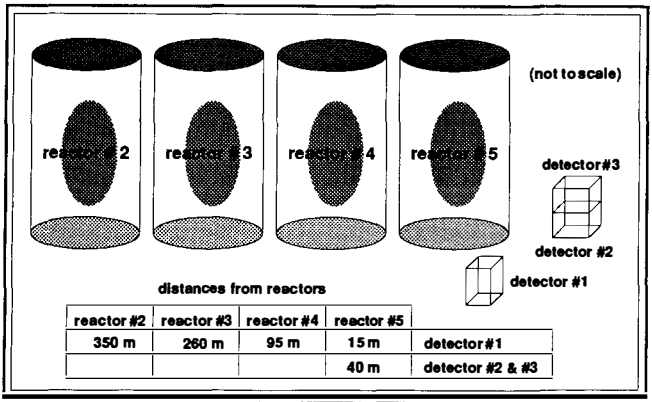


Figure 1. Set up

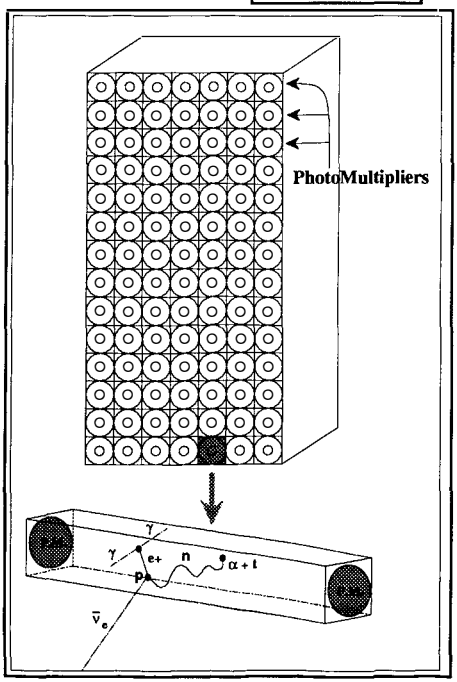


Figure 2. A detector

2. DETECTOR PROPERTIES.

The scintillator is used as a target for the antineutrinos; it detects the positron, thermalizes the neutron and detects the neutron via $^6Li + n_{th} \rightarrow \alpha + t$ giving rise to a sharp line of 530 keV electron equivalent energy. The mean life time of the neutrons coming from antineutrinos is about 30 µs in our detectors and the mean distance between the positron and the neutron detection points is roughly 5 cm. The localization of the antineutrino events combined to the short correlation time between

the positron and the neutron allows us to reduce our accidental background. Discrimination between different particle types is made using pulse shape discrimination (PSD) (figures 3 and 4) [ref 1 and 2].

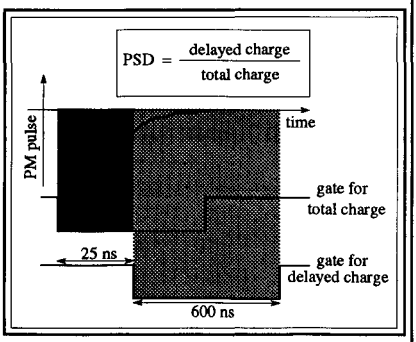


Figure 3. PSD definition

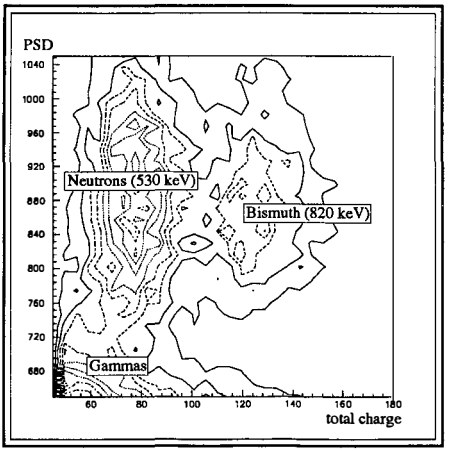


Figure 4. Heavy particle PSD vs total charge of a Bismuth enriched sample of events

3. ON-LINE ANTINEUTRINO SIGNATURE.

We require a 100 μs delayed coincidence between a "neutron" pulse and a "positron" pulse in neighboring cells; a "neutron" pulse following the "positron" pulse defines the on-line "neutrino" trigger,whereas a "neutron" pulse followed by a "positron" pulse defines the on-line "accidental" trigger.

4. CALIBRATIONS.

An Am/Be source emits fast neutrons and 4.4 MeV gammas, with which we carry out:

- a neutron calibration of each cell every month,
- a gamma calibration of each cell every other week.

A spark gap connected to a bunch of 2x98 optical fibers (one optical fiber for each PM) is used to:

- daily monitor the changes in the PM gains and the liquid properties, allowing us to daily interpolate the source calibrations.

In this way we manage to daily monitor for each cell:

- the neutron energy peak position,
- the neutron PSD,
- the positron PSD,
- the energy calibration (to better than 1%).

5. STATUS OF THE LIQUID SCINTILLATORS.

In detector # 2 & 3, the liquids are stable but contaminated with Uranium to a level of 1 ppb [ref 3]. For detector # 1,the liquid has no Uranium contamination problem but is slowly degrading (2 to 4 % per month on the number of photoelectrons/MeV); however due to our daily monitoring it is possible to correct for this degradation.

6. AVAILABLE DATA FOR DETECTOR # 1.

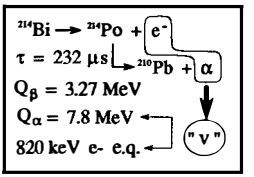
- ON data — 15 m (reactor #5) — > 100 000 antineutrinos
  95 m (reactor #4) — 1800 antineutrinos
- OFF data

run	reactor#5 (15m)	reactor #4 (95m)	reactor #3 (260m)	# days
POF1	OFF	ON	ON	47
POF2	OFF	OFF	ON	6
POF3	OFF	OFF	OFF	40

7. BACKGROUND.

The background is measured and substracted using reactor OFF data. There are mainly two types of identified background:

- accidentals: they are measured on-line using a special trigger "neutron" before "positron",
- correlated, coming from:
  - fast neutrons from cosmic ray μ stops (rejected by veto counters and positronPSD)
  - Bismuth correlated events (rejected by software cuts)



8. PRELIMINARY ANALYSIS OF DETECTOR #1 DATA.

8.a. Analysis cuts.

The analysis cuts used were the following:

- positron energy: E(e+) > 1 MeV,
- positron PSD: PSD(e+) within +- 3 σ of the mean,
- 2.2 σ elliptical cut on (neutron PSD: PSD(n), neutron energy: E(n))
- time difference between neutron and positron: Δt(n,e+) < 60 μs,
- longitudinal distance between neutron and positron: Δz(n,e+) < 35 cm,
- multi-dimensional cut [ref 6].

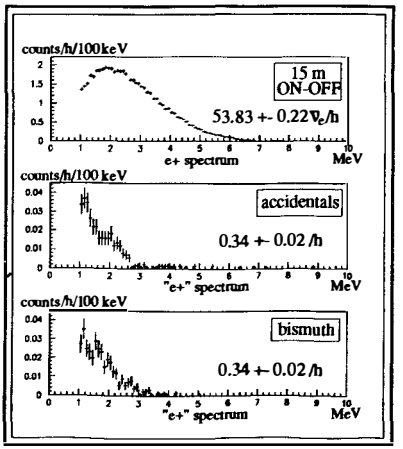


Figure 5. Positron spectra for 15 m data

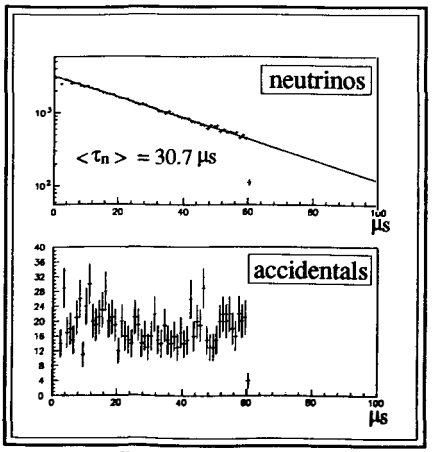


Figure 6. Time distributions

8.b. 15 m data.

These data were taken while reactor #5 was ON. Figure 5 shows the energy spectra of neutrino, accidental and Bismuth events. The number of detected antineutrinos is found to be 53.83 +- 0.22 /h after cuts (§ 8.a). Figure 6 shows the Δt(n,e+) distribution for neutrino and accidental events. Figure 7 shows the very good agreement between the distribution of the number of neutrinos detected along the 14 lines of the detector and the same distribution generated with our Monte-Carlo. The first and last lines are underpopulated due to neutron leaks; the decrease between top and bottom is due to a solid angle effect.

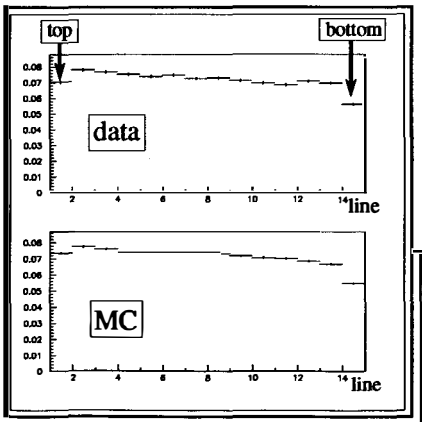


Figure 7. Neutrino events along lines

8.c. 95 m data.

These data were taken while reactor #5 was OFF and reactor #4 was ON. The number of detected antineutrinos is found to be 1.23 +- 0.13 /h after cuts (§ 8.a). Figure 8 shows the energy spectrum of neutrino events and the ratio between the energy spectra of neutrinos coming from 15m and those coming from 95m. This ratio is in good agreement with the difference $\Delta\Omega$ in solid angle between these two distances. Figure 9 shows the 90 % C.L. contour extracted from the comparison of the 95 m and the 15 m data, taking into account a conservative 10 % systematic error.

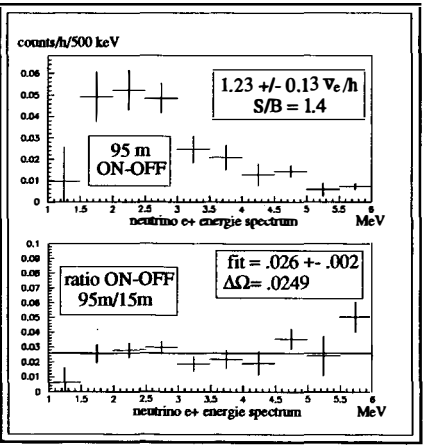


Figure 8. Data for 95 m

8.d. Background data.

The background components, using the 6 days of run POF2 OFF data are the following:

- correlated	Bismuth (estimated)	0.26 (+- 0.05) /h
	ν's from reactor #3	0.15 /h
	remaining background	0.42 (+- 0.12) /h
	total	0.83 +- 0.11 /h
- accidentals		0.25 +- 0.05 /h

9. CONCLUSIONS.

9.a. detector #1.

The data taking is now over. The analysis needs to be finalized including:

- cuts optimization,
- even better background rejection:
  - better estimate of Bismuth contamination,
  - use interpolated neutron calibrations,
  - use shape of "neutron" pulses as measured by 100 MHz Flash Adcs [ref 4],
- lower the systematics (normalization, cuts efficency ...).

For the 15m data, as we have a high statistics, it will be possible to discriminate between Schreckenbach and Tengblad measured spectra [ref 5]. For 95m data a slightly better contour should come out.

9.b. detector #2 and 3 (40m data).

Data have been taken from June 91 on. Data acquisition should continue till summer 92, including 3 months of "OFF data" from April to June. The expected 90 % C.L. contour for the comparison between these data and the 15m data is shown in Figure 9. A final analysis, combining the source spectrum (0m) and the 15, 40, 95 m data should be performed before the end of 92.

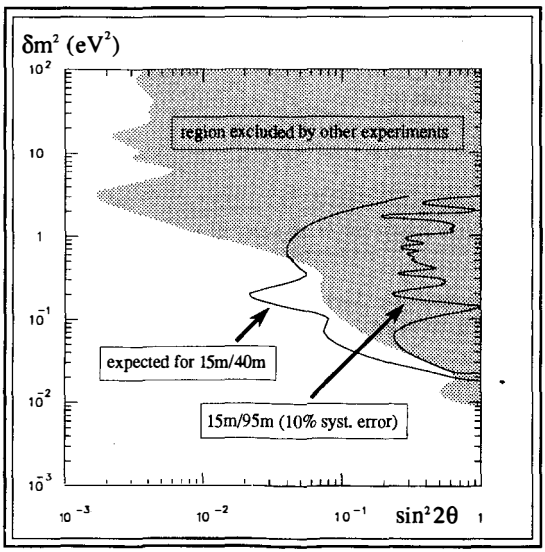


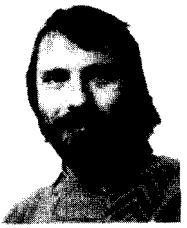
Figure 9. Contours (90 % C.L.)

10. REFERENCES

[1] S. Ait-Boubker et al. NIM A277 (1989) 461.
[2] J. Collot in Proc. of the 23rd Rencontre de Moriond - Moriond Workshops - (Editions Frontières 1988) 153.
[3] M. Avenier in Proc. of the 26th Rencontre de Moriond - Moriond Workshops - (Editions Frontières 1991) 111.
[4] R. Aleksan et al. NIM A273 (1988) 303.
[5] K. Schrenckenbach et al. Phys. Letters 160B (1985) 325.
O. Tengblad et al. Nucl. Phys. A503 (1989) 136.
[6] J.P. Cussonneau - Thèse de doctorat - April 1992 - LPC - Collège de France - Paris (France)

RECENT ARGUS RESULTS ON TAU PHYSICS

Wolfgang Funk
Institut für Hochenergiephysik, Universität Heidelberg *



Abstract

Some recent ARGUS results on τ physics are presented †: The decay $\tau^- \to \nu_\tau \pi^- \pi^+ \pi^-$ is shown to be dominated by a_1 production. Using this decay, we have updated our determination of the ν_τ-helicity. Furthermore, we have tested CVC by exploiting the τ decay into $\pi^- \pi^0$ which is dominated by ρ production. These decays, together with the leptonic ones, indicate that the consistency and the completeness problems in τ decays are still alive. A search for 29 ν-less τ decay channels doesn't give an explanation for these problems. Finally, we have updated our limit on the τ - neutrino mass.

*supported by the German Bundesministerium für Forschung und Technologie under the contract number 055HD21P

†References to a specific charged state are to be interpreted as also implying the charge conjugate state.

The Decay $\tau^- \rightarrow \nu_\tau \pi^- \pi^+ \pi^-$

To investigate the decay τ into 3 charged pions, we have selected events with the so called "1 versus 3 topology". After background subtraction one gets the three pion mass distribution shown in figure 1 which is nicely described by the model of Isgur et.al. [2]. Therefore, and due to the fact that one observes a clean ρ - signal in the two pion mass spectrum (figure 2), one concludes that a_1 production dominates this channel. The background is modelled by the "like sign" two pion mass distribution (full line). Using the model of Isgur et al. we have determined mass and width of the a_1 ($m(a_1) = (1.211 \pm 0.007)\mathrm{GeV/c^2}$, $\Gamma(a_1) = (0.466 \pm 0.021)\mathrm{GeV/c^2}$).

For the decay τ into a_1 we get a branching ratio of $BR = (7.06 \pm 0.10 \pm 0.50)\%$. Adding the $K\pi\pi$ and $KK\pi$ branching ratios to our value, we can compare our result with those of ALEPH [3], CELLO [4], and other experiments [5], and calculate a world average of $(7.8 \pm 0.3)\%$.

The ALEPH [3] and CELLO [4] collaborations determined slightly higher values as well for this channel as for τ goes to $\pi^-\pi^0\pi^0$. When a_1 dominance is assumed both channels are expected to have the same branching ratio because of isospin arguments. Concerning the completeness problem, these two channels cause the major discrepancy between ALEPH and CELLO on one, and ARGUS and the rest of the world on the other hand, which I want to discuss below.

The decay τ into a_1 provides the possibility to measure the sign of the ν_τ - helicity due to the fact that there are two possibilities to form the ρ. This leads to an interference term containing the information on the sign. In 1990, ARGUS has published a value for $\gamma_{AV} = \frac{2g_A \cdot g_V}{g_A^2 + g_V^2}$ of $1.14 \pm 0.34 \pm^{0.34}_{0.17}$ [7], where $g_{A,V}$ is the weak axial and vector coupling constant. We have meanwhile improved our statistics, and have determined the ratio of the D and S wave amplitudes in the decay $a_1 \rightarrow \rho^0\pi^-$ at the mass of the a_1 by analyzing the Dalitz plot projections, and obtained for this ratio a value of -0.11 ± 0.02. With these improvements we get a new value for γ_{AV} of $1.25 \pm 0.23 \pm^{+0.15}_{-0.08}$, where the Standard Model predicts 1. This is a clear proof for the lefthandedness of the ν_τ.

The decay $\tau^- \rightarrow \nu_\tau \pi^- \pi^0$

Selecting τ - pairs where both $\tau's$ decay into ρ, we are able to extract a branching ratio for this channel independent of other ones. In addition we look for $\rho's$ from "1 versus 3 topologies". The result of $(22.6 \pm 0.4 \pm 0.9)\%$, together with those of ALEPH [3], CRISTAL BALL [6], CELLO [4], and other experiments [5] leads to a world average of $BR(\tau \rightarrow \nu_\tau \pi^- \pi^0) = (23.1 \pm 0.6)\%$. This channel can be used to perform a test of the CVC hypothesis which means relating the decay τ into ρ to e^+e^- annihilation at low energies. The CVC prediction with the e^+e^- data from Novosibirsk [8] gives a value

$\frac{Br(\tau\to\nu_\tau\pi^-\pi^0)}{Br(\tau\to\nu_\tau e^-\bar{\nu}_e)} = 1.32 \pm 0.05$, which agrees perfectly with our measurement of 1.31 ± 0.07. Futhermore, we have applied the CVC prediction to the $\pi^-\pi^0$ mass spectrum, where the spectral function $v_1(q^2)$ of the ρ is proportional to $|F_\pi|^2$ for the $(I = 1)$ component. Using Kühn's + Santamaria's fit of F_π^2 to the e^+e^- data [8] we get a good description for the background subtracted spectrum (figure 3). Acceptance and resolution are taken care of in the fit function. This proves that CVC works well and we have dominant ρ-production in this channel.

Leptonic τ decays

For the determination of the leptonic branching ratios we have used the following two data samples:

$e^+e^- \to \tau^+\tau^- \to (3prong)\ \ \nu_\tau\ \ e^-\bar{\nu}_e(\mu^-\bar{\nu}_\mu, h^-)$

$e^+e^- \to \tau^+\tau^- \to \nu_\tau e^-\bar{\nu}_e\ \ \bar{\nu}_\tau\mu^+\nu_\mu$,

leading to the product of branching ratios: $B_e \cdot B_{3prong}$, $B_\mu \cdot B_{3prong}$, $B_h \cdot B_{3prong}$, $B_e \cdot B_\mu$. Calculating ratios of the product $BR's$ one is able to determine B_e and B_μ independently of B_{3prong}. We obtain $B_e = (17.3 \pm 0.4 \pm 0.5)\%$ and $B_\mu = (17.2 \pm 0.4 \pm 0.5)\%$ giving a world average of $(17.8 \pm 0.2)\%$ and $(17.4 \pm 0.2)\%$ for e and μ, respectively, when using all available results (LEP collaborations, CLEO, CELLO and others). There is impressive agreement of all experiments.

Consistency problem

The consistency problem states that the exclusively measured BR into electrons does not equal that one derived from the τ - lifetime:

$$B_e^{excl.} \neq \frac{\tau_\tau}{\tau_\mu} \cdot \left(\frac{m_\tau}{m_\mu}\right)^5 \cdot \left(\frac{G_\tau}{G_\mu}\right)^2 \qquad (1)$$

With the exclusively measured $BR's$ for the decays:

$\tau^- \to \nu_\tau e^-\bar{\nu}_e$, $\tau^- \to \nu_\tau\mu^-\bar{\nu}_\mu$, $\tau^- \to \nu_\tau(\pi^- + K^-)$, $\tau^- \to \nu_\tau\rho^-$, $\tau^- \to \nu_\tau\dot{K}^{*-}$, $\tau^- \to \nu_\tau\pi^-\pi^+\pi^-\pi^0$, $\tau^- \to \nu_\tau\pi^-\pi^0\pi^0\pi^0$ one can, with the help of theory, calculate B_e for each channel. For example, the CVC prediction for the ρ channel was shown above. Averaging all $B_e's$ coming from exclusive τ decay studies, one gets a mean value of $(17.91 \pm 0.13)\%$. Secondly, one can deduce B_e by comparing the decay τ into e with the decay μ into e (equation(1)). Using the latest world average for the τ lifetime $((302.7 \pm 5.6)\ fs)$, including the results of CLEO, the four LEP experiments, and the PDG value, one obtains $(18.91 \pm 0.36)\%$. Thus, there is an apparent discrepancy of 2.6 standard deviations between B_e from exclusive channels and B_e from τ lifetime and τ mass. The

latter one is, by the way, dominated by DELCO measurement of 1978 [9].
Possible solutions of the consistency problem could be: A wrong B_e from exclusive measurements. There are, however, different exclusive channels which agree very well with each other. Next, there could be a mistake in the τ lifetime or mass determinations. A third solution of the consistency problem would be the existence of a light and a heavy τ -neutrino which mix. Finally, violation of lepton universality would explain the consistency problem.

Completeness problem

The second problem in τ physics is the completeness problem which means that the sum of all exclusive branching ratios found is not 100%. Table 1 summarizes the branching ratios for all channels: theory, world average, and the ARGUS and ALEPH values [3]. The theoretical predictions are normalized to the measured electronic branching ratio. Since ARGUS has not measured the decay into $\pi^-\pi^0\pi^0$, $BR(\tau^- \to \nu_\tau\pi^-\pi^+\pi^-) = BR(\tau^- \to \nu_\tau\pi^-\pi^0\pi^0)$ was used assuming isospin invariance and a_1 dominance in this channel. The theoretical prediction based on CVC and hadron production in e^+e^- annihilation is used for $BR(\tau^- \to \nu_\tau\pi^-\pi^0\pi^0\pi^0)$. In the ALEPH column, for $BR(\tau^- \to \nu_\tau K^{*-})$ the world average was inserted. Both the world averages and the ARGUS measurements indicate that the completeness problem is still alive, whereas ALEPH and also CELLO [4] don't see any problem. The major differences between ALEPH and ARGUS are the channels into three charged pions and into one charged and two neutral pions as discussed before. The differences in analysis between ALEPH and CELLO on one hand, and ARGUS and the rest of the world on the other hand, is that ALEPH and CELLO determine branching ratios for all channels in a simultanous fit, without, however, constraining the sum of all branching ratios to 1, due to differing efficiencies. ARGUS determines the branching ratios independently channel by channel. The reason for the experimental discrepancies is not yet clear, but is most likely connected to the method applied.

Search for neutrinoless τ decays

Taking the completeness problem for real, one explanation would be the existence of unexpected decay channels, for instance neutrinoless τ decays. ARGUS has performed a search for 29 such channels with a data sample corresponding to 387 pb^{-1}. Table 2 summarizes the upper limits of ARGUS and those obtained previously by other experiments. We don't see the solution of the completeness problem here.

τ decay mode	Theory [%]	World average %	ARGUS [%]	ALEPH [%]
$e\nu\bar{\nu}$	17.8	17.8 ± 0.2	17.3 ± 0.6	18.1 ± 0.6
$\mu\nu\bar{\nu}$	17.3	17.4 ± 0.2	17.2 ± 0.6	17.4 ± 0.5
$(\pi + K)\nu$	11.4	12.0 ± 0.3	11.7 ± 1.0	12.6 ± 0.6
$\rho\nu$	23.5	23.1 ± 0.6	22.6 ± 1.0	24.6 ± 1.1
$K^*\nu$	1.1	1.4 ± 0.1	1.2 ± 0.3	(1.4)
$"\pi^+\pi^-\pi^-"\nu$	(11.7)	7.8 ± 0.3	7.4 ± 0.5	9.1 ± 0.7
$"\pi^-2\pi^0"\nu$	(11.7)	8.2 ± 0.5	(7.4 ± 0.5)	10.2 ± 1.1
$"\pi^+\pi^-\pi^-\pi^0"\nu$	4.3	5.3 ± 0.3	5.4 ± 0.6	5.0 ± 0.7
$"\pi^-3\pi^0"\nu$	1.1	1.3 ± 0.4	(1.1)	1.6 ± 0.6
Σ	100	94.2 ± 1.1	91.3 ± 3.3	100.0 ± 2.2

Table 1: Branching ratios for major τ decays (see text)

Measurement of m_{ν_τ}

In 1985, ARGUS has published an upper limit of 70 MeV/c^2 on the τ - neutrino mass by investigating the energy spectrum of the three pion mode [14]. With an integrated luminosity of 197 pb^{-1} we have published in 1988 an improved upper limit of 35 MeV/c^2 at 95% *C.L.* using the invariant mass distribution in the 5π channel [15]. With all data available in 1991 (387 pb^{-1}) we have redone this analysis and found the same upper limit (figure 4). This is due to the fact that there is no change at higher values in the 5π invariant mass spectrum which is the sensitve region for determining the ν_τ mass.

Summary

The decay τ into three charged pions is dominated by a_1 production. By analyzing this decay we are able to prove that the ν_τ is lefthanded. Furthermore, we have tested the CVC hypothesis by exploiting the decay τ into $\pi^-\pi^0$ which is dominated by the ρ. New ARGUS values for B_e and B_μ have been presented. The recent measurement of exclusive branching ratios by ARGUS keeps the consistency and completeness problems in τ decays alive. We have also measured 29 upper limits for ν -less τ decays and updated the determination of the ν_τ -mass, but the limit has not improved.

Upper limits $[10^{-5}]$ at 90% CL					
decay channel	MARK II	ARGUS 86	Crystal Ball	CLEO	ARGUS 91
$\tau^- \to e^-e^+e^-$	40	3.8		2.7	1.3
$\tau^- \to e^-\mu^+\mu^-$	33	3.3		2.7	1.9
$\tau^- \to e^+\mu^-\mu^-$				1.6	1.8
$\tau^- \to \mu^-e^+e^-$	44	3.3		2.7	1.4
$\tau^- \to \mu^+e^-e^-$				1.6	1.4
$\tau^- \to \mu^-\mu^+\mu^-$	49	2.9		1.7	1.3
$\tau^- \to e^-\pi^+\pi^-$		4.2		6.0	2.7
$\tau^- \to e^+\pi^-\pi^-$				1.7	1.8
$\tau^- \to \mu^-\pi^+\pi^-$		4.0		3.9	3.6
$\tau^- \to \mu^+\pi^-\pi^-$				3.9	6.3
$\tau^- \to e^-\rho^0$	37	3.9			1.9
$\tau^- \to \mu^-\rho^0$	44	3.8			2.9
$\tau^- \to e^-\pi^+K^-$		4.2		5.8	2.9
$\tau^- \to e^+\pi^-K^-$				4.9	2.0
$\tau^- \to \mu^-\pi^+K^-$		12		7.7	11.0
$\tau^- \to \mu^+\pi^-K^-$				4.0	5.8
$\tau^- \to e^-K^{*0}$	130	5.4			3.8
$\tau^- \to \mu^-K^{*0}$	100	5.9			4.5
$\tau^- \to e^-\gamma$	64		20		12
$\tau^- \to e^-\pi^0$	210		14		17
$\tau^- \to \mu^-\gamma$	55				3.4
$\tau^- \to \mu^-\pi^0$	82				4.4
$\tau^- \to e^-\eta$			24		6.3
$\tau^- \to \mu^-\eta$					7.3
$\tau^- \to \bar{p}\gamma$					29.0
$\tau^- \to \bar{p}\pi^0$					65.5
$\tau^- \to \pi^-\gamma$					28.0
$\tau^- \to \pi^-\pi^0$					37.0
$\tau^- \to \bar{p}\eta$					129.0

Table 2: Upper limits on branching ratios of neutrinoless τ decays compared with the results of other experiments [10,11,12,13].

References

[1] F.J.Gilman and S.H.Rhie, Phys. Rev. **D31** (1985) 1066.

[2] N.Isgur, C.Morningstar and C.Reader, Phys. Rev. **D39** (1989) 1357.

[3] D.Decamp et.al. (ALEPH), CERN-PPE/91-186 (1991).

[4] H.J.Behrend et.al. (CELLO), Z. Phys. **C46** (1990) 537.

[5] Particle Data Group, Phys. Lett. **B239** (1990) 1.

[6] D.Andreasyan et.al. (Cristal Ball), Phys. Lett. **B259** (1991) 216.

[7] H.Albrecht (ARGUS), Phys. Lett. **B250** (1990) 164.

[8] J.H.Kühn and A.Santamaria, Z. Phys. **C48** (1990) 445.

[9] W.Bacino (DELCO), Phys. Rev. Lett. **41** (1978) 13.

[10] K.G.Hayes et al. (MARK II), Phys. Rev. **D25** (1982) 2869.

[11] H.Albrecht et al. (ARGUS), Phys. Lett. **B185** (1987) 228.

[12] S.Keh et al. (Crystal Ball), Phys. Lett. **213B** (1988) 123.

[13] T.Bowcock et al. (CLEO), Phys. Rev. **D41** (1990) 805.

[14] H.Albrecht et.al. (ARGUS), Phys. Lett. **B163** (1985) 404.

[15] H.Albrecht et.al. (ARGUS), Phys. Lett. **B202** (1988) 149.

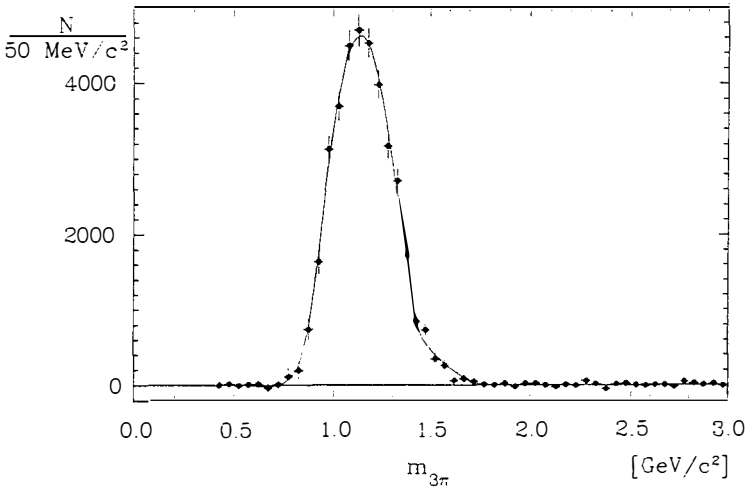


Figure 1: Background subtracted $m_{\pi^-\pi^+\pi^-}$ spectrum. The curve shows the prediction by the model of Isgur et.al. [2].

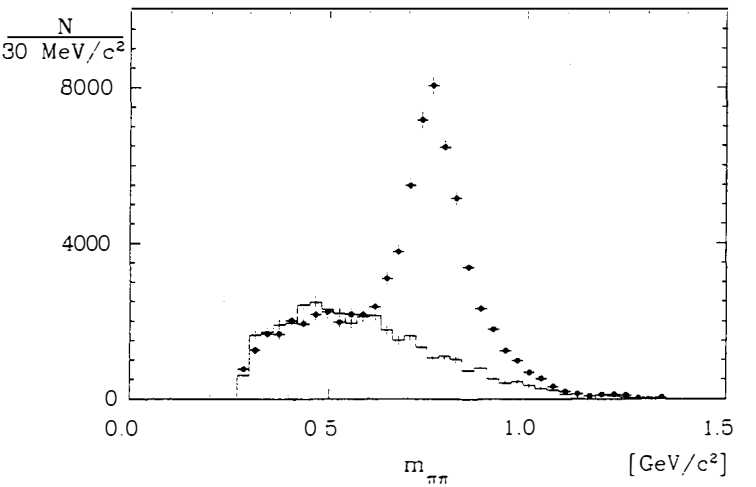


Figure 2: Invariant mass spectrum of the $\pi^-\pi^+$ combinations. The full line represents the spectrum of "like sign" pions to describe the background.

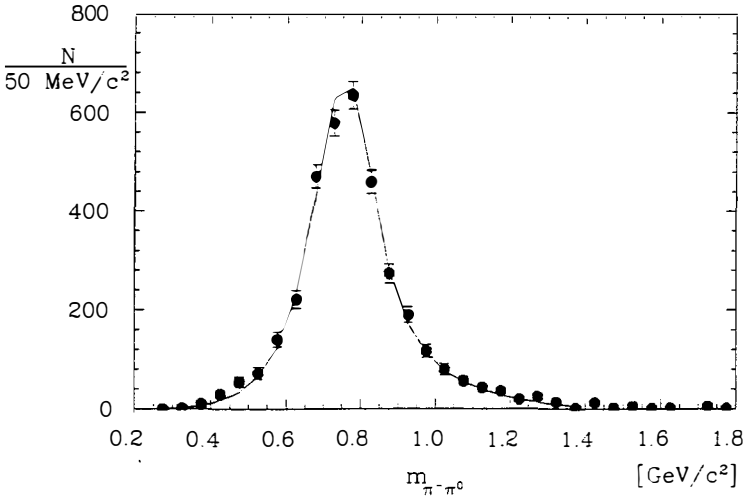


Figure 3: Background subtracted $m_{\pi^-\pi^0}$ spectrum. The curve shows the prediction by Kühn and Santamaria using CVC. Acceptance and resolution are taken into account in the fit function.

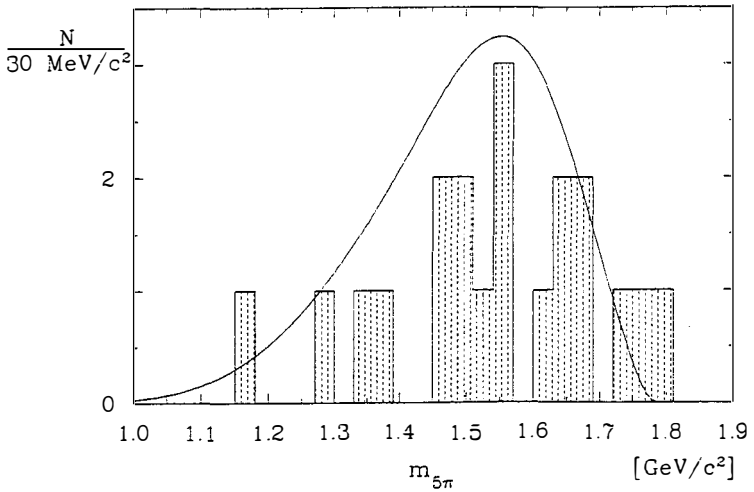


Figure 4: (5π) invariant mass spectrum for the decay τ into five pions. The curve is the prediction for a massless ν_τ

Search for ν_μ - ν_τ oscillations : The NOMAD[1] project

Jacques Dumarchez
LPNHE-Universités Paris 6 et 7
4 Place Jussieu (Tour 33 RdC)
75252 PARIS Cédex 05
France

Abstract

The oscillation mode $\nu_\mu \rightarrow \nu_\tau$ may be the favored channel to probe an effect of massive neutrinos. The NOMAD collaboration has proposed an experiment to look for a ν_τ signal in the CERN wide band neutrino beam. The search relies on good electron identification and efficient kinematical selection, the aim being to improve by one order of magnitude the present oscillation limit obtained with the emulsion technique.

[1] The NOMAD collaboration includes Annecy, Boston, CERN, Dortmund, Dubna, Johns Hopkins, Lund, Melbourne, Ann Arbor, Padova, Paris, Pavia, Pisa, Saclay, Zagreb

1 Introduction

It is generally believed that the ν_τ is likely to have the largest mass among the three neutrino species. A suggested value could be of the order of a few 10 eV when interpreting the missing mass of the Universe in terms of neutrinos. If their mixing angles are not far from what they are in the quark sector, this could give rise to neutrino oscillations at a level observable at present accelerator energies.

To attest the appearance of a ν_τ in a ν_μ beam, the natural way is to visually identify the interaction as producing a τ recognized through its caracteristically short decay length (typically $\sim$ 1mm). This however requires a very high spatial resolution and massive detector. Alternatively we have investigated [1] the possibility to select ν_τ candidates by kinematical criteria.

2 Principle of the search

Apart from a secondary vertex, charged current (CC) ν_τ interactions differ from CC ν_μ or ν_e events by the large missing p_T due to escaping neutrino(s) in τ decays. And indeed the discrimination between ν_τ and ν_e CC events is particularly clear when considering the angle ϕ_1 between the missing p_T and the transverse momentum of the hadronic system, and the angle ϕ_2 between the electron and the hadrons: fig. 1 shows the respective scattered plots of ϕ_1 versus ϕ_2 for the cases of ν_τ and ν_e separately. The basic idea is then for each τ decay channel (in that case $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$) to find a suitable region in the (ϕ_1,ϕ_2) plot plus some additional kinematical cuts to cancel out the background events (here the CC ν_e interactions and all ν_μ NC events where an electron is produced and detected) and leave an appreciable fraction of the possible ν_τ 's.

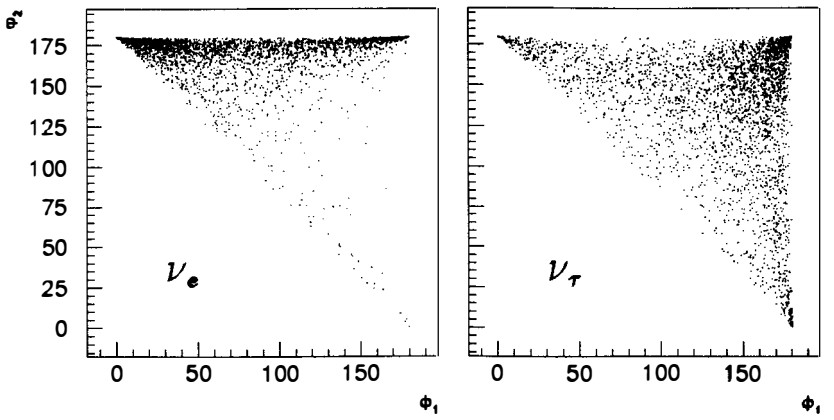


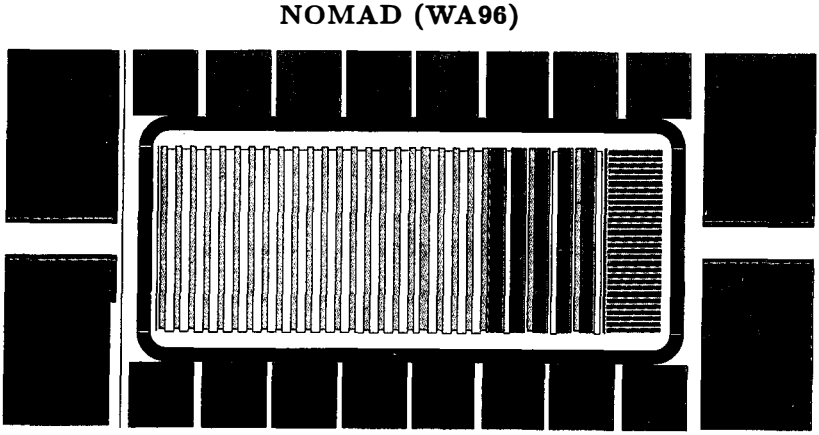
Fig. 1

3 Experimental set-up

To take advantage of the method, the detector has obviously to have good momentum resolution, but also good angular resolution and efficient e/π separation. The technical solution we arrived at is to embed the whole detector in the old UA1 dipolar magnet (fig.2), working with a horizontal field of 0.4T, perpendicular to the beam axis.

The target itself reaches a compromise between the need for high mass and the transparancy necessary for momentum resolution: it consists of 45 big drift chambers (3m x 3m) built with a self-supporting structure made of 1-2cm rohacell sheets, for a total of less than one radiation length. Each chamber has 3 planes of sense wires at angles of 0°, +5°, −5° with respect to the magnetic field direction. With such a configuration it is possible to achieve a precision of $\sigma_y = 200\mu$m per plane and $\sigma_x \sim 2$mm per chamber.

Electron identification is obtained from a transition radiation detector (TRD) and an electromagnetic calorimeter. The TRD is made of 6 modules: the radiators consist of 520 19 μm thick polypropylene foils spaced by 400 μm; the X-ray detectors will have 2 layers of 3m long vertical straw-tubes with diameter 1.6 cm and 25 μm thick walls of alluminized mylar, filled with Xe at atmospheric pressure. According to a simulation program tested on NA31 data, the rejection factor against pions above 2 GeV should reach 1000 with 6 modules. The technical option for the calorimeter is not yet finalized, but will consist of 30 x 30 towers of either lead glass blocks or sandwiches of lead and scintillator tiles read by fibers. The expected energy resolution should be of the order of or less than $10\%/\sqrt{E}$ (E in GeV). Moreover to discriminate between photons and electrons we will have a preshower counter of 1.5 radiation lengths thick lead plate followed by two 1cm wide scintillator strips planes in both x and y directions.



4 Expected results

The experiment has been simulated with the previous detector characteristics and including all possible effects or backgrounds (bremsstrahlung, multiple scattering, Compton, Dalitz decays, knock-on electrons, decays in flight ...). The electron identification combines the TRD signature, the comparison of momentum as measured by the tracking system and energy measured in the calorimeter, and the possible presence of bremsstrahlung photons in a well defined vertical strip. The rejection factor against charged hadrons will then be in excess of 10^6 (a rejection of 10^5 is actually sufficient since it would give 0.4 background events for the full statistics of the proposed run).

Running in the CERN-SPS wide band neutrino beam ($< E_\nu >= 27GeV$) for 2 years ($2.4\ 10^{19}$ protons on target) with a simple interaction trigger would lead to the following number of interactions (the mass of the fiducial target is about 3 tons):

$1.1\ 10^6$ CC ν_μ $\quad$ $3.7\ 10^5$ NC ν_μ $\quad$ 13200 CC ν_e

Different sets of cuts have been defined for the several channels we can study and the following table summarizes the selection efficiencies and number of background events left:

	Selection efficiency	Branching ratio	background events
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$	13.5%	17.8%	4.6
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$	3.9%	17.8%	2.0
$\tau^- \rightarrow \pi^- \nu_\tau$	1.4%	11.0%	0
$\tau^- \rightarrow \rho^- \nu_\tau$	2.3%	22.7%	0
$\tau^- \rightarrow 3\pi \nu_\tau$	8.7%	13.8%	0

It should be stressed that all backgrounds are measurable with the data themselves: kinematically CC ν_e and ν_μ interactions are the same, and backgrounds from NC ν_μ are charge symmetric.

Taking into account the reduction factor for ν_τ cross-section ($\sigma(\nu_\tau)/\sigma(\nu_\mu) = 0.6$) due to the τ mass, the experiment will be able to set a limit on the neutrino mixing angle (in the case of large Δm^2) of

$$\sin^2 2\theta \leq 3.5\ 10^{-4} \text{ at } 90\% \text{ CL}$$

This represents more than one order of magnitude improvement over the present best limit [2].

(More details than on these allowed 3 pages are to be found in the proposal [1])

References

[1] P. Astier et al., CERN-SPSLC/91-21

[2] N. Ushida et al., Phys. Rev. Lett. 57 (1986) 2897.

A CF_4 TPC TO MEASURE THE $\overline{\nu}_e$ MAGNETIC MOMENT AT A NUCLEAR REACTOR

C. Broggini[a], V. Jörgens[b], M. Treichel[b], D. Twerenbold[b], and J.-L. Vuilleumier[b]

[a] I.N.F.N. Laboratori Nazionali del Gran Sasso, 67010 Assergi (AQ), Italy
[b] Institut de Physique, A.-L. Breguet 1, CH-2000 Neuchâtel, Switzerland

ABSTRACT

An experiment is described which offers a significant improvement for the measurement of the $\overline{\nu}_e e^-$ cross section at low energy. The experiment will be sensitive to a neutrino magnetic moment down to a few 10^{-11} Bohr magnetons, on the level of that introduced to explain the solar neutrino puzzle. The detector, to be placed close to a nuclear reactor, is a 1 m^3 Time Projection Chamber surrounded by an active shielding. The key point of the experiment is the use of tetrafluoromethane, CF_4, at the pressure of 5 bar in the TPC.

1 Introduction

One way to explain the observed low flux of ^8B ν_e from the sun in the ^{37}Cl [1] and Kamiokande experiments [2] is the existence of a neutrino magnetic moment of the order $10^{-10} - 10^{-11}$ Bohr magneton units.

The best way of looking for the magnetic moment of the ν_e in a lab experiment seems to be the detailed study of the $\overline{\nu}_e e^-$ elastic scattering. More generally this reaction is fundamental and its study provides information on basic features of the weak interaction. Both charged (CC) and neutral currents (NC) are involved. They are expected to interfere if the NC and CC final state neutrinos are identical, as assumed in the standard model.

The cross section of the $\overline{\nu}_e e^-$ scattering can be divided into two parts: one due to W and Z exchange, which increases linearly with the neutrino energy, and one which comes from the magnetic moment interaction and which increases only logarithmically. To observe the contribution of the second part, one has to use low energy neutrinos. Reactors are copious sources of $\overline{\nu}_e$'s with energies up to 10 MeV, and are ideally suited for such an experiment. The energy spectra are known with good precision, of the order of 3% [3].

Only a few attempts have been made to measure the $\overline{\nu}_e e^-$ scattering [4] [5]. The signal to background ratio was however marginal (~ 0.18 and ~ 0.13) and it was not possible to derive an accurate value for the Weinberg angle $sin^2\theta_W$ from this data.

In order to test the magnetic moment term it is necessary to decrease the threshold on the electron kinetic energy, since it contributes mostly at low T. It therefore appears important to perform a new experiment with a lower threshold detector and a better signal to noise ratio. The signature of the events, one single forward electron, must be fully exploited to suppress as much as possible the background. We believe that a Time Projection Chamber, similar to that used to study $\beta\beta$ decay of ^{136}Xe in the Gotthard underground lab [6], can identify single electrons very well thanks to the good spatial resolution. The end of tracks can be recognized from the enhanced ionization.

2 The test results

The choice of the gas is very important as we need a gas which has at the same time a high density, to maximize the number of target electrons, and a low atomic number Z to minimize the multiple scattering and to allow for the reconstruction of the electron direction. This way it is possible to have a simultaneous measurement of signal plus background events in the forward direction, and of background events in the backward direction. Background can thus be measured on-line, while the reactor is on.

The best way to fulfill our requirements is to use tetrafluoromethane, CF_4, also known as carbon tetrafluoride or freon-14, at the pressure of 5 *bar* [7]. CF_4 has a density of 3.68 $gr \cdot l^{-1}$ (at 1 *bar* pressure and 15 $^\circ C$ temperature) and thus a very high electron density ($1.06 \cdot 10^{21}$ cm^{-3}) and also a good radiation length: $X_o = 35.9$ $gr \cdot cm^{-2}$. Another advantage is that CF_4 does not contain hydrogen, and this suppresses the background reaction $\overline{\nu}_e + p \rightarrow e^+ + n$.

The capability of drifting electrons in CF_4 has been demonstrated in the past over

distances of a few centimeters only [8] [9]. In the experiment we plan to do, we would like to have a threshold of 500 keV or less on the electron kinetic energy and a count rate due to weak interaction of $\sim$ 10 events per day. This is achievable with a 1 m^3 TPC filled with CF_4 at the pressure of 5 bar. Consequently, it is necessary to reach an electron attenuation length of at least $\sim$ 1 m. In order to check if, and how, this value can be obtained, we built a prototype TPC, a cylindrical stainless steel vessel of 20 cm diameter and 30 cm height [10]. All the materials used were selected according to their compatibility with CF_4. The gas is circulated continuosly through an Oxisorb filter to remove oxigen and through a cold trap (-95 $^\circ C$) to remove water and possible freon contaminations

With an ^{55}Fe X-ray source, which we could move inside the TPC, we measured an electron attenuation length longer than 6 m, with a drift field of 120 $V \cdot cm^{-1} \cdot bar^{-1}$.

3 The detector and the signal

The detector will be placed at a distance of 18.6 m underneath the core of the Bugey pressurized water reactor (2800 $MWth$).

The core of the detector we propose to build is a 1 m^3 TPC filled with CF_4 at 5 bar. The gas volume will be defined by a vessel made from lucite which will support the cathode, the read-out planes, and the field shaping rings. We choose lucite because of its radioctive cleaness. The lucite vessel will be immersed in a stainless steel tank filled with a mineral oil based liquid scintillator ,$\sim$ 50 cm thick, which will serve to veto the cosmic muons and as anti-Compton detector. Top and bottom lid of the tank will be instrumented with photomultipliers. The liquid scintillator and the steel vessel will also serve as passive shielding but, since we are above ground, they will not be sufficient. Working outside the containment building we would have to surround the detector by various additional shielding layers, iron + water + B_4C, as was done for the oscillation experiment in Gosgen[3].

The expected rates at the distance of 18.6 m from the core of the 2800 $MWth$ reactor have been calculated assuming W and Z exchange only.

Table 1.

$T(MeV)$	Efficiency (Contained)	Events/day 18.6 m	background
0.25-0.5	0.9	5.4	-
0.5-1	0.85	5.0	2.5
> 1	0.70	4.2	0.2

These rates have to be compared to background rates, which are estimated [10] taking into account direct cosmic muon hits and muon decay events, activation of F and C in inelastic scattering of high energy muons, unstable nuclei production due to μ^- capture, neutron activation, natural ^{40}K, ^{232}Th and ^{238}U activity in the detector. The dominant background is due to the Compton scattering of the γ's produced by the natural activity in the thick walls of the vessel. This background can be estimated from the one measured in the Xe TPC in the Gotthard lab. Taking into account the anti-Compton capability of

the liquid scintillator around the core TPC we obtain the rates given in table 1. These estimates should be considered as upper limits. We hope to learn from the operation of the Xe TPC how to make the CF_4 TPC cleaner.

Considering the signal rates in table 1 a statistical error of 3 % should be achievable in the bin $0.5 < T(MeV) < 1$ in one year of measuring time. Combined with a systematic error of 4 %, essentially from the reactor spectrum, we get a total error of 5 %. This gives a sensitivity to μ_ν around $3 \cdot 10^{-11}$, a factor 10 better than in previous experiments [4] [5]. With some luck the threshold may be lowered more, say down to 300-350 keV, and a sensitivity around $2 \cdot 10^{-11}$ seems then achievable.

The status of the project is the following: a collaboration has been formed between INFN-Gran Sasso, ISN-Grenoble and the Universities of Münster,Neuchâtel, Padova and Züirich and the experiment has been partially financed. We are doing further tests with the prototype TPC in Neuchâtel and making another TPC of the same size but with a plastic vessel in Gran Sasso. We plan to finish the tests in a few months and then to start with the construction of the detector which should be ready for the end of next year.

4 Conclusions

We described a gas detector which offers significant improvements for measurements of the $\overline{\nu}_e e^-$ elastic cross section and which is sensitive to a magnetic moment of the neutrino on the level of that introduced to explain the solar neutrino puzzle.

The distinguishing features of the detector are a low threshold for the electron kinetic energy and track reconstruction capability. The key point of the experiment is to use a 1 m^3 TPC filled with CF_4 at the pressure of 5 bar.

References

[1] K. Lande, Proc. Neutrino 1990 Conf., , Geneva, K. Winter editor

[2] T. Totzuka, Proc. Neutrino 1990 Conf., , Geneva, K. Winter editor

[3] G. Zacek et al., Phys. Rev. D 34(1986)2621

[4] F. Reines et al., Phys. rev. Lett. 37(1976)315

[5] G.S. Vidyakin et al., J. Moscow Phys. Soc. 1(1991)85

[6] M.Z. Iqbal et al., NIM A259(1987)459

[7] C. Broggini, Proc. of the 4th San Miniato Topical Seminar, (1990) 499

[8] L.G. Christophorou et al., J. Chem. Phys. 80(1984)6150

[9] B. Schmidt, NIM A252(1986)579

[10] C. Broggini et al., NIM A311(1992)319

VERY HIGH ENERGY NEUTRINOS FROM ACTIVE GALACTIC NUCLEI.

Hinrich Meyer*
FB Physik
University of Wuppertal
5600 Wuppertal1, Germany

Abstract

Data on the horizontal flux of muons obtained with the Frejus detector in 1984-1988 are used to set a limit on the very high energy (>10^5 GeV) neutrino flux from active galactic nuclei proposed by Stecker et al. The upper limit determined from our data is a factor of 10 lower than the prediction.

* For the FREJUS Collaboration: I. Physikalisches Institut der RWTH Aachen, L.A.L. Orsay, LPNHE Ecole Polytechnique Palaiseau, DPhPE CEN-Saclay, Bergische Universität-Gesamthochschule Wuppertal.

I. INTRODUCTION

At least year's Moriond meeting here in Les Arcs M. Salamon presented a paper on *A Cosmic Background of High Energy Neutrinos from Active Galactic Nuclei* [1]. This paper started a search for the very high muon flux to be expected mostly in the horizontal direction using the data on muons obtained from 1984 –1988 with the FREJUS detector. Neutrinos with an energy spectrum and intensity as predicted to arise from acceleration of protons to very high energy in active nuclei of distant galaxies would have important implications. It would be of great interest for astrophysics, since particle acceleration and interactions has long been proposed as an important ingredient for the understanding of the very energetic processes in the cores of active galactic nuclei. Aside from that the neutrinos would provide a beam with energies way beyond present day accelerators. It could be used to search for unusual phenomena in very high energy neutrino physics like the "inverse W–decay" process $\bar{\nu}_e + e^- \rightarrow W^-$. Even more interesting would be the detection of a process that violates B+L, (baryon number B and lepton L), and simultanously produces with high multiplicity the vector bosons W and Z as well as Higgs particles. The B+L violating process was proposed in 1974 by G. `t Hooft to arise within the standard model, however at an undetectable rate[2]. Estimates for the cross-section at high energies have been worked out in detail recently[3] and measurable rates have been predicted in particular if the AGN-neutrino flux would exist[4]. Further, such neutrinos could serve as a beam for experiments on tomography of the earths interior. I think the points reaised here are of sufficient interest to immediately pursue a search for this very high energy neutrino flux.

II. VERY HIGH ENERGY NEUTRINO DETECTION

Neutrino cross sections has been measured up to a few hundred GeV only. At higher energies the cross sections can be calculated rather reliably up to energies of about 10^8 GeV [5]. At even higher energies the uncertainties in nucleon structure functions at low x, $< 10^{-4}$, limit the accuracy of the calculation. In the near future new measurements of structure functions should become available from the two experiments at HERA - H1 and ZEUS - to sufficiently reduce

this uncertainty. The neutrino cross section reaches about 1 nb at 10^6 GeV and the earth is no longer completely transparent to neutrinos, in fact neutrino absorption becomes strong enough such that the detectable flux of neutrino induced muons is essentially in the horizontal direction. The horizontal muons also have very high energy at the detector such that energy loss of the muons in the detector can be used to discriminate the AGN neutrino flux from background due to atmospheric neutrinos. For the calculation of the neutrino induced muon flux we proceed as follows; the atmospheric neutrino flux at low energies is taken from Gaisser et al.[6] and for the high energy part, >10 GeV from Volkova[7]and the AGN neutrino flux is taken from the paper of Stecker et al.[8]. For the neutrino cross section we refer to Reno and Quigg (Ref. 5) and we have used the structures functions of nucleons as given in Eichten et al.[9]. The energy loss of the muons is parametrised following the calculations of Lohmann et al.[10] using the properties of the rock in the FREJUS tunnel area for the energy loss coefficients. The integral muon energy spectrum at the detector obtained from those calculations is shown in Fig. 1 .

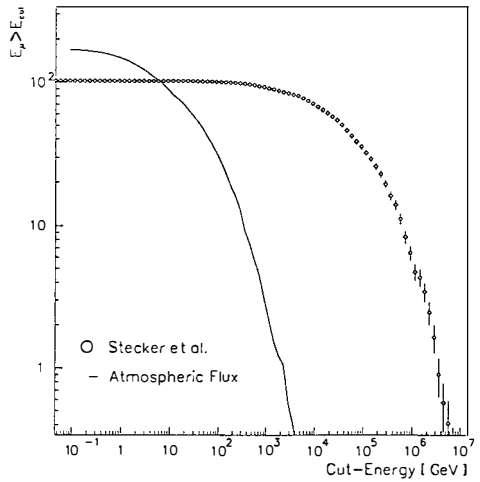


Figure 1.The integral flux of muons with energy > 1 GeV from atmospheric neutrinos (full line) and from AGN neutrinos (points) at the detector level without acceptance cuts.

The rate corresponds to the effective luminosity of the FREJUS dataset but not yet applying any cuts due to the direction of the muons. Note that muons from AGN - neutrinos have very high average

energy, of order 10 TeV at the detector. The crossover between atmospheric- and AGN-neutrinos occurs at about 10 GeV muon energy.

III. THE FRÉJUS-DETECTOR AND -DATA SET

The FREJUS detector took data for about 4 years, from February 1984 till September 1988. The detector of size (6x6x12)m^3 was an iron calorimeter with high spatial and energy resolution[11]. The independent horizontal and vertical views of events consisted of 450.000 flash chamber pixels each and 15.000 geiger tubes used for triggering. The orientation of the detector planes and iron absorber plates was vertical for reasons of mechanical stability. Horizontal muons therefore penetrate the detector perpendicular to the planes and have an average length of about 11 m at an effective density for the detector of 1.95 g/cm^3. The calculation of the muon energy loss in the detector follows the procedure given in Lohmann et al. (Ref10). The response of the detector to electromagnetic and hadronic particles has been calibrated in test beams at the electron-synchrotrons at DESY and the University of Bonn (Ref.11). For the purpose of this paper the hadronic energy loss of muons is treated like the electronmagnetic energy losses, this introduces little uncertainty to the observable energy loss since the response of the FREJUS calorimeter to hadrons and electrons is similar and the hadronic energy loss of muons is only at the level <10 % of the total energy loss (Ref.10).

During its lifetime the FREJUS detector accepted about 10^6 triggers, and in about 500.000 of them muons were recorded. The other half consisted of accidental triggers. Contained neutrino events have a very low rate and only about 250 events have been recorded in the total data sample. The muon sample contains about 470.000 single muons, 6000 muons that range out in the detector (stopping muons) and 15.000 events with more than muon (parallel multi muons) the details of which have already been published[12] The single muon events have been reconstructed to determine the zenith- and azimuth-angle of the muons and to extrapolate back through the structure of the surrounding rock, determining the depth of rock penetrated (as usual in units of 100 g/cm^2). This way a vertical depth intensity curve is obtained and shown in Fig. 2.

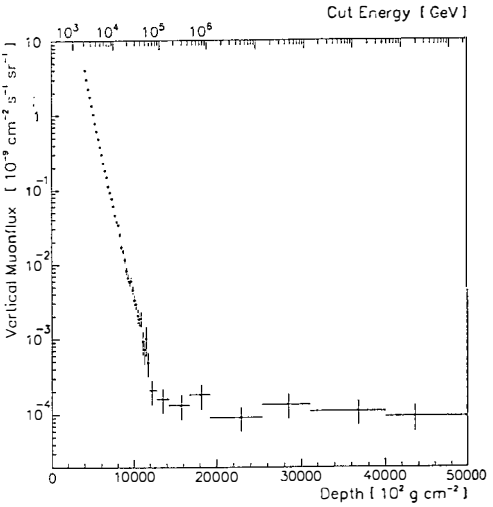


Figure 2. The vertical muon flux as measured in the FREJUS detector.

The muon flux at given zenith angle θ is renormalised to vertical by multiplication with a factor cos θ, since the muon flux from pion and kaon decays in the atmosphere is enhanced by 1/cos θ (at zenith angle θ) in good agreement with our data.

As can be seen from Fig. 2, for rock depth larger than 13000 m.w.e. (4.75 km of rock) the muon flux from pion and kaon decay is absorbed to a level, that neutrino induced muons dominate. This is in very good agreement with data from earlier experiments. Note that in Fig.2 the muon flux from neutrinos is also weighted by a factor cos θ and therefore the data points cannot be directly compared to calculations. For a quantitative comparision to the expectation from the atmospheric neutrino flux (Ref.5,6) we proceeded as indicated in the previous paragraph. The direction of the momentum of the muons cannot be determined in the FREJUS detector since timing in the Geiger counters was not precise enough. Muons around the horizontal direction however are safely identified as neutrino induced since, irrespective of their direction the depth penetrated is larger than 4.75 km. We find that the horizontal muon flux observed in the FREJUS detector is (3.7 ±0,5) $x10^{-13}$ (cm^{-2} sec^{-1} sr^{-1}) for $|\cos\theta|<0.3$, while the flux expected from atmospheric neutrinos is (2.7 ±0.2)$x10^{-13}$ (cm^{-2} sec^{-1} sr^{-1}), about 1.5 standard deviations less than our observation. From this (insignificant) difference between data and expectation no useful limit on the AGN-neutrino flux can be derived. However using

the measurable muon energy loss in the FREJUS detector a very significant limit is obtained. The expected energy loss (in MeV/meter) calculated for the muon spectra as shown in Fig. 1 are presented in Fig. 3, with neutrino induced muons from AGN (full line), from the atmosphere (dashed line) and the measured energy loss for muons as observed near the horizontal direction in the FREJUS detector (dotted line).

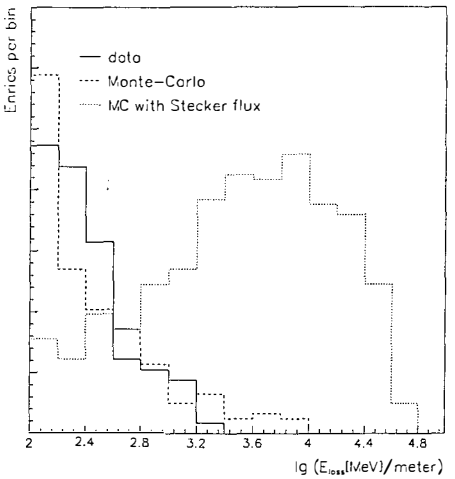


Figure3. The expected energy loss of neutrino induced muons (dashed and dotted lines) as compared to the observed energy loss (full line) for horizontal muons in the FREJUS detector.

We set at cut at 2000 MeV/m where no event survives and 1.1 events (using 12 x the luminosity of our experiment in the MC-calculation) are expected from the atmospheric flux.

The expectation for a AGN-neutrino induced muon flux above this energy loss cut is 25.8 events! On this basis a 90 % c.l. upper limit on the AGN flux can be set (2.3 events) which is a factor of 10 (25.8/2.3) below the expectation of (Ref.8).

This result from the FREJUS detector has important implications on expected neutrino signals for the very large neutrino detectors under construction now[13]. But - at least for me, even more important are the implications for a possible detection of the B + L violating process (Ref.3), indeed loosely speaking 10 times better detectors are required than have been considered (Ref. 4) on the basis of the AGN-flux of Stecker et al. (Ref.8).

Acknowledgements

It is a great pleasure to thank Tran Thanh Van and his crew for a most enjoyable meeting. Discussions with my colleagues from the FREJUS collaboration, in particular K. Daum, W. Rhode and Y. Wei were essential for this analysis. This work was supported by BMFT under 55WT84P.

References:

[1] M.H. Salamon, p. 233 in Proceedings of the XXVI th RENCONTRE DE MORIOND, Les Arcs, Savoie, France, ED. FRONTIERES 1991.

[2] G. `tHooft, Phys. Rev. **D14**, 3432, 1976.

[3] A. Ringwald, Nucl. Phys. **B330**, 1, 1990.

[4] D. A. Morris and R. Rosenfeld, Phys. Rev. **D44**, 3530, 1991.
L. Bergström et al. Phys. Lett. **276**, 231, 1992.

[5] M. H. Reno and C. Quigg, Phys. Rev. **D37**, 657,1988.

[6] T. Gaisser et al. Phys. Rev. **D38**, 85, 1988.

[7] L. Volkova, Sov. J. Nucl. Phys. **31**, 784, 1980.

[8] F. Stecker et al., Phys. Rev. Lett. **66**, 2697, 1991.

[9] E. Eichten et al., Rev. Mod. Phys. **56**, 578,1984 and **58**, 1065, 1986.

[10] W. Lohmann et al. CERN 85-03, March 1985.

[11] Ch. Berger et al. Nucl. Instr. Methods **A262**, 463, 1987

[12] Ch. Berger et al. Phys. Rev. **D40**, 2163, 1989.

[13] Third Int. Workshop on NEUTRINO TELESCOPES, Venezia 1991.

THE HEIDELBERG-MOSCOW DOUBLE BETA EXPERIMENT WITH ENRICHED ^{76}GE: FIRST RESULTS

A. Balysh [b], M. Beck [a], S.T. Belyaev [b], F. Bensch[a], J. Bockholt [a], A. Demehin [b], J. Echternach[a], A. Gurov [b], G. Heusser [a], M. Hirsch[a], H.V. Klapdor-Kleingrothaus[a], I. Kondratenko [b], V.I. Lebedev [b], B. Maier [a], A. Müller[c], M. Osterloh[a], F. Petry [a], A. Piepke[a], U. Schmidt-Rohr[a], H. Strecker [a] and K. Zuber[a]

[a]Max-Planck-Institut für Kernphysik, 6900 Heidelberg, Germany
[b]Kurchatov Institute, Moscow, Russia
[c]Istituto Nazionale di Fisica Nucleare, 6700 Assergi, Italy

Abstract

Two ultra low background detectors of 3.7 kg active mass, made from isotopically enriched ^{76}Ge, have been operated for 1.45 kg·y in the Gran Sasso underground laboratory in Italy. This corresponds to an ^{76}Ge exposure of 16.6 mol·y. New half life limits for the neutrinoless double beta ($\beta\beta0\nu$) decay of ^{76}Ge to the ground state ($T_{1/2} > 1.6 \cdot 10^{24}$ y with 90% c.l.) and the first excited state ($T_{1/2} > 4.0 \cdot 10^{23}$ y with 90% c.l.) of ^{76}Se are reported.
Neglecting right-handed admixtures to the weak interaction Majorana neutrino masses above 1.2 eV can be excluded by this experiment with 90% c.l.
The background characteristic of the experiment is discussed. First results for the two neutrino double beta ($\beta\beta2\nu$) decay are presented.

Presented by: A. Piepke

1. Introduction

The investigation of rare processes allows physics to test some of its basic principles. An example for this is the investigation of double beta decay, the rarest known nuclear decay mode. Especially the lepton number violating neutrinoless decay mode is at present the most sensitive tool to distinguish massive Dirac from Majorana neutrinos. In the frame of GUT theories the B-L symmetry is tested. This is due to the fact, that the decay rate depends besides the leptonic phase space integral and the nuclear matrix element also on "non-standard model physics" as the effective Majorana mass of the neutrinos and a right-handed admixture to the weak interaction.
While the observation of the $\beta\beta0\nu$-decay would prove the existence of massive Majorana neutrinos, its nonobservation yields at present the most stringent limits on the above parameters, since the nuclear physics part of the decay rate is well understood [1].
The two neutrino double beta ($\beta\beta2\nu$) decay, which is a concurrent decay channel, does not require any non-standard physics and has been directly observed for several nuclides (see [2]). Its experimental investigation serves as a test of our understanding of nuclear structure with respect to the calculation of $\beta\beta$-matrix elements-although only to a limited extent for the $\beta\beta0\nu$-matrix elements (see [1]).

2. Experimental procedure

In this experiment the $\beta\beta$-decay of ^{76}Ge is studied. This isotope has a favourable matrix element, a relatively high Q-value (E=2038.56 keV [3]) and can be converted into high resolution detectors. The main advantage of the described experiment is the use of 16.9 kg isotopically enriched Ge. The isotopical abundance of ^{76}Ge is 86% compared to only 7.8% in natural Ge. Goal of this experiment is to test Majorana neutrino masses into the sub-eV range.

Table 1: Technical parameters of the detectors.

	Det#1	Det#2	Det#3
Total mass [kg]	0.985	2.878	2.474
Active mass [kg]	0.93	2.76	2.35
Source strength [mol ^{76}Ge]	10.6	31.6	26.8
FWHM at 1.3 MeV [keV]	2.5	2.4	2.0*)
Threshold [keV]	110	7	-
Measuring time [d]	250.8	107.7	-

*) Measured in a test cryostat

The isotopical composition of det.# 1 has been verified by AMS in order to show, that the enriched material is not diluted by natural Ge during the purification and crystal growing process.

The use of enriched Ge in form of semiconductor detectors allows to utilize large amounts of source strength in a small number of detectors with consequently low sensitivity to background radiation.
Currently two enriched detectors with an active mass of 3.69 kg are in operation. A third one is already working in a test cryostat and will be installed in its final

cryostat in 1992. The technical parameters of the working detectors are summarized in table 1. A fourth crystal of 3.4 kg weight has been recently grown. We expect another ~2.9 kg detector from this crystal.
The enriched detectors are operated in separate shields. The shield of detector#1 has already been described in [4]. Detector#2 is operated together with a natural Ge-detector in a shield of 10 cm LC2-grade Pb and 20 cm of Boliden Pb. Due to its modular construction from standard bricks it is possible to operate up to four big detectors in this shield. The set-up is continuously flushed with ultra pure nitrogen to remove the airborn radon. Both set-ups are located in the Gran Sasso underground laboratory of the INFN in Italy. Its shielding thickness of 3500 m w.e. suppresses the cosmic radiation effectively.

3. The Background

The experience made with det.#1 helped to improve the construction of det.#2. Though using the same construction materials a further background reduction was possible, mainly by more effective elimination of surface contaminations. The exposure to the cosmic radiation was minimized even further by only sea level transportation and deep shielding in between.
In the background spectra of both enriched detectors the characteristic peaks of the cosmogenic isotopes ^{22}Na, ^{54}Mn, ^{57,58,60}Co and ^{65}Zn, the man made activity ^{137}Cs and the natural activities ^{40}K, ^{232}Th and ^{238}U can be identified. For all peaks det.#2 has a lower count rate than det.#1. The reduced overall background is of large importance for the detection of the continuous spectrum of the $\beta\beta2\nu$-decay.
Isotopes decaying via EC within the Ge crystal itself can be identified through their characteristic peaks shifted by the K x-ray energy. Since det.#1 was air transported twice across the atlantic ocean and det.#2 exclusively on surface, we can deduce the equilibrium activities of these isotopes on surface and during flight. The production rates during the flight can be unfolded from the background measurement of det.#1, since the exposure history of both detectors is known. Table 2 contains the measured absolute activities (only statistical errors are quoted) of the cosmogenic isotopes ^{54}Mn, ^{57,58}Co and ^{65}Zn in both enriched crystals, as well as the estimated equilibrium activities on surface and during flight. The latter one has not been calculated for the relatively shortlived isotope ^{58}Co for which the details of the exposure are more important. It should be noted, that the backgrounds of the described detectors are low enough to measure μBq activities.

Table 2: Activities identified inside the Ge.

Isotope	Activity [μBq/kg]		Equil. Act. [μBq/kg]	
	Det.#1	Det.#2	Flight	Surface
^{54}Mn	11 ± 3	4 ± 1	$4 \cdot 10^3$	27
^{57}Co	16 ± 2	3 ± 1	$6 \cdot 10^3$	19
^{58}Co	33 ± 15	7 ± 2	-	14
^{65}Zn	87 ± 12	23 ± 4	$31 \cdot 10^3$	127

From the results contained in tab. 2 we conclude that air transportation of ultra low background Ge should be avoided.
The isotopes contained in tab. 2 can only be formed by spallation reactions.

The biggest improvement factors from det.#1 to det.#2 of 5.9 and 6.6 were achieved for the isotopes ^{40}K and ^{137}Cs, respectively.

4. Results

Table 3: Measured detector backgrounds

Detector	Background [counts/keV·y·kg]		
	100-2700 keV	1465-1700 keV	2000-2080 keV
#1	10.80 ± 0.08	1.35 ± 0.10	0.57 ± 0.11
#2	6.53 ± 0.06	0.86 ± 0.07	0.28 ± 0.07
1 + 2	8.40 ± 0.05	1.07 ± 0.06	0.41 ± 0.06

The $\beta\beta0\nu$-decay of ^{76}Ge to the ground state of ^{76}Se would result in a peak at E=2038.56 keV. The decay to the first excited state ($J^{\pi}=2^{+}$) can only be observed, if the deexcitation γ-quantum of 559.1 keV leaves the crystal without any interaction, leading to a line at 1479.46 keV. The escape probability was calculated to be 37.0% for the combined data of both detectors. The average background around these energies defines the quality and sensitivity of the used detectors. They are given in table 3 together with the integral count rate.

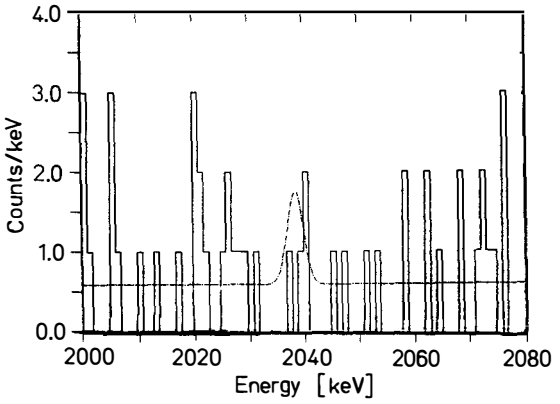


Fig. 1: Energy range around the hypothetical $\beta\beta0\nu$-peak. Also plotted is the excluded signal at 90% c.l.

Constraints on physical parameters are evaluated from the combined data. These data contain no evidence for a $\beta\beta0\nu$-decay, neither to the ground nor to the first excited state.

For the ground state decay we exclude 4.24(2.31) $\beta\beta$-events with 90%(68%) c.l. by using the method recommended by the Particle Data Group for a Poisson-distributed signal superimposed to background [5]. This corresponds to a half life limit of $T_{1/2}>1.61(2.96)\cdot10^{24}$ y with 90%(68%) c.l. For the decay to the first excited state the deduced limit is $T_{1/2}>4.09(7.34)\cdot10^{23}$ y with 90%(68%) c.l.

In fig. 1 the energy interval around 2038 keV is plotted together with the effect excluded with 90% c.l. Only 4 events are contained in the $3\cdot\sigma$-interval around the hypothetical peak after a ^{76}Ge exposure of 16.6 mol·y.

Using the matrix elements of ref.[1] we deduce upper limits for the effective Majorana neutrino mass, the right-right-handed and the left-right-handed admixture to the weak interaction: $\langle m_\nu \rangle < 1.4$ eV, $\langle \eta \rangle < 1.2 \cdot 10^{-8}$, $\langle \lambda \rangle < 2.2 \cdot 10^{-6}$. These are the most stringent limits on these non-standard model parameters ever reported. Neglecting right-handed admixtures to the weak interaction we deduce: $\langle m_\nu \rangle < 1.2$ eV (all limits with 90% c.l.).

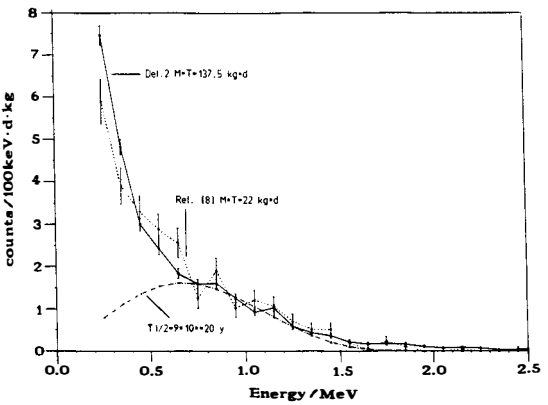


Fig.2: Plotted is the continous part of the spectrum of det.2 (solid curve), the data of ref. [8] (dotted curve) and the calculated $\beta\beta 2\nu$-spectrum for a half life of $9 \cdot 10^{20}$ y.

The unambiguous identification of the continuous energy spectrum of the $\beta\beta 2\nu$ - decay is difficult. References [6, 7, 8] are reporting evidence for this decay of ^{76}Ge. Their half lives are ranging from 9 to $11 \cdot 10^{20}$ y. In order to compare these results with our measurement we have plotted the spectrum of ref. [8] (using also isotopically enriched Ge) in fig. 2 together with our data taken in 1992 (to allow the decay of the short lived background components) with det. # 2. The used measuring time is 137.5 kg·d.

The well defined peaks at 122, 351, 511, 609, 662, 810, 1125, 1173, 1332 and 1462 keV were removed. In this way we take advantage of the good peak to Compton ratio of the big detector, which concentrates the background into the peaks. 326 counts or 5.0% of the events were rejected in this way. No subtraction of any estimated continuous background component has been applied.

From 750 keV to 1650 keV both experiments show essentially the same count rates. They are in agreement with the also plotted $\beta\beta 2\nu$-spectrum with a half life of $9 \cdot 10^{20}$ y.

This finding does not neccessarily imply a confirmation of the $\beta\beta 2\nu$-decay. Although not in conflict with the positive results of the other groups alternative hypothesises might be formulated:

1) Both independent experiments have only the isotopically enriched Ge in common which comes from the same supplier. The similarities in the count rate and shape of the spectra should therefore be caused by the Ge itself. Due to its unmatching spectral shape a possible ^{68}Ge contamination cannot account for the discussed effect. The effect could however be caused by a combination of weak activities present in the crystal.

2) Both experiments show over a wide energy range very similar background conditions. Above 2 MeV the statistical significance of ref. [8] (22 kg·d) is too

low to allow a direct comparison with our data. These similarities could be due to the fact, that in both experiments the cryostats are made from Cu while the shields are exclusively made from Pb. A comparable purity of the used materials would consequently result in similar background spectra.
Since the spectral shape is the only criterion to identify the $\beta\beta 2\nu$-decay in Ge detector experiments, we will test possible alternative background combinations whether they can account for the measured spectrum before claiming a confirmation of the $\beta\beta 2\nu$-decay of ^{76}Ge.

5. Conclusion:

The exceptional source strength of 42 mol of ^{76}Ge and the extremely low background makes the described experiment to the so far most sensitive probe of a nonzero Majorana neutrino mass. The source strength of the full experiment will be ~100 mol. Assuming a moderate improvement of our measured background to 0.1 c/keV·y·kg the sensitivity of the complete set-up will be about $3 \cdot 10^{25}$ y or in units of the neutrino mass ~0.2 eV within 5 years of measurement. Even for the most sophisticated experiments either working or being planned on other enriched isotopes like ^{136}Xe and ^{100}Mo (see contributions in this volume) it will be difficult to attain the sensitivity reachable in this experiment. However they are of great importance to reach some systematics on this nuclear decay mode.

This work is supported by: Bundesministerium für Forschung und Technologie der BR Deutschland and State Committee of Atomic Energy of the USSR.

[1] A. Staudt, K. Muto and H.V. Klapdor-Kleingrothaus, Europhys. Lett. 13, (1990) 31
K. Muto, E. Bender and H.V. Klapdor, Z. Phys. A 334, (1989), 187
[2] M.K. Moe, Nucl. Phys. B (Proc. Suppl.) 19, (1991) 158
[3] J.G. Hykawy, J.N. Nxumalo, P.P. Unger, C.A. Lander, R.C. Barber, K.S. Sharma, R.D. Peters and H.E. Duckworth, Phys. Rev. Lett. 67 (1991) 1708
[4] M. Beck, J. Echternach, J. Gebhard, G. Heusser, H.V. Klapdor-Kleingrothaus, M. Krause, A. Piepke, U. Schmidt-Rohr, H. Strecker, K. Zuber, A. Balysh, S.T. Belyaev, A. Demehin, A. Gurov, I. Kondratenko, V.I. Lebedev and A. Müller, Massive Neutrinos Tests of Fundamental Symmetries, in: Proc. of the XXVIth Rencontre de Moriond (Les Arcs, January 1991), ed. O. Fackler, G. Fontaine, J. Trân Thanh Vân (Editions Frontieres, 1991) p.91
[5] J.J. Hernández et al., Phys. Lett. B 239 (1990) III.36
[6] A.A. Vasenko, I.V. Kirpichnikov, V.A. Kuznetsov, A.S. Starostin, A.G. Djanyan, V.S. Pogosov, S.P. Shachysisyan and A.G. Tamanyan, Mod. Phys. Lett. A 5, (1990) 1299
[7] H.S. Miley, F.T. Avignone III, R.L. Brodzinski, J.I. Collar and J.H. Reeves, Phys. Rev. Lett. 65, (1990) 3092
[8] F.T. Avignone III, R.L. Brodzinski, C.K. Guerard, I.V. Kirpichnikov, H.S. Miley, V.S. Pogosov, J.H. Reeves, A.S. Starostin and A.G. Tamanyan, Phys. Lett. B 256, (1991) 559

NEW LIMITS $T_{1/2}$ FOR DOUBLE β^- DECAY OF ^{116}Cd AND FOR $0\nu\beta^+$/ELECTRON CAPTURE OF ^{106}Cd

F.A. Danevich, V.V. Kobychev, V.N. Kouts, V.I. Tretyak and Yu. Zdesenko
Institute for Nuclear Research of the Ukrainian Academy of Sciences, 252028 Kiev, Ukraine

Presented by Yuri Zdesenko

ABSTRACT

The experiment (~7000 h) with three scintillation crystals $^{116}CdWO_4$ (19; 14 and 13 cm^3) enriched by ^{116}Cd to 83% and natural one $CdWO_4$ (57 cm^3) has been made at the Solotvina Underground Laboratory. The limits $T_{1/2}$ have been set for the double beta decay of ^{116}Cd to ground and excited states of ^{116}Sn and for $0\nu\beta^+$/electron capture of ^{106}Cd:

$0\nu2\beta$ (g.s.- g.s.)	-	$1 \cdot 10^{22}$ y	(90% CL)
$0\nu2\beta$, M^o (g.s.- g.s.)	-	$1 \cdot 10^{21}$ y	(90% CL)
$2\nu2\beta$ (g.s.- g.s.)	-	$1 \cdot 10^{19}$ y	(99% CL)
$0\nu2\beta$ (0^+ - 2^+)	-	$5 \cdot 10^{20}$ y	(90% CL)
$0\nu2\beta$ (0^+ - 0^+)	-	$1 \cdot 10^{20}$ y	(90% CL)
$2\nu2\beta$ (0^+ - 2^+)	-	$1 \cdot 10^{19}$ y	(99% CL)
$2\nu2\beta$ (0^+ - 0^+)	-	$6 \cdot 10^{18}$ y	(99% CL)
$0\nu\beta^+$/EC (^{106}Cd)	-	$1 \cdot 10^{19}$ y	(90% CL)

The energy released in the 2β transition ^{116}Cd - ^{116}Sn is 2802(4) keV. To search this process $^{116}CdWO_4$ scintillators enriched by ^{116}Cd to 83% has been used [1]. Samples of metallic ^{116}Cd were obtained at the State Foundation (619 g; 91%). A cadmium wolframite single crystal with a full mass of 510 g has been grown by the Czochralski technique from a composite material prepared by sintering a mixture of the oxides WO_4 and ^{116}CdO. The crystal was cleaved into five samples and the best from them (19; 14 and 13 cm^3) were used in the measurements. The number of ^{116}Cd nuclei in these samples is 2.18×10^{23}; 1.61×10^{23} and 1.42×10^{23} respectively. The energy resolution of the crystals with photomultiplier tube (PMT) XP2412 were 12 - 13% for γ-line of the ^{137}Cs with energy of 0.661MeV.

All searches have been performed in the Underground Laboratory of the INR which was built in the operating Solotvina salt mine. The Laboratory is situated 430 m deep from the earth surface that corresponds of 1000 m water-equivalent overburden which reduces the muon flux by a factor of 10^4. Due to low radioactive contamination of the salt the natural γ-background in Solotvina Laboratory is 10-100 times lower than in other Underground Laboratories [2].

The experiments were carried out with two kinds of the installation. In first set up two crystals (enriched - 19 cm^3 and natural - 57 cm^3) were coupled to PMT's through the light pipes 9 cm long. The active shielding consisted of four well-type CsI(Tl) crystals with diameter 15 cm and full length 40 cm. The $^{116}CdWO_4$, $CdWO_4$ and CsI(Tl) scintillators were surrounded by the high purity mercury (8 cm thick), lead (23 cm) and polyethylene (24 cm). The plastic scintillator ($110 \times 110 \times 6$ cm^3) was operated as antimuon shielding. The measurements were made during 2600 h [3, 4].

The next set up consisted two enriched crystals $^{116}CdWO_4$ (19 and 14 cm^3) which were coupled to PMT's (XP2412) through the quartz light pipes 25 cm long. The detectors were

shielded by the OFHC copper (3.5 cm thick), mercury (8 cm), lead (23 cm) and polyethylene (24 cm). The background rate in these conditions was little less than early. Therefore it is possible to sum the last spectra which were collected during about 3000 h to the previous ones [3, 4]. The results of the adding are shown in Fig. 1 and give for enriched crystals 850 hxkg (measuring time multiplied by mass), for natural one - 1392 hxkg. The peaks presented in the spectra are caused mainly by ^{40}K and ^{208}Tl from PMT's and ^{137}Cs. There are two differences in these spectra: increase in counts below 320 keV for natural crystal and the small peak at the energy about 800 keV for enriched crystals. The first difference can be attributed to the fourth-forbidden β decay of ^{113}Cd which is contained in the natural crystal (half life about 10^{16} y). Second one can be caused by α-particles from intrinsic radioactive contamination of the enriched scintillators with the level of about $5 \cdot 10^{-11}$ g/g (^{235}U) or $3 \cdot 10^{-10}$ g/g (^{238}U).

Using the data accumulated by enriched crystals during 850 hxkg (Fig. 1) the half life limits for double β decay of ^{116}Cd to ground state of ^{116}Sn have been set:

$0\nu 2\beta$	(g.s.- g.s.)	-	$7 \cdot 10^{21}$ y	(90% CL);
$0\nu 2\beta$, M^o	(g.s.- g.s.)	-	$1 \cdot 10^{21}$ y	(90% CL);
$2\nu 2\beta$	(g.s.- g.s.)	-	$1 \cdot 10^{19}$ y	(99% CL).

Since in the measurements with an active shielding the background spectra have been registered in coincidence with the CsI(Tl) counters as well as in anticoincidence it is possible to search the 2β decay of ^{116}Cd to the excited states of ^{116}Sn and β^+/electron capture of ^{106}Cd which must be followed by γ-rays with energies of 1294 keV ($0^+ - 2^+$), 463 and 1294 keV ($0^+ - 0^+$) and 511 keV (β^+/E capture). Taking into account the numbers of nuclei, time of the measurements and background rates in the corresponding energy regions of the coincidence spectra the limits $T_{1/2}$ were deduced for the 2β decay of ^{116}Cd to the excited states of ^{116}Sn and for

$0\nu\beta^+$/electron capture of ^{106}Cd:

$0\nu2\beta$ $(0^+ - 2^+)$	-	$5 \cdot 10^{20}$ y	(90% CL);
$0\nu2\beta$ $(0^+ - 0^+)$	-	$1 \cdot 10^{20}$ y	(90% CL);
$2\nu2\beta$ $(0^+ - 2^+)$	-	$1 \cdot 10^{19}$ y	(99% CL);
$2\nu2\beta$ $(0^+ - 0^+)$	-	$6 \cdot 10^{18}$ y	(99% CL);
$0\nu\beta^+$/EC	-	$1 \cdot 10^{19}$ y	(90% CL).

Comparing our experimental result for $0\nu2\beta$ decay of ^{116}Cd with theory [5] we find the limit of the neutrino mass less than 8 eV.

The experiment is still running now with three enriched crystals (19+14+13 cm^3) and after the replacement of the PMT's XP2412 by the PMT's FEU-110 with the lower radioactive contamination. The energy resolution of the crystals is about 7.8% at 2.6 MeV and background rate around 2.8 MeV (within FWHM energy interval) is equal 2.3 counts/y/keV/kg. With 540 hxkg of a data taking the present limit $T_{1/2}$ is 10^{22} y (neutrino mass less than 7 eV). After one year of measurements it is possible to improve this limit till $3 \cdot 10^{22}$ y ($m_\nu < 4$ eV). Now we are preparing the big scale experiment with $^{116}CdWO_4$ crystals (10 kg of the ^{116}Cd) in which for $0\nu2\beta$ decay of the ^{116}Cd the limit $T_{1/2} > 10^{25}$ y can be reached ($m_\nu < 0.2$ eV).

REFERENCES

1) F.A. Danevich et al. Lett. JETP **49** (1989) 417
2) Yu.G. Zdesenko et al. Proc. of Int. the Symp. on Underground Physics, Baksan Valley, 1987 (Moscow: Nauka, 1988) 291
3) F.A. Danevich et al. Abstract of the14th EPS Conf. Rare nuclear decays and fundamental processes, Bratislava, 1990 (Frankfurt: EPS) **141** (1990) 29
4) Yu.G. Zdesenko J. Phys. G: Nucl. Part. Phys. **17** (1991) S243
5) A. Staudt, K. Muto and H.V. Klapdor-Kleingrothaus Europhysics Letters **13** (1990) 535

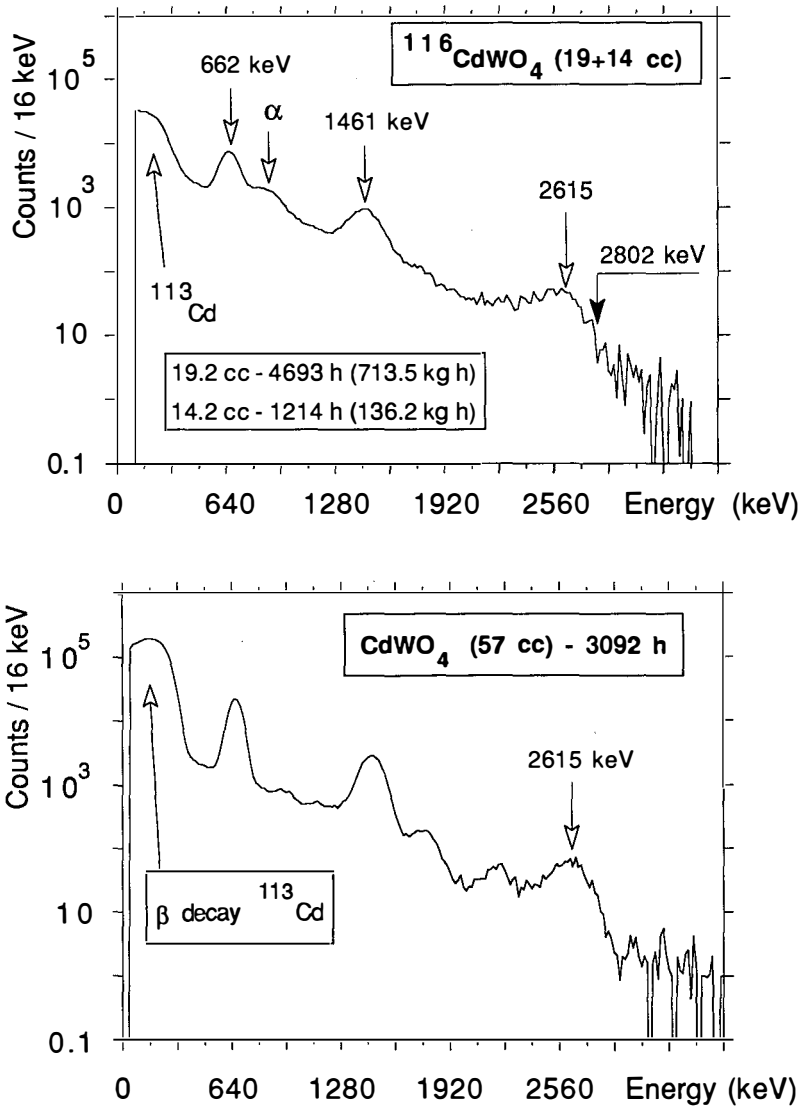


Fig. 1. Background spectra of the enriched (upper) and natural crystals

The Neutrino(less) Experiment with MOlybdenum: 2^{nd} prototype

J.E. Campagne
LAL, Centre d'Orsay, Bat. 200, 91405 Orsay Cedex, France
representing the NEMO collaboration *

Abstract

The NEMO-II detector has been running successfully since August 1991. After 2000 hours no 2e-event has been observed with an energy greater than 2.3MeV on a very pure copper foil. Simultaneously, the level of pollution of a natural molybdenum foil not specialy purified has been measured, and the results are in good agreement with a germanium γ-spectroscopy measurement. Delayed emission of α particles has also been observed using $e\gamma\alpha$-events.

*CEN Bordeaux, CRN Strasbourg, INR Kiev, LAL Orsay, LPC Caen.

1 Introduction

The NEMO collaboration plans to built an experimental set-up managing 10 kg of 99% enriched ^{100}Mo to study $\beta\beta(0\nu)$ decay, in order to probe effective neutrino mass of the order of 0.1 eV.

To reach this goal improvements have to be made in two different fields. First, ultra low radioactivity detectors [1] and sources [2] have to be built to observe counting rates as low as expected in double beta decay. Second, in order to increase the signal/noise ratio, new types of set-up must be developed with multiple selection criteria, such as tracking, time of flight signatures, and detection of X and γ rays, in addition to the measurement of energy in a large solid angle.

2 The NEMO-II set-up

The NEMO-II detector (Fig.1) has been built in order to gain experience in the second field mentioned in the introduction, and to build on the results obtained to date [1] from the 64 cell detector. All the materials used in this detector have been carefully chosen for their low radioactivity level, except the photomultipliers. The detector is presently running in the Fréjus Underground Laboratory (4800 m.w.e depth).

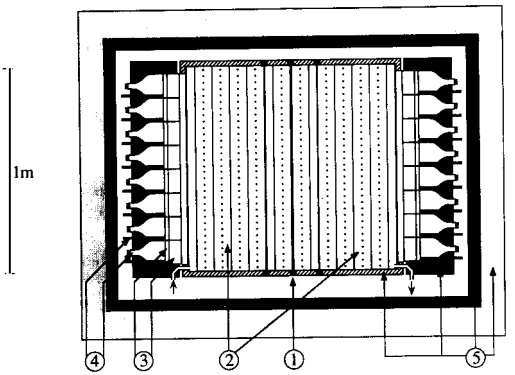


Figure 1: The NEMO-II set-up

1) Central source foil.
2) Tracking volume 1 m^3 with alternate horizontal and vertical Geiger cell planes.
3) Phoswich of 2cm plastic scintillator NE110 and 2.5 mm of CsI(Na).
4) 3" Photomultiplier RTC XP2312 with associated fibre for laser calibration.
5) Shielding: copper frame 2 cm (hatched area), lead 5 cm (dark), iron 20 cm (grey).

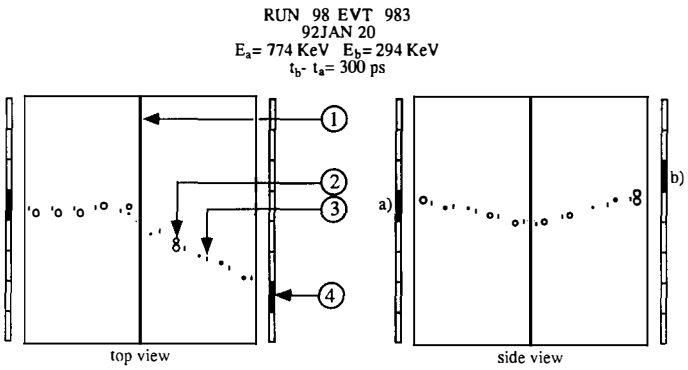


Figure 2: Typical 2e-event

1) 1 m × 1m plane of the source foil.

2) Circle with radius corresponding to the drift time in a Geiger cell. The trajectory of the particle is tangential to the circles.

3) Longitudinal position deduced from the plasma propagating time in the Geiger cell.

4) Plastic scintillator fired giving energy and time measurements.

2.1 The tracking device

On either side of the central source are arranged 5 modules, each containing 64 Geiger-cells, of which half are orientated vertically, half horizontally in two separate planes. The octogonal cells have 30 mm in diameter and 1 m in length. On both ends there is a copper ring used as a cathodic probe. The 1 m^3 detector volume filled with a mixture of helium gas and 2% of ethyl alcohol at atmospheric pressure, is enclosed by sheets of aluminized Mylar (36 μm Mylar, 0.08 μm Al) which are in turn protected by a layer of Rohacell (expanded plexiglass 1 mm thick, density 50 kg/m^3). This tracking device provides a 3D-reconstruction (Fig.2) of charged tracks with σ_T = 500 μm and σ_L = 5 mm for the transverse and longitudinal precisions, respectively.

2.2 Energy and t.o.f measurements

On either side of the tracking volume a plane of 64 counters (12 × 12 cm^2 each) consisting of 2 cm of NE110 plastic scintillator and 2.5 mm of CsI(Na) coupled to 3" PMTs is used first to measure the energy deposit by electrons or γ/X rays, and second to record the e/γ arrival time. The electronic threshold of the leading edge discriminator corresponds to about 50 keV deposit in the plastic scintillators.

The energy calibration is achieved using different kinds of e/γ sources (^{207}Bi, ^{137}Cs and ^{88}Y) and it yields a FWHM of 16% at 1 MeV. The photomultiplier gain stability is checked daily using a laser/light-fibers system. During 4 months of running, the relative gain variation has been less than 5%.

For t.o.f measurement, which is a particularity of our detector, amplitude corrections must be made. This is because the apparent time difference between a 100 keV electron signal and a few MeV electron signal is as large as the t.o.f of a 100 keV electron along 50 cm ($\approx$ 3 ns). Using high energy Compton electrons, a time resolution of 400 ps at 1 MeV has been measured.

2.3 Detection of X rays and α particles

To face the problem of natural radioactive decay simulating double beta events, one should reduce as far as possible the radium isotopes concentration in the source by radiochemistry, but one must also provide for the detection of the remaining impurities. This can be done by tagging X rays and α particles.

Taking into account the design of the phoswich, the computed efficiency of X rays detection inside the CsI(Na) is 70% at 80 keV. As for α particles, we record events occurring between 12.0 μs and 1 ms after an e/γ coincidence; in addition, the tracks left by α particles characteristically stop within the tracking volume. The α particle detection efficiency is only about 5.5%, mainly due to the thickness of the source foil.

3 Data acquisition and first analysis

Usually the trigger requires the firing of at least 2 plastic counters within 50 ns, followed by 4 Geiger modules within 2.5 μs. Since we started taking data the trigger rate has shown a very good stability 0.231 ± 0.004 Hz.

Since August 1991, the central source foil has been divided into two halves. One half consists of a very pure copper foil (1.0×0.3 m^2, 70 μm thick); the other half consists of an unpurified natural molybdenum foil (1.0×0.3 m^2, 50 μm thick). The purpose of this exercice was to determine the impurity level of the molybdenum foil and to compare the result with a germanium γ-spectroscopy analysis. So far the detector has accumulated more than 10^6 triggers in about 2000 hours of continuous running.

3.1 The NEMO-II data

The detector allows us to observe 2e-events like the one shown in (Fig.2). An electron is defined as a track associated to a scintillator with an energy deposit greater than 200 keV. The rate before any cut is about 70 h^{-1}. To decrease the high level of background coming from Compton electrons due to the ^{40}K isotope concentration ($\approx 10^5$ dpm/kg; 1 dpm = one decay per minute) present in the photomultiplier glasses, we make a rough cut based on criteria of t.o.f to separate events of external origin from events originating from the source foil. We reduce the background noise (≈ 0.8 h^{-1}) without loss of signal.

The X-position of the vertices of the remaining events is shown in (Fig.3). It is clear that most of the events come from the molybdenum foil. This corresponds to $\approx$ 22000 events/m^2/y. A simulation using GEANT[3] based on the geometry of the NEMO-II together with the elementary processes and a simplified ^{214}Bi decay scheme

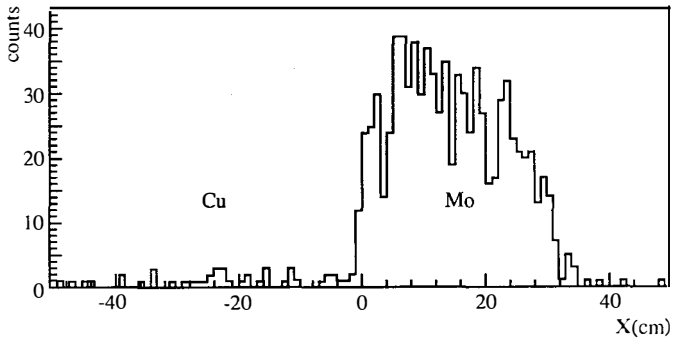


Figure 3: Horizontal vertex position on the central plane of 2e-events after t.o.f cut. The positions of the two foils are [-30 cm, 0 cm] and [2 cm, 32 cm] for the copper and the molybdenum, respectively. During 930 hours there were 723 events (31 events) on the molybdenum foil (copper foil).

gives a figure of $\approx$ 200 dpm/kg for this number of events. A similar analysis done with eγ-events gives the same order of magnitude of pollution with 10^5 events/m^2/y observed.

It is also quite interesting to look at e/γ coincidences followed by α particle emission which is a good signature of the ^{214}Bi $\rightarrow$ ^{214}Po $\rightarrow$ ^{210}Pb decay scheme. We have been monitoring this kind of event for 400 hours and the delay time histogram (Fig.4) of the α particles shows a decay structure with $\tau \approx 185$ μs which is in good agreement with the last cascade time constant (164 μs).

3.2 Comparison with germanium γ-spectroscopy

The useful isotope activities of the molybdenum foil measured using a 100 cm^3 "p" type HPGe crystal in the Fréjus Laboratory are: 90 ± 10 dpm/kg for the ^{214}Bi (^{238}U chain) and 7 ± 2 dpm/kg for the ^{208}Tl (^{232}Th chain). It is obvious that this industrial quality foil is not suitable for a double beta decay experiment, but it is encouraging to note that the ^{214}Bi concentration deduced using NEMO-II data is in quite good agreement with this germanium measurement.

3.3 High energy events

The 31 events left on the very pure copper (during 930 hours) are due to external electrons escaping the t.o.f cuts, and to external photon Compton interactions in the foil followed by Møller scattering. In any case, there is no event with an energy greater than 2.3 MeV ($Q_{\beta\beta} = 3.03$ MeV for the ^{100}Mo). On the molybdenum foil, above 2.7 MeV one predicts from the germanium γ-spectroscopy results about 6 events (3% efficiency assumed) which is consistent with the 3 events observed.

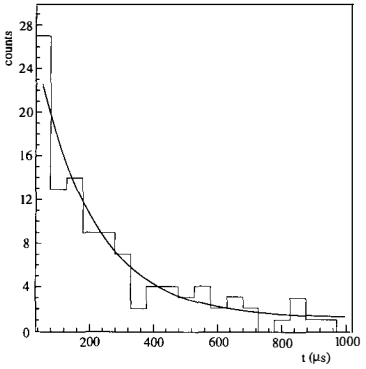


Figure 4: Time delay distribution of α particles occuring in $e\gamma\alpha$-events tagging $^{214}Po \rightarrow {}^{210}Pb$ decay. A maximum likelihood fit with a flat background hypothesis gives a decay time value $\tau = 185^{+55}_{-40}$ μs which is in good agreement with the table value of 164 μs .

4 Conclusion

The NEMO-II detector, second stage of a long term project of $\beta\beta(0\nu)$ study of the ^{100}Mo isotope, has been running successfully since August 1991. The events are very clean and for the most part easy to interpret. An exercise has been carried out using a natural molybdenum foil not specially purified. It turns out that the level of pollution of this foil, as measured using NEMO-II 2e-events and $e\gamma$-events, is in good agreement with a germanium γ-spectroscopy. Moreover, the decay of Polonium into Lead by delayed emission of α particles has also been observed using $e\gamma\alpha$-events.

After 2000 hours no event has been observed with an energy greater than 2.3 MeV on a very pure copper foil. We plan to introduce in the near future a very pure natural molybdenum foil[1] which is more suitable for studying the background in the 3.03 MeV region where $\beta\beta(0\nu)$ decay possibly takes place.

References

[1] D. Dassie et al., Nucl. Instr. and Meth. A309 (1991) 465-475.

[2] D.Blum et al., Phys. Lett. B275 (1992) 506-511. and Yu. G. Zdesenko, ibid.

[3] Release 3.14 of the GEANT 3. version, R. Brun et al, CERN DD/EE/84-1.

[1]CLIMAX Corporation USA and ITEP Moscow.

New Limit on Neutrinoless Double Beta Decay in ^{136}Xe with a Time Projection Chamber

F. Boehm[a], K. Gabathuler[c], H. Henrikson[a],
D. Imel[a], M.Z. Iqbal[a], V.J. Jörgens[b],
L.W. Mitchell[b], B.M. O'Callaghan-Hay[a], J. Busto[b],
M. Treichel[b], J.-C. Vuilleumier[b],
J.-L. Vuilleumier[b], H. Wong[a], P. Fisher[a]

a) California Institute for Technology, Pasadena, CA 91125, USA
b) Institut de physique, Université de Neuchâtel, CH–2000 Neuchâtel, Switzerland
c) Paul Scherrer Institute, CH–5232 Villigen PSI, Switzerland

Presented by J.-C. Vuilleumier

Abstract

A Xenon TPC with an active volume of 207 liters has been built to study neutrinoless double beta decay in ^{136}Xe. Data were taken in the Gotthard Underground Laboratory with 5 atm of xenon enriched to 62.5 % in ^{136}Xe. A total of 6218 hours of data has been taken. Data analysis is in progress and new limits are presented.

Introduction

Neutrinoless double beta decay provides a sensitive probe for lepton number violation and, in particular, Majorana neutrino mass as well as right-handed weak currents[1]. The implications of this so far unobserved nuclear decay have stimulated intense experimental efforts in recent years[2].

Experimental Setup

We have built a Time Projection Chamber (TPC) to study double beta decay in ^{136}Xe (transition energy 2.48 MeV), in both the 0ν and 2ν channels[3]. The schematic diagram of the experimental set-up is shown in Figure 1. The TPC has a cylindrical active volume of 207 litres. The operating pressure is 5 atm, with an admixture of 3.9% methane to increase the drift velocity and to suppress diffusion of the secondary electrons. Xenon enriched to 62.5% ^{136}Xe is used, giving a total of 1.6×10^{25} ^{136}Xe atoms in the active volume. There are 168 readout channels, with 3.5 mm pitch, in each of the X and Y axes. The detector has been built with low background materials, and is shielded by 20-30 cm of lead. The experiment is being conducted at the Gotthard Underground Laboratory, with a 3000 meter-water-equivalent overburden which attenuates the muon flux by a factor of 10^6.

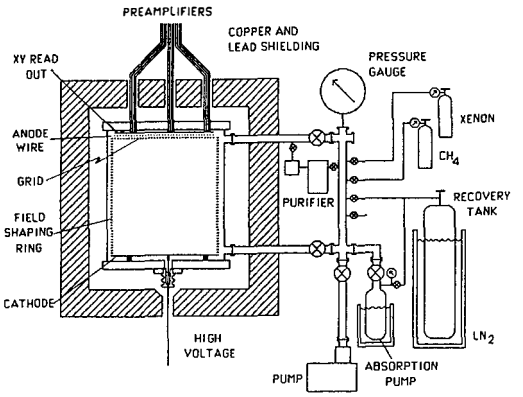


Figure 1: A schematic diagram of the time projection chamber with the associated setup.

Track Reconstruction

The track reconstruction capability of the TPC provides a powerful means of background rejection. A double beta decay event is identified as a continuous trajectory with the characteristic "end features": large angle multiple scatterings and increased charge depositions (charge "blobs"), at both ends. A typical "two-electron" event, which exhibits such features, is depicted in Figure 2a. Such events can be distinguished easily from those due to alpha particles, cosmic rays, single electrons and multiple Compton scattering. Figure 2b shows a beta decay (that is, a single electron with charge blob at only one end) with the emission of an alpha particle at the same (X,Y) co-ordinate 50 μs later. This event is due to the cascade $^{214}\mathrm{Bi} \rightarrow {}^{214}\mathrm{Po} + \mathrm{e}^- + \bar{\nu}_e$ ($\mathrm{Q} = 3.28\,\mathrm{MeV}$, $\mathrm{T}_{\frac{1}{2}} = 19.7\,\mathrm{min}$), followed by $^{214}\mathrm{Po} \rightarrow {}^{210}\mathrm{Pb} + \alpha$ ($\mathrm{Q} = 7.8\,\mathrm{MeV}$, $\mathrm{T}_{\frac{1}{2}} = 164\,\mu\mathrm{s}$), and is evidence of trace radon emission in the system. This cascade can be singled out by looking for the $\sim$100 μs post-trigger alpha activity after an initial single electron event. The energy of an event is obtained by integrating the anode signals over the drift time.

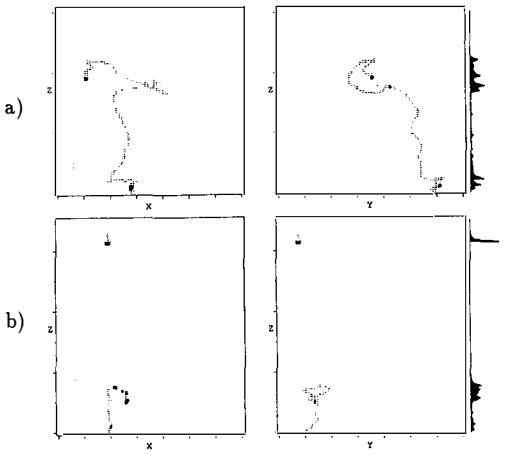


Fig. 2. A typical (a) "two-electron" event, and (b) beta decay with delayed alpha emission.

The energy resolution and calibration has been studied with various gamma sources. A 10-15% variation in charge multiplication across the effective area of the anode plane is observed. To correct this effect, the anode plane is sub-divided into 45 squares and a gain variation map is made from a measurement with a ^{137}Cs source. The anode signals are then compensated at each time bin, based on the (X,Y) co-ordinate of the event at that time. A notable improvement on the energy resolution of 6.6% FWHM at 1.6 MeV is subsequently achieved.

Results

The "two-electron" spectrum from 6218 hours of data is shown in Figure 3. The measured background level at the 0ν range is 0.01 counts $\mathrm{keV}^{-1}\,\mathrm{kg}^{-1}\,\mathrm{yr}^{-1}$. Considering a resolution of 6.6% FWHM at the 0ν transition energy, and an exponential background from 2000 keV to 2650 keV together with a constant background from 2650 keV to 3000 keV, we obtain a 90% C.L. exclusion curve as indicated by the solid line.

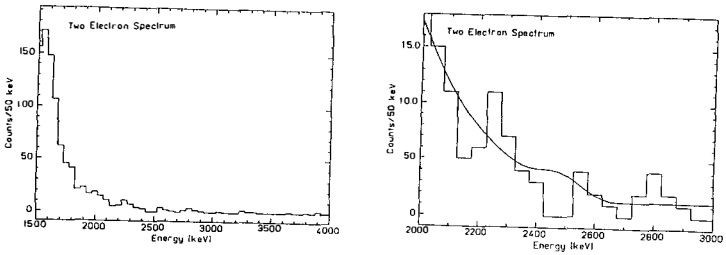


Fig. 3. Energy spectrum for "two-electron"event.

Folding in the respective detector (25% and 21%) and analysis efficiencies (81% and 64%) for the mass mechanism and right-handed currents modes, we obtain the following 90(68)% C.L. half-life limits for the $0^+ \rightarrow 0^+$ transition:

$$T^{0\nu}_{\frac{1}{2}}(\langle m_\nu \rangle) > 4.1(7.8) \times 10^{23}\ \text{years}\,,$$

and

$$T^{0\nu}_{\frac{1}{2}}(\mathrm{RHC}) > 2.2(4.2) \times 10^{23}\ \text{years}\,.$$

These limits represent a factor of 30 improvement over existing ones[5,6].

The limit for the Majorana neutrino mass parameter thus deduced depends on which nuclear matrix element calculation one adopts. To illustrate the range of sensitivities, the limits deduced from the calculations by the Caltech[6] and Heidelberg[7] groups are tabulated in Table 1. For comparison, the best ^{76}Ge 0ν half-life limit[8] as well as the projected sensitivity of this experiment with 3 years of data are also shown.

Table 1. Limits on Majorana neutrino mass parameter from the ^{76}Ge and ^{136}Xe experiments, adopting calculations from the Caltech and Heidelberg groups.

Isotope (run time)	$T^{0\nu}_{\frac{1}{2}}(\langle m_\nu \rangle)$ / yr (at 68% C.L.)	$\langle m_\nu \rangle$ / eV	
		Caltech	Heidelberg
^{76}Ge (3 years)	$> 8.8 \times 10^{23}$	$< 2.4 - 4.8$	< 1.4
^{136}Xe (6218 hours)	$> 7.8 \times 10^{23}$	$< 2.2 - 2.8$	< 1.7
^{136}Xe (projection for 3 years)	$(>)3.3 \times 10^{24}$	$(<)1.1 - 1.4$	$(<)0.8$

Owing to the theoretical uncertainties, it is essential that double beta decay investigations are pursued in various isotopes (preferrably with different ranges in Z). The experiment described here is at present the only one which has achieved the $0\nu\beta\beta$ sensitivities of ^{76}Ge with a different isotope. The simultaneous tracking and calorimetry capabilities of the TPC, using the source as its detector medium, prove to be ideal for double beta decay studies. The technique of using the difference in dE/dx to measure the directionality of low energy electron can be a great asset to other experiments at nuclear energies (like the measurement of the neutrino magnetic moment using reactor neutrinos[9]).

References

1. See, for example, F. Boehm and P. Vogel, *Physics of Massive Neutrinos* Cambridge University Press, Cambridge (1987).
2. H.T. Wong, Ph. D. Thesis, Caltech (1991).
3. H.T. Wong *et al.*, Phy. Rev. Lett. **67** 1218 (1991).
4. A.S. barabash *et al.*, phys. Lett. **B223(2)**, 273 (1989).
5. E. Bellotti *et al.*, Phys. Lett. **B266**, 193 (1991).
6. J. Engel, P. Vogel and M.R. Zirnbauer, Phys. Rev. **C37**, 731 (1988).
7. A. Staudt *et al.*, Europhys. Lett. **13(1)**, 31 (1990).
8. J.G. Hykawy at al., Phy. Rev. Lett. **67**, 1708 (1991)
9. C. Broggini *et al.*, Nucl. Instrum. Methods. **A311**, 319 (1992)

Acknowledgements

This work was supported by the Fonds National pour la Recherche Scientifique and by the US Department of Energy.

Search for Double Beta Decay of ^{130}Te with a Cryogenic Thermal Detector and Simultaneous Detection of Light and Thermal Signals in a Cryogenic CaF_2 Detector

presented by Luigi Zanotti

A. Alessandrello, V. Bashkirov(*), C. Brofferio(**), D.V. Camin, O. Cremonesi, E. Fiorini, G. Gervasio, A. Giuliani, M. Pavan, G. Pessina, E. Previtali, L. Zanotti

Dipartimento di Fisica - Università di Milano e INFN - Sezione di Milano, Italy
(*) Moscow Engineering Physical Institute - Moscow, Russia - only for the work on CaF_2
(**) Laboratori Nazionali del Gran Sasso - INFN, Assergi (L'Aquila), Italy

Abstract

Two recent results obtained with low temperature thermal detectors are presented.

First a search for double beta decay of ^{130}Te is described; a low temperature thermal detector consisting in a single TeO_2 crystal of 73 g has been operating for 868 hours in the Gran Sasso laboratory. Performances of the detector and results of the measurement are described; a lower limit of 2×10^{21} years (90% C.L.) was obtained for neutrinoless double beta decay of ^{130}Te.

Then a new approach to the study of double beta decay of ^{48}Ca is presented. The technique is a combination of scintillation and thermal pulses detection, studied to reject α particles background which is relevant for such decay, whose transition energy is very high ($Q = 4271$ keV). Simultaneous detection was successful in a measurement performed with a 2 g $CaF_2(Eu)$ crystal at $T = 60$ mK. Future evolution of the technique is discussed.

1 - Introduction

Double beta decay (DBD) of ^{130}Te has been observed in experiments performed by the geochemical technique, which is not able to distinguish between the conventional two-neutrino (2ν) decay and the lepton number violating neutrinoless (0ν) decay. The world average of results, as presented by O. Manuel in Bratislava (1990)[1], yields a half-lifetime $\tau_{1/2} = 0.8 \times 10^{21}$ y; it is therefore of some interest to understand which is the contribution of the 0ν channel to DBD of that isotope, if any. Since transition energy is high ($Q = 2528.8 \pm 0.8$ keV, in between the 2615 keV γ line of ^{208}Tl and its Compton edge) and isotopic abundance is 34 % in natural Tellurium, a direct detection technique, in which source and detector coincide, can reach a good sensitivity to the process, considering that, according to recent theoretical calculations[2,3], a more favorable nuclear merit factor (phase space $\times$ | nuclear matrix element |^2) than in the case of ^{76}Ge is obtained.

^{48}Ca, even if its isotopic abundance is very low (0.187 %), is also a good candidate for the study of 0ν DBD, since its Q-value is very high ($Q = 4271 \pm 4$ keV), well outside natural radioactivity background, and recent calculations, based on the QRPA approach[4], show that its nuclear merit factor is also favorable in comparison with ^{76}Ge, while recent experimental results[5] are still far from lifetime limits that would give a sensitivity to neutrino effective mass comparable to that of ^{76}Ge. At such a high transition energy it is however possible that α decays from radioactive natural chains (mainly ^{238}U) induce a background in the energy region of interest; it is therefore necessary to develop some technique to identify such decays in the energy spectrum of the measurement. A possible new technique will be described in the following.

2 - The technique

The idea of detecting particles through the measurement of heat released in a medium was put forward by E. Fiorini and T. Niinikoski in 1984[6]. A crystal of pure dielectric and diamagnetic material has a thermal capacity given by:

$$C = 1944 \frac{m}{A} \left(\frac{T}{T_D} \right)^3 \qquad J\ K^{-1}$$

where m = mass, A = molecular weight, T = temperature, T_D = Debye temperature. If T is low enough (100 - 10 mK) and T_D is high enough, then C can be so low that even the energy deposited by, for instance, an electron of 1 MeV, can produce a detectable raise in the crystal temperature ($\Delta T = \frac{E}{C}$), which can be transformed in an electrical signal by a suitable thermistor.

The detector used in this work is a 73 g TeO_2 crystal ($T_D = 275$ K), whose theoretical thermal capacity at 20 mK is 0.5 nJ K^{-1}, corresponding to a $\Delta T = 300$ μK for a 1 MeV

energy deposition. Theoretical resolution with such detectors can be extremely good, namely of 10 eV (R.M.S.) in the above quoted case.

3 - Experimental setup

The 73 g crystal is mounted in a suitable frame (OFHC Copper) and held by 68 spring loaded tips to compensate for different thermal contraction of materials and minimize vibrations; thermal to electrical signal transducer is a NTD thermistor, provided by E. Haller (Berkeley), which is glued on the crystal surface by a two component epoxy glue. The signals are read out via a cryogenic differential voltage sensitive preamplifier and recorded on computer by a CAMAC digitizing system.

Crystal and its frame are mounted in a dilution refrigerator that was specially built by Oxford Instruments Company using materials selected for low radioactivity content; when possible OFHC Copper was used instead of stainless steel.

The refrigerator, whose cooling power is 1000 μW at 100 mK, can reach a base temperature of 5.5 mK and it is installed in a building in Hall A of INFN Gran Sasso Underground Laboratories (LNGS), located near the town of L'Aquila, in Central Italy.

The mountain overburden guarantees a reduction of cosmic ray flux to 1 $\mu\ m^{-2}\ h^{-1}$, and neutron flux to about 10^{-6} neutrons $cm^{-2}\ s^{-1}$; more-over the refrigerator is shielded against environmental radioactivity by a 10 cm thick Lead shield, is closed in a plastic bag flushed with Nitrogen to reduce Radon contaminations and is surrounded by a Faraday cage to suppress electromagnetic disturbances of local origin.

The base temperature for the crystal is 12 mK, which raise to 16 mK when bias is applied to thermistor (25 mV on a 438 MΩ resistance).

Detector performances are shown by Fig. 1 in which a comparison is reported with a conventional Ge diode; in the calibration spectrum with a ^{232}Th source energy resolution is ~ 6 keV practically independent of energy.

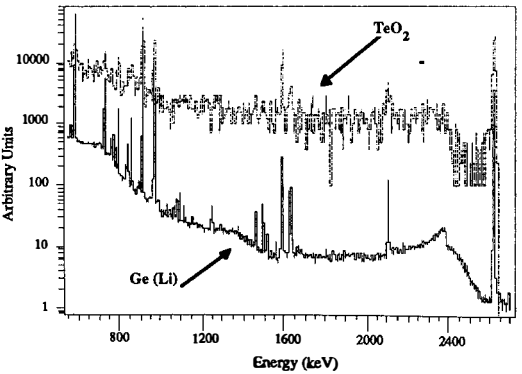


Fig. 1 - Calibration spectra of Ge(Li) and TeO_2 detectors

4 - The measurement

After having collected data for 868 hours (the spectrum is shown in Fig. 2) a preliminary background analysis is already possible.

All relevant γ lines have been identified and most of them can be attributed to ^{222}Rn daughters, showing that there is still a Radon source in the surroundings of the detector; ^{232}Th lines appear only at high energy (2.6 MeV), indicating a possible source external to the cryostat. Moreover a clear peak appears at E = 5413 keV, which can be attributed to a contamination of ^{210}Po in the crystal; considering that α particle energy is E_α = 5305 keV, the difference to 5413 keV is due to full detection of nuclear recoil, which is an interesting feature in view of possible Dark Matter experiments with this technique. The activity from ^{210}Po corresponds to 7.3 ± 0.4 mBq/kg.

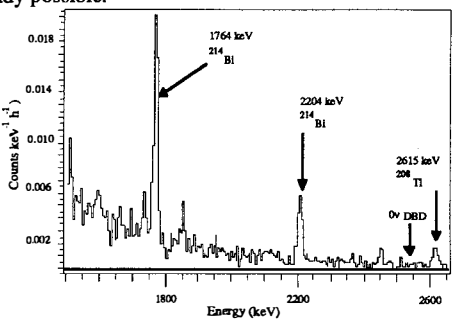


Fig.2 - Spectrum after 868 h of measurement

Considering a 10 keV window centered on the 0v DBD energy of ^{130}Te at 2528 keV, a maximum likelihood procedure allows to quote the following result after 868 h of measurement:

$$N = -2.4 \,{}^{+\,3.0}_{-\,1.1} \text{ events}$$

Therefore, after renormalizing conservatively the likelihood function to the physical region (N > 0), we obtain the following result:

$$\tau_{1/2}(0\nu\,;\, 0^+ - 0^+) > 2 \times 10^{21} \text{ y} \quad (90\% \text{ C. L.})$$

which implies that neutrinoless DBD contribution to the geochemical result previously quoted is excluded at a level greater than 40 %.

The above quoted result must be considered as preliminary; to improve we have to collect higher statistics, to decrease backgrounds at least in the energy region of interest and to use larger crystals. A possible goal would be the use of an array of 4 crystals of 400 g each, with which it could be possible to reach a sensitivity of ~ 10^{23} y, in the limit of 0 events. Finally it must be stressed that the device described above is also a very sensitive γ ray detector, owing to the high Z of Tellurium.

5 - The CaF_2 experiment: heat pulse detection tests

Scintillation technique has been, until now, the most direct way to study ^{48}Ca DBD; much better resolutions, as shown before, can however be achieved by using a low temperature thermal detector. In this case a detector consisting in a CaF_2 crystal, whose Debye temperature is rather high (T_D = 400 K) can be used. However ^{238}U α lines can appear at energies around 4.2 MeV, and a contamination level as low as 10^{-11} at/at would give about 1000 ^{238}U decays/year in a 30 g $^{48}CaF_2$ crystal. This problem can be circumvented considering that, since the light yield of α's is about 1/5 of that of electrons at the same energy, it is possible to discriminate α background by collecting scintillation pulses in coincidence with thermal pulses from the crystal at low temperature (< 100 mK); the ratio of scintillation to thermal pulses would be different for α's and electrons of the same energy. A preliminary measurement was performed at T = 60 mK on a pure CaF_2 crystal of ~ 2 g, showing a resolution of 50 keV at 1.6 MeV; the measurement was repeated with a crystal of the same mass doped with 0.01 % Eu, which is commonly used to produce scintillation, in order to check possible side effects of the dopant on the thermal capacity of the crystal, and the same result was obtained.

6 - The CaF_2 experiment: scintillation tests

Scintillation of CaF_2 at low temperatures was also tested. Crystals of ≈ 2 g with different doping levels (from pure to 0.07 % Eu content) were cooled down to 4 K with a 100 mm^2 silicon photodiode glued on one face, while the crystal was covered by a teflon light reflector, and irradiated by a ^{241}Am α source. Results of such test showed that light yield at 420 nm decreases until T ≈ 50 K and then stabilizes at a level of 50 % (in the case of a 0.03 % doping) of that detected at room temperature, while pure CaF_2 light is no more detectable below such temperature. A further test was performed on the crystal doped at 0.03 % level, only changing the reflective coating to aluminum 10 μm thick not to spoil the thermal properties of the CaF_2; the same result as before was obtained down to 30 mK with a light yield for 5.5 MeV α particles of ≈ 4500 photoelectrons, corresponding to ≈ 4000 photoelectrons/MeV for electrons.

7 - The CaF_2 experiment: coincidence measurement

After having performed the above quoted tests, the final setup was mounted in a dilution refrigerator in Milano, in order to test the simultaneous detection of light and heat. The same CaF_2(0.03 % Eu) crystal that was used before was glued to two 100 mm^2 AME photodiodes by means of a special two component optical glue, and firmly mounted with spring loaded tips. Photodiodes were parallel-connected and DC coupled to a GaAs charge preamplifier operating at 4 K. On a side of the crystal a NTD Ge thermistor was glued to collect the thermal pulses; since a room temperature preamplifier was used a large parasitic

capacitance gave thermal pulse risetimes of ≈ 10 ms. Nevertheless at T ≈ 20 mK it was possible to collect data in coincidence between the two channels, even if the characteristic signal collection times in the two channels differ by about 3 orders of magnitude (thermal pulses t ≈ 10 ms, scintillation pulses t ≈ 10 μs). In the scatter plot reported in Fig. 3 coincidence events are plotted according to their pulse height in the thermal channel (Y axis) versus scintillation channel (X axis); two features appear: a cluster of events that correspond to ^{241}Am α particles, and events grouped around a straight line, which correspond to photons from a ^{232}Th source. It is therefore proved that simultaneous detection of light and heat pulses at low temperatures can be performed; this implies that CaF_2 scintillation can be detected down to 20 mK and that doped CaF_2 crystals can be successfully used as thermal detectors with good resolutions.

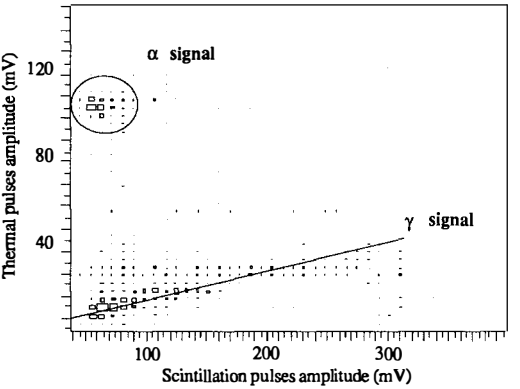


Fig. 3 - Scatter plot of thermal pulse height versus scintillation pulse height; pulses are produced by α particles from ^{241}Am and γ rays from ^{232}Th.

Preliminary tests show that a new kind of mounting is possible, in which the light detector (photodiode) is not optically (and hence thermally) connected to the crystal and the light reflector is just a white paint on the inner walls of the copper mount of the crystal (which is always kept in position by spring loaded tips); such a configuration should improve the thermal response of the system that would work as in a purely thermal detection setup. Such tests have been performed on a 70 g crystal, using a AME photodiode of 900 mm^2 down to about 20 K; the system will be moved soon into the dilution refrigerator for lower temperature and coincidence test.

References

[1] O. K. Manuel, J. Phys. G: Nucl. Part. Phys. **17** (1991) S221
[2] K. Muto et al., Z. Phys. **A334** (1989) 187
[3] J. Engel et al., Phys. Rev. **C37** (1988) 731
[4] J. Suhonen, University of Jyväskylä, Finland, JYFL 25/91
[5] K. You et al., Physics Letters **B265** (1991) 53
[6] E. Fiorini, T. O. Niinikoski, N.I.M. **224** (1984) 83

UPDATED REVIEW OF 17 keV NEUTRINO EXPERIMENTS

Douglas R.O. Morrison
C E R N
Geneva, Switzerland.

Abstract

A brief review of the history of 17 keV neutrinos up to August 1991 is given and it is noted that experiments using tritium are unreliable because of atomic effect and BEFS fluctuations. Recent experimental results are presented. The question of neutrinos from Supernova 1987A, `shape correction factors' in magnetic spectrometer experiments, and the problems of Si(Li) detectors with backscattering etc. are discussed – in each case the evidence is not favourable for 17 keV neutrino. Overall the balance of experiments is strongly against there being any heavy neutrino near 17 keV.

1. Introduction

A review of experiments on a possible 17 keV neutrino up to August 1991 has already been given [1]. In this present review the main results up to that time are recalled briefly and new results and future experiments are considered in more detail.

2. History of 17 keV neutrino, 1985 to Aug. 1991.

In 1985 J.J. Simpson reported [2] that in the decay spectrum of electrons from tritium, there was a deviation from the theoretical prediction, being a kink 17 keV below the peak energy, E_0 of 18.7 keV. He interpreted this as a new heavy neutrino of 17 keV mass with a 3% branching ratio. However, other experiments at Princeton, Alzitoglou et al. [3], Tokyo, Ohi et al. [4], ITEP, Apalikov et al. [5], BARC/TIFR, Datar et al. [6], Caltech, Markey and Boehm [7], and Wark and Boehm [8], Chalk River, Hetherington et al. [9], and, at CERN, using internal bremsstrahlung, Borge et al. [10], all showed no evidence for a 17 keV neutrino and gave upper limits on a branching ratio much lower than 3%, (being 0.4%, 0.3%, 0.17%, 0.6%, 0.25%, 0.25%, 0.35%, 0.9% resp., at least 95% Confidence Limit).

Several theoretical papers, Lindhard and Hansen [11], Druckarev and Strikman [12] and more recently Koonin [13], showed that the region of the tritium spectrum near one keV where Simpson claimed to observe a deviation, had major problems with atomic effects and fluctuations due to Beta Environmental Fine Structure [13]. It was concluded that tritium could not safely be used to study a possible 17 keV neutrino. The experiments [3–9] used sources (^{35}S or ^{63}N) which had higher end-points so that these atomic effects did not apply to them. In what follows, experiments with tritium will not be seriously considered as they are unreliable for studying the 17 keV question.

Thus in the period 1986 to 1988, it was generally considered that the experimental evidence, plus theory, was strongly against the existence of any heavy neutrino near 17 keV.

However in 1986, Simpson [14] strongly criticized the null experiments and said some of them actually showed in their published data evidence for a kink corresponding to a 17 keV neutrino. In particular he presented a graph of the results of Ohi et al. [4] which appeared to show data consistent with a 17 keV neutrino with 2% mixing; he said that only the region near the expected kink should be considered and defined this as ± 4 keV. However, the interval chosen was in fact, +4.1 to -4.9 keV which included favourable fluctuations and excluded unfavourable fluctuations; furthermore the fit chosen is inconsistent with the end-point – this is shown more fully in ref. [1]. Also Ohi et al. took considerably more data [15] which show no fluctuation in favour of a 17 keV neutrino. A similar analysis by Simpson of the early Hetherington et al. data [9] was also shown to be a statistical fluctuation [16]. Simpson [14] also noted a deviation from a straight line in a plot in the Apalikov paper [5] but this is inconsistent with the actual spectrum published and a re-analysis gives an upper limit of 0.16% with 90% CL [17].

It was suggested by Simpson [14,18] that the null experiments used magnetic spectrometers and these had large correction factors making them unreliable. He was correct in saying that all magnetic spectrometer experiments up to 1989 needed to introduce a `scale factor' or `shape correction factor' to correct for imperfections, and it is useful that he made this comment and brought the subject

to people's attention. However, the essential question is `Do these scale factors introduce an anti-kink in the spectrum which would cancel out the kink that a 17 keV neutrino would introduce?' Inspection of the scale factor curves shows them to be smooth and no evidence for an anti-kink has been found in any experiment. Further there are many different designs of magnetic spectrometers and different shapes and dimensions, so that it would be a remarkable coincidence if all these scale factors had anti-kinks at exactly (E_0 - 17) keV.

In 1989 interest was revived when in the Phys. Rev. of April 1st appeared two papers from Guelph, one by Simpson and Hime [18] claiming a neutrino of (16.9 ± 0.4) keV with mixing of (0.73 ± 0.11)% using ^{35}S as a source thus avoiding the atomic effects problems of tritium (the other paper, by Hime and Simpson, used a tritium source) Then in 1991 Hime and Jelley [19] in a new experiment found with a ^{35}S source, evidence for a (17.0 ± 0.4) keV neutrino with a mixing of (0.84 ± 0.08)% . A major change from the earlier work at Guelph was the use of baffles to ensure the electrons from the source impinged almost normally on the detector. Further confirmation came from Berkeley where Sur et al. [20] using a ^{14}C source found evidence for a (17 ± 2) keV neutrino with a mixing of (1.4 ± 0.5)%. Also in Zagreb, Zlimen et al. [21] found weak evidence(two standard deviations) for a 17 keV neutrino with 1.6% mixing. On the other hand at Caltech, Becker et al. [22] found no evidence for a 17 keV neutrino and gave an upper limit for the mixing of 0.6%.

This series of reports re-awakened interest in 17 keV neutrinos and many experiments were started and a large number of theoretical papers were written, some investigating the consequences of the existence of such a particle and others describing the problems for such a particle to be consistent with current ideas and knowledge of astrophysics and cosmology.

J.J. Simpson [23] noted that in Supernova 1987A, there was a gap in time between the early and late neutrinos recorded by Kamiokande and IMB detectors. Assuming this is caused by a 17 keV neutrino which decays to a normal neutrino giving a time delay, he calculated the mean life of the 17 keV neutrino to be (0.6 to 1.6).10^4 s.

3. New experimental results

A further sign of the great interest was that a three-day meeting on 17 keV neutrinos was held in Berkeley in December 1991 where some new data were presented plus many proposed experiments.

Later in 1991, a paper by Hime and Jelley [24] gave further evidence for a neutrino of (16.75 ± 0.35 ± 0.15) keV mass and a mixing probability of (0.99 ± 0.12 ± 0.18)% but this time using a different source, ^{63}Ni.

The Berkeley group of Sur et al. [25] have increased their statistics and their new value is (17.1 ± 0.6) keV with a (1.2 ± 0.3)% mixing, that is the significance has increased to four standard deviations. The group leader, Eric Norman's conclusions [26] are (1) they have evidence for a kink in their data and, (2) they have no interpretation of this kink.

A major new experiment at Caltech has been reported at this Moriond conference by Tom Radcliffe [27]. They find no evidence for a 17 keV heavy neutrino with an upper limit of 0.45% mixing at 90% CL, and they exclude a 0.85% admixture at the 99.9% confidence level. To check that if a kink were there, they would have detected it, they inserted a thin (17 micron) aluminium foil

covering about 10% of the source, which with corrections, was expected to give a spectral distortion resembling a 1% mixture heavy neutrino. Analysing the resultant spectrum assuming a kink, gave a 15.6 keV `neutrino' with a 2.5% mixing. The `scale factor' for their spectrometer was shown and is completely smooth with small errors and no sign of an anti-kink at 17 keV below the end-point. The experiment is continuing.

Schreckenbach et al. have presented results [28] from a magnetic spectrometer of an unusual design at ILL, Grenoble. They claim to have no shape corrections in the region expected for a 17 keV neutrino. They use a ^{177}Lu source. They find no evidence of a 17 keV neutrino with an upper limit of 1% with 90% CL. They expect to reduce this to about 0.1 to 0.3% soon.

A joint Oxford/NPL experiment [29] using the NPL magnetic spectrometer, has found no evidence for a 17 keV neutrino. They find that they need a scale factor for their spectrometer and are continuing the experiments in the hope of removing this factor.

The experiment with the most impressive statistics (2.4 billion events over 1800 energy bins!), is by Ohshima et al. [30] of INS, Tokyo. They use a magnetic spectrometer and checked that the scale factor is smooth and has no anti-kink. They find no indication of any significant kink in their data. They give an upper limit of 0.1% with 95% confidence for the mixing probability of a 17 keV neutrino.

A question that is frequently asked is why do people who find heavy neutrinos always find them at 17 keV and not at some other mass. An Argentine group, DiGregorio et al. [31] from Tandar, report that in an experiment studying the inner bremsstrahlung of ^{71}Ge, they find a best-fit mass of (13.8 ± 0.18) keV with a mixing of (0.8 ± 0.25)% and they exclude a mass of 17.2 keV with a 0.85% mixing with 96% confidence. Thus observation of other masses is possible.

4. Special subjects

4.1 Supernova 1987a

Simpson [23] has suggested that there is a gap in time in the events recorded by Kamiokande and by IMB. The actual time distribution is shown in fig. 1 of ref. [32], normalized to a common time zero. It may be noticed that the later IMB events fall in the gap of the Kamiokande events. Further it is important to note that the events must be weighted by their efficiency of observation which is very high near the threshold. The thresholds are 20 eV for IMB and 7.5 eV for Kamiokande. The weights are 5 and 7 times for the two IMB events which fill the gap. Some typical weights are marked against the points in this figure. It may be seen that there is no significant gap in the arrival time of neutrinos from Supernova 1987A.

However it is interesting to note that after Simpson [23] justifies his calculation by the existence of a gap, he later says that assuming the result obtained, calculation shows that there should not be a gap.

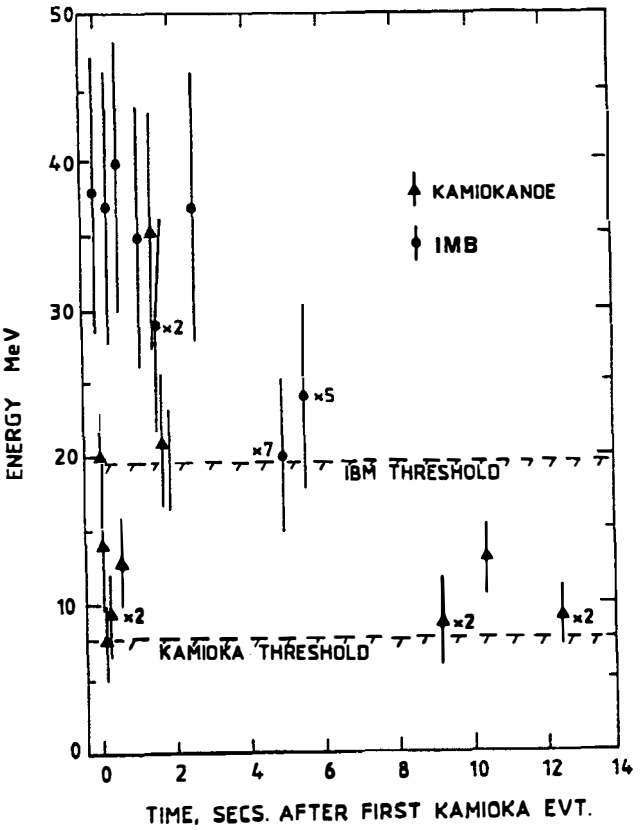


Fig. 1: Plot of energy versus time for the Kamiokande and IMB events, taking time zero for each experiment as the time of their first event. Some typical weights estimated are shown as numbers against the event.

4.2 Null experiments associated with magnetic spectrometers

Many have noted the claim [14,18,19] that null results are associated with experiments using magnetic spectrometers and these spectrometers require `shape correction factors' which are of the same magnitude as the effect searched for. Three points can be made:

a) There are experiments which do **not** use magnetic spectrometers and which do **not** find a 17 keV neutrino. The experiments of Borge et al. [10] and DiGregorio et al. [31], both use internal bremsstrahlung, and, more importantly, the experiment of Ohi et al. [4] uses Si(Li) detectors as do those finding a 17 keV neutrinos. Instead of using a single detector where there is a major problem of backscattering (between 13 to 30%), they have a better design where they use two Si(Li) detectors close together with the source in between - one is taken as the main detector and the other as the veto - see refs. [1,4]. Ohi et al. give an upper limit of 0.3%.

b) Thanks to Simpson, who correctly pointed out the existence of `scale factors', the groups using magnetic spectrometers now consider their correction factors very carefully. They point out that they are smooth in the region of interest and could not conceal a kink. Simpson has commented on the high-statistics experiment of Hetherington et al. [9] noting that the correction factor was of the same magnitude as the effect needed to see a 17 keV neutrino. However, recent analysis [16] shows that the actual correction used increased with electron energy instead of decreasing as required by Simpson – in addition the shape was quite different from that required by him.

c) In general the addition of a magnetic field allows the momentum of the electrons to be measured and controlled. This extra control results in a lower background and thus a better experiment.

4.3 Problems of Si(Li) detectors

The major problem of Si(Li) detectors is the large and uncertain amount of backscattering. In the 1989 experiment at Guelph, Simpson and Hime [18] used no baffles between the source and the detector. They state that this is similar to the case of diffuse electrons and assume the value of 30% for the amount of backscattering (whereas they say that if the electrons had been incident normally there would be 15% backscattering); further, they make the important assumption that the rate of backscattering is independent of the energy. In the 1991 experiments, Hime and Jelley [19, 24] used apertures in front of the source and of the detector and an intermediate baffle to try and make the electrons strike the detector normally and they state that there is no energy dependence. Remembering that they claim an accuracy of one part in a thousand for the mixing, it is interesting to check if there is an energy dependence. In the thesis of A. Hime [33], a graph is shown of published data for the % backscattering for normal electrons and a value of (13.5 ± 0.2)% is quoted. In addition are shown three experimental points they have measured giving an average of (12.9 ± 2.4)%, but the three points suggest an energy dependence. Also over the energy ranges for ^{35}S and ^{63}Ni there is a significant energy variation. Further, if the apertures work perfectly, the electrons are up to 10 degrees from normal and this would also give a significant energy dependence. Further, the intermediate baffle is unfortunately placed as its flat edge is visible to both the source and the detector. The existence of an energy dependence is important as this could appreciably reduce the significance of any effect and would increase the error making it more than one part in a thousand.

In a paper presented at this conference, Piilonen and Abashian [34] list eight possible sources of energy-dependent errors in the Hime and Jelley experiment [19], and have been able to calculate five of them – these five would account for a third of the effect claimed.

In future it would be safer to fit the electron energy spectrum with an additional term of the form $(1 + aE)$, where E is the electron energy and a is a parameter.

It is possible that the explanation of the kink could be some other effect not mentioned here.

5. Future experiments

So far we know of 17 experiments in progress or proposed. Four are taking further data (Radcliffe et al., at Caltech, Wark et al. at Oxford/NPL, Norman et al.,

at Berkeley, and Schreckenbach et al., at ILL). Two new experiments are taking data (Pomansky et al., at Baksan, and Mortara et al., at Argonne/Berkeley). Seven other experiments are starting and of these four are using tritium; some of these use interesting new techniques. It is interesting to note that 9 of the experiments are in the Bay Area.

6. Summary of experimental results

The main experimental results are summarised in fig. 2. It may be seen that there are four experiments which observe an effect with a mixing of about 0.8 to 1.0%. There are 10 experiments which observe no effect and which have upper limits which are less than 0.6% with 90% confidence, that is, the upper limits are significantly less than the positive values claimed. Further, some have very low upper bounds, in particular Ohshima et al. [30] have an upper limit of 0.1% with 95% confidence.

The criticisms of the experiments showing no effect, have been shown not be applicable. Worries have been presented about the use of Si(Li) detectors because of the large amount of backscattering (13% to 30%), and possible energy variations which could be serious compared to the the error of one part in a thousand claimed.

It is concluded that the balance of experimental results is strongly in favour of there being no 17 keV neutrino.

7. Conclusions

1) The original 1985 experiment of Simpson using tritium, is shown to have no significance and the value of 17 keV obtained is arbitrary.

2) There are four significant experiments (Simpson and Hime [18], Hime and Jelley [19], Hime and Jelley [24] and Sur et al. [25]) which show evidence in favour of a 17 keV neutrino. There are 13 significant experiments which find no effect; of these 10 have upper limits on the amount of mixing which are significantly less than the mixing claimed. Further, there are several high-statistics experiments which have upper limits on mixing far below the values claimed.

3) Generally the use of magnetic fields in an experiment is a major advantage as they allow better control and improve signal-to-noise ratios. Most magnetic spectrometers have `scale factors' which correct imperfections. An anti-kink, which would cancel out a real kink indicating a 17 keV neutrino, has been looked for but in every case the scale factors are smooth and give no such indication.

4) Si(Li) detectors have a major problem of backscattering (13% to 30%) and energy dependence which seem not to have been completely studied yet.

5) The experiments finding no evidence for a 17 keV neutrino have been performed by many different groups and a variety of techniques. The experiments supporting a 17 keV neutrino involve rather few people and a narrow range of techniques.

8. Overall conclusion

One cannot say with 100% certainty that a 17 keV neutrino exists. One cannot say that the probability of it existing is 0%. But the balance of experimental evidence is strongly against a 17 keV neutrino existing.

Acknowledgments

It is a pleasure to thank many people for helpful discussions, in addition to those listed in the references are A. Hime, N. Jelley, D.H. Perkins, and T. Yamazaki.

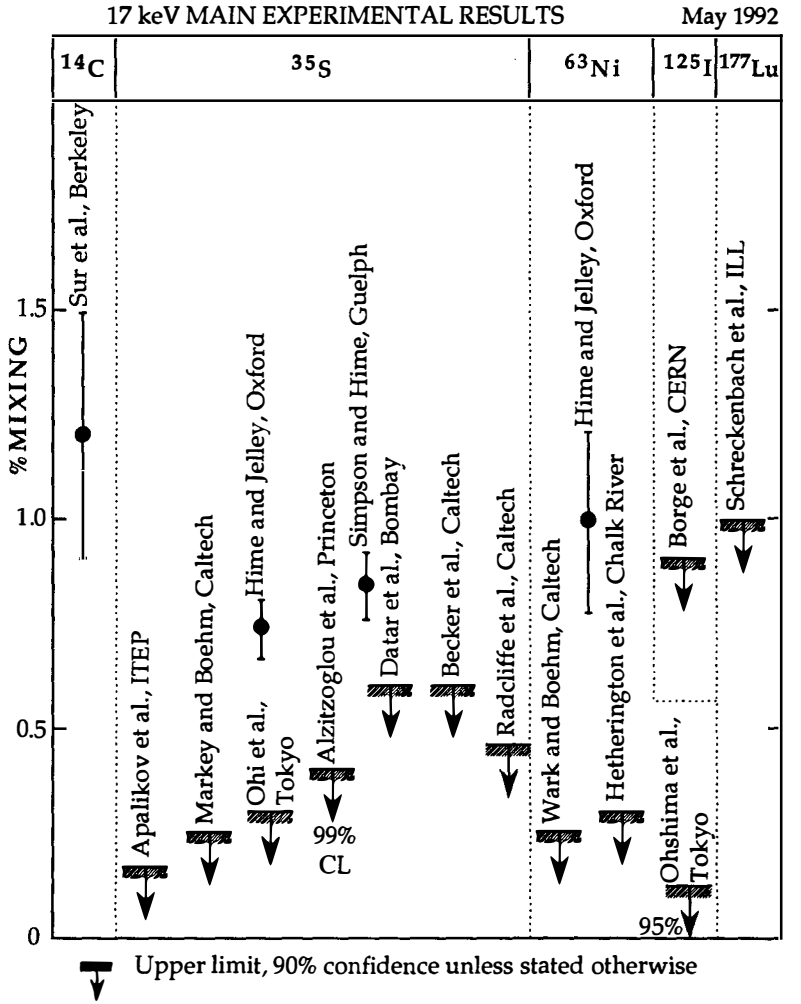


Fig. 2: Main experimental results on % mixing of 17 keV neutrinos. In addition DiGregorio et al. [31] find a best fit of 13.8 keV with 0.8% mixing and exclude a 17.2 keV neutrino having 0.85% mixing with 96% confidence

References

1. D.R.O. Morrison, Inv. talk at *Joint Lepton-Photon Symposium and Europhysics Conference on High Energy Physics*, Geneva, July 1991 and preprint CERN-PPE/91-140.
2. J.J. Simpson, *Phys. Rev. Lett.* **54** (1985) 1891.
3. T. Alzitoglou et al., *Phys. Rev. Lett.* **55** (1985) 799.
4. T. Ohi et al., *Phys. Lett.* **160B** (1985) 322.
5. A.M. Apalikov et al., *Pis'ma Zh. Eksp. Teor. Fiz.* **42** (1985) 233 [*JETP Lett.* **42** (1985) 289].
6. V.M. Datar et al., *Nature* **318** (1986) 547.
7. J. Markey and F. Boehm, *Phys. Rev.* **C32** (1985) 2215.
8. D. Wark and F. Boehm, *Proc. Symp. on Nuclear Beta Decays and Neutrinos*, Osaka, 1986, Eds. T. Kotani et al. (World Scientific, Singapore, 1986).
9. D.W. Hetherington et al., *Phys. Rev.* **C36** (1987) 1504.
10. M.J.G. Borge et al., *Phys. Scr.* **34** (1986) 591.
11. J. Lindhard and P. G. Hansen, *Phys. Rev. Lett.* **57** (1986) 965.
12. E.G. Druckarev and M.I. Strikman, *Phys. Lett.* **186** (1987) 1.
13. S.E. Koonin, *Nature* **354** (1991) 468-469.
14. J.J. Simpson, *Phys. Lett.* **174B** (1986) 113, and *Proc. Moriond Workshop on Massive Neutrinos in Particle Physics and Astrophysics*, Les Arcs, 1986, Eds. O. Fackler and J. Tran Thanh Van (Editions Frontières, Gif-sur-Yvette, 1986), p. 565.
15. T. Ohi, Thesis, Univ. of Tokyo (1986), (priv. comm. from T. Yamazaki).
16. D.W. Hetherington and P.L. Graham (private commumication).
17. V. A. Lyubimov, (private commumication).
18. J.J. Simpson and A. Hime, *Phys. Rev.* **D39** (1989) 1825.
19. A. Hime and N. Jelley, *Phys. Lett.* **257B** (1991) 441.
20. B. Sur et al., *Phys. Rev. Lett.* **66** (1991) 2444.
21. I. Zlimen et al., *Phys. Rev. Lett.* **67** (1991) 560.
22. H.-W. Becker et al., Caltech preprint 63-605 (1991).
23. J.J. Simpson, *Phys. Lett.* **269B** (1991) 454.
24. A. Hime and N. Jelley, Oxford report OUNP-91-21 (1991).
25. B. Sur et al. , *Proc. Workshop on the 17 keV Neutrino Questions*, Berkeley, 18-20 Dec. 1991 (transparencies, Center for Particle Astrophysics, Berkeley, 1992), p. 219.
26. E. Norman, (private commumication).
27. T. Radcliffe et al., This conference, and Caltech report Calt-63-629 (1992).
28. K. Schreckenbach (private commumication), and S. Schoenert same Proc. as ref. 25, p. 27.
29. D. Wark et al., same Proc. as ref. 25, p. 72 and private communication.
30. T. Ohshima et al., (private commumication).
31. D.E. DiGregorio et al., same Proc. as ref. 25, p. 347.
32. D.R.O. Morrison, *Usp. Fiz. Nauk* **156** (1988) 719 and *Proc. XVIII Int. Symposium on Multiparticle Dynamics*, Tashkent, 1987, eds. I. Dremin and K. Gulamov, p. 755, and preprint CERN/EP 88-933.
33. A. Hime, Thesis, Oxford Univ. (1991).
34. L. Piilonen and A. Abashian, this conference and private communication.

NEW LIMITS ON THE 17 keV NEUTRINO

T. J. Radcliffe*, M. Chen, D. A. Imel**, H. Henrikson and F. Boehm
California Institute of Technology
Pasadena, California 91125

ABSTRACT

We have performed a new measurement of the beta decay spectrum of ^{35}S, using the Caltech double-focusing, iron-free beta spectrometer. Several important sources of systematic error present in previous magnetic spectrometer experiments have been identified and controlled. A simulation code for the spectrometer has been written, and initial runs have given insight into aspects of the spectrometer shape factor. Most importantly, we have been able to induce a kink in the beta decay spectrum of ^{35}S by artificial means, which allows us to demonstrate "positively and directly by experiment" that we are sensitive to features of the kind for which we are searching. No evidence for a 0.85% admixture of a 17 keV neutrino is seen at the 99.9% confidence level.

---

* present address: Medical Physics,MCTRF,100 Olivia St. Winnipeg, Manitoba, Canada;
** present address: JPL 300-319 Pasadena, CA 91109

Introduction

Several studies[1-6] of the shapes of beta spectra using magnetic spectrometers were conducted following Simpson's claim[7], in 1985, to have found evidence for a 17 keV neutrino in the beta spectrum of ^{3}H. The Caltech studies of the 17 keV neutrino with the $\sqrt{2\pi}$ spectrometer began with the work of Markey and Boehm[2] in 1985 finding no evidence for the admixture of a 17 keV neutrino to the usual light neutrino in the beta decay of ^{35}S). This result agrees well with the other spectrometer experiments quoted above confirming the absence of a heavy neutrino in the mass range of 5 - 50 keV. (For a brief review see Boehm[8] and Boehm and Vogel[9]). Encouraged by the renewed announcements by Hime and Simpson[10] and Hime and Jelley[11] we have repeated our experiment with improved techniques. An early account was presented at last year's Moriond Workshop[6].

The present report describes results from recent studies of the beta decay spectrum of ^{35}S employing direct measurements of the magnetic field. Careful attention was paid to account for possible systematic effects.

Experimental Details and Results

Spectrometer and Detector. The 35 cm radius, iron-free double-focusing $\sqrt{2\pi}$ magnetic spectrometer at Caltech (momentum resolution 0.27%) employed at its focus a silicon surface barrier detector cooled by Peltier elements to approximately 5 C. Its energy resolution (figure 1) between 120 - 160 keV is $\approx$ 4 keV FWHM. The detector has an active surface layer of 300 μm and an area of 4 mm x 25 mm. A pulser was used to monitor the stability of the electronics.

Source Preparation. The spectrometer was calibrated using a 100 μCi ^{57}Co source vapour deposited onto a 0.9 micron thick mylar backing. Scattering from the backing is less than 1%. Two ^{35}S sources were made using a technique that is a combination of vapour deposition and aqueous deposition. A layer of barium was first deposited on the gold-coated mylar through a mask in the shape of the desired source (in this case, 2 mm x 20 mm). An aqueous solution of ammonium sulfate containing ^{35}S was then brought into contact with the barium coated region of the foil. After the solution had been in contact with the surface for about an hour, it was drawn off using a micro-pipette. This process, with a yield of about 40%, allowed us to make sources of 3 and 7 mCi strengths.

Stability. We directly monitored the magnetic field with two fluxgate magnetometers. A Bartington Instruments magnetometer and high field probe were used to measure the vertical component of the spectrometer field. A Walker Scientific magnetometer monitored the horizontal component of the ambient field, outside of the spectrometer coils. The stability of the spectrometer field was better than 30 ppm over fifteen minutes.

We observed a secular variation in the ambient field consistent with typical fluctuations in the Earth's magnetic field. We also measured irregular but large ($\approx$ 1% of the Earth's field) variations whose apparent origin is local to the building. These field fluctuations are almost certainly the cause of the non-statistical scattering observed by Becker et al.[6].

Calibration and Detector Response. We used the ^{57}Co conversion lines to determine the spectrometer calibration and its response function. The conversion electron spectrum is shown in figure 2.

Data. Two data sets were acquired, both using the 3 mCi source. In the first (run A), the beta spectrum was measured from 130 keV to the endpoint. In the second data set (run B), a 10 keV region around the kink for a 17 keV neutrino was scanned with high statistics. Each data point was corrected for source half-life and dead-time. Background counts in the detector, measured with the spectrometer set above the ^{35}S endpoint, were subtracted from each data point.

The counts observed at a given momentum setting were taken by summing the counts from the detector spectrum, including those in the backscattered tail, down to a fixed fraction (20%) of the peak energy. Published parameterizations[12] of surface barrier detector response indicate that the fraction of counts excluded by this method has less than 0.3% energy dependence over the region of interest. This procedure has significantly reduced the size of the required shape corrections to the spectra.

Analysis and Results. The corrected binned data were analyzed using various fits to the allowed beta decay spectral shape for ^{35}S. That shape included Fermi function and radiative corrections provided by Petr Vogel. The response function of the spectrometer, determined from the ^{57}Co K-conversion lines, was convoluted with the theoretical spectrum and the result used to fit the two data sets.

Data run A, from 130 - 165 keV, gives a fit with χ^2/DOF = 66.1/50, when the endpoint and normalization are allowed to vary as free parameters and the neutrino admixture and shape corrections are fixed at zero (figure 3a). A significantly better fit (χ^2/DOF = 36.5/48), according to the F-test, is obtained by allowing linear and quadratic shape factor corrections as free parameters. It should be noted that the magnitude of these corrections over the energy range of interest is less than 1%. Data run A is shown in figure 3b, plotted as the fitted residual to the spectral shape without heavy neutrino emission (0% admixture). The results of various fits which attempt to include the emission of a 17 keV neutrino are shown in Table 1. In these fits, the endpoint, normalization, linear and quadratic shape factor terms were all allowed to vary freely. These results place an upper limit on the admixture of a 17 keV neutrino of 0.45% at the 90% confidence level and rule out a 0.85% admixture at the 99.3% confidence level.

Data run B features a high statistics study of the beta spectrum around 420 keV/c, the momentum threshold for the emission of a 17 keV neutrino. Fitting this data without shape corrections for a 0% admixture results in a χ^2/DOF = 27.9/30 with a fitted endpoint of 167.483 ± 0.007 keV. Over this limited region, no shape correction is required for a good fit. In attempting to include a 0.85% admixture, the data give a poor fit of χ^2/DOF = 57.5/30, representing a better than 5 σ exclusion of this admixture. Nonetheless, because knowledge of the spectrometer response was shown to be incomplete in run A by the requirement of a shape correction, one can admit a linear shape term as a free parameter in the fit to this data. The fit to 0% admixture is slightly improved to χ^2/DOF = 25.7/29, with an endpoint of 167.547 ± 0.043 keV, while the fit to a 0.85% admixture gives χ^2/DOF = 36.4/29 and a linear correction three times as large as is required in the previous fit. The results of these fits are shown in Table 2. They place an upper limit on the 17 keV neutrino admixture of 0.42% at the 90% confidence level and in particular rule out a 0.85% admixture at the 99.9% confidence level. Figure 4a shows data run B plotted as the residual to the fitted spectrum without heavy neutrino emission. Figure 4b shows the same data with fits normalized above the threshold for 17 keV neutrino emission. One sees in this plot that the data can indeed admit a small linear correction, but that it does not easily accomodate the kink from a heavy neutrino.

Combining the statistical exclusion from the two independent runs and neglecting correlated systematic errors, we find that our data give a best fit admixture to a 17 keV neutrino of -0.21% ± 0.23%, consistent with no mixing. Thus, we can place a combined 90% CL upper limit on the admixture of a 17 keV neutrino to be: $U^2 < 0.27\%$. For all the quoted exclusions, the prescription of the Particle Data Group[13] for estimating parameters constrained to lie within a bounded physical region (positive U^2) was used.

Synthetic Kink. In order to demonstrate that spectrometer experiments can detect a kink such as that produced by a hypothetical 17 keV neutrino we have carried out the following auxiliary experiment.

We have introduced a kink of known size and energy into the spectrum by masking off roughly 10% of the source with a 17 micron aluminum foil, which causes a minimum energy loss of 16 keV for 150 keV electrons passing through it. Taking the shape as estimated by an EGS4 calculation of electron transport through the foil, we predict a spectral distortion induced by this foil that resembles heavy neutrino emission with a best fit admixture of about 1%. This is lower than the fraction of the source area covered because the spectral distortion created by energy loss is quite different from the kink expected for a heavy neutrino, with which we fit. In addition, some fraction of the electrons incident on the foil are scattered out of the transmission aperture of the spectrometer or backscattered towards the source, reducing the magnitude of the distortion. Analyzing a spectrum taken over 48 hours, with the 7 mCi source and the aluminum foil,

gave a best fit heavy neutrino fraction of 2.5%, with a mass of 15.6 keV, demonstrating the sensitivity of the spectrometer (figure 5). It is expected that a more detailed calculation, which includes the spectrometer transmission of scattered electrons, will improve the agreement.

Conclusion

Motivated by criticism of earlier work, we have made a precise measurement of the beta decay spectrum of ^{35}S. We have identified systematic effects that may have influenced the limits on the 17 keV neutrino from the Caltech magnetic spectrometer reported previously. We have measured the spectrometer magnetic field directly for the first time, and have demonstrated positively and directly by experiment that we are sensitive to spectral features similar to the kink caused by a 17 keV neutrino.

From two independent data sets, we rule out a 0.85% admixture for the 17 keV neutrino at the 99.3% and 99.9% confidence level, respectively.

The origin of the feature so clearly seen by Simpson and by Hime et al. in the beta spectra of ^{3}H and ^{35}S, with both internal and external sources, with different detector materials, and with careful control over systematic effects, remains to be explained.

We wish to thank Petr Vogel for generous advice, and the USDOE for support.

References

1) T. Altzitzoglou et al., Phys. Rev. Lett. 55, 799 (1985).
2) J. Markey and F. Boehm, Phys. Rev. C32, 2215 (1985).
3) D. Hetherington et al., Phys. Rev. C36, 1504 (1987).
4) D. Wark and F. Boehm, In Nuclear Beta Decays and Neutrino, Proc. Osaka Symposium 1986, Ed. T Kotani et al., World Scientific, 391 (1986).
5) A. Apalikov et al., JETP Lett., 42, 289 (1985).
6) H.-W. Becker et al., Proceedings of the Eleventh Moriond Workshop, Les-Arcs, Savoie, France, Jan. 26 - Feb. 2 (1991), Ed. O. Fackler and Tran Thanh Van, Paris: Editions Frontiers; and Calt 63-605, Caltech Report, unpublished (1991).
7) J.J. Simpson, Phys. Rev. Lett., 54, 1891 (1985).
8) F. Boehm, in The Vancouver Meeting, Particles and Fields 91, p. 956, Ed. D. Axen et al. Singapore: World Scientific.
9) F. Boehm and P. Vogel, Physics of Massive Neutrinos, Second Edition, Cambridge University Press, Chapter 4 (1992).
10) A. Hime and J.J. Simpson, Phys. Rev. D39, 1837 (1989).
11) A. Hime and N. Jelley, Phys. Lett. B257, 441 (1991).
12) A. Damkjaer, Nucl. Instrum. Methods 200, 377 (1982).
13) Particle Data Group, Phys. Lett. B239, III.35 (1990).

	M [keV]	$u^2 \times 100$	Q [keV]	χ^2/ν
without shape factor	-	0.0	167.610 ± 0.007	66.1/50
with shape factor	-	0.0	167.563 ± 0.030	36.5/48
	17	0.85	167.512 ± 0.030	45.2/48
	17	0.4	167.539 ± 0.029	39.7/48
	17	-0.58 ± 0.45	167.598 ± 0.041	34.8/47
	18	-0.54 ± 0.45	167.600 ± 0.041	35.0/47
	16	-0.65 ± 0.48	167.595 ± 0.038	34.5/47

Table 1: Results of Fitting to Data Run A

	M [keV]	$u^2 \times 100$	Q [keV]	χ^2/ν
without shape factor	-	0.0	167.483 ± 0.007	27.9/30
	17	0.85	167.562 ± 0.007	57.5/30
with shape factor	-	0.0	167.547 ± 0.043	25.7/29
	17	0.85	167.368 ± 0.042	36.4/29
	17	0.4	167.463 ± 0.043	28.3/29
normalized above	-	0.0	167.498	39.1/32
kink	17	0.85	167.498	236.2/32

Table 2: Results of Fitting to Data Run B

Figures

Figure 1. Typical electron spectrum in the silicon surface barrier detector for mono-energetic incident electrons. For this spectrum, an electron momemtum of p = 420 keV/c was selected by the spectrometer. Approximately 17% of the incident electrons are backscattered, depositing less than their full energy in the detector.

Figure 2. Momentum calibration of the spectrometer from the directly measured magnetic field, using a ^{57}Co internal conversion electron source.

Figure 3. Data run A (a) fit to a beta spectrum without massive neutrino emission before shape correction. (b) after fitting with linear and quadratic shape factor terms. The solid curve represents the hypothetical shape of a best fit experimental spectrum with a 0.85% admixture, 17 keV neutrino.

Figure 4. Data run B (a) fit to a beta spectrum without massive neutrino emission. The solid curve again represents the best fit experimental spectrum with 0.85% mixing probability to a 17 keV neutrino. (b) fit with the normalization of the experimental spectrum only to data above 421 keV/c. The solid curve represents a spectrum with a 0.85% admixture, 17 keV neutrino.

Figure 5. Synthetic kink induced in the beta spectrum by a 17 μm thick aluminum foil. The data clearly show the spectral distortion induced by the foil. The solid curve superimposed on the plot is the hypothetical experimental spectrum from a 15.6 keV neutrino with a 2.5% admixture.

Figure 1

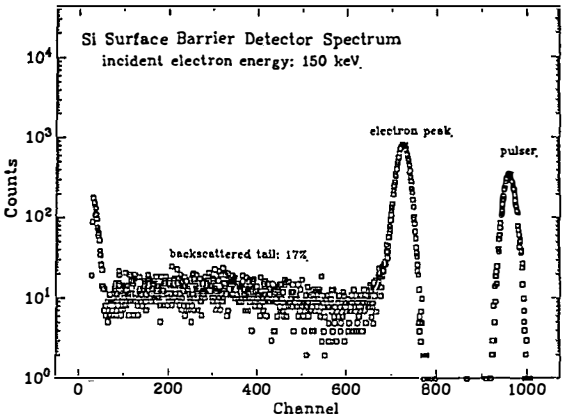


Figure 2

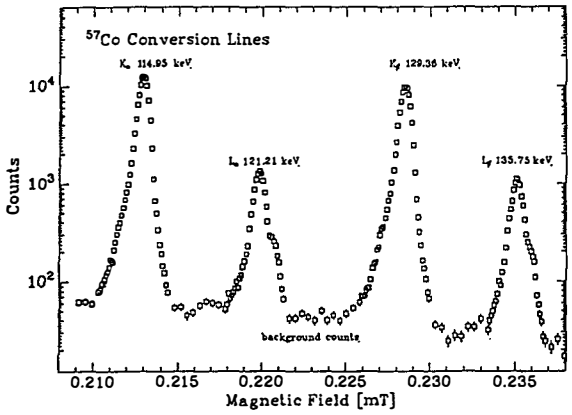


Figure 3a - Data Run A, before shape factor

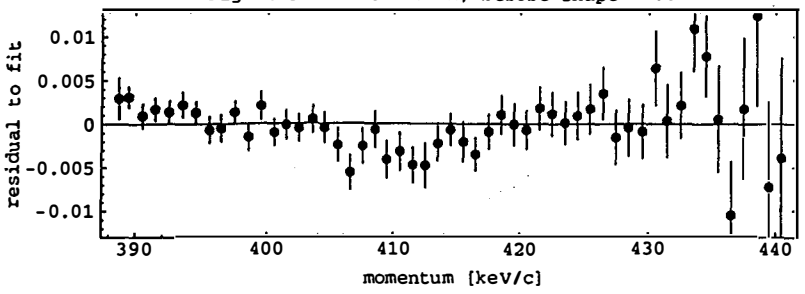


Figure 3b - Data Run A

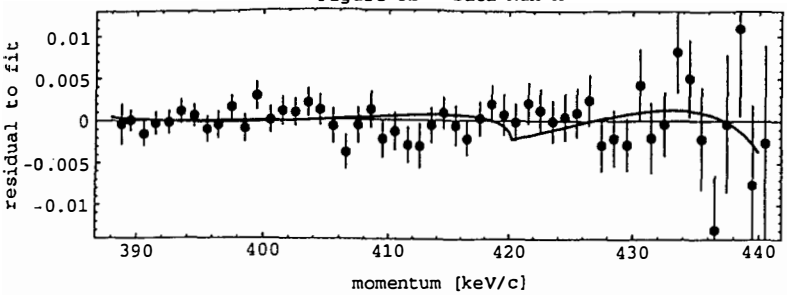


Figure 4a - Data Run B

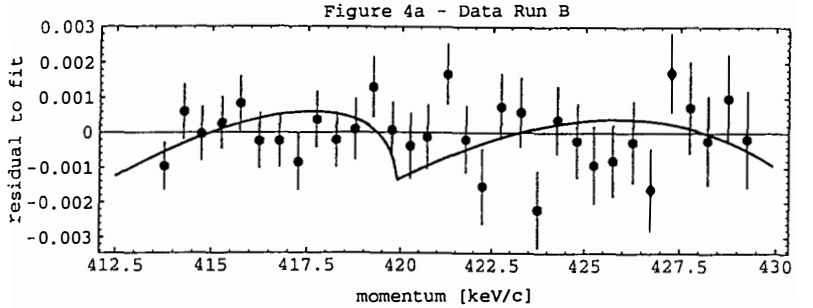


Figure 4b - Data Run B, normalized above kink

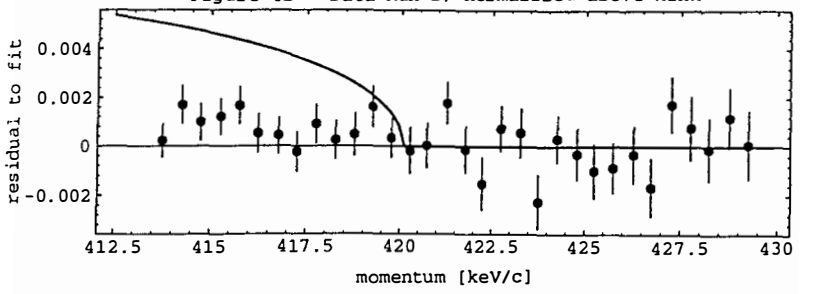
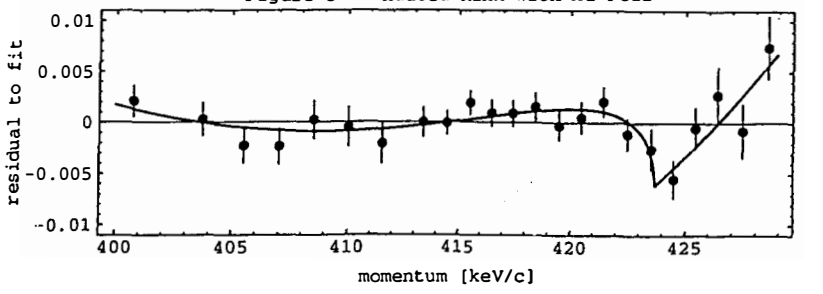
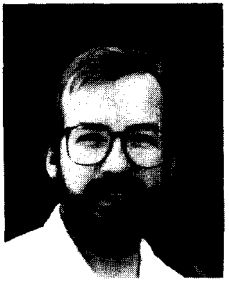


Figure 5 - Induced Kink with Al Foil



On the Strength of the Evidence for the 17 keV Neutrino

Leo Piilonen and Alexander Abashian
Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061



We have performed an independent reanalysis of the 1991 ^{35}S beta decay experiments at Oxford that were asserted to provide evidence for a 17 keV component of the electron antineutrino. Although the form of the response function used in the original analysis did not properly account for all of the processes in the Oxford spectrometer, the distortion seen in the original analysis persists even when these effects are incorporated, although it is no longer consistent with a 17 keV neutrino.

Existence of a heavy component of the $\bar{\nu}_e$, with mass $17\,\mathrm{keV}/c^2$, was first claimed by Simpson[1] in 1985 as necessary to account for an excess of low-energy electrons in measurements of the beta decay spectrum of ^{3}H implanted in a Si(Li) detector. Many experiments since then have denied[2–9] or confirmed[10–13] the existence of this 17 keV neutrino. A lengthy review of these measurements was recently completed by Hime.[14]

We have reanalyzed the 1991 ^{35}S data of Hime and Jelley[12] to see if the distortion in the beta spectrum can be accounted for by a more mundane effect than a 17 keV component of $\bar{\nu}_e$. Since we did not have access to the raw data, we transcribed the relative deviations—(DATA-FIT)/FIT—for runs 1 and 2 from the published graphs. Here, DATA refers to the counts in each 606 eV bin after subtraction of ambient background and unrejected-pileup events and a subsequent shift up by eV_B to account for the voltage bias $V_B = 1.529\,\mathrm{kV}$ on the detector contact, while FIT refers to the massless-neutrino beta spectrum with $Q = 167.013$ (167.016) keV for run 1 (2), convolved with the spectrometer response function described in detail in Hime's thesis.[15]

Figures 1a–c show the relative deviations of the transcribed data after shifting the Q-value of the FIT spectrum up to 167.072 keV in order to move to zero the relative deviations near the endpoint. The solid curves in these figures, as in Hime and Jelley's paper, represent the relative deviations that would be expected if a synthetic data set containing a 0.9% branch for a 17 keV component of $\bar{\nu}_e$ were compared to this new FIT spectrum. The curve has been deliberately left out of Figure 1c to avoid guiding the reader's eye. While the combined data is certainly consistent with the 17 keV hypothesis, there is simply not enough statistical precision near 150 keV to see an unmistakable kink.

If some other more mundane explanation could be found for the distortion in these data, the principle of Occam's razor would clearly prefer it to the two-component neutrino hypothesis. In the following sections, we re-examine the methodology used by Hime and Jelley in their analysis of the ^{35}S data to explore the possibility of an alternate hypothesis that could account for their measurements.

Ambient Background: The energy spectrum $N_B(E)$ of the ambient background was measured by Hime and Jelley in separate runs before and after the ^{35}S data. This 1.3 Hz background was *not* due to cosmic rays, whose rate in a detector of $\sim 1\,\mathrm{cm}^2$ projected area was only 0.02 Hz and whose mean, most probable and lowest energy losses in 0.16 cm of silicon were 620, 325 and 240 keV, respectively, well above the range of interest. We note also that the ambient background rate in the later ^{63}Ni beta spectrum measurement[16]

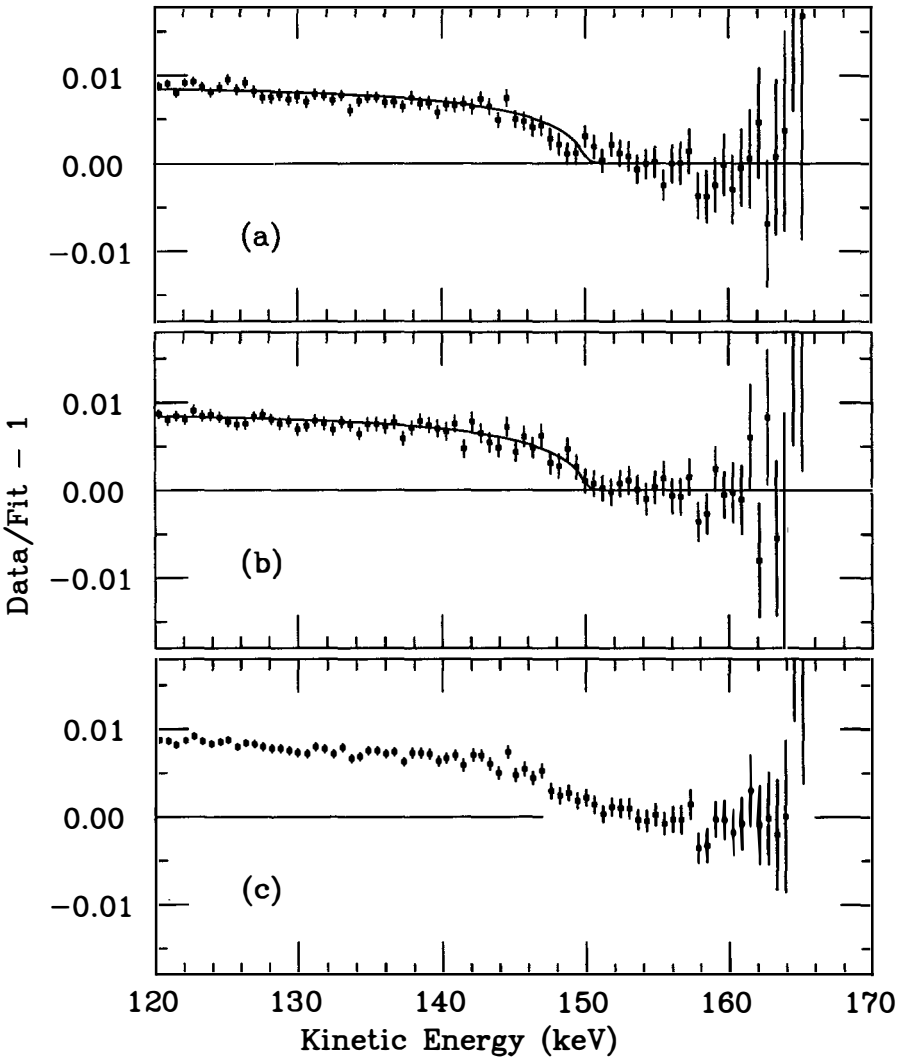


FIGURE 1. Relative residuals of Hime and Jelley's DATA from (a) run 1, (b) run 2, and (c) combined, when compared to a massless-neutrino FIT spectrum with $Q = 167.072$ keV. The solid curves are what would be expected for a two-component neutrino with $m_2 = 17$ keV and $\sin^2\theta = 0.009$ were analyzed in this way.

with the same spectrometer fell by a factor of three to 0.42 Hz. This background must have been due to some local (radioactive?) source whose intensity varied with time. The energy spectrum, from the ^{63}Ni preprint, is adequately represented between 25 and 120 keV by a falling exponential with a scale of $\lambda = 0.0189\,\mathrm{keV}^{-1}$ and an extrapolated intensity in a 151.5 eV bin near 150 keV of ~ 178 counts for ^{35}S run 1 (or 0.14% of the signal). The ^{35}S signal falls below this ambient background at energies above 164.9 keV; thus, an over- or under-subtraction of the ambient background from the raw data would have distorted the **DATA** spectrum only within ~ 2 keV of the endpoint.

Unrejected Pileup: The unrejected-pileup spectrum $N_p(E)$—due to random coincidences within the 340 ns resolving time of the pileup inspection logic—was derived by Hime and Jelley from the raw ^{35}S data by convolving the beta spectrum $N_{raw}(E)$ with itself between $E_{min} = 1.4$ keV and the biased endpoint $Q_B \equiv Q - V_B \sim 166.5$ keV:

$$N_p(E) = \int_{\max(E_{min},E-Q_B)}^{\min(E,Q_B)} N_{raw}(\varepsilon) N_{raw}(E-\varepsilon)\,d\varepsilon\,. \qquad (1)$$

This spectrum was then rescaled, along with the ambient background spectrum, so that their sum matched the raw data above the ^{35}S endpoint, and then subtracted from the raw data to yield the net ^{35}S beta spectrum (denoted by **DATA** in the introduction). The ^{35}S signal falls below this unrejected-pileup background for energies above 163.5 keV; thus, an over- or under-subtraction of the unrejected-pileup from the raw data would have distorted the data only within ~ 3.5 keV of the endpoint. The systematic errors associated with not using zero for the lower limit of the convolution (which would have changed the shape of the pileup spectrum below $Q_B + E_{min}$) and not iterating the convolution procedure after background subtraction were found to be negligible—assuming there were no high-intensity spikes in the raw data below E_{min}.

There is another potentially serious problem related to pulse pileup, however. From the discussion in Hime's thesis, the data acquisition system was dead for about 25.7 μs after the arrival of an event in run 1 (about 10% deadtime at 4100 Hz), whereas the pileup inspector was active only for the first 18 μs. There was thus a 3.1% (2.3%) probability in run 1 (2) that a second pulse arrived in the last 7.7 μs after the pileup inspector turned off but before the ADC was ready to begin a new conversion. It is not clear if this second pulse was discarded, recorded correctly, or recorded with a defective (reduced) pulse height. The last case would result in the appearance in the **DATA** spectrum of a ghost of the ^{35}S beta spectrum at lower energy characterized by the typical pulse height defect. The probability of this process was only 0.3% (1%) in the ^{57}Co (^{109}Cd) conversion line sources because of

their reduced strength, and non-existent in the ^{63}Ni beta spectrum because of the longer pileup inspection gate.

Theoretical Spectrum: We have compared Hime and Jelley's method for computing the screened Fermi function with a direct calculation,[17] as prescribed by Behrens and Janecke,[18] that solves for the relativistic electron radial wave function in the presence of a finite-size nucleus and an atomic electron cloud parameterized by a Hartree-Fock potential. They agree to within $\pm 8 \times 10^{-5}$ between 100 to 170 keV, so this cannot account for the distortion that they saw.

In their 1991 paper, Hime and Jelley neglected the radiative corrections[19] to the beta decay process in calculating their FIT spectra. Hime covered this point later in his thesis, where he said that including them in the BEST FIT reduced $\sin^2\theta$ from 0.0084 to 0.0078. The radiative corrections modify the theoretical beta spectrum by about 1.5% between 120 and 165 keV, and at first blush would be expected to cause an energy-dependent distortion of comparable magnitude in the relative deviations (DATA-FIT)/FIT if they were left out of the FIT spectrum, and by implication would have reduced $\sin^2\theta$ by about one-third when properly included in the FIT. However, this energy-dependent distortion is fairly well masked in the interval 120–164 keV if the endpoint of the FIT is adjusted downward by just 15 eV—the change in the phase space factor then compensates for the neglect of the radiative corrections except within 3 keV of the endpoint. This conclusion has been verified in fits of many synthetic data sets having statistics comparable to Hime and Jelley's runs, in which the data was generated with radiative corrections then compared to a fit spectrum that neglected them. Thus, for Hime to have seen *any* change in the best fit value for $\sin^2\theta$ when radiative corrections were included in the BEST FIT is remarkable, and was most likely due to the fitting procedure's attempt to better match the somewhat questionable data near the endpoint.

Response Function: We performed a detailed EGS4[20] simulation of the Hime and Jelley spectrometer to get an accurate determination of all of the contributions to the electron response function as well as their dependence on energy. In this simulation, we made two modifications to EGS4 that permitted us to study the passage of electrons through extremely thin regions, and a third modification that allowed us to model the transport of soft x-rays in the gold contact and silicon of the detector.

First, the EGS4 multiple scattering routine for electrons was replaced with a routine based on Mott elastic scattering from screened nuclei. This routine duplicated the Molière distribution for a large number of scatterings but also generated the correct angular dis-

tribution for single or plural scattering.

Second, the *average* restricted electron energy loss due to ionization in aluminum, copper, gold, mylar, or silicon was replaced with a routine based on the work of Talman[21] that simulated the inelastic scattering from the electrons in the atomic shells of each element. This routine's energy loss distribution had the same average as the value used by EGS4, while giving a more realistic spectrum of the individual energy losses suffered by electrons in traversing sub-micron distances.

Finally, in the detector only, our energy loss routine and the EGS4 routines for Møller scattering and the photoelectric effect were modified to permit the explicit creation and propagation of K, L and M x-rays above 1 keV[22] so that we could monitor the recapture of gold x-rays in the active volume and the escape of silicon K x-rays.

We simulated the histories of electrons at six different energies for both geometries of the spectrometer. We threw 2.5×10^7 events in each case, or 3×10^8 in all, of which $\sim 3.8 \times 10^5$ entered the Si(Li) detector. From these entering events, we were able to deduce the probabilities of occurrence for the energy-lowering scattering processes in the spectrometer, as well as the distributions of post-scattering energies for electrons that suffered any one of these processes. The energy dependence of the probabilities determined what fraction of electrons were removed from the full-energy peak of the response function at each energy in the ^{35}S beta spectrum, while the distribution of scattered energies determined where (if anywhere) these electrons reappeared in the beta spectrum. Hime and Jelley considered only ionization in the detector's gold contact and back-diffusion from the Si(Li) as the relevant scattering processes. Our simulation indicated a more complicated response.

The acceptance of the spectrometer for unscattered electrons was 0.001593 ± 0.000003 (0.000941 ± 0.000003) for the geometry of run 1 (2), independent of energy. The energy-dependent acceptances, relative to the unscattered acceptance, after scattering from the source backing (alone or in combination with another structure), from either the source or detector aperture (or both), and from the anti-scatter baffle (alone or in combination with either aperture) are plotted in Figure 2. These events are what constitute the bulk of the low-energy tail seen in Hime and Jelley's ^{57}Co response function curves, rather than events suffering large ionization energy loss in the detector contact. Although the individual components have a strong energy dependence, it is remarkable that the overall detection probability for any scattered electrons is very nearly constant over the range 120–170 keV. This may explain the success of Hime and Jelley's use of a "flat-tail" response function in analyzing their run 2 data.

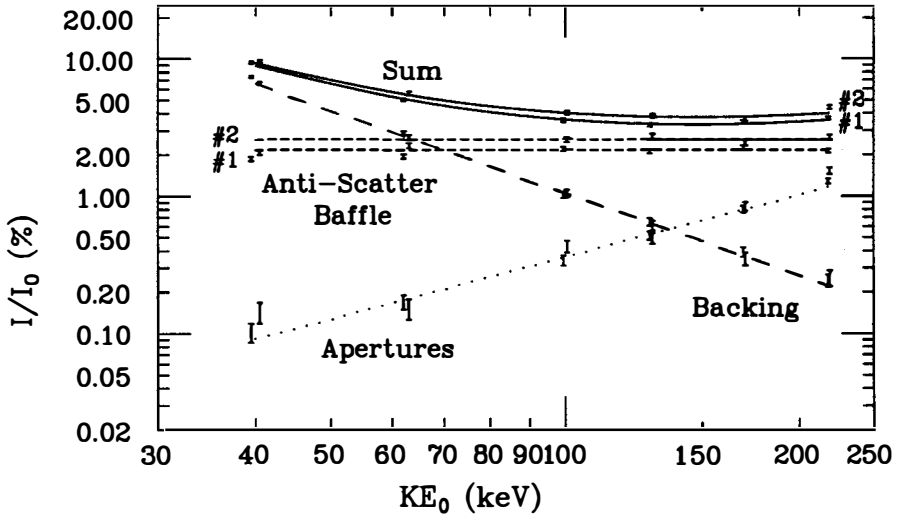


FIGURE 2. The acceptances for electrons after scattering from the spectrometer structures, relative to the constant unscattered-electron acceptance. The fit curves are: $105.34/T_\circ^2$ (source backing); $3.477\times10^{-6}\cdot T_\circ^{3/2}$ (apertures); 0.0216 and 0.0260 (anti-scatter baffle for runs 1 and 2). $T_\circ$ is the pre-scattered kinetic energy in keV.

The normalized scattered-energy spectra following scattering from the source, aperture(s), and baffle can be parameterized by

$$N_S(T,T_\circ) = \frac{a_S}{T_\circ}e^{-\lambda_S x} + \frac{b_S}{T_\circ}\delta(x) \tag{2}$$

$$N_A(T,T_\circ) = \frac{a_A}{T_\circ}e^{-\lambda_A x} \tag{3}$$

$$N_B(T,T_\circ) = \frac{1}{T_\circ}e^{-\lambda_B x}\left(a_B x + b_B x^2 + c_B x^3\right) \tag{4}$$

with $T_\circ$ and T denoting the pre- and post-scattering energies in keV and $x \equiv 1 - T/T_\circ$ the fractional energy loss. The coefficients are $a_S = 9.109$, $b_S = 0.1666$, $\lambda_S = 10.93$, $a_A = 4.351$, $\lambda_A = 4.291$, $a_B = 284.2$, $b_B = -1533$, $c_B = 3511$, and $\lambda_B = 13.15$.

The probability that an electron of velocity βc incident on the detector's gold contact would backscatter elastically out of the detector was found to be $0.000147(1-\beta^2)/\beta^4$ for a detector surface of 250 Å of gold and 0.1 μm of dead silicon.[23] The systematic effect of ignoring this inefficiency in Hime and Jelley's ^{35}S analysis was small, since the backscatter probability was less than 0.2% for energies above 120 keV. However, their ^{63}Ni analysis should have included this effect.

Electrons passing through the detector surface into the active volume nearly always suffered some ionization energy loss, although this process typically involved scattering from a valence electron. The post-ionization energy distribution could be written as

$$D_{ion}(T, T_o) = [1 - P_{ion}(T_o)]\,\delta(T - T_o) + P_{ion}(T_o) I_{ion}(T, T_o) \tag{5}$$

where the simulation gave

$$P_{ion}(T_o) = A_0 + \frac{A_1}{T_o} + \frac{A_2}{T_o^2} + \frac{A_3}{T_o^3} \tag{6}$$

$$I_{ion}(T, T_o) = \left(\frac{B}{T_o - T}\right)^2 \left[a_I + b_I \left(\frac{B}{T_o - T}\right) + c_I \left(\frac{B}{T_o - T}\right)^2\right] \quad [T_o - T \geq B] \tag{7}$$

with T_o and T in keV, $A_0 = 0.4982$, $A_1 = 58.85$, $A_2 = -2564$ and $A_3 = 39311$, $B = 0.050\,\text{keV}$, $a_I = 3.275\,\text{keV}^{-1}$, $b_I = 29.147$ and $c_I = -28.40$, The tail of this ionization distribution fell off much more rapidly than in the corresponding distribution used by Hime and Jelley.

We used the parameterization from Hime's thesis for the distribution of energies deposited in the Si(Li) for electrons that back-diffused out of the detector (our simulation gave the same distribution). However, we used the empirical formula of Tabata[24] for the probability of back-diffusion, which varied by 2% for incident energies between 120 and 167 keV.

Finally, we used Hime and Jelley's energy dependence for the width of the gaussian that was used to smear the energy deposited in the Si(Li) into the energy recorded by the ADC. The net response to monoenergetic electrons predicted by our simulation is in excellent agreement with the ^{57}Co and ^{109}Cd conversion line measurements of Hime and Jelley.

When we reanalyze their **DATA** with a **FIT** massless-neutrino beta spectrum convolved with *our* form for the response function detailed above, we still find a serious disagreement between the combined **DATA** and this hypothesis, as shown in Figure 3 where the Q-value of the **FIT** spectrum was chosen by using only those points between 158 and 164 keV. However, there is even less evidence of a kink at 150 keV than there was in Figure 1c. (We speculate that the form of the response function used by Hime and Jelley was overly broad on the low-energy side of the full-energy peak and that this caused a serious distortion in their **FIT** spectrum within a few keV of the endpoint.) The relative residuals appear to fall

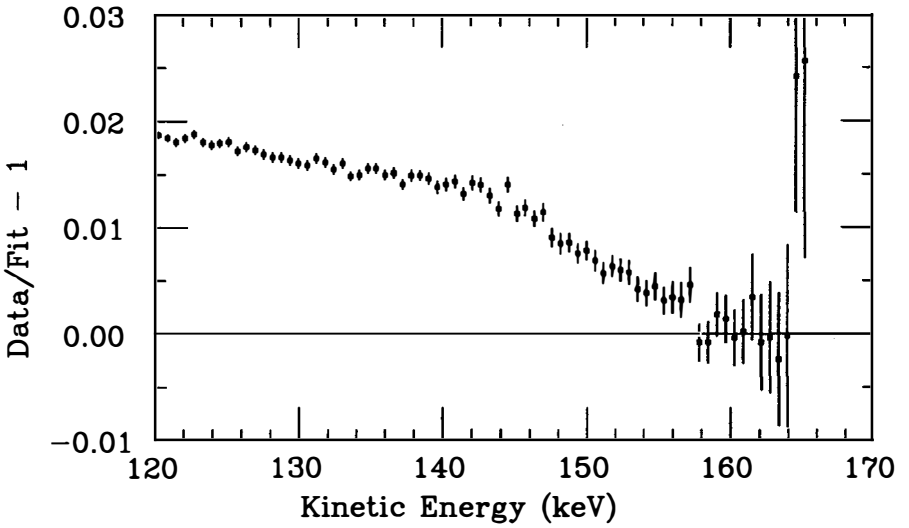


FIGURE 3. Relative residuals of Hime and Jelley's combined DATA when compared to a massless-neutrino FIT spectrum with $Q = 167.140\,\mathrm{keV}$ (to match the points between 158 and 164 keV) that was convolved with *our* response function.

on two separate straight lines that cross at 142 keV. It would be interesting to see if this is an artifact of some small non-ideal behaviour, such as a differential nonlinearity, in the data acquisition system.

Note that the quality of the fit can be improved somewhat by including linear and quadratic shape factors in the massless-neutrino FIT of the form typically employed in magnetic spectrometer experiments. In doing so, however, one is hard-pressed to find a suitable explanation for the need to include such terms in the fit.

Conclusion: We agree with Hime and Jelley that there is a serious distortion in their ^{35}S data, though we cannot pinpoint any definite cause for it. We believe that if the original data is reanalyzed by Hime and Jelley with a more realistic electron response function such as we have derived in our simulation, then the consistency of this distortion with a two-component neutrino hypothesis (with $m_2 = 17\,\mathrm{keV}$) will disappear.

We are indebted to A. Hime, D. Morrison and M. Bahran for valuable discussions on this topic.

REFERENCES

1. J. J. Simpson, Phys. Rev. Lett. **54**, 1891 (1985).
2. T. Altzitzoglou *et al.*, Phys. Rev. Lett. **55**, 799 (1985).
3. V. M. Datar *et al.*, Nature **318**, 547 (1985).
4. A. Apalikov *et al.*, Pis'ma Zh. Eksp. Teor. Fiz. **42**, 233 (1985) [JETP Lett. **42**, 289 (1985)].
5. T. Ohi *et al.*, Phys. Lett. **B160**, 322 (1985).
6. J. Markey and F. Boehm, Phys. Rev. C **32**, 2215 (1985).
7. D. W. Hetherington *et al.*, Phys. Rev. C **36**, 1504 (1987).
8. H. W. Becker, D. A. Imel, H. Henrikson, V. Novikov and F. Boehm, *Search for Evidence of the 17 keV Neutrino in the β-spectrum of* ^{35}S, (California Institute of Technology preprint, 1991).
9. M. Bahran and G. R. Kalbfleisch, *Limit on Heavy Neutrino in Tritium Beta Decay*, (University of Oklahoma preprint, 1991); also submitted to the Joint International Lepton-Photon Symposium and Europhysics Conference on High Energy Physics, Geneva, Switzerland, July 1991.
10. A. Hime and J. J. Simpson, Phys. Rev. D **39**, 1837 (1989).
11. J. J. Simpson and A. Hime, Phys. Rev. D **39**, 1825 (1989).
12. A. Hime and N. A. Jelley, Phys. Lett. **B257**, 441 (1991).
13. B. Sur *et al.*, Phys. Rev. Lett. **66**, 2444 (1991), and Proceedings of the Conference on Intersections between Nuclear and Particle Physics, Tucson, Arizona, May 1991 (to be published).
14. A. Hime, *Pursuing the 17-keV Neutrino*, to be published in Mod. Phys. Lett. A.
15. A. Hime, D. Phil. thesis, Oxford report OUNP-91-20 (1991).
16. A. Hime and N. A. Jelley, Oxford report OUNP-91-21 (1991).
17. L. Salkind, Senior thesis (unpublished), Princeton University (1981).
18. H. Behrens and J. Janecke, *Numerical Tables for Beta-Decay and Electron Capture*, edited by H. Schopper (Springer-Verlag, Berlin, 1969).
19. A. Sirlin, Phys. Rev. **164**, 1767 (1967).
20. W. R. Nelson, H. Hirayama and D. W. O. Rogers, SLAC Report 265 (1985).
21. R. Talman, Nucl. Instr. and Meth. **159**, 189 (1979).
22. Fluorescent yields and Coster-Kronig transition probabilities from W. Bambynek *et al.*, Rev. Mod. Phys. **44**, 716 (1972), E. McGuire, Phys. Rev. **A3**, 587 (1971), and J. Scofield, At. Data and Nucl. Data Tables **14**, 121 (1974); Auger emission rates from M. H. Chen *et al.*, At. Data and Nucl. Data Tables **24**, 13 (1979).
23. Typical silicon dead layer from G. F. Knoll, *Radiation Detection and Measurement*, (John Wiley and Sons, New York, 1979), p. 487.
24. T. Tabata, R. Ito and S. Okabe, Nucl. Instr. and Meth. **94**, 509 (1971)

PHYSICS OF A 17 KEV NEUTRINO*

Boris Kayser
Division of Physics, National Science Foundation
Washington, DC 20550 USA

Abstract

The possible 17 keV neutrino, if real, cannot be ν_μ but could be essentially ν_τ. Relic 17 keV neutrinos from the big bang must have disappeared, through a non-Standard-Model decay or annihilation process, before the present epoch. If we assume that the 17 keV neutrino is not a Dirac neutrino of the conventional kind, then we are led to picture it as a Dirac neutrino of the unconventional Zeldovich-Konopinski-Mahmoud kind. It is then an amalgam of ν_τ and $\overline{\nu_\mu}$.

* This is a brief summary of the talk presented at the Workshop. A more complete account of the physics discussed there and further references may be found in the article on which the talk was based, being prepared for Reviews of Modern Physics.

If the possible 17 keV neutrino should turn out to be real, how would it fit in with what we know about neutrinos? In discussing this question, I will make extensive use of a nice analysis by Caldwell and Langacker,[1] draw on laboratory and astrophysical data, and inject some personal prejudices. Let us recall that in the standard model (SM), extended to include neutrino masses and mixing, the charged-current weak interaction is given by

$$\mathcal{H}_{cc} = \frac{g}{\sqrt{2}} W^{\alpha} \sum_{\substack{f=e,\mu,\tau \\ m}} i \bar{\ell}_{fL} \gamma_{\alpha} L_{fm} \nu_{mL} + \text{h.c.} \tag{1}$$

Here, ℓ_e is the electron, ℓ_μ is the muon, etc., the ν_m are neutrino mass eigenstates, and L is a nontrivial unitary leptonic mixing matrix. Let us refer to the 17 keV neutrino as ν_3. It is claimed[2] that the mass of this mass eigenstate is

$$M_3 \cong 17\,\text{keV} , \tag{2}$$

and that the mixing matrix element L_{e3} connecting it to the electron obeys

$$|L_{e3}|^2 \cong 0.008 . \tag{3}$$

If we assume that leptonic mixing, like quark mixing, is small, then each neutrino of definite flavor, such as ν_e, consists of some dominant mass eigenstate component plus small amounts of the other mass eigenstates. Given the small value of $|L_{e3}|^2$, ν_3 is clearly not the dominant mass eigenstate component of ν_e. A comparison of $|L_{e3}|^2$ to the upper bound on $\nu_e \to \nu_\mu$ oscillation[3] shows that ν_3 cannot be the dominant mass eigenstate in ν_μ either. The ν_3 could be the dominant mass eigenstate in ν_τ, so we shall at times think of it as essentially ν_τ.

If ν_3 is stable, then ν_3 particles created in the hot big bang contribute about 200 times more to the present mass density of the universe than is observationally allowed.[4] Thus, the ν_3 particles from the big bang must have disappeared, either through decay or annihilation, before the present epoch. Standard-model processes cannot lead to decay within the required time,[5] so physics beyond the SM must be involved. It can be shown that disappearance via annihilation would have to involve non-SM physics as well.[4,6]

Is ν_3 a Majorana neutrino (i.e., its own antiparticle) or a Dirac one (i.e., not its own antiparticle)? There is an argument based on the observed neutrino burst from supernova SN1987A and on nucleosynthesis in the early universe that suggests that ν_3 is not a Dirac neutrino, or at least not one of the conventional kind.[7,4] Recent work makes this argument inconclusive.[8] However, the most popular explanation of the lightness of neutrinos is the "see-saw" mechanism, and this mechanism predicts that neutrinos are Majorana particles. Thus, it is still attractive to explore the possibility that ν_3 is a Majorana neutrino.

If ν_3 is a Majorana neutrino, then a nucleus (A,Z) containing Z protons can decay to the nucleus (A,Z+2) by emitting a pair of virtual W^- bosons, which then exchange ν_3 to produce an outgoing pair of electrons. This decay, $(A,Z) \rightarrow (A,Z+2) + 2e^-$, is referred to as neutrinoless double beta decay, or $\beta\beta_{0\nu}$. The amplitude for this process is proportional to an effective neutrino mass, M_{eff}. Including the contributions of all the Majorana neutrino mass eigenstates ν_m which can be exchanged by the virtual W bosons to yield the outgoing electrons, and neglecting CP violation,[9] M_{eff} is given by[10]

$$M_{eff} = \sum_{\substack{\text{All Majorana} \\ \text{mass eigenstates} \\ \nu_m}} CP(m)\, |L_{em}|^2\, M_m \ . \tag{4}$$

Here, M_m is the mass of ν_m, CP(m) is its intrinsic CP parity (which can be either odd or even), and L_{em} is defined by Eq. (1). So far, $\beta\beta_{0\nu}$ has not been seen, and the experimental upper limit on its amplitude,[11] combined with calculations of the nuclear matrix element for the reaction,[12] yields the bound

$$|M_{eff}| \lesssim 2 \text{ eV}. \tag{5}$$

Now, if ν_3 is a Majorana particle, then, all by itself, it makes a contribution to M_{eff} of magnitude (cf. Eqs. (2) and (3))

$$|L_{e3}|^2 M_3 \cong (0.008) \times (17 \text{ keV}) \cong 140 \text{ eV} \ . \tag{6}$$

If the constraint (5) is not to be violated, this contribution must be cancelled, to within ~ 1%, by that of some other neutrino or neutrinos!

It may appear that such a cancellation must involve a highly unnatural conspiracy among neutrinos. However, it can actually occur quite naturally as a result of a symmetry or conservation law. The required cancellation of the ν_3 contribution cannot be provided by ν_1, the dominant mass eigenstate in ν_e, because $M_1 < 9$ eV,[13] and $|L_{e1}|^2 \leq 1$. It can be provided by ν_2, the dominant mass eigenstate in ν_μ, if

$$CP(2) = -CP(3) \tag{7}$$

and

$$|L_{e2}|^2 M_2 = |L_{e3}|^2 M_3 \ . \tag{8}$$

These constraints could be satisfied by a ν_2 heavier (not lighter) than ν_3.[1] However, a perhaps more attractive possibility is that they are satisfied as the result of a symmetry which has the effect that $M_2 = M_3$, $|L_{e2}|^2 = |L_{e3}|^2$, and $CP(2) = -CP(3)$.

A Majorana neutrino, being its own antiparticle, consists of just two states, one left handed and one right handed. (If a familiar neutrino such as ν_e is a Majorana particle, then its right-handed state is the one with antifermion-like interactions that we normally refer to as the "antineutrino.") By contrast, a Dirac neutrino, being distinct from its antiparticle, consists (together with that antiparticle) of four states with a common mass: the left-handed and right-handed neutrino, plus the left-handed and right-handed antineutrino. Now, if the constraints (7) and (8) are satisfied by a pair of Majorana neutrinos ν_2 and ν_3 with identical masses, then the 17 keV neutrino, ν, is actually a *Dirac* neutrino. It has four states, coming from the states of the degenerate ν_2 and ν_3. The left-handed state of ν is just $\nu_{\tau L} + \varepsilon\nu_{eL}$, where $|\varepsilon|^2 = |L_{e3}|^2 = 0.008$. This makes ν the dominant mass eigenstate in ν_τ, as allowed by neutrino oscillation data, and allows it to couple to an electron, so that its claimed emission in nuclear β decay is possible. Since at the time of nucleosynthesis the number of neutrino species was strictly less than four,[14] the right-handed state of ν cannot be an *extra*, exotic state, but must be the antiparticle of one of the three known left-handed neutrinos: ν_{eL}, $\nu_{\mu L}$, or $\nu_{\tau L}$. This right-handed state cannot be $(\overline{\nu_e})_R$, the antiparticle of ν_{eL}, because the dominant mass eigenstate in ν_e, ν_1, weighs no more than 9 eV, and ν weighs 17 keV. Nor can it be $(\overline{\nu_\tau})_R$, because then it would

be essentially the antiparticle of the left-handed state, so that we would have a Majorana neutrino, rather than a Dirac one. Thus, it must be $(\overline{\nu_\mu})_R$. Hence the 17 keV neutrino has the structure[15,1]

$$\nu_L = \nu_{\tau L} + \varepsilon \nu_{eL} \tag{9.1}$$

and

$$\nu_R = (\overline{\nu_\mu})_R \ , \tag{9.2}$$

and is not a conventional Dirac neutrino, but one that is referred to as a Zeldovich-Konopinski-Mahmoud (ZKM) neutrino.[15,1]

It is easily shown that in the models where the 17 keV neutrino has the structure (9), (electron number) – (muon number) + (tau number) is conserved.[16] Since neutrinoless double beta decay, $(A,Z) \rightarrow (A,Z+2) + 2e^-$, obviously does not conserve this quantum number, this decay is completely forbidden. Thus, the cancellation among the different contributions to M_{eff} must be complete because of the conservation law.[17]

An elegant possible explanation of the observed deficit of solar neutrinos[18] is provided by the Mikheyev-Smirnov-Wolfenstein (MSW) effect.[19] This effect converts a ν_e, which is produced by nuclear reactions in the solar core and which would be detectable in terrestrial solar neutrino detectors, into an undetectable, or largely undetectable, neutrino ν_x. This ν_x may be the ν_μ, the ν_τ, or some sterile (i.e., non-interacting) neutrino. If the conversion is to occur for neutrinos throughout the solar neutrino energy range of interest, then the masses of ν_e and ν_x must be very close:[20,19]

$$M_{\nu_x}^2 - M_{\nu_e}^2 \lesssim 10^{-6}\,\mathrm{eV}^2 \ . \tag{10}$$

Now, since $\nu_e \sim \nu_1$, $M_{\nu_e} < 9$ eV. Thus, if the 17 keV ν_3 is essentially ν_τ, then clearly ν_x cannot be ν_τ. Furthermore, whether the $\beta\beta_{0\nu}$ constraint (8) is satisfied by a $\nu_2 \sim \nu_\mu$ heavier than ν_3, or by a ν_μ degenerate with ν_τ and forming the ZKM neutrino (9) with it, clearly ν_x cannot be ν_μ either. Hence, to explain the solar neutrino deficit by the MSW effect, one must invent an *extra* neutrino ν_x which is light so that $M_{\nu_x}^2 - M_{\nu_e}^2 \lesssim 10^{-6}\,\mathrm{eV}^2$ can be satisfied, and which is sterile so that it would not have been observed

experimentally, just so that solar electron neutrinos can convert into it.[1,21] We leave the attractiveness of this invention to the reader's judgment.

Interestingly, our exploration of the possibility that the 17 keV neutrino is a Majorana particle led us to a picture in which it is a *Dirac* particle, albeit one of the unconventional ZKM type. Part of our motivation for exploring the Majorana possibility was the observation that the see-saw explanation of the lightness of neutrinos predicts that they are Majorana particles. If the 17 keV neutrino is actually a ZKM Dirac neutrino, can one still explain its lightness, relative to, say, the τ, by a see-saw mechanism? It has been shown that indeed one can,[22] in the sense that one can construct a model of this neutrino in which its mass is inversely proportional to a large but physically reasonable mass, just as in the normal see-saw mechanism. To be sure, the physics of this model is quite different from that of the normal see-saw, but the model involves the same key algebraic structure as the latter.

In summary, the possible 17 keV neutrino does not violate anything that is known about neutrinos. However, if it is not a conventional Dirac neutrino, then it is not easily accommodated within the existing constraints. It will be interesting, to say the least, to see whether this neutrino turns out to be real.

References

1. D. Caldwell and P. Langacker, *Phys. Rev. D* **44** (1991) 823.
2. Experiments reporting evidence for a 17 keV neutrino include those discussed in J. Simpson, *Phys. Rev. Lett.* **54** (1985) 1891; B. Sur *et al.*, *Phys. Rev. Lett.* **66** (1991) 2444; A. Hime and N. Jelley,*Phys. Lett.* **B257** (1991) 441 and Oxford Univ.preprint OUNP-91-21; and I. Zlimen *et al.*, *Phys. Rev. Lett.* **67** (1991) 560. Experiments which do not see such evidence include those discussed in D. Hetherington *et al., Phys. Rev. C* **36** (1987) 1504; H. Becker *et al.*, Caltech preprint CALT-63-605; and M. Bahran and G. Kalbfleisch, Univ. of Oklahoma preprint OKHEP-91-005.
3. B. Blumenfeld *et al.*, Johns Hopkins preprint JHU HEP 1289-1. For a review of neutrino oscillation results, see F. Pierre, Saclay preprint DPhPE 90-19.

4. E. Kolb and M. Turner, *Phys. Rev. Lett.* **67** (1991) 5.

5. A. De Rujula and S. Glashow, *Phys. Rev. Lett.* **45** (1980) 942.

6. E. Carlson, D. Land, and L. Randall, *Phys. Lett.* **B264** (1991) 132.

7. K. Gaemers, R. Gandhi, and J. Lattimer, *Phys. Rev. D* **40** (1989) 309; R. Gandhi and A. Burrows, *Phys. Lett.* **B246** (1990) 149; M. Turner, *Phys. Rev. D* **45** (1992) 1066; G. Raffelt and D. Seckel, *Phys. Rev. Lett.* **67** (1991) 2605; J. Cline and T. Walker, *Phys. Rev. Lett.* **68** (1992) 270. I thank M. Turner, G. Raffelt, and S. Kahana for helpful discussions of these analyses.

8. A. Burrows, R. Gandhi, and M. Turner, Fermilab preprint 92/29-A.

9. The effect of CP violation on M_{eff} is discussed in M. Doi *et al., Phys. Lett.* **102B** (1981) 323.

10. L. Wolfenstein, *Phys. Lett.* **107B** (1981) 77; B. Kayser and A. Goldhaber, *Phys. Rev. D* **28** (1983) 2341; S. Bilenky, N. Nedelcheva, and S. Petcov, *Nucl. Phys.* **B247** (1984) 61.

11. D. Caldwell *et al., Nucl. Phys. B (Proc. Suppl.)* **13** (1990) 547; A. Piepke, talk presented at the XII Moriond Workshop on Massive Neutrinos and Tests of Fundamental Symmetries, Les Arcs, France, 1992; J.-C. Vuilleumier, talk presented at the XII Moriond Workshop on Massive Neutrinos and Tests of Fundamental Symmetries, Les Arcs, France, 1992.

12. W. Haxton and G. Stephenson, Jr., *Prog. Part. Nuc. Phys.* **12** (1984) 409; M. Doi, T. Kotani, and E. Takasugi, *Prog. Theor. Phys. Suppl.* **83** (1985) 1; K. Grotz and H. Klapdor, *Phys. Lett.* **B153** (1985) 1; T. Tomoda *et al., Nucl. Phys.* **A452** (1986) 591; T. Tomoda and A. Faessler, *Phys. Lett.* **B199** (1987) 475; J. Engel, P. Vogel, and M. Zirnbauer, *Phys. Rev. C* **37** (1988) 731.

13. W. Stoeffl, talk presented at the Spring Meeting of the American Physical Society, Washington 1991; H. Robertson *et al., Phys. Rev. Lett.* **67** (1991) 957; E. Holzschuh, talk presented at the XII Moriond Workshop on Massive Neutrinos and Tests of Fundamental Symmetries, Les Arcs, France, 1992.

14. K. Olive *et al., Phys. Lett.* **B236** (1990) 454.

15. Y. Zeldovich, *DAN SSSR* **86** (1952) 505; E. Konopinski and H. Mahmoud, *Phys. Rev.* **92** (1953) 1045; L. Wolfenstein, *Nucl. Phys.* **B186** (1981) 147; S. Petcov, *Phys. Lett.* **B110** (1982) 245; R. Barbieri and L. Hall, *Nucl. Phys.* **B364** (1991) 27; K. Babu, R. Mohapatra, and I. Rothstein, Univ. of Maryland preprint UMDHEP 92-061; A. Smirnov and J. Valle, Valencia preprint FTUV/91-38.

16. See L. Wolfenstein and S. Petcov, Ref. 15.

17. Modification of these models may remove this conservation law but leave in place a symmetry which demotes $\beta\beta_{0\nu}$ from $\mathcal{O}$ [neutrino mass] to $\mathcal{O}$ [(neutrino mass)^3]. The theory of $\beta\beta_{0\nu}$ at this more-suppressed level is studied in W. Haxton, *Phys. Rev. Lett.* **67** (1991) 2431.

18. K. Lande, talk presented at Neutrino 90, Geneva, 1990; Y. Totsuka, in *Neutrino 90*, eds. J. Panman and K. Winter (North-Holland, 1991), p. 69; A. Abazov *et al.*, *Phys. Rev. Lett.* **67** (1991) 3332.

19. S. Mikheyev and A. Smirnov, *Nuovo Cimento C* **9** (1986) 17. See also L. Wolfenstein, *Phys. Rev. D* **17** (1978) 2369.

20. P. Rosen and J. Gelb, *Phys. Rev. D* **34** (1986) 969.

21. A. Smirnov, Valencia preprint FTUV/91-44; J. Valle, Valencia preprint FTUV/91-45.

22. K. Babu and R. Mohapatra, University of Maryland preprint UMDHEP 92-150.

ASTROPHYSICAL (SN1987A) CONSTRAINTS ON THE DIRAC NEUTRINO MASS

J.A. Grifols

Grup de Física Teòrica,
Universitat Autònoma de Barcelona,
08193 Bellaterra (Barcelona), Spain

ABSTRACT

Dirac masses of any neutrino species can be bounded by Supernova cooling arguments. The so obtained upper bound lies in the tens of keV. This is obviously relevant for the case of the 17 keV neutrino. We conclude that general luminosity arguments cannot strictly reject the 17 keV neutrino, but detailed analyses using star evolution codes that include the backreaction of neutrino cooling probably do.

The energy drain in Supernova collapse proceeds mainly through thermal neutrino emission. This is a surface effect. The neutrinos diffuse through the very dense material (neutrino trapping occurs for core densities in excess of $2 \times 10^{11} g/cm^3$) experiencing very many scatterings until they reach the neutrino-sphere where they are subsequently blackbody radiated. The gravitational binding energy released is about 10^{53} erg. Indeed, at most 4×10^{53} erg are emitted and from observational data[1] (IMB, Kamiokande) on SN1987A we know[2] that the energy carried away by neutrinos was larger than 2×10^{53} erg. This energy was radiated in a diffusion time on the order of 1-10 seconds. As a consequence any additional energy drain (besides the one associated to standard left-handed (LH) neutrino emission) should have a luminosity below 2×10^{53} erg/s.

Should neutrinos be massive Dirac particles, then right-handed (RH) degrees of freedom do exist and they should provide for an extra source of SN cooling[3–6]. In fact, RH-neutrinos should be produced in the hot SN core via helicity flip mechanisms. For small enough neutrino mass, RH-neutrinos stream freely out of the star because RH-neutrinos are (almost) sterile. This is a volume effect. The size of the effect is proportional to the neutrino mass squared and the relevant spin-flip processes are: neutrino nucleon elastic scattering, nucleon nucleon neutrino bremsstrahlung and e^+e^- annihilation. The RH-neutrino luminosity associated to the processes that take place in bulk matter of the SN core ($T \sim 60 MeV$, $R \sim 10 Km$) can be easily calculated.

One reaches the conclusion that[5]

$$m_\nu \lesssim 14 \; keV$$

otherwise, the emission power would exceed the limit 2×10^{53} erg/s.

We can check "a posteriori" that neutrinos with $m_\nu < 14 \; keV$ do indeed stream out freely. The mean-free-path in the core is calculated to be

$$8 \times 10^2 \; km \left(\frac{50 \, keV}{m_\nu} \right)^2$$

when the main opacity source $\nu_R N \rightarrow \nu_L N$ is taken into account. Now, it is obvious from this that by increasing m_ν sufficiently we shall reach a diffusion regime and, if m_ν is large enough, the corresponding difussion time will be long enough such that

the luminosity is again below 2×10^{53} erg/s. Hence a *lower* neutrino mass bound (as oposed to the previous *upper* limit) follows. It can be obtained in the following way. First, the ν_R-sphere R is calculated by imposing that the optical depth

$$\tau(r) = \int_r^\infty \frac{dr}{\lambda} \qquad (\lambda = \text{mean} - \text{free} - \text{path})$$

be 2/3 at $r = R$. Next, use Stefan-Boltzmann's law to compute the luminosity. One needs to know for this purpose the temperature at the ν_R-sphere (and we use the temperature profile of ref. 7). The luminosity depends inversely on m_ν^2 and the constraint $L < 2 \times 10^{53} erg/s$ implies then

$$m_\nu \gtrsim 34\ MeV$$

Hence, there is a forbidden mass window between about 14 keV and 34 MeV for the Dirac neutrino mass of any species. Only the lower end of this range is of phenomenological interest. Specially in the light of the 17 keV Simpson neutrino[8]. In this respect, the question arises as to the reliability of the bound. Of course, a (most) important source of uncertainty is our ignorance of the equation of state at supernuclear densities. Allowing for the density to change by a factor of two implies that the core temperature spans the range 30 to 100 MeV (i.e. the typical temperature interval entertained in models of stellar collapse) and leads to a factor ~ 3 uncertainty in the bound.

This bound - based on the luminosity argument - could be somewhat improved but not much. We expect that by including previously neglected effects (like charged current processes in nucleon-nucleon bremsstrahlung) one may reach the limit

$$m_\nu \lesssim O(10\ keV) \times (3^{\pm 1})$$

On the other hand, similar conclusions can be obtained using other arguments. Burrows and Gandhi[6] obtain $m_\nu \lesssim 28\ keV$ based on detailed numerical analysis using proto-neutron star evolution codes which include the backreaction of cooling via massive neutrinos. In particular, they study the effect of cooling by emission of wrong helicity neutrinos on the duration of the detected neutrino bursts. Unless $m_\nu \lesssim 28\ keV$, this duration is intolerably short.

Recently, Turner[9] has reexamined the issue and claims that the absolute limit should be well below 10 keV.

Our own conclusion is that the sole luminosity argument is unable to exclude the 17 keV neutrino and that one needs more elaborate considerations as well as detailed analysis (such as the ones used in refs. 6 and 9) to exclude the 17 *keV* neutrino.

REFERENCES

1. K. Hirata et al., *Phys. Rev. Lett* (1987) 1490;
   R. Bionta et al., *Phys. Rev. Lett* (1987) 1494.
2. D.N. Schramm, in: *1987 Int. Symp. on Lepton and Photon interactions at high energies*, eds. W. Bartel and R. Rückl (North-Holland, Amsterdam, 1988) p. 471.
3. G. Raffelt and D. Seckel, *Phys. Rev. Lett.* **60** (1988) 1793.
4. K.J.F. Gaemers, R. Gandhi and J.M. Lattimer, *Phys. Rev.* **D40** (1989) 309.
5. J.A. Grifols and E. Massó, *Phys. Lett.* **B242** (1990) 77.
6. R. Ghandi and A. Burrows, *Phys. Lett.* **B246** (1990) 149.
7. M.S. Turner, *Phys. Rev. Lett.* **60** (1988) 1797.
8. J.J. Simpson, *Phys. Rev. Lett.* **54** (1985) 1891.
9. M.S. Turner, FERMILAB-Pub-91/136-A, May 1991.

Physics with Neutrons and Atoms

NEUTRON INTERFEROMETRY: A UNIQUE TOOL FOR QUANTUM MEASUREMENTS

Helmut Kaiser and Samuel A. Werner
University of Missouri Research Reactor and Physics Department
Columbia, Missouri 65211 USA

Abstract

Since its successful realization in 1974, neutron interferometry has become an important and unique technique for probing and elucidating fundamental quantum mechanical principles on a macroscopic scale. To date, all neutron interferometry results have been in agreement with the predictions of quantum mechanics, yet they serve to enhance our understanding about the interpretation of this theory. In this paper we present an introduction and overview of this field and describe in some detail the following three experiments carried out in Missouri: (a) Gravitationally-induced quantum interference; (b) Topological effects in quantum mechanics; (c) Coherence effects and interference.

1. Introduction and Overview

Neutron interferometry in the Ångstrom wavelength range generally uses Bragg reflection from crystal planes to split and recombine the beam. Historically, this idea of using a monolithic perfect silicon crystal was first developed for x-ray interferometry by Bonse and Hart in 1964.[1)] Neutron interferometry was first demonstrated by Rauch, Treimer and Bonse[2)] at the TRIGA reactor at the Atominstitut in Vienna in 1974. A year later, the first successful neutron interferometry experiment in the U.S. was performed: the gravitationally-induced quantum interference experiment by Colella, Overhauser and Werner.[3)] Since that time many significant experiments have been carried out, and this single-crystal device has proven to be a marvelous laboratory for probing and elucidating the fundamental quantum mechanical principles of nature.

A list of general references is given in Table 1. In Table 2, we give a compilation of important neutron interferometry experiments to date. The label MURR means University of Missouri, ILL stands for Institut Laue Langevin and MIT for Massachusetts Institute of Technology. A new and more sophisticated neutron interferometer spectrometer is under construction at NIST (National Institute of Standards and Technology) in Washington and will presumably be operational within the next year.

We would like to mention a series of other efforts in de Broglie matter waves interferometry: the first realization of a diffraction grating interferometer for thermal neutrons in 1985[4)] and, more recently, for very cold neutrons.[5,6)]

Table 1. Neutron Interferometry

General References:

1. Neutron Interferometry, Proceedings of an Internat. workshop held June 5-7, 1978 in Grenoble, France at the ILL, ed. U. Bonse and H. Rauch (Oxford University Press, 1979).
2. "Neutron Interferometry," S. A. Werner, Physics Today, December 1980.
3. "Neutron Optics," A. G. Klein and S. A. Werner, Rep. Prog. Phys. 46, pp. 259-335, 1983.
4. "Neutron Optics," S. A. Werner and A. G. Klein, Chap. 4 of Neutron Scattering, Methods of Experimental Physics, Vol. 23A, ed. K. Sköld and D. L. Price (Academic Press, 1986).
5. New Techniques and Ideas in Quantum Measurement Theory, Conf. in honor of Eugene Wigner in NYC, Jan. 21-24, 1986, ed. D. M. Greenberger, Annals of N. Y. Acad. of Sciences, Vol. 480, December 1986.
6. Matter Wave Interferometry, Proceedings of an Internat. Workshop in Vienna, Austria, Sept. 14-16, 1987, ed. G. Badurek, H. Rauch, and A. Zeilinger (North-Holland, 1989).
7. Neutron Optics, Varley F. Sears (Oxford University Press, 1989).

Table 2. Neutron Interferometry (1974-1992)

1. First Test of Si-crystal Interferometer. Vienna (1974)
2. Sign Change of Fermion Ψ During 2π Precession. ILL, MURR (1975, 1976)
3. Gravitationally-Induced Quantum Interference. Ann Arbor, MURR (1975, 1980, 1985, 1988)
4. Neutron Sagnac Effect.
   - Earth's rotation. MURR (1979)
   - Turntable. MIT (1984)
5. Neutron Fizeau Effect.
   - Moving boundaries, ILL (1981, 1985)
   - Stationary boundaries, MURR (1985, 1988)
6. Search for Non-linear Terms in the Schrödinger Equation.
   - LL- 2-crystal interferometer. MIT (1981)
   - Long-λ Fresnel diffraction, ILL (1981)
7. Search for an Aharonov-Bohm Effect for Neutrons. MIT (1981)
8. Measurement of Longitudinal Coherence Length of a Neutron Beam. MURR (1983)
9. Coherent Superposition of Spin States ("Wigner Phenomena"). ILL (1983)
10. Search for Quaternions in Quantum Mechanics. MURR (1984)
11. Quantum Interference in Accelerated Frames. ILL (1983)
12. Search for New Guage Fields--Rotating U Rod. MIT (1983)
13. Precision Measurement of Scattering Lengths ^{149}Sm, ^{235}U, ^{3}He, ^{3}H (Four-Body Nuclear Interaction). ILL, MURR (1975-1985)
14. Observation of the Aharonov-Casher Effect. MURR (1989)
15. Stochastic vs. Deterministic Attenuation of a Neutron Beam. ILL (1987)
16. Neutron Spin-Pendellösung Resonance. MIT (1988)
17. Coherence and Spectral Filtering. MURR (1991)
18. Observation of the Scalar Aharonov-Bohm Effect. MURR (1992)

There are also some very intensive activities in atom interferometry; for a discussion of its current status see a Physics Today 1991 article.[7] However, in all of these cases the beam separation is still rather small (up to a couple millimeters).

2. Perfect crystal interferometers and quantum phase shift

A neutron interferometer for thermal neutrons is usually machined from a monolithic perfect silicon crystal, which produces two widely separated, coherent beams. A schematic diagram of an LLL(Laue-Laue-Laue)-type device is shown in Fig. 1. Two of the interferometer crystals which are presently in use at the University of Missouri Research Reactor are shown in Fig. 2. The three perfect crystal slabs (four in the case of the skew-symmetric interferometer) are cut perpendicular to a set of strongly reflecting lattice planes, typically the (220) planes. The common base ensures the perfect alignment of the crystal planes from slab to slab.

A nominally collimated, monochromatic beam, typically $\Delta\lambda/\lambda \approx 0.01$, is directed from the source to the first slab (point A in Fig. 1), where it is split coherently by Laue diffraction. The two resulting beams are split by the second

silicon slab in the regions near points B and C. The central two of these four beams overlap in the region of point D on the third silicon slab. The two beams leaving the third slab are superpositions of the wavefunction on beam path I and beam path II:

$$I \propto \left| \Psi_I + \Psi_{II} \right|^2 \tag{1}$$

For an exact treatment, the reflection process within each crystal slab has to be calculated by using the dynamical theory of diffraction, but it is not essential for understanding the conceptual operational features. This device is topologically identical to the Mach-Zehnder interferometer of classical optics.

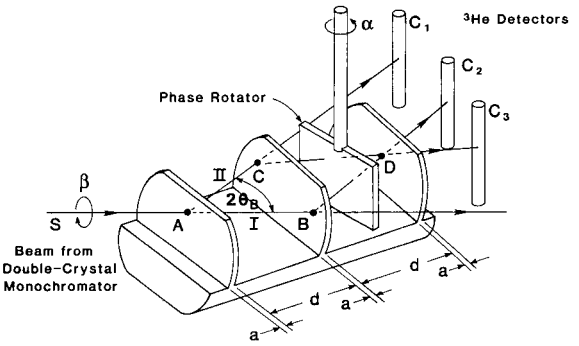


Fig. 1 Schematic diagram of our LLL Bonse-Hart Si-crystal interferometer.

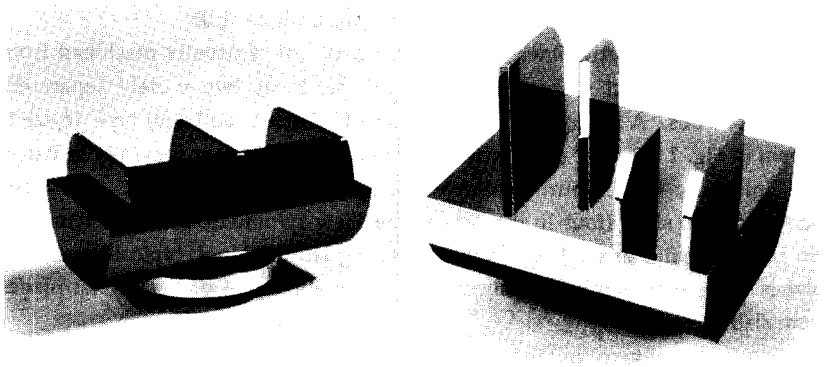


Fig. 2 Photograph of our (a) LLL interferometer and (b) skew symmetric interferometer

If we place a plate of a certain material with thickness D into one of the beams (e.g. beam path I), the phase of the wavefunction for the empty interferometer $\Psi_{0,I}$ will be shifted by an amount $\Delta\phi$

$$\Psi_I = \left| \Psi_{0,I} \right| e^{-i\Delta\phi} \tag{2}$$

This phase shift can be described--analog to light optics--with the index of refraction n for neutrons

$$\Delta\phi = (1 - n)kD = N b_c \lambda D \tag{3}$$

with

$$n = 1 - \lambda^2 \frac{N b_c}{2\pi} \tag{4}$$

where λ is the wavelength of the incident neutron, b_c the coherent nuclear scattering length per atom and N the atom density of the material. Thus with neutron interferometry, the phase of the neutron's wavefunction, Ψ, becomes directly accessible to measurement.

Now, by placing a plate into both beams (as shown in Fig. 1) and changing the optical path length by rotating it in small increments α, one obtains--after recombination of Ψ_I and Ψ_{II}--constructive and destructive interference oscillations, which can be observed in detectors C_2 and C_3. These oscillations, as a function of the overall phase shift $\Delta\phi(\alpha)$, are of the form

$$I_2 = a_2 - b_2 \cos(\Delta\phi(\alpha)) \tag{5}$$

and

$$I_3 = a_3 + b_3 \cos(\Delta\phi(\alpha)) \tag{6}$$

where the constants a_2, a_3, b_2 and b_3 are characteristic to the actual interferometer setup and $b_2 = b_3$ and $a_2/a_3 \approx 2.7$. Thus, the neutron current is swapped back and forth between detectors C_2 and C_3 as $\Delta\phi(\alpha)$ is varied. Because of particle conservation, the sum $I_2 + I_3$ has to be a constant.

The theory of quantum mechanics is based upon potentials, canonical momenta and phase shifts, whereas classical mechanics is founded upon fields, momenta and forces. A quantum phase shift can be caused by any potential, e.g. nuclear, gravitational, magnetic, electric, etc. By using the appropriate Hamiltonian for a particular problem, the canonical momentum can be calculated using Hamilton's equations. In general, then, the phase shift in an interferometer experiment can be calculated by evaluating the line integral of the canonical momentum around the interferometer loop, namely

$$\Delta\phi = \oint \vec{k} \cdot d\vec{s} = \frac{1}{\hbar} \oint \vec{p} \cdot d\vec{s} \tag{7}$$

where $\vec{k}$ is the wave vector, $\vec{p}$ is the canonical momentum and $d\vec{s}$ the vector element of length along paths I and II. We will now apply this fundamental concept in the following three experiments.

3. Gravitationally Induced Quantum Interference

In most phenomena of interest in physics, gravity and quantum mechanics do not simultaneously play an active role. However, the neutron interferometer is sufficiently sensitive to detect the small changes in the neutron wave function induced by the Earth's gravity, and also the phase shift due to the inertial consequences of its rotation (Neutron Sagnac Effect). The Hamiltonian describing the neutron's motion in our laboratory's coordinate frame may be written as

$$\mathcal{H} = \frac{\vec{p}^2}{2m_i} + m_g\vec{g}\cdot\vec{r} - \vec{\omega}\cdot\vec{L} \tag{8}$$

where m_i and m_g are the neutron's inertial and gravitational mass, $\vec{p}$ is the neutron's canonical momentum, $\vec{r}$ is the position vector from the center of the Earth, $\vec{g}$ is the acceleration due to gravity, $\vec{\omega}$ the angular rotation velocity of the Earth, and $\vec{L} = \vec{r}\times\vec{p}$ is the angular momentum of the neutron's motion about the center of the Earth. Using Hamilton's equations with the above Hamiltonian and Eq. (7), one obtains the following phase shifts: (a) when the direction of the incident neutron beam is along the local (horizontal) north-south axis of the Earth and it is also the axis of rotation, the resulting phase shift is

$$\Delta\phi = \Delta\phi_{grav} + \Delta\phi_{Sagnac} = -\frac{m_i m_g g\lambda A \sin\beta}{2\pi\hbar^2} + \frac{2m_i\omega A \sin\Theta_L \cos\beta}{\hbar^2} \tag{9}$$

and (b) when the incident beam and the axis of rotation is vertical, the resulting phase shift is due to Earth's rotation only,

$$\Delta\phi = \Delta\phi_{Sagnac} = \frac{2m_i\omega A \sin\Theta_L \sin\beta'}{\hbar^2} \tag{10}$$

where A is the area enclosed by the beam trajectories in the interferometer, β is the angle between this area and a horizontal plane, λ is the neutron wavelength, Θ_L is the colatitude angle at the experimental site, and β' is the interferometer orientation angle about the vertical axis.

The influence of gravity only (first term in Eq.(9)) was first observed in 1975[3] by turning the interferometer around a horizontal axis as indicated in Fig. 1, which caused one of the beams to travel on a higher gravitational potential than the other. This difference in gravitational potential caused the above

described phase shift $\Delta\phi_{grav}$. The magnitude of the second term in Eq. (9), $\Delta\phi_{Sagnac}$, is only about 2% of $\Delta\phi_{grav}$ for this particular geometry and was just a small correction. Subsequently, a series of increasingly precise experiments have been carried out by our group.[8,9] At present, the observed frequency of oscillation of the interferogram due to gravity is 0.8% lower than theory predicts.[9] It was suggested[10] that this slight discrepancy is due to the different neutron and x-ray beam sizes we used. We intend to investigate this suggestion experimentally this coming summer.

The phase shift shown in Eq.(10), which gives rise to the Coriolis force, was measured in 1980[8] by turning the interferometer around a vertical axis. From symmetry, it is clear that for a beam, which is precisely vertical, there is no phase shift due to gravity. In this case, the Sagnac phase changed according to the orientation of the interferometer with respect to the cardinal directions, as is evident from Eq.(10). There is good agreement between observed and predicted values for the phase shift (see Fig. 3).

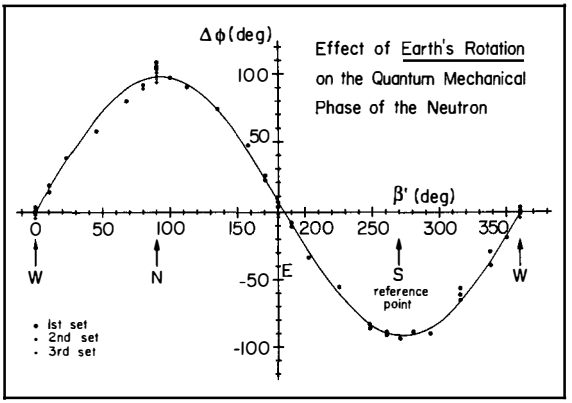


Fig 3. Phase shift due to Earth's rotation. The angle β' specifies the orientation of the interferometer normal area vector with respect to the local N-S axes of the Earth.

The combination of these two experiments represent a quantum mechanical interference measurement of the inertial and gravitational masses.

4. Topological Effects in Quantum Mechanics

In 1959, Aharonov and Bohm (AB)[11] wrote a renowned paper in which they predicted that electrons will suffer phase shifts in passing through regions of space with zero fields but nonzero potentials. They point out the now well-known magnetic (or vector) AB-effect, which concerns the phase shift for

electrons diffracted around a tube of magnetic flux, and the less-known electric (or scalar) AB-effect, where a phase shift for electrons is predicted when they pass through a region of electrostatic potential (such as a scalar potential V=-eU). These types of effects cannot be explained classically and are therefore regarded as prototype experiments in quantum mechanics measurement, since there is no Lorentz force (i.e., no change in velocity) acting on the particles. With the advent of neutron interferometry, the question immediately arose, as to whether an analog effect for neutrons exists.

In 1984, Aharonov and Casher (AC)[12] suggested that electrically neutral particles with a magnetic dipole moment, such as neutrons, should experience a phase shift when they traverse a region of space containing an electric field, e.g. created by a line of electric charge. The AC-effect is regarded as an electrodynamic and quantum mechanical dual of the vector AB-effect, as shown in Fig. 4.

In both the AB and AC cases, the effect is topological in that (a) the actual shape of the paths does not matter, as long as the line of charge (or tube of magnetic flux) is long and fully enclosed by the trajectories (For the AC-effect, however, this is only valid if the axis of quantization, $\vec{\mu}$, is parallel to the axis of the cylinder.); (b) the cross-sectional shape of the line charge (or tube of magnetic flux) is irrelevant; (c) it is independent of the speed of the particle. This

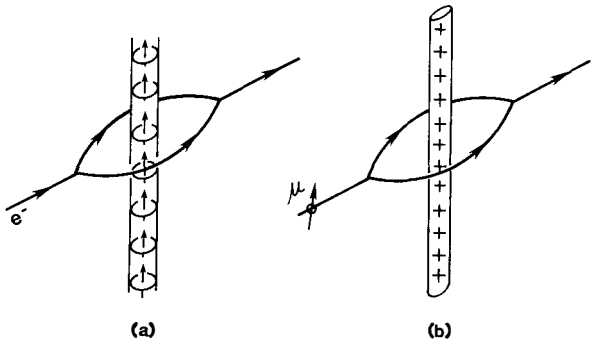


Fig. 4. (a) AB-effect: charged particles diffract around a line of magnetic dipoles (conceptually replacing the tube of magnetic flux). (b) AC-effect as an analog of the above: neutral particles with a magnetic dipole moment (e.g. neutrons) diffract around a line of charge.

topological aspect is important in order to replace a single line of electric charge with an extended, charged electrode system to carry out an experimental test of the AC-effect.

For the arrangement shown in Fig. 4b, the canonical momentum for a neutron of mass m and magnetic moment $\vec{\mu}$ moving with velocity $\vec{v}$ in a region of space containing an electric field $\vec{E}$ is

$$\vec{p} = m\vec{v} + \frac{1}{c}\vec{\mu} \times \vec{E} \tag{11}$$

For a line charge (or an extended electrode) of lineal charge density Λ, the phase shift experienced by a neutron traversing one path relative to the other around the line charge is

$$\Delta\phi_{AC} = \frac{1}{\hbar}\oint \vec{p} \cdot d\vec{s} = \sigma \frac{4\pi\mu\Lambda}{\hbar c} \tag{12}$$

where $\sigma = \pm 1$ for spin up or spin down neutrons.

We carried out such an experiment using our LLL-neutron interferometer and observed a phase shift after the neutrons had traversed a strong electric field (30kV/mm), which was produced by a special electrode system. A schematic of this set-up is shown in Fig. 5. The measured phase shift for our experimental conditions was about 2 milliradians and the agreement with theory is quite satisfactory. For a more detailed description of the actual measuring procedure and results, see ref.[13,14].

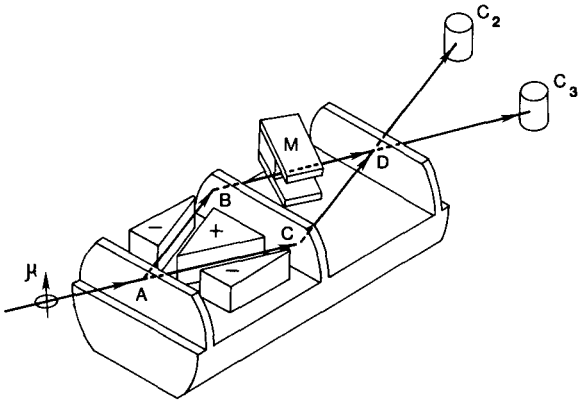


Fig. 5 Schematic for the AC experiment showing the electrode system which applies the electric field in the region traversed by the split beams and in one beam the magnet supplying the vertical magnetic bias field.

Most recently, we carried out a neutron interferometry experiment which must be regarded--as was pointed out by Zeilinger in 1986[15]--as the neutron analog to the scalar AB-effect for electrons. The phase shift in this case is due to the scalar potential $V = -\vec{\mu} \cdot \vec{B}$ and the Hamiltonian is

$$\mathcal{H} = \frac{\vec{p}^2}{2m} - \vec{\mu} \cdot \vec{B} \tag{13}$$

where $\vec{\mu}$ is the magnetic moment, and $\vec{B}$ is the magnetic field. A schematic diagram of the scalar AB experiment for neutrons is shown in Fig. 6. A solenoid is placed in one of the coherent beams and a current pulse $i_2(t)$ is applied while the neutron is in the force-free interior of solenoid 2. The resulting magnetic field $B_2(t)$ gives rise to a phase shift

$$\Delta\phi_{AB} = \frac{\sigma}{\hbar} \int \mu B_2(t)\, dt \tag{14}$$

where $\sigma = \pm 1$ depends on whether the neutron's spin is up or down relative to the magnetic field, the direction of quantization. In the actual experiment, short duration current pulses of 8 μsec width were applied to a suitably designed solenoid placed in beam path II of our skew symmetric interferometer. The coil in beam path I was used to provide a constant bias field for calibration purposes. With this arrangement, and with a time-of-flight measuring technique, we were able to measure phase shifts of $\Delta\phi_{AB} = 0$ up to $3\pi/4$, which was achieved with $B_2(t) = 0$ up to 30 Gauss pulses. As is clear from Eq.(14), this again is a topological effect in the sense mentioned above. For a more detailed write up of the measuring technique and results, see ref.[16].

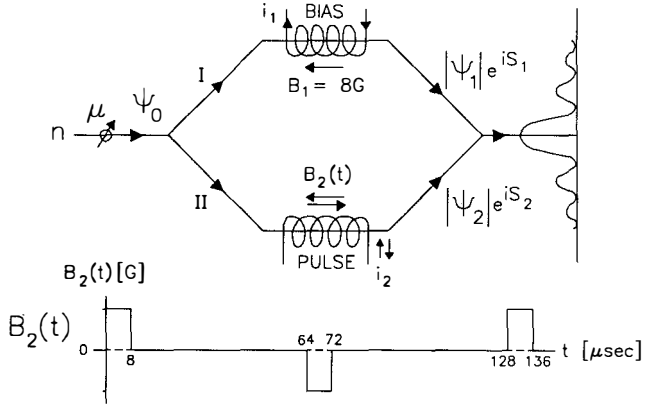


Fig. 6 Schematic diagram of the scalar Aharonov-Bohm experiment for neutrons. The wave forms of the applied pulses are also shown.

We would like to emphasize that the scalar AB-effect for electrons has never been measured due to technical difficulties with existing types of electron interferometers, and that our neutron experiment is the first verification of the scalar AB-effect.

5. Coherence Effects and Interference

In a series of experiments we investigated the longitudinal coherence properties of a thermal neutron beam and the loss and recovery of fringe visibility or contrast.

When a material with a neutron-nuclear optical potential V_{op}

$$V_{op} = \frac{2\pi \hbar^2 N b_c}{m} \quad , \tag{15}$$

(where m is the neutron mass, N is the atom density and b_c the nuclear scattering length of the material) is placed in one beam path of the neutron interferometer, there is a loss of fringe visibility $((I_{max}-I_{min})/(I_{max}+I_{min}))$, the extent of which depends on the longitudinal coherence length. The coherence length of the neutrons depends on the spectral width of their wavelength distribution $g(\bar{k})$ with

$$g(\bar{k}) = \left| a(\bar{k}) \right|^2 \quad , \tag{16}$$

where $a(\bar{k})$ are the amplitudes of the component plane waves. In our experiments, the spectral distribution is not precisely a Gaussian, but can easily be modelled by two closely spaced Gaussians. We have shown that the lost fringe visibility can be restored: (1) by narrowing the spectral distribution; (2) by placing another slab of material (having opposite sign in V_{op}) into the same path of the interferometer.

In the first case, we placed a series of highly polished *Bi* metal slabs in beam path II of our skew symmetric interferometer and placed an analyzer crystal (pressed Si with narrow mosaic width) after the interferometer in the C_3 beam, in order to select out of the spectral distribution $g(\bar{k})$ a window of $w(\bar{k})$ of Fourier components $\bar{k}$. A schematic diagram of the set-up is shown in Fig. 7. Due to the narrower spectral width of the analyzed C_3 beam, the coherence length is increased and some of the fringe visibility is restored. This observed interference phenomenon can be described by the mutual coherence function $\Gamma(D,\alpha)$, which is given by the autocorrelation function of the overlapping wave functions

$$\Gamma(D,\alpha) = \langle \Psi^*(0)\Psi(\Delta\ell) \rangle = (2\pi)^{-1/2} \int g(\bar{k}) \cos\Delta\phi \, d\bar{k} \quad , \tag{17}$$

where $\Delta\phi$ is the overall phase shift (due to the *Bi* sample and the Al phase shifter), D is the sample thickness, α is the phase shifter angle, $\Delta\ell = -1/2(DV_{op}/E)$ is the longitudinal displacement of Ψ_{II} versus Ψ_I and E is the kinetic energy of the neutrons. The relative contrast (contrast for sample-in divided by contrast for sample-out) can be calculated using the magnitude of the

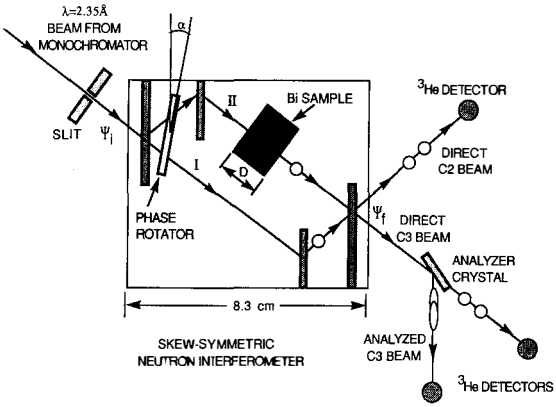


Fig. 7 Schematic diagram of our skew symmetric interferometer. *Bi* sample in beam path II and analyzer crystal (pressed Si crystal) after the interferometer as used in the spectral filtering experiment. The interference contrast is measured by rotating the aluminum phase flag (or phase shifter) in steps α about a vertical axis.

coherence function. The results are summarized in Fig. 8. The contrast falls off with *Bi* thickness D as expected and it clearly shows that the relative contrast for the analyzed beam drops off a lot slower. For D = 8mm *Bi*, there is still 50% contrast, whereas there is none for the direct beam (with no analyzer crystal). We also observed a phase switch of 180° for a certain range of sample thicknesses. This phase switching and the contrast modulations (tail for the direct beam in Fig. 8) can be easily explained by modelling the measured spectral

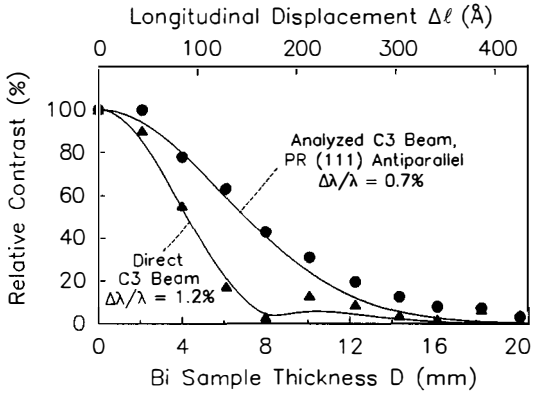


Fig. 8 Plot of the relative contrast in the direct C_3 beam and also in the analyzed C_3 beam as a function of the *Bi* slab thickness.

distribution by two Gaussians. For a more detailed description of the experiment and the results, see ref.[17,18].

In the second case, when we placed different thicknesses of *Bi* or *Ti* samples alone in beam path II, the contrast vanished again with a certain sample thickness, whereas the contrast was recovered, when the same *Bi* ($V_{op}>0$) and *Ti* samples ($V_{op}<0$) were jointly placed in the beam. We showed that there is hardly any loss of contrast up to a combined sample thickness of 45mm. This effect can be interpreted by spatially shifting two wavepackets with respect to each other: the first sample shifts the packet by $\Delta\ell$ and the second sample by $-\Delta\ell$. This quantum mechanical phenomenon is also called "Neutron Phase Echo Effect." See ref.[19] for a more detailed description of the experiment and result.

We also carried out a neutron interferometry experiment in the time domain. A chopper served as a source of neutron intensity pulses, which passed through our skew symmetric interferometer crystal. While travelling the pulses spread, with the faster neutrons tending toward the leading, the slower toward the trailing edge. When placing *Bi* samples of different thicknesses in one beam path of the interferometer, again a loss of fringe visibility could be observed. With a time-of-flight measuring technique, the pulses are divided into time slices, within which the wavelength spectrum is narrower than in the overall pulse. Thus the coherence length in the time slice is increased and some of the fringe visibility can be restored, even though it had disappeared in the overall pulse. We also observed an additional contrast modulation due to the overlap of neighboring pulses. See ref.[20] for a more detailed description of the experiment.

6. Future Plans

Our plans in this field include the following experiments: (1) Gravitational-Induced Quantum Interference--stage III. (2) Effect of the Moon and Sun upon Quantum Interference. (3) Multi-Photon Exchange with a Neutron. (4) Neutron Michelson-Morley experiment. (5) Wheeler-Delayed-Choice experiment.

7. Acknowledgments

This work is supported by the Physics Division of the NSF (Grant No. PHY-9024608), the U.S.-Austria program at NSF (Grant NO. INT-8712122), and Fonds zur Förderung der Wissenschaftlichen Forschung (Projekt No. S4201) in Austria. Author's photograph by M. Seewoster.

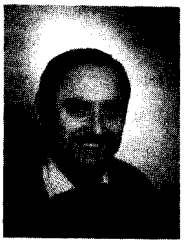
8. References

1. U. Bonse and M. Hart, Appl. Phys. Lett. 6, 155 (1965).
2. H. Rauch, W. Treimer and U. Bonse, Phys. Lett. **47A**, 425 (1974).
3. R. Colella, A.W. Overhauser and S.A. Werner, Phys. Rev. Lett. 34, 1472 (1975).
4. A.I. Ioffe, V.S. Zabiyakin and G.M. Drabkin, Phys. Lett. **A111**,373 (1985).
5. M. Gruber, K. Eder, A. Zeilinger, R. Gähler, and W. Mampe, Phys. Lett. **A140**, 363 (1989).
6. K. Eder, M. Gruber, **A**. Zeilinger, R. Gähler, and W. Mampe, Physica B**172**, 329 (1991).
7. B.G. Levi, Physics Today **44**, #7, 17 (1991).
8. J.-L. Staudenmann, S.A. Werner, R. Colella and A.W. Overhauser, Phys. Rev. **A21**, 1419 (1980).
9. S.A. Werner, H. Kaiser, M. Arif and R. Clothier, Physica B**151**, 22 (1988).
10. H. P.Layer and G.L. Greene, Phys. Lett. **A155**, 450 (1991).
11. Y. Aharonov and Bohm, Phys. Rev. **115**, 485 (1959).
12. Y. Aharonov and A. Casher, Phys. Rev. Lett. **53**, 319 (1984).
13. A. Cimmino, G.I. Opat, A.G. Klein, H. Kaiser, S.A. Werner, M. Arif and R. Clothier, Phys. Rev. Lett. 63, 380 (1989).
14. H. Kaiser, S.A. Werner, R. Clothier, M. Arif, A.G. Klein, G.I. Opat and A. Cimmino, Proceedings of ICAP-12, Ann Arbor, Michigan, July 29-August 3, 1990. AIP Conference Proceedings **233**,"Atomic Physics 12," ed. J.C. Zorn and R.R. Lewis (AIP, New York, 1991), pp. 244-268.
15. A. Zeilinger, "Fundamental Aspects of Quantum Theory", ed. by V. Gorini and A. Frigerio, pp. 311-318, Plenum Publishing Corp. (1986).
16. B.E. Allman, A. Cimmino, A.G. Klein, G.I. Opat, H. Kaiser and S.A. Werner, Phys. Rev. Lett. **68**, 2409 (1992).
17. S.A. Werner, R. Clothier, H. Kaiser, H. Rauch and H. Wölwitsch, Phys. Rev. Lett. 67, 683 (1991).
18. H. Kaiser, R. Clothier, S.A. Werner, H. Rauch and H. Wölwitsch, Phys. Rev. **A45**, 31 (1992).
19. R. Clothier, H. Kaiser, S.A. Werner, H. Rauch and H. Wölwitsch, Phys. Rev. **A44**, 5357 (1991).
20. H. Rauch, H. Wölwitsch, R. Clothier, H. Kaiser and S.A. Werner, Phys. Rev. **A** , to be published in July 1992.

NEUTRON SPIN-ROTATION AND INTERFEROMETRIC MEASUREMENT OF GEOMETRIC PHASES AND THE MAGNETIC AHARONOV - CASHER EFFECT

G. Badurek, S. Wehinger and H. Weinfurter*
Institut f. Kernphysik, Technical University of Vienna, Schüttelstrasse 115,
A-1020 Vienna, Austria
*) Institut f. Experimentalphysik, University of Innsbruck, Technikerstrasse 25,
A-6020 Austria

presented by
Gerald Badurek



Abstract

According to its intrinsic interference nature neutron spin rotation in magnetic fields can be used to measure geometrically induced phase shifts of quantum systems. The most recent experiment of this kind is presented together with a proposal of an explicit interference experiment with spatially separated neutron beams. Furthermore we propose a neutron spin rotation demonstration of the so-called "magnetic" or "scalar" Aharonov-Casher effect".

1. Introduction

In his famous article of 1984 Berry[1] showed that the wave function of a quantum system which evolves adiabatically in a slowly varying environment besides the usual dynamical phase acquires an additional phase shift which is of purely geometric origin. The classical analogue of Berry's phase is the so-called Hannay angle[2], which is of geometric nature, too, and refers to the adiabatic dynamics of classical mechanics, that plays an essential role e.g. in Foucault's pendulum. Both the classical and the quantum concepts of geometric phases are closely related to parallel transport at curved surfaces[3] and gauge field theories[4]. Berry's original phase was specified only for cyclic adiabatic evolutions in parameter space, that means when some external parameters of the underlying Hamiltonian are cyclically varied at sufficiently slow rate. An important generalization of Berry's phase was given by Aharonov and Anandan[5], who considered cyclic evolutions of quantum states in the projective space of rays in Hilbert space. They could show that a geometric phase (AA-phase) appears even for nonadiabatic and noncyclic variations of the Hamiltonian, provided the quantum state itself is cyclic. They also demonstrated that the Aharonov-Bohm effect[6] is a special case of the AA-phase. Subsequently it has been proven that even the condition of cyclic evolution in ray space is dispensable[7]. Geometric phases should hence be observable even for non-cyclic quantum states transported along an open path in parameter space.

Since Berry's pioneering work more than 450 articles were published[8] and evidence of geometric phase shifts was found in a variety of different experiments ranging from the observation of the rotation of the polarization of light in twisted optical fibres[9] to optical interferometry[10], molecular energy level spectroscopy[11], nuclear magnetic resonance[12], electron diffraction[13] and neutron spin rotation in slowly varying magnetic fields[14]. An excellent review of geometric phases in physics is given in Ref. 15.

The above mentioned example of a spin in a slowly changing magnetic field actually became a paradigm in the observation of geometric phase effects. Our own contribution, which will be briefly summarized in Section 2 was a neutron spin rotation experiment with time-dependent magnetic fields[16]. It represents the first demonstration of geometric phase shifts for non-cyclic evolution of a quantum state.

Although the measurement of a phase shift implies the superposition of states, one should note that in most of the experiments reported thus far no explicitely separated quantum states were at hand. Neutron interferometry offers an elegant possibility to perfonn a "genuine" interference experiment with two spatially separated quantum states which can be manipulated completely independent from each other. Following essentially the ideas of Wagh and Rakhecha[17] we propose such an interference experiment with polarized neutrons. Finally we also propose a neutron spin rotation measurement of the "magnetic" Aharonov-Casher effect which has very recently been verified by neutron interferometry.[18]

2. Geometric phase shift of a spin-½ system for non-cyclic evolution

According to Berry's original paper an adiabatic cyclic variation C of the external parameters returns the system to its original state, but with a phase shift that can be factorized in a dynamical part and a

purely geometric phase $\exp\{i\Gamma(C)\}$ which is independent of the rate of evolution

$$|\psi(T)\rangle = \exp\{-\frac{i}{\hbar}\int_0^T E(t)dt\} \cdot \exp\{i\Gamma(C)\}|\psi(0)\rangle \qquad (1)$$

In the case of a spin-½ particle in a magnetic field it turns out that $\Gamma = -\frac{1}{2}\Omega(C)$, where $\Omega(C)$ is the solid angle the field **B** subtends around its variation along C[1]. Let us consider now the special case of a neutron in a magnetic field rotating in the $\hat{x}\hat{y}$-plane perpendicular to the beam axis (see Fig.1). Initially the eigenvectors of the Hamiltonian $\mathcal{H}(t) = -\mu\sigma\mathbf{B}(t)$ are aligned along $\pm\hat{x}$ and evolve in time according to

$$|n(t)\rangle_{\pm} = U(t)|n(0)\rangle_{\pm}, \quad \text{with} \quad U(t) = \exp(-i\omega t\sigma_x/2) \qquad (2)$$

In the adiabatic limit the unitary time evolution operator $U(t)$ approximately diagonalizes the Hamiltonian, too. Therefore the approximate solution of the Pauli equation yields a wavefunction of the form

$$|\psi(\tau)\rangle_{ad} = U(\tau)D(\tau)|\psi(0)\rangle \quad , \qquad (3)$$

where $D(\tau) = \exp(-i\sigma_x\omega_L\tau/2)$ is the exact solution of the adiabatic Hamiltonian in the rotationg frame. It describes the dynamic evolution of the system whereas $U(\tau)$ is characterized by the geometry in parameter space.

If the evolution is not cyclic and the system does not return to its initial state, how can then a phase relation be established between the initial and the final state? Pancharatnam's connection[19] which was shown to be equivalent with the adiabatic connection of states[20] defines a phase difference γ between any two states by the phase of their scalar product $\langle A|B\rangle = \exp(i\gamma)|\langle A|B\rangle|$, provided these states are not orthogonal ($|\langle A|B\rangle| \neq 0$). The states are hence are said to be "*in phase*" if their scalar product is real and positive. According to Pancharatnam's connection the phase change of the wavefunction after some evolution time interval τ follows from

$$\langle\psi(0)|\psi(\tau)\rangle_{ad} = \langle\psi(0)|U(\tau)D(\tau)|\psi(0)\rangle \, . \qquad (4)$$

For our special field configuration the dynamical evolution operator is $D(\tau) = \exp(-i\sigma_x\omega_L\tau/2)$. If the initial state is parallel to the eigenstate $|n_+(0)\rangle$ of the Hamiltonian one obtains

$$\langle n_+(0)|U(\tau)D(\tau)|n_+(0) = \exp(-i\omega\tau_L/2)\langle n_+(0)|U(\tau)|n_+(0)\rangle \, . \qquad (5)$$

Because the eigenvectors are perpendicular to the rotation direction of the field the adiabatic solution describes their evolution along a geodesic arc on the Poincaré sphere, which never leads to a geometric phase shift[7,20]. This can be inferred immediately from $\langle n_+(0)|U(\tau)|n_+(0)\rangle = 2\cos(\omega\tau/2)$,which is always real and positive for $|\omega\tau| < \pi$ (i.e. for the shorter of the two geodesic arcs on the Poincaré sphere connecting the initial and the final state). However, if one starts with an eigenstate $|m\rangle$ of $U(t)$ it follows immediately from the scalar product

$$\langle m|U(\tau)D(\tau)|m\rangle = \exp(i\gamma_m)\langle m|D(\tau)|m\rangle \qquad (6)$$

that in this case the phase shift contains the full geometric information about the evolution. Since $|m\rangle$ remains parallel to the $\hat{z}$-direction during the evolution one can observe the geometric phase effect at any time, and not only for cyclic trajectories. For instance, with $|m\rangle = |+z\rangle$ one thus obtains

$$\langle +z|U(\tau)D(\tau)|+z\rangle = \exp(-i\omega\tau/2)\cos(\omega_L\tau/2) \quad , \qquad (7)$$

where the geometric phase shift $\gamma_+ \equiv \Gamma_+ = -\omega\tau/2$ is just minus half the angle the field direction subtends during the evolution time interval τ. The dynamical phase shift $\delta_+ = \omega_L\tau/2$ is due to the spin precession around **B** with Larmor frequency $\omega_L = 2\mu B/\hbar$.

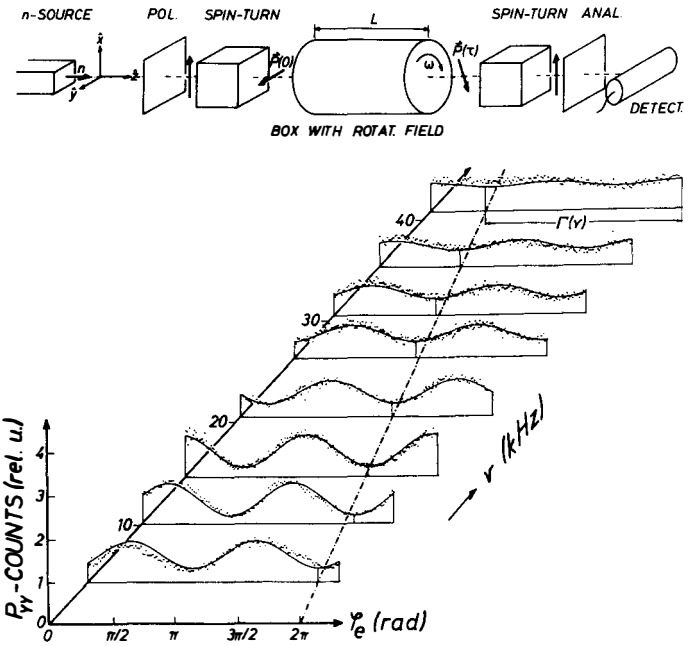


Fig.1: Experimental setup to study the Berry phase for noncyclic evolution and stroboscopically measured frequency dependence of the polarization P_{yy} change along the $\hat{y}$-direction in the adiabatic limit (B=3.8 mT). The geometric phase shift Γ increases linearly with rising field rotation frequency (broken line).

To detect experimentally such a geometrically induced phase shift one can exploit the intrinsic interference character of neutron spin rotation. Each arbitrary spin state of the neutron can be interpreted as a coherent superposition of two orthogonal base states $|\pm z\rangle$ and consequently any phase shift between these two states manifests itself by an orientational change of the neutron polarization vector $\mathbf{P} \equiv \langle\sigma\rangle$.

For our rotating field configuration this means that in the adiabatic limit $\omega_L >> \omega$ the two interfering states experience geometric phase shifts $\Gamma_\pm = \mp\omega\tau/2$ of opposite sign, yielding a total phase difference $\Gamma = \Gamma_+ - \Gamma_- = \omega\tau$ between them. This geometric phase can be separated from their dynamical phase difference $\delta = \delta_+ - \delta_- = \omega_L\tau$ if for given evolution time τ and field amplitude $|B|$ and hence constant dynamical phase shift δ, for instance, the polarization change $P_{yy} \equiv P_y(\tau)/P_y(0)$ along the $\hat{y}$-direction

$$P_{yy} = \cos\Gamma\cos^2(\delta/2) + \cos(2\varphi_e - \Gamma)\sin^2(\delta/2) \qquad (8)$$

is measured stroboscopically as a function of the angular position φ_e of the rotating field at the entrance moment of the neutrons. The broken line in Fig.1 indicates clearly the linear increase of the geometric phase difference between the $|\pm z\rangle$ states with increasing field rotation angle $\omega\tau$.

3. Neutron interferometric measurement of geometric phases

Some years ago Badurek, Rauch and Tuppinger[21] reported on a quantum beating experiment with polarized neutrons where the spin state of each of the two coherent subbeams of a Mach-Zehnder-type neutron interferometer was inverted by resonant neutron-photon interaction within separated oscillatory magnetic fields. Reinterpreting this experiment Wagh and Rakhecha[17] recently proposed to measure geometric phase shifts by means of neutron interferometry. We have invited them to perform together such an experiment at the perfect crystal interferometer instrument S18 of the ILL Grenoble after the repairment of the reactor. Its scheme is depicted in Fig.2. Two identical (static) spin-flippers $SF_{1,2}$ are inserted in the two beam paths inverting the initial $|+z\rangle$ or "up" state to the $|-z\rangle$ or "down" state with respect to the quantization axis defined by a homogeneous magnetic field in the $\hat{z}$-direction which extends over the whole interferometer region. As indicated, these two flippers are rotated about the $\hat{z}$-axis by angles $\pm\frac{1}{2}\beta$ with respect to the wave front.

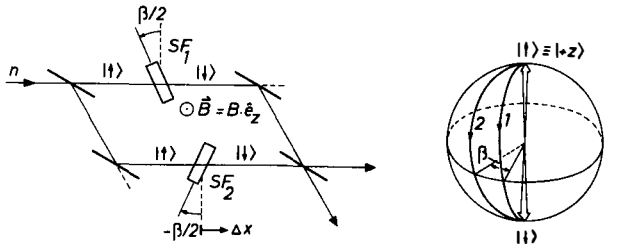


Fig.2: On the observation of the Aharonov-Ananadan phase with a Mach-Zehnder neutron interferometer and two spin-flippers SF_1, SF_2 which are rotated around the $\hat{z}$-axis by angles $\pm\beta/2$. The spin trajectories *1* and *2* of these flippers enclose a solid angle -2β (after Wagh and Rakhecha[17]).

The operator F_2 of the flipping action of flipper SF_2 can therefore be expressed by $F_2 = UF_1U^{-1}$, with

the unitary spinor rotation operator $U = \exp(-i½\beta\sigma_z)$. Since for both flippers $F|+z\rangle = |-z\rangle$ holds per definition, it follows that

$$F_2|\pm z\rangle = \exp(\mp i\beta)F_1|\pm z\rangle \ . \tag{9}$$

Depending on its initial spin state the neutron thus picks up a phase shift $\pm\beta$ of obvious geometric origin on its passage through the interferometer in addition to the dynamical phase, which is due to the excess Larmor precession in one of the interferometer arms. This additional phase shift is just minus half the solid angle 2β subtended by the closed loop formed by the trajectories *1* and *2* on the spin sphere when followed in mutually opposite senses (see Fig.2). The dynamical phase would vary by $\Delta\delta = \omega_L\Delta x/v_n$, where v_n is the neutron velocity, if one of the flip devices is moved a distance Δx along the beam path, but like all the other phase shift contributions (nuclear, gravitational, intrinsic phase etc.) it would not change upon rotation of the flippers around the $\hat{z}$-axis. Hence a separation of the geometric contribution to the total phase could be easily achieved. Evidently, because of the rotation the effective thickness of the flipper coils varies. Consequently their supply current has to be readjusted in order to maintain exact spin reversal.

However, one should notice that in this proposed experiment one will measure the Aharonov-Anandan (AA) phase and *not* the Berry phase, since here only the "system" and not its environment evolves in time and, moreover, this occurs *nonadiabatically.*

4. The magnetic Aharonov-Casher effect

In 1959 Aharonov and Bohm[6] published their famous proposal of electron interference experiments, known since then as the AB effects, which stressed the importance of potentials in quantum mechanics, unlike to classical mechanics where they are considered only as convenient mathemathical tools for calculating electromagnetic fields of force. The AB effects are illustrious examples of quantum-non-locality, because they predict an observable phase shift of the electron's de Broglie wavepacket which depends on fields in regions of space *not* accessible to the interfering electron. Hence there is *no* force acting on the particle, the phase shift is entirely due to non-zero potentials, namely the vector potential $\mathbf{A}(\mathbf{r})$ in the so-called "magnetic" (or "vector") AB effect and the scalar potential φ in the less often cited "electric" (or "scalar") AB effect (Fig.3). Whereas the vector AB effect was verified experimentally long ago[22], until now technical difficulties prevented the realization of the conceptually simpler scalar phenomenon, which requires to raise and subsequently to lower again the electric potential of one of the two Faraday cages placed along the beams in an electron interferometer while the electron wavepacket propagates just inside this cage.

The neutron analogue of the vector AB effect, the so-called Aharonov-Casher (AC) effect[23], refers to the phase shift that is induced by placing an electrically charged wire between the two coherent subbeams

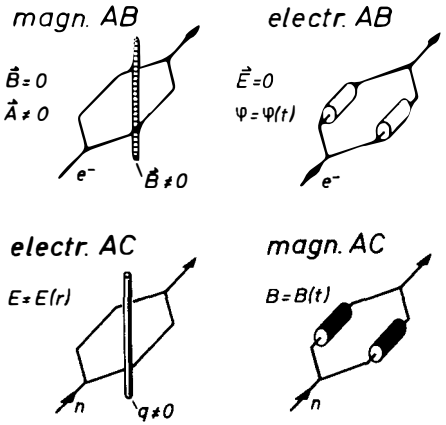


Fig.3: The Aharonov-Bohm (AB) effects and their Aharonov-Casher neutron analogues.

of a neutron interferometer. The first experiment of this kind was reported by Cimmino et al[24]. It is important to note that because of symmetry arguments this "electric" AC effect is of topological nature only if the wire is oriented normal to the interferometer plane[24].

Very recently Allman et al.[18] have realized experimentally also the neutron analogue of the scalar AB effect, which for consistency reasons we shall denote here as the scalar or "magnetic" AC effect (see Fig.3). There the two interfering beams traverse two identical solenoids, one of them biased by a DC current but the other pulsed in such a way that a homogeneous magnetic field B(t) is switched on and off for a certain time interval τ while the neutron is in this field region inside the coil. Then no force acts on the neutron and the phase shift

$$\Delta\phi_{EAC} = \pm\frac{\mu}{\hbar}\int_0^{\tau} B(t)dt \,, \tag{10}$$

where the sign belongs to the two possible spin eigenstates, is nondispersive, i.e. independent of the neutron wave number k. However, in the cited experiment thermal neutrons were used and, though the predicted phase shift was found, it was not possible to confirm unambigously that really no force acted on the neutron. With very cold neutron interferometry[25] this goal should be achievable, however.

Again, neutron spin rotation alone without reference to an explicit interference experiment could serve to demonstrate the essential feature of the scalar AC effect, the non-dispersivity of the phase shift

$$\frac{\partial}{\partial k}\Delta\phi_{EAC} = 0 \,. \tag{11}$$

The proposed experimental arrangement is shown in Fig.4. By means of a $\pi/2$-spin-turn device the polarization of a chopped neutron beam with finite momentum distribution $\Delta k/k$ is oriented orthogonal to the direction of a homgeneous magnetic field produced by a coil of length L. A second spin-turn behind this precession coil rotates the polarization vector back to its initial (and analyzing) direction. If the field is static then the total Larmor precession angle $\phi=\phi(k)$ is dispersive. For sufficiently large L the probability W_+ to find the neutron in the initial spin state will be 50%. However, if the field is switched on and off for a short time interval τ when the (spreading) neutron ensemble is located just in the central region of the coil, the precession phase is equal for all neutrons $\phi=\omega\tau$. In this case W_+ is a function of both ϕ and the polarisation parameter p ($-1\leq p\leq 1$) of the incident beam. For instance, for complete "up" polarization ($p=1$) one obtains $W_+ = \cos^2(\phi/2)$, as clearly seen in Fig.4.

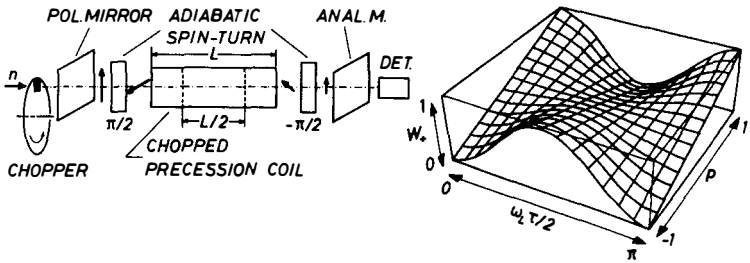


Fig.4: Proposed eperimental neutron spin-rotation setup to demonstrate the magnetic AC-effect. Whereas in a static precession field for the probability W_+ to find the neutron finally in the "up" state is 50% for sufficiently large L and spectral width $\Delta k/k$, W_+ is a function of the Larmor angle $\omega_L\tau$ and the polarization parameter p if the field is properly pulsed during the neutron transit time.

We wish to thank Prof. H. Rauch for valuable discussions and the Austrian "Fonds zur Förderung der Wissenschaftlichen Forschung" for financial support (Projects S42/01 and S42/08).

References

1. M. Berry, *Proc. Roy. Soc. London A*, **392** (1984) 451.
2. J.H. Hannay, *J. Phys. A*, **18** (1985) 221.
3. B. Simon, *Phys. Rev. Lett.*, **51** (1983) 2167.
4. F. Wilczek and A. Zee, *Phys. Rev. Lett.*, **52** (1984) 2111.
5. Y. Aharonov and J. Anandan, *Phys. Rev. Lett.*, **58** (1987) 1593.
6. Y. Aharonov and D. Bohm, *Phys. Rev.*, **115** (1959) 485.
7. J. Samuel and R. Bhandari, *Phys. Rev. Lett.*, **60** (1988) 2339.
8. Citation index.
9. A. Tomita and R.Y. Chiao, *Phys. Rev. Lett.*, **57** (1986) 937.
10. H. Jiao, S.R. Wilkinson and R.Y. Chiao, *Phys. Rev. A*, **39** (1989) 3475.
11. G. Delacrétaz, E.R. Grant, R.L. Whetten, L. Wöste and J.W. Zwanziger, *Phys. Rev. Lett.*, **56** (1986) 2598.
12. R. Tyoko, *Phys. Rev. Lett.*, **58** (1987) 2281; D. Suter, G.C. Chingas,R.A. Harris and A. Pines, *Mol. Phys.*, **61** (1987) 1327.
13. D.M. Bird and A.R. Preston, *Phys. Rev. Lett.*, **61** (1988) 2863.
14. T. Bitter and D. Dubbers, *Phys. Rev. Lett.*, **59** (1987) 251; D. Richardson, A.I. Kilvington, K. Green and S.K. Lamoreaux, *ibid.*, **64** (1988) 2030.
15. *Geometric Phases in Physics*, eds. F. Wilczek and A. Zee (World Scientific, 1989).
16. H. Weinfurter and G. Badurek, *Phys. Rev. Lett.*, **64** (1990) 1318.
17. A.G. Wagh and Rakhecha, *Phys. Lett. A*, **148** (1990) 17.
18. B.E. Allman, A. Cimmino, A.G. Klein, G.I. Opat, H. Kaiser and S.A. Werner, *Phys. Rev. Lett.*, in print; see also H. Kaiser *et al.* in these proceedings.
19. S. Pancharatnam, *Proc. Ind. Acad. Sci. A*, **44** (1956) 247.
20. M.V. Berry, *J. Mod. Opt.*, **34** (1987) 1401.
21. G. Badurek, H. Rauch and D. Tuppinger, *Phys. Rev. A*, **34** (1986) 2600.
22. R.G. Chambers, *Phys. Rev. Lett.* **5** (1960) 3.
23. Y. Aharonov and A. Casher, *Phys. Rev. Lett.*, **53** (1984) 319; A. Zeilinger, in *Fundamental Aspects of Quantum Theory*, Plenum Publ. Corp. 1986.
24. A. Cimmino, G.I. Opat, A.G. Klein, H. Kaiser, S.A. Werner, M. Arif and R. Clothier, *Phys. Rev. Lett.*, **63** (1989) 3803.
25. M. Gruber, K. Eder, A. Zeilinger, R. Gähler and W. Mampe, *Phys. Lett. A*, **140** (1989) 363.

Are Quantum Particles *Really* Particulate?
or
Bell theorems without inequalities for two and three spinless particles[1]

Herbert J. Bernstein
School of Natural Science
Hampshire College
Amherst, MA 01002
U.S.A.

Abstract

The Einstein, Podolsky, and Rosen[2] (EPR) program insisting on local realism was first enunciated in terms of experiments with definite outcomes, performed on a system of two spinless particles. Bell's[3] famous proof that quantum mechanics is incompatible with EPR local realism needed the application of their program to experiments with probabilistic outcomes. Can one find a similar 'Bell's theorem' result, using only experiments with definite outcomes?

We have found such a disproof of local realism for two spinless particles, but it requires a different extension of the EPR approach. This extension adopts an additional assumption beyond those enunciated by EPR, an assumption consistent with a very strongly particulate view of the nature of a quantum mechanical system. Moreover, ***without*** extending the EPR program, one can still prove EPR local realism incompatible with quantum mechanics, using only experiments with definite outcomes, if one considers a three particle system.[4]

1. Introduction

The famous paper of Einstein, Podolsky, and Rosen[2] (EPR) explicitly pursued a program based on perfect correlations of experimental outcomes. These are the definite outcomes that would surely be the same in any form of theory: probabilistic, stochastic, or deterministic, for they involve only probabilities of unity or zero. EPR considered a variety of space-like separated measurements that could be made on two non-interacting particles and introduced a very specific viewpoint concerning local reality. "Since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system" [***locality***]. "If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity" [***reality***].

It has long been known that these assumptions, (together with a third one, completeness of the theoretical description) reasonable as they appear to be, are inconsistent with quantum mechanics. This fact was discovered by Bell, who considered the system introduced by Bohm of a spin-0 particle decaying into two spin-1/2 particles. In the simplest gedanken experiment, when the spin component of one of the decay particles is measured in a certain direction, quantum mechanics predicts that with 100% certainty its partner's spin will be opposite, if measured in the same direction. This behavior is sufficient to allow the use of the EPR locality and reality criteria to establish that this particular spin component of the second particle is an element of reality. By parallel arguments all spin components of both particles are such elements, and their reality must be established at the time of the original decay, even if they cannot all be simultaneously known experimentally. We call such a simple case, where the result of one measurement can be used to predict with 100% certainty the result of another measurement not yet performed, a case of perfect correlation. For two spin-½ particles one may easily construct a deterministic, local, realistic model that can reproduce these perfect correlations, regardless of the direction in which the original spin measurement is made.

Bell[3] realized, however, that for more general measurements, where the spin components of the two particles are measured along different directions, the quantum mechanical behavior cannot be modeled by any local, realistic, deterministic (or stochastic) model, thus upsetting the EPR program. He showed that starting from the EPR program one can derive an inequality, which must be obeyed by *any* such model, but which is violated by quantum theory. His proof refers to the imperfectly correlated cases, when the

particles are measured along different directions. Bell's theorem says nothing about the case of perfect correlations.

In the present paper, we investigate the possibility of applying perfect correlation type arguments to two particles. We find that such an extension exists if one considers, instead of the more familiar Bohm two-spin system, a two-particle interferometer. The important physical difference between these cases is the number of paths from source to detector. Whereas each particle in the Bohm system has only one available path from the source to the polarization analyzer, each particle in the two particle interferometer has two available paths through the interferometer. As we shall see below, within the EPR viewpoint one and only one of these paths is actually taken by any given particle, and the path taken is an element of reality.

However in order to complete the two-particle argument we find that we must augment the original EPR assumptions given in the first paragraph above with the following auxiliary assumption:

> If the path taken by a particle is an element of reality then no entity associated with the particle in any way samples alternate paths (i.e., the particles are truly *particulate*).[5] [***particulate paths***]

This in turn implies that the particle can be affected only by manipulations along its path, and cannot be affected by any manipulations that take place along alternate paths not taken. We believe that our assumption is one possible natural outgrowth of the EPR viewpoint, but in any case its consequences will be spelled out so that one may judge its plausibility. The particulate path (PP) assumption is, in part, compatible with quantum mechanics in the sense that when a particle is experimentally known to be in one beam path, there is zero amplitude for it to be in any other path. However, EPR assert the existence of an established element of reality, connecting the particle to a definite path, even in cases where one has an experimental arrangement that cannot ascertain such information. Of course, quantum theory denies this possibility.

2. Two Particle Proof Using the Particulate-Path Assumption

Consider the two-particle interferometer arrangement[6] of Fig.1, where a central source emits pairs of spinless particles, 1 and 2, in opposite directions. An experimental

apparatus defines two directions, *a-a'* and *b-b'*. Then the initial quantum state of a pair is

$$|\Psi_0\rangle = (|a\rangle_1|a'\rangle_2 + |b\rangle_1|b'\rangle_2)/\sqrt{2} \qquad (1)$$

where ket $|a\rangle_1$ denotes particle 1 in beam *a* , etc.

For the demonstration that the EPR program fails for two particles if it is supplemented by the PP assumption, we will consider a sequence of four different experiments with this arrangement. In our first experiment we establish that for each particle the path taken inside the interferometer is an EPR "element of reality" and, at the fourth experiment, we will find that the "reality" of this internal path necessarily implies a total disagreement (i.e., a disagreement for each pair detected) with the quantum-mechanical predictions of the state (1).

For the first experiment, move the detectors D directly into the beams *a, a' b, b'*. State (1) implies *a-a'* coincidences for half the pairs detected and *b-b'* coincidences for the other half of the pairs. These perfect correlations and the EPR local reality assumptions immediately imply that the path taken, (*a'* or *b'*), by particle 2 is an element of reality, for if particle 1 is found in path *a* [*b*] then particle 2 will, "with certainty," be found in path *a'*[*b'*]. A parallel argument establishes the reality of path (*a* or *b*) for particle 1. In short, according to the EPR view, half the pairs emitted by the source ***really*** take the *a-a'* paths and the other half ***really*** take the *b-b'* paths.

Moreover, and we stress this point, in the EPR view these elements of reality must exist even when the detectors used to establish their existence are removed and one or both beam splitters are replaced, as in the next three experiments. Thus, even if we replace the beam splitter at position G -- and the detectors for particle 2 further downstream, beyond G (as shown on the right in Fig.1) -- the path inside the interferometer *a'* or *b'* can still be determined with certainty by counters in paths *a, b*. Of course these counters cannot predict which detector (+ or -) will be struck by particle 2, so we have not yet established the reality of a dichotomic observable (one giving the distinct results, + or -) which we call "**sign**." We do that next.

For the next three experiments, place beam splitters at C & G in the paths of both particles, and replace all four detectors D behind them, as shown. Then beam *a* passes through the phase shifter α, and *b'* through β, before recombining respectively with beams

b and *a'* in beam splitters at C and G. So the state (1) becomes:

$$|\Psi\rangle = \frac{ie^{i(\alpha+\beta)/2}}{\sqrt{2}}[(|+\rangle_1|+\rangle_2 + |-\rangle_1|-\rangle_2)\cos(\Delta/2) + (|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2)\sin(\Delta/2)], \quad (2)$$

$$where \Delta = \alpha - \beta.$$

In general, Eq.2[7] does NOT imply perfect correlations, but we shall consider only experiments with phase shifts (α and β) of 0 or π. These values **do** produce perfect correlations.

In the second experiment, set the phase shifts $\alpha = \beta = 0$, and again monitor for coincidences. With these phase values, the state (2) becomes

$$|\Psi\rangle = i(|+\rangle_1|+\rangle_2 + |-\rangle_1|-\rangle_2)/\sqrt{2} \quad (3)$$

at the output. State (3) predicts + + coincidences for half the pairs detected, and - - coincidences for the other pairs. There are no + - or - + coincidences. These perfect correlations and the EPR local reality assumptions immediately imply that the **sign** of each particle is an element of reality, for if particle 1 ends up at the + detector, then particle 2 will *always* strike its + detector. And similarly for negative **sign**. Every particle has a real **sign**, and moreover the source emits no pairs of mixed **sign**.

Because the reality of the internal path can be established even after that particle has left the interferometer, we are now in a position to invoke the PP assumption. This assumption implies that if particle 2 took the path *a'*, it cannot be affected by the particular value of the phase shifter β. Similarly, if particle 1 took path *b*, it cannot be affected by the value of the phase shifter after α. We will use these results in what follows.

For the third experiment, set both phase shifters α and β to π. Then by Eq.2 the state at the detectors will be

$$|\Psi\rangle = -i(|+\rangle_1|+\rangle_2 + |-\rangle_1|-\rangle_2)/\sqrt{2} \quad (4)$$

which also produces perfect correlations. This means that *a-a'* [*b-b'*] pairs, which experience the effect of a π phase shift at α [β], remain of the same **sign**. By locality only one of the pair could have been affected -- for the *a-a'* path, the *a* particle by α [the *b'* particle by β] -- so if the **sign** were changed by a π phase shift, some of the same-**sign** pairs produced by the source would have to be changed into opposite **sign** pairs. But none are detected, ever. We conclude that the **sign** is not only real (from results of the second experiment) but also that passing through a π phase shifter doesn't change the **sign** of a particle.

Finally, in the fourth experiment set α to 0 while leaving β at π. Then state (2) is

$$|\Psi\rangle = +(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2)/\sqrt{2}. \quad (5)$$

Quantum mechanics predicts only + - and - + coincidences will occur. But by comparison with previous experiments, we find the extended EPR view predicts something very different. The *a-a'* pairs encounter $\alpha = 0$ and do not encounter the phase shifter in path *b'*. According to the PP assumption, they cannot be influenced by its value β. We see that these pairs will still produce only the same sign coincidences, as in experiment two. Similarly, consider the *b-b'* pairs. They encounter $\beta = \pi$ in beam *b* and by the PP assumption, they cannot be affected by the $\alpha = 0$ setting in path *a*. Therefore as in the third experiment, these *b-b'* pairs also can only produce + + and - - (i.e. same-sign) coincidences. So the EPR view predicts *no* + - or - + coincidences will occur, in complete disagreement with quantum mechanical prediction. This is the announced contradiction with quantum mechanics. Quantum particles are NOT *real*ly particulate (in EPR's sense, as extended by the PP assumption).

3. The Three Particle Proof With No Additional Assumption

Clearly, the additional assumption weakens our proof; moreover, it is rather strongly against the spirit of most interpretations of quantum mechanics. Fortunately there is a proof that the *un*extended EPR local realistic program itself is incompatible with quantum mechanics, using only the case of perfect correlations. It requires however, the consideration of three-particle states. This Greenberger-Horne-Zeilinger (GHZ) proof is the first simple demonstration of a Bell theorem without inequalities. We present here a version based on work by Mermin which uses three spin-½ particles.[8]

Consider a system, much like the two-particle spin-0 system of Bohm mentioned in section 1 above, but whose source produces spin-correlated triplets (rather than pairs) of particles. Let the experimental arrangement result in a set of three spatially separated beams of particles (labeled 1, 2, and 3) each of which impinges on its own Stern-Gerlach apparatus capable of determining one component of the particle spin. We will need to account for both x- and y-components of the spin, so imagine that the three analyzing magnets may each take up either x- or y-axis orientations. For spin-½ particles, of course, every spin measurement results in one of two outcomes, + or -, corresponding to spin along or opposite the axis of the analyzer.

The proof proceeds, as before, by a series of four experiments. The first three establish EPR "elements of reality" for both the x- and y-components of the spin of each particle, while the fourth shows the contradiction. If the state impinging on the three spatially separated spin analyzers is the entangled superposition of spin-z states

$$|\Psi_{triple}\rangle = (|u\rangle_1|u\rangle_2|u\rangle_3 - |d\rangle_1|d\rangle_2|d\rangle_3)/\sqrt{2} \qquad (6)$$

(where $|u\rangle$ represents spin up along the z-axis and $|d\rangle$ represents down), then conditions exist for showing that EPR local realism contradicts quantum mechanics.

For the first three experiments, align two of the machines along the y-axis, and the third along the x-axis, and monitor for triple coincident counts. By symmetry it makes no difference which particle (i.e. which beam in the laboratory) meets the x-axis machine. The pattern of results is always the same. Whenever the two y-axis outcomes agree, the + detector on the x-axis machine fires. Whenever they disagree, the - shows up.[9]

But this pattern of triple coincidences is one of perfect correlation. By putting two of the Stern-Gerlach machines closer to the source than the third, we can predict "with certainty" what **its** outcome will be. Thus to show that the x-spin of the first particle is ***real***, simply put the x-axis machine in beam 1, and y-axis machines in the other two. Then we can predict that whenever the signs of the two y-spin outcomes are the same, the x-spin component of particle 1 will be positive, i.e. pointing along the x-direction; whenever the two spins we measure have opposite sign, the x-spin of particle 1 will be negative. On the EPR view, the x-spin component of particle 1 is an element of reality, and similarly for those of particles 2 and 3 as well.

These same perfect correlations and the EPR local reality assumptions immediately imply the reality of the y-spin for each of the particles as well, for we can set up an x-axis machine and a y-axis machine in the beams containing the other two particles, then predict with certainty what the outcome of a measurement of the y-spin will be. In fact, we can even turn the direction of the x-spin machine at the last minute, and show that whichever of the two (x- or y-spin) variables we *choose* can be predicted with certainty for the remote particle.

Now set all three machines to the x-axis. According to quantum mechanics, the outcome obeys

$$\sigma^x_1 \sigma^x_2 \sigma^x_3 |\Psi_{triple}\rangle = -|\Psi_{triple}\rangle, \qquad (7)$$

i.e. the product of the signs will be negative: an odd number of - outcomes will always appear. What does the EPR view predict? If both x-spin and y-spin are real, then the three particles either have all the same value of y-spin (whether + or -) or one of them differs from the other two. In the first case, all three pairs of particles agree for y-spin, so by the results of the first experiment, all three x-machines should have + counts show up. In the second case, one pair of y-spin values agrees while two disagree. Two negative x-spins should appear, and only one + value. In all cases, the product of the x-spin signs is positive, in direct contradiction with the prediction of Eq.7; EPR local realism, *without* any additional assumptions, contradicts quantum mechanics. We conclude that particles which obey quantum mechanics are NOT *real*y particulate.

REFERENCES

1. This article is based on work done in collaboration with DM Greenberger, MA Horne and A Zeilinger, Phys Rev A (to appear 1992), to which it is heavily indebted.

2. A Einstein, B Podolsky, and N Rosen, Phys Rev 47, p. 777 (1935).

3. JS Bell, Physics 1, #1, 195 (1964).

4. DM Greenberger, MA Horne, and A Zeilinger in M Kafatos, *ed.*, "Bell's Theorem, Quantum Theory, and Conceptions of the Universe," Kluwer Academic, Dordrecht (1989). The GHZ paper. Also, see GHZ (with A Shimony) Am J Phys. 58, 1131 (1990). The GHSZ paper.

5. The name used here emphasizes assuming a strongly PARTICULATE nature of quantum mechanical particles. The paper of ref.1 calls this assumption "emptiness of paths not taken:"

> ...One might initially suspect that this assumption is by itself tantamount to a denial of elementary quantum theory, since the prohibited entities include such subtleties as a deBroglie pilot wave, an Einstein "ghost" wave, any kind of information gathering Mermin fuzzy "cloud", a Wheeler "smoky dragon", or most emphatically, even a normal Schroedinger quantum mechanical amplitude. (We are tempted to call our assumption "the law of the excluded muddle".) That is, the assumption seems to be incompatible with single-particle interference, the most basic wave-mechanical phenomenon.
>
> This is not the case for two reasons. First, in a single particle interferometer one cannot predict which path the particle will take through the interferometer, the path taken therefore is not an element of reality. Second, in the two-particle interferometer where the path *can* be predicted and hence our assumption applies, there is in fact *no* single particle interference.

6. MA Horne & A Zeilinger introduced two-particle interferometry: see Lahti & Mittelstaedt, *eds.*, "Symposium on the Foundations of Modern Physics" World Scientific, Singapore (1985), p.435. See also MA Horne, A Shimony, & A Zeilinger, Phys Rev Lett. 62, 2209 (1989).

7. Here we have chosen the phase convention that reflection at a beam splitter entails a phase factor of i, while transmission does not change the phase at all. The conclusions do not depend on the choice of splitter convention.

8. ND Mermin, Am J Phys. 58, 731 (1991). GHSZ (ref. 4) present a mathematically analogous three-particle interferometer version of this demonstration, but the notation becomes more tedious, so we present only the more perspicuous case, based on spin measurements.

9. Indeed, we can summarize by noting that the product of the signs is always negative. There are four possible outcomes, and each occurs one quarter of the time: all three + detectors count, or the x-axis + detector counts, while both y-axis machines show -; that makes up half of coincidences. For the other half, two y-axis machines disagree (one quarter of the time they show + -, the other quarter - +) and the x-machine negative detector counts. Observe that individual counters all fire randomly and at the same rate, just as the individual counters in Fig.1 did.

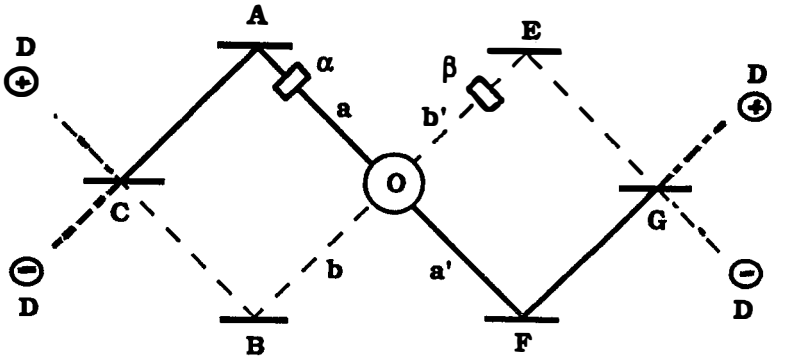


Fig. (1)

Two-Particle Interferometer A pair of particles, 1 & 2, emerges coherently at O, moving along the beams *a*-*a'*and *b*-*b'*, in the state of Eq.1. Beams *a* and *b* are deflected and pass through a beam splitter so that particle 1 ends up at a detector D (labelled + or -) at the left. Similarly for particle 2 at the right. There is one phase shifter, α, along path *a* and another, β, along *b'*. The outgoing state is given by Eq.2. Note that this state implies the two-particle coincidence rates (for + +, + -, - +, or - - counts) depend on α and β. Nonetheless, there is no single particle interference: the count rate at each detector D is independent of α and β.

TESTING THE STANDARD MODEL WITH NEUTRON BETA DECAY

M. S. Dewey, G. L. Greene and W. M. Snow
National Institute of Standards and Technology
Gaithersburg, MD 20899

Abstract

Within the framework of the standard electroweak model, observable weak interactions between baryons and leptons are manifestations of a relatively simple Hamiltonian describing the underlying behavior of quarks and leptons. Due to dynamical simplifications which arise at low energies, nuclear beta decay provides an attractive laboratory for such detailed studies of charged current weak interactions. The beta decay of the free neutron is particularly worthy of study due to the absence of perturbing nuclear effects and the possibility of measuring a variety of physical observables. This paper briefly outlines those features of the standard model which apply to neutron beta decay and describes the relevant phenomenology. It follows with a review of recent experimental results and concludes with a discussion of the constraints which may be set on possible deviations from the Weinberg-Salam model. Particular attention is given to the possible existence of right-handed currents.

1 Introduction

Within the framework of the standard electroweak model, observable weak interactions between baryons and leptons are manifestations of a relatively simple Hamiltonian describing the underlying behavior of quarks and leptons. For most low energy phenomena this essential simplicity is obscured by strong interaction effects. To interpret such phenomena, one must first recognize that the mass eigenstates of quarks are not eigenstates of the weak interaction. Next one must take account of the effects of the perturbation of the quark wave functions within the baryon due to the strong interaction. Finally, for phenomena involving nuclei, one must consider the significant perturbations induced by nuclear binding interactions between constituent nucleons. For a sensitive comparison between weak interaction theory and low energy experiments, it is necessary to find a physical system in which the extraction of detailed knowledge is possible.

Due to dynamical simplifications which arise at low energies, nuclear beta decay provides an attractive laboratory for such detailed studies of charged current weak interactions. The beta decay of the free neutron is particularly worthy of study due to the absence of perturbing nuclear effects. Neutron beta decay is also interesting since it is possible to measure a variety of physical observables (e.g. lifetime, correlation coefficients) which may be compared with one another and with theory.

In the past, measurements of the parameters describing neutron beta decay have lacked sufficient accuracy and internal consistency to provide a precise theoretical test. This situation has dramatically changed within the last few years with the publication of several novel experimental results. In particular, the accuracy in direct determinations of the neutron lifetime τ_n has improved to the few tenths of one percent level. This represents an improvement approaching an order of magnitude in the past decade. Additionally, measurements using vastly different measurement strategies are now in quite good agreement with one another. With the current data, it is possible to set rather stringent limits on possible deviations from the standard model

for weak interactions.

This paper briefly outlines those features of the standard model which apply to neutron beta decay and describes the relevant phenomenology. It follows with a review of recent experimental results. The paper concludes with a discussion of the constraints which may be set on possible deviations from the Weinberg-Salam model. Particular attention is given to the possible existence of right-handed currents.

This paper emphasizes the importance of neutron beta decay measurements for elucidating the nature of the weak interaction. However, it should also be noted that precise measurements of the parameters describing neutron decay are of importance in a variety of other realms. For example, an accurate value for the neutron lifetime is required for detailed calculations concerning big bang nucleosynthesis. [1,2] Among the implications of such calculations are constraints on the amount of universal baryonic matter. Such constraints lead to intriguing suggestions concerning the existence of large amounts of non-baryonic matter in the universe. Additionally, a reliable value for the matrix element describing neutron beta decay is important for the calculation of the solar neutrino flux and for understanding the details of stellar core collapse. [3]

2 Historical Review

While it is possible to begin with a general formulation of the standard model for the electroweak interaction and reduce it to the Hamiltonian describing beta decay, it is instructive to briefly review the historical line of reasoning which originally led to its formulation. This history began with the work of Fermi who, by an analogy to quantum electrodynamics, considered a point like interaction in which a current interacts with itself. [4] For electromagnetism the interaction

between a particle of charge e and an electromagnetic four vector potential A^μ is given by

$$H_{\rm em} = eJ_\mu A^\mu = -e\overline{\Psi}_e\gamma_\mu\Psi_e A^\mu. \tag{1}$$

Fermi proposed that the weak Hamiltonian describing beta decay is of the form

$$H_{\rm weak} = G_{\rm F}\left(\overline{\Psi}_p\gamma_\mu\Psi_n\right)\left(\overline{\Psi}_e\gamma^\mu\Psi_{\bar{\nu}_e}\right), \tag{2}$$

where $\left(\overline{\Psi}_p\gamma_\mu\Psi_n\right)$ and $\left(\overline{\Psi}_e\gamma_\mu\Psi_{\bar{\nu}_e}\right)$ are the hadronic and leptonic currents respectively and $\Psi_{n,p,e,\bar{\nu}_e}$ are respectively the neutron, proton, electron and anti-neutrino spinors. The presence of the γ_μ in Eq. 2 reflects the *vector* current nature of the Fermi Hamiltonian.

Rather soon after its introduction it became clear that a coupling with a purely vector character was not adequate to describe all observed nuclear beta decays. In particular, the so called Gamow-Teller decays require a current having an *axial vector* character. [5] Formally, the addition of such an axial vector coupling is simply a generalization of the original Fermi current. In principle, a current-current Hamiltonian may contain a variety of other couplings in addition to vector and axial vector. The most general expression for such a Hamiltonian in lowest order (assuming lepton conservation, Lorentz invariance and a pointlike interaction) can be written as

$$H_{\rm weak} = G_{\rm F}\sum_i c_i\left(\overline{\Psi}_p O_i\Psi_n\right)\left(\overline{\Psi}_e O_i\Psi_{\bar{\nu}_e}\right) \quad \text{for } O_i = 1, \gamma_\mu, \gamma_5, \gamma_\mu\gamma_5, \sigma_{\mu\nu}. \tag{3}$$

The five operators O_i correspond respectively to scalar (S), vector (V), pseudoscalar (P), axial vector (A) and tensor (T) currents. A careful review of the experimental data revealed that only a vector and axial vector current were required to account for the experimentally observed weak interactions. [6]

The observation of parity violation [7] and the direct determination of the neutrino helicity [8] led to the present interpretation that (within the conventional sign convention) the phases of the vector and axial vector currents are opposite; thus the "$V-A$" designation of the weak interaction. The Hamiltonian for such an interaction is described by two free parameters, conventionally denoted as g_{V}, the vector coupling constant, and g_{A}, the axial vector coupling constant. The Hamiltonian may be written as

$$H_{\mathrm{weak}} = G_{\mathrm{F}}\, \overline{\Psi}_p \gamma_\mu \left(1 - \frac{g_{\mathrm{A}}}{g_{\mathrm{V}}}\gamma_5\right) \Psi_n\, \overline{\Psi}_e \gamma^\mu (1-\gamma_5)\, \Psi_{\bar{\nu}_e}. \tag{4}$$

As will be shown below, a similar result may be obtained from a more modern perspective.

3 The Standard Model in Low Energy Beta Decay

The energy scale for neutron beta decay ($Q = 782\,\mathrm{keV}$) is much less than the mass of the W vector gauge boson ($M_W \approx 80\,\mathrm{GeV}/c$). Thus, the range of the weak interaction will be much smaller than any dynamical length scale in the decay. This implies that the weak interaction will be nearly point like and may be described by a current-current Hamiltonian of the type anticipated by Fermi *with the important distinction that the interacting particles are now quarks and leptons.* The total charged weak current J_μ will be given by the sum of a hadronic component [9–11]

$$j_\mu^{\mathrm{H}} = \overline{d}'\gamma_\mu (1-\gamma_5)\, u + \overline{s}'\gamma_\mu (1-\gamma_5)\, c + \overline{b}'\gamma_\mu (1-\gamma_5)\, t \tag{5a}$$

and a leptonic leptonic component

$$j_\mu^{\mathrm{L}} = \overline{e}\gamma_\mu (1-\gamma_5)\, \bar{\nu}_e + \overline{\mu}\gamma_\mu (1-\gamma_5)\, \nu_\mu + \overline{\tau}\gamma_\mu (1-\gamma_5)\, \nu_\tau. \tag{5b}$$

The Hamiltonian will have the form

$$H_{\text{weak}} = G_{\text{F}}\Big(J_\mu J^{\mu\dagger} + J^{\mu\dagger} J_\mu\Big). \tag{5c}$$

In the standard model, the strength of the weak interaction is related to the mass of the W gauge boson by $G_{\text{F}} = (\sqrt{2}e^2)/(8M_W^2 \sin^2\Theta_W)$, where Θ_W is the weak mixing angle. The primed spinors in Eq. 5a reflect the observation that the mass eigenstates (unprimed) of the quarks are not eigenstates of the weak interaction (primed). The weak eigenstates are related to the mass eigenstates by the Kobayashi-Maskawa mixing matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} c_1 & s_1c_3 & s_1s_3 \\ -s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ -s_1c_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{6}$$

This mixing does not occur for the leptons; thus nuclear beta decay amplitudes are smaller than lepton decay amplitudes by a factor which is historically known as $\cos\Theta_c$, the cosine of the Cabbibo angle ($\cos\Theta_c \equiv c_1 \equiv \cos\theta_1$).

The presence of the factor $(1-\gamma_5)$ again reflects, within the chosen sign convention, the $V-A$ nature of the interaction. The operator $(1-\gamma_5)/2$ is, for massless particles, the projection operator onto a pure state having left-handed helicity; it is for this reason that the weak interactions are often called "left-handed". The equality of the vector and axial vector amplitude reflects the "maximal" parity violating nature of the weak interaction *at the fundamental level.*

It is important to note that the left-handed nature of the standard model does not (at least insofar as it is now understood) represent a manifestation of some essential underlying physical principle. Rather the left-handedness has been put into the theory "by hand" to reflect observed phenomena. Whether or not this handedness is truly fundamental or whether it is

only a manifestation of an essentially symmetric theory at some much higher energy scale is unknown. An important feature of the study of low energy beta decay is its ability to shed light on this issue. This matter will be revisited in section 6.

In order to confront experiment in a detailed fashion it is necessary to apply the Hamiltonian of Eq. 5c to a real physical system. In nuclear beta decay the essential simplicity of Eqs. 5a-5c is complicated by the fact that the quarks are bound within nucleons. The strong quark-quark interaction significantly perturbs the wavefunction of the "decaying" d quark within the neutron. This perturbation can give rise to several additional amplitudes in the expression for the hadronic current in the most general case. However, if the interaction is assumed to be invariant under CP and Lorentz invariance is required, then number of unknown (energy dependent) amplitudes reduces to six and the hadronic current may be written as

$$\begin{aligned} j_\mu^{\mathrm{H}} = c_1 \overline{p} \Big[& f_1(q^2)\gamma_\mu - g_1(q^2)\gamma_\mu\gamma_5 \\ & + i f_2(q^2)\sigma_{\mu\nu}q^\nu + f_3(q^2)q_\mu \\ & - i g_2(q^2)\sigma_{\mu\nu}q^\nu\gamma_5 - g_3(q^2)\gamma_5 q_\mu \Big] n. \end{aligned} \tag{7}$$

In this expression f_1, f_2, f_3, g_1, g_2 and g_3 are the unknown amplitudes. f_2 and f_3 are "second class" currents. In a renormalizable field theory, the implications of the existence of second class currents in beta decay (e.g. non-commutation of the strong and weak gauge groups, possible loss of asymptotic freedom, order α parity violation) are in conflict with the standard model. [12] At present, existing procedures for the calculation of quark-quark strong interactions cannot determine the remaining four unknowns (recall that in the absence of the strong interaction $f_1 = g_1 = 1$ and $f_2 = g_3 = 0$).

It is interesting (and not at all surprising) to note the similarity between Eq. 7 and the most general expression for the phenomenological Hamiltonian given in Eq. 3. The effect of

the strong interaction is a modification of the vector and axial vector amplitudes as well as the introduction of induced tensor and pseudoscalar currents. Since nuclear beta decay is a very low energy phenomenon, only the low energy limit ($q \rightarrow 0$) of Eq. 7 is relevant.

A consequence of the overall gauge invariance of the electroweak theory is the conservation of the vector current (CVC). In the low energy limit CVC requires that $f_1 = 1$. It is important to note that $f_1 = 1$ without regard to the detailed nature of the strong perturbation of the quark wave functions. Thus the coupling to the vector current will be the *same for all nuclear beta decays.* This is the weak interaction analogue to the fact that the electromagnetic coupling constant e is the same for all interactions between charged particles. This point has important consequences for detailed tests of the above theory. Gauge invariance can also be employed to relate the magnitude of the strong perturbation of the electromagnetic structure of the nucleon (i.e. its anomalous magnetic moment) to the magnitude of the induced tensor coupling. This coupling, known as weak magnetism, becomes $f_2(0) = (\mu_p - \mu_n)/(2m_p)$. It can further be shown be shown that $g_3(0)$ is negligibly small at low energy.

With the above simplifications, it is possible to provide an expression for the complete Hamiltonian for neutron beta decay in terms of only two parameters; one describing the coupling to the vector current with magnitude $f_1 = 1$ and the other, g_1, associated with the axial vector current, depends on strong interaction dynamics. This Hamiltonian, usually expressed in terms of two coupling constants g_{V} and g_{A} is then

$$H_{\mathrm{weak}} = G_{\mathrm{F}} \cos\theta_1 \left[\bar{p}\gamma_\mu(1 - \frac{g_{\mathrm{A}}}{g_{\mathrm{V}}}\gamma_5)n\right]\left[\bar{e}\gamma^\mu(1-\gamma_5)\bar{\nu}_e\right]. \tag{8}$$

Thus one recovers the classical form of Eq. 4.

4 Phenomenology of Neutron Beta Decay

From Eq. 8 it is possible to obtain an expression for the differential decay probability dW. This decay probability is usually expressed as

$$dW \propto (g_{\mathrm{V}}^2 + 3g_{\mathrm{A}}^2)F(E_e)\left[1 + a\frac{\vec{p}_e \cdot \vec{p}_{\bar{\nu}}}{E_e E_{\bar{\nu}}} + A\frac{\vec{\sigma_n} \cdot \vec{p}_e}{E_e} + B\frac{\vec{\sigma_n} \cdot \vec{p}_{\bar{\nu}}}{E_{\bar{\nu}}}\right], \tag{9}$$

where $F(E_e)$ is the familiar beta electron energy spectrum, p_e, $p_{\bar{\nu}}$, E_e and $E_{\bar{\nu}}$ are the momenta and kinetic energies of the decay electron and anti-neutrino, and $\vec{\sigma_n}$ is the initial spin of the decaying neutron. [13] The correlation coefficients a, A and B are measurable experimental parameters. If polarization analysis is performed on the decay particles, there are additional measurable correlation coefficients. However little experimental effort has been directed toward such measurements. In addition, Eq. 9 does not incorporate the small energy dependent effects arising from the induced tensor weak magnetism. Though these effects remain slightly below current detection thresholds, appropriate experimental results are slightly adjusted to account for their presumed existence. While Eq. 9 represents phenomenology required by the standard model, there may be additional observables which arise from physics beyond the standard model. In particular there are more complicated correlations which, if non zero, represent a violation of time reversal invariance. This issue is discussed by Wilkerson and others elsewhere in these proceedings.

The correlation coefficients in Eq. 9 may be expressed in terms of the ratio of the free parameters g_{A} and g_{V} in the Hamiltonian of Eq. 8. Defining $\lambda = g_{\mathrm{A}}/g_{\mathrm{V}}$, one obtains

$$a = \frac{1 - |\lambda|^2}{1 + 3|\lambda|^2} \quad A = -2\frac{|\lambda|^2 + \mathrm{Re}\,\lambda}{1 + 3|\lambda|^2} \quad B = 2\frac{|\lambda|^2 - \mathrm{Re}\,\lambda}{1 + 3|\lambda|^2}. \tag{10}$$

It is also possible to derive, from Eq. 8, an expression for the overall decay probability or

neutron lifetime τ_n. This can be shown to be

$$\tau_n^{-1} = \frac{m_e^5 c^4}{2\pi^3 \hbar^7} f_0 (g_V^2 + 3g_A^2), \tag{11}$$

where $f_0 = 1.71465(15)$ is a phase space factor which is determined theoretically. [14] Note that a measurement of τ_n provides a determination of the quadratic sum of $(g_V^2 + 3g_A{}^2)$ while a measurement of any of the correlation coefficients provides a determination of the ratio g_A/g_V. Thus in combination, such measurements provide values for g_A and g_V.

Rather more interesting is the comparison between the above and measurements of selected *nuclear* decay rates ($(ft)_{0^+ \to 0^+}$) which can be shown to arise from a purely vector current interaction (so called superallowed Fermi transitions). The lifetime of such a decay will be given by an expression similar to Eq. 11 with $g_A = 0$. Of course the phase space factor will differ from that of the neutron. As noted earlier, it is a consequence of CVC that the amplitude of the vector current is the same for all nuclear beta decays. Thus a measurement of the lifetime (and the phase space factor) of such a decay provides a determination of g_V. Note that the combination neutron lifetime, neutron decay correlation coefficient and selected nuclear decay rates provides an overdetermined set of data from which the two parameters g_A and g_V may be determined. The interrelationship among these data is displayed in Fig. 1. The standard model requires consistency among the data as reflected in the intersection of the ellipse (τ_n), the diagonal line (g_A/g_V) and the vertical line (g_V).

Should the experimental curves fail to intersect within their errors, this could reflect an error(s) in the measurements (or at least an underestimation of experimental uncertainty), or it could reflect an error in the theoretical reasoning outlined earlier. Thus the consistency of the data provides a test the standard model.

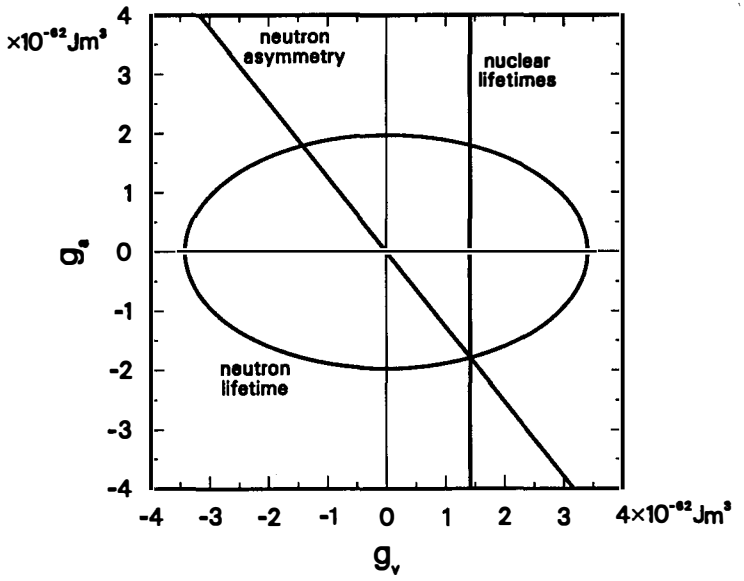


Figure 1: Weak Axial vector coupling constant g_A and vector coupling constant g_V derived from neutron decay parameters and the combined superallowed Fermi transition rate.

5 Experimental Results

This section reviews two classes of neutron beta decay measurements; *direct* determinations of τ_n and measurements of the decay correlation coefficient A from Eq. 9. Measurements of A currently provide the most accurate value of g_A/g_V. Following a brief discussion of the measurements of τ_n the values with the lowest quoted errors are compared with derived *indirect* determinations of τ_n from g_A/g_V, g_V through Eq. 11. [15,16]

Two distinct strategies have been used for the *direct* determination of the neutron lifetime. In "storage" experiments, neutrons are totally confined within a neutron "bottle". Such bottles may be physical containers in which the neutrons are contained by coherent scattering effects at the wall, or they may be magnetic in character with the neutrons confined by the interaction of their magnetic moments with magnetic field gradients. In either case, the experimental strategy involves filling the bottle with a known number of neutrons $N(0)$, storing them for

a fixed period of time, and finally emptying the bottle and counting those that remain. The number of remaining neutrons, *after accounting for all bottle loss mechanisms*, is then fitted to the exponential decay relation $N(t) = N(0)e^{-t/\tau_n}$.

In a neutron "beam" experiment, the differential form of the decay law, $\dot{N} = N/\tau_n$, where N is the mean (and approximately constant) number of neutrons present within a detector and $\dot{N}$ is the rate at which neutrons are decaying within the detector. Typically the neutron number is maintained constant by passing a steady state neutron beam through the detector.

Both methods must overcome serious metrological difficulties in order to reach interesting levels of accuracy. In bottle experiments, all loss mechanisms must be accounted for to an accuracy at least equal to the uncertainty desired for τ_n. This requires quite high precision as the bottle loss rates are typically comparable to the beta decay rate. In beam experiments the absolute neutron number in the detector must be determined as well as the absolute decay rate within the detector. This requires a high level of understanding of the absolute efficiencies of detectors. In spite of these difficulties, a number of novel experimental approaches have produced considerable progress, and several accurate determinations of τ_n have been published within the last 2–3 years.

Notwithstanding some very interesting investigations with magnetic bottles, [17] the most successful storage measurements have employed bottles with material walls. If such a bottle is maintained at high vacuum, the dominant loss mechanism will be associated with wall interactions. One normally assumes an average loss probability *per wall bounce* (for a given neutron velocity). One may therefore account for wall losses by comparing several measurements having equal storage times but a different number of wall collisions.

Mampe and colleagues carried out such a measurement using a neutron bottle having a variable size. [18] The bottle dimensions were changed in such a way that the ratio of bottle surface area to volume could be varied in a controlled fashion. With measurements of the

neutron population remaining in the bottle, *after a fixed number of wall bounces*, but differing lengths of time, the relative loss contribution due to beta decay and wall losses could be accurately estimated. In order to insure that the bottle surface was unchanged with changing bottle geometry, the wall coating (a fluorinated diffusion pump oil) was continuously re-distilled onto the bottle wall. In a different experiment, Serebrov and colleagues examined the bottle storage time as a function of the velocity spectrum of the neutrons. [19] Since the faster neutrons collide with the wall more frequently, Serebrov's method provided a diagnostic for the study of wall losses.

In the most accurate beam experiment to date, a well collimated, low energy neutron beam ($\approx 5\,\mathrm{meV}$) was passed through an electromagnetic Penning trap. [20] Protons from neutron decay occurring in the trap were confined within the trap and later released for detection. This procedure has the advantage of allowing the detection of neutron beta decay with a very high signal to noise ratio (≈ 500). A novel scheme involving a variable length trap was employed to accurately determine the decay volume. The density of neutrons in the beam was measured by passing the beam through a very accurately assayed ^{10}B deposit and counting the resulting alpha particles with a well defined geometry.

The determination of A requires a decay detector capable of identifying the direction in which the beta electron is emitted following neutron decay. Such experiments must utilize *polarized* neutrons so that the direction of the neutron spin is defined. The accuracy of the determination of A is typically limited by the knowledge of the degree of polarization in the neutron beam. The neutron polarization determination poses a particular challenge to the accurate measurement of A.

In the determination of A reported by the PERKEO collaboration (carried out at the Institut Laue Langevin), the neutron beam was polarized by reflection from a magnetic multilayer. [21] This device, known as a *supermirror* provides a very high degree of polarization

($\approx$ 97%). [22] The polarized neutron beam passed axially through a superconducting solenoid. Beta electrons from neutron decay were compelled by the high magnetic field to follow tight helical orbits along the magnetic "field lines". In a region beyond the main detector volume, auxiliary coils deflected the electrons into scintillation detectors. From a knowledge of the neutron polarization and the change in detector count rates upon reversal of the neutron spin, the value of the coefficient A was inferred. A different, more recent measurement of A identified neutron decay by coincident detection of decay protons and electrons. [23] That measurement achieved a very impressive level of accuracy in spite of a disappointingly low ($\approx$ 85%) neutron polarization.

The results of the above experiments are summarized in Table 1. Given are the direct values of τ_n as well as *indirect* values derived from the determinations of A (and therefore g_A/g_V) and the decay rates of the pure Fermi nuclear decays. It is evident from Table 1 that the direct and indirect determinations of τ_n are completely consistent among themselves. However, if the quoted errors are taken seriously, there is a 2.4σ discrepancy between the direct and indirect results.

Table 1: Shown in this table are three recent direct measurements of the neutron lifetime (two bottle experiments and one beam experiment), and three indirect determinations arrived at from measured values of A and the combined pure Fermi superallowed $(ft)_{0^+\to 0^+}$ value. Within each group the agreement is excellent; between the two groups there exists a 2.4σ difference.

Direct Results		Indirect Results	
887.6 $\pm$ 3. s	Mampe *et al*, 1989	897.7 $\pm$ 11. s	1960–1976 results combined
893.6 $\pm$ 5.3	Byrne *et al*, 1989	892.4 $\pm$ 5.8	Bopp *et al*, 1986
888.4 $\pm$ 2.9	Serebrov *et al*, 1990	901.7 $\pm$ 4.4	Erozolimskii *et al*, 1991
888.8 $\pm$ 1.9	Combined result	898.2 $\pm$ 3.4	Combined result

6 Discussion and Summary

A *real* discrepancy between the direct and indirect values for τ_n would imply an error *somewhere* in the line of reasoning described above. While many aspects of the argument may be reconsidered, it is particularly interesting to reconsider the initial assumption that the weak interaction is *purely* left-handed. It is abundantly clear that, at low energy at least, the weak interaction is dominantly left-handed. However that there may be a small right-handed contribution. Such a contribution could be a manifestation of an overall left-right symmetrical theory whose symmetry is broken at low energy. [24–26,16] One such theory posits the existence of two intermediate vector bosons, W_L and W_R, which respectively couple to left and right-handed currents. If these W particles have different masses, the strengths of the left and right-handed couplings will be different. In general, the mass eigenstates W_1 and W_2 for these particles will be a superposition of the weak eigenstates given by

$$\begin{aligned} W_1 &= W_{\mathrm{L}} \cos\zeta - W_{\mathrm{R}} \sin\zeta \\ W_2 &= W_{\mathrm{L}} \sin\zeta + W_{\mathrm{R}} \cos\zeta \end{aligned} \quad , \tag{12}$$

where ζ is a mixing angle which is a characteristic parameter of the theory. One completely parameterizes the theory by this mixing angle and the mass ratio $\delta = (M_1/M_2)^2$ which represents the relative strength of the left and right-handed interactions.

Within the framework of such a theory, τ_n, A and $(ft)_{0^+\to 0^+}$ are now determined by four parameters g_{A}, g_{V}, δ and ζ. With three experimental quantities and four unknown parameters the problem is underdetermined. However, the constraints on the theoretical parameters can be displayed on an exclusion plot of ζ vs δ. Such plots have been widely displayed in the literature. [27–29,16] Such an exclusion plot, based on the five measurements in Table 1 is displayed in Fig. 2.

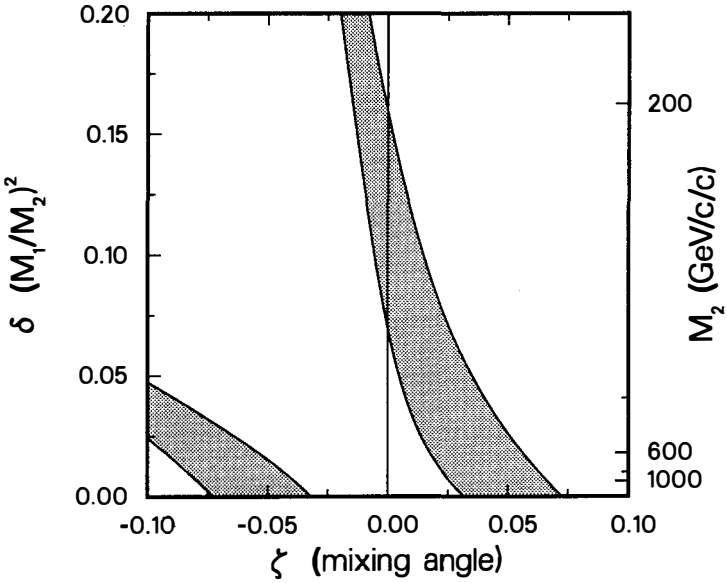


Figure 2: Limits on right-handed currents from neutron decay parameters and superallowed Fermi transition rates. The grey shaded region is allowed at 90% confidence. The origin corresponds to the standard $V - A$ theory.

Note that the origin ($\delta = \zeta = 0$), which represents the standard $V - A$ theory is excluded in this plot. In essence, this exclusion is representative of the disagreement between the direct and indirect values for τ_n seen in Table 1.

In addition to the neutron decay data, there are several additional measurements which bear on the question of the existence of right-handed currents. These include decay parameters of ^{19}Ne, relative beta polarization measurements in nuclear decays, and measurements of muon decay. The overall consistency of these measurements with the $V - A$ model has been investigated by several authors [16] and an interesting pattern emerges.

According to Carnoy *et al*, the complete set of semi-leptonic data suggests, at the 95% confidence level, the existence of an intermediate vector boson with a mass $207 < M_2 < 369\,\mathrm{GeV}/c^2$ which couples to right-handed currents. [16] This is inconsistent with the purely leptonic muon decay data, which is in agreement with the $V - A$ predictions. It is important to

realize that the discrepancy in the semi-leptonic data is overwhelmingly due to one measurement of the neutron A coefficient. [23] Reanalysis without the inclusion of this measurement yields reasonable agreement with the purely left-handed model. As there are only two measurements of A which contribute significantly to the above analysis it is clear that more and improved experimental data is required.

References

[1] D. N. Schramm and L. Kawano, Nucl. Instrum. Methods **A284**, 84 (1989).

[2] T. P. Walker, G. Steigman, D. N. Schramm, K. A. Olive, and H. S. Kang, Astrophys. J. **376**, 51 (1991).

[3] R. J. Gould and N. Guessoum, Astrophys. J. Lett. **359**, L67 (1990).

[4] E. Fermi, Z. Phys. **88**, 161 (1934).

[5] G. Gammow and E. Teller, Phys. Rev. **49**, 895 (1936).

[6] M. Fierz, Z. Phys **104**, 553 (1937).

[7] C. S. Wu, E. Ambler, E. Hayward, D. D. Hoppes, and R. Hudson, Phys. Rev. **105**, 1413 (1957).

[8] M. Goldhaber, L. Grodzins, and A. W. Sunyar, Phys. Rev. **109**, 1015 (1958).

[9] E. D. Commins and P. H. Bucksbaum, *Weak Interactions of Leptons and Quarks*, University of Cambridge, 1983.

[10] B. R. Holstein, *Weak Interactions in Nuclei*, Princeton Series in Physics, Princeton University Press, 1989.

[11] D. Dubbers, Prog. Part. Nucl. Phys. **26**, 173 (1991).

[12] P. Langacker, Phys. Rev. D **15**, 2386 (1977).

[13] J. D. Jackson, S. B. Treiman, and H. W. Wald, Phys. Rev. **106**, 517 (1957).

[14] D. H. Wilkinson, Nucl. Phys. **A377**, 474 (1982).

[15] J. C. Hardy, I. S. Towner, V. T. Koslowsky, E. Hagberg, and H. Schmeing, Nucl. Phys. **A509**, 429 (1990).

[16] A. S. Carnoy, J. Deutsch, R. Prieels, N. Severijns, and P. A. Quin, Is there an indication for the existence of right-handed weak currents in nuclear beta-decay?, 1991, Institut de Physique Nucleaire, Universite Catholique de Louvain Preprint.

[17] F. Anton, W. Paul, W. Mampe, L. Paul, and S. Paul, Nucl. Instrum. Methods **A284**, 101 (1989).

[18] W. Mampe, P. Ageron, J. C. Bates, J. M. Pendlebury, and A. Steyerl, Phys. Rev. Lett. **63**, 593 (1989).

[19] V. P. Alfimenkov et al., Pis'ma Zh. Eksp. Teor. Fiz. **52**, 984 (1990), [JETP Letters **52** 373 (1990)].

[20] J. Byrne et al., Phys. Rev. Lett. **65**, 289 (1990).

[21] P. Bopp et al., Phys. Rev. Lett. **56**, 919 (1986).

[22] O. Schaerpf and N. Stuesser, Nucl. Instrum. Methods **A284**, 208 (1989).

[23] B. G. Erozolimskii, I. A. Kuznetsov, I. V. Stepanenko, I. A. Kuida, and Y. A. Mostovoi, Phys. Lett. B **263**, 33 (1991).

[24] M. A. B. Beg, R. V. Bundy, R. Mohapatra, and A. Sirlin, Phys. Rev. Lett. **38**, 1252 (1977).

[25] B. R. Holstein and S. B. Treiman, Phys. Rev. D **16**, 2369 (1977).

[26] A. S. Carnoy, J. Deutsch, and B. R. Holstein, Phys. Rev. D **38**, 1636 (1988).

[27] A. S. Carnoy, J. Deutsch, T. A. Girard, and R. Prieels, Phys. Rev. Lett. **65**, 3249 (1990).

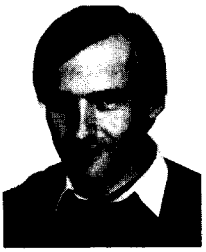
[28] D. Dubbers, W. Mampe, and J. Döhner, Europhys. Lett. **11**, 195 (1990).

[29] Y. V. Gaponov, N. B. Shul'gina, and P. E. Spivak, Phys. Lett. B **253**, 283 (1991).

TEST OF TIME REVERSAL SYMMETRY IN THE DECAY OF POLARIZED ^{8}LI– NEW RESULTS.

M. Allet*, K. Bodek*, A. Converse*, W. Haeberli†, W. Hajdas*, J. Lang*,
M. A. Miller†, R. Müller*,S. Navert,O. Naviliat Cuncic*, G. Spiekerman*,
J. Sromicki* and J. Zejma*

* Institut für Mittelenergiephysik der ETH Zürich, CH-8093 Zürich
† University of Wisconsin, Madison, WI 53706, USA



presented by J. Sromicki

ABSTRACT

The transverse polarization of the electrons emitted in the decay of polarized ^{8}Li has been measured. The magnitude R of the time reversal violating correlation $\vec{J} \cdot (\vec{v} \times \vec{\sigma})$ between the directions of the nuclear spin $\vec{J}$, the velocity $\vec{p}$, and the spin $\vec{\sigma}$ of the electrons has been determined. The present result $R = 0.004 \pm 0.007$ is consistent with time reversal invariance.

1. GENERAL MOTIVATION

Fundamental symmetries and their violations have always played a very important role in physics. Recent interest in this field is related to the development of the Standard Model (SM) of elementary particles and their interactions. Maximum violation of the symmetry with respect to space reflections (P) has its natural explanation in this theory: weak interactions are mediated by left-handed spin 1 gauge bosons; their right-handed counterparts do not exist at all. The situation is entirely different for reflexion of the time axis (T). Although the SM provides room for violation of T symmetry via the Kobayashi- Maskawa (KM) quark mixing scheme, the underlying mechanism for such a violation is not understood. Also, contrary to parity, the experimental information is very scarce. Small effects of CP non- invariance, which is closely related to T violation have been observed only in neutral K meson decays. A further positive experimental evidence could answer the question: does T violation originate in the SM, or is it a signature of a new class of phenomena ?

Some alternatives to the KM picture of T violation have been proposed. E.g. multiplets of Higgs particles (ref.1,2) can induce T violating scalar interactions in addition to the vector and axial vector present in the SM. Massive right-handed bosons with their mixing matrices different from those of their left-handed partners can give rise to effects even in the light quark sector (ref. 3). There exists also the possibility of exotic interactions mediated by e.g. spin-zero leptoquarks or tensor bosons (ref.4,5).

Beta decay provides a good testing ground for these hypotheses. The effects are not obscured by T conserving strong interactions, and a well developed theory can be used in analyses of different experiments and their implications. The results can be directly interpreted in terms of exotic interactions and/or particles, and limits for their strength and masses can be deduced.

The experiment described below is sensitive to the charged tensor interaction. Such an interaction might be present in nonstandard models of the weak forces. For details of the theoretical foundations we refer to our previous report (ref. 6) and references therein. Also, details of the experiment have been extensively reported elsewhere (ref. 6-9). Therefore, we will stress particularly recent improvements and present some aspects of the experiment which were not discussed thoroughly. Finally, we present our new result.

2. THE ^{8}Li EXPERIMENT

We search for T violation by measuring the transverse polarization component in a plane perpendicular to the nuclear spin for electrons emitted in the decay of polarized ^{8}Li nuclei. In the experiment the amplitude R of the triple correlation between the nuclear polarization $\vec{J}$, the velocity of the electron $\vec{v}$, and the polarization of the electron $\vec{\sigma}$ is determined. For our purposes the decay rate for an allowed transition from oriented nuclei can be written as (ref. 10):

$$W \sim 1 + A(\vec{J}\cdot\vec{v}) + R\,\vec{J}\cdot(\vec{v}\times\vec{\sigma}) \qquad (1)$$

where A is the usual parity violating asymmetry parameter. In absence of final state interaction effects, the T- and P- odd R parameter is expected to be zero if time reversal invariance holds.

The largest time reversal violating effects should be seen for mutually orthogonal vectors $\vec{J}$, $\vec{v}$ and $\vec{\sigma}$. In fact, if electrons are emitted in the direction of the nuclear spin, their transverse polarization must vanish by virtue of rotational invariance. A general idea of the experiment is presented in ref. 6. For the present measurements we have constructed a polarimeter of a novel design. It consists of two conical detector rings located above and

below the median decay plane, with their symmetry axis coinciding with the nuclear spin axis. Each ring is subdivided into four segments of triple plastic scintillator telescopes consisting of two transmission detectors and a 4 cm thick stopping detector which measures the energy of scattered electrons. The two transmission detectors are indispensable for the reduction of background to an acceptable level. We use two 90^0 segments placed in forward direction with respect to the incoming beam and two 60^0 segments placed backward for each ring. The polarization of the electrons emitted from the ^{8}Li source (placed in the center of the polarimeter) is analyzed by large-angle Mott scattering from a 35 mg/cm^2 lead foil located in the median plane. The whole system, i.e. electron polarimeter and monitors of the target polarization, consists of 28 detectors utilizing 52 photomultiplier tubes. Fig. 1. shows the location of the scattering foil, the detectors and their shields, as well as some details of the target arrangement.

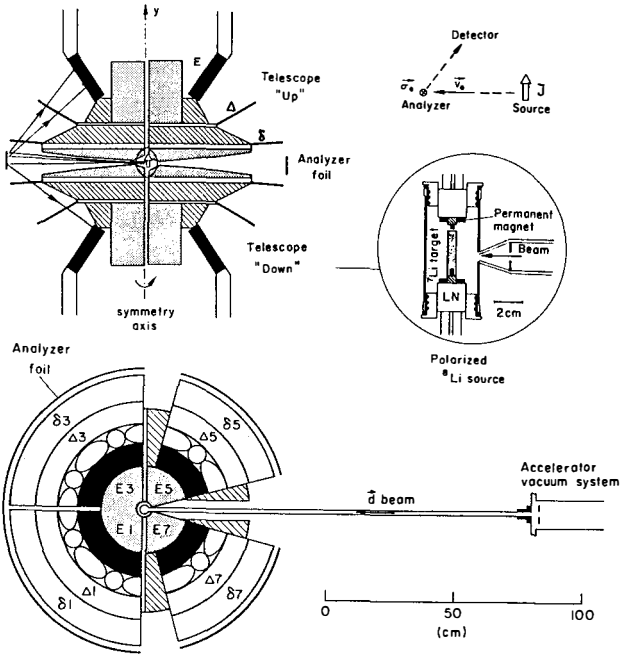


Fig. 1. The main parts of the experimental setup. The electrons are scattered from the analyzer foil and detected in eight triple scintillator telescopes (δ, Δ, E detectors) located in a symmetric up/down arrangement. Upper part: the vertical cross section through the apparatus; lower part: top view. The azimuthal angular coverage amounts to 85%. The desired 360^0 azimuthal symmetry of the apparatus is partly broken at the ϕ angles 90^0 and 180^0 to provide room for support structure and beam pipe assembly. On the top right, the vectors which enter into the definition of the R parameter are shown schematically. Inset: the details of the target chamber with the permanent magnets providing a spin-holding field and the liquid nitrogen cooling (LN) for the ^{7}Li rod. The last section of the beam pipe and housing of the target are on purpose made of low-Z thin material, in order to minimize energy losses and to reduce backscattering of the electrons from other objects than the lead analyzer foil. Recent alterations of the apparatus can be seen by comparing this figure with the corresponding one in the Letter (ref. 9).

^{8}Li nuclei were vector polarized by the polarization transfer reaction $^7\mathrm{Li}(\vec{d},\mathrm{p})^8\vec{L}i$ initiated with 10 MeV polarized deuterons. Technical details concerning the ^{8}Li source, measurement sequence, spin relaxation etc. can be found in refs. (6,8).

3. IMPORTANT SYSTEMATIC EFFECTS

Since the effective analyzing power of the polarimeter amounts to -0.10 and the achieved average polarization of the ^{8}Li source is 0.125 ± 0.005, a statistical accuracy of $1\cdot10^{-4}$ in the asymmetry is needed to achieve an error in the R coefficient of 0.01. In fact, this accuracy must even be slightly better in order to compensate for some losses due to systematic effects, which must be controlled at the level of at least 10^{-5} in the measured asymmetry.

We have considered a number of systematic effects in this experiment. They include: nonuniform illumination of the scattering foil, gain shifts of the photomultipliers, depolarization and multiple scattering in the analyzer foil and the source, scattering in the air, background radiation, accidental coincidences ... Most of these effects and methods used for corrections were discussed previously (ref. 6-9).

By far the most interesting effect is associated with the nonuniform illumination of the scattering foil due to the regular parity nonconserving asymmetry in the beta decay. The intensity distribution N for electrons impinging at the height h on the scattering foil above or below the median plane is given by:

$$N_{\mathrm{d}}^{\mathrm{u}}\left(\alpha=\frac{\mathrm{h}}{\mathrm{r}}\right) \sim 1 + \mathrm{pA}\cos\theta \approx 1 + \mathrm{pA}\alpha \tag{2}$$

where θ is the polar angle with respect to the nuclear spin axis, α is the corresponding angle measured from the median plane ($\alpha=\pi/2 - \theta$, see fig.1), r is the radius of the foil, p is the polarization of the target and A - the decay asymmetry parameter. The upper/lower sign refers to the up (u)/down (d) telescope. Electrons impinging on the upper side of the foil are evidently closer to the upper ring, therefore, even in case of perfect alignment of the detectors, the two corresponding telescopes have different acceptance if $h \neq 0$. Expanding the product of the effective cross section and solid angle $\sigma\Omega$ as a function of the small parameter $\alpha = h/r$, we obtain:

$$(\sigma\Omega)_{\mathrm{d}}^{\mathrm{u}} = (\sigma_0\Omega_0)\left[1 \pm \alpha\frac{1}{(\sigma_0\Omega_0)}\frac{\partial}{\partial\alpha}(\sigma\Omega)|_0\right] \tag{3}$$

where the derivative is taken at $\alpha=0$. The rate measured in each detector is proportional to the integral of the product $N\cdot(\sigma\Omega)$ over the height of the scattering foil. Forming the asymmetry we obtain:

$$\mathrm{ASY} = \frac{[\int N_{\mathrm{u}}(\sigma\Omega)_{\mathrm{u}} - \int N_{\mathrm{d}}(\sigma\Omega)_{\mathrm{d}}]\,\mathrm{d}\alpha}{[\int N_{\mathrm{u}}(\sigma\Omega)_{\mathrm{u}} + \int N_{\mathrm{d}}(\sigma\Omega)_{\mathrm{d}}]\,\mathrm{d}\alpha} = \frac{1}{3}\alpha_0^2(\mathrm{pA})\rho \tag{4}$$

where α_0 is the limiting angle of the foil. The product pA was continuously measured with two polarization monitors mounted at $\theta=40^0$. The parameter ρ:

$$\rho = \frac{1}{\sigma_0\Omega_0}\frac{\partial}{\partial\alpha}(\sigma\Omega)|_0 \tag{5}$$

was measured in an auxiliary series of measurements using a narrow (1 cm) scattering foil. The foil was displaced from the median plane, and the rate was measured as a function of the angle α . This rate, normalized to the rate observed with the foil in the median plane, provides directly the slope parameter $\rho = 3.1 \pm 0.4$.

Since the intensity N changes in step with reversal of the target polarization, the false asymmetry caused by this effect also changes its sign, and therefore simulates a nonzero time reversal coefficient R. This is in contrast to all geometrical imperfections in the positioning of the detectors or in their shapes, which stay the same when the target polarization is reversed, and therefore cancel when spin averaging of the effect is performed ("superratio" method).

4. DATA ANALYSIS AND NEW RESULTS.

We have used two different procedures to evaluate the data. On the one side all the data were cumulatively added and averaged polarization and analyzing powers were used. On the other side, a file by file analysis was performed with 1 MeV energy steps. Below, we present the results with average corrections for these two procedures.

Raw data	-0.012 ± 0.007	-0.012 ± 0.007
Decay associated background, and room background	0.006 ± 0.005	0.007 ± 0.005
Beta decay asymmetry	0.010 ± 0.002	0.010 ± 0.002
Gain shifts of PM's	-0.001 ± 0.002	-0.001 ± 0.002
Accidentals	0.001 ± 0.002	0.001 ± 0.002
Result	0.004 ± 0.009	0.005 ± 0.009

Inspection of the table shows that the effect due to the beta decay asymmetry, discussed above, gives the largest correction to the raw data at present. The only way to reduce this effect is to diminish the height of the scattering foil at the expense of the rate of good events and the signal/background ratio which amounts at present to 15. We note, however that although the value of this correction for the present data is close to the statistical error of the measurement, its uncertainty is well below the uncertainties of the raw data and background subtraction. Therefore our experiment is at present limited by counting statistics.

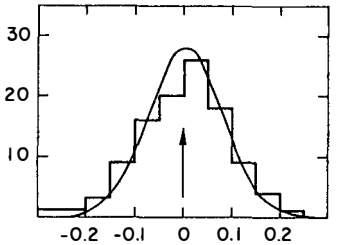


Fig. 2. Distribution of the values of the R-coefficient measured in 2 hour runs by four detector pairs after applying the corrections discussed above.

5. CONCLUSIONS

Taking the mean value of our last measurement and previously published data (ref. 9) we obtain as a present result R= 0.004 ± 0.007. This value is still the subject to a theoretical correction due to final state interactions (FSI). A rigorous calculations taking into account the finite size of the nuclei and other subtle effects were published by P. Vogel. (ref. 11). It was found however, that all these effects contribute at the level of 10 % of the value of the first order correction known since the classic publication by J.D. Jackson (ref.10). Such a correction amounts to 0.0007 in our experiment, and is well below the present accuracy. (We note, that FSI effects are exeptionally small for ^{8}Li decay due to the small nuclear charge and high energy release).

We conclude, that our result is consistent with time reversal invariance. With this result we can set new limits on the imaginary parts of the tensor coupling constants: $Im(C_T+C'_T)/C_A = 0.012 \pm 0.021$.

Recently, the PSI Scientific Committee has approved three weeks of additional beam time for this experiment. This will allow us to improve the error by another factor of three and for the first time break the barrier of 0.01 in a direct measurement of the time reversal violating amplitudes in beta decay.

REFERENCES

1. S. Weinberg, Phys. Rev. Lett. 37(1976)657.

2. T.D. Lee, Phys. Rev. D8(1973)1226, and Phys. Rept. 96.

3. R. N. Mohapatra and J.C. Pati, Phys. Rev. D11(1975)566.

4. P. Herczeg, Fundamental Symmetries in Nuclei and Particles, edited by H. Henrikson and P. Vogel, World Scientific, Singapore (1989) p. 47.

5. W. Buchmüller and D. Wyler, Phys. Lett. 177(1986)377, and D. Wyler, priv. comm. (1991).

6. M. Allet et al., XIth Moriond Workshop on Tests of Fundamental Symmetries, Les Arc (1991), edited by J. Tran Thanh Van, Editions Frontiers M68(1991)301.

7. M. Allet et al., Proceedings 9th International Symposium on High Energy Spin Physics, Bonn, Sept. 1990, edited by K.H. Althoff and W. Meyer, Springer Verlag vol.1(1991)585.

8. M. Allet et al., Intersections Between Particle and Nuclear Physics, Tucson AZ, May 1991), AIP Conference Proceedings 243(1992)977.

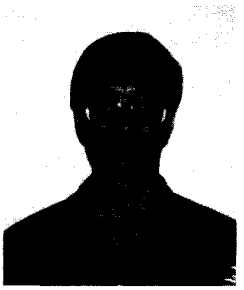
9. M. Allet et al., Phys. Rev. Lett. 68(1992)572.

10. J.D. Jackson, S.B. Treiman and H.W. Wyld, Nucl. Phys. 4(1957)206.

11. P. Vogel and B. Werner, Nucl. Phys. A404(1983)345.

MEASUREMENT OF THE DECAY RATE AND 2γ BRANCHING RATIO OF ORTHOPOSITRONIUM

J. S. Nico*, D. W. Gidley, and M. Skalsey
Department of Physics, University of Michigan, Ann Arbor, MI 48109 USA
P. W. Zitzewitz
Department of Natural Sciences, University of Michigan - Dearborn, Dearborn, MI 48128, USA



ABSTRACT

The results of a measurement of the orthopositronium (o-Ps) decay rate, λ_T, using the vacuum technique are presented. This 230-ppm measurement is consistent with a recent 200-ppm measurement of λ_T in gases and disagrees with theory at the 6.2 sigma level. Experimental searches for alternative o-Ps decay modes that might resolve the discrepancy are reported. A completed search for two-photon decay modes, both QED-allowed and forbidden, shows that they cannot be responsible for the discrepancy at the 233-ppm level.

INTRODUCTION

Quantum electrodynamics (QED) has been rigorously tested in bound-state systems of atomic hydrogen and heavier atoms as well as in the exotic atoms of muonium and positronium. Positronium (Ps), the positron-electron bound state, offers the advantage of being a purely leptonic system, with the electromagnetic interaction being the only interaction present at current levels of experimental and calculational precision. The decay rate of the spin-triplet ground state, or orthopositronium, has been of particular interest due to the existing disagreement between the theoretical and experimental values. The result of the most recent experiment in which Ps is formed in gases[1] is a factor of four more precise (at 200 ppm) than previous measurements and is 9.4 experimental standard deviations higher than the theoretical value. However, a recent measurement in which Ps is formed in an evacuated cavity[2] found at the 1000-ppm level a value one sigma *below* theory and three sigma below the gas result. Neither it nor an earlier 1978 vacuum measurement[3] is precise enough to check rigorously the 9.4 sigma discrepancy. Thus, the new vacuum measurement at a precision comparable to the gas measurement is very attractive because the vacuum measurements are systematically so different from the gas measurement.

The o-Ps decay rate, λ_T, is expressed as a sum of its decay rates into three photons (λ_3), five photons (λ_5), etc.:

$$\lambda_T = \lambda_3 + \lambda_5 + \lambda_7 + ... \tag{1}$$

The effect of the five photon decay has been calculated[4] and found to be of order $10^{-6}\lambda_3$. Thus, to sufficient precision, λ_T is given by λ_3, where

$$\lambda_3 = \alpha^6 \frac{mc^2}{\hbar} \frac{2(\pi^2 - 9)}{9\pi} \left[1 + A(\frac{\alpha}{\pi}) + \frac{1}{3}\alpha^2 \ln \alpha + B(\frac{\alpha}{\pi})^2 + ...\right]. \tag{2}$$

The coefficient A has been calculated independently yielding values of $A = -10.266(11)$[5] and $A = -10.282(3)$[6]. Using the more precise value of A from Reference 6, one obtains $\lambda_3 = 7.03830(7)\ \mu s^{-1}$. We note that B is still uncalculated, and if it is of order 10, the second order term could contribute 50 ppm, or 0.00035 μs^{-1}.

EXPERIMENTAL METHOD AND RESULTS

Our new measurement[7] of λ_T, using a systematically improved vacuum technique, employs a slow positron beam to form positronium in an evacuated (10^{-9} Torr), MgO-lined cavity, similar to the technique used in Reference 3. Several improvements in beam technology have been utilized; in particular, remoderation[8,9], beam gating, and time-tagging[10]. The apparatus (see Reference 7) is divided into two sections: a primary beam that moderates and focuses the positrons from a 50 mCi ^{22}Na source onto a Ni foil remoderator, and a secondary beam that delivers the remoderated positrons into the positronium formation cavity. Secondary electrons produced when the primary beam strikes the Ni foil are detected by a channel electron multiplier array (CEMA) and serve as the start signal for the timing electronics. This time-tagging of the beam eliminates systematics associated with a start detector inside the Ps formation cavity. Since the tagging efficiency is only 15%, the noise from untagged positrons must be eliminated by a normally-closed gate that is triggered open by a start signal. The gated and time-tagged positrons are focused into an aluminum cavity and strike the fumed MgO layer at 700 eV. The MgO, in addition to being an efficient o-Ps formation surface ($\sim$ 25%), minimizes collisional quenching at the cavity walls. The annihilation gammas are detected by two semi-annular fast-plastic scintillators that surround the cavity. Each scintillator is coupled by a light pipe to two PMTs, and the output signals of the four PMTs are mixed to comprise the stop signal.

The lifetime spectrum consists of a histogram of CEMA-start to gamma-stop time interval measurements in the range from 0 to 2 μs (for further details see Reference 7). Beginning 100 ns beyond the prompt peak, the spectrum is fitted by a maximum-likelihood program to a function of the form

$$N(t) = (Ae^{-\lambda t} + B)e^{-Rt} \tag{3}$$

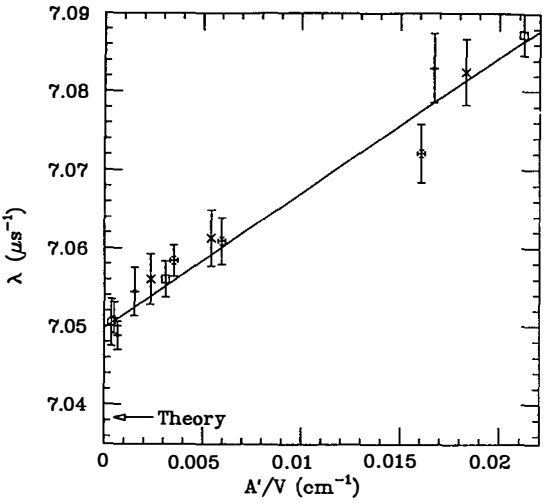


Figure 1. The fitted decay rate as a function of $\frac{A'}{V}$ (where $A' = k_1 k_2 A$) for all cavity geometries. Plotting symbols distinguish the $\frac{S}{V}$ values as shown in Figure 2. The solid line is the intersection of the plane determined by the fit to Equation 4 with the $\frac{S}{V} = 0$ plane. The χ^2 of the fit is only 2 for 10 degrees of freedom, indicating a conservative error bar assignment on the asymptotic value of λ in Equation 3.

where A, B, and λ are the fitted parameters and R is the measured background rate during the gate-closed (beam off) interval. Successively stepping out the start time, t_0, of the fitting program in 50 ns intervals and fitting the remaining channels to Equation 3 reveals a 0.5-1.0% drop in λ as it asymptotically approaches a constant value at $t_0 = 400$-450 ns. Thus, at least one additional lifetime component is present in the histogram. The data were therefore refitted to a modification of Equation 3 that included a second exponential component. We typically obtained fits with acceptable χ^2 beginning at $t_0 = 170$ ns and with fitted values of λ statistically consistent with the asymptotic value from the single exponential fits. The origin of this additional component with a lifetime of roughly 40 ns is not understood. It does not depend directly on any cavity shape parameter, and its intensity is too large to be attributed to the production of excited states of Ps. The most reasonable explanation that appears to satisfy all the observations is that several percent of the o-Ps is trapped in small ($\sim$ 100 Å) pores in the MgO powder. In this, as well as virtually all alternative explanations, the asymptotic value of λ is interpreted as the correct value of λ to be associated with that particular cavity. A systematic test of this interpretation will be discussed in the section on 2 photon decay modes.

The asymptotic value of the fitted decay rate is assumed to be shifted from λ_T linearly by terms that account for the rate of disappearance of o-Ps through the cavity entrance aperture and the rate of annihilation due to collisions with the cavity walls. Thus,

$$\lambda = \lambda_T + c\frac{\bar{v}}{4}k_1 k_2 \frac{A}{V} + \frac{\bar{v}}{4}P_a \frac{S}{V} \tag{4}$$

where V and S are the cavity volume and surface area, respectively, A is the area of the cavity entrance aperture through which the positrons are focused, $\bar{v}$ is the average o-Ps velocity, c is the probability that γ-rays from escaped o-Ps will *not* be detected, and P_a is the probability of annihilation during a wall encounter. The disappearance term contains the parameter k_1 that accounts for reduced o-Ps disappearance through an aperture of nonzero thickness and the parameter k_2 that accounts for the non-uniform particle density that

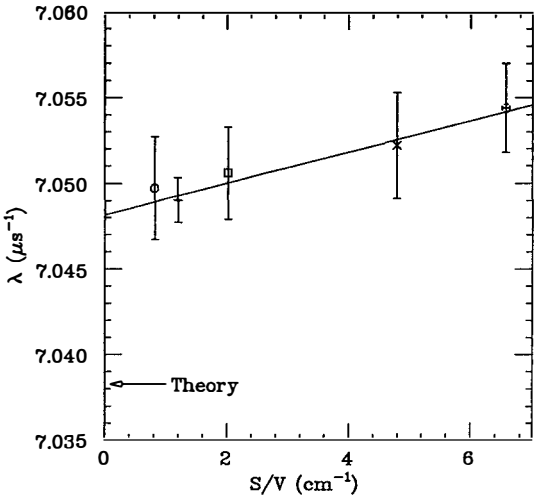


Figure 2. Results of the individual extrapolations to $\frac{A'}{V} = 0$ are shown for each cavity as a function of $\frac{S}{V}$. The solid line represents a linear extrapolation to $\frac{S}{V} = 0$ as our determination of λ_T. The χ^2 of the fit is 0.2 for 3 degrees of freedom.

occurs when the aperture diameter is nearly as large as the o-Ps mean free path in the cavity. The vacuum decay rate can be obtained from Equation 4 as a 2-variable extrapolation to the intercept of a plane in $\frac{A}{V}$ and $\frac{S}{V}$ (see Figure 1). The intercept of equation 4 is $7.0497 \pm 0.0013\ \mu s^{-1}$. Note that the measured λ are within one standard deviation of this intercept, permitting an extrapolation that is almost 2 orders of magnitude smaller than that encountered in gas experiments.

Equation 4 contains the assumption that the slope of the $\frac{A}{V}$ disappearance term will be the same for all cavities. In fact, we see a decrease in this slope at the 1-2 sigma level as $\frac{S}{V}$ increases. We removed this slope restriction by refitting the data for each cavity to Equation 4 without the $\frac{S}{V}$ term. The $\frac{A}{V} = 0$ intercepts for each cavity are then plotted versus $\frac{S}{V}$ as shown in Figure 2. Extrapolation to $\frac{S}{V} = 0$ yields $\lambda_T = 7.0482 \pm 0.0015\ \mu s^{-1}$, or about one sigma lower than the 2-variable extrapolation. We select the result of this double-extrapolation as our determination of λ_T since it is the most free of restrictions on the fitted parameters.

Systematic effects that might alter the shape of the annihilation spectrum and hence shift our determination of λ_T have been considered. A major systematic uncertainty that limited the 1978 experiment is related to the possibility that some of the Ps does not diffuse out of the MgO powder (where the decay rate in the open, intergranular space is roughly 2.5% higher than λ_T[11]). A component with a decay rate that close to λ_T would go undetected as the start time (t_0) of the fitting program is stepped out, but it would shift λ upwards. To search for this effect, a lead gamma ray collimator was constructed to view only the gamma rays originating at or very near the formation surface. We have set a limit on a quantity f_{MgO}, the fraction of o-Ps beyond 350 ns that decays in the formation powder, of $f_{MgO} \leq 0.31\%$. Its effect on the fitted value of λ is then limited to $(2.5\%) f_{MgO}$, or 80 ppm.

The formation of excited states of positronium (Ps*) can also add lifetime components to Equation 3. The 2^3S_1 state is the only one with a formation rate and vacuum lifetime (1.1 μs) that could possibly affect the fitted value of λ_T. The probability that any 2^3S_1 Ps have lifetimes that long is considered to be unlikely based on a previous experiment[1] that determined the cross section for collisional deexcitation of 2^3S_1 states in gases

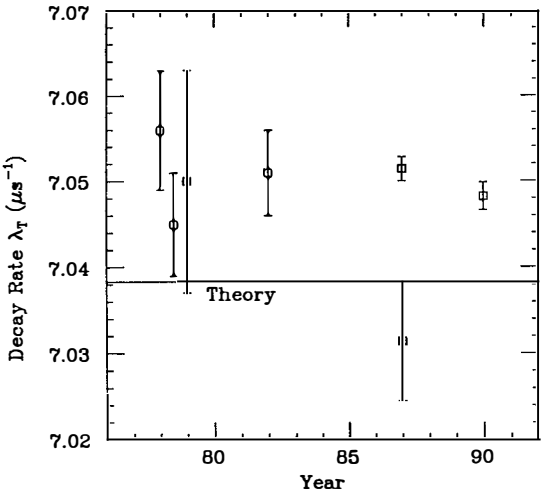


Figure 3. Recent o-Ps decay rate measurements.

to be of order 10^{-14} cm^2. It is improbable that the 2^3S_1 state will survive beyond the first collision with an MgO grain, which occurs within 50 ns in a typical cavity. In view of these considerations, we conclude that the effect of excited state Ps for $t_0 > 400$ ns is negligible.

Using the extrapolation in Figure 2, our result for the 1^3S_1 decay rate is $\lambda_T = 7.0482 \pm 0.0016$ μs^{-1} (230 ppm). The dominant contributions to the uncertainty are the 210-ppm statistical error (that includes the uncertainties in the $\frac{A}{V}$ and $\frac{S}{V}$ extrapolations), the 80-ppm MgO powder systematic, and the time calibration uncertainty of 40 ppm. Figure 3 summarizes the status of the recent measurements of λ_T and includes our new vacuum result. The value does not substantiate the low decay rate reported in Reference 2. It is 6.2 sigma above theory and is in reasonable agreement with the gas measurement of Reference 1.

TWO PHOTON SEARCH

One possible resolution of this discrepancy is that o-Ps may have a small branching ratio for a decay mode not considered in the calculation. A number of exotic decay modes of o-Ps have been searched for: an axion and a single photon[12],[13]; photonless o-Ps decay into non-interacting particles[14] (such as millicharged particles[15]); and o-Ps decay into two gamma rays[16],[17],[18]. Mills and Zuckerman[16] have searched for the 2γ mode as a test of spatial isotropy. The decay of o-Ps into two photons in vacuum would violate angular momentum conservation and charge conjugation invariance. They suggest that o-Ps may interact with some (unknown) field that defines a preferred direction in space. They search for a diurnal variation in the rate of o-Ps decay in SiO_2 as the laboratory frame rotates with respect to this preferred direction. They were thus able to rule out such a large 2γ decay mode unless the preferred direction is oriented within 35° of the earth's axis (*i.e.*, about 80% of all possible directions are excluded).

We have performed a *direct* search for this 2γ decay mode of o-Ps[18]. Our search is a particularly definitive test of the possible contribution of 2γ events to the discrepancy because we use exactly the same experimental conditions under which we measured the o-Ps decay rate. This insures that any hypothesized, non-QED decay modes are present at that level (if they are to account for the discrepancy) regardless of any model-dependent interpretation. Using time-delayed gamma-ray energy spectroscopy to measure the energy of one of the decay photons, we can search directly with a high signal-to-noise ratio for excess 511 keV gamma rays as one indication of o-Ps $\rightarrow 2\gamma$.

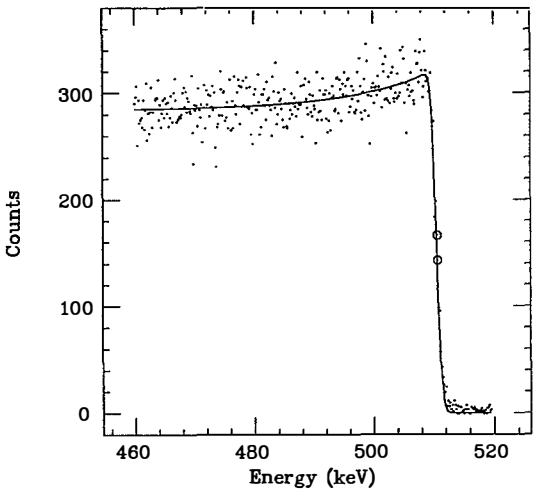


Figure 4. The energy spectrum near 511 keV, with only the background subtracted, for one of the two data runs. Each dot represents counts in a 0.167 keV wide channel. The solid line is the theoretical 3γ continuum spectrum of of Reference 6 after a linear correction and convolution with a 1.7 keV resolution function. It is fitted to the data between 460 and 508 keV as described in the text. The channels on either side of 511.0 keV are circled. With the exception of the summing events above 511 keV, the data are visually consistent with a pure o-Ps continuum with no 2γ peak at 511 keV.

The apparatus is the same as that described for the decay rate experiment with the exception that a Ge detector replaced the fast plastic scintillators. The gamma-ray energy spectrum from 75 keV to 1350 keV is a histogram of events that also fall within a delayed timing window. It mainly consists of the o-Ps continuum spectrum that extends up to 511 keV and some background events throughout the entire range. We are searching for an anomalous 511 keV peak at the edge of the o-Ps continuum and for any events between 511 and 1022 keV that cannot be accounted for by background. We summed 36 histogram channels into 6 keV wide energy windows with the last window centered on 511 keV. The previous 8 windows from 460 keV to 508 keV are fitted to an o-Ps continuum energy spectrum which we then extrapolate to the last window to determine the expected number of continuum o-Ps events in the 511 keV window (see Figure 4). This fitting determines the predicted number of o-Ps continuum events in the 511 keV window that is then subtracted from the measured number of events to obtain the number of detected 2γ decays of o-Ps. We then normalize this number to the total number of o-Ps events observed to obtain the 2γ to 3γ branching ratio, $B_{2\gamma}$.

Two separate runs were acquired 3 months apart during which time the detector and Pb shielding were disassembled and the o-Ps formation cavity was refumed with MgO. Combining these two runs and reducing the error only insofar as the statistics are independent, we obtain

$$B_{2\gamma} = (126 \pm 233) \text{ ppm}$$

Thus at the 233-ppm level we can rule out any 511 keV - 511 keV 2γ decay modes at the 6σ level as accounting for the 1400-ppm decay rate difference between theory and experiment. (An independent experiment reported a similar result[17].) It is very important to note that this limit on $B_{2\gamma}$ includes allowed as well as exotic, QED-violating 2γ decay modes. It therefore substantiates at the 233-ppm level the assumption that the λ asymptotically approaches the correct decay rate near $t_0 = 450$ ns.

CONCLUSION

Our decay rate result of $\lambda_T = 7.0482 \pm 0.0016$ is 6.2 of its standard deviations above the theoretical value. The two photon decay modes considered above cannot be responsible for this disagreement between theory and experiment. Thus, at the present time there are no exotic decay modes that appear to be viable explanations. We do note that a value of the order $(\frac{\alpha}{\pi})^2$ coefficient, B in equation 2, equal to 250 ± 40 would resolve the disagreement. A coefficient that large cannot be ruled out *a priori*, and hence calculation of B is crucial in order to fully confront QED theory at the level of accuracy afforded by the experiments.

This work is dedicated to the memory of our colleague Arthur Rich. We thank the members of the Michigan positron group for helpful discussions. This work is supported by the National Science Foundation under grants PHY8803718 and PHY8818582 and by the Office of the Vice-President for Research of the University of Michigan.

REFERENCES

* Current address: Los Alamos National Laboratory, MS D449, Los Alamos, NM, 87545

1] Westbrook, C. I., Gidley, D.W., Conti, R. S., and Rich, A.: Phys. Rev. A, 1989, 40, 5489

2] Hasbach, P., Hilkert, G., Klempt, E., and Werth, G.: Nuovo Cimento, 1987, 97*A*, 419

3] Gidley, D. W. and Zitzewitz, P. W.: Phys. Lett., 1978, 69*A*, 97

4] Adkins, G. S. and Brown, F. R.: Phys. Rev. A, 1983, 28, 1164; Lepage, G. P., Mackenzie, P. B., Streng, K. H., and Zerwas, P. M.: Phys Rev A, 1983, 28, 3090

5] Caswell, W. G. and Lepage, G. P.: Phys Rev. A, 1979, 20, 36

6] Adkins, G. S.: Ann. Phys. (N.Y.), 1983, 146, 78

7] Nico, J. S., Gidley, D. W., Rich, A., and Zitzewitz, P. W.: Phys. Rev. Lett., 1990, 65, 1344

8] Mills Jr., A. P.: Appl. Phys, 1980, 23, 189

9] Frieze, W. E., Gidley, D. W., and Lynn, K. G.: Phys Rev B, 1985, 31, 5628.

10] Van House, J., Rich, A., and Zitzewitz, P. W.: Orig. of Life, 1984, 14, 413

11] Gidley, D. W.: Ph.D. Thesis, The University of Michigan, Ann Arbor, 1979

12] Carboni, G. and Dahme, W.: Phys Lett., 1983, 123*B*, 349; (1983); Amaldi, U., Carboni, G., Jonson, B., and Thun, J.: Phys. Lett., 1985, 153*B*, 444; Metag, V. *et al.*: Nucl. Phys., 1983, *A*409, 331

13] Asai, S., Orito, S., Yoshimura, K., Haga, T.: Phys. Rev. Lett., 1991, 66, 2440

14] Atoyan, G. S., Gninenko, S. N., Razin, V. I., and Ryabov Yu. V.: Phys. Lett., 1989, 220 B, 317

15] Dobroliubov, M. I. and Ignatiev, A. Yu.: Phys. Rev. Lett., 1990, 65, 679

16] Mills, Jr., A. P. and Zuckerman, D. M.: Phys. Rev. Lett., 1990, 64, 2637

17] Asai, S., Orito, S., Sanuki, T., Yasuda, M., and Yokoi, T.: Phys. Rev. Lett., 1991, 66, 1298

18] Gidley, D. W., Nico, J. S., and Skalsey, M.: Phys. Rev. Lett.: 1991, 66, 1302

IMPROVING THE ELECTRON g-2 MEASUREMENT

R.K. Mittleman, F.L. Palmer and H.G. Dehmelt
University of Washington
Seattle, Washington

ABSTRACT

The measurement and calculation of the anomalous magnetic moment of the electron currently provides the most stringent test of QED. In the past one of the limitations in the experiment has been the small inhomogeneous component of the magnetic field which has been used to detect the magnet resonances of the trapped electron. We have developed a technique which allows us to reduce this inhomogeneous field by as much as two orders of magnitude and have demonstrated the technique with a factor of fifteen reduction, virtually eliminating the broadening of the magnetic lines in our "geonium atom".

INTRODUCTION

In this talk I am going to discuss the current state of the measurement of the anomalous magnetic moment of the electron. This measurement and the accompanying theoretical calculation provide the most stringent test of QED to date. While the current measurement is in agreement with the theoretical calculation, improving this comparison is of great interest, as each new term in the calculation probes new domains of physics. The current measurement of $a_{e(m)} = 1, 159, 652, 185.5(4.0) \times 10^{-12}$[1] agrees with the calculated value of $a_{e(c)} = 1,159,652,214.0(6.7)(27.1) \times 10^{-12}$[2] at the level of about one standard deviation. The first, and smaller error in the theoretical value is the estimated calculation error, while the largest source of uncertainty is due to the experimental uncertainties in the fine-structure constant.

The gyromagnetic ratio of the electron (or any other charged body) is defined as twice the ratio of the spin and cyclotron frequencies, in a uniform magnetic field.

$$g \equiv 2\frac{\omega_s}{\omega_c}$$

Simple Dirac theory predicts that a point fermion will have a gyromagntic ratio of exactly 2. Using the machinery of QED one finds that this value is modified to be slightly larger then two, and the anomalous part of the gyromagnetic ratio is defined to be;

$$a \equiv \frac{g-2}{2} \approx 10^{-3}.$$

The anomaly frequency, which we measure is defined to be the difference between the spin frequency and the cyclotron frequency. These definitions result in the simple formula:

$$a = \frac{\omega_a}{\omega_c}$$

which are the quantities which we measure in our experiment. Our goal is to improve the measurement of the anomaly from its current uncertainty of 4 parts in 10^9. To do this we need to measure both of these frequencies to much better then that, which means that we need to measure: $\omega_a/2\pi$ to ≈ 0.1 Hz (out of 190 MHz) and $\omega_c/2\pi$ to ≈ 100 Hz out of 160 Ghz.

I will start by briefly describing the electron trap and detection scheme. This topic has been extensively cover in the past[3[so I will only try to provide enough background so that the rest of the discussion can be understood. This description will lead into the changes which have been made to reduce the linewidths and some preliminary experimental results. I will finish up by describing the various systematics which must be considered in deriving an anomaly value.

Electron Trap and Detection

The basics of the electron trap **[Fig. 1]** consist of three electrodes which is contained in a vacuum envelope, submerged in liquid helium. Each of these electrodes are hyperbolas of revolution

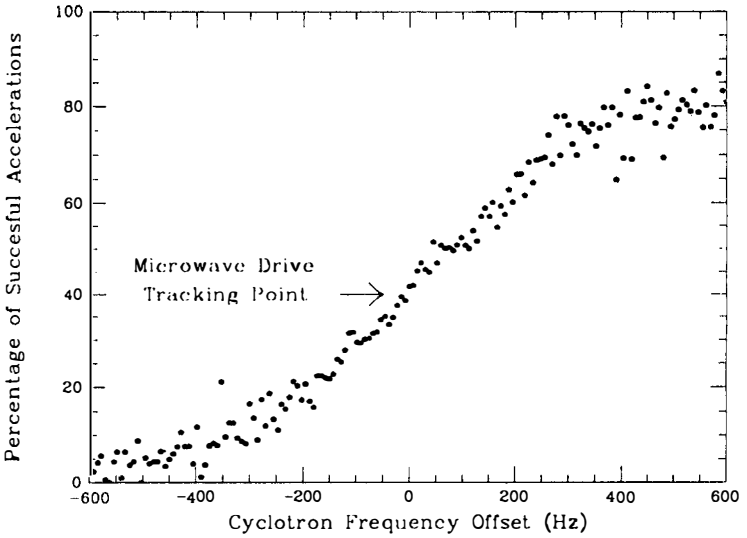


Figure 3. The probability of accelerating the electron to a large cycltron orbit ($n > 50$) versus starting frequency of the microwave drive. The zero is the cyclotron "lock" frequency.

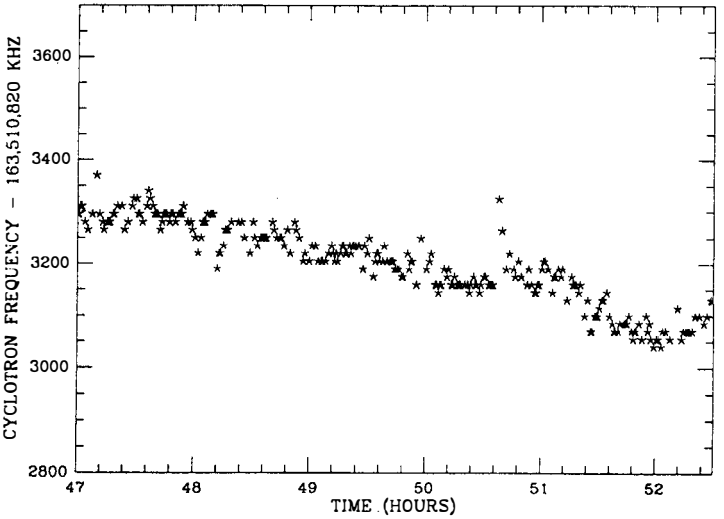


Figure 4. Typical cyclotron "lock" data, showing a width less then 0.5 ppb. The magnetic field shows a drift of approximately 0.2 ppb/hr.

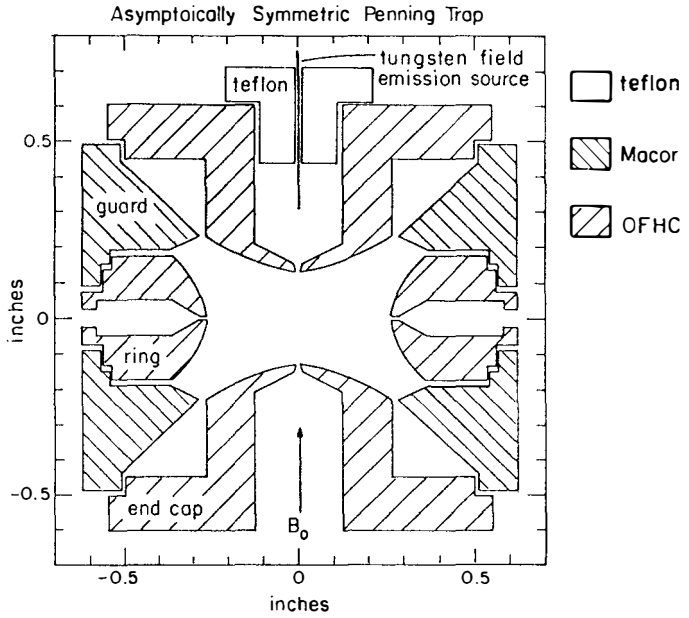


Figure 1. Scale drawing of the Penning trap used in this experiment

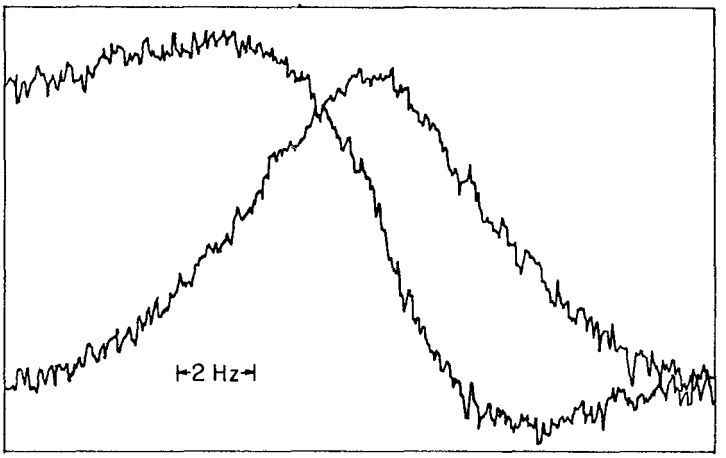


Figure 2. Dispersion and amplitude phases of the single electron axial resonance. The resonance is centered at 62,533,146.5 Hz and is 8.2 Hz wide.

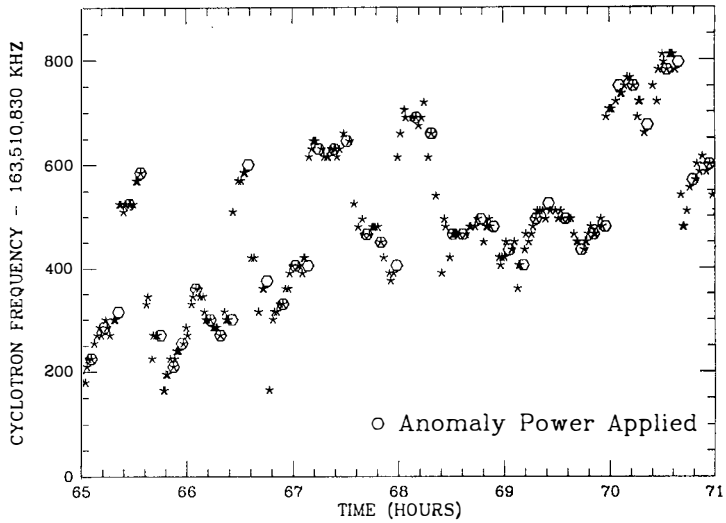


Figure 5. Cyclotron "lock" data with anomaly drive. The open hexagons mark when the anomaly drive was applied for one minute. About one quarter of the attempts resulted in a change of electron spin state.

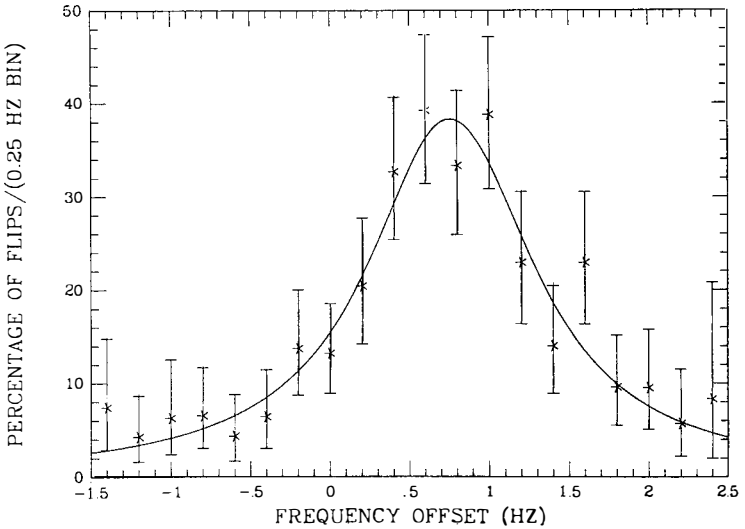


Figure 6. Anomaly data from 900 attempts to flip the electron spin from the high to low energy state. The probability of a change in state is plotted versus the offset from the calculated frequency using the previously published data. The solid line is the best fit from the theoretical formula, and gives and offset of $\omega_a/2\pi = 0.629 \pm 0.068 Hz$.

effects. At this level there are relativistic effects which are comparable to the axial shift due to the magnetic moment-bottle coupling. The dominant relativistic effect is the shifts in resonance frequencies due to the relativistic mass shifts with the energy of the trapped electron. The total axial frequency shift due to changes in the energy in the cyclotron motion is now:

$$\delta\omega_z = \omega_z[\frac{1}{m\omega_z^2}\frac{B_2}{B_o} - \frac{1}{2mc^2}]E_c$$

It turns out that the broadening of the magnetic lines is also proportional to this term, so at some special value of B_2 ($B_2 \approx 5\,\text{gauss/cm}^2$) both the line broadening and coupling (signal) can be tuned to zero. The natural line width of the anomaly line generated by the uncertainty principle and the cyclotron decay lifetime. This arises from the nature of our measurement which is a two photon transition, ($\Delta n = +1, \Delta s = -1$). A compromise between signal and line broadening led us to choose a magnetic bottle, $B_2 \approx 10\,\text{gauss/cm}^2$. The magnetic broadening due to this bottle is about 0.13 of a hertz considerably less then the natural cyclotron linewidth of 0.8 Hz, at our working frequency. As a result of this small bottle size we find that in order to create a detectable shift in the axial frequency of one hertz, we need to excite the cyclotron resonance to the 25th quantum level.

Just as the relativistic effect creates a shift in the axial resonance frequency, it also produces a shift in the cyclotron frequency,

$$\delta\omega_c = \omega_{co}(\frac{-1}{mc^2})E_c,$$

which makes the cyclotron resonance weakly anharmonic, $\delta\omega_c = 215$ Hz for a change of one quanta. This suggests that since a change in spin state changes the energy of the electron by the same amount as a change of one cyclotron level, if the cyclotron frequency can be monitored with sufficient accuracy, we will have a continuous monitor of the spin state of the trapped electron. To achieve this goal of measuring ω_{co} to an accuracy of 0.5 ppb we use a trigger technique. The microwave drive is turned on at ω_{do}, at a very low power. The drive frequency is then swept down ($\approx 40 \times 215 Hz$) and the power is ramped up. If ω_d passes over ω_{co} then electron can be caught and accelerated to very large cyclotron orbits (n's of over 1000 have been reached). If this "micro-synchro-cyclotron" is turned on with ω_{do} sufficiently less then ω_{co}, it cannot capture the electron into an accelerating orbit. So at the appropriate power levels a small change in the initial drive frequency, ω_{do} will determine whether enough energy will be put into the cyclotron orbit to create a detectable shift in the axial frequency. The acceleration probability has a width given by thermal noise, drive noise, etc. In our case the acceleration probability goes from near zero to about 75% in approximately 600 Hz [**Fig. 3**]. The frequency scan is then run many times, at the rate of about once every 3 seconds, and the drive (ω_{do}) is "locked" to a tracking frequency, (ω_t) where the success rate is 40%.

Once we have demonstrated that we have the necessary resolution of the cyclotron resonance [**Fig. 4**] we can attempt to flip the electron spin and measure the anomaly frequency. A series of spin flip attempts is shown in **Figure 5**. The transitions to higher ω_{co} represent transitions from the high energy spin state to the ground state. The probability of this transition as a function of the anomaly drive frequency offset from ω_a, as calculated from the previously measured anomaly value is plotted in **Figure 6**. The solid line is a fit of the theoretical formula[4]. These data fit a line

which are carefully machined to produce an axial harmonic well. When a voltage of $V_o = 10V$ is applied to the ring electrode the potential at the center of the trap is given by:

$$V = -\frac{V_o}{2d^2}(z^2 - \frac{\rho^2}{2})$$

where d is the characteristic trap dimension (0.335 cm). This potential creates a harmonic well in the z (axial) direction at:

$$\omega_z^2 = \frac{eV_o}{md^2} \approx (2\pi \times 62.5\,\mathrm{MHz})^2.$$

The electron is confined radially by a large uniform axial magnetic field ($B_o \approx 58500$ gauss). These fields result in a radial motion which is the combination of two motions, a cyclotron orbit:

$$\omega_c = \frac{eB_o}{mc} \approx 163\ \mathrm{GHz}$$

and a magnetron motion due to the $E \times B$ drift:

$$\omega_m = \frac{\omega_z^2}{2\omega_c} \approx 12\,\mathrm{KHz}$$

which is unimportant to the following discussion and will not be mentioned any further. An important point to note is that the energy corresponding to a change in cyclotron level of one quantum number ($\hbar\omega_c = 6.8 \times 10^{-4} ev$) is larger than the typical energy due to thermal excitation, ($kT = 3.6 \times 10^{-4} ev$). The consequence of this is that the cyclotron motion is in its ground state approximately 85% of the time.

The way which we monitor the electron motion is by connecting a tuned amplifier to one of the endcaps. By driving the electron at the axial resonance frequency and detecting the synchronous component of the signal the axial resonance can be easily seen [**Fig. 2**]. This signal is the only way in which we monitor the electron, consequently in order to detect the magnetic resonances they must be coupled to the axial motion. The initial method was to add a small "magnetic bottle" in addition to the large axial trapping field. This consists of an axial magnetic field proportionate to z^2 which is typically created by a small iron ring wrapped around the trap in the center plane of the trap. The magnetic moment of the oscillating electron couples to this field creating a small change in the axial restoring force and resonance frequency proportional to the electron magnetic moment,

$$\omega_z = w_{zo} - \frac{B_2\mu_z}{m\omega_{zo}} \quad \text{where}$$

$$\mu_z = -(\frac{e\hbar}{mc})n - \frac{g}{2}(\frac{e\hbar}{mc})s.$$

For a large enough magnetic bottle (B_2) a change in spin state or a change of cyclotron quantum number (n) of one, can be detected directly. In practical terms this works out to a bottle large enough to give a shift of 1 Hz out of 62 MHz for a change of one quantum level. This inhomogeneity has the unfortunate effect of asymmetrically broadening out the magnetic lines. In order to overcome this we have greatly reduced the size of the B_2 term, such that the bottle shift is comparable to relativistic

which is essentially Lorentzian in shape and shows only minimal asymmetrical line broadening from the magnetic bottle.

Although we have managed to narrow the anomaly resonance by about a factor of five, we can not yet claim to have improved the g-2 value. This is due to various systematics involved in the experiment, which we are currently working on. The majority of the systematics involve shifts of the various resonance frequencies as a function of an applied drive. In general these shifts are fairly easily dealt with by repeating the experiment a number of times with different drive levels. Another important systematic is the difference between the cyclotron tracking frequency, ω_t and the cyclotron ground state resonance frequency, ω_{co}. To determine the relationship between these two frequencies we apply a short "π – pulse" immediately before the normal cyclotron excitation cycle. When the frequency of the pulse is equal to ω_{co} it should maximally affect the electron acceleration process and give us a very accurate measure of ω_{co}.

The last major systematic is that due to cavity shifts. The fact that there is a conducting cavity around the electron produces small shifts in the resonant frequencies. In particular ω_{co} can be shifted by ten parts in 10^{12} if we happen to be working near a mode of the cavity[5,6]. The shifts in ω_s are much smaller and can be estimated using a simple semi-classical argument, which has been checked using QED[7]. In trying to calculate the frequency shifts due to the mode structure of a hyperbolic trap the perturbations induced by trap imperfection dominate, and an experimental solution to this problem must be found. It is possible to crudely measure the mode structure of the trap over a large frequency range, using a bolometric technique. When combined with high precision measurements of cyclotron lifetime and anomaly value in a narrow frequency range around our working conditions[6] we expect to be able to greatly improve the measurement of the anomalous component of the magnetic moment of the electron.

CONCLUSIONS

We believe that we have found a way to significantly increase the accuracy of the measurement of the anomalous component of the electron magnetic moment. If the corresponding measurement of the fine structure constant and the QED calculation can be improved to match then this test of QED can be improved by almost an order of magnitude. Until that time the g-2 measurement can be converted into the highest precison measurement of α. The improvement in the anomaly linewidth also suggests that we should be able to make a corresponding improvement in the comparision between the electron and the positron[8].

We thank the National Science Foundation for support of the SEPARIS project.

1. R.S. Van Dyck, Jr., P.B. Schwinberg, and H.G. Dehmelt, The Electron, pp239-293 (1991). Kluwer Academic Publishers.
2. T. Kinoshita, W.B. Lindquist, Phys Rev. D **42**, pp636-655 (1990);
3. L.S. Brown and G. Gabrielse, Rev. Mod. Phys. **58**, pp233 (1986); R.S. Van Dyck, Jr., P.B. Schwinberg and H.G. Dehmelt, Phys. Rev. **D 34**, pp722-736 (1986).
4. L.S. Brown, Annals of Physics **159**, pp62-98 (1985).
5. L.S. Brown, G. Gabrielse, J. Tan and K.C.D. Chan, Phys. Rev. A **37 No 11**, pp4163-4171 (1988).
6. H.G. Dehmelt, R.S. Van Dyck and F.L. Palmer, Proc. Natl. Acad. Sci. USA **89**, pp1681-1684 (1992). (1992)
7. D.G. Boulware, L.S. Brown and T. Lee, Phys. Rev. D **32 No 3** pp729-735 (1985).
8. R.S. Van Dyck, Jr., P.B. Schwinberg, and H.G. Dehmelt, Phys. Rev. Lett. **59 No 1**, pp26-29 (1987).

ATOMIC ENERGY LEVEL SHIFTS IN CONFINED SPACE

E. A. Hinds
Physics Department Yale University, New Haven Ct., USA.

When atoms are placed inside a cavity, the energy levels and radiation rates are altered. If the cavity is sufficiently small, the level shift can be understood as a van der Waals interaction between the instantaneous electric dipole moment of the atom and its multiple electrical images in the cavity walls. By contrast, if the walls are sufficiently distant, and if the atom is in its ground state, one expects a Casimir interaction between the atom and the modified electromagnetic vacuum within the cavity. Recent and ongoing experiments at Yale University are observing and testing these cavity QED effects in the laboratory.

Introduction

We are investigating the properties of atoms confined to a small space by plane metallic or dielectric boundaries separated by ~1μm or less. Inside such a cavity or waveguide, the spectrum of the electromagnetic field modes is strongly modified for wavelengths that are comparable with the physical dimensions. Thus it is possible by a suitable choice of cavity geometry to influence the fluctuations of the quantized radiation field and hence to change both the spontaneous decay rate and Lamb shift. Since the energy-level shifts are position-dependent, there are also forces between the atom and the cavity walls. The early references are all theoretical, but this field of research has now progressed to the laboratory and is the subject of several recent review articles.[1,2,3)]

A good paradigm for the study of perturbative cavity QED level shifts is the simple problem of an atom in front of a plane, perfectly conducting mirror surface.[4,5,6)] Fig. 1 shows the calculated energy shift of a sodium atom in the ground state hyperfine sublevel ($F = 2, m_F = 2$) as a function of distance from a plane mirror of high reflectivity. The distance scale is determined by the autocorrelation or "memory" time of the atomic electric dipole, which in this case is of order $\lambda_0/2\pi c$, where $\lambda_0 \approx 600$nm is the free-space wavelength of the 3S-3P resonance line. At distances small compared with $\lambda_0/2\pi$, the level shift asymptotically approaches the Lennard-Jones van der Waals form[7)],

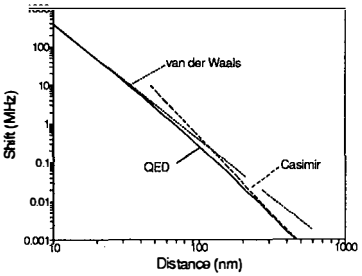


Figure 1.

$$\delta_{inst} = -\frac{1}{4\pi\varepsilon_0}\frac{d_\rho^2 + 2d_z^2}{16z^3}, \tag{1}$$

where d_ρ^2 and d_z^2 are the mean square electric dipole moments of the atom parallel and normal to the mirror surface. This is just the interaction of the instantaneous atomic dipole moment with its electric image in the mirror. Of course, this dipole approximation is good only when z, the distance to the mirror, is much greater than the radius of the atom.

At distances much greater than $\lambda_0/2\pi$, the time taken for the electromagnetic field to travel to the mirror and back is longer than the memory time of the fluctuating atomic dipole and the van der Waals interaction vanishes. Nevertheless, something still happens at this long range - the level shift approaches the Casimir-Polder form

$$\delta_{Casimir} = -\frac{1}{4\pi\varepsilon_0}\frac{3\hbar c}{8\pi z^4}\alpha_{stat}, \tag{2}$$

where α_{stat} is the scalar electric polarizability. This energy shift can be understood as a change in the Stark shift produced by the vacuum - i.e. a change in the Bethe contribution to the Lamb shift - resulting from the modified vacuum field. At a given distance z from the mirror, the spectrum of the vacuum field is modified strongly only for frequencies below c/z. At the long-range limit that we are considering here, (propagation delay exceeds memory time) the affected part of the vacuum spectrum is at lower frequency than all the atomic frequencies and is effectively static. Hence the appearance of the static electric polarizability in eq. 2 is quite natural. An excellent review of Casimir forces in this and other systems is given in *Physics Today* by Spruch.[8)]

The situation is a little different when the atom is in an excited state. At distances less than the characteristic $\lambda_0/2\pi$ (typically λ_0 is the wavelength of the strongest E1 transition), the level shift is once again given by the instantaneous Lennard-Jones van der Waals potential. However, at longer range the main effect is the emission of a photon which reflects from the mirror and is then reabsorbed; the interaction is precisely that of an antenna with its own coherent image in the mirror. The shift is oscillatory with an amplitude that decreases only as 1/z, the amplitude of the far field, exactly as one would expect from classical physics.

Our experiments are conducted in a parallel-plate waveguide, for which mathematics is more complicated,[9)] but the basic physics is unchanged. For a ground-state sodium atom at the center of a narrow waveguide of width $L < \lambda_0/2\pi$, the energy shift is due to the instantaneous van der Waals interaction between the atom and its multiple images in the mirrors and is well approximated by

$$\delta_{inst} = -\frac{1}{4\pi\varepsilon_0}\frac{\zeta(3)}{4L^3}(3d_\rho^2 + 8d_z^2). \tag{3}$$

where $\zeta(3) \approx 1.20$ is the Riemann zeta function. At the other extreme, $L \gg \lambda_0/2\pi$, the sodium ground state level shift is the modified Lamb shift or Casimir interaction and it has the form

$$\delta_{Casimir} = -\frac{1}{4\pi\varepsilon_0}\frac{11\pi^3\hbar c}{90L^4}\alpha_{stat} \tag{4}$$

Related Experiments

Experiments to demonstrate and measure these, or closely related, potentials have a long history, which started in the early 1950s with attempts to measure long-range forces between macroscopic neutral bodies, both metal and dielectric.[10] The earliest of these measurements stimulated the development by Lifshitz and coworkers[11] of a macroscopic theory of dispersion forces which accounts for the many-body aspects of the interaction by means of a frequency-dependent complex refractive index. The predictions of this theory tend to be somewhat qualitative since the frequency-dependent refractive index is usually only partially known. As a result of all this work, experiment and theory are now generally in agreement within a factor of two.

A more recent approach has been to measure[12] the energy intervals between helium Rydberg levels and to compare the results with theory.[13] The comparison can be regarded as a test to see that the van der Waals interaction between two charged systems, the Rydberg electron and the He^+ core, is retarded.[14] In this experiment, the mean separation betwen Rydberg electron and core (~5 nm) is approximately one sixth of the characteristic wavelength for the He^+ core (~30nm), so the range is an intermediate one where the polarization interaction is neither instantaneous nor fully retarded. In addition, of course, there are other much stronger interactions due to the fact that the electron and He^+ core are charged. At the moment, experiment and theory differ by several (experimental) standard deviations and by several times the Casimir-Polder term[13] but the disagreement is ten times larger if retardation effects are omitted entirely, providing strong evidence for retardation in the electromagnetic interaction of charged systems.

The experiments most closely related to our own measurements are those in which atoms or molecules were deflected by a metallic surface.[15] In all these experiments, the deflections were unobservably small except for the few atoms that came within 50 - 100 nm of the surface. Since this is short range in terms of the dipole memory time of the atoms or molecules studied, the deflections were due to the (almost) instantaneous van der Waals coupling, not to the retarded Casimir-Polder interaction. In the cases where experiment could be compared with *ab initio* theory, unexplained discrepancies were found at the factor-of-two level. In contrast, our recent spectroscopic measurements[16] (described below) confirm the van der Waals part of the theory at the level of a few percent. This may well be a consequence of the very different frequencies involved in the dipole fluctuation spectrum. In the case of Rydberg atoms the dipole fluctuations are predominantly at microwave frequencies where the gold mirrors are well described as perfect reflectors. In that case the computation of the van der Waals force does not involve any surface physics. The ground state atoms on the other hand fluctuate most strongly at optical frequencies, corresponding to the resonance transitions, where the mirrors are not perfect. In that case it is necessary to model the surface involved in the experiment in some detail in order to obtain agreement between theory and experiment.

Direct Measurement of the van der Waals Interaction Between an Atom and its Images in a Micron-sized Cavity

Here we describe the first direct and precise measurement of the van der Waals coupling between an atom and a metallic surface, in which the interaction is quantitatively studied as a function of controlled separation and of the electronic state of the atom. This work was carried out in collaboration with V. Sandoghdar, C. I. Sukenik and S. Haroche.

Figure 2a shows the main features of the experiment. An atomic beam of sodium effuses from an oven at 200°C into a 10^{-7} torr vacuum and enters the space between two plane, gold-coated mirrors (flatness ~5 nm), which is 8mm long in the beam direction and 3cm high. The mirrors are pressed against each other to form a wedge whose width is zero at the bottom, while the width at the top can be varied by pushing on the mirrors with screws in

order to compress a 3μm nickel foil spacer acting as a spring. We set the width of the gap by illuminating the cavity with monochromatic light (mercury green, λ=546nm) and adjusting the screws until one of the interference fringes coincides with the laser beams.

Approximately half a millimeter before the cavity exit, the atoms are excited by two superimposed laser beams, one yellow (λ = 589nm) and the other blue (λ~425nm), either to the $10D_{5/2}$ state or to one of the states nS, with n=10-13. The relevant energy levels are shown in fig.2b. The laser beams are focused to a 300μm waist and are carefully aligned to be normal to the mirrors, and therefore to the atomic beam, in order to avoid any first order Doppler effect. Excited atoms leave the cavity and enter a detector where they are field-ionized and counted using a channel electron multiplier. An excitation spectrum is obtained by scanning the wavelength λ_2 while recording the rate of detected ions. The essence of the experiment is to study the shifts of the nS excitation spectra at various mirror spacings, the $10D_{5/2}$ spectra being used to monitor stray electric fields as discussed below.

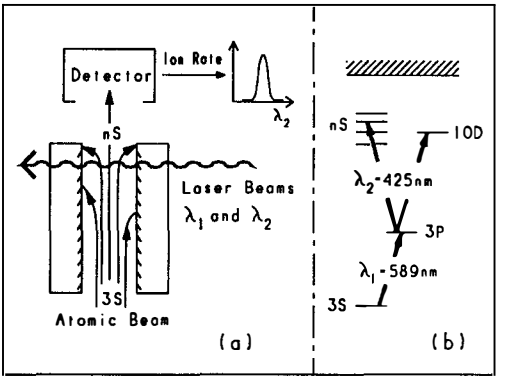


Figure 2

Figure 3 shows the van der Waals potential for the 13S state of sodium in a parallel-plate cavity 1μm wide due to the infinite series of images resulting from successive field reflections. The interaction energy is negative and diverges at each mirror surface. At the center of the gap, this potential reaches a maximum given by eq. (3). The nS states of this study have an electronic extension of order

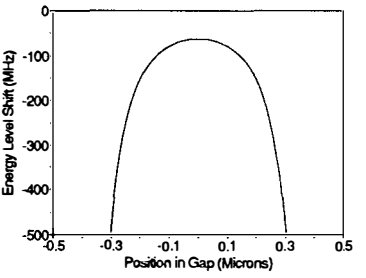


Figure 3.

10 nm and their characteristic emission and absorption wavelengths λ_0 for the (nS $\rightarrow$ n'P) transitions are in the 50 - 150 μm range. Thus the van der Waals formula (3) holds for L between ~100 nm and ~10,000 nm and is valid for the full range of gap sizes we have explored. Moreover, our spectra are sensitive only to the shifts of the excited nS levels since the shifts of the 3S and 3P states are completely negligible in comparison.

At first sight, the potential shown in Fig. 3 suggests very asymmetric spectral lines with a high frequency edge due to atoms at the center of the gap and a long tail on the low frequency side from atoms passing closer to either mirror. In fact, the lines are typically only 20 - 30 MHz wide, indicating that the atoms we detect are excited only near the center of the gap over a region of width 0.2L - 0.3L. This is due to a combination of several effects. First, before laser excitation, atoms far from the center are deflected by the atom-cavity interaction onto the mirrors where they stick and are therefore removed from the beam. Second, the excited atoms experience a much larger attraction to the walls and cannot escape from the cavity to be detected unless they are sufficiently close to the center at the point of laser excitation.Finally, the detected atoms contribute to the spectrum in proportion to the inverse of the slope of the potential at their position, which again favors those excited at the center. Excitation spectra were observed for four nS states (n=10 - 13) for each of the seven cavity widths L corresponding to the third through ninth fringes of the green light. In Fig. 4, graphs (a)-(e) show some of the 13S spectra in successively narrower cavities. The peak on the right of each trace is a reference line obtained on unperturbed atoms in an auxiliary atomic beam. Looking at the peak on the left, the van der Waals shift is clearly seen.

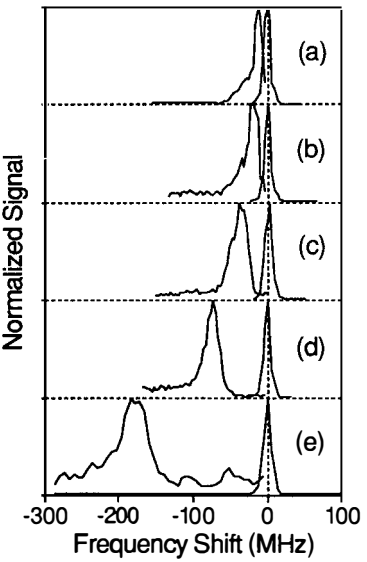


Figure 4.

Three main systematic effects must be taken into account when analyzing these spectra. First, in spite of the strong spatial selection of the atomic beam, the peaks of the spectral lines lie at a slightly lower frequency than they would if all the atoms were exactly at the center of the cavity. A simple numerical model shows that this offset δ_{offset} grows with linewidth and reaches a maximum of -3.3 MHz when the observed linewidth is larger than 1.5Γ, Γ being the free space linewidth which is 11 MHz in our experiment.

A second small correction is required because of the Stark shift δ_{Stark} of the nS states due to an unavoidable stray electric field in the cavity. This electric field produces much larger Stark shifts of the very polarizable $10D_{5/2,m_J=5/2}$ sublevel, which we measure to determine the required correction. In order to find the nS Stark shift δ_{Stark}, we subtract the calculated van der Waals shift of the $10D_{5/2,m_J=5/2}$ level from the measured shift of that level and divide by the ratio of electric polarizabilities $\alpha(10D_{5/2,m_J=5/2})/\alpha(nS)$. This ratio is calculated to vary from 24 for n=13 to 150 for n=10, a result that we checked by recording Stark spectra in known electric fields outside the cavity. We note that due to the large values of these ratios, the Stark correction is actually very insensitive to the assumed value of the $10D_{5/2,m_J=5/2}$ van der Waals shift. The largest nS Stark shift is found in the 13S state and in the smallest gap, where $\delta_{measured} = -178$MHz while $\delta_{Stark} = -6$MHz, corresponding to a stray field of 44 V/cm. In every case the correction is less than 10% of the van der Waals shifts observed. In order to reach this low level of stray electric field, we find it important to bake the mirrors at 120°C for several hours after placing them in the vacuum system. The electric field is then found to vary roughly as L^{-2}, which is consistent with a patch effect due to the crystalline structure of the evaporated gold surface.

The last systematic correction concerns our measurement of the cavity width using the green light fringes. If the phase of the reflection coefficient is written as $\pi + \phi$, the width of the cavity at the m^{th} fringe is $(m - \phi/\pi)\lambda/2$. When ϕ is estimated using 0.32+2.22i for the complex refractive index of gold at 546 nm, we find that the cavity should be narrower than $m\lambda/2$ by 72 nm. Direct measurements on a similar cavity using an electron microscope show that the actual correction is 85 ± 15nm.

The crosses in fig. 5 show the observed van der Waals shifts, $\delta_{measured} - \delta_{offset} - \delta_{Stark}$, versus gap width for each of the four nS levels. The width of the cross (±21 nm) reflects the total uncertainty in the mirror spacing, while the height indicates a ±2 MHz uncertainty in frequency shift due to residual Doppler shift and measurement errors. The lines correspond to the Eq. (1), without any free parameters.

This experiment provides the first direct and precise check of the Lennard-Jones formula. The result clearly verifies both the z^{-3} law and the proportionality to the mean square atomic dipole. We see that the boundary around the atom affects not only the natural widths of the levels, as was demonstrated in earlier Cavity-QED studies,[3] but also shifts the energy levels by an amount which, in the distance range of our experiment, is given precisely by the Lennard-Jones theory. This shift results from the near-field instantaneous coupling of the atomic dipole with its images in the mirrors.

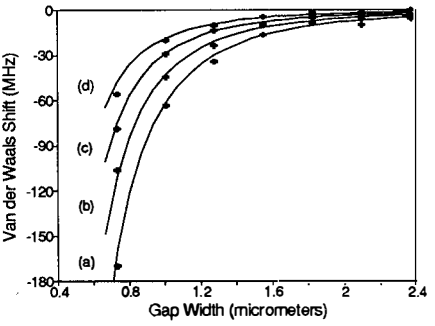


Figure 5

Progress Towards Measuring the Casimir Force

Between an Atom and a Micron-Sized Cavity

There is also an ongoing attempt in our laboratory to measure the long-range retarded Casimir force on a ground-state atom in a cavity. This work is in collaboration with C. I. Sukenik, V. Sandoghdar, M. G. Boshier and D. Cho. We do not use a spectroscopic method because any transition to an excited level is affected by the much larger shift and linewidth of that level. Instead, we observe the deflection of a sodium atomic beam due to the Casimir force as the atoms pass between two plane parallel plates, separated typically by 2μm.

Figure 6 shows the calculated energy level shift of the sodium ground state in a waveguide of width 2μm ($\ll \lambda_0/2\pi$). We have taken the mirrors to be perfect since the reflectivity of gold is indeed extremely close to unity for the low-frequency fields ($\omega < c/L$) that are relevant in determining the Casimir force. We have also taken the temperature to be zero in this calculation, neglecting the presence of thermal photons at frequencies below $kT/\hbar$. A "back of the envelope" estimate of the fractional correction due to finite temperature gives

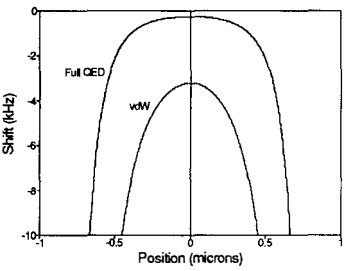


Figure 6.

$$\frac{\delta(T)-\delta(T=0)}{\delta(T=0)} \approx \left(\frac{kTL}{\hbar c}\right)^4, \tag{5}$$

which amounts only to 1% when $L = 2\mu m$. The experiment we are conducting here is to measure the Casimir interaction in cavities 1 – 2μm wide where thermal photons will not have a large effect (although we would be able to observe thermal forces in a cavity of different geometry). For comparison, fig. 6 also shows the much stronger van der Waals potential that would exist if there were no retardation of the electromagnetic field.

Referring once again to fig. 2, sodium atoms from our effusive atomic beam source fly into the 8 mm-long waveguide whose width is adjusted from a few hundred nanometers to ten microns. In this experiment, the two laser beams ($\lambda_1 = 598nm$ and $\lambda_2 \approx 420nm$) are positioned outside the cavity so that atoms emerging from the cavity are excited to the 13D level by the two steps 3S-3P-13D and can then be field-ionized and counted in a channel electron multiplier. The number of transmitted atoms is measured as a function of the cavity width, which we can determine with 20 nm accuracy as described above. When this gap is large, the transmission is expected to be close to geometrical since the atomic trajectories are nearly straight, except for those atoms within 100 nm or so of the mirrors. However, as the width of the cavity is decreased, the Casimir forces becomes stronger and the atoms must

travel more nearly along the ridge of the cavity potential if they are to make it through. Thus the transmitted fraction falls below the geometrical transmission in a way that is sensitive to the shape and strength of the atom-cavity interaction potential.

Figure 7 shows the transmission of atoms through the cavity as a function of mirror spacing for three cases, (a) no force, (b) force predicted by QED (essentially the Casimir force), (c) van der Waals force, which is unphysical since it ignores retardation. These curves were computed using a Monte Carlo program, which takes into account the known geometry and velocity distribution of our thermal beam, to calculate a large number of classical particle trajectories in the various potentials. Also shown are our preliminary results which show a clear preference for the retarded Casimir potential over the instantaneous van der Waals potential. We do not know how accurate this experiment can ultimately be made but it can certainly be much improved and promises to yield the first accurate measurement of this Casimir potential.

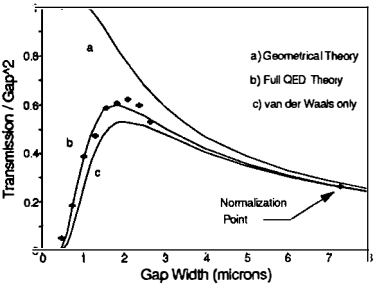


Figure 7.

One very important aspect of this experiment is the control of stray electric fields, which has been one of our main achievements over the last two years. We have learned to probe the electric field by observing excitation spectra of the highly polarizable nD states and this has allowed us to develop techniques for building cavities in which the stray electric field is small. By means of a baking procedure, which cleans the surface or at least distributes the surface contamination uniformly, we are able to reduce the stray field to the patch field level, determined by the size of the surface grains. We have also learned how to make mirrors in which the patches are as small as possible and how to avoid annealing them. As a result we are now able to make cavities in which the electric field at the center of a 1μm gap can be directly measured and is as small as 20V/cm.

The acceptable level of stray electric field is determined by comparing the Stark shift with the Casimir shift. Both are proportional to the electric polarizability α_{stat}, so the comparison is simple - using eq. 4 one finds at the center of a 1μm cavity that the two are

equal when the electric field is 464V/cm. Both shifts also scale as $1/L^4$ because the average patch field varies as $1/L^2$. Hence the control we have achieved over stray electric fields is already sufficient to make the Stark shifts smaller than 1% of the Casimir shifts at all relevant gap widths.

This experiment already provides very clear qualitative evidence for an atom-cavity interaction at ranges where the classical van der Waals interaction is very strongly retarded. In addition, these preliminary data show that the measured transmission probabilities are consistent with the expected Casimir interactions and inconsistent with a simple unretarded van der Waals potential. Further improvements in this work are expected soon.

References

1. Walther H., Physica Scripta **T23,** 165 (1988).
2. Haroche S. and D. Kleppner, Physics Today, **42,** 24 (1989).
3. Hinds, E.A., Advances in Atomic, Molecular and Optical Physics **28**, 237 (1991).
4. Barton G., Phys. B **7,** 2134 (1974).
5. Meschede D., W. Jhe and E. A. Hinds, Phys. Rev. A **41**, 1587 (1990).
6. Hinds, E. A. and V. Sandoghdar, Phys. Rev. A **43**, 398 (1991).
7. Lennard-Jones J. E. Trans. Faraday Soc. **28**, 333 (1932).
8. Spruch L. Physics Today **39,** 37 (1986) .
9. Barton G., Proc. Roy. Soc. Lond. A **410,** 141 (1987); *ibid* p.175.
10. Derjaguin, B.V., Y.I. Rabinovich and N.V. Churaev, Nature **272**, 313 (1978).
    Israelachvili J. N. *Intermolecular and Surface Forces* (Academic Press, 1985).
    Sparnaay M. J., in *Physics in the Making* edited by A. Sarlemijn and M. J. Sparnaay p.235 (North Holland, Amsterdam 1989).
11. Lifshitz E. M., Sov. Phys. JETP **2**, 73 (1956).
    Dzyaloshinskii I. E., E. M. Lifshitz and L. P. Pitaevskii, Adv. Phys **10**, 165 (1961).
12. Hessels, E.A., P.W. Arcuni, F.J. Deck, and S.R. Lundeen, submitted to Phys. Rev. A (1992).
13. Goldman S. P., and G. W. F. Drake, Phys. Rev. Lett. **68,** 1683 (1992).
14. Au C. K., G. Feinberg, and J. Sucher, Phys. Rev. Lett. **53**, 1145 (1984); Phys. Rev. A **43**, 561 (1991).
15. D. Raskin and P. Kusch, Phys. Rev. **179** 712 (1969); A. Shih, D. Raskin and P. Kusch, Phys. Rev. A **9** 652 (1974); J. Bardon and M. Audiffren, J. Phys. Lett. **44**, 883 (1983); A. Anderson et al., Phys. Rev. A **37** 3594 (1988).
16. V. Sandoghdar, C. I. Sukenik and E. A. Hinds and S. Haroche, to be published in Phys. Rev. Lett. June (1992).

Physics and Computers

Devices which Transfer Electrons One-by-One

Ph. Lafarge, Ph. Joyez, H. Pothier, A.N. Cleland, T. Holst,
D. Esteve, C. Urbina and M.H. Devoret

Service de Physique de l'État Condensé
Orme des Merisiers
CEA Saclay
91191 Gif-sur-Yvette Cedex
France

Abstract

This paper provides an introduction to the field of single-electron devices, which are electronic circuits across which electrons may be transferred one by one. The paper will describe experiments on the single-junction box and the four-junction box, devices in which electrons were transferred to a metallic island and their presence detected using a single-electron electrometer. The single-junction box, the simpler of the two devices, allows one to control the thermally-averaged number of excess electrons on the metallic island, whereas with the four-junction box one can actually control the transfer of individual electrons.

In 1911, Millikan[1] showed that the charge on a drop of oil is always an integer multiple of the electron charge e; however, the discreteness of the electron charge is rarely evident in electronic circuits, appearing only as shot noise in the current flowing through a vacuum tube or a semiconductor diode. Screening effects in metals and semiconductors, as well as thermal noise and the spread of the electron wavefunctions, makes the conduction process appear fluid and continuous.

The first realization that the conduction process can be strongly affected by the discreteness of the electronic charge came through experiments on granular metallic thin films, as discussed by Gorter[2]. Experiments had shown that the current-voltage (I-V) characteristics of these films exhibit a region of low conductance at small bias voltages V, and that the conductance of this region increases with temperature T. Gorter postulated that the conduction in these films was due to the hopping of electrons from one metallic grain to the next. If a grain has a self-capacitance C, then the energy E to place one electron on the grain is $E = e^2/2C$. If the capacitance is small enough that the energy satisfies $E >> k_B T$, then the flow of electrons across the grain will be inhibited, resulting in a small value of conductance. When the temperature is increased to a value larger than the largest value of E in the conduction path across the film, this effect will disappear.

In 1986 Averin and Likharev[3] launched the field of single-electron devices, by developing some of the fundamental theoretical ideas in the field, and by proposing that it was possible to intentionally fabricate devices in which the conduction process is directly affected by the discreteness of the electron charge. The fabrication of such devices became possible though the development of nanolithographic techniques, where the use of the scanning electron microscope (SEM) as a lithography exposure tool allows the patterning of thin metal films with dimensions on the order of 100 nm; such small dimensions allow one to meet the charging condition $e^2/2C >> k_B T$, if the measurements are carried out on a dilution refrigerator at a temperature well below 1 K. The single-electron devices which have been developed to date all have a few basic elements in common; they consist of metal wires which are interrupted either by pure capacitors, which allow no electrons to pass across, or by metal-insulator-metal tunnel junctions, which are simply capacitors whose two plates are placed so close together (about 1 nm) that electrons may cross from one plate to the other by quantum-mechanical tunneling through the insulating barrier. When the two metals are in the

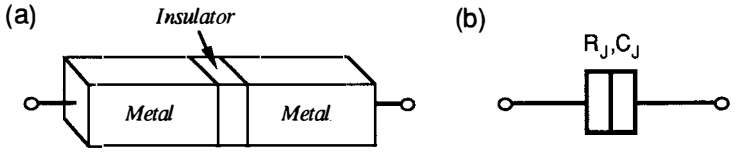


FIGURE 1. (a) Schematic representation of a tunnel junction; the insulator is typically about 1 nm thick. (b) Electrical symbol used to represent a tunnel junction, with tunnel resistance R_J and capacitance C_J.

normal state (as opposed to the superconducting state), and are biased with a constant voltage V, the rate of electron tunneling through the insulating barrier is quite accurately proportional to the voltage: the number of states available for an electron to tunnel into increases linearly with voltage, and by Fermi's rule the rate increases proportionally. The current I through the junction is therefore given by $I = V/R_J$, where the parameter R_J is called the junction resistance. A schematic drawing of a junction and the electronic symbol used to designate it are shown in Figs. 1(a) and (b); note that the geometry of the junction implies that it has a non-negligible capacitance C_J. A requirement on the tunnel junction is that its resistance must be sufficiently high to meet the criterion $R_J >> h/e^2 \approx 25.8\ \text{k}\Omega$, as otherwise the electrons are not sufficiently localized on one side of the barrier as opposed to the other[4].

The Single Junction Box

The most elementary example of a single-electron device, the single junction box, is shown in Fig. 2. The series combination of the capacitor and the tunnel junction forms a metallic island, consisting of the right plate of the capacitor C_S, the left plate of the tunnel junction, and the wire connecting them, as outlined in the figure. The metallic island is designed to have a very small capacitance $C = C_J + C_S$, on the order of 1 fF; in order to place a single electron on this island, one needs to provide an energy $e^2/2C$ corresponding to a temperature of about 1 K. If this device is cooled to about 20 mK by using a ^{3}He-^{4}He dilution refrigerator, the number of electrons on the island will be solely determined by the bias potential U applied across the capacitor and the tunnel junction, and as this potential is increased from zero, the number of electrons will increase one by one in order to minimize the electrostatic energy.

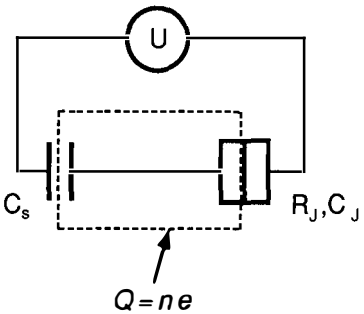


FIGURE 2. Electrical circuit for the single-junction box, consisting of a voltage source U, bias capacitor C_S and a tunnel junction; the number of electrons n on the island is a discrete integer.

More specifically, one can write the electrostatic energy E of the complete circuit in terms of the number of electrons n on the island and the charge $Q = C_S U$:

$$E_n = \frac{[\, n(-e)+Q \,]^2}{2(C_S+C_J)} - \frac{Q^2}{2C_S} \quad . \tag{1}$$

The energy as a function of Q with n fixed is a parabola, as expected, and as one varies n across its allowed integer values one finds a family of parabolas. The system will always be in the state of lowest energy, at least at zero temperature. The value of n will therefore increase by one whenever the value of Q crosses the values $-3e/2$, $-e/2$, $e/2$, etc. The effect of finite temperature will be to make the measured value of n correspond to the thermal average $\langle n \rangle$ over all states E_n, as long as the measurement is over a time much longer than the inverse electron tunneling rate $\Gamma^{-1}(n \rightarrow n \pm 1)$, which is at most of order $R_J C_J$, or less than 1 ns. In general, the transfer rate of electrons across a junction is given by[5]

$$\Gamma = \frac{1}{e^2 R_J} \frac{\Delta E}{1 - \exp(-\Delta E / k_B T)} \quad , \tag{2}$$

where ΔE is the change in total electrostatic energy, including the work done by the external voltage sources.

The single junction box presents an experimental system in which all of the fundamental elements of this type of device can be found. The authors therefore performed an experiment to demonstrate that these elements were understood, by measuring the variation in the electron number n as a function of the bias voltage U; this was to be done by measuring the voltage developed across the tunnel junction due to the charge Q_J on its capacitance C_J. It can easily be shown that the thermally averaged charge $\langle Q_J \rangle$ on the junction is given by

$$\langle Q_J \rangle = \frac{C_J}{C_S+C_J}(\langle n \rangle(-e) + Q) \quad . \tag{3}$$

The major difficulty in measuring $\langle Q_J \rangle$ is that by connecting leads of any significant length to the island of the box, one adds a very large (~ 1 pF) stray capacitance in parallel to the junction capacitance, and in doing so enormously reduce the difference in energy between the states n and $n\pm1$. The resolution of this problem was to use an electrometer with very small input capacitance, as described below.

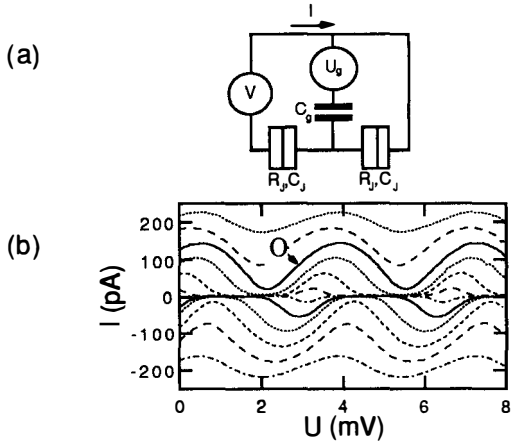


FIGURE 3. (a) Circuit diagram for the electrometer. It is usually operated by holding the bias voltage V constant and modulating the gate voltage U_g, while monitoring the current I. (b) Response I vs. U_g for the electrometer, measured at various bias voltages V, at a temperature of 20 mK. The point O indicates the optimal bias values for U_g and V.

The Electrometer

The single-electron device which can be used as an ultra-sensitive electrometer was first demonstrated by Fulton and Dolan[5]; this device is also known as a single-electron transistor. This device, shown in Fig. 3(a), consists of two series-connected junctions biased with a voltage V, with the metallic island formed between the two junctions also capacitively coupled through the capacitor C_g to a gate voltage U_g. The junctions have tunnel resistances $R_J >> R_K \approx 25.8\text{k}\Omega$ and capacitances C_J of the order of 1 fF; the gate capacitance C_g is typically about 0.1 fF. The detailed behaviour of this device may be worked out with equations similar to those for the single-junction box; we will simply sketch out how this device is used. If the current I flowing through the junctions is measured as a function of the gate voltage U_g for a number of different bias voltages V, one finds the family of curves shown in Fig. 3(b). The current is periodic in the charge q on the capacitor C_g with period e,

and the modulation amplitude of the current varies with voltage V, with the maximum variation $\Delta I \approx e/R_J C_J$ for charge variations $\Delta q = e/2$ appearing for bias voltages $V \approx e/2C_J$. The optimum bias point, with the greatest sensitivity $\Delta I/\Delta q$, is indicated by the point O in Fig. 3(b).

The usefulness of this device can immediately be seen from the characteristics shown in Fig. 3(b): small variations Δq in the charge on C_g give rise to easily measured variations in the current I flowing through the junctions. With standard room-temperature amplifiers, the limits on the noise in the measurement of I enable us to measure equivalent variations in the charge of about $10^{-4}\ e/\sqrt{\text{Hz}}$ at 1 kHz, when the electrometer is biased at the point O; in other words, if we average the current measurement for 1 second, we can measure a charge variation of $10^{-4}\ e$. The major difficulties in using such a device are its high input impedance (formed by the small gate capacitance C_g), and the resultingly large voltages developed at the gate: a charge of e on C_g produces a voltage of about 1 mV with respect to the electrometer island. Note that in standard applications, one usually couples the charge to be measured through a second gate capacitor C_g'; the capacitor C_g is used to bias the electrometer so that it is operating at the optimum bias point O.

Experimental Investigation of the Single Junction Box

The measurement of the single junction box was carried out by coupling an electrometer to the box with a coupling capacitance of about 80 aF; the circuit diagram is shown in Fig. 4(a). The electron box was made with two junctions connected in parallel, equivalent to a single junction with twice the capacitance and half the tunnel resistance; this was done so that the junctions could be tested prior to the measurement. The electron box and electrometer were fabricated on a Si chip using SEM lithography and shadow evaporation of Al, as discussed in Ref. (7), and the device was mounted inside a Cu shield on the mixing chamber of a dilution refrigerator; details of the measurement can be found in Ref. (8). The Al could be driven into the normal state with a 0.5 T superconducting magnet. The basic measurement was to bias the electrometer at its optimal point O, and then observe the current flowing through the electrometer as a function of the bias voltage U on the box. From the electrometer transfer function $I(q)$, the charge on the box island could be determined, and the results are plotted in Fig. 4(b) for both the normal and superconducting states of the device; also plotted are the theoretical predictions from Eqs. (1) and (2).

The sawtooth oscillations in the junction charge $<Q_J>$ in the normal state are fairly close to the predictions; the charge jumps by nearly e for every increase of e in the bias charge $C_S U$ (the jumps are not quite e due to the factor $C_J/(C_s+C_J)$ and thermal averaging). The discrepancy between the theory and experiment can be accounted for by taking a system temperature of 60 mK instead of the temperature of 20 mK indicated by the refrigerator thermometry; however, the source of this increased temperature is not understood. What is more surprising is the observation that the sawtooth oscillations in the superconducting state appear identical to those in the normal state: one would naïvely expect that the oscillations would be $2e$-periodic due to the Cooper pairing of the superconducting electrons. The e-periodicity implies that there is an unknown source of unpaired electrons in the superconding device.

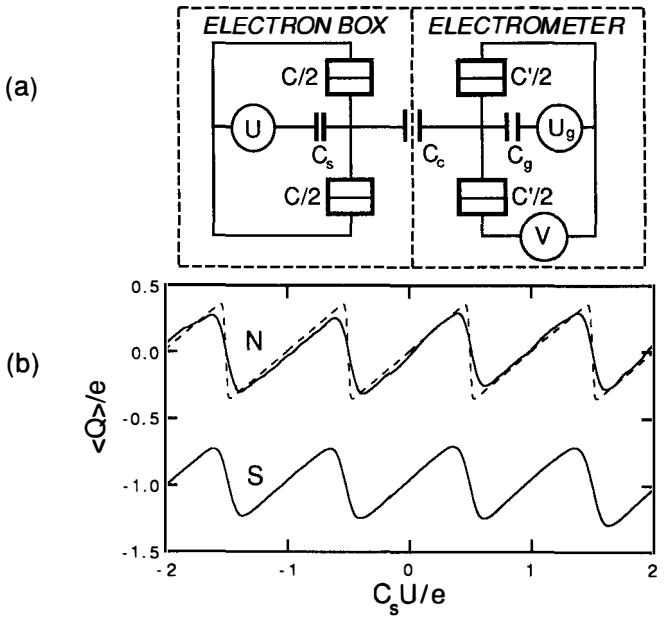


FIGURE 4. (a) Circuit diagram for the measurement of the charge on the electron box with a weakly coupled electrometer; the capacitor C_c has capacitance of 80 aF. (b) Average charge $<Q>$ on the island of the box as a function of the box bias U, in the normal (N) and superconducting (S) states. The theoretical prediction for the N curve is shown as a dotted line, calculated at the experimental temperature of 20 mK.

The Four Junction Box

The single junction box is appealing in its simplicity and in the ability to observe the change in the average charge by one electron on the island; however, the experiment does not measure the rate of tunneling of electrons into and out of the island, as given by Eq. (2), and it also does not allow the observation of the transfer of single electrons into and out of the island, only the change in their thermally-averaged number. In order to measure both of these quantities, the authors developed an experiment in which the single junction of the box was replaced by four series-connected junctions, as shown in Fig. 5(a). Each of the islands *a*, *b*, and *c* formed between two junctions was coupled through a capacitor $C_i = 80$ aF to a voltage source U_i, and the island *p* formed by the leftmost junction and the gate capacitor C_S was capacitively connected through $C_C = 240$ aF to the electrometer, as in the case of the single junction box. The role of the voltage sources U_i is simply to tune out the unavoidable backgrouund charges induced on the islands, and for the sake of this discussion may be ignored. The box state is then specified by the value of the bias voltage U and the number of electrons on each of the islands, $\{n_p,n_a,n_b,n_c\}$. Note that in the circuit diagram, the island *p* is coupled through a capacitance $C_S = 80$ aF to the voltage U, but also has a larger capacitance $C_p = 2$ fF to ground. The energy to add one electron to the island *p* is therefore significantly less than to add an electron to any of the islands *a*, *b*, or *c*. At the bias voltage $U = 0$, the stable state is that with no charge on any island, and as the voltage is increased, the state with one electron on the island *p* drops in energy until at $C_SU = e/2$, it has energy equal to the initial state. For values of U larger than this, the state with one electron on *p* has lower energy than the initial state and the states with one electron on *a*, *b*, or *c*. At this point it becomes possible for an electron to pass from the right side of the circuit to the island *p*, tunneling by way of the virtual states with one electron on each of the intermediate islands: in other words, the electron "co-tunnels" across the four junctions to the island *p*. The rate for the co-tunneling process was worked out by Averin and Odintsov[9]. For the experiment described here, for values of C_SU near $e/2$ the rate of co-tunneling can be well below 1 Hz, and individual tunnel events should therefore be easily observable.

The experimental details of the measurement are very similar to those for the single junction box, and will be published elsewhere[10]. In Fig. 5(b) we display the current through the electrometer as a function of time, with the bias voltage U across the box junctions held fixed. The

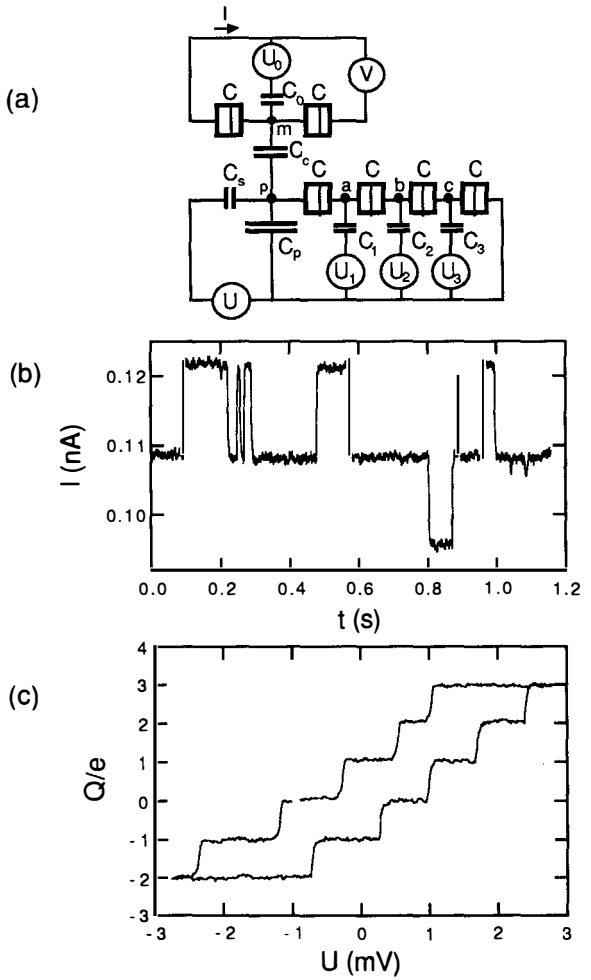


FIGURE 5. (a) Circuit diagram for the 4-junction box, consisting of four junctions connected in series, and an electrometer to measure the charge on the box island. (b) Measurement of the current I through the electrometer as a function of time; the telegraph signals are due to single electrons hopping on to and off of the island p. (c) Measurement of the charge Q on the box island as a function of the bias voltage U applied across the junctions.

sudden changes of the electrometer current are due to single electrons hopping across the four junctions to or from the island p. We have therefore succesfully observed the transfer of single charges to and from a metallic island. In Fig. 5(c) we show the charge Q on the island as a function of the bias voltage U; the voltage was swept through a triangle wave with a 0.4 s period, and the

hysteresis in the value of charge Q is due to the slow tunneling rate of electrons, which causes the actual state of the system to lag the stable state determined by the instantaneous value of U.

Detailed comparison of the tunnel rate predicted by the theory and that found in the experiment indicates a serious discrepancy: the measured rate scales roughly with bias voltage U as predicted by the theory, but the overall rate is 10^5 times higher than expected, even if the theory is calculated using a temperature of 50 mK as opposed to the refrigerator temperature of 20 mK. Adding additional radiofrequency filters, as well as injecting microwave noise into the leads connected to the experiment, had no effect on the tunneling rate. The discrepancy is not at present understood.

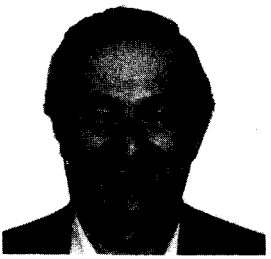
Measurements in the superconducting state were performed as well, and just as in the case of the single junction box, only single electron events were observed. A very similar experiment has been performed by Fulton *et al.*[11], in the superconducting state only, and their measurements yielded very similar results to our own. Despite the large discrepancy with the theory, this experiment has indicated that metrological measurements involving the counting of electrons could quite possibly be performed, with errors of the order of one mis-transferred electron per second.

References

1. R.A. Millikan, Phys. Rev. **32**, 349 (1911).
2. C.J. Gorter, Physica **17**, 777 (1951).
3. D.V. Averin and K.K. Likharev, J. Low Temp. Phys. **62**, 345 (1986).
4. W. Zwerger and M. Scharpf, Z. Phys. B **85**, 421 (1991).
5. U. Geigenmüller and G. Schön, Europhys. Lett. **10**, 765 (1989).
6. T.A. Fulton and G.J. Dolan, Phys. Rev. Lett. **59**, 109 (1987).
7. L.J. Geerligs, V.F. Anderegg, P. Holweg, J.E. Mooij, H. Pothier, D. Esteve, C. Urbina, M.H. Devoret, Phys. Rev. Lett. **64**, 2691 (1991).
8. P. Lafarge, H. Pothier, E.R. Williams, D. Esteve, C. Urbina, M.H. Devoret, Z. Phys. B **85**, 327 (1991).
9. D.V. Averin and A.A. Odintsov, Phys. Lett. **A140**, 251 (1989).
10. P. Lafarge, P. Joyez, H. Pothier, A.N. Cleland, T. Holst, D. Esteve, C. Urbina, M.H. Devoret, Compte Rendus de l'Academie des Sciences (France), to be published.
11. T.A. Fulton, P.L. Gammel, L.N. Dunklebereger, Phys. Rev. Lett. **67**, 3148 (1991).

FINITE NATURE

EDWARD FREDKIN
DEPARTMENT OF PHYSICS
BOSTON UNIVERSITY
BOSTON, MA, 02215, USA



ABSTRACT

A fundamental question about time, space and the inhabitants thereof is "Are things smooth or grainy?" Some things are obviously grainy (matter, charge, angular momentum); for other things (space, time, momentum, energy) the answers are not clear. Finite Nature is the assumption that, at some scale, space and time are discrete and that the number of possible states of every finite volume of space-time is finite. In other words Finite Nature assumes that there is no thing that is smooth or continuous and that there are no infinitesimals. If finite nature is true, then there are certain consequences that are independent of the scale.

Introduction

The development of science, from ancient times to the present, has been a series of nearly unbroken steps where one concept after another has moved out of the shadows of doubt and uncertainty and into the light of accepted scientific fact. The atomic hypothesis[1], whether matter is made up of atoms, is only one of many *atomic* hypotheses. So far every such question, discrete versus continuous, about a property of our world either remains undecided or it has been decided as discrete (atoms, electricity, light, angular momentum, etc.). It is hard to imagine the proof that some property will never admit to a finite description, no matter how fine grained. On the other hand, what is interesting is that so many concepts once thought of as continuous are now accepted as discrete. Finite Nature assumes that that historical process will continue to a logical conclusion where, at the bottom, everything will turn out to be atomic or discrete, including space and time.

The prime implication of Finite Nature is that every volume of space-time has a finite amount of information in it. Every small region of space-time (a cell) must be in one of a small number of states. If the future state of that cell is always a function of the space-time neighbors of the that cell, then we are lucky because there is a branch of mathematics, Automata Theory[2], that deals with such systems. Automata Theory is concerned with the behavior of all systems that evolve from one state (out of a finite number of states) to another, on the basis of a rule that takes into account: (1) the current state and (2) an input state. In the case of physics the input state comes from the states of other cells in the space-time neighborhood. There are versions of Automata called Cellular Automata[3] (CA) that are just what we're discussing.

Once we understand the rule of some CA, we name the states. If there are 3 states per cell, we can call them "1, 2 and 3" or "a, b and c"; it doesn't matter. Given any rule, we can find other rules, with other numbers of states per cell, that evolve in a manner isomorphic to the original system. Strangely enough, there is a wonderful chain of logic that proves that a system with just 2 states per cell, the Bank's[4] rule, with a 2-D neighborhood (4 neighbors; north, south, east and west), starting from an appropriate initial condition, can evolve in a way isomorphic to any other CA; no matter how many states per cell, no matter how high the dimensionality of the space, no matter how complex the rule, no matter how the neighborhood is defined! Of course the Bank's rule may take more cells and more time to do the simulation, but it can do it exactly because it's universal.

Automata, Cellular Automata and Finite State Machines are all forms of computers. A computer is universal if it can be programmed to simulate any other target computer. The simulating computer must have enough memory for two things: (1) to represent the state of the target computer and (2) to hold a program that can operate on that state to mimic the behavior of the target computer. Every ordinary computer is universal. The pioneers of automata theory were interested in what could be done by machines with arbitrarily large memories, so a common but unnecessary interpretation of "universal computer" is a computer with an infinite memory. We extend the idea to include computers with finite memory. A Macintosh, can *exactly* simulate the behavior of the fastest super computer if it has only a little bit more disk memory than the super computer. We pay a lot more money for a super computer not for what it can do, but for how *fast* it does it.

Given Finite Nature, what we have at the bottom is a Cellular Automaton of some kind. The first thing to wonder about is "Is it Universal?" The normal way to show that a computer is universal is to demonstrate a program that can simulate a computer known to be universal. The fact that the laws of physics allow us to build ordinary universal computers is proof of what must be a very fundamental law of physics: "The fundamental process of physics is computation universal."

Uncertainty is at the heart of quantum mechanics. Finite Nature requires that we rule out true, locally generated randomness because such numbers would not, in this context, be considered finite. The reason is that there is no way to create, within a computer, a truly random number that is orthogonal to everything in the computer. On the other hand, another kind of randomness appears in CA where the values of many of the bits are so influenced by distant events as to be essentially orthogonal to any local process. The deterministic nature of finite digital processes is different in that it is unknowable determinism. From within the system an observer will never be able to know very much about the true microscopic state of that system. Every part of space is computing its future as fast possible, while information pours in from every direction. The result is the same as caused by the apparent randomness of quantum mechanical processes.

What is Happening at the Bottom

We imagine that the cells that underlie physics are very small. Lets assume that the distances between such cells is somewhere between a Fermi, 10^{-13}cm, and Planck's length; $\sqrt{(\hbar G/c^3)} = 1.6 \times 10^{-33}$cm. If it's down near Planck's length, particles such as electrons, would be enormous entities consisting of structures spread out over volumes

containing more than 10^{60} bits. The momentum information would be spread out over volumes greater than 10^{75} bits. Because of superposition only a tiny fraction of the information in that volume would actually represent the momentum of the particle. While it is difficult to give strong arguments as to what the scale should be, many interesting consequences are independent of the scale. We call such informational models of physics "Digital Mechanics". At first glance it seems that Digital Mechanics[5] must be at odds with things we know about physics. However, efforts to reconcile the two make steady progress.

The utility of Digital Mechanics rests on the question of Finite Nature. If Finite Nature is true, then we may be able to derive QM from Digital Mechanics. If Finite Nature is false, then QM cannot be derived from Digital Mechanics. If Digital Mechanics ever becomes a successful model, it might allow for the development of a purely mechanistic, deterministic substrate for Quantum Mechanics.

The advantage of Digital Mechanics is that simple universal digital systems can do anything that can be done by any digital system. This is also the disadvantage. It is hard to think about the properties of the members of a class when each member can do everything. The field of Computer Science has very few examples of useful or meaningful analytic solutions as to what some digital system will or won't do. On the contrary, there is a celebrated proof that, in general, there are no analytical shortcuts that can tell the future state of some general computation any quicker than doing the computation step by step (this is the so called "halting problem" for Turing Machines[6]). There are normally no solutions in closed form. There is not yet any good hierarchy of concepts that express complex behavior in terms of simpler behavior, as is done in physics.

On the other hand, good programmers synthesize programs that are hierarchical. Programmers start with simple subroutines and build on them to create more complex programs and data structures. Its just that most any old computational structure can produce complex behavior that defies any analysis other than running the program to see what happens. A typical Macintosh, with an 80 megabyte hard drive is a universal computer with about $10^{200,000,000}$ states; it is clearly capable of some pretty complex misbehavior.

There are exceptions; for example the use of cellular automata for lattice gas models[7]. In this case very simple digital systems behave, in the limit, like systems easily modeled by differential equations.

The Finite State Machine

In order to better understand Finite Nature, we can look to examples of simple systems with similar properties. A digital computer, exclusive of its input and output facilities (with just a processor and memory), has many of the same properties that we find in Finite Nature. In particular, if we concentrate on a computer hardware technique called *single clock*, it will be easy to see why. *Single clock* means that when the computer is in a state, a single clock pulse initiates the transition to the next state. After the clock pulse happens, various changes propagate through the computer until things settle down again. Just after each clock pulse the machine lives for a short time in the messy real world of electrical transients and noise. At the points in time just before each clock pulse the computer corresponds exactly to the workings of a theoretical automaton or Finite State Machine.

The main point is that a computer is an embodiment of a system where Finite Nature is true (just before each clock pulse)[8]. Every such system has the property that it evolves by progressing from one state to the next state. Computers and all other such finite systems would eventually cycle around a closed trajectory if they ran long enough. Reversible systems are always on such a closed cycle, while an irreversible system can start down a path that will not repeat until finally entering a closed cycle. This is the informational viewpoint, as contrasted from the matter and energy viewpoint.

A computer needs to be able to perform three primitive functions: it must remember, communicate and compute. If we turn tables, systems of particles can model computers. An example is the Billiard Ball Model[9] of computation; an idealized 2-D billiard ball table with balls in motion that can exactly mimic the behavior of any computer. Particles remember by existing, communicate by moving and compute by interacting. A good understanding of what's happening reveals that communicating and remembering are the identical process (related by a space-time transformation) from the perspective of the physics involved.

INFORMATION

What is information? We know that "information" sometimes means "The meaning associated with data..." (Donald Knuth) but in this context we are referring to a scalar quantity that is a measure related to quantities like the $\log_2$(number of states) of some system. Finite Nature means that the world is made up of digital information. We understand "digital" exactly (made of, or making use of digits; symbols representing small integers) but we don't yet have a complete understanding as to how to measure or calculate with the scalar quantity named "information". This is not an easy problem, because

information, like energy or work, comes in different forms and thus it is hard to give a simple definition. Kinetic, potential, heat, electrical, chemical and others are all forms of energy; any of which can be transformed to other forms. Momentum, on the other hand is much simpler. There is really nothing else that momentum can be transformed into.

Despite not having a good definition of what information is, we postulate a new law; Conservation of Information. In this context information is the stuff that enables a reversible system to, if time were reversed, proceed backwards along the exact path that it took going forwards. All reversible systems have a conserved quantity; $\log_2(n)$, where n is the number of different states that the system progresses through in one cycle. (All finite state reversible systems evolve along a closed cycle.) This number is constant, independent of where the system is on that cycle and is equal to the quantity of information conserved.

Consider a reversible cellular automaton with a local neighborhood consisting of the seven cells which are east, west, north, south, up, down and past in relation to a particular cell. Finally, we have a transition rule, F, which defines for every possible combination of the states of the neighborhood cells, what state the particular cell should become.

$$C_{x,y,z,t+1}=F(C_{x-1,y,z,t}, C_{x+1,y,z,t}, C_{x,y+1,z,t}, C_{x,y-1,z,t}, C_{x,y,z+1,t}, C_{x,y,z-1,t}, C_{x,y,z,t-1})$$

For the system to be reversible, there must be another function, G, so that:

$$C_{x,y,z,t-1}=G(C_{x-1,y,z,t}, C_{x+1,y,z,t}, C_{x,y+1,z,t}, C_{x,y-1,z,t}, C_{x,y,z+1,t}, C_{x,y,z-1,t}, C_{x,y,z,t+1})$$

If these two criteria are met (the functions F and G exist), then we can say that the system conserves information, i.e., it is a law that information is conserved; no information ever gets lost. It may not be obvious as to what has happened to some information, but all of the information must be there in one form or another.

By observing the behavior of a non-trivial, reversible cellular automata, it becomes apparent that there is something almost magical about the way it evolves. What is hard to appreciate are the consequences of rules that appear to collapse information (a lot of information becoming just a little bit of information), yet nothing gets lost. If each cell has 3 states then 3^7 or 2187 possible states go into the computation of F. The resulting value of F is just one of three possible states. This seems to involve the loss of information; but each of the

neighbors is also involved in the computation of the states of other neighbors.

In ordinary physics, we can easily imagine perfect reversibility based on the fact that no effect ever gets lost because no matter how infinitesimal it becomes, it is still there. This approach makes it clear that the reason that an asteroid might hit the earth 5 years from now might be a direct consequence of the gravitational effect of a rock that rolled down a hill on the moon 2 billion years ago, delicately affecting the orbits of the Earth and all of the asteroids. But with the extreme quantization of the digital process, we have something more akin to the "For the want of a nail, the shoe was lost; for the want of the shoe, the horse was lost..." The results can be similar.

The Consequences of Finite Nature

The following is a list of some of the more obvious consequences of Finite Nature:

1. The *fundamental process* that underlies physics must be computation universal. Since the test of a computation universal process is that one can program up a universal machine within it, the existence of computers proves that the *fundamental process* is universal.
2. Things don't just happen; everything is a simple consequence of the *fundamental process.* Viewed from the informational perspective, the idea that a particle just has a momentum makes no sense. If the momentum is reasonably precise, meaning that it is one of a large number of possible momenta, then it contains a substantial amount of information, and that information must have a means of its representation.
3. The most fundamental stuff of physics is not the *cells* and *bits.* It is the result of the *fundamental process* run in the *cells* with the *bits.*
4. The future is computed as a function of the present and (most likely) the past. One way or another, it is very likely a second order system.
5. Matter has inertia because it takes some kind of process (like from a field) to change the momentum information (the information that governs the motion of the particle).
6. A particle accelerates when the environmental information, representing a field or a photon, interacts to change the information associated with the motion of the particle.
7. Information must have a means of its representation. If we had the right kind of magic microscope, we should

always be able to see the digits that represent whatever information is present. Information is never "just there".

8. What cannot be programmed cannot be physics. This is a very important idea. If a process cannot be programmed on a particular universal computer, despite having enough memory and time, then it cannot be programmed on any computer. If we can't program it on an ordinary computer, Finite Nature implies it can't be a part of physics because physics runs on a kind of computer.
9. An informational process cannot represent motion without a reference metric; finite nature demands the existence of a metric. This is a difficult idea. The implication is that the computational substrate of quantum mechanics must have access to some kind of metric in order to create inertial motion. Whether or not higher level processes in physics can access that metric is another question. It is hard to see why such access need be forbidden, since every particle must access the metric in order to move. Of course it is also clear that experiments done so far fail to detect any such metric. The argument for asserting the metric is that inertial motion requires an informational process that can take information that represents the velocity and use it to change the position of the particle. There must be information available that allows the process to know the velocity relative to something (the metric) that persists throughout the space that the particle will travel through. Without access to a constant, fixed space-time metric or its informational equivalent, there is no way to create an informational process that can involve particles moving independently in rectilinear motion
10. Energy, time, spin and other properties of the world cannot be programmed (made to work in an automata) without the existence of a metric.
11. Digital Mechanics is deterministic with unknowable determinism.
12. In general, physics is computing the future as fast as it can.
13. Some special cases (like the motion of the planets) can be computed faster than physics, allowing for predictions.
14. Most microscopic things, like quantum mechanics, cannot be computed on an ordinary computer faster than real time. From our perspective, the information bath that all events are immersed in will always be beyond our computing (in general). Thus, we call events dependent on that information "random".

15. The speed up theorem limits the detail with which we can know the future.
16. Physics is a consequence of the *fundamental bits* being engaged in the *fundamental informational process.*
17. The *Engine* is the computer that runs the *fundamental informational process* that runs physics.
18. The *Engine* and the *fundamental informational process* are computation universal.
19. Any universal *Engine* will do (given enough *memory*).
20. Physics is totally independent of the characteristics of the *engine*, so long as *it* is universal.
21. Everything mentioned in items 1 through 20 above that is in italics, is not a part of physics and is not to be found in this universe. Rather, it is stuff that exists somewhere, where, perhaps, the rules may be different. If we imagine using a computer to simulate air flow in a wind tunnel, we can think of that simulated airflow as a little universe following simple laws of physics. The computer is not made up of the air flowing in the simulated wind tunnel; the computer is not in the simulated wind tunnel, it is somewhere else.

Tasmania, or Physics out of Biology

On a recent trip to Tasmania, I visited a tree farm. The thousands of trees they planted every week were all identical clones. From a single cell of what was hoped to be a super tree, additional cells are grown in a tissue culture, and after a series of steps they have a forest of clones. It is one thing to contemplate the wonder that a little seed can carry the complete design of a tree, but it is another thing to contemplate that almost any part of the tree has the whole design of the tree in it! Imagine you catch someone trying to scrape of a few cells from his seat in a Boeing 747; with the idea of starting them off in a tissue culture and eventually growing 747 clones. What's wrong with the idea? It can work for trees and frogs! A floppy disk might contain the entire design of an integrated circuit memory chip; but why can't we make more memory chips by immersing the floppy in a nourishing solution instead of sending the floppy to a chip foundry?

The answers have to do with two laws about information; (1) "All information must have a digital means of its representation." We must be able to find in a single cell of the tree every bit of the information about the design of the tree. The second law involves process; (2) "An informational process transforms the digital representation of the state of a system into its future state." That a seed contains the design for a tree is not sufficient to transform the seed into the tree. The information in the seed must be acted upon and transformed like data

and programs are acted upon and transformed in the memory of a computer. The information in the seed is processed and transformed, along with other environmental information and material, in order to grow into the tree. The floppy may have all of the information about the memory chip, but it needs the environment of the chip foundry to be transformed into memory chips, in the same way that the seed needs fertile earth, sunshine, rain, air and time in order to be transformed into a tree.

Every physical system is subject to the same two laws!

---

[1] Dalton, John, "New System of Chemical Philosophy" (part I, 1808; part II, 1810)

[2] Minsky, Marvin, "Computation, Finite and Infinite Machines" (Prentice Hall, Englewood Cliffs, NJ, 1967)

[3] von Neuman, John, "Theory of Self-Reproducing Automata" (edited and completed by Arthur Burks), University of Illinois Press (1966)

[4] Banks, Edwin, "Information Processing and Transmission is Cellular Automata," Doctoral Dissertation and Technical Report MAC TR-81, MIT Project MAC (1971).

[5] Fredkin, Edward, "Digital Mechanics", Physica D, (1990) 254-270 North-Holland

[6] Turing, Alan, "On Computable Numbers with an Application to the Entscheidungsproblem", Proceedings London Math. Soc., series 2, 43 (1936), 544-546

[7] Frisch, Uriel, Brosl Hasslacher and Yves Pomeau, "Lattice-Gas Automata for the Navier-Stokes Equation" Phys. Review Letters 56 (1986), 1505-1508

[8] For a description of computation based on ideas from physics, see: Fredkin, Edward, "A Physicist's Model of Computation" Proceedings of the XXVIth Recontre de Moriond (1991) 283-297

[9] Fredkin, Edward, and Toffoli, Tommaso, "Conservative Logic," International Journal of Theoretical Physics 21 (1982), 219-253

ENVIRONMENT - INDUCED DECOHERENCE AND THE TRANSITION FROM QUANTUM TO CLASSICAL

Wojciech H. Zurek
Theoretical Astrophysics
T-6, Mail Stop B288
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

The environment surrounding a quantum system can, in effect, monitor some of the systems observables. As a result, the eigenstates of those observables continuously *decohere* and can behave like classical states, since their superpositions are rendered unobservable by the *effective superselection rule* induced by such interaction. This process is a key ingredient of the transition from quantum to classical. It suggests an explanation of the apparent reduction of the wavepacket in course of quantum measurements.

Introduction

Quantum mechanics works exceedingly well in all practical applications. No example of conflict between its predictions and the experiment is known at present. Without quantum physics we could not explain behavior of the solids, the structure and function of DNA, the color of the stars, the action of lasers or the properties of superfluids. Yet, well over half a century after its inception, the debate about the relation of quantum physics to the familiar physical world continues. How can it be that a theory which can account for everything we can measure with precision is still deemed lacking?

The only "failure" of quantum theory is its inability to provide a natural framework which could accommodate our prejudices about the workings of the Universe. States of quantum systems evolve according to the *deterministic, linear* Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi> = H|\psi> \ . \tag{1}$$

That is, just as in classical mechanics, given the initial state of the system and its hamiltonian H, one can, at least in principle, compute the state at an arbitrary time. This deterministic evolution of $|\psi>$ has been verified in carefully controlled experiments. Moreover, there is no indication of a border between quantum and classical at which Eq. (1) would fail. There is, however, a very poorly controlled experiment with results so tangible and immediate that it has an enormous power to convince: Our perceptions are often difficult to reconcile with the predictions of Eq. (1). Why? Given almost any initial condition the Universe described by $|\psi>$ evolves into a state in which many of the alternatives never seen to coexist in our world are simultaneously present. Moreover, while the ultimate evidence for the choice of one of the alternatives resides in the elusive "consciousness", there is every indication that the choice occurs much before the consciousness ever gets involved, and that, once made, it is irrevocable. Thus, at the root of our unease with the quantum is the clash between the principle of superposition — the consequence of the linearity of Eq. (1) — and the everyday classical reality in which this principle appears to be violated.

The problem of measurement has a long and fascinating history. A first widely accepted explanation of the emergence of the definiteness of a single outcome from the multitude of potentialities was proposed by Niels Bohr[1], who insisted that a classical apparatus is necessary to carry out measurements. Thus, quantum theory was not to be universal. The key feature of this approach (which has come to be known as the Copenhagen Interpretation, CI for short) was the dividing line between quantum and classical. Bohr has emphasized that the border must be mobile, so that even the "ultimate apparatus" — human nervous system — could be measured and analyzed as a quantum object, providing that a suitable classical device was found to carry out the task.

In the absence of a crisp criterion to distinguish between quantum and classical — to locate the border between the apparatus or observer and the measured system

— an identification of the "classical" with a "macroscopic" has been often tentatively accepted. Inadequacy of this approach has become apparent as a result of relatively recent developments: A cryogenic version of the Weber bar — gravity wave detector — must be treated as a quantum harmonic oscillator even though it can weigh a ton[3]. Non-classical squeezed states can also describe oscillations of suitably prepared electromagnetic fields with macroscopic numbers of photons[4]. Superconducting Josephson junctions have quantum states associated with currents involving macroscopic numbers of electrons, and yet they can tunnel between the minima of the effective potential[5].

If macroscopic cannot be always safely placed on the classical side of the boundary then may be there is no boundary at all? Many Worlds Interpretation is one of the names (or misnomers) of an approach to the measurement problem which claims to do away with the boundary[6]. MWI was developed in the fifties by Hugh Everett III with the encouragement of John Archibald Wheeler. All of the MWI Universe is described by quantum theory. Superpositions evolve forever according to the Schrödinger equation. Each time a suitable interaction takes place between any two quantum systems, the wavefunction of the Universe splits, developing ever more "branches."

Everett's work was initially almost unnoticed. It was taken out of the mothballs over a decade later by Bryce De Witt[7] to upgrade its status from virtually unknown to very controversial. MWI (or, rather Many Universes Interpretation) is a natural choice for a quantum cosmologist. Quantum cosmology describes the whole Universe by means of a state vector. There is nothing more macroscopic than the Universe. It can have no *a priori* classical subsystems. There can be no observer "on the outside". *Classicality* would have to be — at least in this context — an *emergent* property of the selected observables or systems.

At a first glance, the two interpretations — MWI and CI — have little in common. CI demands an *a priori* "classical domain" with a a border which enforces a classical "embargo" by letting through just one potential outcome. MWI aims to abolish the need for the border altogether. Every potential outcome is accommodated by the ever proliferating branches of the wavefunction of the Universe. The similarity of the difficulties faced by these two viewpoints becomes nevertheless apparent when the obvious question — "Why do I, the observer, perceive only one of the outcomes?" — is put forward. Quantum theory, with its freedom to rotate bases in the Hilbert space, does not even clearly define what set of states of the Universe corresponds to the "branches". Yet, our perception of reality with alternatives — not coherent superposition of alternatives — demands an explanation of when, where, and how the decision of what is actually perceived by the observer is made. In this sense, MWI in its original version does not abolish the border, but pushes it all the way to the boundary between the physical Universe and consciousness. Needless to say, this is a very uncomfortable place to do physics.

In spite of the profound nature of the difficulties and the relative lack of a break-

through for quite some time, recent years have seen a growing consensus that a progress in dealing with the measurement problem is being made. The key fact which was known almost since the inception of quantum theory, but its significance for the transition from quantum to classical is being recognized only now has to do with an uncontroversial remark: Apparatus and observer are never isolated from their environments. Therefore — as was emphasized already by H. Dieter Zeh[8] — they should not be expected to follow Schrödinger's equation, which is applicable only to a closed system. As a result — proponents of the "environment-induced decoherence" suggest[8-11] — systems usually regarded as classical, suffer (or benefit) from the natural loss of quantum coherence which "leaks out" into the environment. The resulting "decoherence" cannot be ignored in the discussion of the problem of the reduction of the wavepacket: It imposes, in effect, the required "embargo" on the potential outcomes by allowing the observer to maintain records of alternatives, to be aware of only one of the branches — one of the decoherent histories in the nomenclature of Murray Gell-Mann and James Hartle[11].

The aim of this paper is to explain the physics and thinking behind this approach. The reader should be warned that this writer is not a disinterested witness to this development, but rather, one of the proponents[9]. I shall, nevertheless, attempt to paint a fairly honest picture and focus on the difficulties and questions rather than on the accomplishments.

Correlations and Measurements

A convenient starting point for the discussion of the measurement problem, and, more generally, of the emergence of classical from quantum is the analysis of quantum measurements due to John von Neumann[12]. In contrast to Bohr, who has assumed at the outset that the apparatus must be classical (thus forfeiting claims of quantum theory to universality), von Neumann has analyzed the case of quantum apparatus. I shall reproduce his analysis for the simplest case: a measurement on a two-state system $\mathcal{S}$ (which can be thought of as spin $\frac{1}{2}$) with the result recorded by a quantum two - state (one bit) detector.

Hilbert space of the system $\mathcal{H}_\mathcal{S}$ shall be spanned by the orthonormal states $|\uparrow>, |\downarrow>$, while the states $|d_\uparrow>, |d_\downarrow>$ will spann $\mathcal{H}_\mathcal{D}$ of the detector. Two-dimensional $\mathcal{H}_\mathcal{D}$ is the absolute minimum needed to record the possible outcomes. One can devise a quantum detector (see Refs. 9 and 10) which "clicks" only when the spin is in the state $|\uparrow>$;

$$|\uparrow>|d_\downarrow> \rightarrow |\uparrow>|d_\uparrow> , \tag{2}$$

and remains unperturbed otherwise[8,13].

Before the interaction the system was — I shall assume — in a pure state $|\psi_\mathcal{S}>$;

$$|\psi_\mathcal{S}> = \alpha|\uparrow> + \beta|\downarrow> , \tag{3}$$

with the complex coefficients satisfying $|\alpha|^2 + |\beta|^2 = 1$. The composite system starts as:

$$|\Phi^i> = |\psi_\mathcal{S}>|d_\downarrow> . \tag{4}$$

Interaction results in the evolution of $|\Phi^i>$ into a correlated state $|\Phi^c>$:

$$|\Phi^i> = (\alpha|\uparrow> + \beta|\downarrow>)|d_\downarrow> \Longrightarrow \alpha|\uparrow>|d_\uparrow> + \beta|\downarrow>|d_\downarrow> = |\Phi^c> . \quad (5)$$

This essential and uncontroversial first stage of the measurement process can be accomplished by means of a Schrödinger equation with an appropriate interaction. It might be tempting to halt the discussion of measurements with Eq. (5). After all, the correlated state vector $|\Phi^c>$ implies that if the detector is seen in the state $|d_\uparrow>$, the system is guaranteed to be found in the state $|\uparrow>$. Why ask for anything more?

The reason for dissatisfaction with $|\Phi^c>$ as a description of a completed measurement is simple and fundamental: In the real world only one of the alternative outcomes of the measurement is left on the scene. And even when we do not know what was the outcome, we do know, at least, what are the alternatives, and we can safely act as if only one of them has occurred. As we shall see in the next section, this classically natural assumption cannot be justified for a system described by $|\Phi^c>$. If $|\Phi^c>$ is not good enough to describe an apparatus correlated with a measured quantum system, how else can an observer (who has not yet consulted the detector) express his ignorance about the outcome without giving up his certainty about the "menu" of the possibilities? Quantum theory provides a right formal tool for the occasion: *Density matrix* can be used to describe the probability distribution over the alternative outcomes.

Von Neumann was well aware of these difficulties. Indeed, he was willing to introduce *ad hoc* "process 1" which would take pure, correlated state $|\Phi^c>$ into an appropriate mixture: In addition to the unitary evolution given by Eq. (1) there should be — von Neumann postulated[12] — a nonunitary *reduction of the state vector* which makes the outcomes independent of one another by taking the pure-state density matrix;

$$\rho^c = |\Phi^c><\Phi^c| = |\alpha|^2|\uparrow><\uparrow||d_\uparrow><d_\uparrow| + \alpha\beta^*|\uparrow><\downarrow||d_\uparrow><d_\downarrow|$$

$$+ \alpha^*\beta|\downarrow><\uparrow|d_\downarrow><d_\uparrow| + |\beta|^2|\downarrow><\downarrow||d_\downarrow><d_\downarrow| , \quad (6)$$

and canceling the off-diagonal terms which express quantum correlations, so that reduced:

$$\rho^r = |\alpha|^2|\uparrow><\uparrow||d_\uparrow><d_\uparrow| + |\beta|^2|\downarrow><\downarrow||d_\downarrow><d_\downarrow| \quad (7)$$

emerges. Why is the reduced ρ^r easier to interpret as a description of a completed measurement than ρ^c? After all, both ρ^r and ρ^c contain identical diagonal elements. Therefore, both outcomes are still potentially present. So what — if anything — was gained at the substantial price of introducing a nonunitary "process 1"?

The Question of Preferred Basis: What Was Measured?

The key advantage of ρ_r over ρ_c is the possibility of a probability interpretation: The density matrix ρ^r can be used in its classical probability distribution role to describe

the alternative states of the composite spin-detector system with the *classical* correlations. Von Neumann's "process 1" and Bohr's "border" serve similar purpose, although "process 1" still leaves all of the alternatives. Nevertheless, when the off-diagonal terms are absent one can safely maintain that the apparatus as well as the system are each separately in a definite but unknown state, and that the correlation between them still exists in the preferred basis defined by the states appearing on the diagonal. By the same token two halves of a split coin are classically correlated. Thus, holding an unopened envelope containing one of them we can be sure that its state is either "heads" or "tails" (and not some superposition of the two), and that the second envelope contains the other, matching alternative.

By contrast, it is impossible to attach such a "classical ignorance" interpretation to ρ^c. In particular, the detector could not have yet even "decided" on the set of the alternatives! This can be illustrated in a dramatic fashion by choosing $\alpha = -\beta = \frac{1}{\sqrt{2}}$, so that the density matrix ρ^c is a projection operator constructed from the correlated state;

$$|\Phi^c> = (|\uparrow> |d_\uparrow> - |\downarrow> |d_\downarrow>)/\sqrt{2} \, . \tag{8}$$

This state is invariant under the rotations of the basis. For instance, instead of the eigenstates $|\uparrow>, |\downarrow>$ of $\hat{\sigma}_z$ one can rewrite $|\Phi^c>$ in terms of the eigenstates of $\hat{\sigma}_x$:

$$|\odot> = (|\uparrow> + |\downarrow>)/\sqrt{2} \, , \tag{9a}$$

$$|\otimes> = (|\uparrow> - |\downarrow>)/\sqrt{2} \, . \tag{9b}$$

This immediately yields:

$$|\Phi^c> = (|\odot> |d_\odot> \; - \; |\otimes> |d_\otimes>)/\sqrt{2} \, , \tag{10}$$

where;

$$|d_\odot> = (|d_\downarrow> - |d_\uparrow>)/\sqrt{2} \, , \tag{11a}$$

$$|d_\otimes> = (|d_\uparrow> + |d_\downarrow>)/\sqrt{2} \, , \tag{11b}$$

are perfectly "legal" states of the quantum detector. Therefore, the density matrix $\rho^c = |\Phi^c><\Phi^c|$ could have many different states of the subsystems on the diagonal.

This should not come as a surprise. State vector $|\Phi^c>$ is — save for a slightly different notation — same as the wavefunction of a pair of correlated spin $\frac{1}{2}$ systems in the Bohm's version of the Einstein-Podolsky-Rosen paradox[16]. And the key point of the experiments which show that nonseparable quantum correlations violate Bell's inequalities is to demonstrate that states of the two spins in a system described by $|\Phi^c>$ cannot exists — and not just are unknown — before the "real" measurement[17]. We conclude that when a detector is quantum, *a superposition of the records is a record of a superposition* of outcomes since the off-diagonal terms are present in the density matrix.

Missing Information and Decoherence

Unitary evolution condemns every closed quantum system to "purity." Yet, if the outcomes of a measurement are to become independent events, with consequences which can be explored separately, a way must be found to dispose of the excess information. In the previous sections quantum correlation was analyzed from the point of view of its role in acquiring information. Here I shall discuss the flip side of the story: Quantum correlations can also disperse information throughout the degrees of freedom which are, in effect, inaccessible to the observer. Interaction with the degrees of freedom external to the system — to which we shall summarily refer to as the environment — offers such a possibility.

Reduction of the state vector, $\rho^c \Longrightarrow \rho^r$, decreases the information available to the observer about the composite system $\mathcal{SD}$. This can be quantified by computing its entropy $S = -Tr\rho \lg \rho$ before and after reduction;

$$\Delta S \;=\; S(\rho^r) \;-\; S(\rho^c) \;=\; -(|\alpha|^2 \lg |\alpha|^2 \;+\; |\beta|^2 \lg |\beta|^2)\;, \qquad (12)$$

Entropy must increase, as the initial state described by ρ^c was pure, $S(\rho^c) = 0$, and the reduced state is mixed. Information gain — the objective of the measurement — is accomplished only when the observer interacts and becomes correlated with the detector in the already precollapsed state ρ^r. Information loss is needed if the outcomes are to become classical, to be used as initial conditions to predict the future.

To illustrate the process of the environment-induced decoherence consider a system $\mathcal{S}$, a detector $\mathcal{D}$, and an environment $\mathcal{E}$. The environment is also a quantum system. Following the first step of the measurement process — establishment of a correlation, Eq. (5) — the environment similarly interacts and becomes correlated with the apparatus:

$$|\Phi^c> |\varepsilon_0> = (\alpha|\uparrow> |d_\uparrow> \;+\; \beta|\downarrow> |d_\downarrow>)|\varepsilon_0>$$

$$\Longrightarrow\; \alpha|\uparrow> |d_\uparrow> |\varepsilon_\uparrow> \;+\; \beta|\downarrow> |d_\downarrow> |\varepsilon_\downarrow> = |\Psi> \;. \qquad (13)$$

The final state of the combined $\mathcal{SDE}$ "von Neumann chain" of correlated systems extends the correlation with the state of $\mathcal{S}$ beyond the $\mathcal{SD}$ pair. When the states of the environment $|\varepsilon_i>$ corresponding to states $|d_\uparrow>, |d_\downarrow>$ of the detector are orthogonal; $<\varepsilon_i|\varepsilon_{i'}> = \delta_{ii'}$, the detector-system combination is described by a density matrix obtained by ignoring the information in the uncontrolled (and unknown) degrees of freedom:

$$\rho_{\mathcal{DS}} \;=\; Tr_{\mathcal{E}}\, |\Psi><\Psi| \;=\; |\alpha|^2|\uparrow><\uparrow||d_\uparrow>< d_\uparrow| \;+\; |\beta|^2|\downarrow><\downarrow||d_\downarrow>< d_\downarrow|\;. \qquad (14)$$

This is precisely the reduced state density matrix von Neumann has called for. Now, *a superposition of the records* — of the states of the detector — *is no longer a record of a superposition.* A preferred basis of the detector — sometimes called the "pointer basis" for obvious reason — has emerged. Moreover, we have obtained it — or so it appears

— without having to appeal to anything beyond the ordinary, unitary Schrödinger evolution.

Preferred basis of the detector — or for that matter, of any open quantum system — will be selected by the dynamics. Not all aspects of this process are completely clear. It is, however, certain that the detector-environment interaction hamiltonian plays a decisive role. In particular, when the interaction with the environment dominates, eigenspaces of the observable Λ which commutes with the interaction hamiltonian;

$$[\Lambda, H_{int}] = 0 \tag{15}$$

invariably end up on the diagonal of the reduced density matrix[9]. This commutation relation has a simple physical interpretation: It guarantees that the *pointer observable* Λ will be a constant of motion of the interaction hamiltonian. Thus, when a system is in the eigenstate of Λ, interaction with the environment will leave it unperturbed.

In the real world spreading of the quantum correlations is practically inevitable. For example, if in the course of the experiment described above a photon had scattered from the spin carrying atom while it was traveling along one of its two alternative routes, this would have resulted in a correlation with the environment and would have necessarily led to a loss of quantum coherence. The density matrix of the $\mathcal{SD}$ pair would have lost its off-diagonal terms. Moreover, given that it is impossible to catch up with the photon, such loss of coherence would have been irreversible. As we shall see below, irreversibility could also arise from more familiar, statistical causes: Environments are notorious for having large numbers of interacting degrees of freedom, making extraction of the once lost information as difficult as reversing trajectories in the Boltzmann gas.

Reduction of the State Vector: How Long Does It Take?

A tractable model of the environment is afforded by a collection of harmonic oscillators, or, equivalently, by a quantum field. Excitations of the field scatter off the particle. The resulting "ripples" spreading away from the scattering site contain information about the location, shape, orientation, etc. of the scatterer. Moreover, the structure of the wavefronts excited in the process of scattering will depend on the instantaneous location of the scatterer: Therefore, the state of the field carries information about its trajectory.

A boat, traveling on a quiet lake, or a stone which fell into water will leave such an imprint on its surface. Our eyesight relies on the perturbation left by the objects on the preexisting state of the electromagnetic field. Hence, it is hardly surprising that an imprint is left whenever two quantum systems interact, even when "nobody is looking", and even when the lake is stormy and full of pre-existing waves, and the field is full of excitations — i. e., in equilibrium at some finite temperature. "Messy" initial states of the environment make it difficult to decipher the record, but do not interfere with its existence.

A specific example of decoherence — a harmonic oscillator interacting with a scalar field φ (which can be regarded as a collection of harmonic oscillators) through the hamiltonian:

$$H_{int} = \epsilon x \dot{\varphi} \ , \tag{16}$$

has been extensively studied[18–21]. The conclusion is particularly easily formulated in the so-called "high temperature limit," in which only the effect of thermal excitations of the field φ is taken into account, and the effect of zero-point vacuum fluctuations is neglected.

Density matrix $\rho(x,x')$ of the harmonic oscillator in the position representation evolves in this case according to the *master equation*:

$$\dot{\rho} = \overbrace{\underbrace{-\frac{i}{\hbar}[H_o,\rho]}_{\dot{p}=-FORCE=\nabla V}}^{von\ Neumann\ eq.} - \overbrace{\underbrace{\gamma(x-x')(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'})\rho}_{\dot{p}=-\gamma p}}^{relaxation} - \overbrace{\underbrace{\frac{2m\gamma k_B T}{\hbar^2}(x-x')^2\,\rho}_{classical\ phase\ space}}^{decoherence} \tag{17}$$

which can be obtained by first solving exactly the Schrödinger equation for the harmonic oscillator plus the field, and then tracing out the degrees of freedom of the field. Above, H_o is the harmonic oscillator hamiltonian, and γ is the relaxation rate.

I do not intend to enter into a detailed analysis of Eq. (17). It is, nevertheless, useful to write it down explicitly to demonstrate that it naturally separates into three distinct terms, each of them responsible for a different aspect of the transition from quantum to classical, as we have tried to indicate in Eq. (17) with the appropriate overbraces and underbraces. The first term — the von Neumann equation — generates reversible classical evolution of the *expectation value* of any observable with a classical counterpart regardless of the form of ρ (this is the Ehrenfest's theorem). The second term causes dissipation. Relaxation rate $\gamma = \eta/2m$ is proportional to the viscosity $\eta = \epsilon^2/2$ caused by the interaction. It causes the decrease of the average momentum and loss of energy. The last term also has a classical counterpart: It is responsible for fluctuations — random "kicks" imparted by the environment in Brownian motion. We shall see this in more detail in the next section.

For us, however, the effect of the last term on quantum superpositions will be of much greater interest. I shall show that it *destroys quantum coherence* — eliminates off-diagonal terms responsible for quantum correlations between *spatially separated* pieces of the wavepacket. It is therefore responsible for the classical structure of the phase space, as it converts superpositions into mixtures of localized wavepackets which, in the classical limit, turn into the familiar points in phase space. This effect is best illustrated by an example. Consider a coherent superposition of two gaussians, with the wavefunction $\chi(x) = (\chi^+(x) + \chi^-(x))/\sqrt{2}$ given by the sum of:

$$\chi^{\pm}(x) = <x|\pm> \sim \exp[-\frac{(x \pm \frac{\Delta x}{2})^2}{4\delta^2}] \ . \tag{18}$$

For the case of wide separation ($\Delta x >> \delta$) the corresponding density matrix $\rho(x,x') = \chi(x)\chi^*(x')$ has four peaks: Two on, and two off the diagonal (Fig. 5). Quantum coherence is due to the off-diagonal peaks for which x and x' are very different. With their disappearance position emerges as an approximate preferred basis.

The effect of the last term of Eq. (17) on the diagonal peaks is small: Along the diagonal $x = x'$, so that the last term, which is proportional to $(x - x')^2$, is small. By contrast, for the off-diagonal peaks $(x - x')^2 \approx (\Delta x)^2$ is the square of the separation, Fig. 4. Therefore, the off-diagonal terms decay at the rate[21]; $\frac{d}{dt}(\rho^{+-}) \sim 2\gamma m k_B T/\hbar^2(\Delta x)^2\rho^{+-} = \tau_D^{-1}\rho^{+-}$. It follows that the quantum coherence will disappear on a *decoherence timescale*;

$$\tau_D \simeq \gamma^{-1}(\frac{\lambda_{dB}}{\Delta x})^2 = \tau_R(\frac{\hbar}{\Delta x\sqrt{2mk_BT}})^2 , \qquad (19)$$

which, for macroscopic objects, is typically much less than the relaxation time $\tau_R = \gamma^{-1}$. Above, $\lambda_{dB} = \hbar/\sqrt{2mk_BT}$ is the thermal de Broglie wavelength. For a system with $m = 1g$, room temperature $T = 300K$ and $\Delta x = 1\ cm$ the ratio is $\tau_D/\tau_R \sim 10^{-40}$! Thus, even if the system was exceedingly well isolated so that its relaxation rate was equal to the age of the Universe $\sim 10^{17}s$, quantum coherence would have been destroyed in $\tau_D \sim 10^{-23}s$. For an electron ($m_e \sim 10^{-27}g$), and the ratio between the two timescales changes so that for atomic distances the two rates are comparable. Hence, macroscopic nature of the object is certainly crucial in facilitating the transition from quantum to classical.

Classical Limit of Quantum Dynamics

Schrödinger equation was deduced from classical mechanics in the Hamilton-Jacobi form. Thus, it is no surprise that it yields classical equations of motion when $\hbar$ can be regarded as small. This fact — along with the kinship of quantum commutators with the classical Poisson brackets, Ehrenfest's theorem, and Bohr's correspondence principle are all a part of the standard lore, to be found in the textbooks. Thus, the problem with establishing the quantum-classical correspondence has to do not with the equations of motion, but, rather, with the states. Quantum mechanics is formulated in Hilbert space, which can accommodate localized wavepackets with sensible classical limits as well as the most bizarre superpositions. By contrast, classical dynamics happens in phase space.

To facilitate the study of a transition from quantum to classical it is convenient to employ *Wigner transform* of a wavefunction $\psi(x)$;

$$W(x,p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{ipy/\hbar}\psi^*(x+\frac{y}{2})\psi(x-\frac{y}{2})dy \qquad (20)$$

Using $W(x,p)$ one can express quantum states as functions of position and momentum.

Wigner distribution is real, but it can be negative. Hence, it cannot be regarded as a probability distribution. Nevertheless, when integrated over one of the two variables,

it yields the probability distribution for the other (i. e. $\int W(x,p)dp = |\psi(x)|^2$). For a minimum uncertainty wavepacket,

$$\psi(x) \;=\; \pi^{-\frac{1}{4}}\delta^{-\frac{1}{2}}\exp\{-(x-x_0)^2/2\delta^2 + ip_0x/\hbar\}\;,$$

Wigner distribution function is a gaussian in both x and p:

$$W(x,p) = \frac{1}{\pi\hbar}\exp\{-\frac{(x-x_0)^2}{\delta^2} - \frac{(p-p_0)^2\delta^2}{\hbar^2}\}\;. \tag{21}$$

The system it describes is (in contrast to the other conceivable state vectors) localized in both x and p. Nothing else Hilbert space has to offer is closer to a point in the classical phase space. Indeed, one could formally take this limit by letting $\hbar$ and δ simultaneously to zero (although in our Universe with $\hbar \simeq 1.055 \times 10^{-27} erg\ s$ physical motivation for such a step is far from clear!).

Wigner distribution in the phase space is easily generalized to a mixture:

$$W(x,p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{-\infty} e^{ipy/\hbar}\rho(x-\frac{y}{2}, x+\frac{y}{2})dy\;, \tag{22}$$

where $\rho(x,x')$ will be here the reduced density matrix of the oscillator discussed above.

The phase space nature of the Wigner transform suggests a strategy in exhibiting classical behavior: Whenever $W(x,p)$ is a mixture of localized wavepackets (such as in Eq. (21)), it can be regarded as a classical probability distribution in the phase space. However, when the underlying state is truly quantum, its Wigner distribution function will have alternating sign. This property alone will make it impossible to regard it as a probability distribution in phase space. For the superposition of the two gaussians discussed in the last section:

$$W(x,p) \sim (W^+ \;+\; W^-)/2 \;+ \frac{1}{\pi\hbar}\exp\{-\frac{p^2\delta^2}{\hbar^2} - \frac{x^2}{\delta^2})\}\cdot\cos(\frac{\Delta x}{\hbar}p)\;, \tag{23}$$

where W^+ and W^- are Wigner transforms of χ^+ and χ^- of the previous section. An mixture of two gaussian wavepackets would have been described by the same two gaussians, but without the last oscillating term.

Equation of motion for $W(x,p)$ of a harmonic oscillator coupled to an environment can be obtained from Eq. (17) for $\rho(x,x')$:

$$\frac{\partial W}{\partial t} \;=\; \underbrace{-\frac{p}{m}\frac{\partial}{\partial x}W + \frac{\partial V}{\partial x}\frac{\partial}{\partial p}W}_{\textit{Liouville equation}} + \underbrace{2\gamma\frac{\partial}{\partial p}pW}_{\textit{friction}} + \underbrace{D\frac{\partial^2 W}{\partial p^2}}_{\textit{decoherence}}\;. \tag{24}$$

The three terms of this equation correspond to the three terms of Eq. (17). The last term results in the diffusion of $W(x,p)$ in momentum at the rate given by $D = 2m\gamma k_BT = \eta k_BT$.

The first term is easily identified as a classical Poisson bracket $\{H,W\}$. That is, a familiar classical probability density in the phase space $\varrho(x,p)$ evolves according to:

$$\frac{\partial \varrho}{\partial t} \;=\; -\frac{\partial \varrho}{\partial x}\frac{\partial H}{\partial p} \;+\; \frac{\partial \varrho}{\partial p}\frac{\partial H}{\partial x} \;=\; \{H,\varrho\} \;=\; L\varrho \tag{25}$$

where L stands for the Liouville operator. Thus, classical dynamics in its Liouville form follows from quantum dynamics at least for the harmonic oscillator case, for which Eqs. (17) and (24) have been rigorously derived.

Recovering classical equations of motion is a necessary but insufficient ingredient of the classical limit: We also must obtain the correct structure of the classical phase space by barring all but the probability distributions of well-localized wavepackets. The last term in Eq. (25) has precisely the right effect on non-classical $W(x,p)$. A superposition of two spatially separated pieces of the wavepacket results in a sinusoidal modulation of the Wigner distribution in the momentum coordinate. Oscillating term $\cos(\frac{\Delta x}{\hbar}p)$ is an eigenfunction of the diffusion operator $\frac{\partial^2}{\partial p^2}$ in the last term of Eq. (24). As a result, the modulation of W in will be washed out by diffusion at a rate $\tau_D^{-1} = -\dot{W}/W = (D\frac{\partial^2}{\partial p^2}W)/W = 2m\gamma k_B T(\Delta x)^2/\hbar^2$: Negative valleys of $W(x,p)$ will fill in on a timescale of order τ_D, Eq. (19). In the example described here $W(x,p)$ will retain just two peaks, which now correspond to the two classical alternatives. Superpositions of momenta will also decohere once the resulting difference in velocities has spread out the wave function.

The ratio of the decoherence and relaxation timescales depends on $\hbar^2/m$ (see Eq. (19)). Therefore, when m is large and $\hbar$ small, τ_D can be nearly zero — decoherence can be nearly instantaneous — while, at the same time, the motion of the classical, compact wavepackets (points in the phase space) in the smooth potential becomes reversible. This idealization is responsible for our confidence in classical mechanics, and, more generally, in the belief in classical reality.

The discussion above demonstrates that decoherence and the transition from quantum to classical (usually regarded as esoteric) is an inevitable consequence of the ubiquitous phenomenon of relaxation. True, our considerations were based on a fairly specific model — a harmonic oscillator in a heat bath of harmonic oscillators. However, this is often a reasonable approximate model for many more complicated systems. Moreover, our key conclusions — such as the relation between the decoherence and relaxation timescales, Eq. (19) — do not depend on any specific features of the model. Thus, one can hope that the viscosity and the resulting relaxation always imply decoherence, and that the transition from quantum to classical can be always expected to take place on a timescale of the order of the above estimates.

Quantum Theory of Classical Reality

Classical reality can be defined purely within the context of physics as a combination of classical laws with classical states of macroscopic objects. In the past few sections we have seen how this combination emerges from the substrate of quantum physics: Open quantum systems are forced into states described by localized wavepackets. They obey classical equations of motion, although with damping and fluctuations of possibly quantum origin. What else is there to explain?

The origin of the question about the interpretation of quantum physics can be traced to the clash between predictions of Eq. (1) and our perceptions. It is therefore useful to conclude this paper by revisiting the source of the problem — our awareness of definite outcomes. If these mental processes were essentially unphysical, there would be no hope of formulating and addressing the ultimate question (Why do we perceive just one of the quantum alternatives?) within the context of physics. Indeed, one might be tempted to follow Wigner in giving consciousness the last word in collapsing the state vector[23]. I shall assume the opposite. That is, I shall examine the idea that the higher mental processes all correspond to well-defined, but, at present, poorly understood information-processing functions which are being carried out by physical systems, our brains.

Described in this manner, awareness becomes susceptible to physical analysis. In particular, the process of decoherence we have described above is bound to affect the states of the brain: Relevant observables of individual neurons — including chemical concentrations and electrical potentials — are macroscopic. They obey classical, dissipative equations of motion. Thus, any quantum superposition of the states of neurons will be destroyed far too quickly for us to become conscious of the quantum "goings on". Environment-induced superselection applies to our own "state of mind".

One might still ask why the preferred basis of neurons becomes correlated with the classical observables in the familiar Universe. It would be, after all, so much easier to believe in quantum physics if we could train our senses to perceive nonclassical superpositions. The selection of the available interaction hamiltonians is limited, and must have constrained the choices of the detectable observables. There is, however, another process which must have played a decisive role: Our senses have evolved not for the purpose of verifying quantum mechanics. Rather, they have developed in the process in which "survival of the fittest" played a central role. There is no evolutionary reason for perception when nothing can be gained from prediction. And only classical states which are robust in spite of decoherence have predictable consequences. Indeed, classical reality can be regarded as nearly synonymous with predictability.

Natural sciences were built on a tacit assumption: Information about the Universe can be acquired without changing its state. The ideal of the "hard science" was to be *objective*, to provide a description of *reality*. Information was regarded as unphysical, a mere record of the tangible, material Universe, an inconsequential reflection, existing beyond and essentially decoupled from the domain governed by the laws of physics. Quantum theory has put an end to this Laplacean dream about a mechanical Universe. Observers of quantum phenomena can no longer be just passive spectators. Quantum laws make it impossible to gain information without changing the state of the measured object. The dividing line between what is and what is known to be has been blurred forever. While abolishing this boundary, quantum theory has simultaneously deprived the "conscious observer" of a monopoly on acquiring and storing information: Any correlation is a registration, any quantum state is a record of some other quantum state.

When correlations are robust enough — the record is sufficiently indelible — familiar classical "objective reality" emerges from the quantum substrate. Moreover, even a minute interaction with the environment, practically inevitable for any macroscopic object, will establish such a correlation: The environment will, in effect, measure the state of the object, and this suffices to destroy quantum coherence. Resulting decoherence plays, therefore, a vital role in facilitating the transition from quantum to classical.

Acknowledgment

I would like to thank John Archibald Wheeler for many inspiring and enjoyable discussions on "the quantum", Murray Gell-Mann and Jim Hartle for useful exchanges of ideas on the subject of decoherence, and, in particular, on the consistent histories[11,15] approach to the interpretation of quantum mechanics.

This paper has evolved from an early version of a manuscript which — after revisions — has appeared in *Physics Today*.[24] The main difference between them is in the level of presentation. *Physics Today* version has fewer equations and several figures — it is more "popular". My contribution is included in these proceedings with the kind permission of AIP.

References

1. N. Bohr, *Nature* **121**, 580 (1928), reprinted in Ref. 2.
2. J. A. Wheeler and W. H. Zurek, eds., *Quantum Theory and Measurements* (Princeton University Press, Princeton, 1983).
3. Braginsky, V. B., Vorontsov, Y. I., and Thorne, K. S., *Science* **209**, 547-557 (1980), reprinted in Ref. 2; Caves, C. M., Thorne, D. S., Drewer, R. W. P., Sandberg, V. D., and Zimmerman, M., *Rev. Mod. Phys.* **52**, 341-392 (1980).
4. Teich, M. C. and Saleh, B. E. A., *Physics Today* **43** (Jun.) 26 (1990).
5. Leggett, A. J. et al., *Rev. Mod. Phys.* **59**, 1 (1987); Caldeira, A. O. and Leggett, A. J., *Ann. Phys.* (N.Y.) **143** (1983); Tesche, C. D. *Ann. N. Y. Acad. Sci.* **480**, 36 (1986).
6. Everett, H., III, *Rev. Mod. Phys.* **29**, 454-462 (1957); J. A. Wheeler, *Rev. Mod. Phys.* **29**, 463 (1957).
7. DeWitt, B. S., *Physics Today* **23** (Sept.) 30-35 (1970); DeWitt, B. S. and Gráham, N., eds., *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, 1973).,
8. Zeh, H. D., *Found. Phys.* **1**, 69-76 (1970), reprinted in Ref. 2.
9. Zurek, W. H. *Phys. Rev.* **D24**, 1516 (1981); *Phys. Rev.* **D26**, 1862 (1982).
10. E. P. Wigner, p. 43 in *Quantum Optics, Experimental Gravitation, and the Measurement Theory*, Meystre, P., and Scully, M. O., eds. (Plenum Press, New York, 1983); Joos, E., and Zeh, H. D., *Zeits. Phys.* **B59**, 223-243 (1985); Haake, F. and Walls, D. F., in *Quantum Optics* IV, J. D. Harvey and D. F. Walls, eds. (Springer, Berlin 1986); G. J. Milburn and C. A. Holmes, *Phys. Rev. Lett.* **56**, 2237 (1986).

11. M. Gell-Mann and J. B. Hartle, p. 425-458 in *Complexity, Entropy, and the Physics of Information*, Zurek, W. H., ed. (Addison-Wesley, Redwood City, 1990); J. B. Hartle, in *Quantum Cosmology and Baby Universes*, Coleman, S., Hartle, J. B., Piran, T., and Weinberg, S., eds. (World Scientific, Singapore 1991).
12. von Neumann, J., *Mathematische Grundl. der Quantenmechanik* (Springer, Berlin 1932); English translation by R. T. Beyer, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton 1955), partly reprinted in Ref. 2.
13. Wigner, E. P. *Am. J. Phys.* **31**, 6 (1963) contains the original design of the "RSG"; see Scully, M. O., Englert, B. G., and Schwinger, J., *Phys. Rev.* **A40**, 1775 (1989), and references therein for recent discussion of "one bit detector".
14. Wheeler, J. A. (see, for example, p. 182 in Ref. 2) proposed "delayed choice experiments" in which a decision of what to measure has "... an unavoidable effect on what we have a right to say about the already past history...". For an experimental realization see Alley, C. O., Jakubowicz, O. G., and Wickes, W. C. in *Proc. 2nd Int. Symposium of Foundations of Quantum Mechanics*, M. Namiki et al., eds. (The Physical Society of Japan, Tokyo, 1987).
15. Griffith, R., *J. Stat. Phys.* **36** 219 (1984); see Omnès, R., in *Annals of Physics* **201**, 354-447, (1990) as well as in *Rev. Mod. Phys.* **64**, 339 (1992) for an overview and additional references.
16. Einstein, A., Podolsky, B., and Rosen, N., *Phys. Rev* **47**, 777 (1935); Bohm, D., *Quantum Theory*, Sections 15-19, Chapter 22 (Prentice Hall, Englewood Cliffs, N.J., 1951), reprinted in Ref. 2.
17. Aspect, A., Grangier, P., and Rogier, G. *Phys. Rev. Lett.* **47**, 91 (1982); Aspect, A., Dalibard, J., and Rogier, G., *Phys. Rev. Lett.* **49**, 1804 (1982).
18. Feynman, R. P., and Vernon, F. L., *Annals of Physics* **24**, 118 (1963).
19. Dekker, H., *Phys. Rep.* **80**, 1 (1981).
20. Caldeira, A. O., and Leggett, A. J., *Physica* **121A**, 587 (1983); *Phys. Rev.* **A31** 1057 (1985).
21. Zurek, W. H., in *Frontiers of Nonequilibrium Statistical Physics*, Meystre, P., and Scully, M. O., eds. (Plenum, New York 1986); Unruh, W. G., and Zurek, W. H., *Phys. Rev.* **D40**, 1071 (1989).
22. Wigner, E. P., *Phys. Rev.* **40**, 749 (1932).
23. Wigner, E. P., in *The Scientist Speculates*, I. J. Good., ed., p. 284 (Heineman, London, 1961), reprinted in Ref. 2.
24. Zurek, W. H., *Physics Today* **44** (Oct.) 36-44 (1991).

Quantum Cryptography

Charles H. Bennett
IBM Research Division, T. J. Watson Research Center
Yorktown Heights, NY 10598, USA.

Quantum mechanics, through the use of non-orthogonal states and EPR-type correlations, makes possible a varitey of classically impossible information-processing feats of an adversarial or cryptologic nature. These include 1) the possibility in principle of making banknotes that are physically impossible to counterfeit, 2) a practical communications channel on which it is impossible even in principle to eavesdrop without a high risk of detection by the channel's legitimate users, and 3) a technique by which two users can make joint decisions based on separately-held private information while maintaining its confidentiality to the maximum extent possible.

The violation of Bell's inequality by EPR correlations is famous as a way in which quantum systems escape the limitations of classical statistics. In other respects, eg the uncertainty principle, quantum systems are subject to limitations a classical system would not have. These fundamental limitations, especially the impossiblity of reliably distinguishing among non-orthogonal states, in turn have positive consequences of some practical importance in the adversarial setting of quantum cryptography, where two or more users depend on quantum mechanics to prevent or limit the flow of information to one another.

Narrowly defined, cryptography is the art of mathematically transforming messages into an unintelligible form or *cryptogram* from which the original message can be recovered by applying the inverse transformation when the cryptogram reaches its intended destination. This is possible in general only if the sender and/or receiver possess some secret information or *key* beforehand, not shared by their adversaries, which makes the transformation from cryptogram to original message easy for the receiver to perform, but difficult or impossible for the adversary.

This role of key is illustrated by a perfectly secure but cumbersome cryptosystem called the *Vernam cipher* or *one-time pad*. To use this cipher the sender and receiver must have agreed beforehand on a key consisting of a secret random bit string as long as the message they subsequently wish to transmit. The sender then produces the cryptogram by taking the XOR (bitwise sum modulo 2) of the message and key. The receiver can then reconstruct the original message by taking the XOR of the cryptogram with the same secret key. It is easy to see that a Vernam cryptogram is utterly uncorrelated with its message, except for having the same length, and thus it effectively hides the meaning of the message from an adversary ignorant of the key. The prefect security of this scheme is undermined if the same key is ever used to encrypt two different messages, since the adversary could then obtain the XOR of the two messages by taking the XOR of the two cryptograms. In redundant languages such as English, the XOR of two messages typically contains enough information to reconstruct both messages (the term one-time pad derives from the practice of furnishing the key on a pad of paper, each sheet of which was supposed to be used once then carefully destroyed, if necessary by eating it)

Nowadays cryptography is understood to include other useful feats of adversarial information processing, such as *authentication*, *digital signatures*, and *discreet function evaluation*. Authentication techniques allow the sender of a message to certify to the receiver (eg by an accompanying check-sum that only the sender and receiver know how to calculate) that the message is genuine and has not been altered in transit. This is particularly useful for electronic funds transfers. Digital signatures are an especially strong form of authentication which, unlike an authenticating check-sum, cannote feasibly be forged even by the receiver of a message. They therefore enable the receiver to prove to a third party, eg a judge, that the message came from the claimed sender, and has not been altered in transit. Discreet function evaluation, which is usually called "oblivious circuit evaluation" in the literature, comprises techniques by which two parties, each with secret data they wish not to disclose, can nevertheless cooperate so as to learn the value of an agreed-on public function of their private data, while preserving its confidentiality to the maximum extent possible.

Traditional cryptographic approaches to all these problems are purely mathematical, and do not depend on the physical form of the cryptograms or other data involved.

They assume that digital data transmissions can be accurately monitored and digital documents accurately copied by anyone who has access to them. However, when data is encoded in non-orthogonal states of a quantum system, it is neither possible to make an exact copy ("clone") of the system [1], nor to gain even partial information about the system without disturbing it [2].

Perhaps the simplest example of a classically unrealizable informational effect using nonorthogonal states is Wiesner's quantum money[3]. This money would have no need of the warning printed on French money, of life imprisonment for counterfeiters, because the crime would be impossible to commit. A quantum banknote contains about 20 spin-$\frac{1}{2}$ nuclei or other two-state quantum systems each kept in in a field-free environment so that its spin direction does not change signficantly in time. When the banknote is manufactured, the spins are installed in a secret sequence of the four nonorthogonal states $\uparrow$, $\downarrow$, $\leftarrow$, and -$\rightarrow$, the choices being made independently and randomly for each spin of each banknote. Each quantum banknote also has a unique serial number printed on it in ordinary ink, and this number is carefully recorded, along with that banknote's secret list of installed spin states, in a book distributed to banks but not the general public. The σ_z eigenstates $\uparrow$ and $\downarrow$ comprise one orthonormal basis for the 2-dimensional Hilbert space and the σ_x eigenstates $\leftarrow$ and $\rightarrow$ comprise another, with the property that each state of one basis has equal-length projections into all states of the other basis. Wiesner called such bases *conjugate* by analogy to conjugate variables. States of an orthonormal basis can be reliably distinguished by an appropriate measurement, but if that measurement is applied to states another basis conjugate to the first, the measurement result is utterly random and all information about the state before measurement is lost. Wiesner showed that in a Hilbert space of $2^{(N-1)^2/2}$ dimensions it is possible to find N mutually conjugate bases; in particular in a 2-dimensional Hilbert space, three mutually conjugate bases are possible, for example the σ_z σ_x and σ_y eigenstates .

When a quantum banknote is presented at a bank, the bank consults the book to determine the list of installed spin states for that banknote, then verifies that all spins are still in their original states by measuring each spin in its correct basis, in other words by performing a measurement of σ_z on each original $\uparrow$ or $\downarrow$ spin, and a measurement of σ_x on each $\leftarrow$ or $\rightarrow$ spin. If all the spins pass this test, the money is accepted; otherwise it is rejected, and the police are notified.

A counterfeiter attempting to copy a quantum banknote faces the impossible task of accurately measuring the spins without knowing which of the four nonorthogonal states they are in. It can readily be shown [3] [4] that whatever measurements the counterfeiter makes on a good banknote's spins to guide his choice of spins to install on a counterfeit banknote, each resulting counterfeit spin will fail inspection with probability at least $\frac{1}{4}$. Thus the probability of an entire forged banknote passing inspection is $\frac{3}{4}^{20}$ or about 0.3 per cent.

Quantum bankontes, while perfectly sound in principle, are not practical, because it is infeasible to reduce environmental fields to the extent necessary to avoid having the installed spins decay too rapidly for the money to be useful even in a country with very rapid inflation. Of course other particles, such as polarized photons, could be used, but they are even harder to store than polarized nuclei.

Photons are much better suited for information transmission than for information

storage. This fact, obvious in retrospect, suggested the possibility of an eavesdrop-detecting communication channel, based on a dynamical analog of quantum banknotes. Such a channel might employ single photon states comprising two conjugate bases, for example the horizontal and vertical states of the *rectilinear basis* and the 45° and 135° states of the *diagonal basis*.

If the sender (henceforth "Alice") and receiver ("Bob") had agreed secretly beforehand that the quantum transmission would consisting of a particular random sequence of such photons, the receiver could verify that the correct sequence had been received, and therefore that little or no eavesdropping had taken place on the quantum transmission. This would be a direct analog of quantum banknotes, with Alice playing the role of the mint and Bob the role of the bank where the money is tested. Of course, in the context of a communications channel, it is rather pointless to use the channel to send a message which the receiver already knows. To make the channel useful, Alice and Bob agree on the exact polarizations to be sent only for *some* time slots (to test for eavesdropping), leaving the remaining slots free to transmit meaningful data that has not been agreed on beforehand. The slots used for the eavesdropping test would have to be a random subset unknown to the eavesdropper ("Eve"), who could otherwise simply refrain from eavesdropping at those times, and thereby escape detection. For each of the remaining slots, Alice and Bob would agree beforehand on a random one of the four polarizations to represent a 0 data bit, with the understanding that the opposite polarization of the same basis would represent a 1 bit. Bob would thus know what measurement to perform in each time slot to receive the data correctly.

At first this quantum scheme might seem to have no advantages over the plain Vernam cipher. Both require cumbersome prior agreement on a large body of secret information. Moreover, since the Vernam cipher already effectively hides the meaning of its messages, there would seem to be no further purpose served by guaranteeing that no one has eavesdropped on the unintelligible cryptograms. However, the ability to detect eavesdropping does confer a significant advantage: if Alice and Bob transmit a message and find that no eavesdropping has occurred, they may safely reuse the same pre-agreed secret information to transmit subsequent messages, and can continue doing so until they detect evidence of eavesdropping. This is in sharp contrast to the classical one-time pad, which cannot ever be reused, because it must always be assumed that the cryptogram has been intercepted by Eve. A more practical version of the eavesdrop-detecting channel, using dim coherent light flashes instead of single photons, and able to function despite a small amount of noise (eg photomultiplier dark counts) was described by Bennett, Brassard, Breidbart, and Wiesner in [20]. However, the major disadvantage still remains: the need for prior agreement on a body of secret information.

In [6] Bennett and Brassard overcame this disadvantage, by showing that provably secret communication between two parties who share no secret information initially is possible if the quantum channel is supplemented by a classical *public channel* such as newspapers, whose contents becomes known to everyone but cannot be suppressed or altered. The scheme of [6], called quantum key distribution (QKD), has its immediate goal not the transmission of a meaningful message, but rather of a random secret key, known by Alice and Bob but by no one else. Once this is achieved, Alice and Bob can of course use the key as a one-time pad to transmit a meaningful message of the same length through the public channel. We now describe how Alice and Bob, knowing no

secret information initially, can use the quantum and public channels together to agree on a shared random bit string, with the assurance that it is unknown to anyone else.

Alice begins by sending Bob a random sequence of the four kinds of polarized photons, and Bob performs random measurements on these, measuring the rectilinear polarization of some photons and the diagonal polarization of others. Thus Bob performs the wrong measurement half the time, obtaining a random result uncorrelated with what Alice sent. In the remaining instances (ignoring for the moment inefficiency or noise in his detectors) Bob obtains an accurate measurement of the state sent by Alice. Unfortunately, Bob doesn't know which of his bit positions are good and which he has spoiled by making the wrong measurement.

The next step is for Alice and Bob to engage in a circumspect discussion over the public channel which allows them to discard the spoiled data without disclosing any of the polarizations of the good data. Bob begins the discussion by telling Alice which measurement basis, rectilinear or diagonal, he used in each time slot. Alice replies by telling him which of these bases were correct, and finally Alice and Bob both discard the data from all the remaining time slots, in which Bob used the wrong basis. Once the wrong-basis measurements have been discarded, Alice and Bob are left with a random sequence of photon polarizations on which they ought to agree, if the quantum signal has not been disturbed in transit, eg by eavesdropping or jamming. For each of these photons, although the basis has been disclosed publicly, the polarization within that basis (eg whether a rectilinear photon was vertical or horizontal) has not been disclosed, and so should constitute one bit of shared secret information between Alice and Bob. This ostensibly shared information, which we shall call the "provisional key", can be expressed as a bit string by assigning the bit value 0 to horizontal or 45° photons, and 1 to vertical or 135° photons.

If the quantum transmission has been significantly disturbed by eavesdropping, there will be errors, causing Alice's and Bob's version of the provisional key to differ. Fortunately the public channel can be used to test for such differences, prompting Alice and Bob to reject the provisional key if differences are found. After such a failure, Alice and Bob must restart the whole procedure again from the beginning with a fresh quantum tranmission, in hopes that it will be found free of eavesdropping. On the other hand, if no differences are found, that is evidence that the quantum tranmission has not been subjected to significant eavesdropping attempts, and therefore that the provisional key is unknown to any third party. Alice and Bob then accept the provisional key as valid, and, when the need to send a secret message arises, use it as their one-time pad for transmit the message secretly over the public channel. Whenever they run short of key bits, Alice and Bob repeat the quantum key distribution scheme to replenish the supply.

In more detail, Alice and Bob test for differences in their provisional keys by using the public channel to agree on a random subset of the bit positions in the provisional key, then publicly disclosing the parity of this random subset. They then discard the first bit of the agreed-on subset, to neutralize the advantage Eve may have gained through overhearing its parity discussed on the public channel. It can readily be seen that if Alice's and Bob's versions differed in any way, their calculated random subset parities would disagree with probability $\frac{1}{2}$. Therefore, by repeating the random-subset parity test a modest number of times (say 20), using a different random subset each time, Alice and Bob will be assured with probability 99.9999% that they share the same key, at the

cost of making it 20 bits shorter. Conversely, if any of the 20 random subset parities disagrees, the two key versions must differ, and ought to be rejected.

A practical version of QKD has been described and an experimental prototype constructed [4] which uses dim ($\approx$ 0.1 photon) incoherent light flashes instead of single photons, and operates over a distance of 30 centimeters (since photons propagate in vacuo without loss or depolarization, this range could be extended arbitrarily, eg by sending a diffraction-limited beam through a straight pipe of appropriate dimensions). The practical implementations employ more sophisticated public transformations [7],[4] of the provisional key, and are able to arrive at a correspondingly reduced number of highly secret final key bits even when the provisional key has been somewhat corrupted by detector noise and somewhat compromised by a small amount of eavesdropping.

Quantum key distribution is remarkable because on the one hand, unlike quantum banknotes, it is implementable in practice, and on the other hand it solves a cryptographic problem for which there is in principle no classical solution. All classical techniques for key distribution between parties who share no secret information initially [8][9], are insecure in principle, since they can be broken by an adversary with sufficiently great computing power, and they may be insecure even in practice, if certain apparently hard mathematical problems (eg factoring large integers, or the discrete logarithm problem) turn out to be easier than suspected.

The security of quantum key distribution depends on the fact that Eve cannot exactly monitor the signal Bob receives from Alice through the quantum channel. Maurer[10] has recently proved a related and very general result, in the context of several parties listening to one another through classical noisy channels supplemented by public discussions. Under these circumstances, if Bob and Eve listen to Alice through separate channels whose noise is to some extent independent (or if Alice, Bob, and Eve listen to a common random source through three independent noisy channels), then Alice and Bob can agree on a key that is secret from Eve, even if Eve's channel is much less noisy than Bob's or Alice's. Thus secure key distribution is possible not only through quantum channels, but in other situations where, for one reason or another, it can be guaranteed that Eve has incomplete information on the signal passing between Alice and Bob.

An alternative implementation of quantum key distribution developed by Ekert [11] following a suggestion of Deutsch [12] uses Einstein-Podolsky-Rosen (EPR) pairs instead of single particles. The pair is produced by a source between Alice and Bob, so that Alice can measure one member of each pair and Bob the other member. Subsequently Alice and Bob engage in a public discussion of their measurement results (which violate Bell's inequality) to test for eavesdropping and, if none is found, derive a secret key. This approach is essentially equivalent to single-particle quantum key distribution, because when Alice measures her member of the EPR pair, that is equivalent to preparing the other member (on its way to Bob) in a random state corresponding to the result of her measurement. In the single particle version, Alice would simply have prepared a particle in that state and sent it to Bob. The relation between EPR and non-EPR realizations of quantum key distribution is discussed further in [2].

Although the most familiar example of the EPR effect involves two-particle states with non-classical spin correlations, it has been known for some time [13][14][15] that other two-particle states can be prepared, for example by parametric down-conversion,

which exhibit entirely analogous correlations in *phase* that can be observed interferometrically. Recently Franson [16] and Ekert et. al. [17] have pointed out that these correlations too can be used for key distribution. Bennett [18] discusses one-particle non-EPR interferometric analogs of these schemes.

Key distribution is a cryptographic task with a definite cold war flavor, a means for two clear allies, Alice and Bob, to defeat the eavesdroping attempts of an unambiguous adversary, Eve. In the arena of civil commerce and post-cold-war diplomacy, however, cryptography is often called upon to handle the more delicate situation in which two or more parties wish to cooperate to achieve some common informational goal even though they do not entirely trust one another.

The situation is nicely illustrated by a bank check, or its elecronic equivalent certified by a digital signature. The payer wishes to cooperate with the payee in transferring an agreed upon amount of money, but neither party entirely trusts the other. The digital signature gives the payee the assurance that the payer cannot later deny having signed the check, while the payer receives the assurance that the payee cannot alter the check to a larger amount. Quantum techniques have no known application to digital signatures, which continue to be based entirely on public key cryptography, and therefore could be broken by the discovery of a rapid algorithm for the underlying mathematical problem (eg factoring).

However, quantum techniques involving non-orthogonal states do have an application to the more general post-cold-war cryptographic task of discreet function evaluation, or oblivious circuit evaluation as it is commonly known. Here the goal is to allow Alice and/or Bob to evaluate an agreeed-upon function $f(x, y)$ of two private inputs, where Alice knows x and Bob knows y, while maintianing as much confidentiality as possible. In other words, when they are done, Alice should know no more about y than can be inferred from knowing her own input x and the value of the function $f(x, y)$, and similarly for Bob. Clearly this goal can be achieved by having Alice and Bob confide their inputs to a third party, who evaluates the function and returns the value f, but this requires that Alice and Bob trust the third party to calculate f correctly and never to reveal the secrets x and y later.

Discreet function evaluation can be important even for very simple functions. The so-called "dating problem" consists of revealing to both Alice and Bob the logical AND of their affections for each other, without unnecessary indiscreet revalations, such as allowing Bob to find out that Alice likes him when he doesn't like her. The need to discreetly evaluate more complex functions arises in commerce and government. The first technique for discreet evaluation of a general function $f(x, y)$ was introduced by Yao [19] in 1986, using public key cryptographic methods related to digital signatures, but of course such methods depend on unporven mathematical assumptions such as the difficulty of factoring.

Current quantum techniques for attacking the general problem are descended from a useless-seeming special case of discreet function evaluation solved in Wiesner's early paper [3], a scheme enabling Alice to multiplex two messages m_1 and m_2 into a single quantum transmission from which Bob can recover either message at his choosing, but not both, and Alice does not learn which message he received. This may be regarded as a discreet function evaluation in which Alice's secret $x = (m_1, m_2)$ is the pair of messages,

Bob's secret y is which one he wants (say the second, $y = 2$), and the value of the function $f(x, y) = m_y$, ie the chosen message m_2, is revealed to Bob, while Alice is left knowing no more than she did in the beginning.

Wiesner's, scheme, as subseqently improved by Breidbart[20] begins by having Alice expand each of her messages using an error-correcting code of sufficient redundancy to recover the original message even if the expanded message is received with some errors. Alice then represents the expanded form of m_1 by a stream of rectilinear photons (or more generally dim light pulses) and the expanded m_2 by a stream of diagonal pulses. However, rather than sending the two streams, she sends a stream of pulses polarized in the intermediate directions $22\frac{1}{2}^\circ$, $67\frac{1}{2}^\circ$, $112\frac{1}{2}^\circ$ and $157\frac{1}{2}^\circ$. For example if the first bit of expanded m_1 was 1, represented by a vertical pulse, and the first bit of expanded m_2 was 0, represented by a 45° diagonal pulse, the first bit of the multiplex would be transmitted as a $67\frac{1}{2}^\circ$ pulse. To recover message m_1 Bob uses reads the entire transmission in rectilinear bases, thereby receiving the expanded m_1 with 15 per cent (more precisely $\sin^2 22\frac{1}{2}^\circ$) errors, then uses the error-correcting code to correct the errors and thereby recover the original m_1 correctly. Similarly, to receive m_2 Bob reads the entire transmission in diagonal bases, and then corrects the errors to recover m_2, but in either case, reading one message destroys the other. If Bob greedily attempts to receive both messages, (eg by reading some bits rectilinearly and others diagonally) he only succeeds in making the error rate too high for the error-correcting code, so neither is recovered correctly.

After Wiesner's multiplex was invented, but before it was published, Even, et. al [21] introduced essentially the same idea (implemented by mathematical means depending on the difficulty of factoring) under the name "one-out-of-two oblivious transfer" as a component of a contract-signing protocol. Later Kilian [22] showed that any discreet function evaluation can be reduced to oblivious transfers, without any further assumptions. Thus any means, physical or mathematical, of achieving oblivious transter can be built into a protocol for the more useful goal of discreet function evaluation. Recent quantum oblivious transfer schemes [23] repair defects in Wiesner's original scheme, such as the ability of a cheating Bob to gain partial information about both messages, which would have prevented it from being used directly in protocols for discreet function evaluation. Crepeau and Kilian [24] have explored versions of oblivious transfer based on certifiable noise sources other than quantum mechanics.

Existing quantum oblivious transfer schemes suffer the theoretical weakness of assuming that Bob measures Alice's light pulses one at a time and combines the results of these measurements classically. They thus may be insecure against a coherent attack (utterly infeasible in practice), in which Bob would measure the entire sequence of pulses as a vector in a large Hilbert space spanning all possible transmissions. They also suffer from the practical weakness of rather inefficient mathematical reductions connecting oblivious transfer with discreet function evaluation, requiring many thousands of light pulses to be sent and received to implement a single instance of, say, the dating problem. However, the difficulty and expense of sending quantum transmissions over long distances, which is the chief practical disadvantage of quantum key distribution, is not a serious disadvantage here, since discreet function evaluation still makes sense for negotiating parties sitting in the same room. If its mathematical efficiency can be sufficiently improved, discreet function evaluation may turn out to be the most important application of quantum cryptography.

References

[1] W.K. Wootters and W.H. Zurek, "A single quantum cannot be cloned," *Nature*, **299**, 802-803 (1982).

[2] C.H. Bennett, G. Brassard, and N.D. Mermin *Phys. Rev. Lett.* **68**, 557-9 (1992).

[3] Wiesner, S., "Conjugate coding", manuscript written *circa* 1970, unpublished until it appeared in *Sigact News*, Vol. 15, no. 1, 1983, pp. 78 – 88.

[4] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, *Journal of Cryptology*, **5**, 3, (1992).

[5] Bennett, C. H., G. Brassard, S. Breidbart and S. Wiesner, "Quantum cryptography, or unforgeable subway tokens", *Advances in Cryptology: Proceedings of Crypto '82*, August 1982, Plenum Press, pp. 267 – 275.

[6] Bennett, C. H. and G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, Bangalore, India, December 1984, pp. 175 – 179.

[7] Bennett, C. H., G. Brassard and J.–M. Robert, "Privacy amplification by public discussion", *SIAM Journal on Computing*, Vol. 17, no. 2, April 1988, pp. 210 – 229.

[8] Diffie, W. and M.E. Hellman, "New Directions in Cryptography," *IEEE Trans. Info. Theory* **IT–22**, 644-654 (1976).

[9] Rivest, R.L., A. Shamir and L.E. Adelman, "A method for obtaining digital signatures and public key cryptosystems," *Commun. ACM* **21**, 120-126 (1978).

[10] Maurer, U., *Proc. 23'rd ACM Symposium on Theory of Computing*, p. 561, (1991).

[11] A. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).

[12] D. Deutsch, *Proc. Roy. Soc.* **A 400**, 97 (1985).

[13] R. Ghosh, and L. Mandel, Phys. Rev. Lett. **59**, 1903, (1987).

[14] J. D. Franson, Phys. Rev. Lett. **62**, 2205, (1989).

[15] M. A. Horne, A. Shimony, and A. Zeilinger, Phys. Rev. Lett. **62**, 2209 (1989).

[16] J. D. Franson private communication.

[17] A.K. Ekert, J. Rarity, P. Tapster (European Science Foundation Conference, Sept. 1991, Davos, Switzerland).

[18] C.H. Bennett, "Quantum Cryptography Using any Two Non-Orthogonal States," *Phys. Rev. Lett.* (to appear May 1992).

[19] Yao, Andrew C., "How to Generate and Exchange Secrets," *Proc. 27'th IEEE Symposium on Foundations of Computer Science*, 162-167 (1986).

[20] Bennett, C.H., Brassard, G., Breidbart, S. and Wiesner, S, *Advances in Cryptology—Proceedings of Crypto–82*, pp. 267-275, Plenum, New York (1983).

[21] Even, S., O. Goldreich and A. Lempel, "A randomized protocol for signing contracts", *Advances in Cryptology: Proceedings of Crypto '82*, August 1982, Plenum Press, pp. 205 – 210.

[22] Kilian, J., "Founding cryptography on oblivious transfer", *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, May 1988, pp. 20–31.

[23] Bennett, C. H., G. Brassard, C. Crépeau, and M.-H. Skubiszewska "Practical Quantum Oblivious Transfer" *Advances in Cryptology—Crypto-91 proceedings*, edited by J. Feigenbaum, Lecture Notes in Computer Science vol. 576, pp. 351-366 (Springer, Berlin Heidelberg, 1992).

[24] Crépeau, C. and J. Kilian, "Achieving oblivious transfer using weakened security assumptions", *Proceedings of 29th IEEE Symposium on the Foundations of Computer Science*, 42–52 (1988).

Tests of Special and General Relativity

TESTING GENERAL RELATIVITY WITH THE NEUTRAL KAON SYSTEM

G. Chardin
DAPNIA/SPP, Centre d'Etudes de Saclay, F-91191 Gif/Yvette, France

Abstract : The arguments favouring gravitation as the "Master Arrow of Time" are briefly reviewed and the possibility that CP violation observed in the neutral kaon system may be explained by a violation of the Equivalence Principle is discussed.

Introduction

"Quid est tempus ? Si nemo a me quaerat, scio; si quaerenti explicare velim, nescio !"
ST. AUGUSTINE, *"Confessions"*, Book XI.

When asked about the origin about time-asymmetry, most physicists mumble a few (almost indistinct) words like "entropy" or "Boltzmann theorem". Surprisingly, a remarkably small number of high-energy physicists seem to be aware of the developments of the last twenty years about time-asymmetry. Most of these developments [Bek73, Haw75, Pen79, Zeh89] point in the direction of gravitation as the "Master Arrow of Time" and it is therefore extremely surprising to note that the only evidence for a microscopic time-asymmetry —the so-called CP violation in the neutral kaon system— is considered to be unrelated to the other "arrows of time". Indeed, the very existence of the neutral kaon system has been used, on the contrary, as a means to constrain the difference of gravitational interaction between matter and antimatter [Goo61, Ken90].

Every ten years or so [Gol67, Gal74, Dav74, Pen79, Zeh89], an attempt is made to gather in a unified context the various arrows of time which can be observed or conjectured in Nature —the decay of the K°-meson, quantum-mechanical observations, general entropy increase, retardation of radiation, psychological time, cosmological expansion and the black holes versus white holes puzzle. Retrospectively, it appears that the timing of the reviews on time-asymmetry was rather ironical : for example, "The Nature of Time" [Gol67] is based on a workshop held in May-June 1963, a few months before the experimental discovery of CP violation [Chr64]. Similarly, "The Physics of Time-Asymmetry" [Dav74] misses the findings of Bekenstein [Bek73] and Hawking [Haw75, 76] on the fundamental relation of entropy and the Second Law of Thermodynamics with gravitation by a few months.

Under the demoralizing influence of Kobayashi and Maskawa, the "explanation" of CP violation in the standard model has been essentially reduced to the counting of free parameters in a 3x3 unitary matrix, and most authors have considered the neutral kaon system as irrelevant to the question of time asymmetry, with a few notable exceptions. The reader who would like to further explore the subject of time-asymmetry should refer to [Gol67], [Gal74], [Dav74], [Pen79], [Lan82] and [Zeh89]. Although largely self-consistent, this paper can also be understood as complementary to [Cha90] and [Cha92].

Gravitation and the Master Arrow of Time

"We have reached a remarkable conclusion. The origin of all thermodynamic irreversibility in the real universe depends ultimately on gravitation. Any gravitating universe that can exist and contains more than one type of interacting material must be asymmetric in time, both globally in its motion, and locally in its thermodynamics."
P.C.W. DAVIES, *"The Physics of Time-Asymmetry"*, 1974

What are the arguments for insisting that Gravitation is at the origin of the Master Arrow of Time ? The first argument comes from the structure of the solutions of General Relativity : massive bodies dissipate their energy while collapsing and this dissipation will only end when the massive body will have completely evaporated. The example is well-known of the satellite which increases its speed when being dragged by the atmosphere —a consequence of the surprising "minus" sign in the expression of the virial theorem— and it was realized by Chandrasekar that

sufficiently massive bodies will contract without limit, resulting in "black holes". Now, it may seem at first sight that we have found just the opposite of what we intended to demonstrate : the classical picture of a black hole provides us with an object where no dissipation is possible since even light is trapped in it. But it follows from the work of Bekenstein [Bek73], refined by Hawking [Haw74], that even black holes must evaporate by thermal emission —the so-called Hawking radiation— a process imposed by the consistency with quantum mechanics. The fate of very massive structures seems to be of little relevance to the future of most celestial bodies, but again quantum mechanics allows the tunneling, albeit incredibly slow, of any massive body into a black hole and its subsequent evaporation in thermal, structureless radiation. In this respect, gravitation has been known for twenty years to be linked with time-asymmetry, a relation exemplified by the generalized Second Law of Thermodynamics formulated by Bekenstein.

A second and more subtle argument is provided by the work of Penrose [Pen65] and Hawking [Haw67] : if we assume the validity of General Relativity, the existence of a past singularity in our universe is unavoidable and Penrose has conjectured through his "Weyl curvature hypothesis" [Pen79] that any past singularity —the Big Bang— is associated with a low gravitational entropy which Penrose associates with a vanishing Weyl curvature, whereas future singularities like black holes would be associated with a huge Weyl curvature and entropy. This brings us to an important comment concerning the CPT theorem : most theorists would not pay any attention to your pet theory if it violated CPT symmetry. But as should be well known, the possibility of demonstrating the CPT theorem depends on the topology of spacetime : the unavoidable existence of past singularities (assuming the validity of General Relativity) makes it very doubtful that the CPT theorem can be demonstrated without modification for gravitation. In other words, boundary conditions cannot be eliminated. This point can be expressed more dramatically using the following remark : the dimensionless quantities entering the quantization of electroweak and strong interactions are *local* quantities —the coupling constants; as is well known, no such local expression has been found in the case of gravitation (the theory is non renormalizable). Instead, the only dimensionless quantity is here a *non-local or global* quantity : entropy or information (the presence of the Boltzmann constant in the usual expression of entropy is irrelevant; its physical content is expressed in bits, and can be measured (at least in principle) through the reversible coupling of a black hole with a gas of photons contained in a reflecting box [Zur82]). It is then probable that in a most fundamental sense, the quantization of gravity has been expressed twenty years ago through the Bekenstein entropy relation : $S = M^2$.

Reexpressing the Good argument

"It is hard to believe that Nature is not, so to speak, 'trying to tell us something' through the results of this delicate and beautiful experiment, which has been confirmed several times."
R. PENROSE, in *"General Relativity, an Einstein Centenary Survey"*, 1979

In the following, we will be defending the provocative idea that CP violation as observed in the neutral kaon system can be entirely due to a violation of the Equivalence Principle. As a starting point, let us use Morrison's antigravity [Mor58], a gross violation of the equivalence principle where antimatter is assumed to "fall up", the total force on a static $e^+ e^-$ pair, e.g., being zero. Note that such a definition for antigravity cannot be valid for ultrarelativistic particles : a photon is its own antiparticle (if we disregard its polarisation) and is known [Pou60; Ves76] to follow the

equivalence principle. We will address the question of the relativistic expression of antigravity when discussing the compatibility with the Second Law.

It is "well known" that antigravity violates the sacrosanct CPT symmetry, contradicts the results of the Eötvös-Dicke-Panov experiments, excludes the existence of the long-lived component in the neutral kaon system in the presence of the Earth gravitational field, violates energy conservation and implies vacuum instability. Except for the last one, these arguments have been reviewed critically by Nieto and Goldman [Nie91] and, although we disagree with the solutions proposed by these authors, we refer the reader to their recent review for a critical discussion of these impossibility arguments (see also [Cha90] and [Cha92]). Here, we will only insist on the Good argument and on the apparent violation of energy conservation.

In 1961, *three years before the first experimental observation of CP violation,* Good [Goo61] observed that antigravity would impose that the K_L, a linear combination of K° and $\overline{K}^\circ$, would regenerate a K_S component. Good estimated the phase shift which would develop between the K° and $\overline{K}^\circ$ components from the energy difference due to the gravitational potential ϕ_G; he supposed that the phase factor between the two components would oscillate as

$$\exp(2i\, m_K \phi_G\, t/\hbar).$$

Since the potential energy of a kaon in the Earth gravitational field is ≈0.4 eV, which is 10^5 larger than the energy splitting between the K_S and K_L eigenstates, Good concluded that antigravity was excluded. But, as Good himself had noticed, there is no obvious reason why one should use the Earth potential; why not use instead the Sun, or the Galactic potential which would give even more stringent limits on the difference of acceleration between matter and antimatter ?

In fact, gauge theories teach us that potential themselves are not observable, but only potential differences. Let us then try to restate the Good argument independently of absolute potentials. A kaon has a rest energy of ≈500 MeV, its size is then $\approx \hbar/m_K c \approx 0.4$ fm. Obviously, if there is a mechanism which separates the s quark from the $\overline{s}$ quark (and/or d from $\overline{d}$) by more than the size of the kaon during the mixing time $\Delta\tau$ introduced by weak interactions, which continuously transform $s \leftrightarrow \overline{s}$ (and $d \leftrightarrow \overline{d}$), $\Delta\tau \approx \pi\ (\hbar/\Delta mc^2)$, a large K_S amplitude (typically 0.5) will be regenerated. Assuming antigravity, let us estimate the time needed for the s quark to separate from the $\overline{s}$ to a distance that would induce a regeneration of the K_S component such as observed in CP violation. This time is given by the equation

$$gt^2 \approx \varepsilon \times (\textit{kaon size}) = \varepsilon \times \hbar/m_K c. \qquad (1)$$

where ε is the CP violating parameter.

This gives numerically

$$t \approx 3\ 10^{-10}\ \text{s} \approx 1.7\, \hbar/\Delta mc^2$$

In other words, the time needed for antigravity to generate the amount of regeneration observed in CP violation is just about equal to the mixing time imposed by weak interactions. Expressed independently of absolute potentials, the Good argument, far from excluding the possibility of antigravity, provides an indication that CP violation may be explained by some kind of antigravity ! From equation (1), we are then led to the following approximate expression for the ε parameter

$$\varepsilon = O(1)\, \hbar m_K g / \Delta m^2 c^3 = O(1) \times 0.88\ 10^{-3} \qquad (2)$$

which is the only dimensionless quantity linear in the kaon mass which can be built from the above quantities. This expression has a simple physical interpretation : it means that the energy to lift a kaon by $\Delta z = \hbar/\Delta mc$ is of the order of $\varepsilon\ \Delta mc^2$. It is difficult to believe that Good himself would not have proposed the previous expression for ε if CP violation had been observed at the time when he devised his argument. For a different expression of the argument leading to equation (2), see also [Cha90]. Fischbach [Fis80] had also noted that the quantity $\hbar m_K g\ /\Delta m^2 c^3$ is of the order of the experimental value of ε.

Vacuum instability and black hole radiation

"The information thus solicited makes physics and comes in bits. The count of bits drowned in the dark night of a black hole displays itself as horizon area, expressed in the language of the Bekenstein number"."
J.A. WHEELER, in *"Complexity, Entropy and the Physics of Information"*, 1990

As noted previously, a classical argument against Morrison's antigravity, devised by the author himself [Mor58], is that energy conservation must be violated by this process. But if, as we suggest, the behaviour of the neutral kaon system is linked to the failure of time invariance under (*continuous*) time translations (instead of the *discrete* symmetry usually considered), we may expect some problems with energy conservation : from Noether's theorem, a conserved quantity named energy can only be defined whenever invariance under time translations is respected. Any process defining an "arrow of time" may then forbid the construction of energy as a conserved quantity (instead, as we shall see, *entropy* will become the relevant quantity). Such an apparent violation of energy conservation is already present in the process of black-hole evaporation : while trying to find a static solution for a system including a black-hole Hawking [Haw74], to his surprise, discovered that no such solution existed; instead, the black hole appeared to behave like a grey body with a temperature $k_B T_{bh} = \hbar g/2\pi c$, where g is the surface gravity of the black hole and k_B is the Boltzmann constant. A similar example is provided by the Unruh effect [Unr76], related to the previous effect : a uniformly accelerated detector in vacuum measures a non zero temperature proportional to its acceleration, $k_B T_u = \hbar g/2\pi c$. Here also, it may seem at first sight that there is violation of energy conservation : a detector is excited *and* there is emission of photons.

The apparent violation of energy conservation by antigravity can be evidenced in the following way : if one had antigravity, the total force on an $e^+ e^-$ pair would be zero and it would be possible to carry adiabatically an $e^+ e^-$ pair at a higher altitude Δz at no cost. Annihilating the $e^+ e^-$ pair into two photons, we would propagate downwards the two photons to the lower altitude, and there recreate the $e^+ e^-$ pair. After the fall in the gravitational field, however, we know from the experiments of Pound and Rebka [Pou60] and Vessot and Levine [Ves76] that each of the two photons has gained an energy $mg\Delta z$.

It then seems that we have the possibility to realize a *cyclic* process resulting in the extraction of a photon from the gravitational field. Namely, an $e^+ e^-$ pair is annihilated at the altitude Δz, and the annihilation photons transported to the lower altitude where a pair is recreated with the emission of a photon of energy $2mg\Delta z$; the $e^+ e^-$ pair is then carried back, at zero energetic cost, at the altitude Δz. It then seems that we have obtained "something for nothing" and that we have realized the construction of a perpetual engine. Certainly, we should not expect to solve the energy crisis with such a process, but it appears that, even if there are no "real" particles

present, the vacuum becomes unstable in the presence of a gravitational field. Let us find the typical energy of photons created in virtual creation-annihilation processes in the vacuum. From the Heisenberg inequalities, a virtual creation-annihilation process will probe spacetime over a length scale $\Delta z \approx \hbar / mc$, where m is the mass of the propagated particle. Assuming antigravity, the typical energy of a photon which can be extracted from the gravitational field is then: $\Delta E \approx m\, g\, \Delta z \approx \hbar g / c$. The associated wavelength is $\lambda \approx c^2/g$, approximately one light-year at the surface of the Earth, and is therefore almost unobservable. The attentive reader will have noticed, however, that a similar expression of a "vacuum temperature" enters the Bekenstein-Hawking radiation of black holes. *The vacuum instability induced by antigravity then appears very similar to the very effect which introduces the second law and time-asymmetry in the realm of general relativity.* If black-hole evaporation is acceptable, why should we reject a priori a similar effect induced by antigravity ?

The thermal character of the radiation in the case of antigravity can be qualitatively justified by the following remark : Boyer [Boy84] had noted that if one applied "naively" the equivalence principle, there was an irreducible lowest temperature attainable for any experiment at the surface of the Earth. Since an object at rest on the Earth does not follow its geodesic path and is therefore accelerated, its temperature would be given by the Unruh expression $k_B T_u = \hbar g / 2\pi c$. At first glance, this seems to be a disaster since no really static situation would exist : not only black-holes, but *any* massive object would evaporate. The conventional viewpoint is that such an evaporation is impossible since energy must be conserved [Gri87] but the argument is circular since it presupposes that gravitation is not at the origin of time-asymmetry. Antigravity is just the tidal effect on the vacuum needed to induce the temperature suggested by the "naive" expression of the equivalence principle.

Information is Physical

"The law that entropy always increases holds, I think, the supreme position among the laws of Nature. If someone points out that your pet theory of the universe is in disagreement with Maxwell's equations – then so much the worse for Maxwell's equations. But if your theory is found to be against the second law of thermodynamics, I can give you no hope; there is nothing for it but to collapse in deepest humiliation."
A.S. EDDINGTON, *"The Nature of the Physical World"*, 1928

In the previous paragraph, we have considered the possibility of extracting photons from the gravitational field and the corresponding evaporation of any massive structure. It seems at first sight that the machine can be reversed and the process used to gradually remove photons initially present in a box. If it were possible to hide such photons at no cost, the Second Law of Thermodynamics would be violated. We want here to suggest that when the information cost of following the annihilation photons on curved spacetime is taken into account, the process cannot be used to violate the Second Law.

Imagine a Maxwell-Wheeler demon trying to violate the Second Law using antigravity. He (or she) intends to hide photons in the vacuum at the minimal cost. Certainly, he must realize his experiment in a box since otherwise he would loose the annihilation photons much too often. The Heisenberg inequalities impose then that the box be large enough to accommodate the wavelength of the photons that he is trying to remove. Calling Δz_{Heis} the minimal size of the box, and λ the photon wavelength, the demon must certainly choose a box such that :

$$\Delta z_{Heis} \geq \lambda = hc / 2mg\Delta z_{Heis}$$

or

$$\Delta z_{Heis} \geq \sqrt{\frac{hc}{2mg}} \approx r\sqrt{\frac{r_c}{r_s}} \qquad (3)$$

where r_s is the Schwarzschild radius.

The demon must then try to absorb the photons in the most economical way : if he measures too often the position and momentum of the annihilation photons during his attempt to rebuild the $e^+ e^-$ pair, the information cost of such a measurement will more than compensate the disappearance of a third spectator photon. Performing the proposed experiment in the most efficient way requires a *passive* focusing system such as an ellipsoid where the source of positronium would be at the lower focus and the $e^+ e^-$ pair would be reconstructed at the upper focus of the ellipsoid. It then seems that the demon can perform his work completely passively and at no cost. The demon should not forget, however, that he is working on curved spacetime. The amount of information which is needed to perform his experiment can be estimated to be given by the dynamical Kolmogorov-Sinai entropy, conjectured by Pesin [Pes77] to be related to the sum of the positive Lyapounov exponents of the system. The information cost is irreducibly positive if the system is chaotic. Since the orbits of photons in a Schwarzschild geometry are integrable, it seems that the system is not chaotic and that the information cost can be reduced to zero. Thus, it seems at first glance that the transport of photons can be ignored in the entropy balance. In fact, the demon is not dealing with orbits which are infinite but piecewise due to the reflexions needed to reconstruct the original $e^+ e^-$ pair. From the point of view of dynamical system theory, we are left with a billiard enclosing a volume with negative curvature.

The chaotic behaviour of the dynamical system that we are considering is not known but can be justified by the following qualitative argument : a small deviation from a billiard with a regular shape, without chaos, often leads to a chaotic billiard [Zas85]. Here, the billiard without entropy production is the ellipsoid and the perturbation is the curvature of the piece of Schwarzschild metric enclosed by the focusing billiard. Whereas the backfolding of trajectories is the ultimate source of stochasticity, the Lyapounov divergence will be dominated by the curvature. We can then guess that the typical path length defining the Lyapounov exponent for the annihilation photons is given by the Riemann curvature,

$$\Delta z_{Lyap} = O(1)\sqrt{\frac{c^2 r^3}{GM}} = O(1)\, r\sqrt{\frac{r}{r_s}} \qquad (4)$$

This parameter means that, given a semi-classical orbit of length l, the information cost that our demon will have to pay to follow the trajectory with a constant precision is approximately given by $l/\Delta z_{Lyap}$ (in bits).

It then appears that using an ellipsoid with the minimal size imposed by Heisenberg inequalities, of typical volume $(\Delta z_{Heis})^3$, our demon would have to pay an information cost (in bits) to eradicate one photon in the fundamental mode (approx. one bit)

$$\Delta z_{Heis}\, \frac{(\Delta z_{Heis})^3}{(\alpha r_c)^3} / \Delta z_{Lyap} = \frac{r^{5/2}}{r_c\, r_s^{3/2}\, \alpha^3} > 1 \qquad (6)$$

where r_c is the Compton radius, if one uses the two obvious conditions that $r \geq r_c$ and $r \geq r_s$.

Certainly, our argument is still unsatisfying : in particular, the notion of quantum chaos itself is only very partially understood, and our argument is semiclassical. There

are in addition a number of fundamental constraints that our demon must take into account : when trying to reset the system, he should care about the limited lifetime of the positronium $\tau_{pos} = 2r_c/c\alpha^5$, (for a disintegration in two photons) and, due to the Unruh effect, always manipulate it with an acceleration much less than

$$g_{max} = \frac{c^2}{\Delta z_{Heis}} \qquad (5)$$

The argument suggests, however :

— that the Second Law can be preserved when the dynamical information cost is taken into account

— and that the fundamental conserved quantity in the problem is entropy or information, a fundamental notion first stressed by Bekenstein [Bek72, 73].

Experimental tests

"Can the recently discovered irreversibility in the decay of a neutral kaon be linked with gravitational and cosmological asymmetries ?"
B. GAL-OR, in*"Modern Developments in Thermodynamics"*, 1974

The previous discussion suggests that we should reconsider the possibility of antigravity; they provide a new motivation for two types of experiments : firstly, direct measurements of the gravitational mass of antiparticles. The PS-200 experiment [Bev86], at the Low-Energy Antiproton Ring (LEAR) at CERN, currently in the stage of installation, proposes to measure, with a precision of the order of 1%, the difference between the gravitational masses of proton and antiproton. However, since the quarks inside the proton carry only a small fraction of its mass, the most clear-cut test of an explicit violation of the equivalence principle would be measuring the gravitational mass of the positron, a tremendously difficult measurement (for a discussion about the experimental problems raised by the measurement of the gravitational mass of antiparticles, see [Dar92]).

A second set of experiments involves the most precise interferometric system at our disposal : the neutral kaon system. It is interesting to note that *all the precision measurements of the ε parameter have involved ultrarelativistic horizontal kaon beams*, with one notable exception. The only experiment for which the angle θ with the horizontal is such that $\gamma \sin\theta$ is not negligible compared to 1 (or, equivalently, for which the energy difference associated with the change in altitude of a kaon during the mixing time of weak interactions is comparable to $\varepsilon\ \Delta m\ c^2$) has led to a determination of the η_{+-} parameter equal to $(2.09 \pm 0.02)\ 10^{-3}$, in disagreement by 9 standard deviations from the world average [Aro83]. In this last experiment, the neutral beam used made an angle of $6.25\ 10^{-3}$ rad with the horizontal, at a typical energy of 70 GeV (average $\gamma \sin\theta \approx 1$). It seems therefore crucial to test whether this non standard result is just due to a coincidence or whether the ε parameter really depends on the direction and momentum of the neutral kaon beam relative to the Earth. In particular, the PS-195 experiment [Adi85], also at LEAR, and currently taking data, should, for the first time, determine with a precision better than 1% the ε parameter for an isotropic kaon "beam" and using kaons in the momentum range [400, 700] MeV/c, generated from the annihilation at rest of a proton with an antiproton. An experiment to measure ε with an inclined kaon beam at Brookhaven is also being studied [Fis92].

Clearly, our hypothesis is equivalent to a regeneration effect and requires that the ε' parameter is zero. In addition, the ε parameter in the B system is predicted to be of the

order of 10^{-5}–10^{-6}, from the $m/\Delta m^2$ scaling. These two predictions differ from the Standard Model predictions and provide two further tests of our hypothesis.

Conclusions

In this paper, we have tried to demonstrate that the arguments against antigravity should be reconsidered and that the neutral kaon system, the most sensitive interferometric system at our disposal, is ideal to test the existence of antigravity. We have shown that the Good argument, when expressed in the most natural way, does not exclude antigravity; on the contrary, the Good argument shows that, for a kaon at rest, antigravity would provide just the amount of anomalous regeneration that we observe in the neutral kaon system and that we call CP violation.

Perhaps the most disturbing aspect of antigravity is its apparent violation of energy conservation. But we have stressed that an extremely similar thermal evaporation occurs in the Bekenstein-Hawking radiation of black holes. This leads us to consider that the key quantity when dealing with antigravity is entropy or information. Twenty years ago, the search for the compatibility with the Second Law of Thermodynamics led Bekenstein to the discovery of his relation between mass and entropy, and to the thermal radiation of black holes. In the same spirit, we propose to use the compatibility with the Second Law to establish the relativistic expression of the ε parameter measured in the kaon system, reflecting the divergence of trajectories between a particle and its antiparticle in a gravitational field.

Experimentally, twenty eight years after the experimental discovery of "CP" violation, the CP-LEAR experiment may provide the first precision measurement of the epsilon parameter using an isotropic kaon beam (in a limited momentum range, however). This measurement, together with the measurement of the gravitational mass of the antiproton, is an essential test of General Relativity.

— Acknowledgments : I am extremely grateful to J. Lo Secco for telling me, five years ago, that any undergraduate student knew that measuring the gravitational mass of the antiproton was stupid because a violation of the Equivalence Principle would imply a violation of energy conservation.

References

[Adi85] L. Adiels et al., CERN Report, CERN/PSCC/85-6 (1985).
[Aro83] S.H. Aronson, G.J. Bock, H.Y. Cheng and E. Fishbach, Phys. Rev D **28** (1983) 476, 495.
[Bek72] J.D. Bekenstein, Nuovo Cimento Lett. 4 (1972) 737.
[Bek73] J.D. Bekenstein, Phys. Rev. D 9 (1973) 3292; Physics Today **33** (1980) 24.
[Bev86] N. Beverini et al., CERN Report, CERN/PSCC/86-2 (1986).
[Boy84] T.H. Boyer, Phys. Rev. D **29** (1984) 1096; Sci. Am. **253** (1985).
[Cha90] G. Chardin, in: *CP Violation in Particle Physics and Astrophysics*, ed. J. Tran Thanh Van (Editions Frontières, Gif-sur-Yvette, 1990) p. 377.
[Cha92] G. Chardin and J.M. Rax, (1992) Phys. Lett.B, in press.
[Chr64] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, Phys. Rev. Lett. **13** (1964) 138.
[Dar92] T.W. Darling, F. Rossi, G.I. Opat and G.F. Moorhead, Rev. Mod. Phys. 64 (1992) 237-257.
[Dav74] P.C.W. Davies, *The Physics of Time-Asymmetry* (Surrey University Press, London, 1974).
[Fis80] E. Fischbach, in: Cosmology and Gravitation, eds. P. G. Bergmann and V. de Sabbata (Plenum, New York, 1980), p. 359.
[Fis92] E. Fischbach, private communication.
[Gal74] *"Modern Developments in Thermodynamics"*, ed. B. Gal'Or, (Israel University Press, Jerusalem, 1974)
[Gol67] *"The Nature of Time"*, ed. T. Gold, (Cornell University Press, Ithaca, 1967)
[Goo61] M.L. Good, Phys. Rev. **121** (1961) 311.
[Gri87] L.P. Grishchuk, Ya.B. Zel'dovitch, L.V. Rozhanskii, Sov. Phys. JETP 65(1) (1987) 11.
[Haw67] S.W. Hawking, Proc. R. Soc. Lond. A **300** (1967) 187; S.W. Hawking and R. Penrose, Proc. R. Soc. Lond., A **314** (1970) 529.
[Haw74] S.W. Hawking, Nature **248** (1974) 30; Comm. Math. Phys. **43** (1975) 199.
[Haw76] S.W. Hawking, Phys. Rev. D **13** (1976) 191.
[Ken90] I.R. Kenyon, Phys. Lett. B **237** (1990) 274.
[Lan74] *"The Enigma of Time"*, ed. P.T. Landsberg, (Adam Hilger, Bristol, 1982)
[Mor58] P. Morrison, Am. J. Phys. **26** (1958) 358.
[Nie91] M.M. Nieto and T. Goldman, Phys. Rep. **205**(5) (1991) 221.
[Pen65] R. Penrose, Phys. Rev. Lett. **14** (1965) 57
[Pen79] R. Penrose, in: *General Relativity, an Einstein Centenary Survey*, eds. S.W. Hawking and W. Israel, (Cambridge University Press, Cambridge, 1979) p. 281.
[Pes77]] Ya.B. Pesin, Russ. Math. Surveys **32** (1977) 55.
[Pou60] R.V. Pound and G.A. Rebka, Jr., Phys. Rev. Lett. 4 (1960) 337; R.V. Pound and J.L. Snider, Phys. Rev. **140** (1965) B788.
[Unr76] W.G. Unruh, Phys. Rev. D **14** (1976) 870.
[Ves76] R.F.C. Vessot and M.W. Levine, in: *Proc. of the 2nd Frequency Standards and Metrology Symposium*, ed. H. Hellwig, (1976) p.659.
[Zas85] G.M. Zaslavsky, *Chaos in dynamic systems*, (Harwood Academic Publishers, Chur, 1985).
[Zeh89] H.D. Zeh, *The Physical Basis of the Direction of Time*, (Springer Verlag, Berlin, 1989).
[Zur82] W.H. Zurek, Phys. Rev. Lett. 49 (1982) 1683

THE EXPONENTIAL POTENTIAL AND GALACTIC DYNAMICS or HOW TO DISPOSE OF THE MISSING MATTER MATTER

Donald H. Eckhardt
Phillips Laboratory
Hanscom AFB, MA 01731
USA

ABSTRACT

A hypothetical two parameter exponential potential explains all of the anomalous kinematics of galaxies and galactic clusters without need for the myriad *ad hoc* missing mass models currently in vogue. It also explains much about the scales and structures of galaxies and galactic clusters. It conforms with special relativity and does not conflict with Newtonian dynamics on the scale of the Solar System. There is circumstantial evidence that this force is somehow related to the weak force.

1. MOTIVATION FOR THIS STUDY

Using conventional physics, we can explain the sizes and shapes of stars but we cannot explain the sizes and shapes of galaxies. Is conventional physics enough to do this? Observed motions of stars in our Galaxy, both parallel and perpendicular to the Galactic plane, are inconsistent with observed and inferred mass distributions. Observed rotations of other galaxies are also inconsistent with independent mass estimates. Moreover, applying the virial theorem to galactic cluster data leads to mass estimates that are grossly inconsistent with estimates from mass to luminosity ratios.

Hypotheses to explain these inconsistencies are only hypotheses, and they are incomplete. Missing mass models that explain stellar motions perpendicular to the Galactic plane do not explain anomalously high rotations about the Galactic axis. And nobody has ever detected any missing mass. Modifying Newtonian gravity leads to contrived and awkward alternatives, MOND[1] (Modified Nonrelativistic Dynamics) being a prime example. Neither missing mass nor modified Newtonian gravity explains the sizes and shapes of galaxies and clusters.

One can hypothesize additional forces that are consistent with special relativity, distinct from gravitation, and only significant at galactic scales. These forces are applicable at galactic scales because the Compton wavelengths of their bosons have galactic scales. Pairing attractive and repulsive forces derived from Yukawa potentials that have identical coupling constants (except for sign) and slightly different boson masses results in a net force that can be derived from an exponential potential.[2] Because the coupling constants have identical magnitudes, the net force is negligible at less than galactic dimensions. The charge for these forces can be mass or something close to mass (baryon number, for example) so that the net force masquerades as a perturbation to gravitation.

2. THE EXPONENTIAL POTENTIAL

The massless graviton has spin two; its potential is

$$V_2 = -\frac{GM}{r} .$$

Hypothetical spin zero (attractive) and spin one (repulsive) forces are derived from the Yukawa potentials

$$V_0 = -\frac{\alpha GM}{r} \exp(-\mu_0 r)$$

and

$$V_1 = \frac{\alpha GM}{r} \exp(-\mu_1 r)$$

where α is a dimensionless coupling coefficient. Let $\mu_1 > \mu_0$, and define

$$\lambda^{-1} = \mu = \frac{\mu_0+\mu_1}{2},$$

$$\mu_1 = \mu + \delta\mu/2,$$

and

$$\mu_0 = \mu - \delta\mu/2.$$

The net potential for this hypothetical forces is

$$V_E = V_0 + V_1 = -\frac{\alpha GM}{r}\left[\exp[-(\mu-\delta\mu/2)r] - \exp[-(\mu+\delta\mu/2)r]\right] =$$

$$-\alpha GM \exp(-\mu r)\, 2\, \frac{\sinh(\delta\mu\, r/2)}{r} =$$

$$-\alpha\, \delta\mu\, GM \exp(-\mu r) \quad (\text{for } \delta\mu << 1).$$

Define the dimensionless parameter $\gamma = \alpha\, \delta\mu/\mu$. Then

$$V_E = -\gamma\mu\, GM \exp(-\mu r)\ .$$

For a point source, the (inward) specific forces are

$$\partial V_2/\partial r = GM/r^2 \text{ and } \partial V_E/\partial r = \gamma\mu^2\, GM \exp(-\mu r),$$

and their ratio is

$$\frac{\partial V_E/\partial r}{\partial V_2/\partial r} = \gamma(\mu r)^2 \exp(-\mu r)\ .$$

This ratio of specific forces could be greater or less than one, depending on the values of the parameters. The specific force gradients are

$$\partial^2 V_2/\partial r^2 = -2GM/r^3 \text{ and } \partial^2 V_E/\partial r^2 = -\gamma\mu^3 GM \exp(-\mu r),$$

and their ratio is

$$\frac{\partial^2 V_E/\partial r^2}{\partial^2 V_2/\partial r^2} = \frac{1}{2}\gamma(\mu r)^3 \exp(-\mu r) = \frac{1}{2}\mu r\, \frac{\partial V_E/\partial r}{\partial V_2/\partial r}\ .$$

This ratio of specific force gradients could also be greater or less than one. Nevertheless, in a vacuum surrounding an infinite flat sheet mass source at z = constant,

$$\partial^2 V_2/\partial z^2 = 0 \text{ and } \partial^2 V_E/\partial z^2 \neq 0 \text{ (in general)},$$

so the ratio $\partial^2 V_E/\partial z^2 / \partial^2 V_2/\partial z^2$ is infinite almost everywhere.

Constraints on the maximum value of $\partial V_E/\partial r / \partial V_2/\partial r$ come from laboratory experiments, Solar System kinematics and the tracking of deep space probes. The deep space probes present the tightest constraints. At r = 35 AU = 1.7×10^{-7} kpc where the gravitational acceleration due to the Sun is 500 μGal, the anomalous acceleration is less than 5 μGal.[3] For $\mu r << 1$, $\exp(-\mu r) \approx 1$ and

$$0.01 > \partial V_E/\partial r / \partial V_2/\partial r \approx \gamma(\mu r)^2 = 3 \times 10^{-14}\gamma\mu^2$$

where the unit for μ is kpc^{-1}. Then

$$\lambda > 1.7 \times 10^{-3}\sqrt{\gamma}\ \text{pc}.$$

If GM is set to unity, $\partial V_2/\partial r = r^{-2}$ and $\partial V_E/\partial r = \gamma\mu^2 \exp(-\mu r)$. Figure 1 compares $\partial V_2/\partial r$ and $\partial V_E/\partial r$ for $\mu = 1$ and various values of γ. The ordinate is $\log_{10}(\partial V/\partial r)$ and the abscissa is $\log_{10} r$. For $\gamma > 1.85$, there is a region in the vicinity of $r = \lambda$ where $\partial V_E/\partial r / \partial V_2/\partial r > 1$ and, therefore, the exponential force dominates; moving away from this region, either toward or away from the source, eventually $\partial V_E/\partial r / \partial V_2/\partial r < 1$ and the Newtonian force dominates.

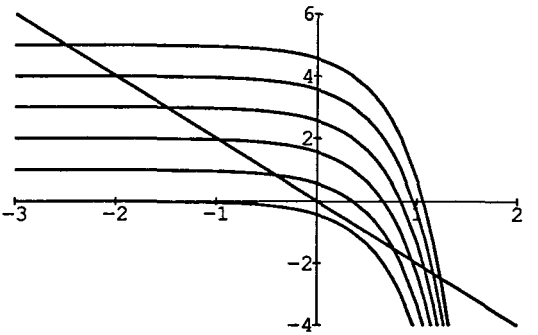


Figure 1: This log-log plot of $\partial V_E/\partial r$ (for $\mu = 1$ and $\gamma = 1$, 10, 100, 1000, 10000 & 100000) and of $\partial V_2/\partial r$ shows the region near $r = \lambda$ where $\partial V_E/\partial r > \partial V_2/\partial r$ and the exponential force dominates the Newtonian force.

Suppose that the exponential potential is commensurate with or greater than Newtonian gravitation over the range 20 kpc to 6 Mpc so that it plays an important role in the dynamics of galaxies and galactic clusters. Then $\partial V_E/\partial r$ exceeds $\partial V_2/\partial r$ over 2.5 orders of magnitude in range. From Figure 1, it can be seen that $\gamma \sim 1000$.

3. MOTION PERPENDICULAR TO THE GALACTIC PLANE IN THE VICINITY OF THE SUN

The Klein-Gordon and Poisson equations that govern the Yukawa and Newtonian potentials are

$$\nabla^2 V_0 = 4\pi\alpha G\rho + \mu_0^2 V_0,$$
$$\nabla^2 V_1 = -4\pi\alpha G\rho + \mu_1^2 V_0, \text{ and}$$
$$\nabla^2 V_2 = 4\pi G\rho .$$

Therefore

$$\nabla^2 V = \nabla^2[V_0+V_1+V_2] = 4\pi\alpha G\rho + \mu_0^2 V_0+\mu_1^2 V_1.$$

Because $V_0 < 0$, $V_1 > 0$, and $\mu_1 > \mu_2$, the contribution to the additional term due to a mass source is positive close to the source, the same as the density term.

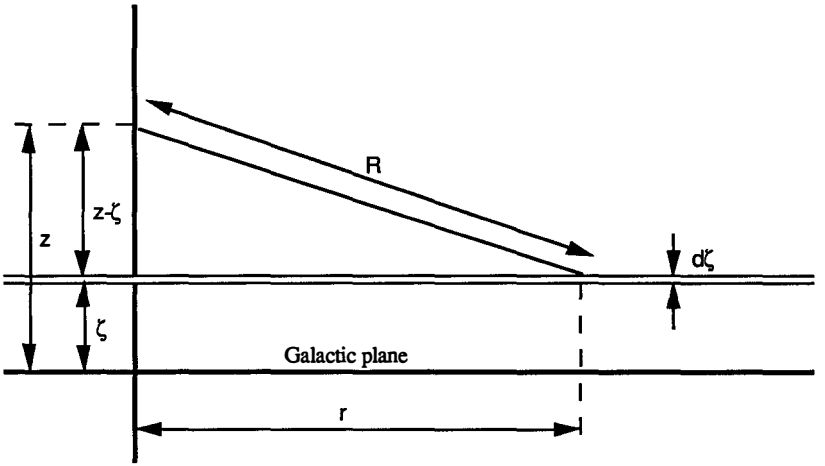


Figure 2: The term $\mu_0^2 V_0+\mu_1^2 V_1$ at distance z above the Galactic plane is calculated by integration, taking into account a normally distributed mass density model in the interval $-\infty \le \zeta \le \infty$.

In the conventional (r, ϑ, z) cylindrical coordinate system, let $R = \sqrt{r^2+(z-\zeta)^2}$. The geometry is depicted in Figure 2. A constant density flat layer of thickness $d\zeta$ at distance ζ "above" the Galactic plane perturbs the Poisson equation at distance z above the plane by

$$dV_E = -\alpha G\,\rho(\zeta)d\zeta \int_0^{2\pi}\int_0^{\infty} \exp[-\mu R]\, r\, dr\, d\vartheta =$$

$$-2\pi\alpha G\,\rho(\zeta)d\zeta \int_{z-\zeta}^{\infty} \exp[-\mu R]\, R\, dR =$$

$$2\pi\gamma G\,\rho(\zeta)d\zeta\,[\lambda + z - \zeta]\exp[-\mu(z-\zeta)].$$

Then

$$d[\mu_0^2 V_0(z)+\mu_1^2 V_1(z)] = d\,[d^2 V_E/dz^2] = 2\pi\,\gamma G\,\rho(\zeta)d\zeta\,\mu[1-\mu(z-\zeta)]\,\exp[-\mu(z-\zeta)]\,.$$

Let the density model be

$$\rho(\zeta) = \rho(0)\,\exp\left[-\tfrac{1}{2}(\zeta/\sigma)^2\right]$$

so that $\sqrt{2}\sigma$ is the scale height. Take $\sigma = 0.65/\sqrt{2}$ kpc = 0.46 kpc.[4]

For $z \geq \zeta$:

$$\rho(0)\,\exp\left[-\tfrac{1}{2}(\zeta/\sigma)^2\right]\,[1-\mu(z-\zeta)]\,\exp\left[-\mu(z-\zeta)\right] =$$

$$\rho(0)\,\exp\left[\tfrac{1}{2}(\mu\sigma)^2\right]\exp(-\mu z)\left[\,[1+(\mu\sigma)^2-\mu z]+\mu\sigma\left[\frac{\zeta-\mu\sigma^2}{\sigma}\right]\right]\exp\left[-\frac{1}{2}\left[\frac{\zeta-\mu\sigma^2}{\sigma}\right]^2\right].$$

Integrate this over ζ using

$$\int_{-\infty}^{z}\exp\left[-\frac{1}{2}\left[\frac{\zeta-\mu\sigma^2}{\sigma}\right]^2\right]d\zeta = \sigma\sqrt{2\pi}\,\Phi(z/\sigma-\mu\sigma)$$

where

$$\Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}\exp(-x^2/2)\,dx = \frac{1}{2}\left[1+\mathrm{erf}\,(u/\sqrt{2})\right]$$

is the normal distribution integral. Also use

$$\mu\sigma\int_{-\infty}^{z}\exp\left[-\frac{1}{2}\left[\frac{\zeta-\mu\sigma^2}{\sigma}\right]^2\right]\left[\frac{\zeta-\mu\sigma^2}{\sigma}\right]d\zeta = -\mu\sigma^2\exp\left[-\frac{1}{2}\left[\frac{z-\mu\sigma^2}{\sigma}\right]^2\right].$$

For $z \geq \zeta$:

$$\rho(0)\,\exp\left[-\tfrac{1}{2}(\zeta/\sigma)^2\right]\exp\left[-\mu(\zeta-z)\right] =$$

$$\rho(0)\,\exp\left[\tfrac{1}{2}(\mu\sigma)^2\right]\exp(\mu z)\left[[1-(\mu\sigma)^2-\mu z]+\mu\sigma\left[\frac{\zeta+\mu\sigma^2}{\sigma}\right]\right]\exp\left[-\frac{1}{2}\left[\frac{\zeta+\mu\sigma^2}{\sigma}\right]^2\right].$$

Integrate this over ζ using

$$\int_z^\infty \exp\left[-\frac{1}{2}\left[\frac{\zeta+\mu\sigma^2}{\sigma}\right]^2\right] d\zeta = \sigma\sqrt{2\pi}\,[1 - \Phi(z/\sigma+\mu\sigma)]$$

and

$$\mu\sigma\int_z^\infty \exp\left[-\frac{1}{2}\left[\frac{\zeta+\mu\sigma^2}{\sigma}\right]^2\right]\left[\frac{\zeta+\mu\sigma^2}{\sigma}\right] d\zeta = \mu\sigma^2 \exp\left[-\frac{1}{2}\left[\frac{z+\mu\sigma^2}{\sigma}\right]^2\right].$$

Some terms cancel out because

$$-\mu\sigma^2 \exp(-\mu z) \exp\left[-\frac{1}{2}\left[\frac{z-\mu\sigma^2}{\sigma}\right]^2\right] + \mu\sigma^2 \exp(\mu z) \exp\left[-\frac{1}{2}\left[\frac{z+\mu\sigma^2}{\sigma}\right]^2\right] = 0\,,$$

leaving as the final result

$$\mu_0^2 V_0(z) + \mu_1^2 V_1(z) = 4\pi G\rho(0)\, f(z, \mu, \gamma, \sigma)$$

where

$$f(z, \mu, \gamma, \sigma) = \gamma\mu\sigma\sqrt{\pi/2}\exp[\tfrac{1}{2}(\mu\sigma)^2]\big[[1 + (\mu\sigma)^2 - \mu z] \exp(-\mu z)\, \Phi(z/\sigma - \mu\sigma) +$$

$$[1 - (\mu\sigma)^2 - \mu z] \exp(\mu z)\, [1 - \Phi(z/\sigma + \mu\sigma)]\big]\,.$$

For small $\mu\sigma$, $\mu z = \mu\sigma \cdot z/\sigma$ is also small provided that z/σ is not large. Then

$$f(z, \mu, \gamma, \sigma) \approx \gamma\mu\sigma\sqrt{\pi/2}\left[\exp(-\mu\sigma \cdot z/\sigma)\, \Phi(z/\sigma) + \exp(\mu\sigma \cdot z/\sigma)\, [1 - \Phi(z/\sigma)]\right]$$

and

$$f(0, \mu, \gamma, \sigma) \approx \gamma\mu\sigma\sqrt{\pi/2}\,.$$

If $f(0, \mu, \gamma,\ 0.46\ \text{kpc}) = 1$, then this term would have the same effect on the gravity gradient as doubling the mass density on the Galactic plane (more or less making up for the supposedly missing mass). That is,

$$\left.\frac{dg(z)}{dz}\right|_{z=0} = \nabla^2 V(z)\Big|_{z=0} = 4\pi G\rho(0)\,[1 + f(0, \mu, \gamma, 0.46\ \text{kpc})] \rightarrow 8\pi G\rho(0).$$

[g(z) is defined as positive "downward."] This is satisfied if $\lambda = \gamma\sigma\sqrt{\pi/2} = 580\,\gamma$ pc.

4. ANOMALOUS ROTATIONS OF SPIRAL GALAXIES

The exponential potential V_E can be used to account for otherwise anomalous stellar motions perpendicular to the Galactic plane. Next it will be used to rationalize anomalously

high galactic rotation rates. To approach the problem for thin galaxies, consider a galaxy model in which the major fraction of the mass is in its nucleus and the remaining mass is distributed in an infinitesimally thin layer on the galactic plane in such a way that the density is a function only of the distance from the galactic center. The magnitude of the exponential potential decay range λ is much larger than galactic scale. Without resorting to missing mass, Newtonian gravity cannot by itself explain the flat rotation curves typical of spiral galaxies. At the distance r from the (point) mass M_N of the galactic nucleus, the inward specific force on a star due to V_E is $\gamma\mu^2 GM_N \exp(-\mu r) = \gamma\mu^2 GM_N$. If the star is in orbit about M_N, its acceleration equals the specific force; and if the orbit is a circle and the mass of the disk can be neglected, the star's velocity is $\sqrt{GM_N/r}\sqrt{1+\gamma(\mu r)^2}$. On comparing this with rotation curves of sample galaxies[5], $\sqrt{1+\gamma(\mu r)^2} = 2$ at $r = 30$ kpc. Suppose that $\lambda = 0.58\gamma$ kpc as estimated above. Then setting $\gamma(r/\lambda)^2 = 3$ gives $\gamma = 900$ and $\lambda = 520$ kpc. The mean mass of the exponential potential boson pair is $m = \hbar/\lambda c = 1.2 \times 10^{-38}$ GeV/c^2 which is ~ 0.7×10^{-40} times the mass of Z or W boson. The apparition of the reciprocal of Dirac's large dimensionless number adds some appeal to this model. Figure 3 is a plot of the V_2 and V_E forces for the model. The force due to the exponential potential exceeds that of gravity over the range 18 kpc < r < 6.1 Mpc.

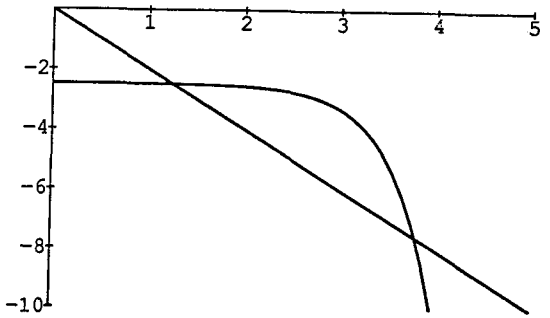


Figure 3: This log-log plot of $\partial V_E/\partial r$ (for $\lambda = 520$ kpc and $\gamma = 900$) and of $\partial V_2/\partial r$ shows the region where $\partial V_E/\partial r > \partial V_2/\partial r$. The abscissa is labeled with the logarithm of the range in kpc.

The next step in the model consists of allowing for the galactic disk. A method is presented below for calculating the effect of variable density concentric rings in the disk. Qualitative aspects of the disk solutions are discussed but more detailed modelling requires disk mass models, probably derived from neutral hydrogen and from mass/luminosity estimates. In general, the addition of the disk effects tends to supplement the effect of the exponential potential of the nucleus in flattening out the velocity distance curves.

Consider the exponential potential at a point that is at the distance r from the center of a thin wire-like circular mass source ring of radius a. The geometry is shown in Figure 4. Let the mass of the ring be $M = 2\pi a\rho_{1d}$ where ρ_{1d} is the 1-D density - the mass per unit length of the

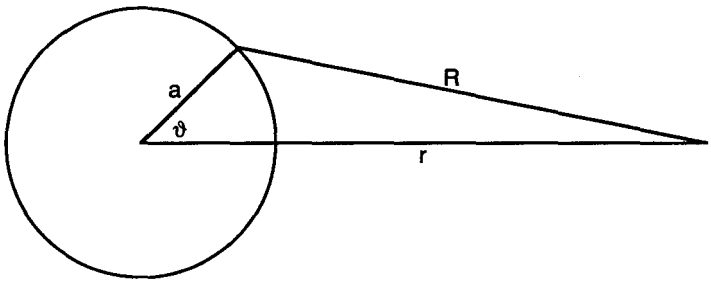


Figure 4: The mass source for the exponential potential is a thin uniform density circular ring of radius a. The potential is evaluated at distance r from the center of the ring. In general, r can be less than, equal to, or greater than a.

wire. The potential is

$$\rho_{1d}\,\gamma G\,\mu\,a\,P(r, a, \mu)$$

where

$$P(r, a, \mu) = -\int_0^{2\pi} \exp[-\mu R]\, d\vartheta$$

and

$$R = \sqrt{a^2 + r^2 - 2ar\cos\vartheta}.$$

Because $\lambda >> R$, the truncated expansion $\exp(-\mu R) = 1 - \mu R + \mu^2 R^2/2$ is entirely adequate for estimating the integral. Thus

$$P(r, a, \mu) = -2\pi + \mu\int_0^{2\pi} R\, d\vartheta - \frac{\mu^2}{2}\int_0^{2\pi} (a^2 + b^2 - 2ar\cos\vartheta)\, d\vartheta.$$

Use

$$\cos\vartheta = 2\cos^2(\vartheta/2) - 1 = 2\sin^2\varphi - 1$$

where $(\vartheta - \pi)/2 = \varphi$. Then

$$\int_0^{2\pi} R\, d\vartheta = 4\int_0^{\pi/2} \sqrt{(a + r)^2 - 4ar\sin^2\varphi}\, d\varphi = 4(a + r)\, E(\kappa)$$

where

$$\kappa = \frac{4ar}{(a + r)^2},$$

$$\Delta = \sqrt{1 - \kappa \sin^2 \varphi}\ ,$$

and

$$E(\kappa) = \int_0^{\pi/2} \Delta \, d\varphi$$

is the complete elliptic integral of the second kind. Also required will be

$$K(\kappa) = \int_0^{\pi/2} \Delta^{-1} \, d\varphi$$

which is the complete elliptic integral of the first kind. Note that

$$K + 2\kappa E' = \int_0^{\pi/2} \Delta^{-1} \, d\varphi - \int_0^{\pi/2} \kappa \sin^2\varphi \, \Delta^{-1} \, d\varphi = \int_0^{\pi/2} \Delta^2 \, \Delta^{-1} \, d\varphi = E$$

so that

$$2\kappa E' = E - K,$$

a result which will be needed shortly. Using the above relations,

$$P(r, a, \mu) = -2\pi + 4\mu(a + r)\, E(\kappa) - \pi\mu^2(a^2 + r^2).$$

Note the symmetry on interchanging a and r. The inward acceleration g(r, a) at distance r from the origin is the derivative of the potential with respect to r. Use

$$\frac{\partial \kappa}{\partial r} = \left[\frac{1}{r} - \frac{2}{a + r}\right]\kappa = -\frac{1}{a + r}\left[1 - \frac{a}{r}\right]\kappa, \quad r \neq 0,$$

so

$$\frac{\partial}{\partial r}[(a + r)E(\kappa)] = E(\kappa) - \left[1 - \frac{a}{r}\right]\kappa E'(\kappa) =$$

$$E(\kappa) - \frac{1}{2}\left[1 - \frac{a}{r}\right][E(\kappa) - K(\kappa)] = \frac{1}{2}\left[1 - \frac{a}{r}\right]K(\kappa) + \frac{1}{2}\left[1 + \frac{a}{r}\right]E(\kappa).$$

The acceleration at distance r from the origin due to a radius a ring is $\rho_{1d}\, \gamma G\, \mu a\, Q(r, a, \mu)$, where

$$Q(r, a, \mu) = \frac{\partial P(r, a, \mu)}{\partial r} =$$

$$2\mu\left[[1 - a/r]K(\kappa) + [1 + a/r]E(\kappa) - \pi\mu a\right].$$

A three dimensional plot of $aQ(r, a, 1/520\ kpc^{-1})$ is given in Figure 5 in which the independent variables a and r vary between 0 and 50 kpc. The force is everywhere inward, even for $a > r$.

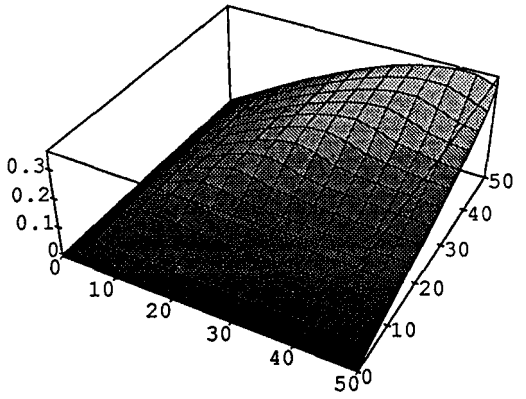


Figure 5: The weighting function $aQ(r, a, 1/520\ kpc^{-1})$ for determining the inward specific force on a star at distance r from the center of a galactic disk is plotted as a function of r and a.. The source of the force is a massive ring with radius a. The r and a coordinate unit is kpc. This and several other figures were plotted using Mathematica.[6]

5. THE VIRIAL THEOREM FOR GALACTIC CLUSTERS AND SUPER-CLUSTERS

The crude model presented explains away the missing matter in the vicinity of our Sun and the missing matter in rotating spiral galaxies. Because the exponential force is larger than the force of gravitation out to a distance of 6 Mpc, it should be very important - if not dominant - in the dynamics of galactic clusters. Let T be the kinetic energy of a cluster of N galaxies (each galaxy considered to be a point mass); let U_2 be the total gravitational potential energy of the cluster; and finally let $\vec{F}_n$ be the exponential force on galaxy n and let $\vec{r}_n$ be its position vector. Then the virial theorem states that

$$2\overline{T} + \overline{U_2} + \sum_{n=1}^{N} \overline{\vec{r}_n \cdot \vec{F}_n} = 0$$

where the $\overline{\text{overbar}}$ denotes the time average. The total exponential potential energy is

$$U_E = -\gamma\mu G \sum_{\substack{m,n \\ m\neq n}} M_m M_n \exp[-\mu|\vec{r}_n - \vec{r}_m|],$$

so

$$\vec{r}_n\cdot\vec{F}_n = -\vec{r}_n\cdot\partial U/\partial\vec{r}_n = -\gamma G \sum_{\substack{m \\ m\neq n}} M_m M_n \vec{r}_n\cdot\frac{\vec{r}_n - \vec{r}_m}{|\vec{r}_n - \vec{r}_m|}\,\mu^2\exp[-\mu|\vec{r}_n - \vec{r}_m|].$$

For every term

$$-\gamma G\, M_m M_n\, \vec{r}_n\cdot\frac{\vec{r}_n - \vec{r}_m}{|\vec{r}_n - \vec{r}_m|}\,\mu^2\exp[-\mu|\vec{r}_n - \vec{r}_m|]$$

in $\vec{r}_n\cdot\vec{F}_n$, there is a corresponding term

$$-\gamma G\, M_m M_n\, \vec{r}_m\cdot\frac{\vec{r}_m - \vec{r}_n}{|\vec{r}_n - \vec{r}_m|}\,\mu^2\exp[-\mu|\vec{r}_n - \vec{r}_m|]$$

in $\vec{r}_m\cdot\vec{F}_m$, and their sum is

$$-\gamma G\, M_m M_n\, \mu^2|\vec{r}_n - \vec{r}_m|\exp[-\mu|\vec{r}_n - \vec{r}_m|]$$

which (remember that $\gamma\mu$ = constant) is also equal to

$$-\mu\frac{\partial}{\partial\mu}[-\gamma\mu G M_m M_n \exp[-\mu|\vec{r}_n - \vec{r}_m|]].$$

Therefore,

$$\sum_{n=1}^{N}\overline{\vec{r}_n\cdot\vec{F}_n} = -\mu\overline{\frac{\partial U_E}{\partial\mu}},$$

and the virial theorem is

$$\overline{2T + U_2 - \mu\frac{\partial U_E}{\partial\mu}} = 0.$$

The estimation of the total potential energy of a spherical cluster can be approached by constructing the cluster by integrating outward over successive shells - like putting an onion back together from the inside out. As each subsequent shell is added, the decrease in potential energy is calculated so the total potential energy can be calculated by integration. The change in potential energy caused by adding a shell of radius a, thickness da, and density $\rho(a)$ is just $4\pi\rho(a)\,a\,da$ times the change in potential energy caused by a unit point mass at radius a. For the simple model of a uniform density sphere, an alternative integration is convenient; the geometry is depicted in Figure 6. The solutions for U_2 and $-\mu(\partial U_E/\partial\mu)$ are

$$U_2 = -4\pi^2 G\rho^2\int_0^a\int_0^{2r} R\,(2 - R/r)\,dR\,r^2dr = (16/15)\pi^2 a^5 G\rho^2$$

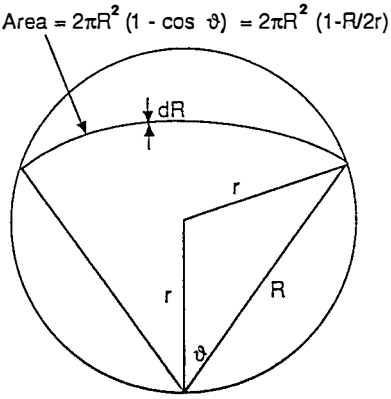


Figure 6. Geometry for the integration of the exponential and gravitational potential energies of a uniform sphere.

and

$$-\mu\,(\partial U_E/\partial\mu) = 4\pi^2\gamma\mu^2 G\rho^2 \frac{\partial}{\partial\mu}\int_0^a \int_0^{2r} R^2\,(2 - R/r)\,\exp(-\mu R)\,dR\,r^2 dr =$$

$$8\pi^2\gamma\mu^{-4}G\rho^2\left[[15 - 6(\mu a)^2 + 2(\mu a)^3] - [15 + 30\mu a + 24(\mu a)^2 + 10(\mu a)^3 + 2(\mu a)^4]\,\exp(-2\mu a)\right].$$

The ratio of the two terms is

$$-\mu\frac{\partial U_E}{\partial\mu}\,U_2^{-1} = \frac{15\,\gamma\mu}{2}\left[15(\mu a)^{-5} - 6(\mu a)^{-3} + 2(\mu a)^{-2} -\right.$$

$$\left.[15(\mu a)^{-5} + 30(\mu a)^{-4} + 24(\mu a)^{-3} + 10(\mu a)^{-2} + 2(\mu a)^{-1}]\,\exp(-2\mu a)\right].$$

The terms with the factor $\exp(-2\mu a)$ are negligible except for small clusters. This ratio, which is plotted in Figure 7, equals or exceeds unity over the range 14 kpc $\leq r \leq$ 83 Mpc and it equals or exceeds 100 over the range 160 kpc $\leq r \leq$ 7.4 Mpc.

A uniform density cluster is not especially realistic, but taking the ratio of the two potential energies somewhat ameliorates this shortcoming. Whatever the model, however, it is clear that V_E swamps V_2 in importance in applying the virial theorem for galactic clusters and superclusters. Incidentally, at a sufficient number of "skin depths" (λ = 520 kpc) inside the boundary of a uniform density cluster, the force due to V_E is negligible compared with gravity. The force due to the exponential potential is a surface force at scales larger than about 10 Mpc. It should behave like a surface tension, organizing large galactic clusters into sheets similar to the way that surface tension contributes to soap bubble films.

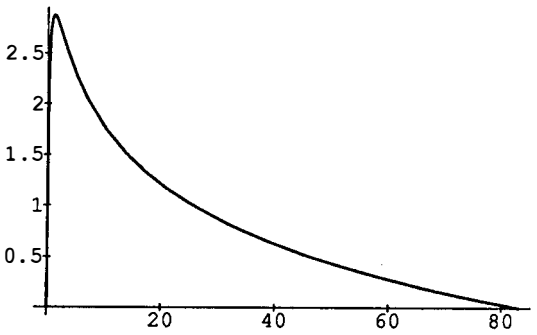


Figure 7. The logarithm of the ratio of the potential energy of the nominal exponential potential model to the gravitational potential energy is plotted as a function of the radius, in Mpc, of a uniform density sphere.

6. CONCLUSIONS

If we can accept an exponential potential of galactic scale, we can resolve the missing matter matter. An exponential potential model conforms with special relativity and it appears to agree well with all observations; no special circumstances are required beyond the very existence of the exponential potential. The exponential potential seems to be able to account for the distinctively different dimensions that are typical of galaxies and galactic clusters. The leading competing explanation for galactic kinematics calls for the existence of invisible matter that is distributed according to very special models, and a model of invisible matter that explains one phenomenon is not robust enough to explain another: a different sort of invisible matter model is required to explain rotation in the plane of our Galaxy from the sort of model that is required to explain motion perpendicular to the plane. Moreover, the amount of invisible matter required to hold galactic clusters together is "astronomical." The very existence of an enormous amount of matter should be tested by looking for it. Not being able to see it seems to be pretty good evidence that it is not there. A theory is tested by seeing how well it conforms with all the empirical evidence available. If there are no conflicts, it is viable. The exponential potential appears to be quite viable.

The two parameters λ and μ of the exponential potential have been estimated from the scale height $L_1 = \sqrt{2}\sigma$ of our Galaxy and the ratio $F = f(0, \mu, \gamma, \sigma)$ of its ostensibly missing mass to the observed mass in the vicinity of the Sun, and from the radial distance L_2 in a typical spiral galaxy at which the observed rotation velocity is $n = 2$ times that calculated from its observed mass by Newtonian theory alone. In essence, the equations used are

$$\lambda = \frac{FL_2^2}{\sqrt{\pi}(n^2 - 1)}$$

and

$$\gamma = \frac{1}{\pi(n^2 - 1)}\left[\frac{FL_2}{L_1}\right]^2.$$

The theory presented demands tests that use ample astronomical data in more detailed models than the simple nominal model of this paper. Its two parameters can be refined from the nominal estimates or the theory can be falsified. Because the ratio of the mass of a weak force boson to an exponential potential boson is $\sim 10^{40}$, an inquest into the nature of this "new" force appears warranted. For example, extending Dirac's speculations[7] on the significance large dimensionless numbers in nature leads to the conjecture that the two forces were commensurate in the early universe.

REFERENCES

1. M. Milgrom, *Ap.J.* **270**, 365 [Paper I], 371 [Paper II], 384 [Paper III](1984); I confess that I never got around to reading Paper IV (1985).
2. E. Fischbach, C. Talmadge and D. Krause, *Phys. Rev. D* **43**, 460 (1991). [I note here that my co-conspirators (Jekeli and Romaides) and I tried modeling tower and mine gravity data using attractive and repulsive Yukawa terms with slightly different coupling coefficients and slightly different boson masses. We found that (take your choice): 1, the coupling coefficients are identical and the boson masses are identical in the geophysical window so that the effects completely cancel out each other or, 2, there are no effects. Philosophers inform me that Occam's razor favors the latter alternative.]
3. J. D. Anderson and E. M. Standish, Jr., in *The Galaxy and the Solar System*,R. Smoluchowski, J. N. Bahcall and M. S. Matthews, eds., The University of Arizona Press, 1986. Probably a tighter constraint exists now, but it would still be utterly insignificant.
4. J. N. Bahcall, *Ap. J.* **287**, 926 (1984).
5. K. G. Begeman, A. H. Broeils and R. H. Sanders, *Mon. Not. R. Astr. Soc.* **249**, 523 (1991).
6. S. Wolfram, *Mathematica: a system for doing mathematics by computer*, Addison-Wesley Publishing Co., Inc, 2nd ed., 1991. Mathematica has also been used to derive many of the equations in this paper.
7. P. A. M. Dirac, *Nature* **139**, 323 (1937); P. A. M. Dirac, *Proc. R. Soc Lond.* A **165**, 199 (1938).

LUNAR LASER RANGING
and other travelling light experiments

Christian Veillet
Observatoire de la Côte d'Azur
Av Copernic - F 06130 GRASSE

ABSTRACT

Lunar Laser Ranging (LLR) is a technique providing one of the most precise information on the dimensions of the space close to the Earth. From a few meters twenty years ago, the accuracy of the Earth-Moon distance monitoring is now 2 to 3 cm, and could reach few millimeters in 1995. Unfortunately, the network of stations ranging the Moon is still very small (two sites only are operational), but it could grow in the future with new observatories equipped for LLR. The modeling of the Earth-Moon system has to meet the quality of the observations, and is sensitive to the gravitation theory used. Such a theory can then be checked against the real world seen. LLR is up to now the best test of the strong equivalence principle.

Other experiments involving light pulses travelling in space are currently used or planned for time synchronisation purposes. The accuracy which can be reached by the new equipments could allow tests of various effects of gravitational fields on the travel path of the light.

1. Laser ranging : a time measurement ...

Laser ranging is not really a range measurement. A laser pulse is sent to a remote retroreflector, which sends back (in the same direction as it arrived) part of the incident light. The start time of the pulse from the laser is recorded, as well as the arrival time of the returning light. The interval between these two events is the round trip time of the light between the station and the target. It does not really correspond to a distance, as both the transmitting telescope and the target moved during the travel time, but, by multiplying it by the speed of light, one obtains a distance equivalent quantity easy to use for a clear understanding of the precision. One nanosecond round trip time is 15cm in distance, and determining the Earth-Moon distance with a 5mm accuracy means measuring the round trip time (roughly 2.5s) with a 30ps accuracy !

2. More targets on the Moon than stations on the Earth !

For many years since 1969 and the Apollo XI landing on the Moon, McDonald Observatory (Texas, USA) has been the main source of LLR data. At that time, one used ruby laser whose pulse length was around 3 ns. This pulse length is a very important factor in LLR, as it is the fundamental source of uncertainty of the technique. In spite of the 10^{18} or 10^{19} photons sent in every laser pulse, one can detect one photon only over 100 to 1000 shots. It is impossible to know if this returning photon comes from the beginning or the end of the transmitted pulse. Shorter is this pulse, better is the round trip time determination, or the equivalent range. The accuracy was around 15cm at the end of the ruby era. The lasers used now have a 300-400ps pulse length, which permits to reach, by averaging on a few minutes, an accuracy of 2 to 3 cm.

Another LLR station at CERGA (Grasse, France) started getting data in 1981, and since 1984 joined McDonald on an operational basis in the LLR network. A new station replaced the 2.7m telescope at McDonald in 1984, using a YAG laser as CERGA did two years later. Since that time, CERGA replaced the Texan station by the number of observations from the four targets available on the Moon (Apollo 11, 14 and 15, Lunakhod 2). Fig. 1 gives an idea of the data collection for the last years. Another station worked very efficiently from Haleakala (Maui, Hawaii), but had to give up ranging the Moon in 1991 due to a lack of funds from NASA.

3. For various scientific applications

The scientific goals of LLR are roughly the same as those established in the late

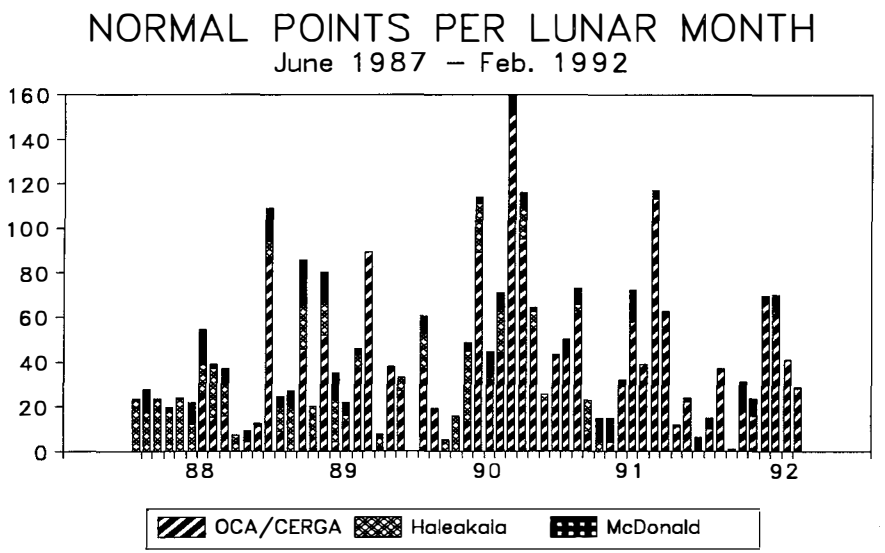


Fig. 1 *Histogram of the LLR data per lunar month. Each data is a normal point, the round trip time of a light pulse leaving the Earth at a given time and reflected on one of the four targets available on the Moon, as determined from 10 to 100 photons detected over 10 to 30 minutes.*

sixities at the beginning of the Apollo 11 experiment. The computation of the round trip time of the laser pulse is made from three vectors linking succesively the station to the Earth's center of mass, Moon's to Earth's centers of mass, and reflector to Moon's center of mass. Earth's dynamics, Lunar motion and Lunar rotation are thus the three main domains of applications of Lunar laser Ranging. (Alley et al., 1965 ; Alley & Bender, 1968). The accurate modeling of the Earth-Moon system needed for accounting for the LLR measurements means that one can check the gravitation theory used, as the result of the computation depends strongly on this theory.

Table 1 gives the present uncertainty of the main parameters of the Earth-Moon system as determined by the JPL analysis group.

4. Which future for the LLR technique ?

A severe limitation to the accuracy of the LLR measurements comes from the atmosphere where the light does not travel as in vacuum ! The knowledge of the time lost in the air on the way to the Moon or back to the Earth is unfortunately inaccurate, as it depends on a model of the atmosphere computed from the meteorological parameters on the ground at the station location. Varying with the elevation of the

Lunar librations	± 10 milliarcsec
Lunar Rotational Dissipation parameter	0.00113 ± 0.00002
Lunar Love Number k_2	0.028 ± 0.003
C/MR^2	0.391 ± 0.002
Lunar Euler Equation Inertial Parameters β' and γ'	$\pm 1 \times 10^{-3}$% and 1×10^{-2}%
Dynamical Equinox	±5 milliarcsec
Obliquity of the Ecliptic	±2 milliarcsec
Parametrized Post-Newtonian Parameter β	$\pm 1.1 \times 10^{-3}$
Geodetic precession	±0.4 milliarcsec/yr
Limit on $\dot{G}/G$	$\pm 2 \times 10^{-11}$/yr

Table 1 - *Present Lunar Laser Ranging results as obtained at JPL (Bender et al 1992)*

target, the uncertainty is never better than 5mm, and can reach 1.5 to 2cm when the Moon is low... The present uncertainty of the LLR measurements is not yet at that level, and the upgrades planned at the OCA/CERGA station for the three years to come should bring the accuracy down to the atmospheric limitation. If ranging is made for a decade at that level of accuracy, one could perform an Eötvös-type experiment using the Earth-Moon system with an accuracy of 5×10^{-13}, as the laboratory experiments have quoted accuracies between 10^{-11} and 10^{-12} (Nordtvedt, 1988). With the present data, LLR can be also considered as the β coefficient experiment (β parametrizes the nonlinearity of the gravitational theory), which gives the best constraint on the nonlinear structure of gravitational theory. The upgrades undertaken at the CERGA/ OCA LLR station, as well as the possible contribution of other sites, will still improve the quality of these results.

5. Other travelling light experiments

Instead of using laser pulses for determining the *distance* between two objects by the round trip time, one can use it for transfering time information between two remote clocks. Such an experiment, LASSO, is currently done between France and the USA, using a flying oscillator (implemented on a geosynchronous satellite Meteosat 3/P2). Originally planned for comparing clock at the 1ns level, it is able to reach an accuracy better than 100ps, assuming that the calibrations at both ends of the laser link are made at that level. For realizing that, the satellite is equipped with a laser pulse

detector linked to the oscillator, and with a retroreflector. A laser pulse is sent from a ground station, is detected onboard of the satellite, and an echo is received back at the station. Start and return time at the station and arrival time on the satellite allow to link the flying clock to the one at the station. If two sites work at the same time, their clocks will be linked through the flying clock which vanishes in this time transfer experiment! Initially designed for comparing the various time transfer techniques, the same concept can be used for checking the behaviour of the light path in a gravitational field.

By improving the components of the experiment (shorter laser pulses, better timing, more stable flying clock, faster light detectors), one can fly a new generation LASSO package in the Solar System, for exemple on the orbit of Mars, and observe the variation of the one way flight time as the light goes close to the Sun. The limit could be only one solar radius from the surface of the Sun, as there is no perturbation by the solar corona as it is for comparable experiments using radio wavelengths. The improvement on the γ parameter through the Shapiro effect (Shapiro, 1964) could be of one or two orders of magnitude compared to the current uncertainties from VLBI observations. The feasability of such an experiment, TROLL (Testing Relativity in Orbit by Laser Light), is being studied at CERGA and CNES.

6. Conclusion

Using light flight time in space has not reached yet its limits in accuracy. Many improvements are being done on lasers, clocks and detectors. The determination of the parameters derived from the techniques using light pulses travelling in space should be better and better within the ten years to come !...

Bibliography

Alley, C.O, Bender, P.L., 1968, in *Continental drift, secular motion of the pole and rotation of the Earth,* Proc. IAU Symp. 32, ed. W. Markowitz & B. Guinot, Reidel, Dordrecht

Alley, C.O, et al., 1965. J. Geophys. Res. **70**, 2267.

Bender, P.L., et al, 1992. Lunar K-band transponder program.*Workshop on early robotic missions to the Moon, Lun. Plan. Inst., Houston, Texas, Feb. 1992*

Nordtvedt, K., 1988. Lunar Laser Ranging and laboratory Eötvös-type experiments. *Phys Rev. D*,**37**, 1070

EXPERIMENTAL RESULTS OF THE GRAVITATIONAL WAVE EXPERIMENT OF THE ROME GROUP

Fulvio Ricci

Dipartimento di Fisica, Universita' di Roma "La Sapienza", Rome, Italy
Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

Abstract

We briefly review the basic features of our Weber type graviational wave antenna of the second generation, the EXPLORER. Analysing the data with an optimum linear filter it is possible to reach a sensitivity for short bursts of gravitational waves limited by the thermal noise of the detector and by the noise of the electronic amplifier.

Since July 1990 the Explorer antenna, located in Geneva at CERN, is running in steady operation at 2 K, with a strain sensitivity $h < 10^{-18}$ sufficent to detect galactic events in coincidence with similar antennas.

The new Ultra low temperature antenna NAUTILUS is completed and it has been cooled for a first test run at a temperature of 95 mK.

The Rome group started its activity on the detection of Gravitational Waves (G.W.) in 1970 having as final target the construction of a Weber type resonant antenna cooled at ultra low temperature (mK). This detector should be sensitive to gravitational wave pulses emitted from the Virgo cluster of galaxies and it should take data in coincidence with similar detectors installed at the Louisiana State University (LSU) and at the Stanford University.

For impulsive signals of a time duration $\tau_g = 1$ ms, the strain sensitivity[1] is expressed by the following formula:

$$h \simeq 3 \cdot 10^{-16} \frac{1\ Kpc}{R} \sqrt{\frac{M_{GW}}{M_\odot}} \tag{1}$$

where R is the distance between the detector and the source expressed in Kpc and the mass ratio $M_{GW}/M_\odot$ is the fraction of solar masses converted in gravitational radiation energy.

Events with $(M_{GW}/M_\odot) = 10^{-2}$ correspond to $h \simeq 3 \cdot 10^{-18}$ for galactic signals ($R \simeq 10\ Kpc$), while for the Virgo Clusters signals ($R \simeq 10\ Mpc$) we have $h \simeq 3 \cdot 10^{-21}$ and we expect an event rate of the order of 1-10 events/month.

The main intermediate milestone of this research activity is to put in operation over a long time period antennas of second generation working at liquid helium temperatures. These new massive antennas started to operate about 10 years later (Frascati 1980, Stanford 1980, Rome 1986, LSU 1987), but the overlap time of data taking was short and the orientation of the antennas was far from the optimum condictions[2] for the detection of G.W. pulses.

Infact, a basic problem of the gravitational wave experiment is the continous operation with good and stable performances. As regards as the second generation dectector of the Rome group installed at CERN, the antenna named EXPLORER, a big effort has been done to improve its stability over a long period. This antenna[3] is an alluminium alloy cylinder with a mass of 2270 kg and a length of 3 m. The bar is cooled with liquid helium in superfluid regime, and it is oriented approximately parallel to the LSU cryogenic detector. We monitor the vibration status of the bar at its first longitudinal vibration mode of $\nu_b = 915.69$ Hz, using a resonant capacitive transducer mounted at one end of the bar. One plate of the capacitive transducer has its firts flexural mode as equal as possible to the bar frequency. A good model of this system is a series of two harmonic oscillators. Two peaks, ν_+ and ν_-, due to the two normal modes of such a double harmonic oscillator system, appear in the Fourier spectrum of the antenna output.

The capacitive transducer is biased at constant electric charge. The typical value of its electrostatic field is $6 \cdot 10^6$ V/m. The frequency of the transducer, ν_t, depends also from the bias electric field and then, by monitoring the frequency change of the two normal modes ν_+ and ν_-, we can measure the static electrical losses of the transducer. These losses are of the order of $10^{-4}\ (day)^{-1}$.

The transducer output is applied to a superconducting transformer and then to the input coil of a dc Squid amplifier in order to match the output transducer impedence with the low input impedence of the Squid. The Squid is used in a classical flux-locked configuration to

stabilize its operation point. The Squid characteristics, i.e. the family of the Squid voltage output functions versus the magnetic flux input, are not smooth. Therefore the Squid performances depend on the operation point that is determined by the bias dc current and the locking of the feedback loop. In this configuration the typical value of the dc Squid flux noise is $\Phi_n \simeq 2 \cdot 10^{-6}\ \Phi_0/\sqrt{Hz}$, with $\Phi_0 = 2 \cdot 10^{-15}$ Weber.

The Squid output is sent to several double lock-in amplifiers. Two of them are used to demodulate the signal input at the frequencies ν_+ and ν_- respectively. The outputs of the two lock-in amplifiers are the orthogonal components $x_\pm$ and $y_\pm$ of the input signal at each of the two normal mode of the antenna - transducer system. The four signals, $x_\pm$ and $y_\pm$, are sampled every 290.8 ms and stored together the Universal Time information in the memory of a computer (μVAX 3800).

The calibration of the apparatus is performed by sending a voltage pulse to a small piezoelectric ceramic glued on the center of the bar, that gives a known amount of mechanical energy to the bar.

An online monitor of the detector background noise is performed applying a simple algorhythm to the data. From the mean value of the quantity

$$E_\pm = \frac{1}{K_\pm}[x_\pm{}^2 + y_\pm{}^2] \tag{2}$$

where $K_\pm$ is a normalization constant, we deduce the equipartition temperatures of the two modes that should be, in a first approximation, equal to the thermodynamical temperature measured with a germanium thermometer. In our system the heating back-action effect of the mechanical modes due to the Squid noise is expected to be negligeable.

The search of short bursts is performed by applying the Wiener- Kolmogoroff (WK) filter to the data. This is done by assuming that the noise in the detector is the superimposition of the white noise Φ_n, due to the Squid , and the Lorentzian noise of the Brownian motion of the two normal modes ν_+ and ν_-. In particular the WK filter requires the kwnoledge of $\Gamma_\pm$, the spectral ratios between the wideband noise and the narrowband noise at the mode frequencies, and $\tau_\pm$, the typical time constants of the modes . We evaluate for each mode the quantities

$$w_{x\pm}(t_i) = \Sigma_j x_\pm(t_i - t_j) f(t_j) \tag{3}$$

and

$$w_{y\pm}(t_i) = \Sigma_j y_\pm(t_i - t_j) f(t_j) \tag{4}$$

where $x_\pm(t_i)$ and $y_\pm(t_i)$ are the $x_\pm$ and $y_\pm$ data sampled at the time t_i, and the coefficients $f(t_j)$ are the weights of the filter. These coefficients decrease exponentially both for data taken in the past and in the future of t_i with a typical time constant

$$\tau_{3\pm} = \tau_\pm \sqrt{\frac{\Gamma_\pm}{1 + \Gamma_\pm}} \tag{5}$$

The two quantities

$$\Delta E_\pm(t_i) = \frac{1}{K_{w\pm}}[w_{x\pm}(t_i)^2 + w_{y\pm}(t_i)^2] \tag{6}$$

are an estimate of the variation of the mode vibration energy; the average values of $\Delta E_{\pm}$ are the effective temperatures $T_{eff\pm}$ of the two modes. The overall effective temperature of the system is obtained by the combination

$$\frac{1}{T_{eff}} = \frac{1}{T_{eff+}} + \frac{1}{T_{eff-}} \tag{7}$$

The strain sensitivity h of the detector , i.e. the minimum detectable amplitude of gravitational wave, is a function of T_{eff} and τ_g:

$$(h\tau_g)_{min} = \frac{L}{v^2}\sqrt{\frac{k_B T_{eff}}{M}} \tag{8}$$

where k_B is the Boltzmann constant, L, M and v are the length, the mass and the sound velocity of the bar.

The last run of data taking[5] started on July 1990 and operated nearly continously until the December 16^{th} of 1991 at a sensitivity level $h < 10^{-18}$. After one year and half , we paused the run in order to recover the efficiency of the active carbon traps, which are located in the vacuum insulation of the cryostat. In figure 1, we report the mean performances of the EXPLORER antenna during the last 6 months. Each point of the plot is the h_{min} value obtained from the average of one hour of data, assuming $\tau_g = 1$ ms. We notice that most of the time the sensitivity is well below the $h = 10^{-18}$ treshold. This figure concerns all recordered data. It is possible to improve the signal to noise ratio by excluding single events or time periods in which the data are correlated to those of the seismo-acoustic and electromagnetic noise sensors of the laboratory. As an example of this selection procedure, in fig. 2 the experimental distribution of ΔE data taken from December 13^{th} 1991, 04:45 UT to December 16^{th} 1991,04:51 UT is shown. The distribution obtained using the information of the auxiliary channel is shown in the same illustration. This selection does not affect the slope of the exponential distribution (in the logarithmic scale), because it is due to the intrinsic noise of the detector. The reduction effect of the distribution tail is a relevant point in the search of coincidences with other antennas. Let me point out that we have an additional tool to select the events. A gravitational wave signal releases energy in both the modes in a well defined ratio, instead external disturbances of certain categories affect one mode more than the other. Then, the events can be selected on this condiction both in the case of one antenna running or in the case of a coincidence experiment.

In the last six months of 1991, EXPLORER was operating in coincidence with the L-SU antenna that has similar resonance frequency and sensitivity. Moreover, starting from November 1991, our smaller antenna of M=400 kg and of ν_b = 1800 Hz, located in Frascati , is running with an effective temperature of the order 20 mK. The coincidence analysis of the data taken with these antennas, with an energy sensitivity 1000 times better than the room temperature ones, is just started.Our plan is to mantain the operation of these detectors as long as possible.

The long data taking of the EXPLORER is a milestone. Our scientific target is to detect events from the Virgo cluster of galaxies. A new system, NAUTILUS, equipped with a ^{3}He -

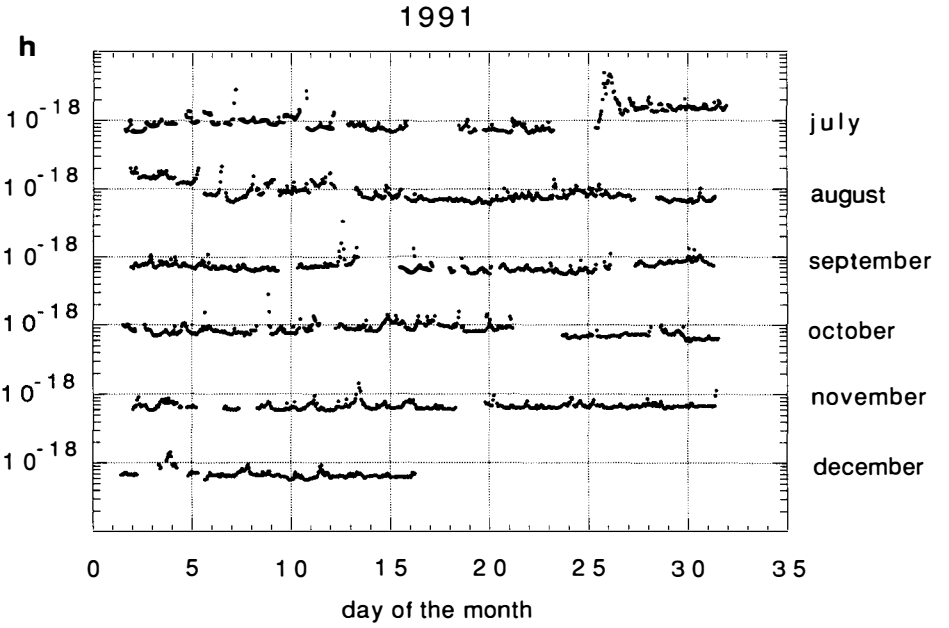


fig. 1 - The EXPLORER sensitivity in the time period July 1991 - December 1991

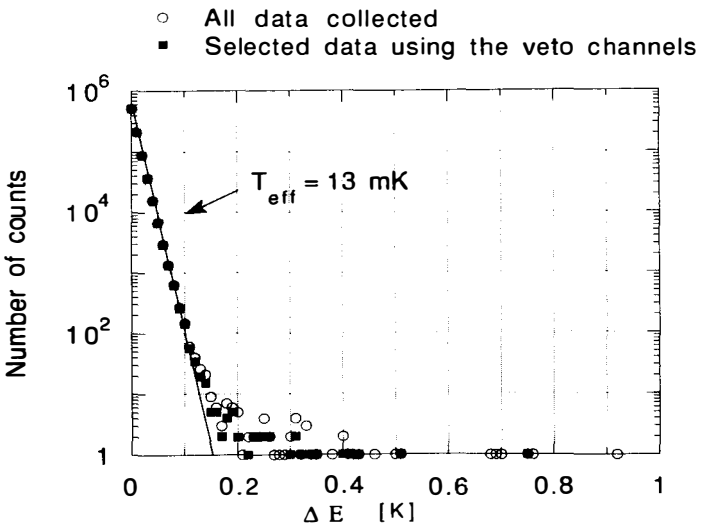


fig. 2 - Experimental distribution of the EXPLORER data taken in the time period 13 -16 december 1991

^{4}He dilution refrigerator, is in developement. The bar of the new detector has the same length and mass of that of the EXPLORER.

In Febrary 1991 we tested for the first time the complete cryogenic apparatus[4)] cooling the 2300 kg cylinder at 95 mK . Due to a technical problem concernig the still chamber of the dilution refrigerator, the final temperature was achieved with a refrigeration power near to the minimum value. In the future, a change of the transducer configuration is planned in order to approach the quantum limit of the detector.

In conclusion, EXPLORER is taking data looking for events emitted in our Galaxy.

The NAUTILUS antenna is just born. It will extend our sensitivity up to the Virgo cluster where thousands galaxies are available as gravitational wave sources. A soon as NAUTILUS becomes operational we will move it to Frascati and operate both EXPLORER and NAUTILUS in coincidence.

References

(1) K.S. Thorne in "*300 Years of Gravitation*", S.W. Hawking, W. Israel editors
(2) E. Amaldi et al. *Astronomy and Astrophysics* **216**, 325 (1989)
(3) E. Amaldi et al. *Nuovo Cimento* **C 9**, 829 (1986)
(4) P. Astone et al. *Europhysics Letters* **16**, 231 (1991)
(5) P. Astone et al. "*Long Term Operation of the EXPLORER Cryogenic G. W. Detector*" in preparation

TENSOR-MULTI-SCALAR THEORIES OF GRAVITATION AND STRONG-FIELD TESTS OF GENERAL RELATIVITY

Gilles Esposito-Farèse
Centre de Physique Théorique – CNRS Luminy, Case 907
F 13288 Marseille Cedex 9 (France) *
Institut des Hautes Etudes Scientifiques, 35 route de Chartres
F 91440 Bures-sur-Yvette (France) **

Abstract : This talk summarizes a work in collaboration with Thibault Damour. We introduce a generic class of theories where gravity is mediated by one tensor field together with several scalar fields. We study the predictions of such theories both in the weak-field and in the strong-field regimes, and both for quasi-stationary and for radiative effects. We show that a wide sub-class of these theories can have strictly the same post-Newtonian limit as general relativity, while being very different from it for strong fields. As an illustration, we exhibit a specific tensor-bi-scalar theory which is not only consistent with all solar-system experiments, but also with the $\dot{P}$-$\dot{\omega}$-γ test in the binary pulsar PSR 1913+16. This underlines that new tests of relativistic gravity in the strong-field regime are necessary to constrain the space of allowed theories.

* Permanent address
** Address in 1991-92

1. Introduction

Solar system experiments have confirmed to one part in 10^3 the post-Newtonian predictions of general relativity in the quasi-stationary-weak-field regime[1,2]. However, many aspects of this theory cannot be tested in such a limit, which would correspond in one dimension to approximate a function $f(\varepsilon)$ by the parabola $f(0) + \varepsilon f'(0) + \frac{1}{2}\varepsilon^2 f''(0) + O(\varepsilon^3)$. For relativistic theories of gravitation, the small parameter ε entering this expansion represents both the strength GM/Rc^2 of the gravitational field [which measures how much the metric $g_{\mu\nu}$ differs from the flat one diag$(-1,1,1,1)$] and its rapidity of time-variation $(v/c)^2$. In order to test Einstein's theory beyond its first post-Newtonian level, three different regimes can therefore be distinguished : (i) rapidly-varying weak fields (gravitational wave experiments), (ii) quasi-stationary strong fields (motion of compact bodies like neutron stars or black holes) and (iii) radiative effects in strong fields (gravitational radiation of compact-body systems). The best way to point out the specific predictions of general relativity in these regimes will be to compare it with alternative theories of gravity. However, it will not be possible to generalize for strong fields the Parametrized Post-Newtonian (PPN) formalism[3], since an infinite number of parameters would be necessary to describe all non-perturbative effects in the compactness GM/Rc^2. We shall instead restrict our attention to a particular *class* of theories, which is wide enough to allow the study of many new effects in these regimes, but also arises naturally in current theoretical attempts at unifying fundamental interactions or at quantizing gravity. Indeed, we shall consider theories where the gravitational interaction is mediated by one tensor field (the "graviton") together with an arbitrary number of scalar fields, which can be either the "compactons" of Kaluza-Klein theories, the "dilaton" already present in the extra-dimensional space of string theories, or the scalar degrees of freedom which appear in higher-order extensions of Einstein's theory. These "tensor-multi-scalar" theories are also the most natural generalizations of the (one-scalar) theory of Jordan-Fierz-Brans-Dicke[4-7], which has played a very useful rôle in the past in suggesting new tests of general relativity, and notably of the strong equivalence principle (universality of free fall of self-gravitating bodies)[8-12]. However, this tensor+1 scalar theory contains only one parameter [which is now well constrained by experimental data[13]: $1/\omega < 2\times10^{-3}$] and *all* its predictions in both weak-field and strong-field conditions fractionally differ from those of general relativity by quantities of order $1/\omega$. Therefore, its usefulness as a contrasting theory of gravity is now essentially reduced to zero. On the contrary, we will see below that a large class of tensor-*multi*-scalar theories can pass all the existing solar-system tests and still be very different from general relativity in the strong-field regime, as we show in detail in the article 14. In this reference, we use our explicit analytical results to define a set of parameters which chart the space of theories beyond the first post-Newtonian level (*i.e.* which generalize the PPN parameters β and γ), and which can be used to interpret binary pulsar data as well as to suggest new tests of relativistic gravity. Because of space constraints, we will be able to quote below only some of the main results of reference 14.

2. Action and post-Newtonian limit

In view of the precise experimental verifications of the local isotropy of space and of the weak equivalence principle (universality of free fall of laboratory-size bodies), we consider only "metric theories", in which matter is coupled to some universal tensor $\tilde{g}_{\mu\nu}$. As was first pointed out by Pauli and Fierz[5], the metric $g^*_{\mu\nu}$ which appears in the Einstein-Hilbert action can differ from this "physical metric" $\tilde{g}_{\mu\nu}$ by a conformal factor depending on the scalar fields φ^a :

$$\tilde{g}_{\mu\nu} = A^2(\varphi^a) g^*_{\mu\nu} \ . \tag{1}$$

The action describing our general class of tensor-multi-scalar theories will therefore be written as

$$S = S_{g*} + S_\varphi + S_m \ , \tag{2}$$

where

$$S_{g*} = \frac{c^4}{4\pi G_*} \int \frac{\mathrm{d}^4x}{c} \sqrt{g^*} \frac{R^*}{4} \tag{3}$$

is the Einstein-Hilbert action describing the gravitons (spin 2),

$$S_\varphi = -\frac{c^4}{4\pi G_*} \int \frac{\mathrm{d}^4x}{c} \sqrt{g^*} \frac{1}{2} g^{\mu\nu}_* \gamma_{ab}(\varphi^c) \partial_\mu \varphi^a \partial_\nu \varphi^b \tag{4}$$

is the kinetic term of the scalar fields (spin 0), and

$$S_m = S_m[\psi_m, \tilde{g}_{\mu\nu}] \tag{5}$$

is a functional of the matter fields ψ_m and of the physical metric $\tilde{g}_{\mu\nu}$. In all these equations, a star index refers to the Einstein-Hilbert metric $g^*_{\mu\nu}$. In particular R^* is its curvature scalar (with the sign conventions of reference 15), and $g^* \equiv |\det g^*_{\mu\nu}|$. Note that the metric $\gamma_{ab}(\varphi^c)$ appearing in (4) is non-necessarily flat, and therefore that S_φ describes in general a non-linear σ-model. A potential $\int \mathrm{d}^4x \sqrt{g^*}\, B(\varphi^a)$ could also be added to this action in order to describe massive or self-interacting scalars, but we shall not consider such a term in this talk (see reference 14 for more details). Note also that the conformal factor $A(\varphi^a)$ hidden in $\tilde{g}_{\mu\nu}$ in equation (5) defines the coupling of matter to the scalar fields.

The dimensionful constant G_* appearing in (3) and (4) is the bare gravitational constant. It differs from the physical one $\tilde{G}$ (the effective coupling constant between two test masses) because the gravitational interaction is mediated by both $g^*_{\mu\nu}$ *and* the scalar fields φ^a. We show in reference 14 that

$$\tilde{G} = G_* A^2 (1 + \alpha_a \gamma^{ab} \alpha_b) \ , \tag{6}$$

where

$$\alpha_a \equiv \frac{\partial \ln A(\varphi)}{\partial \varphi^a} \ . \tag{7}$$

Hence the physical gravitational constant $\tilde{G}$ is the sum of two terms : $G_* A^2$ is due to the exchange of a graviton between the two test masses, and $G_* A^2 \alpha_a \gamma^{ab} \alpha_b$ corresponds to the exchange of the scalar fields [α_a appears at each matter–scalar vertex, and the inverse metric γ^{ab} corresponds diagrammatically to a scalar propagator]. We also

derive the post-Newtonian limit of these theories in reference 14, and we find that the PPN parameters are given by

$$\tilde{\gamma} - 1 = -2\frac{\alpha^2}{1+\alpha^2}\ , \tag{8}$$

$$\tilde{\beta} - 1 = \frac{1}{2}\frac{\alpha^a \beta_{ab} \alpha^b}{(1+\alpha^2)^2}\ , \tag{9}$$

where α^2 is a simplified notation for $\alpha_a \gamma^{ab} \alpha_b = \alpha_a \alpha^a$, and

$$\beta_{ab} \equiv D_a \alpha_b = D_a D_b \ln A \tag{10}$$

is the covariant derivative of α_b with respect to φ^a, which appears at each scalar–matter–scalar vertex. In order to satisfy solar-system experiments, which constrain both $|\tilde{\gamma} - 1|$ and $|\tilde{\beta} - 1|$ to be lower than 10^{-3} at the 1σ level, we will therefore consider theories such that $\alpha^2 \approx 0$ and $\alpha^a \beta_{ab} \alpha^b \approx 0$. Two very different cases can be distinguished :

a – *the case of a positive (or negative) definite σ-model metric γ_{ab}*

In that case, the constraint $\alpha_a \gamma^{ab} \alpha_b \approx 0$ means that all the components of the vector α_a almost vanish, and therefore that the scalar fields are almost not coupled to matter. Hence *all* the predictions of the theory (even in strong fields) are constrained by solar-system experiments. This is precisely the reason why the tensor+1 scalar theory of Jordan-Fierz-Brans-Dicke [which has obviously a positive definite σ-model metric $\gamma_{11} = 1$] is not useful for testing relativistic gravity beyond the post-Newtonian level, as already stressed in the introduction.

b – *the case of a hyperbolic signature for γ_{ab}*

In that case, the constraint $\alpha_a \gamma^{ab} \alpha_b \approx 0$ means only that α_a is almost a "null" vector, *i.e.* is almost on the "light cone" of the σ-model manifold, but the magnitude of its components is not constrained. Similarly, $\alpha_a \beta^{ab} \alpha_b \approx 0$ means only that the vector $\beta^{ab}\alpha_b$ lies almost in the hyperplane normal to α_a, but the magnitude of β^{ab} is not constrained. Therefore, in the case where the σ-model metric γ_{ab} has a hyperbolic signature, solar-system experiments only impose directional properties on α_a and β_{ab}, without limiting the strength of the coupling of the scalar fields to matter. The simplest example of a theory which satisfies strictly $\tilde{\beta} = \tilde{\gamma} = 1$, *i.e.* which has the same post-Newtonian limit as general relativity, is a tensor-*bi*-scalar theory such that $\gamma_{ab} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\alpha_1 = 0$, $\alpha_2 = 1$ and $\beta_{11} = 0$ [Indeed $\alpha_a \alpha^a = 2\alpha_1 \alpha_2 = 0$ and $\alpha_a \beta^{ab} \alpha_b = \alpha_2 \beta_{11} \alpha_2 = 0$]. We shall come back to this model in the next section.

3. Strong-field effects and comparison with binary pulsar data

We derive in reference 14 the Lagrangian describing the motion of N compact bodies, including only the first relativistic orbital corrections in $(v/c)^2$, but taking fully into account the self-gravity effects in GM/Rc^2 [where R denotes the radius of a compact body]. We find that it can be written in the same form as in the post-Newtonian limit, but that the effective gravitational constant $\tilde{G}$ and the PPN parameters $\tilde{\gamma}$ and $\tilde{\beta}$ are

now replaced by body-dependent parameters [the labels A, B, C refer to the compact bodies]

$$\widetilde{G}_{AB} \equiv G_* A^2 \left[1 + (\alpha_A \alpha_B)\right] , \tag{11}$$

$$\gamma_{AB} - 1 \equiv -2 \frac{(\alpha_A \alpha_B)}{1 + (\alpha_A \alpha_B)} , \tag{12}$$

$$\beta^A_{BC} - 1 \equiv \frac{1}{2} \frac{(\alpha_B \beta_A \alpha_C)}{\left[1 + (\alpha_A \alpha_B)\right]\left[1 + (\alpha_A \alpha_C)\right]} , \tag{13}$$

where $(\alpha_A \alpha_B) \equiv \alpha^A_a \gamma^{ab} \alpha^B_b$ and $(\alpha_B \beta_A \alpha_C) \equiv \alpha^a_B \beta^A_{ab} \alpha^b_C$. In these formulae, the σ-model vector α^A_a is a body-dependent generalization of equation (7)

$$\alpha^A_a \equiv \alpha_a + \frac{\partial \ln \widetilde{m}_A}{\partial \varphi^a} , \tag{14}$$

where $\widetilde{m}_A$ denotes the physical (inertial) mass of body A, and similarly $\beta^A_{ab} \equiv D_a \alpha^A_b$ generalizes equation (10). Note that the masses of compact bodies depend in general on the scalar fields, since their values modify the laws of gravitational physics and thereby the internal structure of the bodies. We show in reference 14 how the derivatives $\partial \ln \widetilde{m}_A / \partial \varphi^a$ can be evaluated for neutron stars and for black holes.

We also derive in this reference the expression of the energy loss due to gravitational radiation in tensor-multi-scalar theories, but it is too long to be displayed here. Let us quote only that besides the usual quadrupole formula of order $O(c^{-5})$ which already exists in general relativity, this energy loss involves now additional contributions due to the emission of helicity-0 (scalar) waves, in particular a dipolar radiation term which starts at order $O(c^{-3})$, a monopolar term of order $O(c^{-5})$, and a new quadrupolar term of order $O(c^{-5})$.

The observation of binary pulsars provide a unique opportunity for testing the strong-field regime of gravity, since pulsars are neutron stars whose compactnesses GM/Rc^2 are typically of order 0.2 (as compared to 10^{-6} for the Sun). In particular, the system PSR 1913+16 has been continuously observed since its discovery[16] in 1974, and it is now possible to determine very precisely the values of three observable quantities : (i) the time variation $\dot{P}$ of the orbital period [fractional accuracy of 0.4%], (ii) the periastron shift $\dot{\omega}$ [fractional accuracy of 4×10^{-6}], and (iii) the time-dilation parameter γ which measures both the Doppler shift due to the motion of the pulsar and the red-shift due to the presence of its companion [fractional accuracy of 7×10^{-4}]. The comparison of these experimental data with the predictions of a given theory leads to the so called "$\dot{P}$-$\dot{\omega}$-γ test", that general relativity passes with flying colors. However, we are going to illustrate on an example that it can also be passed by a class of particular tensor-multi-scalar theories. Let us consider the tensor-bi-scalar theories introduced at the end of section 2, which satisfy all solar-system tests since they have the same post-Newtonian limit as general relativity. We shall restrict our attention to those which do not predict any dipolar radiation of order $O(c^{-3})$, because this dominant term has already been studied in the literature[17] and we want to focus on the new phenomena related to tensor-multi-scalar theories. The simplest non-trivial

class of such theories corresponds to the coupling function

$$\ln A(\varphi_1, \varphi_2) = \varphi_2 \left(1 - \varphi_1 + \mu\varphi_1^2\right) + \frac{1}{3!}\beta'\varphi_1^3 + \frac{1}{4!}\beta''\varphi_1^4 \,, \qquad (15)$$

where the constant $\mu \approx 1.1$ is a pure number which has been chosen in order to suppress the dominant dipolar radiation, and which can be expressed in terms of the integral $\int_0^\pi (\sin^3 x/x)\mathrm{d}x$. The real parameters β' and β'' entering $\ln A$ are strong-field analogues of the usual PPN parameters [they are in fact second and third post-Newtonian parameters respectively]; general relativity corresponds to the values $\beta' = \beta'' = 0$. We show in reference 14 that the $\dot{P}$-$\dot{\omega}$-γ test constrains these two parameters to a long strip in the (β', β'') plane, which includes general relativity but also a continuous set of alternative theories. To illustrate how much they can differ from Einstein's theory, let us quote that in the model corresponding to $\beta' = 6$ and $\beta'' = 36$, the effective gravitational constant $\tilde{G}_{AB}$ between the two neutron stars is 47% larger than the Newtonian gravitational constant $\tilde{G}$ measured in the solar system. This model predicts therefore very large deviations from general relativity in the strong field regime, although it is consistent with all solar-system experiments as well as with the $\dot{P}$-$\dot{\omega}$-γ test in the binary pulsar PSR 1913+16. This is due to the fact that this test mixes in an indistinct manner quasi-stationary strong field effects ($\dot{\omega}$, γ) and radiative ones ($\dot{P}$). It is therefore necessary to combine several tests, having different physical content, to really constrain the space of allowed relativistic theories of gravity[18,19]. See the contribution of Thibault Damour to the present proceedings for a discussion of the maximum number of tests which can be extracted from binary pulsar data[20].

References

1. Will C M 1981 *Theory and experiment in gravitational physics* (Cambridge University Press)
2. Shapiro I I 1990 in *General Relativity and Gravitation, 1989,* edited by Ashby N, Bartlett D F and Wyss W (Cambridge University Press) 313–30
3. Will C M and Nordtvedt K 1972 *Astrophys. J.* **177** 757
4. Jordan P 1949 *Nature* **164** 637
   ——1955 *Schwerkraft und Weltall* (Vieweg, Braunschweig)
5. Fierz M 1956 *Helv. Phys. Acta* **29** 128
6. Jordan P 1959 *Z. Phys.* **157** 112
7. Brans C and Dicke R H 1961 *Phys. Rev.* **124** 925
8. Nordtvedt K 1968 *Phys. Rev.* **169** 1014 and 1017
9. Nordtvedt K 1968 *Phys. Rev.* **170** 1186
10. Eardley D M 1975 *Astrophys. J.* **196** L59
11. Will C M and Eardley D M 1977 *Astrophys. J.* **212** L91
12. Will C M 1977 *Astrophys. J.* **214** 826
13. Reasenberg R D *et al.* 1979 *Astrophys. J.* **234** L219
14. Damour T and Esposito-Farèse G 1992 *Class. Quant. Grav.,* in press
15. Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (Freeman)
16. Hulse R A and Taylor J H 1975 *Astrophys. J.* **195** L51
17. Will C M and Zaglauer H W 1989 *Astrophys. J.* **346** 366
18. Damour T and Schäfer G 1991 *Phys. Rev. Lett.* **66** 2549
19. Taylor J H, Wolszczan A, Damour T and Weisberg J M 1992 *Nature* **355** 132
20. Damour T and Taylor J H 1992 *Phys. Rev. D,* in press

NEW TESTS OF GENERAL RELATIVITY

Thibault Damour
Institut des Hautes Etudes Scientifiques
91440 Bures sur Yvette, France
and DARC, Observatoire de Paris–CNRS
92195 Meudon Cedex, France

Abstract

Until now, most experiments have succeeded in testing relativistic gravity only in its extreme weak–field limit. Because pulsars are neutron stars, with surface gravitational potentials $GM/c^2R = 0.2$, physical systems including pulsars necessarily involve strong–field conditions. Recent work showing how to extract strong–field tests from binary pulsar data is summarized. The application to actual experimental data is presented, including in particular the obtention of two new strong–field tests from the timing data of the recently discovered system PSR 1534+12.

Up to now, most tests of general relativity have been performed in the solar system, i.e. in conditions where the gravitational field is extremely weak and slowly changing. Some of these tests (notably relativistic effects in the motion of planets, delay of electromagnetic signals passing near the Sun and limits on the violation of the strong equivalence principle in the Earth–Moon system) have reached a high accuracy, and have shown that the "post–Newtonian" effects (fractional corrections of order $(v/c)^2 \sim GM/c^2r$ to the Newtonian description of gravity) agree with the general relativistic predictions within a fractional accuracy of about 2×10^{-3} (for reviews see [1]).

However, these results tell us something about the structure of the gravitational interaction only in the limit of weak and quasi–stationary fields (the largest relativistic self potential in the solar system is $GM_\odot/c^2R_\odot = 2 \times 10^{-6}$, and the ratio of time derivatives to space derivatives of the gravitational field is of order $v/c \sim 10^{-4}$). Apart from solar system tests, we have some constraints on the large spatio–temporal scale behaviour of gravity from cosmological data. But the corresponding tests are either of poor accuracy or highly model dependent.

The discovery in 1974 [2] of the binary pulsar PSR 1913+16 opened up an entirely new testing ground for relativistic gravity. For the first time, it was possible to have data on systems which comprise strong gravitational fields. Indeed, the relativistic self gravitational potential of a pulsar (i.e. a spinning neutron star) is about $GM/c^2R = 0.2$, which is not very far from the black hole limit $(GM/c^2R)_{\rm BH} = 0.5$. Since 1979, the timing data on PSR 1913+16 have allowed one to measure three independent observables (beyond the easily measured "Keplerian" parameters: orbital period P_b, eccentricity e and projected semi–major axis $x \equiv a_1 \sin i/c$). These three "post–Keplerian" parameters are: the secular advance of the periastron $\dot{\omega}$, a time dilation parameter γ, and the secular change of the orbital period $\dot{P}_b$. Now, any given relativistic theory of gravity makes a specific prediction for the values of $\dot{\omega}$, γ and $\dot{P}_b$ as functions of the well–measured Keplerian parameters and of the (a priori unknown) inertial masses, m_1 and m_2, of the pulsar and its companion. In graphical terms, the measurement of the three post–Keplerian parameters $\dot{\omega}$, γ and $\dot{P}_b$ defines (when interpreted within the framework of a specific theory of gravity) three different curves in the m_1, m_2 plane. Thereby one gets *one* test of the theory, according to whether the three curves meet at one point, as they should. General relativity passes this $\dot{\omega} - \gamma - \dot{P}_b$ test with flying colors [3] [4] (present accuracy 5×10^{-3} [5]). This test is our first probe of strong–field (in $\dot{\omega}$ and γ) and radiative (in $\dot{P}_b$) effects in relativistic gravity. However, it mixes both types of effects in an indinstinct way. In fact, as recently shown [6], one can construct examples of alternative gravity theories that pass the $\dot{\omega} - \gamma - \dot{P}_b$ test, as well as all known tests of gravity in the weak–field limit, while differing substantially from general relativity in the strong–field regime.

This unsatisfactory situation raises the following question: Is it possible to extract other tests of relativistic gravity from binary pulsar data, especially tests that would probe the quasi–stationary strong–field aspects of the gravitational interaction, without mixing of radiative effects ? This question is made especially timely by the recent discovery of two new "relativistic" binary pulsars with short period, highly eccentric

orbits, namely PSR 2127+11C [7] (which however is timable with unsufficient accuracy to provide new tests), and PSR 1534+12 [8] (the latter system is timed with better accuracy than PSR 1913+16 and is therefore of exceptional interest). Recently Damour and Taylor [9] have addressed the above question. They have presented a general phenomenological framework, the "parametrized post–Keplerian" (PPK) formalism, designed to extract the maximum possible information from pulsar timing and pulse–structure data. In brief, they show that as many as 19 (post–Keplerian) phenomenological parameters can be extracted, under favourable conditions, from binary pulsar measurements. These 19 observables can give access to 15 tests of relativistic gravity. [The difference $19 - 15 = 4$ comes from subtracting four, a priori unknown, kinematical parameters; the two inertial masses m_1, m_2 and the two Euler angles, λ, η describing the direction of the pulsar spin axis]. Ref. [9] discussed also the theoretical significance of these new tests by deriving, within the framework of generic boost–invariant theories, expressions linking the observable post–Keplerian parameters to m_1, m_2, λ and η. In generic, non–Einsteinian, theories these expressions will differ from the corresponding general relativistic expressions because of strong–field effects. A recent detailed study of the predictions of tensor–multi–scalar theories of gravitation [6] has succeeded in working out explicitly the influence of the strong–self–gravity of the pulsar and its companion on the observable post–Keplerian parameters.

Having shown that there are in principle many interesting new tests extractable from binary pulsar data, the next question is to know whether these tests can really be performed with the present noise–limited pulsar data. Ref. [9] showed that the recently discovered system PSR 1534+12 should, with presently available data, give access to two new strong–field tests of relativistic gravity. This prediction has been fully confirmed by the work of Ref. [5] which, indeed, succeeded in extracting 4 post–Keplerian parameters from the 1534+12 data. These parameters are: the secular advance of the periastron $\dot{\omega}$, the time dilation parameter γ, and the "range" r and "shape" s of the Shapiro time delay caused by the companion. All four parameters depend only on m_1 and m_2 (and not on λ and η). Therefore, simultaneous measurement of $\dot{\omega}$, γ, r and s provide already $4 - 2 = 2$ new tests of strong–field gravity. At present, the accuracy of these new tests is somewhat low ($\sigma_r/r = 21$ %), but it should steadily improve as more data are accumulated. Moreover, numerical simulations carried out in Ref. [9] show that extended and/or improved observations of both PSR 1913+16 and PSR 1534+12 should give access in the future to further post–Keplerian (PK) parameters: In 1913+16, one should be able (with improved data) to measure 3 new separate PK parameters (r, s and δ_θ), while in 1534+12 one should be able (with more data of the present quality) to measure $\dot{P}_b$, as well as to detect the effects of the relativistic spin–orbit coupling.

Finally, by combining the phenomenological approach of Ref. [9] with the alternative–theory approach of Ref. [6], one can nicely discuss the theoretical content of any of the above tests in terms of constraints on allowed deviations away from the "correct" theory of gravity. More precisely, the work of [6] show that (under some conditions) it is possible to describe strong–field effects in generic tensor–multi–scalar theories by means of a sequence of "theory parameters":

$$\gamma_1,\ \beta_1\ ;\ \beta_2,\ \beta',\ \beta'',\ \beta_3,\ (\beta\beta'),\ldots \tag{1}$$

The first two parameters in the list (1) are equivalent to the well–known (weak–field) post–Newtonian parameters $\gamma - 1$ and $\beta - 1$. [All parameters (1) are normalized so as to vanish in general relativity]. The further parameters β_2, β', β'',... represent deeper layers of structure of the relativistic gravitational interaction which have been left unprobed by solar system tests. In pictorial terms, each parameter in the list (1) represents an independent direction away from general relativity in a space of alternative theories of gravitation, and the theory parameters β_2, β', β'',... provide a chart for the domain of strong gravitational fields. A first, two–dimensional exploration of this domain has been pioneered by the work of Ref. [5]. There, use has been made of a specific two–parameter class of tensor–bi–scalar theories, $T(\beta', \beta'')$, introduced in Ref. [6], to explore the directions associated with the parameters β' and β''. The β', β'' plane provides a common ground within which one can interpret data coming from several different pulsars. More precisely, the data from each pulsar define a certain allowed region within the β', β'' plane, namely the subset of theories which are compatible with the data at the, say, 90 % confidence level. Ref. [5] used data from three different pulsars: PSR 1913+16, PSR 1534+12 and PSR 1855+09. (For the latter, they made use of an idea of Ref. [10] in which it was pointed out that the observation of a low–eccentricity long–period binary pulsar with a white–dwarf companion sets a probabilistic limit on the difference in free–fall acceleration of a neutron star and a white dwarf in the galactic field). The three pulsars define three different allowed regions in the β', β'' plane: (i) a thin (nearly parabolic) strip corresponding to the single (0.5 %–accurate) $\dot{\omega} - \gamma - \dot{P}_b$ test in 1913+16, (ii) a wide potato–shaped region corresponding to the two new (low precision) $\dot{\omega} - \gamma - r - s$ tests in 1534+12, and (iii) the vertical strip $-1.6 < \beta' < +1.5$ [6] corresponding to the $e - P_b$ test in 1855+09. When combining these three independent allowed regions one reaches the following conclusions:

a. The three allowed regions do admit a non empty common intersection, and general relativity [i.e. the point $(\beta', \beta'') = (0, 0)$] lies well inside this intersection region.

b. At the 90 % confidence level the theory parameters β' and β'' are constrained to lie in a thin parabolic segment whose projections on the β', β'' axes are roughly $-1.1 < \beta' < 1.5$, $-1 < \beta'' < 6$.

In conclusion, general relativity has passed with complete success several new, deep and sensitive experimental tests. Among these tests the ones associated with the pulsar data PSR 1534+12 and PSR 1855+09 concern the quasi–stationary strong–field regime of relativistic gravity, without mixing of radiative effects.

References

[1] C.M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, Cambridge, 1981); "The confrontation between general relativity and experiment: a 1992 update", Int. Journ. Mod. Phys. A (1992), in press.

[2] R.A. Hulse and J.H. Taylor, Astrophys. J. (Letters) **195**, L51 (1975).

[3] J.H. Taylor and J.M. Weisberg, Astrophys. J. **345**, 434 (1989).

[4] T. Damour and J.H. Taylor, Astrophys. J. **366**, 501 (1991).
[5] J.H. Taylor, A. Wolszczan, T. Damour and J.M. Weisberg, Nature **355**, 132 (1992).
[6] T. Damour and G. Esposito–Farèse, Class. Quantum Grav. (1992), in press; see also the contribution of G. Esposito–Farèse to these Proceedings.
[7] S.B. Anderson et al., Nature **346**, 42 (1990).
[8] A. Wolszczan, Nature **350**, 688 (1991).
[9] T. Damour and J.H. Taylor, Phys. Rev. D. **45**, 1840 (1992).
[10] T. Damour and G. Schäfer, Phys. Rev. Lett. **66**, 2549 (1991).

Status of the LAGEOS III experiment to measure the gravitomagnetic field

Ignazio Ciufolini
CNR-IFSI-Via G. Galilei-CP 27
00044 Frascati, Italy

Abstract

The LAGEOS III experiment will test the existence of the so-called gravitomagnetic field, predicted by the Einstein general theory of relativity but, so far, never directly measured. The experiment will use LAGEOS, Laser Geodynamics Satellite, together with LAGEOS III, a non-polar, laser-ranged satellite with the same orbital parameters of LAGEOS but a supplementary inclination. In this paper we briefly describe the project and its present status.

LAGEOS III

In 1917-18 Thirring and Lense showed the existence, near a rotating body, of a "new" gravitational field due to the rotation of the body and later called, by its formal analogy with electromagnetism, the gravitomagnetic field. They also derived the "Lense-Thirring nodal drag" of a test particle orbiting a rotating central body; of course this effect is not present in the classical gravity theory. Since then, many experiments have been discussed and proposed in order to measure the gravitomagnetic field.

• In 1984 it was proposed [1] to measure the gravitomagnetic field using two non-polar laser ranged satellites, of LAGEOS type, with supplementary inclination. LAGEOS (Laser Geodynamics Satellite) is a high altitude, small cross sectional area-to-mass ratio, spherical, laser-ranged satellite, launched by NASA in 1976. Its semimajor axis is $a = 12270$ Km and the inclination $I = 109.94$. The accuracy in tracking its orbit is of about 1 cm over 5900 Km of altitude. The LAGEOS gravitomagnetic, Lense-Thirring, nodal precession is $\dot{\Omega}^{LT}_{LAGEOS} = 31\, milli\ arcsec/year$. Position measurements of laser ranged satellites, of LAGEOS type, are accurate enough to detect this very tiny gravitomagnetic precession. Unfortunately, the Lense-Thirring precession cannot be extracted from the measured value of the LAGEOS nodal precession, because of the uncertainty in the theoretical classical precession due to the quadrupole and higher mass moments of Earth: $\dot{\Omega}^{Classical}_{LAGEOS} \cong 126°/year$. However, a new satellite, of LAGEOS type: LAGEOS-III, [1], with supplementary inclination: $I_{LAGEOS\ III} \cong 70°$, will have a classical precession equal in magnitude and opposite in sign to that of LAGEOS. By contrast, since independent of the inclination, the Lense-Thirring precession will be the same, both in magnitude and sign for the two satellites. Therefore, from the measured nodal precessions of LAGEOS and LAGEOS III will be possible to measure the Lense-Thirring effect.

An important problem in the LAGEOS III experiment was to identify and to quantify all the error sources that may affect the measurement of the Lense-Thirring precession.

• In 1988, a preliminary error analysis [2], [3] identified a list of potential error sources and gave a preliminary error budget. Among the potential error sources of the LAGEOS III experiment there

are: orbital injection errors; errors from uncertainties in the harmonics expansion of Earth potential, in particular from uncertainties in the dynamical part of Earth field, i.e. in the modeling of solid and ocean Earth tides; errors from non-gravitational perturbations: direct solar radiation pressure, Earth albedo, satellite eclipses, anisotropic thermal radiation, infrared radiation, atmospheric drag...; errors from the uncertainties in the determination of the LAGEOS orbital parameters and in particular in the determination of the inclination I and of the nodal longitude Ω, relative to an asymptotic inertial frame.

This preliminary error analysis [2] showed that an upper bound to the total statistical error, over the period of the node of $\simeq 3$ years, was about 10% of the Lense–Thirring effect to be measured.

• In May 1988, NASA and ASI, Italian Space Agency, formed study groups for the purpose of performing a comprehensive analysis of the experiment and a comprehensive numerical simulation. We summarize here the main results of the 1989 ASI–NASA study [4], [5].

• Neutral and charged particle drag explains part of the LAGEOS semimajor axis decay. To study the effect on the LAGEOS node worst possible cases were considered. The result of the investigation was that, even under the most extreme conditions, the secular nodal drag due to particle drag will never exceed a few parts in 10^{-3} of the Lense–Thirring effect.

• The LAGEOS nodal perturbations due to anisotropic re-emission of Earth infrared radiation and re-emission of sunlight modulated by satellite eclipses by Earth were studied. Infrared radiation from Earth is absorbed by the LAGEOS retroreflectors, and due to their thermal inertia, the re-emission of the heat causes an along-track acceleration opposite to the satellite motion. This effect explains part of the average secular decrease of the LAGEOS semimajor axis of about 1 mm per day. The corresponding effect on the node was investigated and it was found a maximum cumulative secular nodal precession of about 2% of the Lense–Thirring effect, corresponding to the worst possible case of spin axes orientation of LAGEOS and LAGEOS III. In the case of heating by sunlight modulated by eclipses, the largest secular and long periodic terms can be at most $\simeq 1\%$ of the Lense–Thirring effect, corresponding to the worst values of some thermal parameters considered. These figures refer to the maximum effect of the thermal thrust; however: the orientation and the rotation rate of the spin of LAGEOS and LAGEOS III can be measured and modeled and the knowledge of thermal and optical properties can be substantially improved through various measurements. Therefore, because of partial cancellations of the overall effect, over a period of a few years, the sum of the unmodeled secular nodal drifts of LAGEOS and LAGEOS III due to thermal thrust should be less than 1% of the Lense–Thirring effect.

• Concerning Earth's albedo, the error in the Lense–Thirring measurement, due to Earth-reflected radiation pressure on LAGEOS and LAGEOS III, would not exceed a few percent of the gravitomagnetic effect. However, improvements are achievable by modeling Earth's albedo. Furthermore, many of the terms in the expression of the nodal precession due to Earth albedo give, for the two satellites, nodal precessions equal in magnitude but opposite in sign, thus substantially reducing the uncertainty due to Earth's albedo. Similarly, concerning the LAGEOS nodal precession from direct radiation pressure due to Earth infrared radiation. Since the main latitudinal effect is equal in magnitude but opposite in sign for LAGEOS and LAGEOS III and since the effect is small, the error in modeling the direct effect of Earth's infrared radiation on the LAGEOS node is substantially negligible.

• All the other non gravitational effects are negligible in comparison with the Lense–Thirring effect.

• The study confirmed that the largest source of error is due to uncertainties in the harmonics expansion of Earth potential, in particular to its dynamical part, i.e. to solid and ocean Earth tides. The geopotential error was estimated to be at the most 5% of the Lense–Thirring effect.

• A covariance analysis [5] confirmed a $1-\sigma$ error of 8% or less, considering a (large) injection error of 0.1° in the inclination and with no improvements in the knowledge of the various parameters and in the measurement errors.

• Blind tests [5] have given a root-mean square difference between the values arbitrarily assumed for the frame dragging effect and the values recovered from the simulated data of 8%.

• Therefore, the covariance analysis and the blind tests have confirmed the previous [2] comprehensive error analysis of a maximum error of 10% or less.
• In 1991 McDonnell Douglas reported a study on the injection accuracy achievable in the orbit of LAGEOS III using a special Delta II rocket: the $3-\sigma$ injection error corresponds to an error of less than 1% in the LAGEOS III experiment; furthermore the actual error of the past Delta II two-stage orbits was always well within the $1-\sigma$ predicted error; therefore the injection errors should be negligible.
• To minimize the errors in modeling the thermal thrust, accurate measurements of the LAGEOS spin axis orientation and spin rate are currently being studied; they may be done using three different methods: microwave Doppler, infrared Doppler, and optical glints from front reflectors surfaces. Accurate measurements of thermal and optical properties of LAGEOS-type satellites have also been studied.
• Therefore, considering improvements due to the modeling of thermal thrust and radiation pressure effects from LAGEOS III-LAGEOS I thermal, optical and spin orientation measurements; improvements in the knowledge of Earth gravity field and of Earth solid and ocean tides from the data of various recent and near future Earth-satellites; that the combination of data from LAGEOS I and LAGEOS III will be supplemented by the near future LAGEOS II data; and by analyzing data corresponding to periods multiple than three years (the nodal period), since the previous figures (8%) were obtained corresponding to a period of 3 years only and by assuming no future improvements (however the accuracy in the modeling of the various perturbations and in the measurements with laser ranging and VLBI are improving steadily), it should be possible, over an interval corresponding to a few nodal periods, a measurement of the Lense-Thirring effect with an error of 4% or less.
• The LAGEOS III satellite will be launched with a Delta II rocket (in a NASA manifest the launch was scheduled in the second quarter of 1994) in collaboration beween ASI, Italian Space Agency, NASA and US Air Force.

References

1. I. Ciufolini, *Phys. Rev. Lett* **56**, 278 (1986), submitted October 1984.
2. I. Ciufolini, *Int. J. of Mod. Phys. A*, **4**, 3083 (1989), submitted April 1988.
3. I. Ciufolini, in *NASA Workshop on Relativistic Gravitation Experiments in Space*, Annapolis, Maryland, June 1988, NASA Conference Publication 3046.
4. I. Ciufolini et al. *NASA/ASI study on LAGEOS III* Appendix **B** , (30 September 1989).
5. B. Tapley and I. Ciufolini, *Results of a joint NASA/ASI study on LAGEOS III, CSR 89-3* (30 September 1989), see also: J. C. Ries, R. J.Eanes, M. Watkins and B. D. Tapley, *NASA/ASI study on LAGEOS III* appendix **A** , (30 September 1989).

Tests of Newtonian Gravity

THE FIFTH FORCE IN ELECTROMAGNETISM AND GRAVITY

Harry Kloor, Ephraim Fischbach, and Carrick Talmadge

Physics Department, Purdue University, West Lafayette, IN 47907, USA

ABSTRACT

Recent interest in the "fifth force" has focused on new intermediate range interactions whose presence could lead to apparent deviations from the predictions of Newtonian gravity. Here we draw attention to the possibility that the *same* forces could also lead to apparent anomalies in electromagnetic interactions, for appropriate values of the range and coupling constant. We contrast the phenomenology of searches for gravitational and electromagnetic signals for the fifth force, and we present new experimental limits for each case.

I. Introduction

There is at present considerable experimental and theoretical interest in the possibility of deviations from Newton's law of gravitation,[1–4] motivated in part by the suggestion of a possible "fifth force".[5] The similarity between Newton's law and Coulomb's law naturally leads to the question of whether there may exist an electromagnetic analog of the "fifth force", a possibility which has been raised recently by Bartlett and Lögl (BL).[6] Our viewpoint in the present paper is that there may exist a new interaction mediated by an ultralight boson, whose mass μ is sufficiently small that $\lambda = \hbar/\mu c$ is of macroscopic dimensions, and which has no intrinsic connection to either gravity or electromagnetism. However, for appropriate values of λ and coupling strength f, this interaction can "compete" with either gravity or electromagnetism, or with both simultaneously, depending on the characteristic distance scale of the experiment. Thus by combining data from both gravitational and electromagnetic systems we can obtain limits on f for values of λ that extend from the astrophysical scale ($\lambda \cong 10^{12}$ m), to the subatomic scale ($\lambda \cong 10^{-15}$ m).

In order to simultaneously study the limits arising from electromagnetic and gravitational interactions, we require a uniform parametrization of the effects of the new couplings that we are probing. In the simplest theories these couplings arise from the exchange of scalar and vector fields, and the phenomenology of the interactions that arise from such fields has been intensively studied in gravitational systems. However, an analogous framework for electromagnetism does not yet exist, and so we begin by developing the formalism needed to describe possible deviations from Coulomb's law and other electromagnetic phenomena. We assume that the new fields have a small non-zero mass, so that the potential energy of two interacting particles 1 and 2 will have the typical Yukawa behavior $\sim exp(-r/\lambda)/r$ where $r = |\vec{r}_1 - \vec{r}_2|$ is the separation of the particles. One possibility is that the new field is a vector ($J^P = 1^-$) which couples to electric charge, in which case the resulting interaction is the well known theory of "massive electrodynamics".[7] The potential energy of two charges in the combined presence of this interaction and conventional electrodynamics is then given by

$$V(r) = \frac{Q_1 Q_2 e^2}{r}(1 + \beta e^{-r/\lambda}) \equiv V_C + \Delta V, \tag{1}$$

where $V_C = Q_1 Q_2 e^2/r$. In Eq. (1) Q_1 and Q_2 are the charges of the two objects (in units of e), and $\beta = f^2/e^2$ is a dimensionless constant which fixes the strength of the new interaction in terms of e^2, where $e^2/\hbar c \equiv \alpha_{em} \cong 1/137$ is the usual electromagnetic fine structure constant.

Another possibility is that the new field is a scalar ($J^P = 0^+$). As we discuss below, the phenomenology of an interaction based on scalar exchange is of interest because of the rich variety of possibilities for the effective charge (i.e., the analog of Q) in such theories. To construct the simplest version of such a theory we consider the Lagrangian

$$\mathcal{L} = (f_p \bar{\psi}_p \psi_p + f_n \bar{\psi}_n \psi_n + f_e \bar{\psi}_e \psi_e)\phi, \tag{2}$$

where ϕ is the scalar field and f_p, f_n, f_e are real constants which describe the coupling of this field to protons, neutrons, and electrons respectively. Apart from the reality of these constants, there are no other *a priori* constraints that we need impose. In particular, the product $f_e f_p$ can be either positive or negative without affecting the Hermiticity or Lorentz invariance of the coupling in Eq. (2). If a and b denote any of the species p, n, e then the potential energy arising from the interaction of a and b via ϕ-exchange has the form

$$V_{ab} = -f_a f_b \frac{e^{-r/\lambda}}{r}, \tag{3}$$

where $\lambda = \hbar/\mu c$ and μ is the ϕ mass. Note that the sign of V_{ab} depends on the product $f_a f_b$, but not on whether a is a particle or the corresponding antiparticle, as is evident from the Lorentz structure of the coupling in Eq. (2).

We now study the phenomenological implications of various choices for f_p, f_n, and f_e. Consider for example the case

$$f_p = f_e \equiv f, \quad f_n = 0. \tag{4}$$

Then,

$$\mathcal{L} = f(\bar{\psi}_p \psi_p + \bar{\psi}_e \psi_e)\phi. \tag{5}$$

The charge Q' which determines the strength of the coupling of ϕ to a sample of bulk matter is then given by

$$Q' = f(Z - L), \tag{6}$$

where Z and L denote the number of protons and electrons respectively, with $L = +1$ for e^- and -1 for e^+. Another simple possibility is

$$f_p = -f_e \equiv f, \quad f_n = 0, \tag{7}$$

which corresponds to

$$Q' = f(Z + L). \tag{8}$$

The expression for Q' in Eq. (6) is interesting, because it behaves as the conventional electric charge Q for ordinary matter. However, this charge does not reverse sign for a coupling to antimatter, in contrast to the electromagnetic charge, so that the $e^+ - p$ and $e^- - p$ forces would have the same sign for the scalar coupling.

Evidently there are many other choices for the constants f_p, f_n, and f_e which lead to interesting phenomenological models. What is important about all of these models, as well as "massive electrodynamics" discussed above, is that they are well-defined physical theories, in

which unambiguous predictions can be made for the outcome of any experiment. Moreover, by assuming that these interactions coexist with conventional Newtonian gravity and conventional electrodynamics, one arrives at a framework in which the interpretation of experimental results is quite different from what it would be in the alternative "single component" picture. This is the framework in which it is assumed that Coulomb's law for the interaction of charges 1 and 2 has the form

$$\vec{F}(r) = \frac{Q_1 Q_2 e^2 \hat{r}}{r^{2+\delta}}, \tag{9}$$

where δ is a parameter to be determined from experiment. In contrast to the present formalism, and that of BL, there is little (if any) theoretical motivation for supposing that deviations from Coulomb's law would be of the form given in Eq. (9). We will henceforth assume the parametrization in Eq. (1) for the case of massive electrodynamics, and a similar form for scalar exchange. In the latter case we write specifically,

$$V(r) = \frac{Q_1 Q_2 e^2}{r} - \frac{Q'_1 Q'_2 f^2 e^{-r/\lambda}}{r}, \tag{10}$$

where $Q_{1,2}$ are the usual electromagnetic charges, and $Q'_{1,2}$ are the scalar charges, as in Eqs. (5) and (9). It is convenient to recast Eq. (10) in a form similar to Eq. (1) by writing

$$V(r) = \frac{Q_1 Q_2 e^2}{r}(1 + \beta_{12} e^{-r/\lambda}), \tag{11a}$$

$$\beta_{12} \equiv -\left(\frac{Q'_1}{Q_1}\right)\left(\frac{Q'_2}{Q_2}\right)\frac{f^2}{e^2}. \tag{11b}$$

We can thus use Eq. (1) generically (for all models of the new interaction), bearing in mind that β might actually depend on the detailed compositions of the test masses 1 and 2 through $Q_{1,2}$ and $Q'_{1,2}$. For the case of a new coupling coexisting with gravity we use the conventional parametrization,

$$V(r) = \frac{-G_\infty m_1 m_2}{r}(1 + \alpha_{12} e^{-r/\lambda}). \tag{12}$$

Here $m_{1,2}$ are the masses of the test objects, and α_{12} is given by

$$\alpha_{12} \equiv -\left(\frac{Q'_1}{\mu_1}\right)\left(\frac{Q'_2}{\mu_2}\right)\frac{f^2}{G_\infty m_H^2}, \tag{13}$$

where G_∞ is the Newtonian gravitational constant, $\mu_{1,2} = m_{1,2}/m_H$, and $m_H = m({}_1H^1)$ is the mass of atomic hydrogen. For many applications the relevant dynamical quantity is the force $\vec{F}(r)$ rather than the potential, and from Eq. (1) we have

$$\vec{F}(r) = -\vec{\nabla}V(r) = \frac{Q_1 Q_2 e^2 \hat{r}}{r^2}[1 + \beta(1 + r/\lambda)e^{-r/\lambda}]. \tag{14}$$

In analyzing various experimental tests for the presence of the additional terms in Eq. (14), it is useful to exhibit the expression for $\vec{F}(r)$ in two limiting cases:

$$\vec{F}(r \gg \lambda) = \frac{Q_1 Q_2 e^2}{r^2}\hat{r}, \tag{15a}$$

$$\vec{F}(r \ll \lambda) = \frac{Q_1 Q_2 e^2}{r^2}(1+\beta)\hat{r}. \tag{15b}$$

We see that in both of these extreme cases Coulomb's law holds exactly, albeit with different values of the electric charge, namely e and $e(1+\beta)^{1/2}$ respectively. The fact that the usual $1/r^2$ law is recovered in both limiting cases has important experimental consequences. For example, it follows from Eqs. (15a) and (15b) that tests for deviations from Gauss' law lose their sensitivity when the characteristic size of the apparatus is either much bigger or much smaller than the assumed value of λ. Since this is not the case for the conventional single component parametrization in Eq. (9), we see that the interpretation of a given experiment in the present formalism can be quite different from what is was earlier.

Thus far our discussion of deviations from Coulomb's law has paralleled that used to describe possible deviations from Newton's law of gravitation arising from the putative "fifth force". However, there are important differences between electromagnetism and gravity, notwithstanding the formal similarities between Coulomb's and Newton's laws. To illustrate some of these we note from Eqs. (15a) and (15b) that the presence of the additional interaction (which is proportional to β) can be inferred by exhibiting two systems of different size having different fine structure constants, $e^2/\hbar c$ and $e^2(1+\beta)/\hbar c$ respectively. The corresponding gravitational experiment, however, could not be used in the same way to infer the coupling strength α (the analog of of β) for the fifth force. This can be seen by noting that the overall strength of the gravitational interaction for two objects 1 and 2 is determined by the product $G_\infty m_1 m_2$. In the presence of a fifth force, the analog for gravity of Eq. (15b) would be proportional to the factor $G_\infty(1+\alpha)m_1 m_2$. It follows that in order to distinguish between G_∞ and $G_\infty(1+\alpha)$ one must know m_1 and m_2 independently of G_∞, and this is difficult to do for large gravitational sources. This problem is circumvented in electromagnetism because charge is ***quantized***: Q_1 and Q_2 are known to be ***integers*** (in units of e), whereas m_1 and m_2 are unconstrained. This simple observation has important experimental consequences.

The preceding observation points out another fundamental difference between tests of Coulomb's law and Newton's law. For electromagnetism quantum systems exist which can be used to test Coulomb's law over distances from approximately 10^{-13} cmto 10^{-6} cm. By contrast, gravitational effects in quantum systems are generally too small to be detected, the notable exception being neutron interferometer experiments. Even for such such experiments the smallest distance scales that are probed are on the order of cm, which is the characteristic size of the interferometer used. Stated another way, what allows gravity to be detected in these systems in the first place is that they are essentially macroscopic. However, over the same

distance scales other macroscopic detectors (e.g. torsion balances) provide more stringent tests of Newton's law. Although numerous suggestions have been advanced over the years to study gravitational effects over smaller distance scales, none of these has proved practicable to date. The fact that tests of Coulomb's law can be carried out in atomic systems leads to still another practical difference from gravity: Whereas the classical orbit of a planet about the Sun provides information for a limited range of distance scales, the quantum motion of an electron about the nucleus probes a much broader region, due to the quantum-mechanical " smearing" of the electron wavefunction. Since the electron has a nonzero probability of being at the nucleus, atomic systems can usefully test Coulomb's law down to distances of order 10^{-13} cm, whereas the motion of a satellite or planet tests Newton's law only in a much narrower range of distances away from its semimajor axis.

It follows from the preceding discussion that, where appropriate, a full quantum mechanical treatment of the non-Coulomb contribution ΔV will likely give results which are qualitatively different from those obtained from the semiclassical treatment of BL. Indeed, one of the implications of this argument is that the quantum treatment should provide more stringent constraints on β by virtue of the fact that a quantum mechanical electron can probe regions that are not accessible classically. One of the motivations for the present paper is to focus on the new features that arise in searches for an electromagnetic fifth force, when the quantum behavior of the new interaction is fully incorporated.

In what follows we will present limits not only for the Yukawa model of Eq. (1), but also for the related exponential model.[8] This model is of interest because it describes the case of two nearly-cancelling Yukawas, which is a theoretically attractive alternative to the case of a single Yukawa. In the exponential model, the potential energy $V_E(r)$ assumes the form

$$V_E(r) = \frac{Q_1' Q_2' f^2 e^{-r/\lambda}}{\lambda}, \tag{16}$$

which contrasts to the conventional Yukawa potential $V_Y(r)$ upon which Eq. (1) is based. In the combined presence of electromagnetism and such a new force, the potential $V(r)$ can be written in the form

$$V(r) = \frac{Q_1 Q_2 e^2}{r}[1 + \beta \frac{r}{\lambda} e^{-r/\lambda}] \tag{17}$$

and the corresponding force is given by

$$\vec{F}(r) = -\vec{\nabla} V(r) = \frac{Q_1 Q_2 e^2}{r^2} \hat{r}[1 + \beta \frac{r^2}{\lambda^2} e^{-r/\lambda}]. \tag{18}$$

In Sec. II below we illustrate the method of obtaining limits on β from an analysis of hydrogenic systems. In Sec. III we summarize the constraints obtained from this and other systems, the details of which will be presented elsewhere.

II. Experimental Constraints on Deviations From Coulomb's Law

As noted in the Introduction, we wish to consider limits on the strength parameter β in Eq. (1), which describes the deviations from Coulomb's law. This can be done with some precision over the atomic, laboratory, and planetary scales. We illustrate our methodology by examining the constraints on Coulomb's law at the atomic scale which follow from the consistency of different determinations of the Rydberg constant, $\mathcal{R}$. Details on inferring the constraints for other distance scales will be published elsewhere. For a Coulomb potential, the theoretical expression for the energy $\Delta\mathcal{E}_0(n,m)$ released when an electron in a hydrogen atom undergoes a transition $n \rightarrow m$ is given to lowest order by

$$\Delta\mathcal{E}_0(n,m) \equiv \mathcal{R}_\infty hc\left[\frac{1}{n^2} - \frac{1}{m^2}\right], \tag{19a}$$

where[9]

$$\mathcal{R}_\infty = \frac{e_\infty^2}{2hca_0} = 109737.31534(13)\,\mathrm{cm}^{-1}, \tag{19b}$$

$$a_0 = \left(\frac{h}{2\pi m_e e_\infty}\right)^2, \tag{19c}$$

and

$$\Delta\mathcal{E}_0 = \Delta\mathcal{E}_{exp} - \Delta\mathcal{E}_{qed}. \tag{19d}$$

Here $\Delta\mathcal{E}_{exp}$ is the experimentally measured value, and $\Delta\mathcal{E}_{qed}$ denotes the calculated contributions from higher-order quantum electrodynamic effects. When a Yukawa or exponential potential is added to the Coulomb potential, the modified expression for $\Delta\mathcal{E}$ can be derived by treating the Yukawa or exponential potential as a first-order perturbation on the hydrogen Hamiltonian. Using perturbation theory, the energy $\Delta\mathcal{E}(n,m)$ released during the transition $n \rightarrow m$ is

$$\Delta\mathcal{E}(n,m) = \Delta\mathcal{E}_0(n,m) + \beta hc\mathcal{R}_\infty[\Delta E_{n,l} - \Delta E_{m,l'}], \tag{20a}$$

where for the Yukawa potential

$$\Delta E_{n,l} = \Delta E^Y_{n,l} \equiv \langle\Psi_{n,l}|(e^{-r/\lambda}/r)|\Psi_{n,l}\rangle\left(\frac{1}{hc\mathcal{R}_\infty}\right), \tag{20b}$$

and for the exponential potential

$$\Delta E_{n,l} = \Delta E^E_{n,l} \equiv \langle\Psi_{n,l}|(e^{-r/\lambda}/\lambda)|\Psi_{n,l}\rangle\left(\frac{1}{hc\mathcal{R}_\infty}\right). \tag{20c}$$

The effective Rydberg $\mathcal{R}_{n,m}$ is defined through the relation

$$\Delta\mathcal{E}(n,m) \equiv \mathcal{R}_{n,m}hc\left[\frac{1}{n^2} - \frac{1}{m^2}\right], \tag{21}$$

and from this and Eq. (20a) it is straightforward to show that

$$\mathcal{R}_{n,m} = \mathcal{R}_\infty[1 + \beta(\frac{\Delta E_{n,l} - \Delta E_{m,l'}}{n^{-2} - m^{-2}})]. \tag{22}$$

We see that $\mathcal{R}_{n,m}$ now depends upon the principal quantum numbers of the transition $n \to m$ from which it is measured. Defining

$$G(n,m) \equiv (\Delta E_{n,l} - \Delta E_{m,l'})/(n^{-2} - m^{-2}), \tag{23}$$

Eq. (22) can be cast in the form

$$\mathcal{R}_{n,m} = \mathcal{R}_\infty[1 + G(n,m)\beta]. \tag{24}$$

From the above relation, it directly follows that

$$\beta = (\Delta\mathcal{R}/\mathcal{R}_\infty)[G(N,M) - G(n,m)]^{-1} \tag{25a}$$

where

$$\Delta\mathcal{R} = \mathcal{R}_{N,M} - \mathcal{R}_{n,m} \;\; ; \quad N > M \geq n > m. \tag{25b}$$

These equations hold for both the Yukawa and exponential models. The integrals arising in Eqs. (20b) and (20c) can be done analytically, and the constraints which follow from Eq. (25a) are shown graphically in Figs. 1 and 2, for the Yukawa and exponential models respectively. We use as experimental input the values of $\mathcal{R}$ determined by Beausoleil, *et al.*[10] from the transition $n = 2 \to n = 1$, and by Julien, *et al.*[11] from the transition $n = 10 \to n = 2$. Also shown in Fig. 1 are the earlier results for the Yukawa model which were obtained by BL,[6] who used a semiclassical approximation in place of the full quantum mechanical treatment we have given here. In the semiclassical approximation, the additional Yukawa potential is treated as a constant by using the Bohr relation $r = a_0 n^2/Z$, so that for hydrogen ($Z = 1$),

$$V_Y(r) \to \quad V_Y^{SC}(n) = -(e^2/a_0 n^2)[\beta e^{-a_0 n^2/\lambda}], \tag{26a}$$

and

$$\Delta\mathcal{E} = \Delta\mathcal{E}_0 + [V_Y^{SC}(m) - V_Y^{SC}(n)]. \tag{26b}$$

where a_0 is the Bohr radius, and Z is the nuclear charge. As can be seen from Fig. 1 the semiclassical approximation used by BL and the quantum-mechanical treatment used here, agree in the regime where $\lambda \gg a_0$. However, for $\lambda \lesssim a_0$ the quantum-mechanical treatment gives more stringent constraints than those obtained from the semiclassical method, and this is precisely the effect of the spread of the wavefunction in quantum mechanics that we discussed earlier. A more detailed comparison of these two approaches will be presented elsewhere, along with a detailed analysis of the sensitivity of different experiments as a function of λ.

III. Summary and Conclusions

We have developed in Sec. I a unified framework for describing the deviations from both Coulomb's law of electrostatics and Newton's law of gravitation produced by an additional

scalar or vector field. In Sec. II we illustrated this methodology, by considering the constraints implied by the consistency of different determinations of the Rydberg constant. These are the curves denoted by (2) in Figs. 1 and 2. As discussed previously, the effect of including the full quantum-mechanical treatment is to make the Rydberg constraints even more restrictive than in the semi-classical treatment of BL, shown by the curves denoted by (1). For lack of space, we have not considered the other constraints more fully here, but the detailed derivation of these will be presented elsewhere. Briefly, however, curves (3), (4), and (5) are derived from the Cavendish method, and curve (6) is obtained from a study of the Earth's magnetic field. Finally, curve (7) is inferred from the consistency of different determinations of the electromagnetic fine structure constant over several distance scales. The analogous constraint curves for gravitational systems are shown in Figs. 3 and 4. Due to the existence of a much larger number of experimental results than in the electromagnetic case, we present only the envelope of the constraint curves and not the individual results. Fig. 3 exhibits the constraints on $\alpha_{12} \equiv \alpha$ as a function λ from composition-independent experiments. Fig. 4 exhibits the analogous constraints on $\xi \equiv f^2/G_\infty m_H^2$ obtained from composition-dependent experiments, assuming that $Q' = B =$ baryon number. Further details on these can be found in Refs. 1 & 3.

In conclusion, it is fair to say that the study of possible new forces "competing" with electromagnetism is barely in its infancy. However, even at this relatively early stage it is evident we have much to learn about which systems will give the best constraints on β for different λ. Moreover, we have only begun to explore the implications of combining data on electromagnetic and gravitational systems.

IV. Acknowledgements

We wish to thank Geoffrey Greene and Dennis Krause for helpful discussions on precision electromagnetic tests. This work was supported by the US Department of Energy under contract no.DE AC02-76ER01428, and by the Air Force Geophysical Laboratory under contract F192628-90-K-0010. One of the authors (HK) wishes to thank the *Cray Research* corporation for a travel grant to the *Rencontre de Moriond*, in connection with this work.

References

[1]E. Fischbach and C. Talmadge, *Nature* **356**, 207 (1992).

[2]E. Fischbach, G.T. Gillies, D. Krause, J.G. Schwan, and C. Talmadge, *Metrologia*, **29**, #3 (1992) in press.

[3]E.G. Adelberger, B.R. Heckel, C.W. Stubbs, and W.F. Rogers, *Ann. Rev. Nucl. Part. Sci.* **41**, 269 (1991).

[4]Y. Fujii, *Int. J. Mod. Phys. A* **6**, 3505 (1991).

[5]E. Fischbach, D. Sudarsky, A. Szafer, C. Talmadge, and S. H. Aronson, *Phys. Rev. Lett.* **56**, 3 (1986); *Ann. of Phys. (NY)* **182**, 1 (1988).

[6]D. F. Bartlett and S. Lögl,*Phys. Rev. Lett.* **61**, 2285 (1988).

[7]A. S. Goldhaber and M. M. Nieto,*Phys. Rev. Lett.* **21**, 567 (1968); A. S. Goldhaber and M. M. Nieto,*Rev. Mod. Phys.* **43**, 277 (1971).

[8]E. Fischbach, C. Talmadge, and D. Krause, *Phys. Rev. D* **43**, 460 (1991).

[9]E.R. Cohn, and B. N. Taylor *Phys. Today* **43**, BG 9 (Aug. 1990).

[10]R.G. Beausoleil, D. H. McIntyre, C. J. Foot, E. A. Hildum, B. Couillaud, and T. W. Hansch, *Phys. Rev. A.* **35**, 4878 (1987).

[11]L. Julien, J. C. Garreau, and F. Biraben, in *Laser Spectroscopy VIII,* edited by W. Persson and S. Svanberg (Springer- Verlag, Berlin, 1987), p.8.

[12]G. D. Cochran and P. A. Franken, *Bull. Am. Phys. Soc.* **13**, 1379 (1968).

[13]D. F. Bartlett, P. E. Goldhagen, and E. A. Phillips. *Phys. Rev. D.* **2**, 483 (1970).

[14]E. R. Williams, J. E. Faller, and H. Hill, *Phys. Rev. Lett.* **26**, 721 (1971).

[15]E.G. Adelberger,*et al. Phys. Rev. D* **42**, 3267 (1990).

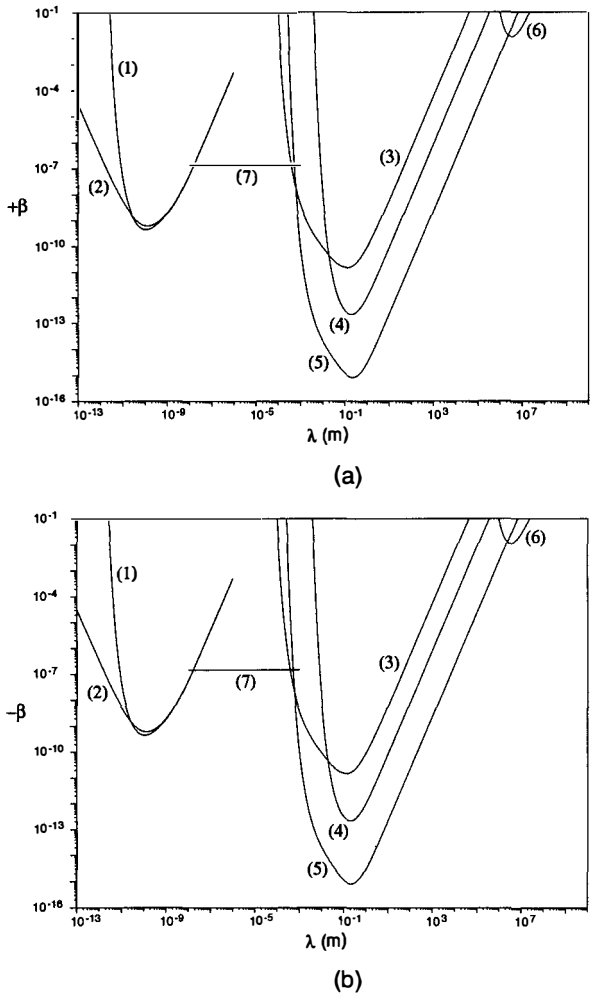


Fig. 1: Limits implied by electromagnetic experiments on the strengths $+\beta$ or $-\beta$, respectively, as a function of the range λ for a fifth force modeled by a Yukawa potential as in Eq. (1). The area above each curve is excluded by experiment. (1) Rydberg constraint obtained from the semiclassical method.[6] (2) Rydberg constraint obtained from the quantum mechanical perturbation method. Both Rydberg curves where constructed using $R_{10,2} = 109737.315711(53)\text{cm}^{-1}$,[11] and $R_{2,1} = 109737.31572(3)\text{cm}^{-1}$.[10] Curves (3),(4), and (5) were calculated using limits provided by Cavendish experiments.[12–14] Curve (6) is obtained from the Schrödinger external field method,[7] using data on the Earth's magnetic field. Curve (7) is obtained from an analysis of the consistency of different determinations of the fine structure constant. Details of our analysis will be published elsewhere.

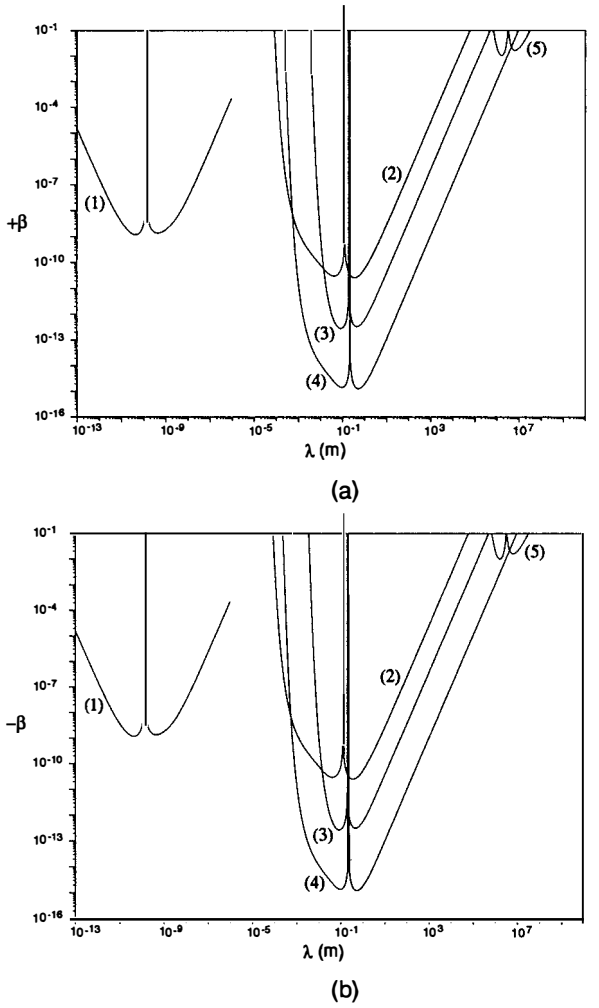


Fig. 2: Limits implied by electromagnetic experiments on the strengths $+\beta$ or $-\beta$, respectively, as a function of the range λ for a fifth force modeled by an exponential potential as in Eq. (17). See text and caption to Fig. 1 for further details.

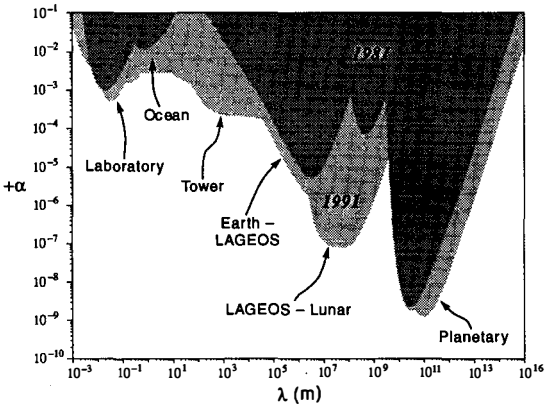


Fig. 3: Constraints on the coupling constant α as a function of the range λ from composition-independent experiments. The dark shaded area indicates the status as of 1981, and the lighter region the current limits. For further details see Ref. 1.

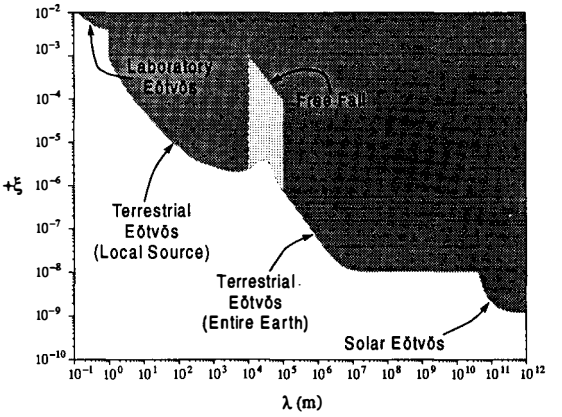


Fig. 4: Constraints on the coupling constant ξ as a function of the range λ from composition-dependent experiments, assuming a coupling to baryon number. With the exception of the region denoted by 'Solar Eötvös', all the constraints shown have been obtained since 1986, and most of the results shown are from the Eöt-Wash experiments.[15] The hatched region denotes values of λ where the limits on ξ depend on detailed models of the Earth and are at present uncertain. The dotted contour is our 'best guess' for the constraint in this region.

PROPOSED INVERSE SQUARE LAW TESTS OF GRAVITY

Michael W. Moore and Paul E. Boynton
University of Washington
Seattle, WA

Presented by
Michael W. Moore

Abstract

We propose two null experiments to test the inverse square law of gravity in the critical range between 10 cm and 100 meters with significantly higher precision than earlier work. The tests use an oscillating torsion pendulum whose configuration is sensitive primarily to the horizontal gradient of the Laplacian of the gravitational potential. A non-zero result will indicate a violation of the Newtonian inverse square law. The reasons for choosing to measure the horizontal gradient of the Laplacian of the gravitational potential are discussed here, along with the proposed design of the torsion pendulum and the source masses.

1. Testing the Inverse Square Law of Gravity

Existing experimental data do not strongly exclude the possibility that static, weak-field, Newtonian gravity is violated for distances less than 1000 km, nor that there might exist some yet-undiscovered, ultra-weak interaction of macroscopic range, coupling approximately to mass. We propose two null experiments to test the inverse square law of gravity in the range between 10 cm and 100 meters, where current experimental constraints are the weakest, with significantly higher precision than earlier work.

Our approach is to characterize the putative violation as a "Yukawa" interaction:

$$U_{12} = -G_\infty \frac{m_1 m_2}{r}\left(1 + \alpha e^{-r/\lambda}\right),$$

where α is the strength relative to gravity and λ is the characteristic range. We then show that for a torsion pendulum experiment, the lowest-order detectable signature of the non-Newtonian character of this interaction is proportional to the gradient of the Laplacian of the interaction potential. We were motivated to pursue this study by the recent work of Fischbach and Talmadge[1].

For $\lambda = 1$ meter, the present 2σ upper limit on α is about 2×10^{-3} (see Figure 1). The major source of uncertainty in these laboratory experiments is metrological. In our proposed experiment the lower multipole gravitational potentials are suppressed by design so that the leading gravitational interaction term is fifth order (i.e., an $\ell = 5$ or 32-pole potential). By contrast, the leading term uniquely associated with the "Yukawa" potential that can be detected with a torsion pendulum is third order (the gradient of the Laplacian). Because this coupling is two orders lower than the leading gravitational term, we have an $(R/r)^2$ signal-to-background advantage over previous experiments, which do not suppress low-order multipoles and rely only on metrological precision. (R/r is the ratio of the characteristic distance to the source mass over the characteristic size of the pendulum mass.) This strategy, which is the same as that used in our present composition-dependent force (CDF) experiments, should enable us to improve the limits on α by as much as one-and-a-half orders of magnitude for both laboratory-scale source masses and for a complementary hillside experiment.

2. Comparison to the CDF Experiment

The experimental technique for detecting a "Yukawa" potential is quite similar to the one we have been using to search for a composition-dependent potential[2]. The essential feature is an ability to measure with high sensitivity an azimuthal variation in potential energy of the pendulum mass, $U_{pend}(\phi)$. The objective of this experiment is to design a source mass and a pendulum mass such that $U_{pend}(\phi)$ is constant for a strictly Newtonian interaction but will exhibit a characteristic $\cos(\phi)$ dependence for a

non-Newtonian interaction, thereby defining a true null experiment. In practice, of course, there will be systematic errors; the most troubling ones mimic the characteristic $\cos(\phi)$ dependence of the "Yukawa" interaction.

In our composition-dependent force search, precisely designed lead masses are employed to cancel the large quadrupole potential of our topographical source mass[2], which, if uncancelled, would produce a false $\cos(\phi)$ signal. In placing such massive lead plates near to the pendulum, however, one must be careful to avoid producing higher gravitational multipoles that yield a detectable $\cos(\phi)$ signature. For our laboratory "Yukawa" experiment, lead plates are used to build the source mass itself, but with the same stringent requirement that offending gravitational potentials must be avoided. The problem of eliminating the higher gravitational multipoles is the same for both experiments, and the expertise we have developed in designing the CDF lead masses is directly applicable to the design of the "Yukawa" lead masses.

In the CDF experiment, the basic configuration that optimizes sensitivity to the non-Newtonian field is obvious — the left and right halves of the pendulum are made from materials of contrasting composition charge. For the "Yukawa" experiment, however, intuition does not serve us in selecting a pendulum design that is optimally and distinctly sensitive to the non-Newtonian aspect of the potential. Insight is gained by examining the multipole expansion of the "Yukawa" potential. Since the traditional multipole expansion applies only to potentials whose Laplacian is identically zero, a generalization is required.

3. Generalized Multipole Expansion

The potential of a source mass can be expanded in a three-dimensional Cartesian Taylor series around the center-of-mass of the pendulum, which we choose as the origin of the coordinate system.

$$U_{pend} = V|_0 m + \left.\frac{\partial V}{\partial x}\right|_0 m_x + \left.\frac{\partial V}{\partial y}\right|_0 m_y + \left.\frac{\partial V}{\partial z}\right|_0 m_z + \frac{1}{2}\left.\frac{\partial^2 V}{\partial x^2}\right|_0 m_{xx} + \frac{1}{2}\left.\frac{\partial^2 V}{\partial y^2}\right|_0 m_{yy} + \frac{1}{2}\left.\frac{\partial^2 V}{\partial z^2}\right|_0 m_{zz} + \left.\frac{\partial^2 V}{\partial x \partial y}\right|_0 m_{xy} + \left.\frac{\partial^2 V}{\partial x \partial z}\right|_0 m_{xz} + \left.\frac{\partial^2 V}{\partial y \partial z}\right|_0 m_{yz} \cdots$$

where the mass moments are defined as $m_x = \int_{pend} \rho_{pend}(x,y,z)\, x\, dx\, dy\, dz$, etc. Defining $V_p(x,y,z)$ as the potential due to a unit point mass (e.g., for Newtonian gravity, $V_p = G/r$,) the derivatives of V can be calculated by integrating over the lead source mass distribution:

$$\left.\frac{\partial V}{\partial x}\right|_0 = \int_{lead} \rho_{lead}(x',y',z')\left(-\frac{\partial}{\partial x'}\right) V_p(-x',-y',-z')\, dx'\, dy'\, dz'$$

For our purposes, however, it is revealing to switch from a Cartesian basis to a spherical harmonic basis. The spherical harmonic bases for the zeroth, first, second, and third mass moments are:

$$\left\{Y_0^0\right\}$$
$$\left\{rY_1^1, rY_0^1, rY_{-1}^1\right\}$$
$$\left\{r^2Y_2^2, r^2Y_1^2, r^2Y_0^2, r^2Y_{-1}^2, r^2Y_{-2}^2, \boxed{r^2Y_0^0}\right\}$$
$$\left\{r^3Y_3^3, r^3Y_2^3, r^3Y_1^3, r^3Y_0^3, r^3Y_{-1}^3, r^3Y_{-2}^3, r^3Y_{-3}^3, \boxed{r^3Y_1^1}, \boxed{r^3Y_0^1}, \boxed{r^3Y_{-1}^1}\right\}$$

The boxed terms make the bases complete, but they are not included in a multipole expansion that describes gravitational, electrostatic, or magnetostatic potentials, for which the Laplacian is identically zero. Yet, the corresponding mass moments are precisely those that couple to the Laplacian and to the derivatives of the Laplacian. Consequently, they are the uniquely non-Newtonian signature by which one is able to detect a "Yukawa" potential. The multipole expansion that includes these non-vanishing Laplacian terms will be referred to as the *generalized* multipole expansion to distinguish it from the traditional Newtonian potential multipole expansion.

The mass moments may be expressed in spherical coordinates as

$$M_m^{nl} = \int_{pend} \rho_{pend}(r,\theta,\phi) r^n Y_m^l r^2 \sin(\phi) dr d\theta d\phi \,.$$

Similarly, the potential derivatives in the Cartesian basis can be recombined into linear combinations, V_m^{nl}, such that the potential energy can be expanded as

$$U_{pend} = \sum_{n,\ell,m} U_m^{n\ell} = \sum_{n,\ell,m} V_m^{n\ell} M_m^{n\ell}, \quad \text{where} \begin{cases} n = \{0,1,2,\ldots\} \\ \ell = \{0,2,\ldots,n-2,n\} \text{ for even } n \\ \ell = \{1,3,\ldots,n-2,n\} \text{ for odd } n \\ m = \{-\ell,-\ell+1,\ldots,\ell-1,\ell\} \end{cases}$$

There are two elegant features of the spherical harmonic basis: it separates Newtonian from non-Newtonian interaction terms, and the angular dependence of a particular $U_m^{nl}(\phi)$ is simply $e^{im\phi}$ (or, in real functions, $\sin(m\phi)$ and $\cos(m\phi)$). This reveals why the Laplacian of the "Yukawa" potential can not be directly detected with a torsion pendulum. The Laplacian is a scalar, and has no angular dependence ($n = 2$, $\ell = 0$, $m = 0$), and therefore will not produce any azimuthal variation in the pendulum potential energy. The lowest order non-Newtonian potential term that can be detected is the horizontal component of the gradient of the Laplacian, which couples to a third order ($n = 3$, $\ell = 1$, $m = \pm 1$) mass moment.

The multipoles exhibit important invariances under arbitrary rotation about any axis: 1) a multipole will transform into the other multipoles with different m values but the same n and ℓ values; 2) the sum of the squares of the magnitudes of all the multipoles with the same n and ℓ is a scalar invariant under rotation; 3) moreover, if all multipoles with a given n and ℓ are equal to zero for any orientation, they will remain zero under an arbitrary rotation. These rules form the basis of the pendulum design.

4. Design of the Pendulum

The pendulum is designed to meet three criteria: 1) the breaking strength of the torsion fiber fixes the total mass of the pendulum; 2) the ($n = 3$, $\ell = 1$, $m = \pm 1$) mass moment should be chosen to maximize sensitivity to the corresponding uniquely non-Newtonian potential term; 3) all the standard gravitational mass moments through *at least* $\ell = 3$ should be canceled to prevent these $m = \pm 1$ *Newtonian* gravitational interactions from mimicking the "Yukawa" interaction. To simplify calculating and machining, cylindrical symmetry was chosen for the pendulum. This has the additional advantage of canceling all but the $m = 0$ moments.

Of course, cylindrical symmetry also cancels the ($n = 3$, $\ell = 1$, $m = \pm 1$) mass moment that we are apparently trying to maximize. We choose instead to maximize the ($n = 3$, $\ell = 1$, $m = 0$) mass moment, which makes the pendulum sensitive to the vertical derivative of the Laplacian, and then to rotate the pendulum 90° (onto its side). Because the three ($n = 3$, $\ell = 1$) mass moments transform like the components of a vector under rotation, the pendulum now becomes sensitive to the horizontal derivative of the Laplacian as desired. If the $m = 0$ *gravitational* mass moments are suppressed by design prior to the 90° rotation, then they remain zero after the 90° rotation as mentioned above.

We have settled upon a preliminary "three ring" design (actually, two rings and a point mass) that has vanishing $\ell = 1$, $\ell = 2$, and $\ell = 3$ Newtonian gravitational mass moments. In the original vertical orientation, the dimensions are:

	Mass (g)	Height (cm)	Radius (cm)
Ring 1	15.43	-2.500	2.500
Ring 2	27.43	0.625	2.500
Ring 3	17.14	1.250	0.000

Table 1 Dimensions of Proposed Pendulum

The actual pendulum's rings will have finite thickness and be connected by an appropriately thin, cylindrical surface. The mass moments of this structure can be readily evaluated and the dimensions modified to keep the $\ell = 1$, $\ell = 2$, and $\ell = 3$ gravitational moments zero. This is identical to the type of calculations we have

performed in designing our composition dipole pendulum. In theory, we can cancel even higher multipole mass moments by adding more rings; however, the more intricate the design the greater is the susceptibility to machining error. At present, we feel there is little to be gained by canceling the pendulum mass moments higher than $\ell = 3$.

In the original, unrotated, vertical orientation, the only uncancelled gravitational mass moments are the ($n = 1$, $\ell \geq 4$, $m = 0$) moments. The 90° rotation will mix the $m = 0$ moments into all other m values; however, our primary concern is the offending $m = \pm 1$ multipoles. After a 90° rotation, the pendulum has mirror symmetry about the xy-plane (i.e., it is invariant under a $z \rightarrow -z$ transformation). This symmetry causes the ℓ = even, $m = \pm 1$ moments to be zero. Thus, the most egregious gravitational mass moment in the design will be the $\ell = 5$, $m = \pm 1$ mass moment, followed by $\ell = 7$, $m = \pm 1$.

5. Design of the Lead Mass

The lead mass is the source of the "Yukawa" ($n = 3$, $\ell = 1$, $m = \pm 1$) potential in our proposed experiment. For simplicity in calculating and machining, we chose 180° sectors of lead, since this maximizes the $m = \pm 1$ signal. Unfortunately, this choice also maximizes the offending $m = \pm 1$ gravitational potentials, but if we give the lead mass mirror-symmetry about the xy-plane ($z \rightarrow -z$ transformation), the ℓ = even, $m = \pm 1$ potentials are canceled. Thus, the primary design goal is to cancel the ℓ = odd, $m = \pm 1$ gravitational potentials by carefully selecting the lower and upper elevations and inner and outer radii of the lead sectors. Naturally, we seek to maximize the strength of the "Yukawa" potential while canceling the gravitational potentials.

The ($n = 1$, $\ell = 1$, $m = \pm 1$) gravitational potential is not a problem. The torsion pendulum is essentially a plumb line, so there can be no horizontal gradient to the gravitational potential. The gravitational potentials that must be canceled, therefore, are the $\ell = 3$, $m = \pm 1$ and the $\ell = 5$, $m = \pm 1$ potentials. Although the $\ell = 5$, $m = \pm 1$ potential is weaker, it is important because the $\ell = 5$, $m = \pm 1$ mass moment was not canceled on the pendulum. Our design of the lead source mass is therefore constrained to cancel the $\ell = 3$, $m = \pm 1$ and the $\ell = 5$, $m = \pm 1$ potentials.

6. Expected Performance of the Experiment.

In our present composition-dependent force experiments, we can measure the frequency of a 447 second period torsion pendulum to 2 parts in 10^8 from the timing of a single oscillation cycle. By integrating over several hundred periods for each of the eight azimuthal equilibrium positions, the period at each position can be measured to about 1 part in 10^9. For $\lambda = 1$ meter, this enables a $\sim 2\sigma$ detection of a value of α as small as 2.5×10^{-4}. Further reduction is problematic because systematic errors due to tilt, translation, and machining errors may become important at this level. A graph of the

detectable α versus λ for $\Delta\omega/\omega = 10^{-9}$ is shown in figure 1. The 2σ envelope of α versus λ plots from previous experiments by other groups is plotted for comparison[3]. For laboratory experiments, constraints are strongest for $\lambda \lesssim 0.1$ meter. For $\lambda > 100$m, there are stringent limits on α from the tower experiments. Observations of orbiting bodies place still stronger limits for $\lambda > 100$ km[3]. Consequently, there is a pronounced gap in experimental limits in the range of $\lambda = 0.1$ to 100 meters. The recent ocean experiment by Zumberge et al.[4] partially fills this gap. A proposed cliff-side version of our experiment should also help place more stringent limits on α in this range.

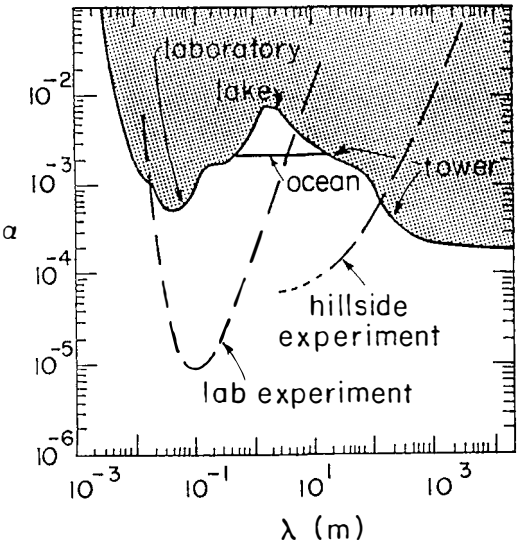


Figure 1. Existing limits on α compared to the two inverse square law violation tests proposed here (shaded area is ruled out.)

7. Using a Cliff as the Source Mass

For λ larger than R, (the characteristic size of the lead mass,) the "Yukawa" signal decreases as λ^{-2}. For a large source mass such as a cliff or hillside and for λ larger than the distance to the source but smaller than the size of the source mass, the "Yukawa" signal decreases only as λ^{-1}. This essentially gives the cliff or hillside a λ/R advantage over a lab source of comparable density. In absolute terms, however, the signal is still decreasing with increasing λ.

A natural remedy for this decreased signal is to increase r, the characteristic size of the pendulum mass. Being proportional to the third derivative, the signal will increase as r^3. Using the Index cliff[2] as a source, increasing the pendulum size by a factor of two, and maintaining the same $\Delta\omega/\omega = 10^{-9}$ sensitivity, we could detect $\alpha = 2.5\times10^{-4}$ for $\lambda = 30$ meters. A graph of α versus λ for such a cliff-side experiment is also plotted in Figure 1.

As in the CDF experiment, lead masses must be used to cancel the ambient gravitational potentials — particularly the quadrupole potential. If the pendulum size is increased by a factor of two, however, the gravitational interactions with the fabrication and placement errors of the lead mass would become troublesome. One solution is to make the lead mass proportionally larger, but this would require a

prohibitive amount of lead. Another solution is to find a site that is half-way up the side of the cliff rather than at the base or top. Being odd in z, the offending $\ell = 2$, $m = \pm 1$ potential can then be canceled without a decrease in the "Yukawa" signal — provided, of course, that the cliff is of the same size and density as the Index cliff. Having significantly reduced the $\ell = 2$, $m = \pm 1$ potential, small lead compensating masses can be designed with substantially reduced higher multipole contributions. We have such a site selected on the Hanford reservation in eastern Washington State:. Rattlesnake Ridge is larger than the Index feature, rising 1000 meters from the desert floor with a slope of nearly 40°.

8. Comparison to Other Experiments

Several laboratory torsion balance tests for a "Yukawa" potential have been carried out[5]. The difference between our proposal and previous experiments (such as that of Hoskins, Newman, and Spero[6]) is, as mentioned above, that the lower gravitational multipole interactions are to be measured and canceled. This allows an $(R/r)^2$ improvement in sensitivity beyond what can be achieved by minimizing machining and placement errors only. Multipole analysis and cancellation have been used previously in composition-dependent force searches, but not in inverse square law tests. Our approach to detecting an inverse square law violation is also new. Unlike the CDF pendulum, the symmetry breaking necessary to produce sensitive "Yukawa" pendulum is quite severe, and the basic design can not be determined by inspection from symmetry considerations. Our design allows a true null test with Newtonian interactions suppressed to unobservable levels at our current sensitivity.

Still other experiments (e.g., Chan, Moody, and Paik[7]) can measure the Laplacian of the potential directly. In theory, measuring the Laplacian instead of its gradient yields a factor λ/R advantage in sensitivity for large λ. In practice, though, these experiments must subtract off the background potential by spatial differencing. As a result, they effectively measure the gradient of the Laplacian, and the resulting λ dependence is the same as ours. The deflection per unit force of their sensing masses, however, is orders of magnitude smaller than the deflection of a similar-mass, thin-fiber torsion balance, and they must contend with the complications that arise from operating at cryogenic temperatures.

A third approach is to measure g, the vertical gradient of the potential, along the ground and up a vertical tower[8,9,10]. By solving the Neumann boundary value problem along the ground, the potential can be analytically continued upwards and the $g_{predicted}$ can be compared to $g_{measured}$. The difference is proportional to the Laplacian of the potential. Since the boundary value problem implicitly assumes $g = 0$ at infinity, *absolute* measurements of g must be made in order to measure the Laplacian. If only relative measurements of g are made, then only the vertical gradient of the Laplacian can be measured. As with the gradiometer mentioned above, the gravimeters used to

measure g directly have a much smaller deflection per unit force than a torsion balance, and consequently have smaller intrinsic sensitivity.

The ocean experiment of Zumberge et al.[4] is essentially an aquatic version of the tower experiments, except that direct measurement of the density of the source mass (the ocean) is possible. This allows them to measure G. In fact, this was the main purpose of the experiment; the limits on inverse square law were a pleasant secondary result. By measuring G over a characteristic range of several kilometers and comparing that result with very accurate laboratory measurements of G over a range of a few centimeters, Zumberge et al. were able to extend their limits on α from kilometers to centimeters. We are proposing to place still lower limits in the range from 0.1 to 100 meters.

Acknowledgments

The authors thank John Guthrie and Ron Legere for substantial contributions to the development of the ideas presented here. This work was supported largely by the National Science Foundation under grant PHY8822705.

References

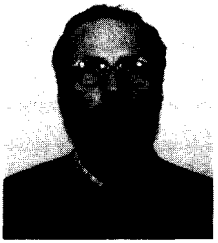
1. E. Fischbach and C. Talmadge, in *New and Exotic Phenomena '90, Proceedings of the X^{th} Moriond Workshop*, Les Arcs, France, January 1990, eds., O. Fackler and J. Tran Thanh Van (Editions Frontières), p.187
2. P.E. Boynton in *International Symposium on Experimental Physics*, Guangzhou, China, August 1987, eds. P.F. Michelson, Hu En-ke, and G. Pizella (World Scientific), p.75
3. E. Fischbach and C. Talmadge, *Nature* **356**, 207 (1992)
4. M.A. Zumberge, J.A. Hildebrand, J.M. Stevenson, R.L. Parker, A.D. Chave, M.E. Ander, and F.N. Spiess, *Phys. Rev. Lett.* **67**, 3051 (1991)
5. J. Faller, E. Fischbach, Y. Fujii, K. Kuroda, H.J. Paik, C. Speake, *IEEE Trans. Inst. and Meas.* **38**, 180 (1989)
6. J.K. Hoskins, R.D. Newman, R.D. Spero, *Phys. Rev. D.* **32**, 3084 (1985)
7. H.A. Chan, MV. Moody, and H.J. *Paik, Phys. Rev. Lett.* **49**, 1745 (1982)
8. C. Jekeli, D.H. Eckhardt, A.J. Romaides, *Phys. Rev. Lett.* **64**, 1204 (1990)
9. C. Speake, T.M. Niebauer, M.P. McHugh, P.T. Keyser, J.E. Faller, J.Y. Cruz, J.C. Harrison, J. Mäkinen, and R.B. Beruff, *Phys. Rev. Lett.* **65**, 1967 (1990)
10. J. Thomas, P. Kasumeyer, O. Fackler, D. Felske, R. Harris, J. Kammeraad, M. Millet, M. Mugge *Phys. Rev. Lett.* **63**, 1902 (1989)

PRELIMINARY RESULTS FROM THE *GAL* EXPERIMENT

S. Carusotto, V. Cavasinni, A. Mordacci, F. Perrone, E. Polacco, G. Zavattini*
Dipartimento di Fisica, Università di Pisa and INFN, Sez. Pisa, Italy

E. Iacopini
Scuola Normale Superiore, Pisa and INFN, Sez. Pisa, Italy

G. Stefanini
CERN, Geneva, Switzerland



Abstract

We report preliminary results for the difference Δg in the free fall acceleration of aluminium and copper, in vacuum, obtained in a differential Galileo's type experiment. Within the sensitivity of the apparatus $\Delta g/g = 7 \times 10^{-10}$ no g-universality violation has been observed.

*Now at Dipartimento di Fisica di Bologna, Italy

1 Introduction

A direct measurement of the difference in the free fall acceleration in vacuum of test bodies with different nuclear composition, is, at present, the only way to verify the fifth force theory [1] at interaction ranges between 10 Km and the Earth radius.

To investigate effects in this range, we proposed [2] [3] a true differential free fall Galileo's type experiment (GAL). The basic method is to measure the angular acceleration $\dot{\omega}$ of a free falling disk, made of two half-disks of different materials. The angular motion of the disk around its axis is measured with the help of a modified Michelson interferometer, in which the two arms terminate in two cube-corner reflectors (CCR), mounted on the rim of the disk.

2 Experimental apparatus

The disk assembly consists of two aluminium holders (see Fig.1) which contain, in their central region, the two half disks (see Fig. 2a) and which keep at 70 mm from the disk axis eight CCR used for the measurement. The axis of the assembly (see Fig.2b), rigidly fixed in one of the two holders, was made in titanium.

The disk geometry has been chosen to minimize the effects due to local gravity gradients. We have mounted 8 CCR to have the possibility of inverting the position of the half disks with respect to the interferometric arms, by simply rotating of 180° the disk assembly around its axis, without breaking the vacuum. The two half disks have been manufactured in such a way to have

1. the same mass (m_{Al}=346.8 g, m_{Cu}=347.0 g);
2. the gravity centers symmetrically located with respect to the disk assembly at the same distance (b=2.412 $\pm$ 0.003 cm);
3. the same moments of inertia in the plane orthogonal to the disk axis ($\Delta J/J < 10^{-3}$).

The experimental apparatus is schematically shown in Fig.3. The light source is a frequency stabilized He-Ne laser (NL-1, from Newport). The disk is installed in a cylindrical stainless-steel vessel, 8 m in length, 270 mm in diameter, evacuated down to 3. $\times$ 10^{-5} Torr. The two windows W3 and W4 allow small translations of the beams, in order to obtain the maximum optical contrast. All windows W1-W4, together with the beam splitter BS1, have been measured to have a prismaticity smaller than 2. $\times$ 10^{-6} rad.

The disk axis is kept in the upper rest position by the horizontal "biellette" B (see Fig.1), and it seats on two horizontal hardened silver steel parallel pins, 2 mm in diameter, fixed on the "biellette" orthogonally to the disk axis. The small flats (2 $\times$ 20 mm) at the end of the axis forbid the disk to run away. The disk axis is inside a suitable housing fixed on a carriage, and the distance between the disk axis and the housing walls is 3 mm in the horizontal plane and 50 mm in the vertical direction.

A measurement cycle is initiated by mechanically releasing the carriage, which starts a free fall. After 10.3 ms the disk is also released. This time interval has been calibrated so as to guarantee no contact between the disk axis and the housing during the measurement. The carriage is guided by vertical rails on the first 30 cm of its path. To release the disk, we use a compressed air actuator, installed outside from the vacuum vessel. The actuator pushes a brass rod fixed on a bellow which, through a suitable fork P inside the vacuum vessel, open the "biellette" fast enough to release the disk with the minimum amount of perturbation. After 4.2 m of free fall, a progressive deceleration is produced on the carriage by braking pads rubbing on spring-loaded stainless steel bars. A soft landing of the disk on the carriage is obtained at the beginning of this braking by a suitable damping catch in the bottom of the axis housing. The whole assembly (disk+carriage) comes at rest in 1.5 m.

A gear is used to pull the carriage with the disk on it to the top of the vacuum vessel and to rearm the "biellette" for another measurement cycle. A full cycle requires about 5 minutes.

During the disk free fall, the signal from the photodiode PD1 is sampled by a 14 bits analog-to-digital converter (SHA1144 + ADC1131K from Analog Devices) at a fixed, stable frequency of 50 KHz ($\delta\nu/\nu < 10^{-8}$), obtained from a thermally stabilized quartz oscillator (HP 3325A frequency synthesizer). The data are stored in a fast buffer memory (LeCroy MM8206A), which is read by a

microcomputer at the end of each measurement and transferred to a host 8300 VAX computer, for off line analysis.

3 Signal analysis

If there is a difference Δg in the free fall acceleration of aluminium and copper, then the disk assembly experiences a torque

$$T = (m_{Al} + m_{Cu})\frac{b\Delta g}{2} \tag{1}$$

and, therefore, there is an angular acceleration of the disk assembly $\dot{\omega}$ given by

$$\dot{\omega} = \frac{T}{J_0 + J_1 + J_2} = K\frac{\Delta g}{g}, \tag{2}$$

where J_0, J_1, J_2 are the moments of inertia in the disk axis direction of the holder + CCR, of the aluminium half disk and of the copper half disk, respectively.

The measurement of $\dot{\omega}$ is performed by observing, during the free fall, the angular motion of the disk around its axis with the help of the modified Michelson interferometer (see Fig. 3) in which the two arms terminate in two CCR fixed on the disk. Every τ=20 μs, we sample the light intensity from the interferometer, which reads

$$I(t) = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\phi(t) \equiv A\cos\phi(t) + B, \tag{3}$$

where I_1 and I_2 are the light intensities in the two interferometric arms and $\phi(t)$ is the phase difference between the two optical paths, given by

$$\phi(t) = \phi_0 + \frac{8\pi R}{\lambda}\sin(\omega_0 t + \frac{1}{2}\dot{\omega}t^2) \approx \phi_0 + \frac{8\pi R}{\lambda}\omega_0 t + \frac{4\pi R}{\lambda}\dot{\omega}t^2 = \phi_0 + 2\pi f t + 2\pi\epsilon t^2, \tag{4}$$

where λ is the light wavelength, R=70 mm is the distance between the disk axis and each CCR, f is the fringe frequency (proportional to ω_0, the disk angular velocity at t=0) and ϵ is the fringe frequency drift, proportional to $\dot{\omega}$, i.e. to the effect to be measured.

We discard the first 7000 samples (14 ms) of $I(t) \equiv I(n\tau) \equiv I_n$ because of the noise induced in the interferometer by the release systems, and we consider only the next 36000 samples (0.72 s). For every fringe period, we fit the amplitude A and the offset B, and then we best-fit these data with a 4th order polynomial, in order to assign to each sample its own value A_n and B_n.

Then, we unfold the interferometric phase $\phi(n\tau) \equiv \phi_n$ using the relation

$$\phi_n = \arccos\left(\frac{I_n - B_n}{A_n}\right) \tag{5}$$

and we fit the phase data ϕ_n with the first twenty Chebyshev orthogonal discrete polynomials $Q_k(n)$[5]. We obtain

$$\begin{aligned} \phi_n &= \phi_0 + 2\pi f\cdot n\tau + 2\pi\epsilon\cdot(n\tau)^2 + noise = \alpha_0 Q_0(n) + \alpha_1 Q_1(n) + \alpha_2 Q_2(n) + \ldots \\ & 0 \le n \le N = 36000 \end{aligned} \tag{6}$$

and it is:

$$\begin{aligned} \phi_0 &= \alpha_0 + \alpha_1 + \alpha_2 \\ 2\pi f\tau &= -\frac{2\alpha_1}{N} + \frac{6\alpha_2}{N-1} \\ 2\pi\epsilon\tau^2 &= \frac{6\alpha_2}{N(N-1)}. \end{aligned} \tag{7}$$

Consequently, the value of the Chebyshev coefficient α_2 is a measurement of ϵ, i.e of the effect we are looking for.

4 Noise evaluation

The values of the coefficients $\alpha_3 \ldots \alpha_{19}$ were used to estimate the random noise affecting the measurement, assuming white noise. This hypothesis was checked with the χ^2 method. We studied the origin of the noise by sampling, f.i., both photodiodes PD1 and PD0, during the same free fall. The coefficients $\alpha_0 \ldots \alpha_{19}$ for the two sets of data where the same, and this means that the interferometric phase noise was indeed the dominant one.

The r.m.s. σ on $\Delta g/g$ coming from this phase noise has been measured to be

$$\sigma \approx 14 \times 10^{-10} \quad \text{per free fall.} \tag{8}$$

5 Systematics

Given the high sensitivity of the measurement, there are many possible sources of systematic effects that have to be taken into account. Some of them (beams misalignment, disk mechanics) have been suitably reduced at the level of 10^{-10} or less [2],[3]. Some others, such as residual gas pressure effects and eddy currents induced during the free fall, have been considered. However, they give rise, predominantly, to a t^3 term in the interferometric phase, i.e. to a sizeable contribution to α_3 in the Chebyshev expansion, which was instead absent.

The most serious source of systematic error in our measurement was the disk precession. During the free fall, the disk axis precesses around the angular momentum. This motion simulates in the interferometric phase the presence of a disk angular acceleration $\dot{\Omega}$ given by

$$\dot{\Omega} = \omega_x \cdot \omega_y, \tag{9}$$

where ω_x and ω_y are the disk angular velocity components (at t=0) in the disk plane (ω_x is the component along the vertical). Typical values are $\omega_x, \omega_y \approx 1.5$ mrad/s, simulating disk angular accelerations $\dot{\Omega} \approx 2.2\ \mu\text{rad}/s^2$, i.e. a $\Delta g/g \approx 400. \times 10^{-10}$. This systematic effect can be accounted for by measuring ω_x, ω_y. For this reason, we fixed inside each of the two aluminium holders, 4 plane mirrors (25 mm in diameter and 3 mm in thickness), disposed parallel to the disk plane. In each of the 4 disk positions, a mirror reflects back, in the horizontal plane, the beam of a He-Ne laser, into a position sensitive photodiode, oriented along the x-y axis and disposed at L=3.3 m from the disk itself. When the disk starts to free fall, we sample every 40 μs the photodiode signals for about 60 ms, before loosing the beam. From these signals, we compute the velocity of the beam spot on the photodiode and, therefore, dividing by 2L, the components ω_x and ω_y.

The uncertainty in the ω_x, ω_y determination was about 0.15 mrad/s, due to the mirror surface quality, which means an uncertainty σ on $\Delta g/g$ of $\approx 60. \times 10^{-10}$ per free fall.

6 Results

We started our set of measurements using a homogeneous aluminium disk, to check the sensitivity of the apparatus and to look for possible spurious effects. We performed measurements in a disk position and in the conjugate one, obtained by rotating the disk around its axis by 180°. The mean value of the two sets of results were subtracted in order to eliminate possible contributions due to local gravity gradients. With a total of 70 measurements, we obtained the following

$$\Delta g/g = (3.2 \pm 9.5) \times 10^{-10}. \tag{10}$$

Then, we started the measurements using copper and aluminium. We use the same above algorithm and, with 63 measurements, we obtained

$$\Delta g/g_{Al-Cu} = (8.5 \pm 9.5) \times 10^{-10}. \tag{11}$$

We reversed the disk and we repeated 65 measurements. We obtain

$$\Delta g/g_{Al-Cu} = (-4.8 \pm 11.2) \times 10^{-10}. \tag{12}$$

By combining together the two above independent values, we get the present GAL limit on the $\Delta g/g$ for aluminium and copper, which is

$$\Delta g/g_{Al-Cu} = (2.9 \pm 7.2) \times 10^{-10}. \quad (13)$$

This result is in quite a good agreement with the one obtained by Kuroda and Mio [4] for the same materials, and using both results together, one reaches the following limit on the g-universality violation in aluminium and copper

$$\Delta g/g = (0.9 \pm 5.2) \times 10^{-10}, \quad (14)$$

obtained in direct free fall Galileo's type experiments.

Aknowledgments

We thank Dr. Ph. Bernard and his group for useful criticism and technical support, and in particular for having provided the vacuum vessel, the pumping system and part of the electronics. We want also to thank B. Smith for having built the circuitry concerning the 14-bits ADC, C. Rosset for having drawn the apparatus and given a substantial help in its design. Finally we wish to thank M. Del Colletto for his invaluable, skilled work on the mechanical realization of the experiment.

References

[1] E. Fischbach, D. Sudarsky, A. Szafer, C. Talmadge and S.H. Aronson, *Phys. Rev. Lett.*, **56** (1986) 3

[2] V. Cavasinni, E. Iacopini, E. Polacco and G. Stefanini, *Phys. Lett. A* **116** (1986) 157

[3] S. Carusotto, V. Cavasinni, E. Iacopini, E. Polacco and G. Stefanini *Proceedings of the XXII Rencontres de Moriond, Les Arcs, January 1988,pag. 523*

[4] K. Kuroda and N. Mio, *Phys. Rev.Lett.*, **62** (1989) 1941 and *Phy. Rev. D*, **42** (1990) 3903

[5] M.Abramowitz and I.A. Stegun, *Handbook of mathematical functions, Pag. 788-792*

Figure captions

- Figure 1. View of the disk assembly and of the release mechanism. B is the "biellette" and P is the fork.
- Figure 2. Mechanical drawing of the half disks (a) and of the disk axis (b).
- Figure 3. Schematic view of the apparatus. DS is the disk, CR is the carriage, BB are the breaking bars, W are the optical windows, BS are the beam splitters, M is the interferometer mirror, L are lenses, PP is a pentaprism and PD are photodiodes.

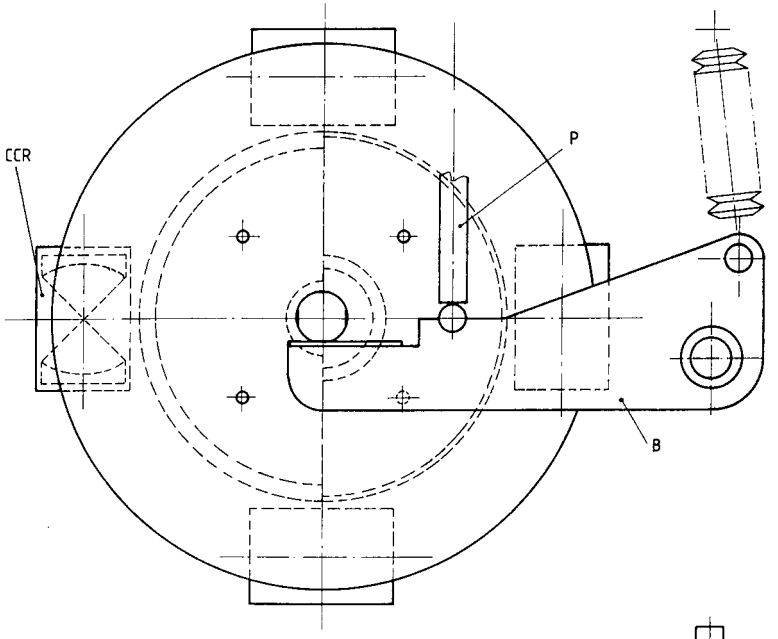


FIG 1

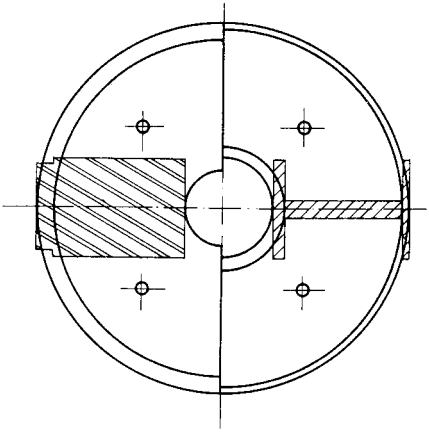


FIG 2A

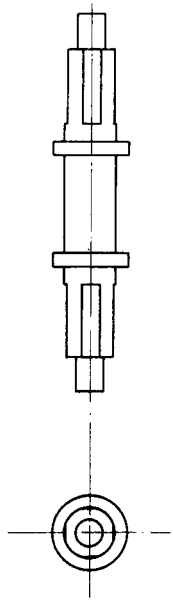


FIG 2B

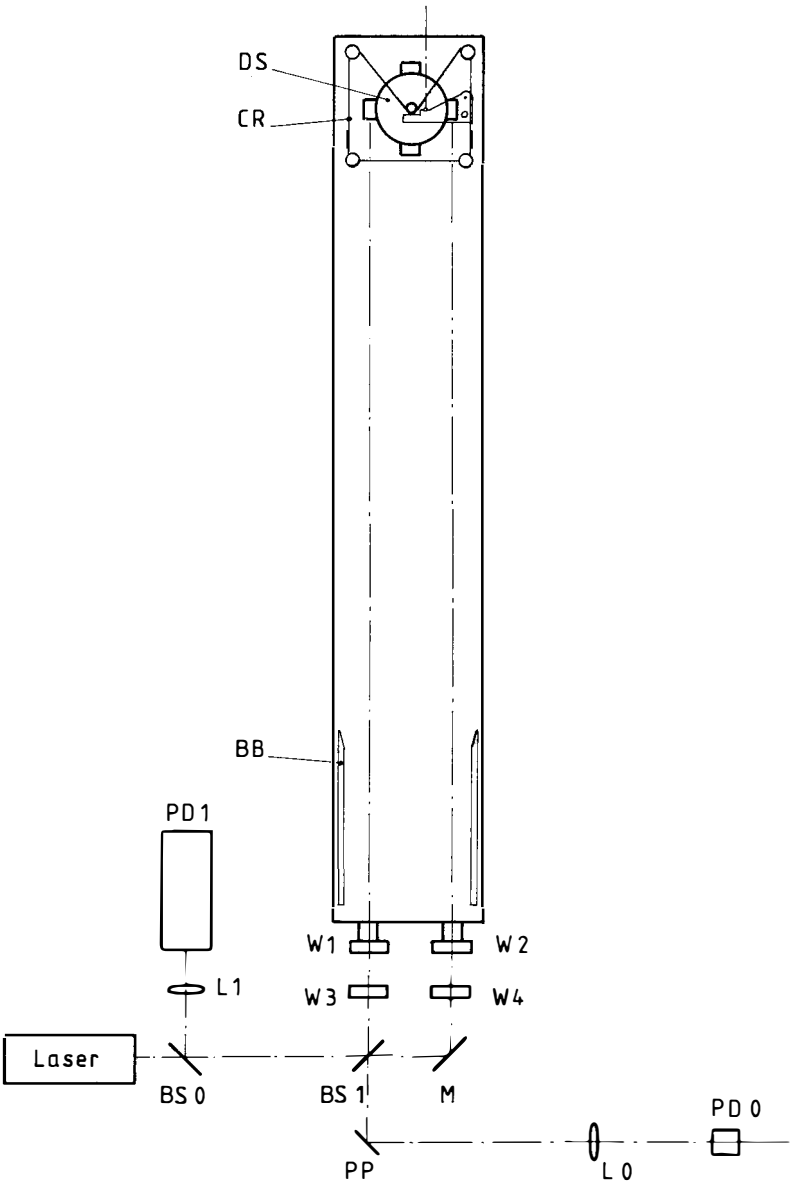
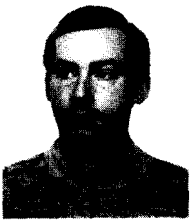


FIG 3

A TEST OF NEWTON'S GRAVITATIONAL LAW IN THE RANGE OF 0.6 M TO 3.6 M

J. Schurr, B. Langensiepen, H. Meyer, H. Piel and H. Walesch

Fachbereich Physik, Universität Wuppertal, 5600 Wuppertal, Germany



ABSTRACT

We have performed an experiment to test the $1/r^2$ dependence of Newton's law of gravitation and to determine the gravitational constant G, using a microwave resonator. The gravitational force of a laboratory test mass acts on this resonator and the resulting change of the resonator frequency is used to determine the gravitational force as a function of distance. In a first series of experiments we have investigated the inverse square law in the range of distances of 0.6 m to 3.6 m and determined the gravitational constant with a relative error of $1.6 \cdot 10^{-3}$. No deviations from Newton's inverse square law or the CODATA value of G were found.

I. INTRODUCTION

Many aspects of Newton's law of gravitation have been investigated experimentally during the last three centuries and the precision of these experiments has increased continuously. Until today the most precise determination of the gravitational constant G is possible by means of the famous Cavendish torsion balance. It has been highly developed in the last century and G can now be determined with a relative error of about 10^{-4}. It is however remarkable, that none of the experiments which have been performed to obtain a precision value for G have at the same time tested the inverse square law. Only a few experiments in the range of a few millimeter up to 10 m have been performed to test the inverse square law using a laboratory test mass. But none of them result in values for the gravitational constant. This is astonishing, because a lot of experiments are limited by systematic influences and the gravitational constant cannot be investigated independent of the $1/r^2$ law of gravitation [1].

The reports of possible intermediate-range deviations from Newton's gravitational law have stimulated a fifth-force discussion during the last five years [2]. This discussion has motivated us to develop a novel gravimeter. The gravimeter is based on a Fabry-Perot microwave resonator. External forces acting on the resonator can be detected due to the shift of its resonant frequency which can be measured with very high accuracy. The gravimeter was designed to measure the gravitational field of a test mass as a function of distance and from this to determine the gravitational constant G. A possible intermediate-range force can be investigated without assuming an explicit dependence on the material in use. Systematic errors which could influence the precision of the gravitational constant will in general result in deviations from the inverse square law. Their dependence on the distance of the interacting masses is a useful information to identify and to eliminate the source of these effects. The very sensitive check of systematic influences allows to enhance the precision of the measurement of the gravitational force and this again allows to enhance the precision of the gravitational constant.

II. THE GRAVIMETER

We have started our experiments to test Newton's law with measurements of the gravitational force of a test mass in a distance of about 10 cm to the resonator [3]. This prototype set-up was used to get an idea of the achievable resolution and accuracy and to develop solutions for various problems.

Based on this experience an improved set-up was designed to measure the gravitational force in the range of 0.6 to 3.6 m [4]. The main part of the

gravimeter consists of two Fabry-Perot mirrors suspended as a pair of pendula (Fig. 1). The pendula have a length l of about 2.6 m and the distance between the two mirrors is about 0.24 m. The gravitational force of a test mass M displaces both pendula and results in a change of the mirror separation b. Therefore, the length of the Gaussian beam inside the resonator changes and results in a change of the resonator frequency f. That is, the gravitational force of the test mass is measured by means of the frequency shift of the Fabry-Perot resonator.

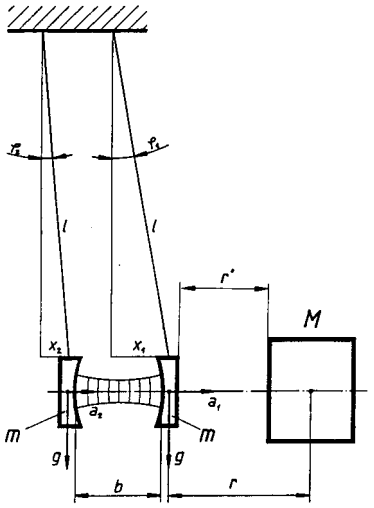


Fig. 1: The principle of the Fabry-Perot gravimeter.

A resonator frequency in the microwave range of approximately 22 GHz is chosen. A high quality factor of about $2 \cdot 10^5$ is obtained experimentally which is limited only by reflection losses of the copper mirrors. Furthermore, in the microwave region it is possible to meet all requirements to transmit the power of a source with appropriate stability to the resonator, to couple them into the resonator in order to excite the desired Fabry-Perot mode, and then to transmit them to a detector system. Furthermore, the limitation of the quality factor due to reflection losses of the copper mirrors can be overcome by using superconducting mirrors. We have demonstrated that at a temperature of 4.2 K a quality factor of $2 \cdot 10^7$ can be obtained with niobium mirrors [5]. Recent results of High-T_c-Superconductors seem to open the possibility for an improved gravimeter at 77 K [6].

Fig. 2 shows the schematic arrangement of the experimental set-up. Its main part, the two copper Fabry-Perot mirrors, are suspended as pendula inside of a vacuum tank which is mounted into a supporting steel construction.

The test mass is a cylinder of 576 kg positioned outside of the vacuum tank. Their dimensions were chosen in a way that the gravitational force between test mass and resonator is nearly the same as the gravitational force of point masses positioned at the centers of gravity. The test mass rests on a special guide rail and glides on rollers which are rotating on ball bearings fixed to this rail. The test mass can be positioned precisely by means of a motor driven spindle. All movable parts of this rail perform a strictly rotational motion and do not contribute to the mass distribution of the movable cylinder. The guide

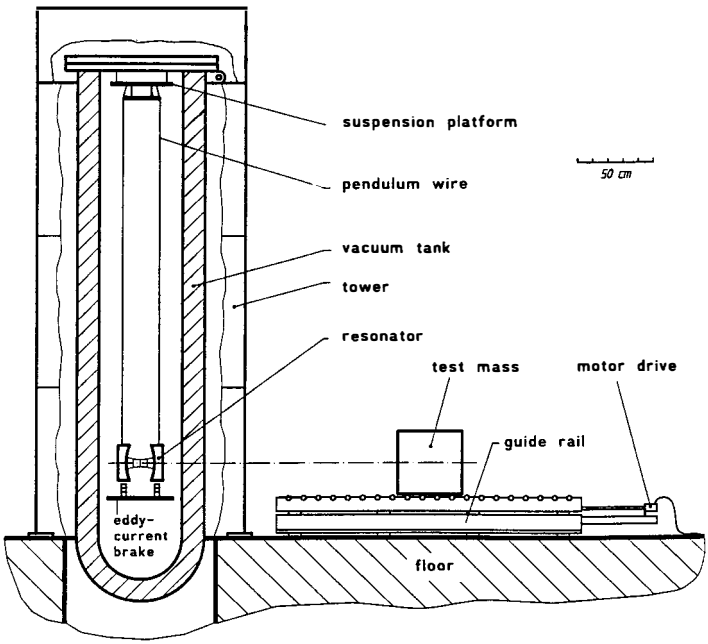


Fig. 2: Schematic arrangement of the experimental set-up.

rail is aligned such that the test mass has the same height as the resonator and moves along a line extending the resonator axis, i.e. the centers of gravity of the pendula.

III. THE EXPERIMENTAL PROCEDURE

The test mass alternates between a position with a distance r to the resonator and a reference position r_{Ref}. It rests in each position for about 15 minutes, a time which is much larger than the time constant of the pendula. The motion of the pendula is damped by eddy-current brakes and they reach their new equilibrium position in a few seconds.

The gravitational force from the moving test mass results in a change of the mirror distance Δb of maximal 12 nm and changes the resonant frequency Δf by approximately 1 kHz. The relation between Δb and Δf is determined from the measured frequency spectrum of the Fabry-Perot resonator with appropriate accuracy. An external force acting on the pendula is related to the change Δb in mirror separation via the eigenfrequency of the pendula which is determined from an additional measurement. This procedure results in a calibration bet-

ween the measured quantity, the change of the resonator frequency, and the quantity of interest, the gravitational force.

To obtain a single measurement of the resonant frequency of the chosen Fabry-Perot mode, the frequency of the microwave generator is swept across the resonance curve and the resonant frequency is calculated by a least-squares fit of the transmitted power to the Lorentzian shaped resonance curve. This procedure is repeated in equidistant time steps and results in a time series of resonant frequencies which is further analysed and processed.

The periodic motion of the test mass results in a modulation of the resonator frequency with nearly rectangular shape. The modulation amplitude is determined by means of a demodulation technique which allows a strong suppression of random noise and thermal drift of the mirror separation. So far we have achieved a resolution of the change of the mirror separation of 1 to 4 pm, depending on the integration time. Therefore, the change in mirror separation Δb can be measured with high resolution and with high precision.

IV. RESULTS

The modulation procedure described above has been repeated with the test mass in different positions (but with the same reference position), and the shift of the resonant frequency has been measured as a function of distance. The distances r' between test mass and resonator are measured from the front side of the test mass to the back of the mirrors (Fig. 1). The results of the shift of the mirror distance Δb versus the distance r' are plotted in Fig. 3.

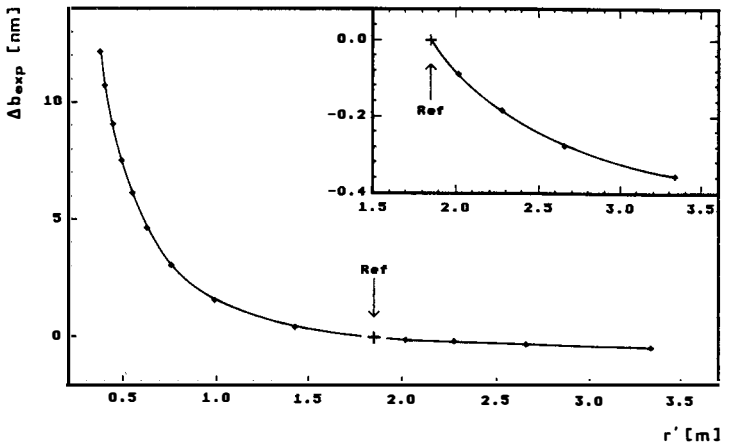


Fig. 3: The measured values of the change Δb in mirror separation due to the gravitational force of the test mass versus the distance r' of the testmass.

The corresponding distances r from the center of the test mass to the center of the pendulum next to the test mass are within the range of 0.6 m to 3.6 m. The residuals of Δb are shown in Fig. 4 in some more detail.

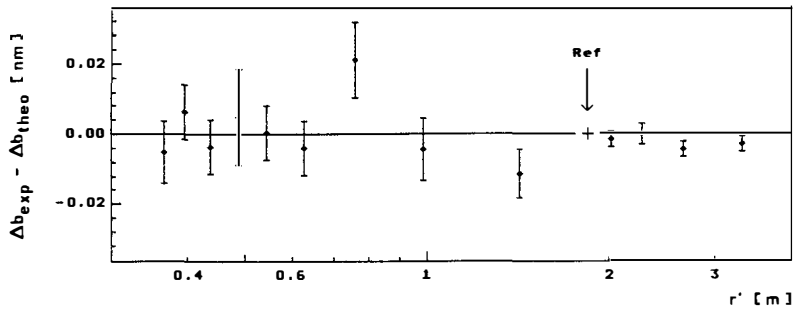


Fig. 4: The difference between measured and calculated values.

The full line in Fig. 3 is a least-squares fit of the theoretical values which are computed using Newton's inverse square law. The data points are normally distributed and no significant deviations from the inverse-square law are observed. From this fit the gravitational constant is determined to be

$$G = (6.6613 \pm 0.0011 \pm 0.0093) \cdot 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$$

and $$\frac{G - G_c}{G_c} = (-1.69 \pm 0.17 \pm 1.39) \cdot 10^{-3} .$$

The first error quoted is the statistical error as determined from the least-squares fit. The second error is the systematic error which has to be taken into account additionally [4]. The value of the gravitational constant G as determined in this experiment agrees with the CODATA-value G_c very well.

REFERENCES

1 G. Gillies, Metrologia 24,1 (1987).
2 Proceedings of the Moriond-Workshop 1986 to 1991, Les Arcs, France, Ed.: O. Fackler and J. Tran Thanh Van, Edition Frontieres, Gif-sur-Yvette.
3 J. Schurr, N. Klein, H. Meyer, H. Piel, H. Walesch, Metrologia 28, 397 (1991).
4 J. Schurr, H. Meyer, H. Piel, H. Walesch, to be published in "Lecture Notes in Physics, Proceedings Bad Honnef 1991", Springer-Verlag 1992.
5 N. Klein, University of Wuppertal, thesis, WUB-DIS 89-3,1990.
6 H. Piel, H. Chaloupka, G. Müller, to be published in "Proceedings of the 4th Int. Symposium on Superconductivity", Tokyo 1991.

This work is supported by the Deutsche Forschungsgemeinschaft, Bonn, Germany, under Grant No. Pi 118 / 4 - 1.

List of Participants

ABBES Mohamed	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE
AKHMEDOV Eugueni	S I S S A Strada Costeria 11 I- 34014 TRIESTE ITALY
BADUREK Gerald	University of Vienna Inst. f. Kernphysik A- 1020 VIENNA AUSTRIA
BALDISSERI Alberto	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
BEL Nicole	Observatoire de Paris-Meudon 5 Place Janssen F- 92195 MEUDON PRINCIPAL Cedex FRANCE
BENNETT Charles	I B M Research YORKTOWN HEIGHTS 10598 USA
BERNSTEIN Herbert	Hampshire College AMHERST Va 01002 USA
BIZZETI Pier Giorgio	Universita di Firenze Istituto di Fisica I- 50125 FIRENZE ITALY
BOUCHEZ Jacques	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
BROGGINI Carlo	Lab. Naz. Del Gran Sasso Strada Statale 17 Bis I - 67010 ASSERGI (L'Aquila) ITALY
CAMPAGNE Jean-Eric	Université Paris Sud LAL F- 91405 ORSAY Cedex FRANCE
CELNIKIER Ludwik	Observatoire de Paris-Meudon 5 Place Janssen F- 92195 MEUDON PRINCIPAL Cedex FRANCE

CHARDIN Gabriel	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
CHARLES François	Univ. Paris VI et VII LPNHE 4, Place Jussieu F- 75005 PARIS Cedex FRANCE
CIUFOLINI Ignazio	C N R Inst. Fisica Spazio Interplanetario 00044 FRASCATI ITALY
CLELAND Andrew	C E N Saclay S P E C F- 91191 GIF SUR YVETTE Cedex FRANCE
CUSSONNEAU Jean-Pierre	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE
DAMOUR Thibault	IHES Inst. Hautes Etudes Scientif. F- 91440 BURES-SUR-YVETTE FRANCE
DANEVICH Fedor	I N R Ukrainian Academy of Sciences U- 252028 KIEV 28 UKRAINE
de KERRET Hervé	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE
DE RUJULA Alvaro	CERN TH Division CH-23 1211 GENEVE SWITZERLAND
DREXLIN Guido	Kernforschungszentrum Institut für Kernphysik D- 7500 KARLSRUHE 1 GERMANY
DUMARCHEZ Jacques	Univ. Paris VI et VII LPNHE 4, Place Jussieu F- 75005 PARIS Cedex FRANCE
ECKHARDT Donald	PL/G P E HANSCOM AFB MA 01730 USA

ESPOSITO FARESE Gilles	IHES Inst. Hautes Etudes Scientif. F- 91440 BURES-SUR-YVETTE FRANCE
FACKLER Orrin	University of California Lawrence Livermore Nat. Lab. LIVERMORE CA 94550 USA
FALLER Jim	University of Colorado JILA BOULDER CO 80309-0440 USA
FAVIER Jean	LAPP Lab. de Physique des Particules F- 74941 ANNECY LE VIEUX Cedex FRANCE
FISCHBACH Ephraim	Purdue University Dept of Physics WEST LAFAYETTE IN 47907 USA
FREDKIN Ed	University of Boston Dept of Physics BOSTON MA 02215 USA
FRIED Herbert	Université de Nice Lab. de Physique Théorique F- 060034 NICE Cedex FRANCE
FUNK Wolfgang	Universität Heidelberg Institut für Hochenergiephysik D- 6900 HEIDELBERG 1 GERMANY
GARRETA Denis	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
GERBIER Gilles	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
GIAMMARCHI Marco	Universita di Milano Istituto di Fisica I- 20133 MILANO ITALY
GONZALEZ-MESTRES Luis	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE

GOUGH Douglas	University of Cambridge Cavendish Lab. - High Energy Phys. Group CAMBRIDGE CB3 OHE UNITED KINGDOM
GREENE Geoffrey	N I S T GAITHERSBURG MD 20899 USA
GRIFOLS Tony	Univ. Autonoma de Barcelona Faculdad de Ciencias E 08193 BELLATERRA Barcelona SPAIN
HELLINGS Ronald	Jet Propulsion Laboratory PASADENA CA 91109 USA
HINDS Edward	Yale University Department of Physics NEW HAVEN CT 06511 USA
HOLZSCHUH Eugen	Universität Zurich Physik Institut CH- 8001 ZURICH SWITZERLAND
IACOPINI Enrico	Scuola Normale Superiore Istituto di Fisica I- 56100 PISA ITALY
JOHNSTON Janet	European Office of Aerospace & Development LONDON N W 1 5TH UNITED KINGDOM
KAISER Helmut	Univesity of Missouri Research Reactor COLUMBIA MO 65211 USA
KAJFASZ Eric	Faculté des Sciences Centre de Phys. des Particules F- 13288 MARSEILLE Cedex 02 FRANCE
KARASIEWICZ Alain	Institut Henri Poincaré Lab. de Physique Théorique F- 75231 PARIS Cedex FRANCE
KAYSER Boris	N S F Division of Physics WASHINGTON DC 20550 USA

KRASTEV Plamen	CERN TH Division CH-23 1211 GENEVE SWITZERLAND
LEFIEVRE Bernard	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE
LESLIE James	Queen's University Dept of Physics KINGSTON Ontario K7L 3N6 CANADA
LOPES Ilidio	C E N Saclay DAPNIA/SAP F- 91191 GIF sur YVETTE Cedex FRANCE
LORENZ Eckhart	Max Planck Inst. für Physik Föhringer Ring, 6 D- 8000 MUNICH 40 GERMANY
MAREL Gérard	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
MARTEMYANOV V.P.	I.V. Kurchatov Inst. of Atomic Energy R 123 182 MOSCOW RUSSIA
MASCHUW Reinhard	Kemforschungszentrum Institut für Kemphysik D- 7500 KARLSRUHE 1 GERMANY
MELLITI Tahar	Institut Henri Poincaré Lab. de Physique Théorique F- 75231 PARIS Cedex FRANCE
MEYER Hinrich	University of Wuppertal FB Physik D- 5600 WUPPERTAL 1 GERMANY
MITTLEMAN Richard	University of Washington Dept of Physics SEATTLE WA 98195 USA
MOORE Michael	University of Washington Dept of Physics SEATTLE WA 98195 USA

MORRISON Douglas	CERN Division PPE CH- 1211 GENEVE 23 SWITZERLAND
NEUMAIER Stefan	Fakultät f. Physik Techn. Univ. München D- 8046 GARCHING GERMANY
NEWMAN Riley	Univ. of California/Irvine Dept of Physics IRVINE CA 92717 USA
NICO Jeffrey	L A N L Physics Division LOS ALAMOS NM 87545 USA
NOLTE Eckehart	Fakultät f. Physik Techn. Univ. München D- 8046 GARCHING GERMANY
NORDTVEDT Kenneth	IHES Inst. Hautes Etudes Scientif. F- 91440 BURES-SUR-YVETTE FRANCE
OBERAUER Lothar	Fakultät f. Physik Techn. Univ. München D- 8046 GARCHING GERMANY
OBOLENSKY Michel	Collège de France Lab. de Physique Corpusculaire F- 75231 PARIS Cedex 05 FRANCE
OYAMA Yuichi	KEK National Lab. for High Energy Physics 305 IBARAKI KEN JAPAN
PAIN Reynald	Univ. Paris VI et VII LPNHE 4, Place Jussieu F- 75005 PARIS Cedex FRANCE
PESSARD Henri	LAPP Lab. de Physique des Particules F- 74941 ANNECY LE VIEUX Cedex FRANCE
PETCOV Serguey	S I S S A Strada Costeria 11 I- 34014 TRIESTE ITALY

PETERSEN Jon Wulff	University of Aarhus Institut of Physics DK- 8000 AARHUS C DENMARK
PIEPKE Andreas	Max Planck Institut für Kernphysik D- 6900 HEIDELBERG GERMANY
PIILONEN Leo	Virginia Polytechnic Inst. & State Univ. BLACKBURG Va 24061-0435 USA
PINSONNEAULT Marc	Yale University Sloane Lab. - Dept of Physics NEW HAVEN CT 06520 USA
POMANSKY Alexander	I N R Baksan Neutrino Obs. 361 609 BAKSAN RUSSIA
RADCLIFFE Tom	Medical Physics, MCTRF WINNIPEG R3E OV9 CANADA
RAMSEY Norman	Harvard University Dept. of Physics CAMBRIDGE MA 02138 USA
RICCI Fulvio	I N F N Universita di Roma "La Sapienza" I- 00100 ROMA ITALY
SCHURR Jürgen	University of Wuppertal FB Physik D- 5600 WUPPERTAL 1 GERMANY
SENJANOVIC Goran	Inter. Centre for Theor. Physics Strada Costeria 11 I- 34 100 TRIESTE ITALY
PERRONE Franco	INFN - Lab. di San Piero S. Piero a Grado I- 56010 PISA ITALY
SROMICKI Jurek	Eidg. Technische Hochschule Institüt fur Hochenergiephysik CH- 8093 ZÜRICH SWITZERLAND

SUKHOTIN S.V.	I.V. Kurchatov Inst. of Atomic Energy R 123 182 MOSCOW RUSSIA
TRAN THANH VAN Jean	Université Paris Sud LPTHE F- 91405 ORSAY Cedex FRANCE
TOSHEV Stoyan	Bulgarian Academy of Sciences Inst. f. Nuclear Res. & Nuclear Energy 1784 SOFIA BULGARIA
TURCK-CHIEZE Sylvaine	C E N Saclay DAPNIA/SAP F- 91191 GIF sur YVETTE Cedex FRANCE
TURLAY René	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
VEILLET Christian	CERGA F- 06130 GRASSE FRANCE
VUILLEUMIER Jean-Claude	Université de Neuchâtel Institut de Physique CH- 2000 NEUCHATEL SWITZERLAND
VYRODOV V.N.	I.V. Kurchatov Inst. of Atomic Energy R 123 182 MOSCOW RUSSIA
WANG Ning	CEN Saclay DAPNIA/SPP F- 91191 GIF sur YVETTE Cedex FRANCE
WEBER Joseph	University of Maryland Physics Department COLLEGE PARK MD 20742 USA
WEYER Gerd	University of Aarhus Institut of Physics DK- 8000 AARHUS C DENMARK
WILKERSON John	L A N L Physics Division LOS ALAMOS NM 87545 USA

YOUNG Albert	Caltech 161-33 PASADENA CA 91125 USA
ZACEK Gabriele	CERN Division PPE CH- 1211 GENEVE 23 SWITZERLAND
ZANOTTI Luigi	Universita di Milano Istituto di Fisica I- 20133 MILANO ITALY
ZDESENKO Yuri	I N R Ukrainian Academy of Sciences U- 252028 KIEV 28 UKRAINE
ZEILINGER Anton	Universität Innsbruck Institut für Experimentalphysik A- 6020 INNSBRUCK AUSTRIA
ZUMBERGE Mark	University of California San Diego LA JOLLA CA 92093-0225 USA
ZUREK Wojciech	Los Alamos National Lab. Theoretical Astrophysics LOS ALAMOS NM 87545 USA