Non-accelerator Particle Physics

Related titles from Institute of Physics Publishing

Gauge Theories in Particle Physics Second edition
I J R Aitchison and A J G Hey

Introduction to Gauge Field Theory Revised edition
D Bailin and A Love

Supersymmetric Gauge Field Theory and String Theory
D Bailin and A Love

Hadron Interactions
P D B Collins and A D Martin

Electron–Positron Annihilation Physics
Edited by B Foster

The Weak Interaction in Nuclear, Particle and Astrophysics
K Grotz and H V Klapdor

Nuclear Physics in the Universe
Edited by M W Guidry and M R Strayer

Nuclei in the Cosmos
F Käppeler and K Wisshak

Physics of the Early Universe
Edited by J A Peacock, A F Heavens and A T Davies

Non-accelerator Particle Physics

H V Klapdor-Kleingrothaus

Max-Planck-Institut für Kernphysik
Heidelberg

and

A Staudt

Bayer AG
Leverkusen

Translated by S S Wilson

Institute of Physics Publishing
Bristol and Philadelphia

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN 0 7503 0305 0 hardback
0 7503 0502 9 paperback

Library of Congress Cataloging-in-Publication Data are available

Simultaneous German edition, *Teilchenphysik ohne Beschleuniger*, published by B G Teubner GmbH, Stuttgart

Consultant Editors for Institute of Physics Publishing: Professor Euan Squires, University of Durham, and Dr Brian Foster, University of Bristol

First published 1995 (hardback)
Reprinted with corrections 1998 (paperback)

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London

Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK

US Editorial Office: Institute of Physics Publishing, The Public Ledger Building, Suite 1035, 150 South Independence Mall West, Philadelphia, PA 19106, USA

Typeset in TEX using the IOP Bookmaker Macros
Printed in the UK by J W Arrowsmith Ltd, Bristol

Contents

Preface to the English edition

Many central questions of particle physics are beyond the capabilities of modern accelerators. They can, however, to some extent be investigated via non-accelerator experiments. These may be purely terrestrial laboratory experiments or they may exploit some of the various connections of particle physics with astrophysics, cosmology and the evolution of the universe. It is particularly relevant to gain a better awareness of such experiments, and of the resulting possibilities, at a time when elementary particle physics is now imposing very extreme requirements on new generations of accelerators. The rapid development and increasing importance of non-accelerator experiments in recent years, in numerous underground laboratories and elsewhere, clearly speaks for itself. In this book we review and discuss the most important of these advances and their consequences for a better understanding of particle physics. In so doing, we hope to fill a gap in the current literature.

Thus, the book also provides an insight into current aspects of modern physics. Stemming from seminars at the University of Heidelberg, it is applicable to undergraduate students of physics and to readers interested in topical questions of modern physics and, in particular, in the close relationships between particle, nuclear and astrophysics.

We are indebted to Dr Martin Hirsch (MPI, Heidelberg) and Dr Irina Krivosheina (Radiophysical Institute, Nishny Novgorod) for their critical readings of the manuscript, to Dr Kai Zuber (University of Heidelberg) for useful discussions, and to Professor Euan Squires (University of Durham) and Dr Brian Foster (University of Bristol) for their reviews of the text and their useful advice. We thank Frau Veronika Traeumer for her untiring technical assistance in the generation of the figures, Dr Stephen S Wilson for his translation from the German original, which has been published by B G Teubner GmbH, Stuttgart, and Mr Jim Revill and Mr Martin Beavis of Institute of Physics Publishing for their faithful collaboration in the publication of this English edition.

H V Klapdor-Kleingrothaus **A Staudt**
Heidelberg/Cologne, March 1995

Acknowledgments

The authors and publisher gratefully acknowledge permission to reproduce previously published material, including many figures and tables taken from the extensive journal literature, as granted by authors and publishers, and as indicated by citations in the captions. They have attempted to trace the copyright holders of all material reproduced from all sources and apologise to any copyright holders whose prior permission might not have been obtained.

The following figures and tables are reproduced from the cited sources, as referenced in the Bibliography, by permission of the copyright holders acknowledged below.

Springer-Verlag GmbH & Co. KG: figures 3.1, 3.2 and 8.4 from [Bür88]; figure 7.5 from [Fei88a]; figure 6.30 from [Avi88].

Cambridge University Press: figures 7.18 and 7.21 from [Bah88]; figures 6.7, 6.45, 7.4 and 7.27 from [Boe92].

Addison-Wesley Publishing Company: figure 2.17 from [Qui83]; figure 3.4 from [Kol90].

Nature (Copyright 1990 Macmillan Magazines Limited): figure 9.4 from [Val90].

Plenum Publishing Corporation: figure 8.7 from [Ahl83].

Scientific American: figure 8.3 from [Car82].

World Scientific Publishing Co Pte Ltd: figure 7.14 from [Pes88]; figure 7.1 from [Kay89].

Annual Review of Nuclear and Particle Science: figure 5.4 (Copyright 1982 Annual Reviews Inc) from [Ram82]; figures 10.4, 10.5, 10.6 (Copyright 1989 Annual Reviews Inc) from [Smi89].

National Academy Press: figures 2.3, 2.5(a) and 2.24 (Copyright 1986 National Academy of Sciences) from [Phy86].

R Piper Gmbh & Co. KG: figures 1.9, 1.11, 2.5(b), 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12 and 2.13 from [Scho89].

Chapter 1

Modern Elementary Particle Theories

Elementary particle physics is primarily concerned with research into the elementary building blocks of our world and the fundamental forces which operate between these objects. This reflects man's ancient search for simple general laws which explain the variety and complexity of the world today.

The simplest approximation to a unified description of nature may be obtained by thinking in terms of elementary particles and the interactions between these. Here, one hopes, in principle, to be able to trace the macroscopic physical and chemical processes back to these basic units. Historically, this doctrine of 'atomism' was first introduced by Democritus (around 460–370 BC). However, the first conclusive proof of the existence of atoms was given much later.

In particle physics, the 'indivisible' atoms, which had already become the fundamental building blocks for many purposes in chemistry and physics, were decomposed further. Atomic physics distinguishes between the atomic shell and the nucleus of an atom. Nuclear physics teaches us that atomic nuclei are constructed from nucleons, namely the protons and the neutrons. According to elementary particle physics, nucleons also have an internal structure which is explained in terms of quarks and gluons, while there are those who suggest that even quarks may be constructed from preons or other subquarks.

Our theoretical understanding of the fundamental particles and interactions is largely based on two important milestones of modern physics, namely quantum field theory (QFT) and the recognition of the importance of symmetry principles.

We now include a brief discussion of the elementary particles and interactions known today, before describing symmetries and the basic ideas of gauge theories and grand unified theories.

1.1 THE ELEMENTARY CONSTITUENTS OF MATTER

1.1.1 Introduction

Elementary particle physics begins with research into neutrons and protons, the constituents of the nucleus, and ultimately leads to ideas about the quark–lepton structure of matter. Our current idea about the structure of matter may be roughly summarized as follows:

(i) Atoms, and molecules built from atoms, form the basis of chemical substances.

(ii) Atoms themselves consist of an electrically positively charged atomic nucleus (diameter $d \simeq 10^{-15}$ m) which is surrounded by the negatively charged electrons of the atomic shell ($d \simeq 10^{-10}$ m).

(iii) The atomic nucleus consists of protons and neutrons, which are collectively known as nucleons.

(iv) However, scattering experiments with high-energy leptons have shown that the nucleons themselves are constructed from even more elementary building blocks. In a simple model each nucleon may be thought of as consisting of three quarks. It is currently thought that the quarks together with the leptons (e, μ, τ, ν_e, ν_μ, ν_τ) are elementary particles in the proper sense, since no internal structures have been detected to date, i.e. they may be considered to be point-like (diameter $d < 10^{-17}$ m).

1.1.2 Leptons and quarks

The elementary constituents of matter, leptons and quarks, are fermions, i.e. particles with a half-integral spin.

Leptons are, in addition to gravity, also subject to the weak, and, if they are charged, to the electromagnetic interaction. They are not involved in the strong interaction. Quarks, on the other hand, are subject to all four forces.

Six leptons are currently known, namely the electron (e^-), the muon (μ^-) and the tauon (τ^-), together with three neutrinos (a neutral neutrino is assigned to each of the three charged leptons; the ν_τ has not yet been detected directly). Thus, we have the following three lepton pairs, arranged in order of increasing mass of the charged leptons:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}. \tag{1.1}$$

Some of their properties are summarized in table 1.1.

Analogously to the leptons, quarks exist in several so-called flavours. Five types of quark have been detected in experiments to date, namely, the up quark (u), the down quark (d), the strange quark (s), the charm quark (c) and the

Table 1.1. Properties of the leptons (L_i = flavour-related lepton number, $L = \sum_{i=e,\mu,\tau} L_i$).

Lepton	$Q[e]$	L_e	L_μ	L_τ	L
e^-	−1	1	0	0	1
ν_e	0	1	0	0	1
μ^-	−1	0	1	0	1
ν_μ	0	0	1	0	1
τ^-	−1	0	0	1	1
ν_τ	0	0	0	1	1

bottom (beauty) quark (b). It is conjectured that there exists a further sixth flavour, the top quark (t) (see chapter 2)[1].

Like the leptons, the quarks may be arranged in pairs, where there is a difference in electrical charge of one unit ($\Delta Q = 1e$) between the two particles of a pair

$$\begin{bmatrix} Q = 2/3e \\ Q = -1/3e \end{bmatrix} \quad \begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}. \tag{1.2}$$

The properties of the five known quarks (and the expected sixth) are summarized in table 1.2.

Table 1.2. Properties of the quarks.

Flavour	Spin	B	I	I_3	S	C	B^*	T	$Q[e]$
u	1/2	1/3	1/2	1/2	0	0	0	0	2/3
d	1/2	1/3	1/2	−1/2	0	0	0	0	−1/3
c	1/2	1/3	0	0	0	1	0	0	2/3
s	1/2	1/3	0	0	−1	0	0	0	−1/3
b	1/2	1/3	0	0	0	0	−1	0	−1/3
t	1/2	1/3	0	0	0	0	0	−1	2/3

I = Isospin, S = Strangeness, C = Charm, Q = Charge, B = Baryon number, B^* = Bottom, T = Top.

The quantum numbers shown in table 1.2 (baryon number B, strangeness S, charm C, bottom (or beauty) B^*, isospin I and the isospin component I_3)

[1] For first experimental indication see [Abe94, Aba95].

satisfy the (extended) Gell-Mann–Nishijima formula

$$Q = I_3 + \tfrac{1}{2}(B + S + C + B^*). \tag{1.3}$$

The given elementary fermions may be arranged in three families (generations):

$$\begin{pmatrix} e^- \\ \nu_e \\ u \\ d \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \\ c \\ s \end{pmatrix} \quad \begin{pmatrix} \tau^- \\ \nu_\tau \\ t \\ b \end{pmatrix} \quad \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \tag{1.4}$$

1st generation 2nd generation 3rd generation 4th generation?

Each family contains two leptons and two quarks. The individual generations essentially differ according to their particle masses (figure 1.1). The first generation contains the lightest quarks, the lightest charged lepton and, by conjecture, the lightest of the three neutrinos. Thus, all stable matter is constructed from members of the first generation.

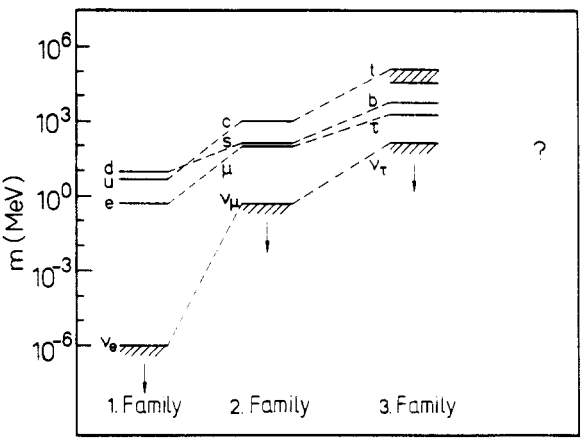


Figure 1.1. Mass spectrum of the known elementary fermions. The dotted lines link the corresponding particles of the different families.

The arrangement into three families reflects the behaviour under the strong, the electromagnetic and the weak interaction. Corresponding members of the different families are equivalent under these different forces. Differences only arise in the case of gravity, as a result of the mass dependence.

We shall return to the structure of the individual families in our discussion of the grand unified theories. The fact that the number of quark families is the same as the number of lepton families, in some sense guarantees that the underlying elementary particle theory is anomaly-free (see [Gro89, 90, 92]).

Since, according to the results from LEP to be discussed in chapter 2, there are only three (light) neutrino types, it is likely that there are only three generations.

It is conjectured that the quarks and leptons discussed above are not the smallest building blocks of matter, but are themselves composed of more elementary particles or subquarks (e.g. preons) (for so-called composite models see e.g. [Moh86a, Schr85b]). The various families could then be interpreted, for example, as different excitation states of bound preon systems. Such speculation is triggered by the historical development to date and by the fact that the current number of elementary particles is again relatively large.

1.1.3 Antiparticles

To every particle there corresponds an antiparticle. Particles and antiparticles have the same mass, the same spin, the same isospin and, if they are unstable, the same lifetime. However, they differ in the sign of all additive quantum numbers (see section 1.3), such as the electrical charge and the baryon number.

The existence of antiparticles with precisely these properties is a fundamental consequence of relativistic quantum field theory (see section 1.3.1). The positron, the antiparticle of the electron, the existence of which was predicted historically from solutions of the Dirac equation with negative energy, was discovered in cloud chamber photographs by Anderson in 1932 [And32]. The antiparticle of a fermion f is often denoted by $\overline{f}$ or f^C, where C denotes the charge-conjugation operator.

1.1.4 Construction of the hadrons from quarks

The collective term 'hadrons' covers all particles which are subject to the strong interaction. They are divided into two classes:

(i) mesons (integer spin (bosons), e.g. $\pi^0, \pi^{\pm}$);
(ii) baryons (half-integer spin (fermions), e.g. n, p, Δ^0).

We now believe that hadrons are constructed from quarks. Experiments involving the scattering of electrons and neutrinos by protons have shown that the proton has a structure and a size of approximately 1 fm.

In the quark model, baryons consist of three quarks (strictly speaking, valence quarks; we shall not go into the complications due to sea quarks here). Quarks (see table 1.2) have a half-integer spin and third-integer charges and baryon numbers. The quark content of protons and neutrons is as follows:

$$p \equiv uud \qquad n \equiv udd. \tag{1.5}$$

Mesons have baryon number $B = 0$; thus, they consist of an identical number of quarks and antiquarks. The simplest assumption of a quark–antiquark pair

already provides a successful description of the classification of the known mesons. For example, one finds that

$$\pi^+ \equiv u\bar{d} \qquad \pi^- \equiv d\bar{u} \qquad K^0 \equiv d\bar{s}. \tag{1.6}$$

In summary, qqq systems may be identified with the baryons, while $q\bar{q}$ systems may be identified with the mesons:

$$\begin{aligned} &\text{Baryons} : qqq \\ &\text{Mesons} : q\bar{q} \end{aligned} \tag{1.7}$$

1.2 THE ELEMENTARY INTERACTIONS

1.2.1 Introduction

Phenomenologically, the interactions between the elementary fermions mentioned above may be traced back to four fundamental forces. In order of decreasing strength, these are the colour interaction (strong interaction between the quarks), the electromagnetic interaction, the weak interaction and gravity. The different strengths are apparent from the physical phenomena. However, a quantitative comparison is difficult since the phenomenologically defined coupling constants are of different dimensions (see table 1.3), while dimensionless quantities are required for a comparison. In addition to these four known interactions, further hypothetical forces are predicted by grand unified theories (GUTs), which try to give a unified description of these phenomenological forces.

1.2.2 The concept of interaction in modern quantum field theories

There now follows a short discussion of the development of the physical concept of forces over the centuries. In the Newtonian theory of gravity (17th century) force is understood to mean the action of two distant bodies on one another (action at a distance). The concept of the field was introduced in electrodynamics. The presence of a charge modifies the space around the charge, giving rise to an electrical field with a certain energy density. The force on a sample charge is generated by the action of the field at the location of the sample charge. The concept of action at a distance is replaced by that of a proximity effect of the field.

With the introduction of quantum mechanics it became necessary to quantize the fields. In modern quantum field theories the interactions are described by the exchange of field quanta. The most important properties of the exchange particles of the four fundamental forces and the hypothetical GUT interaction are listed in table 1.4. The exchanged bosons have a four-momentum vector which

Table 1.3. Phenomenology of the four fundamental forces and the hypothetical GUT interaction.

Interaction	Strength	Range R	Exchange particle	Example
Weak	$G_F \simeq 1.02 \times 10^{-5} m_p^{-2}$	$\approx m_W^{-1}$ $\simeq 10^{-3}$ fm	$W^{\pm}$, Z^0	β-decay
Electro-magnetic	$\alpha \simeq 1/137$	∞	γ	Forces between electrical charges
Strong (nuclear-)	$g_\pi^2/4\pi \approx 14$	$\approx m_\pi^{-1}$ ≈ 1.5 fm	Gluons	Nuclear forces
Strong (colour-)	$\alpha_s \simeq 1$	Confinement	Gluons	Forces between the quarks
Gravity	$G_N \simeq 5.9 \times 10^{-39}$	∞	Graviton?	Mass attraction
GUT	$m_X^{-2} \approx 10^{-30} m_p^{-2}$ $m_X \approx 10^{15}$ GeV	$\approx m_X^{-1}$ 10^{-16} fm	X, Y	p-decay

Table 1.4. Properties of the exchange bosons.

Boson	Interaction	Spin	Mass [GeV/c^2]	Colour charge	Electrical charge	Weak charge
Gluons	Strong	1	0	Yes	0	No
γ	Electromagnetic	1	0	No	0	No
$W^{\pm}$; Z^0	Weak	1	81.8; 91.2	No	± 1; 0	Yes
Graviton	Gravity	2	0	No	0	No
X; Y	GUT	1	$\sim 10^{15}$	Yes	$\pm 4/3$; $\pm 1/3$	Yes

corresponds not to that of a free particle but to that of a particle for which the square of the mass is negative. Thus, we speak of virtual exchange particles.

Interaction processes as exchange processes are often represented graphically as Feynman diagrams (see figures 1.2–1.4). These graphs represent formal rules for calculating the interaction cross sections for the corresponding processes (see e.g. [Ait89, Gro89, 90]).

The elementary components of a Feynman diagram are vertices at which a fermion is associated with an exchange boson (figure 1.2(a)–(c)). There are also boson–boson vertices. Thus, a vertex represents a linkage between different particles at a point in space–time.

Interaction processes are represented by combining two elementary vertices,

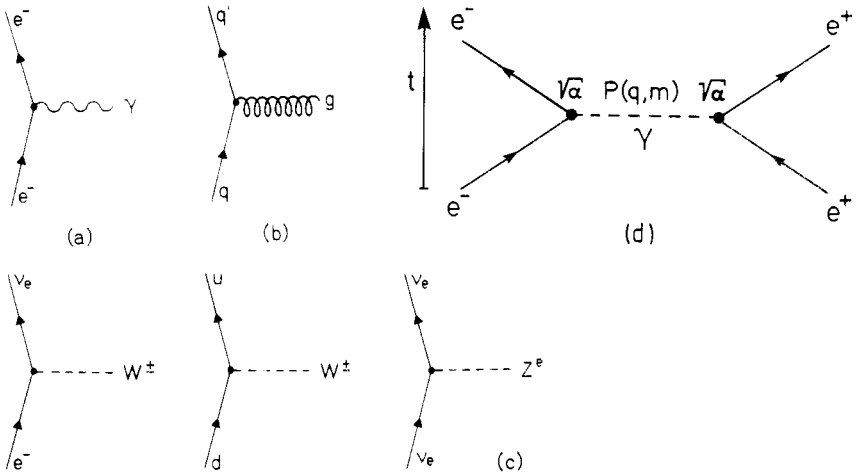


Figure 1.2. Elementary vertices for (*a*) electromagnetic, (*b*) strong and (*c*) weak interactions, and Feynman diagram (*d*) for e^+e^- scattering.

taking into account conservation principles (quantum numbers), where one thinks in terms of a predefined time direction running from bottom to top or from left to right.

Figure 1.2(*d*) shows a simple Feynman diagram. We use this diagram to outline how to estimate the probability of the illustrated e^+e^- scattering. A virtual photon couples to a charged lepton with amplitude $\sqrt{\alpha}$; here, α denotes the fine structure constant, i.e. the coupling constant of the electromagnetic interaction. The probability amplitude for the process as a whole is the product of the amplitudes at the two vertices and a propagator term which describes the exchange of the virtual boson:

$$\text{Probability amplitude} = \text{Coupling} \times \text{Propagator} \times \text{Coupling}. \tag{1.8}$$

For the process shown in figure 1.2(*d*) we have

$$A \propto \sqrt{\alpha} P(q, m) \sqrt{\alpha} \tag{1.9}$$

where $P(q, m)$ is the propagator term. $P(q, m)$ may be derived from a description in terms of a Green's function (see [Nac86]). We find the following dependence on the mass of the field quantum m and the transmitted four-momentum q_μ

$$P(q, m) \propto \frac{1}{q^2 - m^2}. \tag{1.10}$$

Thus, by virtue of (1.9), we have

$$A \propto \frac{\alpha}{q^2 - m^2}. \tag{1.11}$$

The interaction cross section is proportional to the square of the probability amplitude A. We obtain

$$\frac{d\sigma}{dq^2} \propto \frac{\alpha^2}{(q^2 - m^2)^2}. \tag{1.12}$$

If we neglect the mass of the exchange boson in comparison with the transferred momentum, (1.12) becomes the well-known Rutherford scattering cross section

$$\frac{d\sigma}{dq^2} \propto \frac{\alpha^2}{q^4}. \tag{1.13}$$

In the case of the e^+e^- scattering, this approximation is permissible, since the field quantum, the photon, has a vanishing mass. Equation (1.12) indicates that the behaviour of an interaction with distance is closely related to the mass of the field quantum (see section 1.2.3).

In gauge theories, the interaction structures follow from symmetry considerations which, in particular, also determine the number of exchange quanta.

1.2.3 The range of an exchange interaction

In the previous section we have seen that interactions are mediated by the exchange of bosons. These processes can only take place and be understood in the framework of the Heisenberg uncertainty principle. The spontaneous emission of a photon by an electron violates the principle of energy conservation. Thus, the carrier particles, in this case the photon, can only exist for a short time Δt given by

$$\Delta E \, \Delta t \geq \hbar \tag{1.14}$$

before they are annihilated in a second process. Thus, we speak of virtual exchange particles, although these may result in a real effect.

If the observation time is limited to the interval Δt, the energy can only be determined accurately down to

$$\Delta E \simeq \frac{\hbar}{\Delta t} \tag{1.15}$$

i.e. a brief occurrence of a particle with mass

$$m \simeq \frac{\Delta E}{c^2} \simeq \frac{\hbar}{\Delta t c^2} \tag{1.16}$$

is possible.

The range of the force mediated by bosons of mass m can now be estimated if we assume that the exchange quantum essentially moves forward at the speed of light. In the time Δt, it covers a distance

$$R \simeq c\Delta t \simeq c\frac{\hbar}{\Delta E} \simeq \frac{c\hbar}{mc^2}. \tag{1.17}$$

Thus, the range is

$$R = \frac{\hbar}{mc}. \tag{1.18}$$

According to (1.18), the short range of the weak interaction is directly related to the large mass of the intermediate $W^{\pm}$ and Z^0 vector bosons.

Equation (1.18) may be derived in a mathematically more rigorous way from the Klein–Gordon equation. This relativistic wave equation describes scalar ($S = 0$) (and also vector ($S = 1$)) particles (see e.g. [Gro89, 90]). Based on the relativistic energy–momentum relation for a free particle

$$E^2 = p^2c^2 + m^2c^4 \tag{1.19}$$

we obtain the Klein–Gordon equation by carrying out the following replacement, analogous to the 'derivation' of the Schrödinger equation:

$$p \to -\mathrm{i}\hbar\nabla \qquad E \to \mathrm{i}\hbar\frac{\partial}{\partial t}. \tag{1.20}$$

From (1.19), by multiplication with ψ, we obtain

$$\frac{\partial^2\psi}{\partial t^2} - c^2\nabla^2\psi + \frac{m^2c^4}{\hbar^4}\psi = 0. \tag{1.21}$$

Unlike the Dirac spinors, the wavefunction ψ is a scalar. m denotes the rest mass.

A stationary solution of (1.21) is given by

$$\psi = \psi_0\frac{\mathrm{e}^{-\lambda r}}{r} \tag{1.22a}$$

where

$$\lambda = \frac{mc}{\hbar}. \tag{1.22b}$$

As the distance increases, the wavefunction (1.22a) vanishes exponentially. The range is defined by

$$R = \frac{1}{\lambda} = \frac{\hbar}{mc}. \tag{1.23}$$

In 1935, such considerations led Yukawa [Yuk35] to predict the existence of particles with a mass of approximately 200 MeV/c^2 as field quanta of the nuclear

forces. The latter typically have a range in the region of 1 fm, so that from (1.23), we obtain

$$m \simeq \frac{\hbar}{c \times 1 \text{ fm}} \simeq 197 \text{ MeV}/c^2. \tag{1.24}$$

These new particles were actually discovered later in cosmic rays [Lat47, Bjo50]. They are the mesons which we now know as pions ($\pi^{\pm}, \pi^0$).

According to Yukawa's idea, the forces between nucleons come about through the exchange of pions. Under this assumption many phenomena of nuclear physics have qualitatively and quantitatively satisfactory explanations. However, according to the current view, the nuclear forces may be traced back to more elementary processes, namely the colour interaction between the constituents of the nucleons (the quarks) which is mediated by gluons. The classical nuclear forces are essentially only a type of residual interaction analogous to the van der Waals forces which are a residual interaction of the electromagnetic force. But the idea of an exchange interaction survives.

According to (1.18) the range of an interaction is determined by the mass of the exchange quanta. Since the photons, which mediate the electromagnetic interaction, are massless, this interaction has infinite range. The following simple consideration even enables us to derive the $1/r^2$ behaviour of Coulomb's law:

The exchange of a virtual boson results in a change in the momentum Δp. According to the laws of mechanics we have

$$F = \frac{\Delta p}{\Delta t}. \tag{1.25}$$

A massless particle moves at the speed of light c, whence

$$\Delta t = \frac{r}{c}. \tag{1.26}$$

The distance covered, r, is determined by Heisenberg's uncertainty principle

$$r \, \Delta p \geq \frac{\hbar}{2\pi}. \tag{1.27}$$

If we replace the $\geq$ in (1.27) by $\approx$, by virtue of (1.25) and (1.26), we obtain

$$F \propto \frac{1}{r^2}. \tag{1.28}$$

The short-range *strong* interaction may also be traced back to the exchange of massless field quanta (the gluons) between the quarks. The apparent contradiction between the short range and the masslessness of the gluons is resolved by the fact that in addition to the quarks the gluons themselves also have a colour charge and thus interact *among themselves*, in particular, since one-gluon exchange does not occur because the hadrons are colourless (see e.g. [Gro89, 90, 92]).

1.2.4 Phenomenology of the known interactions

Four known forces act between the elementary particles: gravity, electromagnetic interaction, strong interaction, weak interaction.

1.2.4.1 Gravity

The gravity between two elementary particles is linked to their mass. However, for masses of elementary particles this effect is so weak that it can be neglected in most discussions in this book; but it plays a role, for example, at the very high energies of the early universe (see chapter 3). One major problem both in particle physics and in cosmology is that there is as yet no renormalizable quantum theory of gravity.

1.2.4.2 Electromagnetic interaction

The electromagnetic interaction takes place between all electrically charged particles. It is successfully described by a quantum field theory, quantum electrodynamics (QED). In the macroscopic area the masslessness of the photons leads to a $1/r^2$ force law with infinite range. Predicted quantum field effects, such as the Lamb shift and the deviation of the electron g factor from two, have been confirmed in precision experiments.

1.2.4.3 Strong interaction

The strong interaction holds together the nucleons in the atomic nucleus and the quarks in the nucleon. The forces between quarks are explained by analogy with the electromagnetic interaction, where the so-called colour charge replaces the electrical charge.

Measurements of the ratios of the annihilation cross sections of electrons and positrons $\sigma(e^+e^- \rightarrow q\bar{q} \rightarrow \text{hadrons})$ and $\sigma(e^+e^- \rightarrow \mu^+\mu^-)$ show (see e.g. [Gro89, 90, 92]) that each quark has three degrees of freedom which are assigned to the colour-charge states (e.g. red, blue and green). The force between the quarks results from the exchange of field quanta with spin 1, the gluons. A field-theoretical description is given in quantum chromodynamics (QCD), based on an SU(3) symmetry. A colour interaction may change the colour charge but not the flavour of a quark. One consequence of the particular properties of the colour interaction is that all physical systems (hadrons) are outwardly neutral in colour or 'white' (qqq or $q\bar{q}$); in other words, they form a singlet for the colour interaction. The energy of isolated colour states would very probably be infinite. Quarks cannot exist as free particles, but only in bound systems (quark confinement). On the other hand, the colour interaction decreases asymptotically to zero for small distances (asymptotic freedom, see e.g. [Gro89, 90, 92]).

All hadrons (baryons and mesons) are subject to the strong interaction. Typical reaction times are in the region of 10^{-23} s. However, the lifetimes of hyperons under decay into nucleons and mesons are very much longer (e.g. $\tau(\Lambda^0) = 2.63 \times 10^{-10}$ s, $\tau(\Omega^-) = 0.82 \times 10^{-10}$ s). This is because the strong interaction is forbidden by certain selection rules, so that these decays only result from the weak interaction.

1.2.4.4 Weak interaction

The weak interaction takes place between leptons, between leptons and hadrons, and between hadrons. It was first observed and studied in the β decay

$$n \rightarrow p + e^- + \overline{\nu}_e. \tag{1.29}$$

The decay (1.29) is very slow in comparison with the rates of the processes of the electromagnetic and the strong interaction. Moreover, (1.29) violates the invariance under spatial reflection (parity violation, see section 1.3).

The weak interaction is the most universal interaction after gravity. While all particles take part in gravity, the weak force operates on all fermions. It is the only interaction in which both the electrical charge of the fermions involved and their flavour quantum numbers may change. The change in the charge follows from the fact that the $W^{\pm}$ field quanta carry electrical charges.

The classical theory of the weak interaction knew only charge-altering processes such as the β decay and the μ decay (see figure 1.3). Figure 1.3 uses the rule derived from the CPT theorem (see section 1.3) that it is possible to reverse a fermion line from the past into the future provided that, at the same time, the fermion is replaced by its antiparticle ('outgoing antiparticle' ≡ 'ingoing particle').

In the modern theory of the electroweak interaction, the Glashow–Weinberg–Salam theory, which combines the weak and the electromagnetic

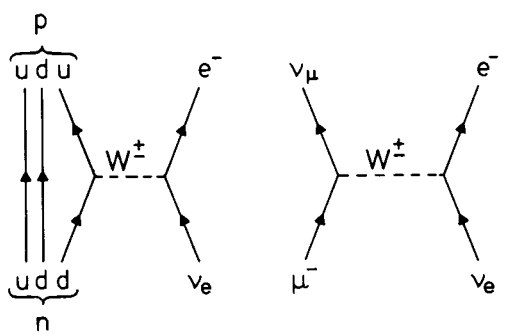


Figure 1.3. Feynman diagrams for β decay and μ decay ('charge-changing currents').

interactions, there exists a neutral Z^0 boson in addition to the charged W bosons. Consequently, there are processes in which the electrical charge of the fermions does not change. In this context we speak of *neutral currents*, which, for example, contribute to $\nu_e e^-$ scattering (see figure 1.4).

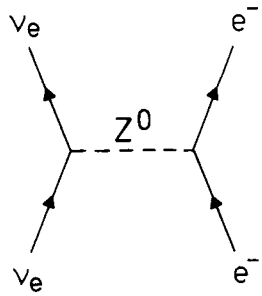


Figure 1.4. The contribution of neutral currents to $\nu_e e^-$ scattering.

The weak interaction associates two pairs of fermions, where the allowed pairings may be divided into two classes as described above. We distinguish between charged and neutral weak currents to which the following lepton pairs may be assigned:

(i) charged weak currents ($W^{\pm}$ exchange)

$$\begin{pmatrix}\overline{\nu}_e\\ e^-\end{pmatrix}\quad\begin{pmatrix}\overline{\nu}_\mu\\ \mu^-\end{pmatrix}\quad\begin{pmatrix}\overline{\nu}_\tau\\ \tau^-\end{pmatrix}\quad\begin{pmatrix}\nu_e\\ e^+\end{pmatrix}\quad\begin{pmatrix}\nu_\mu\\ \mu^+\end{pmatrix}\quad\begin{pmatrix}\nu_\tau\\ \tau^+\end{pmatrix}\tag{1.30a}$$

(ii) neutral weak currents (Z^0 exchange)

$$\begin{pmatrix}\overline{\nu}_i\\ \nu_i\end{pmatrix}\text{ for } i = e, \mu, \tau\qquad\begin{pmatrix}l^+\\ l^-\end{pmatrix}\text{ for } l = e, \mu, \tau.\tag{1.30b}$$

A similar classification holds for the quark sector. However, here, the *Cabibbo mixing* is an additional peculiarity which we shall discuss in section 1.3. The coupling of the vector bosons to the quark states u, c, t and the Cabibbo-mixed states d', s', b' is as follows:

(i) charged weak currents

$$\begin{pmatrix}u\\ \overline{d}'\end{pmatrix}\quad\begin{pmatrix}c\\ \overline{s}'\end{pmatrix}\quad\begin{pmatrix}t\\ \overline{b}'\end{pmatrix}\quad\begin{pmatrix}\overline{u}\\ d'\end{pmatrix}\quad\begin{pmatrix}\overline{c}\\ s'\end{pmatrix}\quad\begin{pmatrix}\overline{t}\\ b'\end{pmatrix}\tag{1.31a}$$

(ii) neutral weak currents

$$\begin{pmatrix}u\\ \overline{u}\end{pmatrix}\quad\begin{pmatrix}c\\ \overline{c}\end{pmatrix}\quad\begin{pmatrix}t\\ \overline{t}\end{pmatrix}\quad\begin{pmatrix}d'\\ \overline{d}'\end{pmatrix}\quad\begin{pmatrix}s'\\ \overline{s}'\end{pmatrix}\quad\begin{pmatrix}b'\\ \overline{b}'\end{pmatrix}.\tag{1.31b}$$

Only the fermion–antifermion pairs given in (1.30) and (1.31) may be coupled together.

The weak interaction has a very rich phenomenology. Most easily experimentally accessible phenomena are decay processes. In addition, the weak interaction also manifests itself in scattering processes, for example in νe^- scattering and νN scattering.

$$\nu_e e^- \to \nu_e e^- \qquad \nu_\mu n \to \mu^- p. \tag{1.32}$$

The processes of the weak interaction may be classified as follows, depending on the type of the particles involved:

(i) *Purely leptonic reactions* involving only leptons, e.g.

$$\mu^- \to e^- \overline{\nu}_e \nu_\mu \qquad \nu_e e^- \to \nu_e e^-. \tag{1.33}$$

(ii) *Semileptonic reactions* involving leptons and quarks (or hadrons), e.g.

$$\begin{aligned} n &\to p e^- \overline{\nu}_e \qquad (d \to u e^- \nu_e) \\ \pi^- &\to \mu^- \overline{\nu}_\mu \qquad (d\overline{u} \to \mu^- \overline{\nu}_\mu) \\ \Lambda &\to p e^- \overline{\nu}_e \qquad (s \to u e^- \overline{\nu}_e). \end{aligned} \tag{1.34}$$

(iii) *Purely hadronic reactions* which do not involve leptons, e.g.

$$\begin{aligned} \Lambda &\to \pi^- p \qquad (s \to \overline{u} d u) \\ K^+ &\to \pi^0 \pi^+ \qquad (\overline{s} \to \overline{u} u \overline{d}). \end{aligned} \tag{1.35}$$

Another feature of the weak interaction is the *e–μ–τ universality*. In all weak processes the muon, the tauon and the corresponding neutrinos occur in exactly the same way as the electron and its associated neutrino. Apart from the mass, e^-, μ^- and τ^- behave in exactly the same way. The different masses of the charged leptons do not affect their properties as far as the weak interaction is concerned. This is referred to as the e–μ–τ universality; its root cause is still a puzzle.

Another peculiarity of the weak force is the helicity structure. Research into the β decay of ^{60}Co first showed that leptons are emitted with a preferred direction of rotation. The direction of rotation is defined by a pseudoscalar operator given by the scalar product of the nuclear spin $\boldsymbol{I}$ and the electron momentum $\boldsymbol{p}_e$. The momentum $\boldsymbol{p}_e$ is a vector. The nuclear spin $\boldsymbol{I}$, on the other hand, like all angular momentum operators, is an axial vector; in other words, it does not change sign under spatial reflection (parity transformation $\boldsymbol{x} \xrightarrow{P} -\boldsymbol{x}$). Consequently, the sign of the quantity $\boldsymbol{p}_e \cdot \boldsymbol{I}$ (pseudoscalar) changes under spatial reflection. If parity is exactly conserved, such a pseudoscalar operator must always vanish.

Wu *et al* [Wu57] investigated the β decay of ^{60}Co and, in particular, determined the expectation value $\langle \boldsymbol{p}_e \cdot \boldsymbol{I} \rangle$ by comparing the rate of the electrons emitted in the direction of the nuclear spin with that of those emitted in the direction opposite to the nuclear spin. It turns out that the rate is dependent on the direction of emission relative to the orientation of the nuclear spin $\boldsymbol{I}$, i.e. the following holds for the measured expectation value:

$$\langle \boldsymbol{p}_e \cdot \boldsymbol{I} \rangle \neq 0. \tag{1.36}$$

The analysis showed that electrons are preferentially emitted with a counterclockwise (left-handed) twist ($\boldsymbol{p}_e \uparrow\downarrow \boldsymbol{I}$) while all antineutrinos are right-handed.

In an equally famous experiment, Goldhaber *et al* showed that neutrinos are left-handed polarized ([Gol58], see also [Vyl84]).

These experimental findings indicate that only left-handed particles and right-handed antiparticles take part in the weak interaction. This means that the weak interaction prefers one sense of rotation against the mirror image; in other words, it is associated with a (maximal) violation of parity invariance.

1.3 SYMMETRIES AND INVARIANCES

1.3.1 Symmetry operations in modern physics

The various interactions between the elementary particles may be summarized and characterized (for a survey see e.g. [Gro89, 90, 92]) using conservation principles and symmetries. The invariance of the elementary processes under symmetry transformations implies a certain structure of the physical laws. In gauge theories (section 1.4) the very existence and structure of interactions are derived from underlying symmetry groups. The symmetries contained in the physical theories are recognizable from the fact that the corresponding equations and the processes these describe are invariant under certain mathematical operations. Every conserved quantity corresponds to an invariance of the equations of motion or the Lagrange function under certain symmetry operations.

In general, we distinguish between *external* and *internal* symmetries. The external symmetries relate to the space–time continuum. The homogeneity of space, the isotropy of space and the homogeneity of time lead to the important conservation principles for momentum, angular momentum and energy (see e.g. [Lan79a, Gre79, 89]). Internal symmetries concern parameters of the particle wavefunctions, e.g. the phase of a wavefunction. The electrical charge is an example of a conserved quantity which goes back to an internal invariance.

We distinguish between *continuous* and *discrete* symmetry operations depending on whether they are parametrized by a real number or an integer, respectively. Examples of discrete symmetries include the invariances under

point reflection (parity) and charge conjugation (C), which lead to multiplicative quantum numbers. Additive quantum numbers such as the electrical charge result from continuous symmetry transformations.

In quantum mechanics, a system is described by a state vector or a wavefunction ψ. The result of a physical measurement is given by the expectation value $\langle\psi|Q|\psi\rangle$ of a corresponding self-adjoint operator Q. The evolution of a system with time may be described by a time-dependent wavefunction $\psi_s(\boldsymbol{r}, t)$ and a time-independent operator Q_0 (Schrödinger picture). The equation of motion (Schrödinger equation) is

$$\mathrm{i}\hbar\frac{\partial\psi_s(\boldsymbol{r}, t)}{\partial t} = H\psi_s(\boldsymbol{r}, t) \tag{1.37}$$

where H denotes the Hamiltonian operator.

A completely equivalent representation is based on a time-independent state function ψ_0, while the evolution with time is described by a time-dependent operator (Heisenberg picture). This Heisenberg representation has the advantage that the correspondence to the corresponding classical variables is explicit. The Heisenberg equation of motion for the operator Q is [Lan79b]

$$\mathrm{i}\hbar\frac{\mathrm{d}Q}{\mathrm{d}t} = \mathrm{i}\hbar\frac{\partial Q}{\partial t} + [Q, H]. \tag{1.38}$$

Thus, the expectation value $\langle\psi|Q|\psi\rangle$ is generally time dependent.

In quantum mechanics, conserved quantum numbers may be defined via the corresponding operators. If the operator Q does not explicitly depend on time ($\partial_t Q = 0$), then it corresponds to a conserved quantity if, and only if, it commutes with the Hamiltonian operator

$$[H, Q] = 0 \tag{1.39}$$

i.e. when H and Q are simultaneously diagonalizable. Under this condition, there exist eigenstates ψ of H, which are also eigenstates of Q

$$H\psi = E\psi \qquad Q\psi = q_0\psi. \tag{1.40}$$

Then $\langle\psi|Q|\psi\rangle$ is a constant of motion and the eigenvalue q_0 is a conserved ('good') quantum number.

Three important *discrete* symmetry operations which we shall discuss in detail are the parity transformation P, the charge conjugation C and the time reversal T.

1.3.1.1 Parity P

The parity is characterized by a multiplicative quantum number belonging to a discrete symmetry operation. The parity operation P defines the spatial reflection

of the coordinates at the origin. While the sign of a polar vector is changed ($\boldsymbol{r} \rightarrow -\boldsymbol{r}$), axial vectors, such as the orbital angular momentum $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$, are unchanged.

The parity operator P is an Hermitian operator which transforms a scalar wavefunction as follows

$$\psi^P(\boldsymbol{r}, t) = P\psi(\boldsymbol{r}, t) = \psi(-\boldsymbol{r}, t). \tag{1.41}$$

A repeated application of P leads back to the initial state, so that $P^2 = 1$. If $\psi(\boldsymbol{r}, t)$ is an eigenstate of P ($P\psi = \pi\psi$) then ψ has either even ($\pi = +1$) or odd ($\pi = -1$) parity. The invariance of a physical system under spatial reflections leads to the conservation principle for parity.

Experimental findings show that the strong and the electromagnetic interaction conserve the parity, provided the particles are assigned an *intrinsic parity* independent of the state of motion. On the other hand, the non-conservation of parity, in other words, a right–left asymmetry of nature, is a characteristic feature of the weak interaction. The left-handedness of the neutrinos is a particularly clear example of the violation of parity. P changes the handedness of a particle and so generates a right-handed neutrino, which is not observed in nature, from a left-handed neutrino.

1.3.1.2 Charge conjugation C

The charge conjugation, or particle–antiparticle conjugation, C, is a discrete internal symmetry involving a transformation of a particle into the corresponding antiparticle, where all additive quantum numbers change sign (see e.g. [Nac86]).

While the C invariance of the strong and the electromagnetic interaction has been verified to within approximately 1% accuracy, the weak interaction violates this symmetry, as, for example, the longitudinal polarization of the electrons and positrons in β decay shows. In fact, there is a preference to emit *left*-handed electrons and *right*-handed positrons. However, the charge-conjugate state of a left-handed electron state would be a *left*-handed positron state

$$|e_L^-\rangle^C = |e_L^+\rangle. \tag{1.42}$$

1.3.1.3 CP conjugation

Although the weak interaction is not invariant under C or P, it exhibits an almost perfect invariance under the composite operation CP, which was first introduced in 1957 by Landau [Lan57]. To date, a violation of the CP invariance has only been observed in the decay of the neutral K mesons [Chr64].

Since the strong interaction conserves the strangeness S, the K mesons K^0 and $\overline{K^0}$ formed in processes of the strong interaction have a well-defined

quantum number S. Since the weak interaction may alter the strangeness, the states K^0 and $\overline{K^0}$ are not eigenstates of the weak interaction. However, states with well-defined CP eigenvalues may be formed by linear combination (see e.g. [Per82])

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K^0}\rangle) \qquad CP = +1 \tag{1.43a}$$

$$|K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K^0}\rangle) \qquad CP = -1. \tag{1.43b}$$

While K^0 and $\overline{K^0}$ are characterized by their generation, K_1 and K_2 are distinguished by their decay. K_1 (K_S) decays with a lifetime $\tau_S = 0.9 \times 10^{-10}$ s into two pions ($CP = +1$), while K_2 (K_L) decays with a lifetime $\tau_L = 0.5 \times 10^{-7}$ s into three rather than two pions (this is a result of the negative CP eigenvalue). The longer lifetime is explained by the less favourable phase space. In a famous experiment in 1964, Christenson, Cronin, Fitch and Turlay [Chr64] showed that the long-lived state may also decay into two pions with a very small probability. If there were exact CP invariance, this decay would be forbidden. The following amplitude ratios define a measure of the CP violation:

$$\eta_{+-} = \frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)} \tag{1.44a}$$

$$\eta_{00} = \frac{A(K_L \to \pi^0\pi^0)}{A(K_S \to \pi^0\pi^0)} \tag{1.44b}$$

where the following values are found experimentally [PDG92]

$$|\eta_{+-}| = (2.27 \pm 0.02) \times 10^{-3} \qquad |\eta_{00}| = (2.33 \pm 0.08) \times 10^{-3}. \tag{1.45}$$

The CP violation has also been detected in the leptonic decay modes of K_L (see [Per82]):

$$K_L \to l^+ \nu_l \pi^- \tag{1.46a}$$

$$K_L \to l^- \overline{\nu}_l \pi^+. \tag{1.46b}$$

The decay widths are not identical for the two processes.

This violation of the CP invariance is connected with a violation of the time-reversal via the CPT theorem discussed in the next section.

1.3.1.4 Time reversal

The invariance under reflection in time means that the laws of nature do not change if time is allowed to run 'backwards' (see figure 1.5). The time-reversal

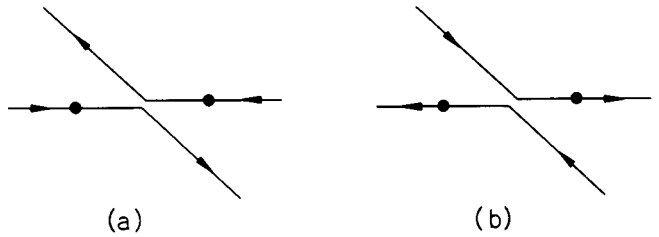


Figure 1.5. Collision of two spheres. Diagram (*b*) shows the time-inverted process for diagram (*a*).

operator T replaces the time coordinate t by $-t$, but leaves the positional coordinates unchanged

$$t \xrightarrow{T} -t \qquad r \xrightarrow{T} r. \tag{1.47}$$

Thus, the velocity $v = \mathrm{d}r/\mathrm{d}t$, the momentum and the angular momentum all change their sign.

The laws of classical physics are invariant under time reversal. For example, the Newtonian equation of motion

$$F = m\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} \tag{1.48}$$

is a second-order differential equation in t which remains unchanged under the mapping $t \to -t$. Maxwell's equations are also invariant under the operation T (for a more detailed discussion, see e.g. [Hol89]).

In quantum mechanics, Schrödinger's equation is invariant under time reversal, if T is defined for a scalar wavefunction as follows [Bet86, Hol89]:

$$T\psi(r, t) = \psi^*(r, -t). \tag{1.49}$$

However, the invariance under time-reversal does not lead to a conserved quantum number which plays a role similar to that of parity. This is because the operator T maps the wavefunction ψ into the complex-conjugate function ψ^*, so that ψ cannot be an eigenstate of T. The invariance under time-reversal thus cannot be tested by searching for decays which contradict the conservation of a 'time parity'. However, there are other ways of checking the T invariance. These include the search for an electrical dipole moment of elementary particles, in particular of the neutron. Another way of testing uses the principle of detailed balance, which is derived from the T invariance and says that the transition probability for a reaction is the same as that for the corresponding reverse reaction for states reflected in time.

1.3.1.5 CPT invariance

Table 1.5 shows the behaviour of a number of physical quantities under the operations P, C and T.

Table 1.5. The behaviour of a number of important physical quantities under C, P and T transformations.

Quantity	P	C	T
Position vector r	$-r$	r	r
Time t	t	t	$-t$
Momentum p	$-p$	p	$-p$
Spin σ	σ	σ	$-\sigma$
Electrical field E	$-E$	$-E$	E
Magnetic field B	B	$-B$	$-B$

The CPT theorem is one of the fundamental theorems of modern quantum field theories [Sch51, Lüd54, Pau55, Lüd57]. It says that all physical interactions are invariant under the composite transformation CPT, where the sequence of the individual operations is arbitrary. The assumptions on which the derivation of this theorem is based are so general that it appears extremely difficult to construct a theory without CPT conservation. In particular, any Lorentz-invariant quantum field theory, which includes local field equations and commutation relations for the fields and satisfies the spin-statistics principle (fermions (half-integer spin): Fermi–Dirac statistics; bosons (integer spin): Bose–Einstein statistics), is automatically CPT invariant (see [Lan79b]).

The CPT invariance guarantees that there exists an antiparticle for every particle which has the same half-life, mass and spin, but opposite additive quantum numbers, such as the charge. As previously mentioned, both C and P invariance are broken (see table 1.6 which gives an overview of the conservation principles). To date, a T violation has only been found indirectly using the CPT theorem via the CP violation in the decays of the neutral K mesons.

1.3.2 Models for a T and CP violation

The CP violation found in the K^0 system and the violation of the T invariance are closely related to each other by the CPT theorem. In this section, we give a brief summary of the theoretical ideas about CP and T violation (for excellent reviews see [Jar89, Tra90]).

In the standard model (see section 1.4) a violation of the time-reversal invariance by a phase of the Kobayashi–Maskawa matrix is possible [Kob73].

Table 1.6. *Overview of the conservation principles.*

Conservation principle	Strong interaction	Electromagnetic interaction	Weak interaction
Energy	Yes	Yes	Yes
Momentum	Yes	Yes	Yes
Angular momentum	Yes	Yes	Yes
B, L	Yes	Yes	Yes
P	Yes	Yes	No
C	Yes	Yes	No
C, P	Yes	Yes	No†
T	Yes	Yes	No‡
CPT	Yes	Yes	Yes

† To date only in the K^0 system.

‡ Follows indirectly from the CP violation and the CPT invariance.

The Kobayashi–Maskawa mixing matrix describes a mixing in the quark sector. From experiments on weak decays with strangeness, Cabibbo deduced that the d quark state in the weak interaction is not identical to the d quark defined by the mass eigenstate, but is a superposition of d and s states (see section 1.3.1)

$$d' = d\cos\theta_C + s\sin\theta_C \tag{1.50}$$

with the Cabibbo angle $\theta_C \approx 13°$. Thus, the $SU(2)_L$ quark doublet is $(u, d')_L$. This transformation 'saved' the universality of the weak interaction. In order to take account of the experimental finding that there are no neutral weak currents which change the quark flavour, it was necessary to introduce another interaction term containing the state s':

$$s' = s\cos\theta_C - d\sin\theta_C. \tag{1.51}$$

The missing partner in the new quark doublet was the c quark, the existence of which was required by these considerations even before it was discovered (GIM mechanism [Gla70]). The extension of the theory to three families leads to the following left-handed quark doublets:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \qquad \begin{pmatrix} c \\ s' \end{pmatrix}_L \qquad \begin{pmatrix} t \\ b' \end{pmatrix}_L. \tag{1.52}$$

The quark states $|d'\rangle$, $|s'\rangle$ and $|b'\rangle$ are mapped by a unitary transformation U

from the mass eigenstates $|d\rangle$, $|s\rangle$ and $|b\rangle$

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad \text{with} \quad U^\dagger = U^{-1} \tag{1.53}$$

where U denotes the Kobayashi–Maskawa matrix

$$U = \begin{pmatrix} U_{ud} & U_{us} & U_{ub} \\ U_{cd} & U_{cs} & U_{cb} \\ U_{td} & U_{ts} & U_{tb} \end{pmatrix}. \tag{1.54}$$

The matrix element U_{us}, for example, gives the relative strength of the transition $u \leftrightarrow s$.

The Kobayashi–Maskawa matrix may be parametrized by three Euler angles and six phase parameters. However, five of these six phases are not observable, so that we obtain only one non-trivial phase δ (see e.g. [Nac86])

$$U = \begin{pmatrix} c_1 & s_1c_3 & s_1s_3 \\ -s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ -s_1s_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix} \tag{1.55}$$

where $c_i = \cos\theta_i$, $s_i = \sin\theta_i$ $(i = 1, 2, 3)$ and $0 \leq \theta_i \leq \pi/2$, $0 \leq \delta \leq 2\pi$. The three mixing angles θ_1, θ_2 and θ_3 replace the Cabibbo angle θ_C in the case where there are only four quarks ($c_1 = \cos\theta_C$). The determination of the individual matrix elements is the subject of intensive research [Ng89]. Since $e^{i\delta}$ is mapped into $e^{-i\delta}$ under the time-reversal transformation, a phase δ different from 0 or π means that T or CP violation may occur. However, no element of the mixing matrix may vanish if CP is to be violated (conditions for the occurrence of CP violation can be found in [Jar85, Jar89]). It has only recently been shown that $U_{ub} \neq 0$ [Alb90, Ful90, Schm90a].

There are a number of models for the CP violation (see below), but the Kobayashi–Maskawa model is the only one within the framework of the currently known physics, in other words, it contains no new physics going beyond the standard model.

The assignment of the physical particles K_S (short-lived) and K_L (long-lived) to K_1 and K_2 of (1.43*a*, *b*) is only correct in the case of exact CP conservation. There are two conceivable possibilities for a CP violating decay of the long-lived K meson into two pions, namely a *direct* and an *indirect* CP violation. The indirect CP violation means that CP violating components are admixed to the wavefunctions of K_L and K_S; in other words, the physical states are linear combinations of the CP eigenstates K_1 and K_2

$$|K_L\rangle = \frac{|K_2\rangle + \epsilon|K_1\rangle}{\sqrt{1 + |\epsilon|^2}} \qquad |K_S\rangle = \frac{|K_1\rangle + \epsilon|K_2\rangle}{\sqrt{1 + |\epsilon|^2}}. \tag{1.56}$$

According to this, all *transitions* in nature would themselves be CP conserving and the violation would be solely and exclusively an effect of the small admixture ϵ of the opposite CP parity to the physical states.

Another possibility involves a *direct* CP violating transition, i.e. the interaction Hamiltonian operator itself contains a CP violating component

$$\langle \pi^0 \pi^0 | H_W | K_2 \rangle \neq 0. \tag{1.57}$$

This contribution is often described by a parameter ϵ'.

Experimentally, one determines the amplitudes of the decays K_L, $K_S \rightarrow 2\pi$. The parameters η_{+-} and η_{00} defined in (1.44) may be expressed in terms of the parameters ϵ and ϵ' (see e.g. [Nac86])

$$\eta_{+-} = \epsilon + \epsilon' \qquad \eta_{00} = \epsilon - 2\epsilon'. \tag{1.58}$$

Until 1987 all experimental data were consistent with the assumption of indirect CP violation, with $\epsilon \sim 2 \times 10^{-3}$ and $\epsilon' \sim 0$. The first indication of a direct CP violation ($\epsilon' \neq 0$) was claimed in 1988 at CERN in Geneva [Bur88]

$$\mathrm{Re}\,(\epsilon'/\epsilon) = (33 \pm 11) \times 10^{-4}. \tag{1.59}$$

However, this result was not confirmed by experiments at the Fermilab in Chicago, where the American group obtained the result [Pat90]

$$\mathrm{Re}\,(\epsilon'/\epsilon) = (-4 \pm 14 \pm 6) \times 10^{-4}. \tag{1.60}$$

In the Kobayashi–Maskawa model discussed above, we expect $|\eta_{+-}| \sim |\eta_{00}| \sim |\epsilon| \sim \sin\theta_2 \sin\theta_3 \sin\delta < 3 \times 10^{-3}$ (see e.g. [Nac86]). In particular, we require $\epsilon'/\epsilon \approx 10^{-2}$ [Gil79]. The T violating phase of the mixing matrix is associated with the heavy-quark sector so that we do not expect any T violating low-energy phenomena. On the other hand, a breaking of the CP invariance should occur in the decay of the B mesons (quark content $B^0 = |\bar{b}d\rangle$).

This Kobayashi–Maskawa picture may be modified in left–right symmetric models in which the electroweak interaction is mediated by the exchange of vector bosons which couple both left- and right-handedly. The fact that, until now, only left-handed weak charged currents have been observed is explained by the much greater mass of these right-handed vector bosons. Left- and right-handed gauge bosons may have different mixing parameters. Although some models postulate that the left- and right-handed Kobayashi–Maskawa matrices are identical,

$$U_L = U_R \tag{1.61}$$

in general U_L and U_R are completely independent of each other. This additional freedom due to the right-handed currents means that T violating phases are also possible in the light-quark sector (see [Hol89]).

Soon after the discovery of the CP violation, Wolfenstein proposed a phenomenological model to explain the observation (superweak model [Wol64]). This model postulates the existence of a superweak force, which mediates transitions between the CP eigenstates $|K_1\rangle$ and $|K_2\rangle$ but should appear practically nowhere else

$$\langle K_1|H_{sw}|K_2\rangle \neq 0. \tag{1.62}$$

This new interaction could change the strangeness by two units ($\Delta S = 2$). The superweak model requires

$$\eta_{+-} = \eta_{00} = \epsilon \text{ and } \frac{\epsilon'}{\epsilon} = 0. \tag{1.63}$$

This approach can be excluded if it is possible to detect a direct CP violation ($\epsilon'/\epsilon \neq 0$) or if a CP or T violation is found in another system.

A phase δ different from 0 and π in the Kobayashi–Maskawa matrix is one possible cause of CP violation in the standard model. Non-perturbative effects in quantum chromodynamics (QCD) could be another source of CP violation [t'Ho76]. The QCD vacuum cannot be uniquely defined, instead there are a number of equivalent degenerate vacuum states. For the actual vacuum, we anticipate a superposition of these degenerate states with random phases $e^{i\nu\theta}$, where ν is an integral topological quantum number [Cal76, Jac76]. Such a vacuum would not be CP invariant. As a measurable effect an electrical dipole moment d_n of the neutron is induced, which should be proportional to θ

$$d_n = (2.7 - 5.2) \times 10^{-16}\theta \; e\,\text{cm}. \tag{1.64}$$

The experimental bounds for d_n lead to very small values of

$$\theta < 3 \times 10^{-10} \tag{1.65}$$

which is difficult to understand if θ is a purely random phase (θ *problem* or *strong* CP *problem*, see section 5.2.5).

The CP violating fraction of the Lagrange density $\mathcal{L}_\theta$ is relatively unimportant for the K^0 system. If one imagines that the usual weak interaction is initially switched off, then the CP eigenstates $|K_1\rangle$ and $|K_2\rangle$ are also eigenstates of C with different signs. Since $\mathcal{L}_\theta$ is even under charge conjugation, there can be no transitions between K_1 and K_2 and thus no state mixing. This could only occur as a result of a joint effect of $\mathcal{L}_\theta$ and the weak interaction. However, it cannot explain the strength of the CP violation (see e.g. [Nac86]). In other words, the search for an electrical dipole moment of the neutron leads to far-reaching conclusions about theoretical concepts, which cannot be derived by studying the K^0 system alone.

1.4 GAUGE THEORIES AND THE STANDARD MODEL

1.4.1 Introduction

At first sight, the weak interaction, electromagnetism and the strong interaction appear to have nothing in common. They differ both in their phenomenology and in their range.

However, through modern gauge theories an order has been successfully imposed in the area of the natural forces. One important realization is that the three fundamental interactions can be derived from the postulate of gauge symmetries. They are all mediated by the exchange of gauge bosons. The reason for the differences in these interactions lies in the symmetry groups which represent the interactions and determine the number of gauge bosons and the form of the force.

In the standard model the theory of the strong interaction, QCD, is characterized by the gauge group $SU(3)_c$. The group $SU(2)_L \otimes U(1)$ determines the structure of the electromagnetic and weak forces which combine to form the electroweak force.

However, in the standard model, which is based on the direct product of the groups

$$SU(3)_c \otimes SU(2)_L \otimes U(1) \tag{1.66}$$

a true unification of all three interactions has not yet been achieved. The $SU(3)_c$ and the $SU(2)_L \otimes U(1)$ transformations are independent of each other and we still have three different coupling constants. A true unification of the three interactions can only be achieved if the three different interaction strengths can be derived from a *single* fundamental coupling constant. This can be realized with a *simple* group. The smallest group satisfying these conditions is $G = SU(5)$ (see section 1.5).

In the standard model the strengths of the interactions are described by dimensionless coupling constants:

$$\begin{aligned} g_3 &= g_s & & SU(3) \\ g_2 &= e/\sin\theta_W & & SU(2) \\ g_1 &= e/\cos\theta_W & & U(1). \end{aligned} \tag{1.67}$$

The parameter θ_W denotes the Weinberg angle which is not predicted by the standard model and must be determined experimentally; e is the elementary electrical charge.

1.4.2 The gauge principle

In preparation for the more detailed discussion, we now give a brief sketch of some basics of gauge theories. The central importance of gauge theories

in modern physics stems from the fact that they are *renormalizable* [t'Ho72, Lee72].

Renormalization is the procedure of eliminating divergences (infinite singularities) which occur, in particular, in calculations of higher-order corrections (so-called radiative corrections). For the formal aspects of renormalizability, we refer readers to the specialized literature (see e.g. [Ait89]).

The gauge principle is based on the fact that both classical physics and quantum theory involve quantities which, in principle, cannot be measured. Equivalent theories, which give the same predictions for all experiments, result for different values of these quantities. Thus, it is possible to 'gauge' a theory by a suitable choice of the non-measurable parameters, in order, for example, to simplify the equations of motion.

In gauge theories these gauge freedoms are not simply viewed as random events but are raised to a general principle. The requirement that there should exist such physically unfixed (gaugeable) quantities is used to deduce the existence and the structure of interactions with the corresponding interaction fields (for introductory descriptions, see [Ait89, Gro89, 90, 92]).

Electromagnetism as an example of a gauge theory. We shall explain this gauge freedom using the example of electromagnetism. The four-potential

$$A^{\mu} = (\phi, \boldsymbol{A}) \tag{1.68}$$

is an example of a quantity which, in principle, cannot be measured. Different values of A^{μ} lead to the same physical fields $\boldsymbol{E}$ and $\boldsymbol{B}$, as we shall see below.

We begin the discussion by restating Maxwell's equations:

$$\operatorname{div} \boldsymbol{E} = \varrho \tag{1.69a}$$

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{1.69b}$$

$$\operatorname{div} \boldsymbol{B} = 0 \tag{1.69c}$$

$$\operatorname{curl} \boldsymbol{B} = \boldsymbol{j} + \frac{\partial \boldsymbol{E}}{\partial t} \tag{1.69d}$$

where ϱ and $\boldsymbol{j}$ satisfy the continuity condition

$$\frac{\partial \varrho}{\partial t} + \operatorname{div} \boldsymbol{j} = 0. \tag{1.70}$$

Instead of the fields $\boldsymbol{E}$ and $\boldsymbol{B}$, it is often advantageous to introduce the vector potential A^{μ}, where

$$\boldsymbol{B} = \operatorname{curl} \boldsymbol{A} \tag{1.71a}$$

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial}{\partial t}\boldsymbol{A}. \tag{1.71b}$$

This choice means that equations (1.69*b*) and (1.69*c*) are automatically satisfied.

Using the definition of a four-current j^μ

$$j^\mu = (\varrho, \boldsymbol{j}) \tag{1.72}$$

the continuity equation (1.70) may be written in the simple form

$$\partial_\mu j^\mu = 0. \tag{1.73}$$

By introducing the field-strength tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{1.74}$$

we may write the Maxwell equations (1.69*a*) and (1.69*d*) in the compact form

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{1.75}$$

The gauge invariance of classical electrodynamics has its origins in the fact that the potential A^μ may take different values for the same predefined physical fields $\boldsymbol{E}$ and $\boldsymbol{B}$.

The specification of a definite value for A^μ is called *choosing a gauge*. Transformations which leave $\boldsymbol{E}$ and $\boldsymbol{B}$ unchanged are called gauge transformations. If we carry out the replacement

$$A^\mu(x) \to A^\mu(x) + \partial^\mu \Lambda(x) \tag{1.76}$$

or

$$\boldsymbol{A}(\boldsymbol{x}, t) \to \boldsymbol{A}(\boldsymbol{x}, t) - \operatorname{grad} \Lambda(\boldsymbol{x}, t) \tag{1.77a}$$

$$\phi(\boldsymbol{x}, t) \to \phi(\boldsymbol{x}, t) + \frac{\partial}{\partial t}\Lambda(\boldsymbol{x}, t) \tag{1.77b}$$

where $\Lambda(\boldsymbol{x}, t)$ is an arbitrary, differentiable, scalar function, the field-strength tensor $F^{\mu\nu}$ and the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are unchanged. Thus, the transformation (1.76) (or (1.77)) is called a gauge transformation.

The gauge freedom may be used to subject the four-potential to a secondary condition. Usually, one applies the so-called *Lorentz gauging*

$$\partial_\mu A^\mu = 0 \tag{1.78}$$

which enables us to write the Maxwell equation (1.75) in the simple form

$$\Box A^\mu = j^\mu. \tag{1.79}$$

Here, the electromagnetic current density j^μ represents the source of the field A^μ.

As previously mentioned, the gauge freedom is raised to a general principle in gauge theories. The internal structure of gauge transformations, and thus the dynamics of the resulting forces, are determined by an underlying symmetry group. As we have seen, electromagnetism can be described in the framework of a gauge theory. Since the detection of the $W^{\pm}$ and Z^0 bosons at CERN, there is no longer any doubt that the electroweak interaction may be derived from the gauge principle. The same holds for the colour interaction in QCD.

The root of the gauge principle is the conservation of the internal symmetry (i.e. the invariance of form) of equations of motion under transformations. In general, we distinguish between *global* and *local* symmetries.

Global transformations change a physical quantity in the same way everywhere; they are independent of the space and time coordinates. As an example, we mention the Schrödinger formulation of quantum mechanics in which the physical states are described by complex wavefunctions $\psi(\boldsymbol{x}, t)$. In this case, a global transformation corresponds to multiplication by a constant phase factor

$$\psi'(\boldsymbol{x}, t) = T\psi(\boldsymbol{x}, t) \equiv \mathrm{e}^{-\mathrm{i}e\varrho}\psi(\boldsymbol{x}, t) \tag{1.80}$$

where ϱ is real and independent of $(\boldsymbol{x}, t)$. The transformed wavefunction ψ' satisfies the same Schrödinger equation as ψ, since differentiation with respect to space and time does not affect the constant phase factor. Thus, the Hamiltonian operator commutes with the transformation operator

$$[H, T] = 0. \tag{1.81}$$

The fundamental equation of quantum electrodynamics, the Dirac equation, is also form-invariant under the global transformation (1.80).

The Dirac equation is the relativistic equation of motion of a free electron. It is given by

$$(\mathrm{i}\gamma^{\mu}\partial_{\mu} - m)\psi(x) = 0 \tag{1.82}$$

where γ^{μ} denote the following 4×4 matrices

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \gamma = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \tag{1.83}$$

and $\boldsymbol{\sigma}$ denotes the Pauli spin matrices. We would refer readers who are unfamiliar with relativistic quantum mechanics to e.g. [Ait89].

Substitution of (1.80) in (1.82) immediately gives

$$(\mathrm{i}\gamma^{\mu}\partial_{\mu} - m)\psi'(x') = 0. \tag{1.84}$$

Noether's theorem ([Noe18], see also [Bjo78, Gro89, 90, 92]) says that for *every* global transformation under which the Lagrange density $\mathcal{L}$ is invariant

there exists a conserved quantity, i.e. a measurable quantity, the value of which does not vary with time.

Examples from classical mechanics include the connection between translation, rotation and time invariance, and the conservation of momentum, angular momentum and energy, respectively. The invariance of the Dirac equation under (1.80) results in the conservation of the electrical charge (see e.g. [Gro89]).

Local (= gauge) symmetries are far more important than the global symmetries which occur in most physical theories. These play a crucial role in the description of the interactions.

In this context, local means that the symmetry transformation may be chosen in a different way at each point in space and time, i.e. (1.80) becomes

$$\psi'(x) = \mathrm{e}^{\mathrm{i}e\varrho(x)}\psi(x) \qquad x = (\boldsymbol{x}, t) \text{ local transformation.} \tag{1.85}$$

The requirement for invariance under a *local* transformation is a much stronger condition than that for invariance under global transformations.

The Dirac equation for free particles is not invariant under local transformations of the form (1.85), i.e. $\psi'(x)$ is not a solution of (1.82). Suppose that $\psi(x)$ is a solution of the Dirac equation, then $\psi'(x)$ satisfies the equation

$$(\mathrm{i}\gamma^{\mu}(\partial_{\mu} - \mathrm{i}e\partial_{\mu}\varrho(x)) - m)\psi'(x) = 0. \tag{1.86}$$

The differentiation of the phase factor gives one term more than in (1.82).

The Dirac equation may be made invariant by introducing an interaction field. This *gauge* field must be transformed simultaneously with the equation (1.85) in such a way that the effect of the additional term in (1.86) is eliminated. The combined transformation of the field ψ and the gauge field is called a *gauge transformation.*

In the case of quantum electrodynamics, the invariance of the Dirac equation is achieved by introducing a gauge field $A^{\mu}(x)$ to which $\psi(x)$ couples with strength determined by the charge e. For this, by analogy with the minimal coupling in classical physics, the derivative ∂_{μ} is replaced by the covariant derivative with respect to the field $A^{\mu}(x)$:

$$\partial_{\mu} \rightarrow D_{\mu} = \partial_{\mu} - \mathrm{i}eA_{\mu}. \tag{1.87}$$

Replacing ∂_{μ} by D_{μ} in the Dirac equation, we obtain

$$(\mathrm{i}\gamma^{\mu}D_{\mu} - m)\psi(x) = 0 \tag{1.88a}$$

$$(\mathrm{i}\gamma^{\mu}\partial_{\mu} - m + e\gamma^{\mu}A_{\mu})\psi(x) = 0. \tag{1.88b}$$

Based on (1.88*b*), for the transformed field $\psi' = e^{ie\varrho(x)}\psi$ we obtain

$$(i\gamma^\mu\partial_\mu - m + e\gamma^\mu A_\mu)e^{-ie\varrho(x)}\psi'(x) = 0 \tag{1.89a}$$

from which it follows that

$$(i\gamma^\mu(\partial_\mu - ieA_\mu - ie\partial_\mu\varrho(x) - m))\psi'(x) = 0. \tag{1.89b}$$

The Dirac equation retains its original form if the gauge field is transformed in an appropriate way

$$A'_\mu(x) = A_\mu(x) + \partial_\mu\varrho(x). \tag{1.90}$$

Substituting (1.90) in (1.89*b*) we obtain

$$(i\gamma_\mu(\partial^\mu - ieA^{\mu'}(x)) - m)\psi'(x) = 0. \tag{1.91}$$

Thus, we have shown that (1.88) is invariant under the gauge transformation

$$\psi'(x) = e^{ie\varrho(x)}\psi(x) \tag{1.92a}$$

$$A'_\mu(x) = A_\mu(x) + \partial_\mu\varrho(x). \tag{1.92b}$$

Physically, the gauge field A_μ is interpreted as a *photon field* in the case of QED.

In general, a local invariance can only be obtained by introducing a gauge field, i.e. an interaction. The properties of the interaction are largely specified by the gauge principle.

A quantization of the gauge field leads to gauge bosons as mediators of the interaction (in the case of QED these are the photons). The gauge fields also satisfy certain equations of motion. For example, the gauge field A_μ is subject to the Klein–Gordon equation, since photons are bosons. In addition, these equations of motion are required to be invariant under (1.92). However, this is only guaranteed if the exchange particles have a vanishing rest mass. This requirement is clearly satisfied for photons. It is a general property of (*unbroken*, see below) gauge theories that they, *in principle*, predict the existence of *massless* exchange fields.

On the other hand, we have already seen that the weakness of the weak interaction is due to the short range which is a consequence of the large mass of the exchange bosons.

The *Higgs mechanism* (see below) provides a possible solution to this apparent contradiction. According to this mechanism the exchange particles are assigned a mass at a later stage by the process of spontaneous symmetry breaking, without perturbing the local (= gauge) symmetry of the theory, or, more precisely, without disturbing the renormalizability.

One major advantage of gauge theories is the fact that they are renormalizable. Nevertheless, their importance is ultimately based solely on

the fact that they have until now provided a successful description of the electromagnetic and the electroweak interaction. The requirement for gauge invariance cannot be derived from fundamental physical principles. It is conceivable that the gauge principle may one day have to be replaced by a more fundamental principle. Kaluza–Klein theories, for example, attempt to trace all interactions back to a single geometrical principle (see chapter 12).

Gauge theories are normally classified according to the properties of the transformation. The structure of a theory may simply be described by the specification of a gauge group $G(n)$ which reflects the symmetry properties. We distinguish between Abelian and non-Abelian gauge groups, depending on whether or not the group satisfies the commutative law (see [Gro89, 90]).

QED is described by the group of the phase factors $e^{ie\varrho(x)}$. This is an Abelian U(1) symmetry. U(1) denotes the group of all unitary 1×1 matrices.

In our dealings with the weak and the strong forces, we shall come across the groups SU(n). S stands for special, i.e. for matrices with a positive determinant ($\det U = +1$) and n denotes the dimension of the matrices. Thus, SU(n) is the group of all unitary $n \times n$ matrices U with $\det U = +1$. In the group descriptor SO(n), O stands for orthogonal, instead of unitary. The number of gauge bosons in the special groups SU(n) and SO(n) is $n^2 - 1$.

The particles involved in an interaction are arranged in *multiplets*, where the members of a multiplet may be transformed into one another. The composition of a multiplet is invariant under the interaction in question. Since the electromagnetic force does not alter the particle type, all particles are singlets under this interaction.

1.4.3 Spontaneous symmetry breaking

The gauge bosons which occur in gauge theories are in principle all massless. The case of photons as the exchange particles of the electromagnetic interaction and that of gluons as mediators of the colour interaction are compatible with this physical picture. However, we know that the gauge bosons of the weak interaction have a rest mass of around 80 to 90 GeV/c^2. If this mass were to be explicitly included in the Lagrange density $\mathcal{L}_W$, the invariance under gauge transformations would be lost.

Spontaneous symmetry breaking provides us with a mechanism for assigning the gauge bosons a mass without destroying the gauge invariance.

We speak of a spontaneous symmetry breaking when the fundamental equations of a system possess a symmetry which the ground state does not have. For example, it is conceivable that the Lagrange density might be invariant under a gauge transformation, although the vacuum, as the state of least energy, does not possess this symmetry.

Suppose that the vacuum $|0\rangle$ is the eigenstate of the Hamiltonian operator H with the least energy $E_{\min}$:

$$H|0\rangle = E_{\min}|0\rangle. \tag{1.93}$$

If the equations of motion are invariant under the transformation U, then we have

$$[H, U] = 0. \tag{1.94}$$

Thus, for the transformed vacuum $U|0\rangle$ we obtain

$$HU|0\rangle = UH|0\rangle = E_{\min}U|0\rangle \tag{1.95}$$

i.e. $U|0\rangle$ is also an eigenstate of H with eigenvalue $E_{\min}$. If we assume that there is a unique vacuum state, it follows from (1.93) and (1.95) that

$$U|0\rangle = |0\rangle \tag{1.96}$$

i.e. under the above assumption, the vacuum is also invariant under U. However, if there are a number of degenerate states $|0\rangle_i$, we then have

$$U|0\rangle_i = |0\rangle_j \qquad i \neq j. \tag{1.97}$$

Thus, the vacuum is no longer necessarily invariant under U even when, as before, H and U commute. Thus, the phenomenon of spontaneous symmetry breaking is closely connected with a degeneracy of the ground state.

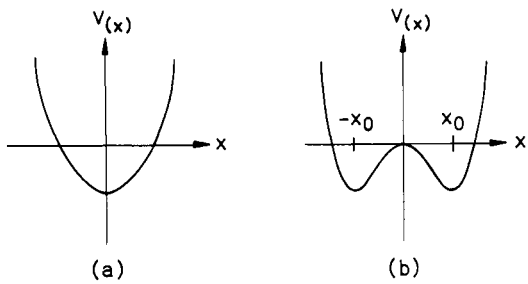


Figure 1.6. Symmetric potentials. The W-shaped potential in (*b*) illustrates the phenomenon of spontaneous symmetry breaking. A particle is in the ground state at either $x = x_0$ or $x = -x_0$. Neither of these equilibrium positions shows the symmetry of the potential under the transformation $x \to -x$.

We illustrate the most important aspects of spontaneous symmetry breaking by means of a simple example (see e.g. [Nac86, Ait89]). We consider a simple W-shaped potential of the form (see figure 1.6)

$$V(x) = -\tfrac{1}{2}\mu^2 x^2 + \tfrac{1}{4}\lambda x^4 \qquad \mu^2, \lambda > 0. \tag{1.98}$$

The potential $V(x)$ is symmetric under the transformation $x \to -x$

$$V(-x) = V(x). \tag{1.99}$$

The equilibrium position is calculated from the condition $\partial V/\partial x = 0$ as

$$x = \pm x_0 = \pm\sqrt{\frac{\mu^2}{\lambda}}. \tag{1.100}$$

In the ground state a particle in this potential is found either at $x = x_0$ or at $x = -x_0$. Neither of these equilibrium positions exhibits the symmetry of the potential under the transformation $x \to -x$. The symmetry is spontaneously broken.

The principle of spontaneous symmetry breaking is a common phenomenon in physics. Important examples include:

(i) the spontaneous magnetization of a solid body below the Curie temperature (ferromagnetism);
(ii) the formation of crystals from liquids below a critical temperature and the condensation of water vapour;
(iii) superconductivity.

In ferromagnetic materials, electron spins above the Curie temperature T_C have a purely statistical distribution. When the temperature decreases below the critical temperature T_C, the spins align themselves uniformly within the Weiss domains, leading to a spontaneous magnetization.

In statistical mechanics, the free energy of a ferromagnet may be written, using the Landau–Ginzburg approximation, as a function of the temperature T and the magnetization M in the form

$$F(M, T) \simeq F_0(T) + \tfrac{1}{2}\mu^2(T)M^2 + \tfrac{1}{4}\lambda(T)M^4 + \ldots. \tag{1.101}$$

This expansion is valid for small magnetizations. For stability reasons, we require $\lambda(T) > 0$. The free energy F is rotation invariant and all directions of magnetization M are equally likely. The minimum of (1.101) is given by

$$\text{grad}_M F(M, T) = 0 \Rightarrow [\mu^2(T) + \lambda(T)M^2]M = 0. \tag{1.102}$$

We must distinguish between two cases:

(α) $\mu^2 > 0$. There is a unique solution $M = 0$. There is no magnetization, F is rotation invariant.
(β) $\mu^2 < 0$. There are two other solutions in addition to $M = 0$

$$|M_0|^2 = -\mu^2/\lambda. \tag{1.103}$$

The induced magnetization destroys the rotation invariance of F.

In the case of spontaneous magnetization, the expectation value of M in the ground state is non-zero. However, only the absolute value of the magnetization is fixed, while the direction is arbitrary.

The critical temperature (Curie temperature), at which the symmetry breaking occurs, is determined by the condition

$$\mu^2(T_C) = 0. \tag{1.104}$$

Similar considerations apply to superconductors.

In the case of crystallization, spontaneous symmetry breaking also occurs when the temperature decreases below a certain value. While the liquid has complete rotational symmetry, crystals have preferred orientations, and thus do not respect all the invariances of the original equations.

All the examples have in common the fact that above a critical temperature (Curie temperature, transition temperature, ...) the complete symmetry is reproduced, while the preferred directions induced by the spontaneous symmetry breaking are randomly distributed.

In principle, both global and local symmetries may be spontaneously broken. However, a spontaneously broken global symmetry is inseparably associated with the existence of a massless scalar field, that of the Nambu–Goldstone boson [Nam60, Gol61].

In the following, we shall only consider the spontaneous breaking of local (= gauge) symmetries, which play the crucial role in the description of the electroweak interaction and in grand unified theories. Only theories which are invariant under local (gauge) transformations can be Lorentz invariant (see e.g. [Nac86]).

In the current quantum-field theoretical description of interactions, the spontaneous symmetry breaking is generated by the Higgs mechanism [Hig64, Kib67, Gun90]. This mechanism requires the introduction of additional scalar fields. These Higgs fields have a non-zero vacuum expectation value, i.e. the state of least energy is attained at a non-zero expectation value of the Higgs field. Thus, in particular, they affect the structure of the vacuum.

As a result of the interaction of the as yet massless fermion and boson fields with the Higgs field, the equations of motion are modified in such a way that fermions and bosons behave as though they had a mass. The coupling of the particle fields to the Higgs field is proportional to the square of the mass. As a result of the finite vacuum expectation value of the Higgs field, fermions and bosons obtain a finite potential energy which is expressed as an effective mass. This mechanism may be compared, in a certain sense, with the effective mass of an electron in a solid body, which differs, at times considerably, from the physical mass of the free electron. In this concept the true mass in the equations of motion is replaced by the effective mass, which takes the complex interaction of the electron with the lattice into account in an approximate manner.

In gauge theories, the fermions and bosons acquire a mass as a result of the interaction with the Higgs field. A manifest physical mass would destroy the gauge invariance of the theory.

1.4.4 The Glashow–Weinberg–Salam model

After our description of the fundamental ideas behind the gauge principle and the Higgs mechanism we shall now give a brief outline of the Glashow–Weinberg–Salam model (GWS model) of the electroweak interaction [Gla61, Wei67, Sal68].

The electromagnetic interaction can be successfully described by a U(1) gauge theory (QED). The initial attempt to handle the weak interaction in an SU(2) gauge theory failed. It was apparent that the two forces could not be described by separate gauge theories.

Finally, Weinberg, Glashow and Salam succeeded in deriving both forces as different components of a single gauge theory. The gauge group of this electroweak interaction is the direct product of the groups U(1) and $SU(2)_L$

$$\mathcal{E}_{\mathrm{GWS}} = \mathrm{SU}(2)_L \otimes \mathrm{U}(1). \tag{1.105}$$

The index L indicates that the interaction described by the group $SU(2)_L$ is purely left-handed. This takes into account the experimental finding that only left-handed particles and right-handed antiparticles take part in the weak interaction. However, we note that neither the $SU(2)_L$ nor the U(1) transformation in (1.105) may be uniquely identified with the weak or the electromagnetic interaction, respectively.

Both symmetries, the $SU(2)_L$ and the U(1), must be spontaneously broken. (In nature only the electromagnetic interaction and the colour interaction occur as consequences of an unbroken symmetry, i.e. with massless exchange bosons.) The 'trick' of the GWS theory is that $SU(2)_L$ and U(1) are spontaneously broken in such a way that a subgroup of the product of the two groups, namely the $U(1)_{\mathrm{EM}}$, results as an *unbroken* symmetry.

According to experimental findings to date, the charged weak currents are purely left-handed. Thus, a suitable theoretical description is obtained by decomposing the electron and neutrino fields into left- and right-handed components. The left-handed components are brought together in a doublet

$$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}. \tag{1.106}$$

The right-handed neutrino component ν_R does not enter into the theory. This corresponds to the fact that a right-handed neutrino has not been observed in experiments to date. The right-handed component of the electron forms a singlet

$$e_R. \tag{1.107}$$

The gauge transformations of the doublet and the singlet are given by:

(i) $SU(2)_L$

$$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}' = \mathcal{U}_{SU(2)} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{1.108a}$$

$$e'_R = e_R \tag{1.108b}$$

(ii) U(1)

$$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}' = \mathcal{U}_{L(1)} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{1.109a}$$

$$e'_R = \mathcal{U}_{R(1)} e_R. \tag{1.109b}$$

For the mathematical formulation of the theory we refer readers, for example, to [Ait89, Gro89, 90, 92]. At this point, we mention just two important aspects:

(i) The invariance of the Lagrange density under local $SU(2)_L \otimes U(1)$ transformations is obtained by forming an appropriate covariant derivative

$$\partial_\mu \rightarrow D_\mu \tag{1.110}$$

with appropriate gauge fields $B_\mu(x)$ and $\boldsymbol{W}_\mu(x)$ of the U(1) and the $SU(2)_L$ transformations which are coupled to the fermions with different coupling constants

$$U(1): \quad g' \quad (B_\mu(x)) \tag{1.111a}$$

$$SU(2)_L: \quad g \quad (W^+_\mu(x), W^-_\mu, W^3_\mu(x)). \tag{1.111b}$$

Since, according to (1.108) and (1.109), left- and right-handed field components are transformed in different ways under gauge transformations, explicit mass terms are forbidden in the Lagrange density.

(ii) The masses of the bosons are generated later via the Higgs mechanism. The coupling constant G_F of the Fermi theory of the weak interaction and g are related as follows

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}. \tag{1.112}$$

The Z^0 boson is understood to be the mixture of a neutral $SU(2)_L$ boson state (W^3_μ) with purely left-handed ($V - A$) coupling and a U(1) state (B_μ)

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W. \tag{1.113}$$

Correspondingly, the following holds for the photon field

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W. \tag{1.114}$$

The Weinberg angle θ_W describes this mixing by which the weak neutral currents, in contrast to the weak charged currents, obtain a right-handed admixture.

The Weinberg angle and the coupling constants g and g' are related as follows:

$$\cos\theta_W = = \frac{g}{\sqrt{g'^2 + g^2}} \tag{1.115}$$

$$\sin\theta_W = \frac{g'}{\sqrt{g'^2 + g^2}}. \tag{1.116}$$

Since the coupling between the photon field A_μ and an electron is given by the elementary electrical charge

$$e = \frac{gg'}{\sqrt{g'^2 + g^2}} = g\ \sin\theta_W \tag{1.117}$$

follows immediately. Thus, from (1.112) we obtain

$$\begin{aligned} m_W &= \frac{g}{2}\left(\frac{\sqrt{2}}{2G_F}\right)^{1/2} \\ &= \frac{e}{2\sin\theta_W}\left(\frac{\sqrt{2}}{2G_F}\right)^{1/2} \end{aligned} \tag{1.118}$$

i.e. for a known Weinberg angle m_W can be calculated (predicted) from G_F.

The mass of the neutral Z^0 is calculated as

$$m_{Z^0} = \frac{m_W}{\cos\theta_W}. \tag{1.119}$$

Just like the masses of the gauge bosons, the masses of the fermions result from a gauge-invariant interaction of the fermions with the same Higgs field. The corresponding coupling constants are free parameters of the GWS model.

In the simplest case of an $SU(2)_L$ Higgs field with four degrees of freedom, we expect the existence of a new physical particle, the *Higgs particle*. The three other degrees of freedom of the Higgs field do not lead to additional particles, but show up in the longitudinal polarization of the $W^\pm$ and Z^0 bosons. Here, we note that massless bosons with S (spin) = 1, such as the photon, have only two degrees of freedom for the transverse polarization, while massive particles have a further degree of freedom for the longitudinal polarization.

In addition to this simple Higgs structure, more complicated cases are also conceivable.

1.4.5 The strong interaction

According to our present understanding, the strong interaction may also be described by a gauge theory, QCD, the basics of which we shall now describe briefly.

The experimental facts may be explained by the existence of a gauge interaction operating between the colour charges of the quarks. The colour interaction is based on the gauge group $SU(3)_c$, where the index c stands for the colour.

Three colour charges (red, green, blue) are introduced, corresponding to the $SU(3)_c$ symmetry. They prevent among other things a contradiction which otherwise would occur between quark wavefunctions and Pauli principle. (For example a state Δ^{++} can now be described by three u quarks in the same spin state.) These additional three degrees of freedom have been confirmed in various experiments (see e.g. [Per82, Gro89, 90, 92]). The three colour charges may be combined into an uncharged (colourless or 'white') singlet state. A singlet state is obtained also from the combination of a charge with its anticharge. Until now, only 'colourless' singlet states have been freely observed in nature (see chapter 10).

According to these model ideas, the strong interaction between hadrons should only be viewed as a residual interaction of the fundamental colour force between the quarks.

To the best of our knowledge the $SU(3)_c$ symmetry, on which QCD is based, appears to occur unbroken in nature. This has the consequence that the gauge bosons, the gluons, do not have a mass. The other known interactions do not couple to the colour charge, so that the $SU(3)_c$ transformations commute with the $SU(2)_L \otimes U(1)$ transformations. Thus, the members of a $SU(3)_c$ multiplet differ only in their colour charge, and not in any of their other quantum numbers. Otherwise, their properties would make them distinguishable by the electroweak interaction and gravity and the $SU(3)_c$ symmetry would be broken.

The fundamental $SU(3)_c$ triplets are

$$\begin{pmatrix} u_r \\ u_g \\ u_b \end{pmatrix} \begin{pmatrix} d_r \\ d_g \\ d_b \end{pmatrix} \begin{pmatrix} c_r \\ c_g \\ c_b \end{pmatrix} \begin{pmatrix} s_r \\ s_g \\ s_b \end{pmatrix} \begin{pmatrix} t_r \\ t_g \\ t_b \end{pmatrix} \begin{pmatrix} b_r \\ b_g \\ b_b \end{pmatrix}. \qquad (1.120)$$

The indices r, g and b denote the three colour charges. Unlike the weak interaction, the colour interaction does not distinguish between left- and right-handed field components.

The requirement for gauge invariance leads to eight gauge fields ($n^2 - 1$, where $n = 3$), the gluon fields G_{ij}. The gluons themselves carry a combination of a charge and an anticharge. For example, G_{rb} has a red charge and a blue anticharge; thus, it transforms a blue quark into a red quark.

The fact that the gluons themselves have colour charges and therefore interact among themselves is the cause of two important properties of QCD: confinement and asymptotic freedom.

In this connection, confinement means that quarks can only exist in bound systems in a colour singlet state. Thus, we do not expect quarks to exist as free particles (see chapter 10). When an attempt is made to extract an individual quark from a neutrally coloured combination, so much energy must be applied that ultimately a new quark–antiquark pair forms.

The sharp increase in the potential energy when a critical distance between the quarks is exceeded is a result of the gluon–gluon interaction. As a result of the colour charge of the gluons each of these gauge bosons acts as the source of new gluons, so that the space between two quarks is filled by a gluon field, with its strength essentially independent of the distance between the quarks. Whence the potential energy increases approximately proportionately to the distance between quarks r.

In contrast, the potential energy for a Coulomb-like law of force tends towards a saturation value. If one wishes to describe the strong interaction by a Coulomb-type law, one has to define an effective coupling constant g_{eff} which is not constant but increases as the distance r increases.

For small distances r between quarks this effective coupling constant tends to zero. The quarks behave essentially like free particles; we therefore speak of *asymptotic freedom*. We shall return to the concept of the effective coupling constants in section 1.5.1 (see also [Gro89, 90, 92]).

1.4.6 The $SU(3)_c \otimes SU(2)_L \otimes U(1)$ group – the standard model

We have learnt that the weak interaction and the electromagnetic interaction merge into the electroweak interaction with the $SU(2)_L \otimes U(1)$ gauge group. At high energy, this symmetry occurs in unbroken form. The $W^{\pm}$ and Z^0 bosons only acquire a mass when the energy decreases below a certain threshold, of the order of 100 GeV, while the photons remain massless. Thus, the weak and the electromagnetic interaction are only distinguishable at lower energies. The colour interaction is described by QCD which is realized as an unbroken $SU(3)_c$ gauge symmetry.

In the standard model, the descriptions of the electroweak interaction and the colour interaction are combined. Since the colour interaction is limited to the quark sector the generators of the $SU(3)_c$ group commute with those of the $SU(2)_L \otimes U(1)$ group. The direct product

$$S = SU(3)_c \otimes SU(2)_L \otimes U(1) \tag{1.121}$$

forms the basis for the standard model.

For very large energies $E \gg 100$ GeV, the full gauge symmetry is realized, up to small corrections. For energies $E \ll 100$ GeV, only the electromagnetic interaction and the colour force remain as unbroken symmetries after the spontaneous symmetry breaking

$$\mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1) \xrightarrow{m_{(Z^0,W)}c^2 \simeq 100\ \mathrm{GeV}} \mathrm{SU}(3)_c \otimes \mathrm{U}_{\mathrm{EM}}(1). \tag{1.122}$$

In the standard model the fermions are divided into three families or generations. The restriction to three families is a result of measurements of the Z^0 width at LEP and at SLC (see chapter 2).

Left-handed particles which take part in the weak interaction are arranged in doublets, while right-handed particles form singlets. In addition we have the antiparticles. In the case of the leptons, only the experimentally known states are represented; for example, right-handed neutrino and left-handed antineutrino states are missing.

The quarks occur in three colours. Since quarks also take part in the weak interaction, left-handed quarks are arranged in doublets and right-handed quarks in singlets. We note that here the mixed states rather than the pure QCD eigenstates play a role. The $\mathrm{SU}(2)_L$ multiplets are

$$\begin{gathered}(u_f)_R, (d_f)_R, (c_f)_R, (s_f)_R, (t_f)_R, (b_f)_R, (e^-)_R, (\mu^-)_R, (\tau^-)_R \\ \begin{pmatrix} u_f \\ d'_f \end{pmatrix}_L, \begin{pmatrix} c_f \\ s'_f \end{pmatrix}_L, \begin{pmatrix} t_f \\ b'_f \end{pmatrix}_L, \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L\end{gathered} \tag{1.123}$$

where $f = r, b, g$.

The mixed states d', s', b' are generated by the Kobayashi–Maskawa matrix U_{KM} (see section 1.3.2)

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U_{\mathrm{KM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.124}$$

One important task of modern physics is the determination of the free parameters of the standard model. These are:

- coupling constants: e, g_s, $\sin\theta_W$;
- boson masses: m_W, m_{Higgs};
- lepton masses: m_e, m_μ, m_τ;
- quark masses: m_u, m_d, m_s, m_c, m_t, m_b;
- parameters of the Kobayashi–Maskawa matrix U_{KM}: three angles θ_i and a phase δ.

m_Z and g and g' may be expressed in terms of the above quantities. The neutrinos are assumed to be massless.

We summarize a number of important features of the standard model below. These are *not* predictions, since these properties essentially form the basis upon which the model was drawn up.

- There are no transitions between leptons and quarks nor between leptons and antileptons or quarks and antiquarks. Thus, the lepton number L and the baryon number B are separately conserved.
- One family contains only the experimentally known left-handed neutrino and the associated right-handed antineutrino.
- Neutrinos are massless.
- The weak interaction (more precisely, the charged weak currents) has a pure $(V - A)$ structure, i.e. the interaction is purely left-handed ('maximal parity violation').
- The coupling constants g, g' and g_s are free parameters.
- The charge of the proton is exactly the same as that of the positron, although the lepton and quark sectors are not coupled.

The standard model *predicts* the following:

- The proton is stable.
- The double-beta decay always involves the emission of two electrons (positrons) *and two* antineutrinos (neutrinos). The neutrinoless double-beta decay ($0\nu\beta\beta$ decay) is forbidden.

The experimental test of these predictions and of the assumption of a vanishing neutrino mass is the subject of chapters 4 and 6.

1.5 MODELS OF GRAND UNIFICATION

1.5.1 Motivation for GUTs

The objective of a Grand Unified Theory (GUT) is to explain the phenomenologically very different electroweak and strong forces, and ultimately also gravity, as the result of a *single* elementary basic principle. In this fundamental approach, it should be possible to derive the different properties of the individual forces at low energies. The standard model described in outline above provides a common description of the electroweak and the strong interaction but not a true unification of all three interactions. The latter would require that the three different interaction strengths were derivable from a *single* fundamental coupling constant. This can be achieved if the corresponding symmetry group $\mathcal{G}$ of the new theory is a *simple group*.

Since the new theory must contain the GWS theory and QCD, we must have

$$\mathcal{G} \supset \mathrm{SU}(2)_L \otimes U(1) \qquad \mathcal{G} \supset \mathrm{SU}(3)_c. \tag{1.125}$$

Moreover, since the $\mathrm{SU}(3)_c$ transformations of QCD and the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)$ transformations are mutually independent (they commute), $\mathcal{G}$ must also contain the direct product

$$\mathcal{G} \supset \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1). \tag{1.126}$$

The simplicity of $\mathcal{G}$ is crucial. A group $\mathcal{A}$ is said to be simple if it does not have a decomposition of the form

$$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \ldots \otimes \mathcal{A}_n. \tag{1.127}$$

This ensures that the theory only contains *one* coupling constant. The smallest group satisfying these conditions is $\mathcal{G} = \mathrm{SU}(5)$.

Thus, in GUT models, it is assumed that the symmetry group $\mathcal{S}$ of the standard model is part of a larger simple group $\mathcal{G}$, which is only 'visible' at higher energies (typically around 10^{15} GeV).

1.5.2 Effective coupling constants

Only one unified interaction with a typical coupling constant may be derived from a simple group. However, this symmetric state only exists at extremely high energy densities.

At the energies which are normally accessible in today's experiments this GUT symmetry is broken, whence the individual interactions appear very different to us. However, they are only different aspects of a single fundamental force. In addition to the known gauge bosons (photon, $W^{\pm}$, Z^0, gluons) we also expect that there exist as yet undiscovered bosons (e.g. X and Y bosons with masses in the area of the GUT energy scales), the numbers and properties of which are determined by the special symmetry group.

The question arises as to how the interaction constants g_s, g, g' (or g_s, G_F and α) may be derived from a fundamental constant g_GUT and whether there are experimental indications of the existence of g_GUT.

An attempt to answer this question leads us to the concept of the effective (or running) coupling constants. The experimentally observed coupling constants are actually not constants; they depend, to a greater or lesser extent, on the distance and the energy. This dependence is a consequence of the vacuum polarization and of other higher-order effects.

The vacuum polarization in QED is an effect of the interaction of photons with virtual electron–positron pairs. An electrical charge polarizes the virtual e^+e^- pairs which are always present in the vacuum. The vacuum essentially reacts like a dielectric. Thus, at large distances it is only possible to see the

sample charge screened by the surrounding polarization charges. This means that the effective charge increases as the sample charge gets nearer, since the screening effect is reduced.

According to these ideas, deviations from the Coulomb law occur when the relative distances are very small, i.e. at very high energies; an example of these deviations is the Lamb shift. Qualitatively, we expect a slight increase in the strength of the electromagnetic interaction as the energy increases.

The true coupling constant, or the bare charge, is not experimentally accessible, since the vacuum polarization and other higher-order effects cannot be excluded in any experiment.

For this reason, the concept of running coupling constants or effective charge was introduced. These effective coupling constants take account of many of the higher-order renormalization effects (for a discussion, see also [Gro89, 90, 92]). These renormalization effects depend on the square of the absolute value of the four-momentum q^2. As a result of the conservation of energy and momentum, q^2 is always negative for a virtual exchange boson. Thus, the positive quantity $Q^2 = -q^2$ is usually used instead of q^2.

In gauge theories, the dependence of the effective coupling constants on Q^2, $g(Q^2)$, for large Q^2 is described by the *renormalization group equation*

$$\frac{\mathrm{d}(g(Q^2))^2}{\mathrm{d}(\ln Q^2)} = bg^4(Q^2) + \mathrm{O}(g^6) \tag{1.128}$$

where

$$b = -\frac{1}{(4\pi)^2}\left(\frac{11}{3}C - \frac{4}{3}T\right). \tag{1.129}$$

If m_f denotes the mass of the heaviest fermion involved in the interaction, then the parameters C and T depend on the gauge group and the particle multiplet for $Q^2 \gg m_f^2c^2$. The solution of the renormalization group equation is given by

$$\frac{1}{(g(Q^2))^2} = \frac{1}{(g(Q_0^2))^2} + b\ln\left(\frac{Q_0^2}{Q^2}\right) \tag{1.130}$$

where Q_0 denotes an arbitrary reference point.

The sign of b is crucial as far as the dependence of the coupling constants on Q^2 is concerned. For a U(1) gauge group b is positive. Thus, we obtain the increase in the coupling constant with increasing energy caused by the screening effect.

Unlike the photon, gluons have a charge and thus interact strongly amongst themselves. In this case, the vacuum polarization leads via virtual quark–antiquark pairs and colour-charged gluons to an increase in the effective coupling strength for large distances or at low energies ($\Rightarrow$ confinement). An analogous

behaviour is found for the weak interaction, since, in this case, the gauge bosons have a weak charge and consequently interact with each other via the weak force. However, the corresponding dependence on Q^2 is much less marked.

Unlike the U(1) symmetry, the SU(2) and the SU(3) gauge groups lead to a negative parameter b. The coupling strengths decrease as Q^2 increases. For the colour interaction we have the following dependence (see e.g. [Lan81, 86])

$$\alpha_s(Q^2) = \frac{12\pi}{33 - 2n_f} \frac{1}{\ln(Q^2/\Lambda^2)} \tag{1.131a}$$

where

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{1.131b}$$

Here, Λ is a scale factor to be determined in the experiment ($\Lambda \simeq 0.3$ GeV) and n_f denotes the number of quark flavours. For $n_f \leq 16$, (1.131a) explains the asymptotic freedom of the quarks, since

$$\alpha_s(Q^2) \to 0 \text{ as } Q^2 \to \infty. \tag{1.132}$$

The sharp increase in $\alpha_s(Q^2)$ as Q^2 becomes small points to the confinement. However, we note that (1.131a) is the result of a perturbation calculation with the expansion parameter α_s, so that the expansion becomes invalid for large α_s. Figure 1.7 shows the schematic graph of the effective coupling constants g, g' and g_s as a function of Q^2.

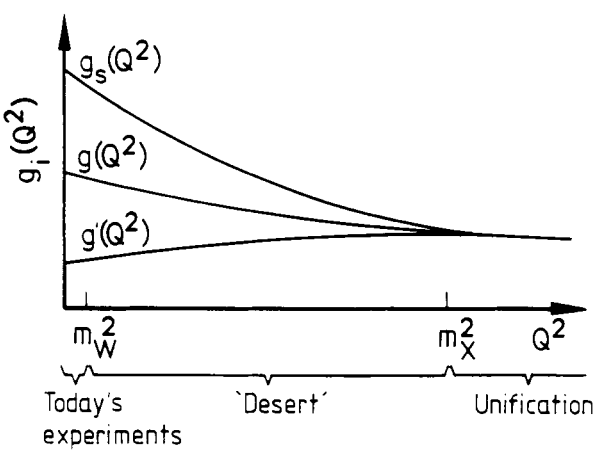


Figure 1.7. Schematic representation of the Q^2 dependence of the effective coupling constants g and g' of the electroweak interaction and g_s of the colour interaction. Above $Q^2 = m_X^2 c^2$ all three coupling constants become one (from [Gro89,90]).

The renormalization group equation may be used to extrapolate coupling parameters known experimentally for low energies to high energies. In fact, one finds that $g(Q^2)$, $g'(Q^2)$ and $g_s(Q^2)$ approach one another as Q^2 increases and finally appear to merge in the region between 10^{15} and 10^{16} GeV. A more precise analysis, taking into account the most recent experimental data, is given in chapter 2.

This energy dependence is an important indication of the existence of a GUT interaction. However, the unification energy at which the phase transition from unbroken to spontaneously broken GUT symmetry takes place lies around 10^{15} to 10^{16} GeV, far beyond the limits of accessibility using terrestrial accelerator installations.

According to the above considerations, the energy dependence of the gauge symmetry may be schematically illustrated as follows (m_X denotes the mass of the X, Y bosons):

$$\begin{array}{l} \mathcal{G} \\ \quad \downarrow E_X \simeq m_X c^2 \simeq 10^{15}\ \text{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1) \\ \quad \downarrow E \simeq m_W c^2 \simeq 100\ \text{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{U}_{\mathrm{EM}}(1). \end{array} \tag{1.133}$$

According to these GUT models, we do not expect any 'new physics' in the broad energy region between 100 GeV and 10^{15} GeV. However, from experience to date, it seems unlikely that no fundamentally new phenomena will become apparent over this enormously broad region. There are also other types of theory which populate this 'desert' with new intermediate mass scales. We shall only mention the right–left symmetric model due to Pati and Salam [Pat74, Moh86a] at this point. This is based on the following symmetry breaking scheme:

$$\begin{array}{l} \mathcal{G} \\ \quad \downarrow ? \\ \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_R \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_{B-L} \\ \quad \downarrow m_{W_R} c^2 \simeq 10^4\ \text{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1) \\ \quad \downarrow m_{W_L} c^2 \simeq 10^2\ \text{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{U}_{\mathrm{EM}}(1). \end{array} \tag{1.134}$$

A right-handed $\mathrm{SU}(2)_R$ gauge group is introduced here, i.e. right-handed fermions also take part in the weak interaction. However, the right-handed W bosons are given a greater mass than their left-handed partners, so that the coupling at low energies is correspondingly weak.

In the following we shall briefly describe a number of other, specific GUT models.

1.5.3 The SU(5) model

The simplest realization of a GUT model with the properties required in section 1.5.1 is the SU(5) model proposed by Georgi and Glashow [Geo74]

$$\mathcal{G} = \mathrm{SU}(5). \tag{1.135}$$

SU(5) is the *smallest* simple group which satisfies the conditions of section 1.5.1.

The fermions are arranged in three families, as usual. In what follows we shall only consider the first family; the extension is obvious.

The SU(5) theory contains the following 15 known left-handed fermions

$$u_g, u_r, u_b, u_g^c, u_r^c, u_b^c, d_g, d_r, d_b, d_g^c, d_r^c, d_b^c, e^-, e^+, \nu_e. \tag{1.136}$$

These are arranged in two multiplets:

$$\bar{5}_L = \begin{pmatrix} d_r^c \\ d_b^c \\ d_g^c \\ e^- \\ \nu_e \end{pmatrix} \qquad 10_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -u_g^c & u_b^c & u_r & d_r \\ & 0 & -u_r^c & u_b & d_b \\ & & 0 & u_g & d_g \\ & \text{anti-} & & 0 & e^+ \\ & \text{symmetric} & & & 0 \end{pmatrix}. \tag{1.137}$$

In addition, there are the 15 right-handed antiparticles.

The SU(5) model involves only the known fermions. There is no room for left-handed antineutrinos or right-handed neutrinos. There are 24 gauge bosons ($(n^2 - 1) = 5 \times 5 - 1$; see e.g. [Gro89, Gro90], section A.5.3), including the X and Y bosons in addition to the 12 known (γ, $W^{\pm}$, Z^0, 8 gluons) bosons (see table 1.4).

As in the GWS model, the breaking of the SU(5) symmetry occurs spontaneously by coupling to Higgs fields. The SU(5) symmetry breaking at 10^{15} GeV could be generated by a 24-dimensional Higgs field. During this, only the X and Y bosons acquire a mass by coupling to the finite vacuum expectation value; the other bosons, like the fermions, remain massless. The breaking of the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)$ symmetry at 10^2 GeV requires a further 5-dimensional Higgs field which is responsible for the mass of $W^{\pm}$ and Z^0 and of the fermions.

We shall now summarize some of the important properties of the minimal SU(5) model:

- The model only contains the two neutrinos ν_L and $(\nu^c)_R$ which are currently known from experiments. Thus, the neutrino is a Majorana particle (see section 1.6).
- Neutrinos are massless. Thus, in a purely left-handed weak interaction a distinction between Dirac and Majorana neutrinos makes no sense (see section 1.6 and chapter 6, see also [Gro89, 90, 92]). The neutrinoless double-beta decay is forbidden.

- The exact equality of the proton and positron charge is a simple consequence of the arrangement of leptons and quarks in a *single* multiplet: the SU(5) multiplets must be electrically neutral, i.e. the charges of the individual fields must sum to zero in each multiplet.
- Since the leptons and quarks are arranged in a single multiplet, they may also be converted into each other. The transitions between quarks and leptons are mediated by the X and Y bosons. Thus, the SU(5) predicts the non-conservation of the baryon number and the lepton number. The instability of baryonic matter (proton decay) is a direct consequence of this.
- The baryon number B and the lepton number L are not conserved separately, but $(B - L)$ is conserved.
- There is a total of 24 gauge bosons. In addition to the known γ, $W^{\pm}$ and Z^0 bosons and the 8 gluons, these are the X and Y bosons with charges $\pm\frac{4}{3}e$ and $\pm\frac{1}{3}e$ and masses in the region of 10^{15} GeV.
- The coupling constants g, g' and g_s can be derived from the universal coupling constant g_5 of the SU(5) gauge group.
- The SU(5) model makes statements about the Weinberg angle. At the unification point
$$\sin\theta_W = \tfrac{3}{8} = 0.375. \tag{1.138}$$
For a comparison with experiments, this value must be converted to laboratory energies. The predictions lie in the region of the values found experimentally; however, there is no agreement within the limits of error. This discrepancy can be eliminated by a supersymmetric extension of the minimal SU(5) model (see also chapter 2).
- Magnetic monopoles with extraordinarily high masses ranging from 10^{16} to 10^{17} proton masses are predicted (see chapter 8).

The experimental results relating to the Weinberg angle, the unification energy and the proton lifetime, which are presented in later chapters, suggest an extension of the minimal SU(5) model.

1.5.4 The SO(10) model

The gauge group SO(10) is another candidate for the sought-after GUT symmetry [Fri75, Geo75]

$$\mathcal{G} = \mathrm{SO}(10). \tag{1.139}$$

SO(10) is a simple group and contains SU(5) as a subgroup

$$\mathrm{SO}(10) \supset \mathrm{SU}(5). \tag{1.140}$$

In addition to the 24 gauge bosons of the SU(5) model this model includes other bosons which lead to a conversion of elements of the $\overline{5}_L$ representation into

elements of the 10_L decuplet and vice versa. While the elementary fermions occur in two different representations ($\bar{5}$ and 10) in the SU(5) model, in the SO(10) model all fermions of the same family belong to the same multiplet. This SO(10) multiplet covers the 15 elements of the SU(5) quintet and the SU(5) decuplet and a further 16th element which forms an SU(5) singlet and does not take part in the SU(5) interactions.

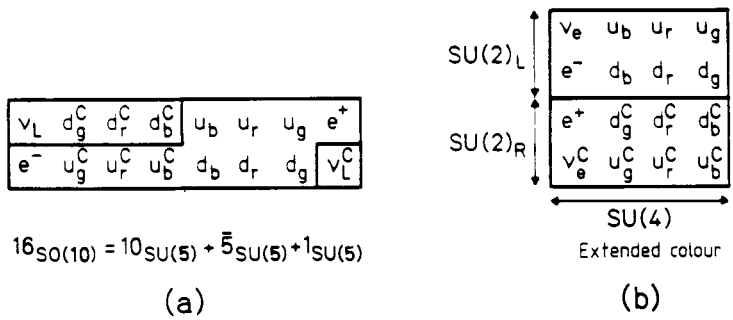


Figure 1.8. Particle contents of the SO(10) multiplet. The 16th element is the as yet undetected right-handed neutrino ν_R or its CP partner ν_L^C. Part (*b*) shows the decomposition of the SO(10) multiplet according to the $SU(4)_{EC} \otimes SU(2)_L \otimes SU(2)_R$ structure (from [Gro89,90]).

This 16th element of the SO(10) multiplet is identified with the right-handed partner ν_R of the experimentally known left-handed neutrino. Figure 1.8 shows the particles contained in the SO(10) model. Since the 16-dimensional fermion multiplet is again purely left-handed, the 16th element is not ν_R itself but the corresponding left-handed antineutrino $(\nu^c)_L$.

ν_R is not involved in any of the SU(5) interactions, and, in particular does not take part in the weak interaction. However, it does participate in a very much weaker interaction mediated by the new SO(10) bosons, which is the right-handed counterpart of the normal weak interaction.

Thus, SO(10) is a right–left symmetric model, indeed it is the simplest such model.

Unlike the SU(5) model, the SO(10) model includes mechanisms which lead not only to the non-conservation of the baryon number B and the lepton number L but also to a non-conservation of $(B - L)$. Only this breaking of the $(B - L)$ symmetry makes, for example, the neutrinoless $\beta\beta$ decay possible (see chapter 6).

Since the SO(10) symmetry contains the SU(5) symmetry it is possible that, for energies $E < m_X c^2 \simeq 10^{15}$ GeV, there may be no difference between SU(5) and SO(10) and that the SO(10) symmetry is broken at an even greater energy.

This would give the following scheme

$$\begin{array}{ll} \mathrm{SO}(10) & \\ \quad\downarrow & E = M_{10}c^2 > m_X c^2 \simeq 10^{15}\ \mathrm{GeV} \\ \mathrm{SU}(5) & \\ \quad\downarrow & E = m_X c^2 \simeq 10^{15}\ \mathrm{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1) & \\ \quad\downarrow & E = m_W c^2 \simeq 10^{2}\ \mathrm{GeV} \\ \mathrm{SU}(3)_c \otimes \mathrm{U}_{\mathrm{EM}}(1). & \end{array} \tag{1.141}$$

However, another breaking scheme which conserves the right–left symmetry also for energies below $M_{10}c^2$ is conceivable [Pat74]

$$\begin{array}{ll} \mathrm{SO}(10) & \\ \quad\downarrow & E = M_{10}c^2 \\ \mathrm{SU}(4)_{\mathrm{EC}} \otimes \mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R. & \end{array} \tag{1.142}$$

The index EC stands for 'extended colour'. We shall not go into the meaning of $\mathrm{SU}(4)_{\mathrm{EC}}$. We see that in this scheme the $\mathrm{SU}(2)_R$ occurs as a counterpart to $\mathrm{SU}(2)_L$. Since the right-handed weak interaction described by $\mathrm{SU}(2)_R$ has not yet been observed, the right-handed $W^{\pm}$ bosons which mediate it must be very heavy. An analysis of the neutrinoless double-beta decay gives the bound [Moh86c, 88b, 88c]

$$m_{W_R} \geq 800\ \mathrm{GeV}/c^2. \tag{1.143}$$

A higher limit is suggested by an analysis of the $K^0\overline{K^0}$ system [Moh88b]

$$m_{W_R} \geq 1.6\ \mathrm{TeV}/c^2. \tag{1.144}$$

An even higher bound comes from a consideration of the neutrino emission in supernovae [Moh91a]. Under certain, relatively restrictive conditions, including the restriction that the mass of the right-handed neutrino should be less than 10–100 MeV/c^2, we have

$$m_{W_R} \geq 23\ \mathrm{TeV}/c^2. \tag{1.145}$$

In addition to the SU(5) and the SO(10) models, there are a number of other models, some of which have even larger gauge groups; we shall not go into these here. For a detailed presentation of the models considered in the above, we refer readers to [Lan81, 88, 93, Moh86a, 88b, 88c, Gro89, 90, 92, Wil93].

1.5.5 Supersymmetric GUT models

Supersymmetry (SUSY) was introduced into elementary particle physics by Akulov and Volkov [Aku72] and Wess and Zumino [Wes74]. The SU(5) and

SO(10) GUT models discussed above considered a symmetry between quarks and leptons. Supersymmetry now produces a symmetry between fermions and bosons (see figure 1.9); fermions and bosons are combined into supermultiplets. This is a completely new type of symmetry, since the particle spin changes under a transformation. Such a theory must be viewed as being as fundamental as the CPT theory which represents a symmetry between particles and antiparticles. Wess and Zumino [Wes74] first formulated a renormalizable theory which contains such a symmetry between fermionic and bosonic degrees of freedom.

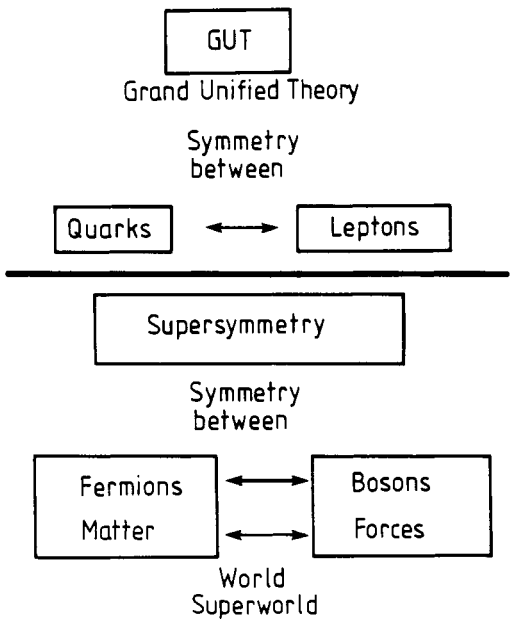


Figure 1.9. Grand unification (GUT): symmetry between quarks and leptons; supersymmetry (SUSY): symmetry between fermions and bosons (from [Scho89]).

Every GUT model based on a normal gauge symmetry may be extended to a supersymmetric version. The symmetry between fermions and bosons is such that every fermion has a bosonic partner in the same multiplet, and vice versa. In the case of an unbroken symmetry, the two partners have the same mass.

We already know a number of fundamental fermions and bosons in nature. Unfortunately, however, it is not yet possible to combine any of these particles into a supermultiplet. Thus, in the current SUSY models the number of particles is practically doubled. For each known fundamental fermion (lepton or quark)

there is a bosonic partner with spin 0 (slepton or squark). Analogously, the known bosons (photon, gluon, W, Z, Higgs) are associated with fermions, which take the ending -ino (photino, gluino, wino, zino, higgsino; see table 1.7).

Table 1.7. The SUSY partners of a number of particles.

Normal particle	SUSY partner	Short notation	Spin
quark	squark	$(\tilde{q})$	0
lepton	slepton	$(\tilde{l})$	0
gluon	gluino	$(\tilde{g})$	1/2
W boson	wino	$(\tilde{w})$	1/2
photon	photino	$(\tilde{\gamma})$	1/2
Higgs	higgsino	$(\tilde{h})$	1/2
graviton	gravitino		3/2

If the supersymmetry were exact, the SUSY partner would be degenerate in the mass. Since this is not the case in experiments (proof of the existence of SUSY particles has not yet been found), the SUSY must be broken. The mass scale at which the supersymmetry is broken should also roughly determine the mass of the superparticles.

Even more fields are needed to generate the breaking of SUSY. Since these fields have 'non-physical' properties, they are only acceptable if they remain decoupled from the observable world. In this context, we speak of a visible and an invisible or *hidden particle sector*. In some models this non-observability of the invisible sector is guaranteed by introducing a multiplicative quantum number, the R parity. R is positive for particles of the visible sector and negative for those of the invisible sector. Thus, the decay of an invisible particle into particles solely in the visible sector is forbidden.

SUSY models are very attractive candidates for models beyond the standard model (see e.g. [Hab93]) since the renormalization properties are very much better than in conventional theories (non-renormalization theorems, see e.g. [Gro89, 90, 92]). Finally, theories with gauged supersymmetry (requirement for *local* SUSY) open a natural route towards the unification of gravity with the other interactions. Local supersymmetric models automatically contain gravity and are known as supergravity (SUGRA) models (see e.g. [Gro89, 90, 92]). In addition to the gauge field of gravity with spin 2, there is a further gauge field with spin 3/2. This describes the SUSY partner of the graviton, the gravitino.

The ultimate objective of a SUGRA theory is to derive all phenomena from one theory with the Planck mass (see section 3.1) as the only parameter. The Planck mass $m_{\mathrm{Pl}} = 1.2 \times 10^{19}$ GeV/c^2 is a characteristic mass of gravity, in

as much as the Newtonian constant of gravity is equal to one in units of m_{Pl}. For energies and momentum transfers of the order of magnitude m_{Pl} gravity dominates all other interactions (see e.g. [Oku87]).

1.5.6 Superstrings

Supersymmetry eliminates a number of divergences which arise in the quantum field theory of non-gravitative forces. Nevertheless, on its own, the combination of supersymmetry and quantum field theory does not yet appear to lead to a consistent quantum theory of gravity. As long as the theory contains point-like objects (fermions) it diverges for energies above the Planck energy (on the problem of quantizing gravity, see e.g. [deWit62]).

It may be possible to bypass this problem by interpreting fermions as one-dimensional extended objects, or *strings*, rather than as point-like objects. Below the Planck mass such a theory should be indistinguishable from a point-like quantum field theory with SUSY. The string nature is first noticeable at mass scales above the Planck mass, when it prevents the divergences caused by the point-like fermions.

The gauge groups SO(32) and $E_8 \otimes E_8$ (see e.g. [Gre85, 86, Schw85]) are particularly attractive. Both groups contain 496 gauge fields and both also contain the 24 generators of the minimal SU(5). These gauge groups would have to be gradually broken down to the SO(10) and SU(5) symmetries so that the superstring theory does not differ from the normal GUT models at smaller energies ($E \leq 10^{15}$ GeV). All the 'low-energy' physics would be contained in a *single* E_8 factor [Gro85a]. An $E_8 \otimes E_8$ superstring theory leads to a further complication. It is only possible in a 10-dimensional geometrical space. However, this must compactify to the observable four space–time dimensions (see also chapter 12). For further literature, see [Wit81, Lee84a, Der86, Sha86, Val86, Dra87, Moh88b, c] and, in particular, [Fre88, Gre87].

The Kaluza–Klein theories, which we shall discuss in chapter 12, provide another way of incorporating gravity.

1.6 THE DESCRIPTION OF NEUTRINOS

The mathematical representation of neutrinos involves a peculiarity which does not exist in the case of charged fermions. In addition to the description by the Dirac theory, there is also an alternative Majorana description. We shall now discuss this ambiguity.

1.6.1 Parity and charge conjugation for neutrinos

We begin with a number of remarks about the relationship between neutrinos and antineutrinos. To date, only left-handed neutrinos ν_L and right-handed antineutrinos $\overline{\nu}_R$ have been detected experimentally. Strictly speaking, $\overline{\nu}_R$ is not the charge-conjugate particle corresponding to ν_L, since spin and momentum remain unchanged by charge conjugation. The operator C does not affect the handedness. In fact, ν_L and $\overline{\nu}_R$ are related by the CP operation. The parity operator P is responsible for the change in sign of the handedness

$$(\nu_L)^{CP} = \overline{\nu}_R. \tag{1.146}$$

The charge-conjugate particle corresponding to ν_L would again have to be a left-handed particle. There are two possibilities for this:

(i) The neutrino ν_L is its own charge-conjugate particle

$$(\nu_L)^C = \nu_L. \tag{1.147}$$

Correspondingly, we would also have

$$(\overline{\nu}_R)^C = \overline{\nu}_R. \tag{1.148}$$

There are only two physically distinguishable states. In this case we speak of a Majorana neutrino.

(ii) The charge-conjugate particle corresponding to ν_L and that corresponding to $\overline{\nu}_R$ are independent particles which have not yet been detected experimentally. This is a four-component theory. In this case we speak of a Dirac neutrino.

For clarity, in the following, we shall write ψ^C to denote the CP-conjugate state of ψ.

In relativistic quantum theories, fermions with half-integral spin are described by spinors with four components

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \tag{1.149}$$

which are solutions of the Dirac equation

$$(\mathrm{i}\gamma_\mu \partial^\mu - m)\psi(x) = 0. \tag{1.150}$$

The solution in terms of plane waves

$$\psi_-(x) = u \cdot \mathrm{e}^{-\mathrm{i}px} = u \cdot \mathrm{e}^{-\mathrm{i}(Et - \boldsymbol{p}\boldsymbol{x})} \tag{1.151a}$$

$$\psi_+(x) = v \cdot \mathrm{e}^{\mathrm{i}px} = v \cdot \mathrm{e}^{\mathrm{i}(Et - \boldsymbol{p}\boldsymbol{x})} \tag{1.151b}$$

gives four independent basic spinor fields u_1, u_2, v_1 and v_2 for $m \neq 0$. Thus, the Dirac field has four degrees of freedom. However, for massless fermions only two of the four basic Dirac spinors are linearly independent of each other.

For positive energies, the solutions $\psi_-(x)$ describe particles, while the solutions $\psi_+(x)$ describe the corresponding antiparticles. The two degrees of freedom u_1 and u_2 for $\psi_-(x)$ and v_1 and v_2 for $\psi_+(x)$, respectively, define the possible spin settings or helicities.

The left- and right-handed components of $\psi(x)$ are easily obtained by applying the projection operators

$$P_L = \tfrac{1}{2}(1 - \gamma_5) \tag{1.152a}$$

$$P_R = \tfrac{1}{2}(1 + \gamma_5) \tag{1.152b}$$

to the spinor ψ

$$\psi_L = P_L \psi \tag{1.153a}$$

$$\psi_R = P_R \psi. \tag{1.153b}$$

An antiparticle state results from the particle state by application of the charge-conjugation operator C. We define the antiparticle field by

$$\psi^C = C\overline{\psi}^T \tag{1.154}$$

where $\overline{\psi} = \psi^\dagger \gamma^0$ and $C = \mathrm{i}\gamma_2\gamma_0$. T denotes transposition.

Thus, it follows that the antiparticle state corresponding to a left-handed state is a right-handed state

$$(\psi_L)^C = (P_L\psi)^C = C\overline{P_L\psi}^T = P_R(C\overline{\psi}^T) = P_R\psi^C = (\psi^C)_R. \tag{1.155}$$

Here, we have used the commutation rules for the γ matrices (see e.g. [Gro89, 90]).

The four degrees of freedom of the Dirac field ψ correspond to the particle and antiparticle states with two possibilities for the handedness in each case ($\psi_L^C \equiv (\psi^C)_L$)

$$\psi_R \quad \psi_L \quad \psi_R^C \quad \psi_L^C. \tag{1.156}$$

In the case of the electron, these four components are actually different, since e_L^- and the CP-conjugate state corresponding to e_R^- (namely $(e_R^-)^C = e_L^+$) have different electrical charges. However, in the case of the neutrino, it is not clear whether ν_R and ν_R^C and ν_L and ν_L^C, respectively, are in fact different, or whether there are only two independent degrees of freedom.

We shall study the difference between Dirac and Majorana neutrinos in more detail below. A detailed presentation can be found in [Gro89, 90, 92, Kay89].

1.6.2 Dirac and Majorana description

In the framework of the weak interaction, one often works with two-component spinors ψ_L and ψ_R, which are also called Weyl spinors. In the language of the second quantization, ψ_L may destroy a left-handed particle or create a right-handed antiparticle, while ψ_L^+ may create a left-handed particle or destroy a right-handed antiparticle. The roles of L and R are exactly reversed for ψ_R. In the interaction-free case we write

$$\psi_L(x) = \int \frac{\mathrm{d}^3 p}{\sqrt{(2\pi)^2 2E}} \left(b_L(\boldsymbol{p}) u_L(\boldsymbol{p}) \mathrm{e}^{-\mathrm{i}px} + d_R^\dagger(\boldsymbol{p}) v_R(\boldsymbol{p}) \mathrm{e}^{\mathrm{i}px} \right). \tag{1.157}$$

Here, b_L and d_R are annihilation operators for L particles and R antiparticles and $b_L^\dagger$ and $d_R^\dagger$ are the corresponding creation operators. The roles of L and R are exactly reversed for $\psi_R(x)$.

A conventional four-component Dirac field $\psi(x)$ may be written as the sum of the left- and right-handed projections or Weyl spinors

$$\psi = \psi_L + \psi_R. \tag{1.158}$$

Thus, we again obtain the customary representation of the Dirac field

$$\psi(x) = \sum_{s=L,R} \int \frac{\mathrm{d}^3 p}{\sqrt{(2\pi)^3 2E}} \left(b_s(\boldsymbol{p}) u_s(\boldsymbol{p}) \mathrm{e}^{-\mathrm{i}px} + d_s^\dagger(\boldsymbol{p}) v_s(\boldsymbol{p}) \mathrm{e}^{\mathrm{i}px} \right). \tag{1.159}$$

After these introductory remarks, we now return to the specific case of the neutrino.

1.6.2.1 Dirac neutrino

Let us assume that the state ν_R exists and is not identical to ν_R^C, so that ν_L, ν_R, ν_L^C and ν_R^C form a four-component Dirac particle and we may write

$$\nu_D = \nu_L + \nu_R \tag{1.160a}$$

$$\nu_D^C = \nu_L^C + \nu_R^C. \tag{1.160b}$$

A mass term in a Lagrange density generally couples fields with different helicities. Thus, in the Dirac case, we may define the following mass term[2]

$$-\mathcal{L}_D = m^D \overline{\nu}_L \nu_R + \text{h.c.} = m^D \overline{\nu}_D \nu_D. \tag{1.161}$$

In this case, the lepton number is conserved, since there are no transitions between ν and ν^C (note that $\overline{\nu}$ stands for $\overline{\nu} = \nu^\dagger \gamma^0$!).

[2] h.c. = Hermitian conjugate.

To help readers understand this, we mention that the equations of motion, such as the Dirac equation, may be derived from the Lagrange density $\mathcal{L}$ using the Euler–Lagrange equations. In this process, the Dirac equation may be divided into two coupled equations for the two handednesses. The coupling is expressed in the Lagrange density $\mathcal{L}$ in such a way, that two fields with different handednesses are coupled in the mass term, where the Dirac mass m^D denotes the coupling strength. For $m^D = 0$, the equations are decoupled, since the handedness is a conserved quantity for massless particles.

1.6.2.2 Majorana neutrino

In the case of the electrically neutral neutrino, Majorana introduced another type of coupling. This two-component theory of the neutrino also resolves the Dirac equation, but assumes that ν_R and ν_R^C are physically indistinguishable. The coupling in the mass term of the Lagrange density is between ν_L and the CP-conjugate ν_R^C and between ν_R and ν_L^C, respectively. There are two Majorana mass terms, corresponding to the coupling of the left-handed and the right-handed fields with their CP conjugates

$$\mathcal{L}_M = \mathcal{L}_M^L + \mathcal{L}_M^R \tag{1.162}$$

where

$$-\mathcal{L}_M^L = \tfrac{1}{2} m_L^M (\overline{\nu}_L \nu_R^C + \overline{\nu_R^C} \nu_L) \tag{1.163a}$$

$$-\mathcal{L}_M^R = \tfrac{1}{2} m_R^M (\overline{\nu}_R \nu_L^C + \overline{\nu_L^C} \nu_R). \tag{1.163b}$$

Both ν_L and ν_R^C and ν_R and ν_L^C may be combined into a two-component Majorana neutrino. The Majorana mass eigenstates ν_1 and ν_2

$$\nu_1 = \nu_L + \nu_R^C \tag{1.164a}$$

$$\nu_2 = \nu_R + \nu_L^C \tag{1.164b}$$

are superpositions of the linearly dependent interaction eigenstates. The Majorana neutrinos ν_1 and ν_2 are their own antiparticles, since, because $(\nu_L)^C = (\nu^C)_R$, it follows that

$$(\nu_1)^C = (\nu_L)^C + (\nu_R^C)^C = (\nu_L)^C + (\nu^{CC})_L = \nu_R^C + \nu_L = \nu_1 \tag{1.165}$$

and, analogously, $(\nu_2)^C = \nu_2$.

Using the states ν_1 and ν_2, we may rewrite $\mathcal{L}_M^{L/R}$ as follows

$$-\mathcal{L}_M^L = \tfrac{1}{2} m_L^M \overline{\nu}_1 \nu_1 \tag{1.166a}$$

$$-\mathcal{L}_M^R = \tfrac{1}{2} m_R^M \overline{\nu}_2 \nu_2. \tag{1.166b}$$

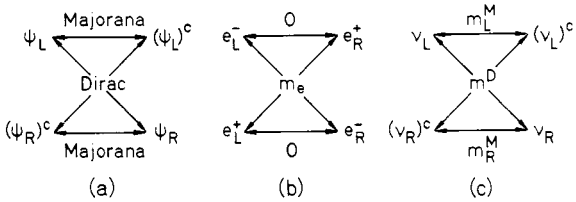


Figure 1.10. Coupling scheme for fermion fields via Majorana and Dirac masses. (*a*) General coupling scheme for left- and right-handed fields and their charge-conjugate fields. (*b*) Coupling scheme for electrons. Because of the charge, only the Dirac coupling arises. (*c*) Coupling scheme for neutrinos. Both Dirac and Majorana coupling may occur for neutrinos only (from [Mut88]).

As a result of the possible transitions between ν_L and ν_R^C, the Majorana coupling violates the lepton number by two units, $\Delta L = \pm 2$. Figure 1.10 again summarizes the coupling possibilities. It is clear that Majorana coupling is not possible for charged fermions. The hypothetical Majorana mass eigenstate of the electron

$$e_M = e_L^- + e_R^+ \tag{1.167}$$

would not have a well-defined charge, while the Dirac states

$$e_D^- = e_L^- + e_R^- \tag{1.168a}$$

$$e_D^+ = e_L^+ + e_R^+ \tag{1.168b}$$

do have a well-defined charge.

We have seen that Majorana neutrinos are characterized by the fact that neutrinos and antineutrinos are identical, so that, unlike the four-component Dirac description, a two-component Majorana description suffices. We now note that, in the case in which the neutrino does not have a mass, and only the left-handed field ν_L is involved in the weak interaction, it is not possible to determine experimentally which of the descriptions is correct (see [Gro89, 90, 92, Kay89])[3].

The Majorana case can only be distinguished from the Dirac case for $m = 0$ if the weak interaction has a right-handed component, since then the two additional Dirac degrees of freedom would also be detectable.

One other important means of distinguishing between Dirac and Majorana neutrinos is based on the magnetic moment (for massive neutrinos). By virtue

[3] This can be illustrated in a somewhat simplified picture. In the case of a purely left-handed interaction the ν_R does not take part in the interaction. If the ν_R is massless, the helicity is conserved, i.e. the ν_R cannot be detected even if it does exist.

of the property $\nu = \nu^C$, Majorana neutrinos cannot have a magnetic moment. According to the CPT theorem, the magnetic moments of the neutrino and the antineutrino should have opposing signs. However, since, by assumption, the two particles are identical, the magnetic moment vanishes (see also chapter 7).

1.6.3 The physical neutrino mass

The masses m^D, m_L^M and m_R^M introduced in the last section are initially only coupling parameters. We shall now relate them to the physical mass. In the case where only one of the above coupling schemes is in place, the coupling parameter corresponds to the mass and the coupling scheme defines the nature of the neutrino (Dirac or Majorana).

In general, the Lagrange density contains all three possible couplings

$$\mathcal{L}_m = \mathcal{L}_D + \mathcal{L}_M^L + \mathcal{L}_M^R. \tag{1.169}$$

We can summarize this in a compact form

$$\begin{aligned} -\mathcal{L}_m &= \overline{\nu}_L m^D \nu_R + \tfrac{1}{2}\overline{\nu}_R m_R^M \nu_L^C + \tfrac{1}{2}\overline{\nu}_L m_L^M \nu_R^C + \text{h.c.} \\ &= \tfrac{1}{2}(\overline{\nu}_L \overline{\nu_L^C}) M \begin{pmatrix} \nu_R^C \\ \nu_R \end{pmatrix} + \text{h.c.} \end{aligned} \tag{1.170}$$

where M is a mass matrix

$$M = \begin{pmatrix} m_L^M & m^D \\ m^D & m_R^M \end{pmatrix}. \tag{1.171}$$

The physical mass of the neutrino is an eigenvalue of the mass matrix M. Applying a suitable unitary transformation, we obtain

$$UMU^\dagger = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \tag{1.172}$$

with eigenvalues

$$m_{1/2} = \frac{m_L^M + m_R^M}{2} \pm \sqrt{\frac{(m_L^M - m_R^M)^2}{4} + m_D^2}. \tag{1.173}$$

A number of special cases are of particular interest:

(i) $m_L^M = m_R^M = 0$. We obtain a *pure Dirac neutrino* with $m = m_D$. This consists of two Majorana states with degenerate masses $m_1 = m_2 = m_D$.

(ii) $m^D \gg m_L^M = m_R^M = \varepsilon$. We obtain two almost degenerate Majorana mass eigenstates with $m_{1/2} = m^D \pm \varepsilon$ and opposing CP quantum numbers. These essentially form a Dirac neutrino with a small L-violating admixture. In this case, we speak of a *pseudo Dirac neutrino* (see [Mut88] and the references cited there).

(iii) $m_L^M = 0$ and $m_R^M \gg m^D$. Diagonalization of the mass matrix

$$M = \begin{pmatrix} 0 & m^D \\ m^D & m_R^M \end{pmatrix} \tag{1.174}$$

to

$$UMU^{-1} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \tag{1.175}$$

where

$$U = \begin{pmatrix} 1 & -m^D/m_R^M \\ m^D/m_R^M & 1 \end{pmatrix} \tag{1.176}$$

gives a heavy and a light neutrino with masses

$$m_1 = m_R^M \left[1 + \left(\frac{m^D}{m_R^M} \right)^2 \right] \approx m_R^M \tag{1.177a}$$

$$m_2 = \frac{(m^D)^2}{m_R^M} \ll m^D. \tag{1.177b}$$

The mass of the light neutrino m_2 differs from the original Dirac mass m^D by a factor m^D/m_R^M. This is the mechanism for generating small neutrino masses proposed by Gell-Mann, Ramond and Slansky [Gel79], Yanagida [Yan79b] and Stech [Ste80]. It is called the *see-saw* mechanism, which is of importance, particularly in right–left symmetric models such as the SO(10) model (see below). A large Dirac mass is reduced by the interplay with an even larger Majorana mass. Thus, at the cost of introducing a second very heavy neutrino, we have a very natural explanation for the smallness of the neutrino masses.

The two neutrinos, the light neutrino and the superheavy neutrino, can be represented by almost pure Majorana mass eigenstates. The heavy neutrino essentially consists of the states which have not yet been observed (right-handed neutrino and left-handed antineutrino).

On the other hand, up to small corrections of the order of m^D/m_R^M, the light neutrino consists of the two experimentally known states.

1.6.4 Neutrinos in GUT models

The neutrino properties depend on the choice of the gauge theory. Since the standard model is based on an $SU(2)_L \otimes U(1)$ symmetry, there are no right-handed neutrinos ν_R, so that Dirac coupling is impossible. A Majorana mass term must also vanish in the standard model, since it results in the violation of the conservation of the lepton number (which is assumed here).

In the following, we shall describe some important neutrino properties in the simplest specific GUT models.

1.6.4.1 *SU(5) neutrinos*

In the minimal SU(5) model, the simplest GUT model, the baryon number and the lepton number are not separately conserved, as in the standard model; however, $(B - L)$ is conserved. The SU(5) model contains only the two experimentally known neutrino degrees of freedom $(\nu_L, \overline{\nu}_R)$. A Dirac mass is not possible, since the right-handed partner of the ν_L for a Dirac coupling of the form $m^D \overline{\nu}_L \nu_R$ is missing. Thus, the neutrino in the SU(5) model is necessarily a Majorana particle. However, in the minimal SU(5) model, the structure of the SU(5) Higgs field does not permit an SU(5)-invariant Majorana coupling, so that the Majorana mass term also vanishes. Thus, the neutrino in the minimal SU(5) model is massless. Since the weak interaction is assumed to be left-handed, it is not possible to distinguish between a Dirac and a Majorana nature.

1.6.4.2 *SO(10) neutrinos*

Massive neutrinos occur naturally in GUT models based on the SO(10) gauge group[4]. The right–left symmetric SO(10) theory contains both the left-handed neutrino field and a right-handed neutrino ν_R (see figure 1.8). ν_L and ν_R may be combined into a Dirac field

$$\nu_D = \nu_L + \nu_R. \tag{1.178}$$

Consequently, the Lagrange density contains a Dirac mass term. The Dirac field (1.178) is arranged in a multiplet together with other fermion fields. It is a consequence of the SO(10) invariance that the Dirac mass of the neutrino cannot be independent from the Dirac masses of the other fermions. Instead, to a first approximation, the Dirac mass term of the neutrino is proportional to the u quark mass [Lan81]

$$m^D_{\nu_e} \approx m_u. \tag{1.179}$$

Such a large neutrino mass contradicts the experimental upper bound for the neutrino mass which is in the region of a few eV. The see-saw mechanism discussed in section 1.6.3 provides one way out of this apparent contradiction.

Since $(B - L)$ is not necessarily conserved in the SO(10) model, Majorana mass terms are also allowed. A Majorana mass of the neutrino would violate the $(B - L)$ quantum number, since it would permit oscillations between neutrinos and antineutrinos ($\Delta L = 2$); however, the baryon number is not affected.

We note that the existence of a Goldstone boson, the majoron, is a consequence of a spontaneous breaking of the global $(B - L)$ symmetry [Chi81, Gel81, Moh91a, Ber92b]. Under certain circumstances, this majoron would be

[4] An actual representation is found in [Moh94].

Table 1.8. Models of the neutrino mass, together with their predictions for the mass of light neutrinos (from [Lan88]). We note that, in most cases m_{ν_e} is equal to the effective neutrino mass $\langle m_{\nu_e} \rangle$ (see sections 6.2.3 and 6.2.5.4).

Model	m_{ν_e}	$\langle m_{\nu_e} \rangle$	m_{ν_μ}	m_{ν_τ}
Dirac	1–10 MeV	0	0.1–1 GeV	1–100 GeV
Pure Majorana (Higgs triplet)	Arbitrary	m_{ν_e}	Arbitrary	Arbitrary
GUT see-saw ($M \approx 10^{14}$ GeV)	10^{-11} eV	m_{ν_e}	10^{-6} eV	10^{-3} eV
Intermediate see-saw ($M \approx 10^{9}$ MeV)	10^{-7} eV	m_{ν_e}	10^{-2} eV	10 eV
$SU_{2L} \otimes SU_{2R} \otimes U_1$ see-saw ($M \approx 1$ TeV)	10^{-1} eV	m_{ν_e}	10 keV	1 MeV
Light see-saw ($M \ll 1$ TeV)	1–10 MeV	$\ll m_{\nu_e}$		
Charged Higgs	< 1 eV	$\ll m_{\nu_e}$		

emitted in the double-beta decay (see sections 6.2.3 and 6.2.4 and [Kla92b, Bec93a, Bur94]).

In the SO(10) model, a 126-dimensional Higgs field could generate both Dirac masses for all fermions and Majorana masses for the neutrino. This field only couples to ν_R and not to ν_L. It is therefore possible to obtain a vanishing m_L^M and a large m_R^M in the mass matrix (1.171).

Thus, for $m_R^M \gg m^D$ the see-saw mechanism comes into play. We obtain a light neutrino and a superheavy neutrino. m_R^M is often identified with a new breaking scale. Since the mass of the light neutrino state is given by

$$m_2 \simeq \frac{(m^D)^2}{m_R^M} \tag{1.180}$$

where m^D denotes the fermion mass (quark or lepton mass) for the same generation, it follows that the mass hierarchy of the neutrinos is given by (see [Moh88b, c])

$$m_{\nu_e} : m_{\nu_\mu} : m_{\nu_\tau} = m_u^2 : m_c^2 : m_t^2 \tag{1.181}$$

or

$$m_{\nu_e} : m_{\nu_\mu} : m_{\nu_\tau} = m_e^2 : m_\mu^2 : m_\tau^2. \tag{1.182}$$

In some models, typical values of m_{ν_e} are in the region from 10^{-1} eV/c^2 to 1 eV/c^2 (table 1.8). However, in general, the predictions vary between 10^{-11} eV/c^2 and 1 eV/c^2, depending on the particular mechanism chosen (see table 1.8 and [Lan88, 92a, Blu92]). There are, however, also other models, which predict highly degenerate masses for the different neutrino flavours instead of the mass hierarchy of table 1.8 [Lee94, Moh94, Pet94, Ioa94].

1.6.4.3 Neutrinos in superstring models

Superstring models tend to predict Majorana neutrinos with masses which are far too large (see e.g. [Moh88b, c]). This is because no Higgs fields occur which could generate a sufficiently large Majorana mass m_R^M for the right-handed neutrino. We therefore need much more complicated mixing mechanisms than those of the see-saw model to explain the smallness of the neutrino mass (e.g. [Val87]). This requires an intermediate mass scale in the TeV region, which, under certain circumstances, could be identified with the SUSY scale (see e.g. [Val93]).

1.7 OUTLOOK

The particle theories outlined above, including the standard model and all its extensions, have in common the fact that the effective symmetry increases with increasing energy. In other words, the underlying idea is a hierarchy of symmetries in which the individual phenomenological forces are split off from a higher symmetry, which incorporates a unification of all forces into an original 'fundamental force' by sequential spontaneous symmetry breaking at decreasing energies (figure 1.11). The fact that the universe passes through these successive stages as it cools after the big bang provides an extremely exciting link between particle physics and cosmology (see chapters 3 and 9). However, we must at least mention that it is also held to be conceivable that a completely chaotic behaviour may occur at high energies and that the symmetries at low energies are quite deceptive (see [Nac86]). We shall illustrate this using an example from solid-state physics. If we consider a piece of amorphous matter (glass) on an atomic scale, we find neither translational nor rotational symmetries, while the piece of matter appears homogeneous and isotropic at the macroscopic level.

In recent years, many of the fundamental assumptions (and predictions) of the particle theories described above have been studied, in particular in accelerator experiments (see chapter 2). In the next decade new accelerators (circular proton colliders and e^+e^- linear colliders which are able to reach the multi-TeV region, see chapter 2) will look at further central questions of these theories, some of which are listed in table 1.9. These include the search for the hypothetical Higgs particle, a most urgent problem which is crucial to the

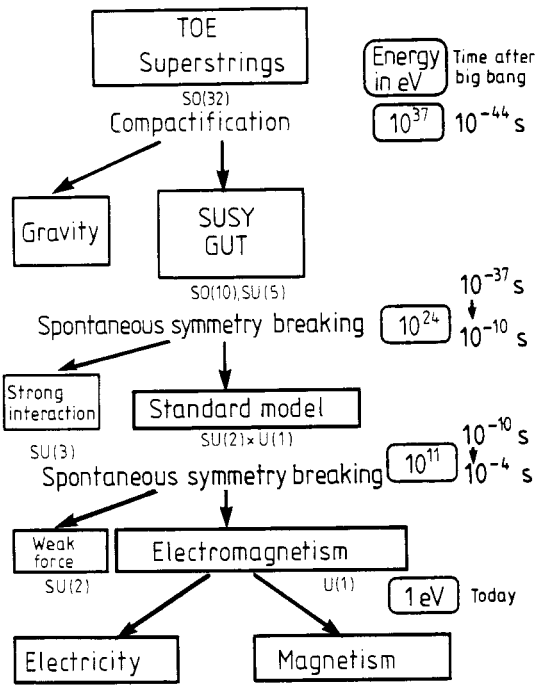


Figure 1.11. Hierarchy of symmetries. Beginning with a highest symmetry (TOE = Theory of Everything, superstrings) involving a unification of all forces into a single unified force, the individual forces are split off by successive spontaneous symmetry breakings at decreasing energies. The various stages in the cooling of the universe after the big bang (according to our present understanding) are passed through successively. The right column gives the times after the big bang (from [Scho89]).

resolution of the mass problem, together with the search for the top quark. In addition, the natural scales for the masses of SUSY partners of the conventional particles, which lie in the region from several hundred GeV to around one TeV, will become accessible. For a detailed discussion of future possibilities, we refer readers to [Jar90, Buc92, Zer92, 93].

However, central questions are reserved for non-accelerator experiments, which provide in key problems the only access to energy scales which cannot be reached on accelerators. Therefore, the importance of non-accelerator particle physics will increase considerably in the coming years. These questions principally include the properties of the neutrino, which play a key role as far as the structure of GUTs is concerned, and which are at the same time closely related to the development of the early universe. We mention here the sphaleron-induced

Table 1.9. Proposed areas of research for future accelerators (from [Zer93]).

	e^+e^- linear colliders	Proton colliders LHC/SSC
Higgs boson standard model	Mass $\leq$ 250 GeV all decay channels	Mass $\leq$ 180 GeV: rare decays Mass $\geq$ 180 GeV: $H \rightarrow ZZ \rightarrow 4l$
Higgs bosons SUSY-extension	Complete coverage of the Higgs parameter space	Partial coverage of the Higgs parameter space
W, Z bosons standard model	Magnetic dipole and electric quadrupole moments Gauge-theoretical structure of the electroweak forces	
Extended gauge theories	New leptons [heavy neutrinos]	New gauge bosons and quarks
Top	Precise measurements of mass, dipole moments, decay current; Higgs and other rare decays	Higgs decays CP violation
Supersymmetry	Sleptons and electroweak Gauginos/Higgsinos	Squarks and gluinos

violation of the conservation of baryon number below the electroweak phase transition and its connection with the neutrino mass [Fuk90, Cam91, Gel92] and also the importance of the neutrino as a candidate for dark matter (see chapter 9).

In see-saw models (sections 1.6.3 and 1.6.4.2) the search for a neutrino mass in double-beta decay or in solar-neutrino experiments in principle probes energy scales in the TeV region up to 10^{15} GeV.

In addition to information about the mass of the electron neutrino, the double-beta decay also provides information about the mass of the superheavy neutrino introduced in right–left symmetric models, about the mass of right-handed W bosons [Moh86c, 88b, 88c, 91a] and about various SUSY model parameters [Moh86b, Hir95].

The observation of neutrinos from supernova explosions provides some of the sharpest bounds on other neutrino properties such as the half-life and the magnetic moment.

A classical example of these experiments is the search for the decay of the proton, which might give the only more or less direct access to the energy scale of grand unification. Other examples include the search for dark matter in the form of WIMPs, axions, SUSY particles, magnetic monopoles, etc. The search for magnetic monopoles at the same time relates to our understanding of the

early universe (inflation) and of particle physics at the energy scales of the GUT symmetry breaking.

The search for particles with a fractional charge contributes to the verification of QCD. Experiments relating to the dipole moment of the neutron and to neutron–antineutron oscillations concern both the standard model and physics beyond that (θ problem, strong CP problem).

Studies of the time dependence of the natural constants are closely related to Kaluza–Klein theories, which could provide ways of embedding gravity in a unified description of the forces.

Chapters 3 to 12 of the present book are devoted to these possibilities offered by non-accelerator experiments which complement future accelerator experiments.

Chapter 2

Accelerator Particle Physics

Particle accelerators and storage rings are the microscopes through which the natural laws in the subatomic area are studied. Together with corresponding detectors they constitute the most important technological aids of high-energy physics.

With the continual progress of accelerator technology, major advances in the identification and study of the basic building blocks of matter and the fundamental interactions have been made in recent decades. Here, use is made of the fact that the de Broglie wavelength λ of a particle decreases as the particle energy increases. For very high energies E the rest mass m_0 may be neglected. Hence, the following relationships hold

$$E \simeq pc = \frac{hc}{\lambda} \Rightarrow \lambda \simeq \frac{hc}{E}. \tag{2.1}$$

The smaller the de Broglie wavelength λ (in other words, the greater the particle energy), the smaller the structures which can be resolved. A resolution of 1 fm ($= 10^{-15}$ m) corresponds, for example, to an electron energy of 1 GeV. On the other hand, higher energies allow the *production* of heavier particles.

Results from scattering experiments with high-energy leptons have shown that all matter is composed of point-like particles (quarks and leptons), for which no substructures can be detected on a scale of 10^{-18} m. Thus, to answer the question of the nature and possible structure of the elementary particles known today further experiments at even higher energy scales are necessary. As we have seen in chapter 1, the central questions relating to the mass of the top quark, the existence of the Higgs particle and of SUSY particles can also only be successfully studied using higher accelerator energies.

Since as energy increases the cross sections for events of interest decrease (see (2.15)), in addition to the energy the luminosity L plays a decisive role as a second parameter in the development of high-energy accelerators. L is a

measure of the beam intensity. For a reaction with a known cross section σ, the luminosity determines the event rate R via the equation

$$R = \sigma L. \tag{2.2}$$

Thus, the energy and luminosity of an accelerator determine the resolving power and the number of observable events. In addition, the development of suitable detectors is crucial to the success of high-energy accelerator experiments.

2.1 ENERGY SCALES OF CURRENT AND FUTURE ACCELERATORS

An accelerator essentially consists of an electron or ion source from which the particles emerge with a low energy, together with the accelerator tubes themselves, in which the particles in high vacuum are accelerated to the desired energy using an electrical field (figure 2.1). Subsequently, the high-energy particles are directed onto a target in order to trigger the reactions under investigation.

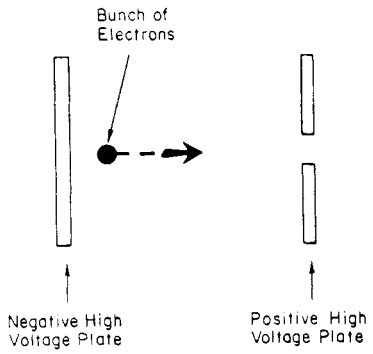


Figure 2.1. The accelerating effect in particle accelerators is achieved using electrical fields, as shown here for a simple example.

In the electrical field $\mathcal{E}$ a charged particle with charge q receives kinetic energy

$$E = q \oint_0^s \mathcal{E} \cdot \mathrm{d}s \tag{2.3}$$

where the contour integral is taken over the whole of the path covered by the particle within the field. The necessary electrical field may be generated by various methods. One technique involves applying a direct voltage U between the ion source and the target, resulting in the energy

$$E = qU. \tag{2.4}$$

Often, the acceleration is produced by alternating electromagnetic fields, where the charged particles move in orbits in a magnetic field and the $\mathcal{E}$ field is always parallel to the orbital direction. Electrons in a betatron are accelerated by means of rotational electrical fields (principle of induction). Another possibility involves the generation of travelling electrical waves. The charges 'ride' on the front edge of these waves and thus at any instant in time meet with an accelerating field in the direction of motion. For details, we refer readers to the corresponding specialized literature (e.g. [Dan74, Wil92]; see also [Scho89]).

In order to achieve very high energies, several acceleration steps are needed. Two different types of accelerator have been developed, namely linear and circular accelerators (see figure 2.2).

A linear accelerator essentially consists of a long vacuum tube containing several consecutive high-frequency acceleration stages. The alternating current must be such that a bunch of particles passing through is always accelerated. Thus, the final beam energy depends on the accelerating voltage and the overall length of the accelerator. These two factors limit the maximum energy achievable.

In circular accelerators, the charged particles are forced by magnetic fields into orbits perpendicular to the magnetic field (Lorentz force $\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}$). The same particle may pass through the accelerator many times, passing an

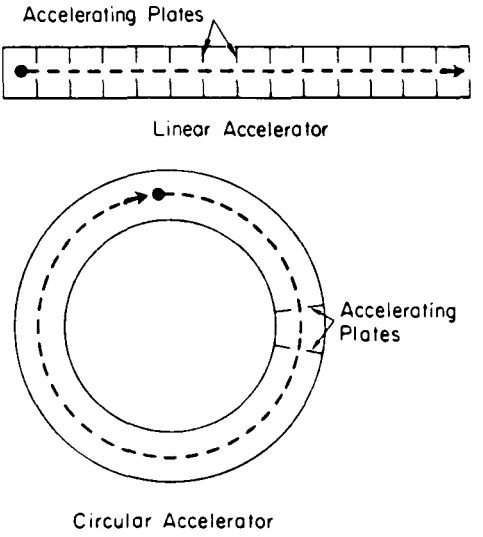


Figure 2.2. Very high end energies are achieved either by employing several successive acceleration steps like that of figure 2.1 (linear accelerator) or by passing through a single acceleration circuit several times (circular accelerator).

acceleration stage each time. High final energies are achieved by a large number of passes through the accelerator.

In standard modern synchrotron rings, the beam passes through a linear accelerator on each circuit. To ensure a constant orbital radius, the magnetic-field strengths must be continuously altered. This places high requirements on the vacuum, the beam control and the focusing, since, overall, the particles cover a very long distance in the accelerator. We shall not discuss the differences between betatrons, cyclotrons and synchrotrons here.

One disadvantage of circular accelerators is the increase in synchrotron radiation emission as the beam energy increases. By forcing the charged particles circulating in the accelerator into an orbit with radius R, they are subjected to a constant acceleration, so that energy is continuously drawn from the beam by electromagnetic emission. At very high energies, the energy supply is ultimately limited by the electromagnetic radiation. The loss of energy per circuit in an orbit with radius R is

$$\Delta E = \frac{q^2}{3\epsilon_0 R}\left(\frac{E}{m_0 c^2}\right)^4 \tag{2.5}$$

where q is the charge of the particle, ϵ_0 is the dielectric constant of the vacuum and m_0c^2 is the particle rest energy. Thus, the power emitted by synchrotron emission is

$$P \sim \frac{E^4}{R^2 m_0^4}. \tag{2.6}$$

For electrons, because of their small rest mass, this effect is considerably greater than for heavier particles such as protons, for which it does not currently play a role. The electron energies achievable in a circular accelerator depend largely on the loss of energy due to synchrotron emission. For example, in the 6.3 km HERA ring (see below), an electron at 30 GeV loses 0.5% of its energy per circuit by synchrotron emission; with 50 000 circuits per second, this figure is substantial.

Until the 1960s practically all accelerator experiments were 'fixed-target' experiments, in which the particle beam was deflected onto a stationary target and the reaction products detected. Usually, the target consisted of hydrogen, in other words, of protons. This technique was used in a wide range of pp and ep experiments to measure the absorption and scattering of beam particles in the target and to study the creation of secondary particles. In addition, beams of secondary particles such as pions, muons, kaons, neutrinos, etc were often used as projectiles for new research.

Fixed-target experiments have the disadvantage that the energy available for the production of new particles is only proportional to the square root of the beam energy E. The remaining energy is not lost, but reappears in the kinetic energy of the secondary particles. Let us briefly consider these relationships in

more detail. For a fixed target, we have

$$\begin{aligned} s = W^2 &= \big((E, pc) + (m_T c^2, 0)\big)^2 \\ &= (E + m_T c^2)^2 - |p|^2 c^2 \\ &= 2E m_T c^2 + (m_T^2 + m_0^2)c^4 \end{aligned} \tag{2.7}$$

where W is the centre-of-mass energy. In high-energy accelerators, the beam energy E is generally very much larger than the particle rest energy $m_0 c^2$ and the rest energy of the target nuclei $m_T c^2$, whence

$$s \simeq 2E m_T c^2. \tag{2.8}$$

This means that in the centre-of-mass system only the energy

$$W \simeq \sqrt{2E m_T c^2} \tag{2.9}$$

is available for reactions. Thus, this energy increases only with the square root of the accelerator energy. On the other hand this technology allows us to generate high-energy secondary beams. Figure 2.3 summarizes the maximum achievable energies of existing fixed-target accelerators.

Particle colliders and storage rings may be used to achieve greater centre-of-mass energies. The collider principle is illustrated in figure 2.4. Two particle bunches circulate in opposite directions and are brought into collision at certain interaction points. Let us consider particles with energies E_1 and E_2 and momenta $p_1 = -p_2$; then

$$s = W^2 = ((E_1, p_1 c) + (E_2, -p_1 c))^2 = (E_1 + E_2)^2. \tag{2.10}$$

For two identical particles with $E_1 = E_2 = E$, the centre-of-mass energy W is given by

$$W = 2E \tag{2.11}$$

in other words, all the collision energy is available for the creation of new particles. Most particles pass through the oncoming contrarotating bunch without colliding, so that the particle bunches circulate in the ring for several hours and may be made to interact time and again. Figure 2.5 shows the maximum energies of various colliders.

Figures 2.5 and 2.6 illustrate the development of beam energy in various accelerator installations since 1930. It is apparent that, approximately every six to seven years, an increase in the highest energies by a factor of 10 has been achieved, with older accelerator types being replaced by newer ones. Since the maximum energy of an accelerator is of crucial importance in the search for new particles, we shall now note an important difference between e^+e^- and

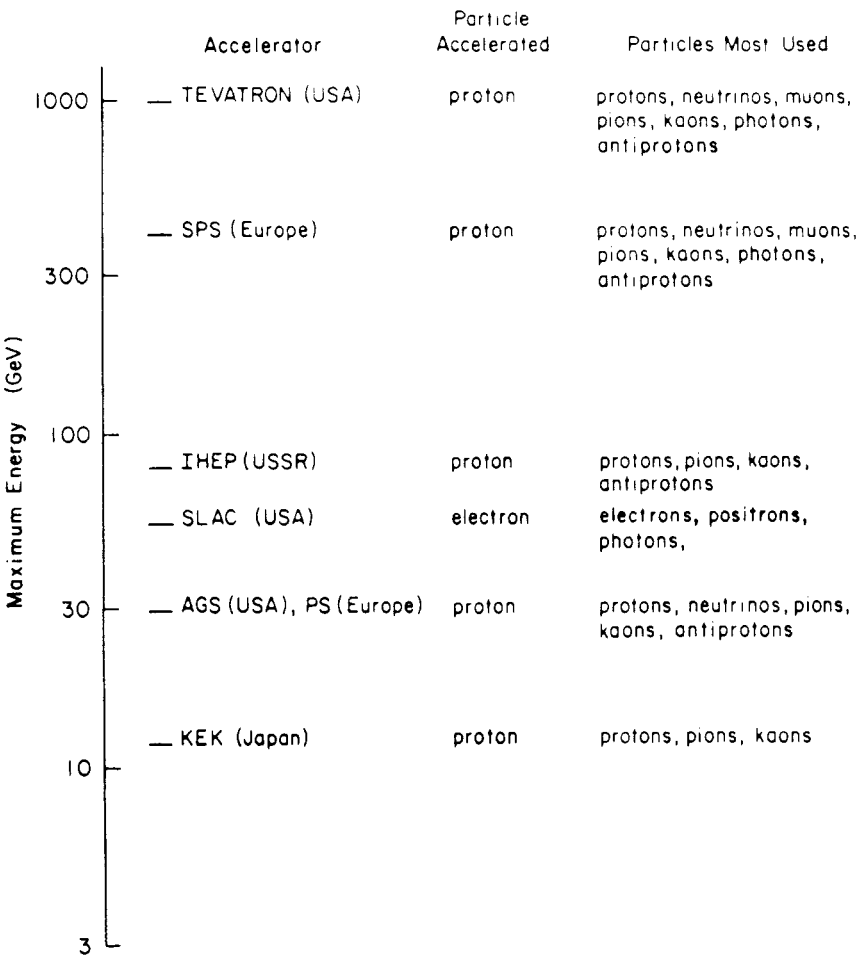


Figure 2.3. Maximal end energies achievable in present-day 'fixed-target' accelerators (from [Phy86]).

hadron colliders. In e^+e^- annihilation, all the energy of (2.11) is available for the production of new particles. On the other hand, in $p\overline{p}$ and pp collisions, a centre-of-mass energy greater than the rest energy of the particle to be created is required. This is because the reaction actually involves the interaction between the individual quarks or gluons of each proton and those of the other proton (antiproton). These constituents of the nucleon carry on average one-sixth of the total energy. It is known from neutrino quark scattering experiments that quarks carry only approximately 50% of the total momentum of a nucleon (the first

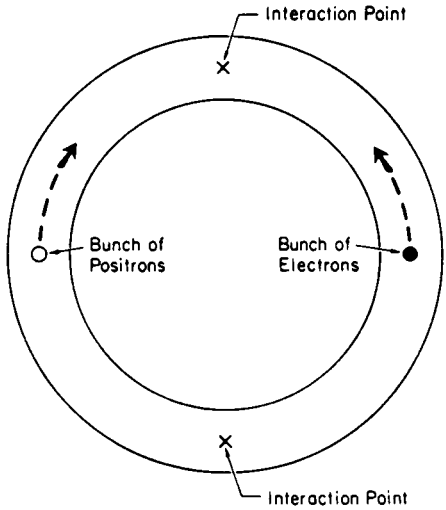


Figure 2.4. Principle of a storage ring (particle collider): two particle bunches travel in opposite directions and are continually brought into collision at specific interaction points.

indirect indication of the existence of gluons). If we assume that three quarks are constituents of the nucleon, then each quark typically carries a fraction $x_p \simeq 1/6$ of the momentum. Because the strong interaction is invariant under charge conjugation, the distribution of antiquarks in the antinucleon is identical to the distribution of quarks in the nucleon. In other words, an antiquark in the $\overline{p}$ also carries only approximately one-sixth of the overall momentum $x_{\overline{p}} \simeq 1/6$. Thus, the centre-of-mass energy in a $q\overline{q}$ collision is given by $\sqrt{s_{q\overline{q}}} = \sqrt{x_p x_{\overline{p}} s} \simeq \frac{1}{6}\sqrt{s}$. This implies that in $p\overline{p}$ and pp collisions a centre-of-mass energy approximately six times greater than the desired rest energy is required.

The 1 TeV energy scale is of interest for the theories discussed in chapter 1, since the top quark and the Higgs particle are expected to have masses under 1 TeV. SUSY particles could also have masses in this area. Thus, the following minimum requirements must be imposed on future accelerators:

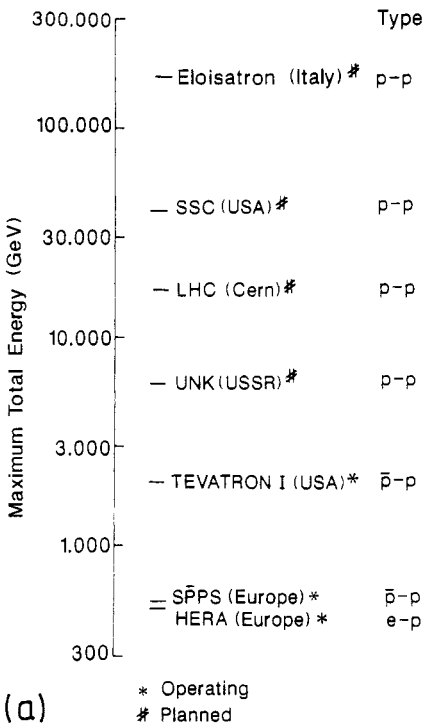
(i) pp or $p\overline{p}$ colliders:

$$W \geq 10 \text{ TeV} \qquad L \geq 10^{32} \text{ cm}^{-2} \text{ s}^{-1} \tag{2.12a}$$

(ii) e^+e^- colliders:

$$W \geq 1 \text{ TeV} \qquad L \geq 10^{32} \text{ cm}^{-2} \text{ s}^{-1}. \tag{2.12b}$$

One disadvantage of storage-ring experiments relates to the reaction rate. In a fixed target the target density and thus also the reaction probability is much



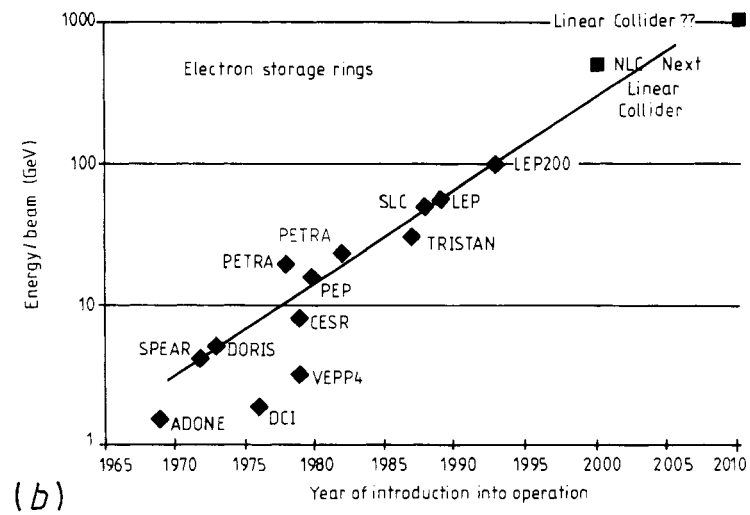


Figure 2.5. Highest energies of (*a*) pp, $p\overline{p}$ and ep storage rings (after [Phy86]) and (*b*) electron–positron storage rings (from [Scho89]).

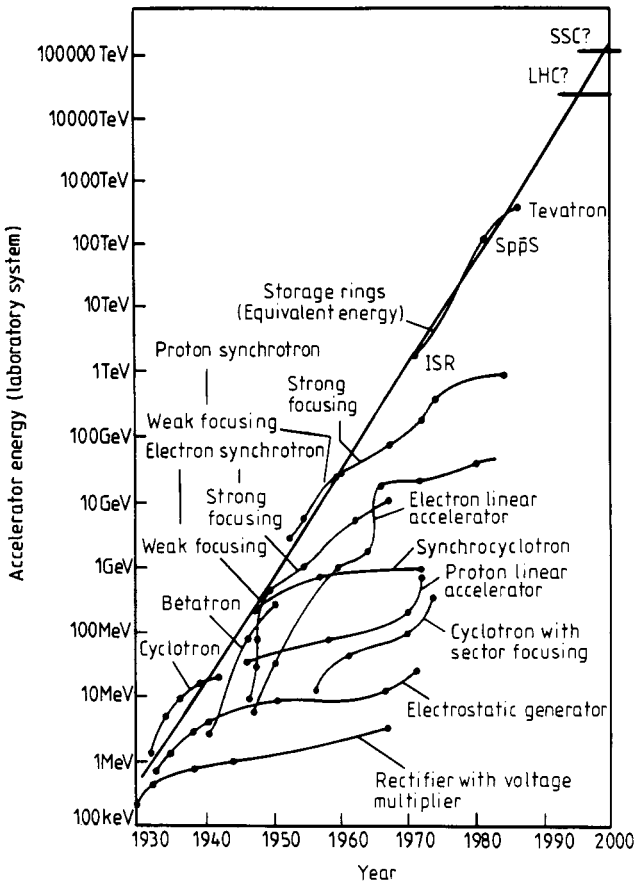


Figure 2.6. Highest energies achieved over the years using different types of accelerators (measured in the laboratory system, log scale). An increase by a factor of 10 approximately every seven years has been achieved, with new accelerator types replacing older ones. For proton storage rings the figure shows the equivalent energies a 'fixed-target' accelerator must have to deliver the same 'usable' energies (from [Scho89]).

higher than in the case of contrarotating particle bunches. The luminosity is a measure of the probability that the particles of the two beams interact with one another. It depends on the particle density and the overlap of the bunches. The relationship between the reaction rate and the luminosity L is given by (2.2). Because of the relatively low density of a beam bunch in comparison with a fixed target, the problem in this accelerator technology is to achieve a

high luminosity. Typical values in modern storage rings are $L \simeq 10^{31}$ cm^{-2} s^{-1} for $p\bar{p}$ and e^+e^- rings.

In fixed-target machines, we have

$$L = \Phi\rho \tag{2.13}$$

where ρ is the particle density in the target and Φ denotes the particle flux in particles per unit time. For contrarotating beams of relativistic particles, L is given by

$$L = f n_B \frac{N_1 N_2}{A}. \tag{2.14}$$

Here, $N_{1/2}$ are the numbers of particles per bunch, n_B denotes the number of bunches per beam, A is the collision cross section and f is the bunch circulation frequency.

Instead of using storage rings, it is possible to induce collisions between beams propagating in opposite, or initially in the same direction (see SLC) in linear accelerators. In this case, the requirements on the particle density in the bunches are still much higher, since, unlike in circular accelerators, the bunches can only interact once.

Storage rings and colliders can only operate with stable charged particles. The following combinations of particles are used (existing or planned installations are given in brackets):

e^+e^- collider (LEP, SLC)
$p\bar{p}$ collider (SPPS, Tevatron I)
pp collider (ISR, LHC)
$e^{\mp}p$ collider (HERA).

It is also possible to induce collisions between heavier ions; however, we shall not discuss the physics of heavy ions here.

Whilst hadron storage rings with beam energies up to approximately 9 TeV are planned (LHC), the technical prerequisites for the implementation of an e^+e^- storage ring with 1 TeV centre-of-mass energy have not yet been met.

Particularly important more recent developments include the completion and bringing into operation of LEP at CERN, the first linear e^+e^- collider SLC (Stanford Linear Collider) and the first e^-p collider HERA [Eis92, Schn94] at DESY.

2.1.1 LEP (Large Electron Positron ring)

LEP is the world's largest e^+e^- storage ring with a circumference of 27 km; it was completed in 1989 (see figures 2.7 and 2.8). The first collisions between electrons and positrons were observed at the beginning of August 1989. Four intensive bunches of approximately 10^{11} electrons and positrons circulate in each

of the opposite directions in a circular vacuum tube ($p \simeq 10^{-10}$ Torr). Along the ring, some 3400 magnets keep the particle beams in their orbit in a relatively weak magnetic field of 0.1 T. More than 1300 quadrupole and sextupole magnets are responsible for the focusing.

LEP is in the first phase of its development and has a maximum energy of 2×50 GeV; it is designed to create large numbers of Z^0 bosons, the properties of which may then be studied. The electron and positron bunches cross at four predetermined points of the ring. A large detector (ALEPH [ALE89], DELPHI [DEL89], L3 [L89] and OPAL [OPA89]) is installed at each of the four interaction points. Some 700 000 Z^0 bosons were detected within the first 12 months of operation, allowing a number of precision experiments.

In the coming years, the LEP energy is to be increased to approximately 2×100 GeV in order to investigate W^+W^- pair creation (LEP 200). Superconducting acceleration cavities are being developed for this purpose.

The study of e^+e^- collisions at centre-of-mass energies above 200 GeV requires the development of new accelerator concepts since, because of the synchrotron emission, the costs of e^+e^- storage rings increase almost quadratically as the energy increases.

One solution is provided by linear colliders, in which the particles are brought to high energies in separate linear accelerators before their collisions are induced. The costs of such linear colliders increase only linearly with the centre-of-mass energy so that this concept represents the more economic solution above 200 GeV.

The first e^+e^- linear collider was built at the Stanford Linear Accelerator Center (SLAC), California.

2.1.2 SLC (Stanford Linear Collider), NLC (Next Linear Collider)

SLC (figure 2.9), like LEP, currently provides a centre-of-mass energy of 100 GeV and is used to study Z^0 physics. In the SLC accelerator, electron and positron bunches of 10^{10} particles are accelerated in a linear accelerator to a final energy of 50 GeV. A magnetic field is used to redirect the bunches into two semi-circular paths at the end of the acceleration path and induce a collision. The interaction area was surrounded by the MARK II detector [MAR89], and is now surrounded by the SLD detector. Whilst approximately 45 000 collisions per second take place at LEP, SLC achieves a repetition rate of 60 per second. Thus, in order to also achieve high Z^0 production rates the particle bunches must be concentrated into very small volumes.

The production rates at SLC are smaller than those at LEP. In the future, one advantage of SLC over LEP is likely to be the greater longitudinal polarization of the beam, which may be relevant for investigation of the helicity structure of the weak interaction.

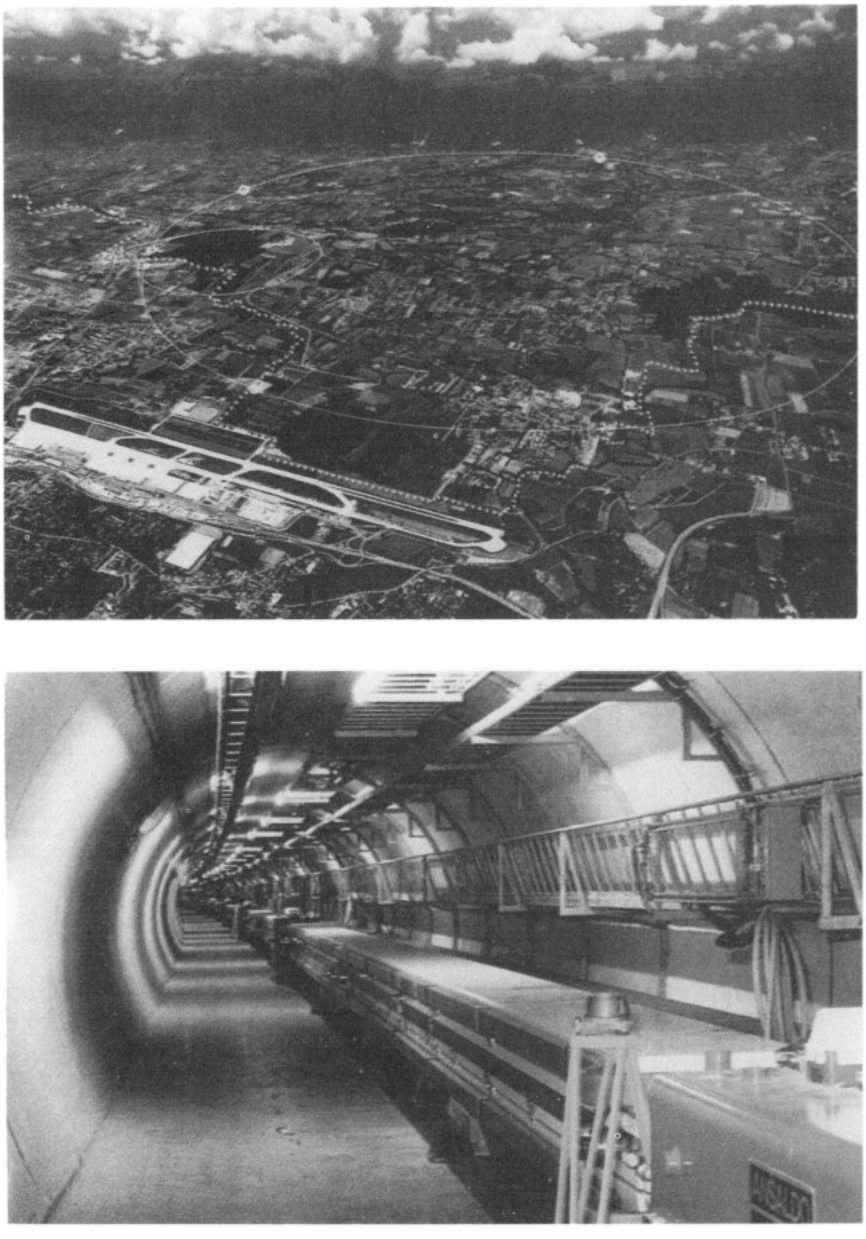


Figure 2.7. Top: aerial photograph of CERN. Geneva airport in the foreground, the French Jura in the background. In between are the LEP tunnel and (left) the SPS. The frontier between Switzerland and France is shown by dots (photo: Swissair). Bottom: view into the LEP tunnel, showing the pairs of dipole magnets (lighter) and the focusing quadrupole magnets. Note the low curvature of the tunnel (photo: CERN) (from [Scho89]).

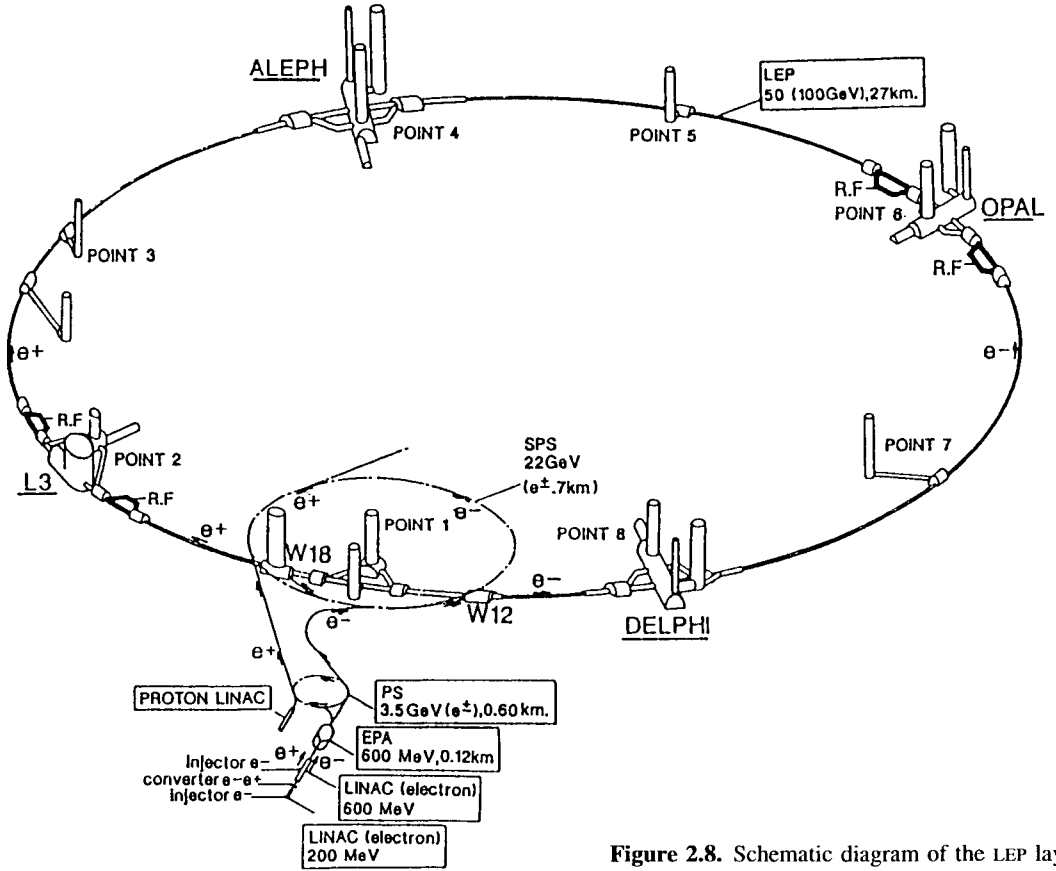


Figure 2.8. Schematic diagram of the LEP layout (from [Scho89]).

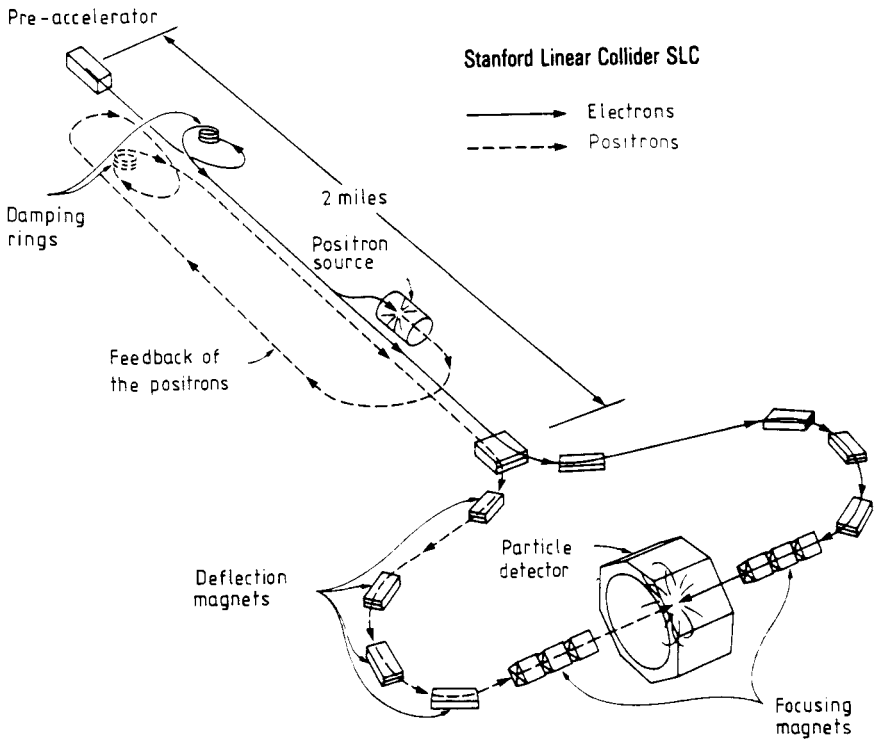


Figure 2.9. Schematic diagram of the SLC layout (Stanford Linear Collider). Electrons and positrons are accelerated shortly behind each other and brought to collision using two half arcs. Particle bunches are initially generated in two damping rings (from [Scho89]).

SLC was the first e^+e^- linear collider to be successfully operated. Based on these experiences, the construction of a new e^+e^- linear collider (NLC = Next Linear Collider) is now under discussion at SLAC. Together with the LHC hadron storage ring (see below) this would permit experiments at 300–500 GeV with very low background, in the search for the Higgs particle and the top quark [Loh92]. Beam energies of 1000–2000 GeV are the long-term objective of such developments in a number of laboratories.

2.1.3 HERA (Hadron–Electron Ring Anlage)

Compared with pure hadron and electron colliders, HERA at DESY in Hamburg (figure 2.10) represents a hybrid form, in which collisions are induced between electrons at 30 GeV and protons at 820 GeV (maximum energies of approximately 600 GeV in the centre-of-mass system). The two HERA rings share a common tunnel of length 6.3 km and diameter approximately 5 m. HERA was

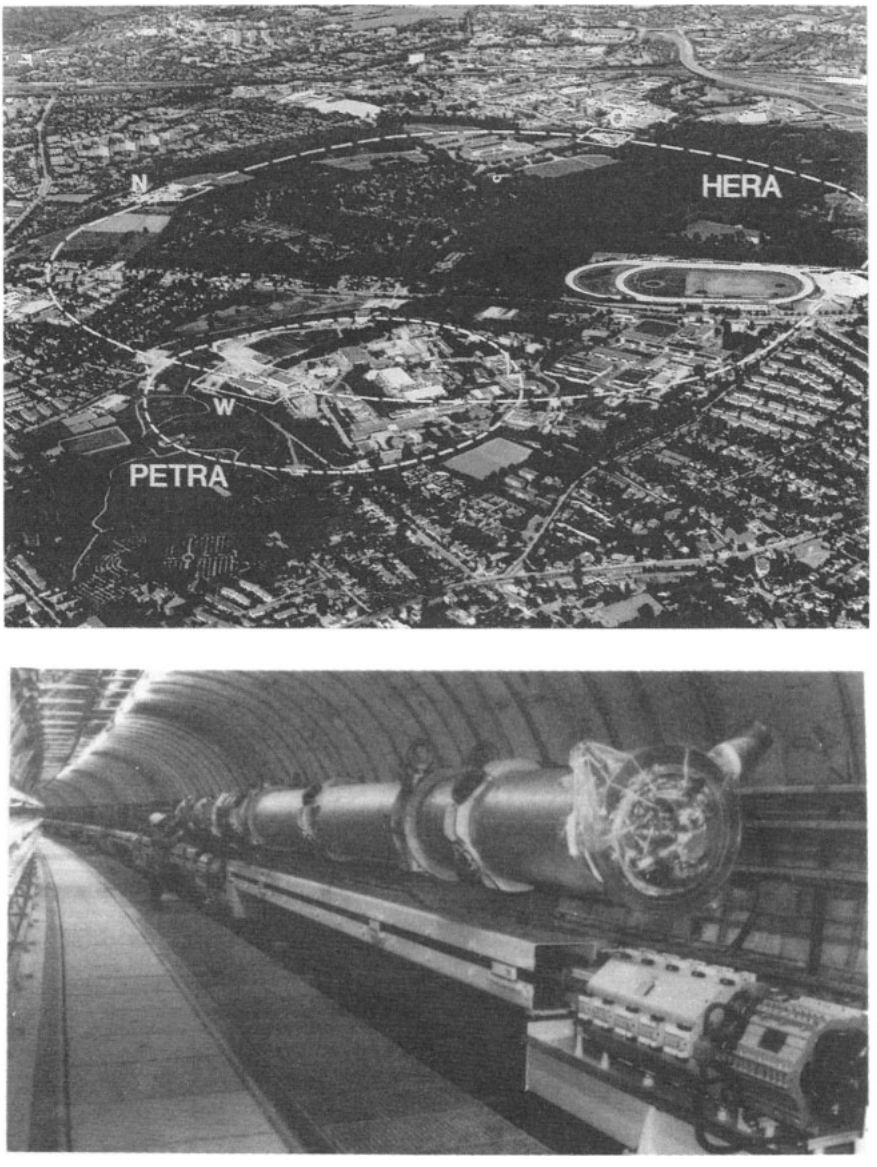


Figure 2.10. Top: aerial photograph of DESY. The PETRA ring is visible and the underground HERA ring is shown schematically. The surface buildings in the four experimental zones (north, south, west and east) can be seen. HERA passes round the Altona public park and the trotting course (photo: DESY) (from [Scho89]). Bottom: view into the HERA tunnel at DESY. The already completed electron ring is shown at the bottom, with a number of the already installed superconducting magnets above it (from [Scho89]).

the first major European installation to use superconductivity and corresponding low-temperature technology on a large scale. Superconducting magnets are used for the proton ring. Superconducting cavities, which will permit an increase in the electron energy in a later phase, are under development for the electron ring [Schm90b, Eis92, Schn94].

2.1.4 LHC (Large Hadron Collider)

Current plans essentially concentrate on the development of one large hadron accelerator, LHC at CERN (see, for example, [Win88, Phy86, Sch89, Rub93]. A further project, SSC in the USA (see below), has meanwhile been cancelled.

The construction of the superconducting proton collider LHC is planned in the existing LEP tunnel (which has the appropriate dimensions, figure 2.11). The existing accelerators at CERN could be used as pre-accelerators. Superconducting dipole magnets may be used to generate magnetic fields of approximately 10 T. However, for this, the magnets must be cooled to 1.8 K using He II. In such a magnetic field, assuming the dimensions of the LEP tunnel, protons may be accelerated to energies of 8–9 TeV and stored and thus a pp centre-of-mass energy of 16–18 TeV may be achieved.

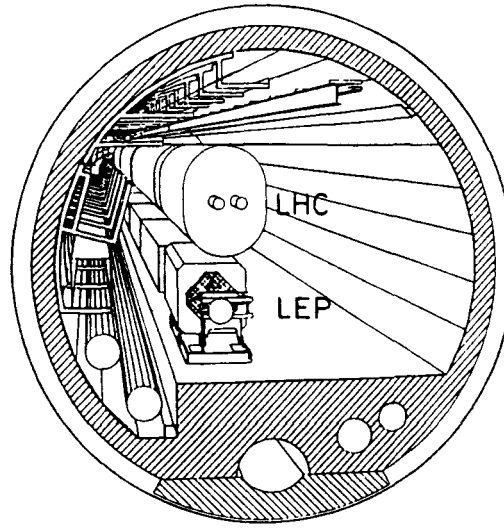


Figure 2.11. Possible layout of a double ring for a proton–proton collider (LHC) above the LEP magnet ring (from [Scho89]).

This would make it possible to study the important energy area of 1 TeV in quark–quark collisions. In addition, a high luminosity of 10^{33}–10^{34} cm^{-2} s^{-1} is sought (for details see [Rub93]).

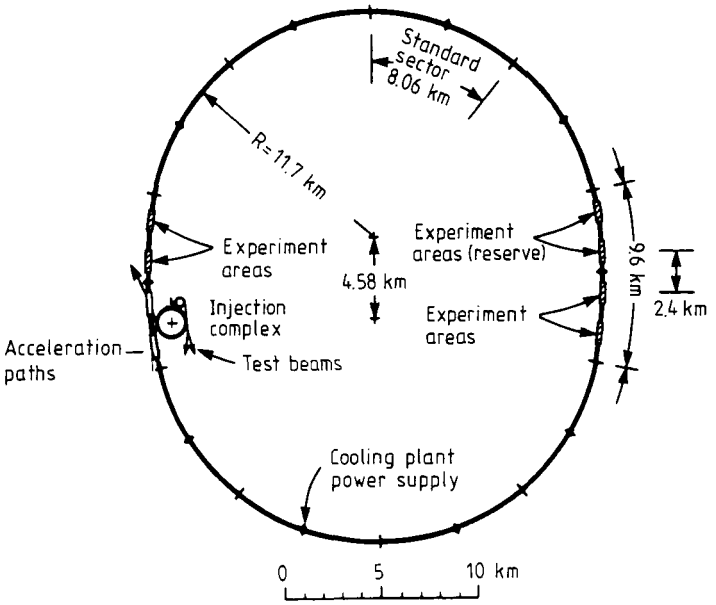


Figure 2.12. Plan of the American SSC (Superconducting Super Collider) with a circumference of around 83 km, which was to be built to the south of Dallas, Texas (from [Scho89]).

2.1.5 SSC (Superconducting Super Collider)

The USA planned to construct the largest pp storage ring to date (figure 2.12). This 83 km long Superconducting Super Collider would have achieved an energy of 2×20 TeV for contrarotating protons; in other words, the maximum total energy should have been 40 TeV. The planned luminosity was 10^{33} cm^{-2} s^{-1} (for details see [Sch93]).

A further project was under way in the former USSR. A 3 TeV proton accelerator (UNK) was planned for use with a fixed target in the first instance and for later extension to a 2×3 TeV pp collider.

2.1.6 Accelerators around the year 2000

Figure 2.13 gives an overview of the world's most important accelerator centres (cf. figure 4.13). In summary, we note that physical research to the end of the 1990s will be supported by three large machines: Tevatron ($p\overline{p}$, beam energy 1000 GeV), LEP (e^+e^-, beam energy 100 GeV) and HERA (e^-p, 820 GeV protons and 30 GeV electrons). We would also mention a number of smaller electron–positron colliders (DORIS at DESY, CESR, SLC) and the LEAR ring at CERN for

antiprotons at very low energies. In addition, many places have plans for a new generation of accelerators, designed to generate extremely high particle fluxes at relatively low energies (K and B meson factories, etc). Further in the future, around the year 2008, an additional large machine, LHC (beam energy 8–9 TeV) will come on stream.

On the other hand, limits of the current technology for constructing accelerators are becoming visible. Because of synchrotron radiation, the size of a circular machine increases approximately with the square of the energy. For example, if it is desired to gain a factor of five in the beam energy in comparison with LEP, a ring approximately 25 times the size of LEP would have to be constructed, corresponding to a circumference of 600–700 km. Thus, the next step after LEP can only be a linear accelerator.

The cross sections for an interaction between quarks and/or leptons decrease as the energy E increases according to the formula

$$\sigma \sim 1/E^2. \tag{2.15}$$

Thus, in order that sufficiently many events may be detected, an increase in the centre-of-mass energy by a factor of 10 requires a simultaneous increase in the luminosity by a factor of 100. This places extreme demands on the beam optics.

2.2 PHYSICS AT THE ACCELERATORS OF THE 1990s AND EARLY 2000s

A major part of our current knowledge of the fundamental interactions and the physics of elementary particles is based on high-energy experiments at large accelerator facilities. It is not our aim to provide an exhaustive description of the successes and milestones of elementary particle physics. Instead, this section is intended as a brief and necessarily incomplete outlook for the possibilities for and limits of physical research on the accelerators of the 1990s and early 2000s.

The minimal standard model with three quark and lepton generations is able to describe a wealth of experimental facts. On the other hand, it raises many questions which require new experiments on high-energy accelerators to confirm the assumptions of the model or to contradict these and thereby discover 'new physics'.

Important open questions, even in the framework of the standard model, relate to the top quark and the Higgs particles which are said to be responsible for the mass of quarks, leptons and the intermediate vector bosons via the Higgs mechanism. Other questions are raised by the various types of Grand Unified Theory (GUT, SUSY, etc) and include, for example, the question of the existence of supersymmetric particles, heavy neutrinos, majorons, etc (see chapters 1, 6, 7 and 9).

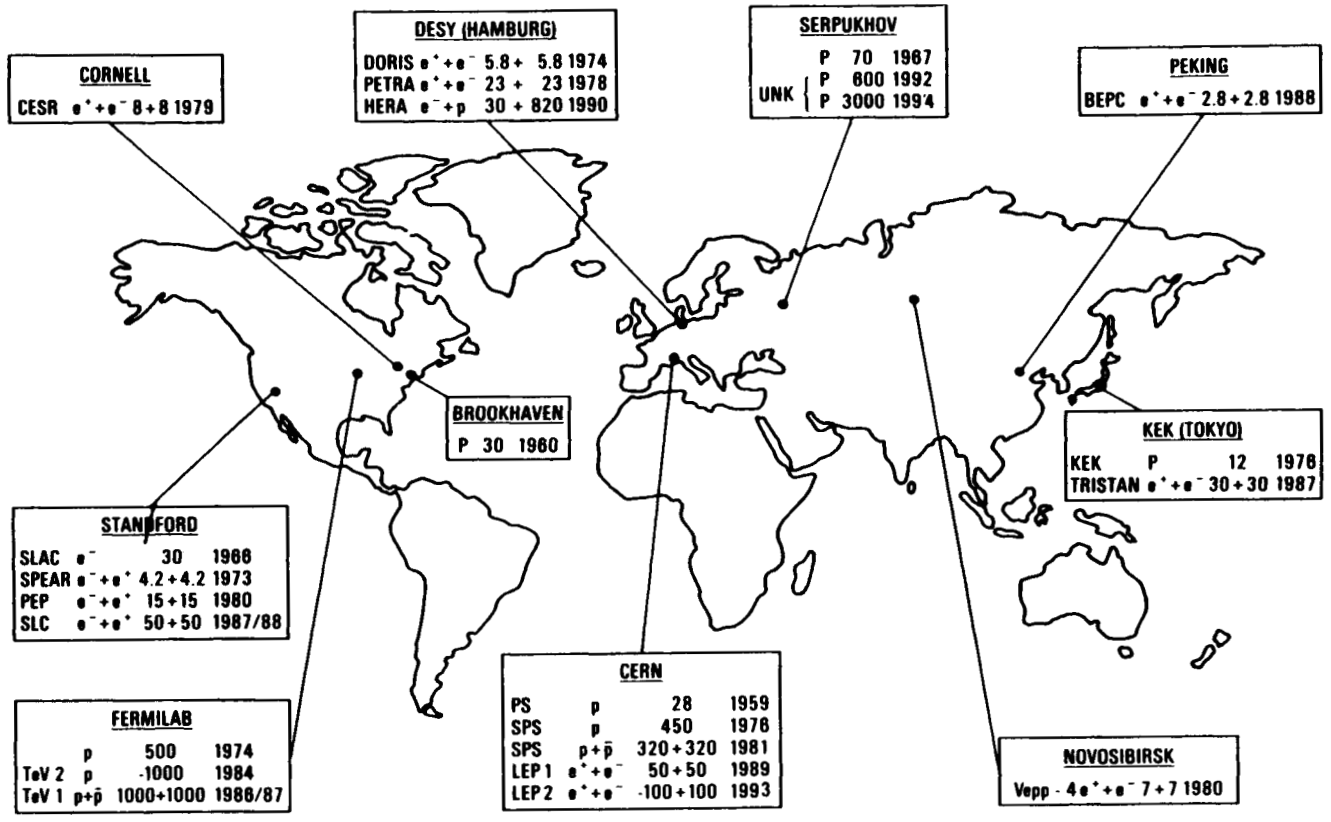


Figure 2.13. Map of the world showing accelerator centres (from [Scho89]).

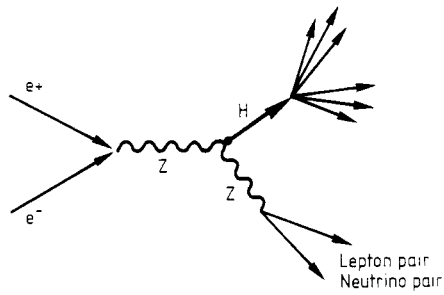


Figure 2.14. Diagram of the generation of Higgs particles as 'bremsstrahlung' from a Z^0, which then decays into a fermion pair.

Thus, there is a clear compulsion to strive for increasingly higher energies. The 100 GeV energy region, defined by the mass of the W^+ and Z^0 bosons, had been predicted from measurements of the mixing parameter $\sin^2\theta_W$ in semi-leptonic neutrino scattering experiments based on the theory of the electroweak interaction.

Higgs masses and masses of supersymmetric particles are expected in the area up to 1 TeV from present particle physics theories. Heavier Higgs particles seem to imply a violation of the unitarity limit at around 1 TeV.

In what follows, we shall summarize some of the current objectives of accelerator high-energy physics. For detailed discussions of the scientific potential of the accelerators discussed in section 2.1 we refer to [Bar87d, Jar90b, Zer92, 93, Buc92] (see also table 1.9 of the previous chapter).

Test of the standard model. Increasingly sensitive tests of the inner consistency and predictive power of the standard model play an important role in high-energy physics (for an overview see [Lan92a, 93, Wil93]). In the past, crucial progress has been achieved by work at LEP and SLC. Since these results are also important in relation to the questions of neutrino physics discussed in later chapters, we shall discuss them in somewhat more detail later (section 2.2.1).

Higgs particles. To date there is no experimental evidence for the Higgs particle, which plays a central role for our understanding of the spontaneous symmetry breaking. LEP, the current Higgs monopole machine, and a future e^+e^- linear collider adding to the LEP energy are the ideal machines for the search for Higgs particles in the energy region to 250 GeV/c^2. Possible reactions are (cf. figures 2.14 and 2.15):

$$e^+e^- \rightarrow Z^0H \tag{2.16}$$

$$e^+e^- \rightarrow \overline{\nu}\nu H \tag{2.17a}$$

$$e^+e^- \rightarrow e^+e^-H. \tag{2.17b}$$

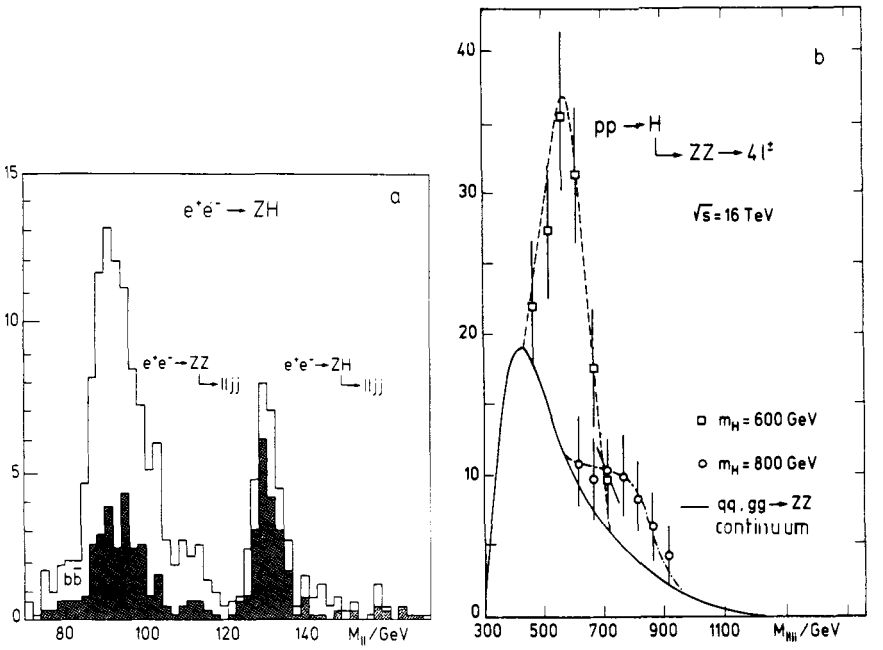


Figure 2.15. Generation of Higgs particles in the intermediate-mass area in e^+e^- linear colliders and in the high-mass area in proton colliders (LHC); simulations from [Jar90b, Zer92] (from [Zer93]).

Higgs particles up to masses of around 1 TeV/c^2 could be detected in new high-energy proton colliders. Here, Z^0 decays of the Higgs particle with subsequent leptonic decays above the threshold $H \rightarrow Z^0Z^0$ permit a particularly efficient search (figure 2.15):

$$pp \rightarrow H \rightarrow Z^0Z^0 + llll. \tag{2.18}$$

For a detailed discussion of possibilities for the Higgs search we refer to [Gun90].

Search for the top quark. LEP measurements give a typical upper bound on the mass of the top quark of

$$m_T < 200 \text{ GeV}/c^2. \tag{2.19}$$

Thus, there is a reasonable hope that the top quark may be detected directly at hadron colliders via the reaction

$$p\bar{p} \rightarrow t\bar{t} + \ldots . \tag{2.20}$$

The centre-of-mass energies of the two most powerful $p\overline{p}$ colliders, the S$p\overline{p}$S at CERN ($\sqrt{s}$ = 630 GeV) and Tevatron I at the Fermilab ($\sqrt{s}$ = 1.8 TeV) already lie in the required region. However, this energy is distributed across the constituents of the colliding protons and antiprotons, so that on average, only approximately one-sixth of the energy is available for particle production. Moreover, the cross section for the creation of a top quark is very small, depending on the mass m_t (see e.g. [Kle88]). For expectations for m_t see section 2.2.1.2. Recently, a group at the Fermilab Tevatron published indications for $m_t = (174 \pm 10^{+13}_{-12})$ GeV/c^2 [Abe94].

Weinberg angle. One other objective is precise measurement of the electroweak mixing angle, the Weinberg angle, $\sin^2\theta_W$, which is a parameter of the standard model, but is predicted by GUT models. Progress is expected both from hadron colliders (from more precise measurements of the mass difference $m_Z - m_W$) and from the extension of the LEP storage ring to energies above the W^+W^- production threshold (LEP II).

Quark mixing. Further precise measurements with heavy-quark flavours are needed to determine the matrix elements of the Kobayashi–Maskawa matrix describing quark mixing.

Search for 'new physics'. In addition to the above work, which partially is already a search for 'new physics', the search for clues to a physics beyond the standard model is a particularly crucial motivation for the construction of new accelerators. By 'new physics' we mean the discovery of new particles, clues to GUT or SUSY structures, to internal structures of quarks and leptons, to the existence of leptoquarks (bound states of quarks and leptons), new vector bosons, supersymmetric particles, etc.

e^+e^- linear colliders are the ideal machines for the search for light sleptons and electroweak gauginos/higgsinos (for whose masses we expect $\sim$ 100–200 GeV/c^2) while proton colliders can generate the colour-carrying SUSY particles (squarks and gluinos with masses in the region of several hundred GeV/c^2) with high rates (figure 2.16) (see [Jar90b, Zer92, Zer93]). Approaches to theoretical descriptions going beyond the standard model (see e.g. [Lan92b, 93, Wil93]) could be based not only on the discovery of new particles (SUSY or Higgs particles, new vector bosons) but also on the study of rare decays, which should be forbidden by conservation principles (e.g. B or L conservation). Possible indications *have* already been found in precision measurements of the strong and the electroweak coupling constants (see section 2.2.1.3).

The accelerator facilities to achieve these goals are either hadron colliders or e^+e^- or ep storage rings or linear colliders, respectively. All accelerators designed to achieve the highest energies currently use particle bunches with the particle combinations described in section 2.1. Hadron accelerators with

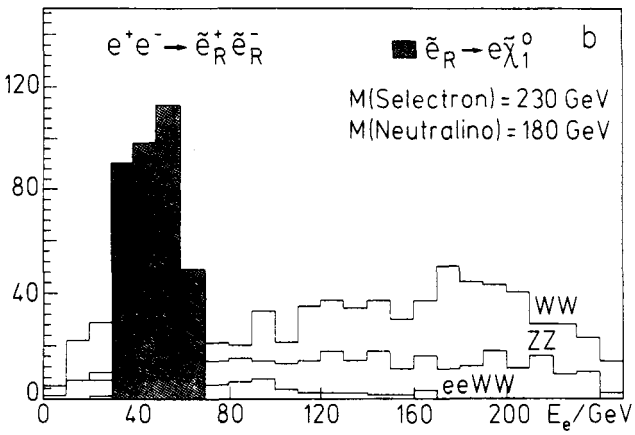


Figure 2.16. Production of supersymmetric particles: selectrons, scalar partners of the electrons, in e^+e^- linear colliders; simulations from [Jar90b, Zer92] (from [Zer93]).

a fixed target still retain their importance as so-called factories for light-quark flavours. This type of accelerator is essentially needed to generate secondary beams of mesons (π, K, ϕ, B), antiprotons and neutrinos. At the current state of accelerator technology, the highest centre-of-mass energies are achieved in hadron storage rings, while the lowest background is achieved in electron machines in which effects of the strong interaction are avoided.

A glance back into history shows that the collider principle has already permitted important discoveries in the past. The proton–proton Intersecting Storage Ring (ISR) at CERN provided the first possibility for investigating pp collisions. Its maximum energy of 63 GeV made it possible to clarify many aspects of the strong interaction going beyond the energies accessible in fixed-target experiments.

The next step, the large proton–antiproton storage ring ($Sp\overline{p}S$) at CERN made it possible to store contrarotating proton and antiproton beams in a single ring. Collisions were observed for centre-of-mass energies up to 600 GeV.

The intermediate $W^{\pm}$ and Z^0 vector bosons were successfully created and detected at the CERN $Sp\overline{p}S$ through $p\overline{p}$ collisions at a centre-of-mass energy of 540 GeV. The fundamental reaction for the creation of W bosons in hadronic collisions is

$$u\overline{d} \rightarrow W^+ \qquad \overline{u}d \rightarrow W^- \tag{2.21}$$

where the quarks and antiquarks are the constituents of the contrarotating protons and antiprotons. Observation of the decays

$$W^- \rightarrow e^-\overline{\nu}_e \qquad W^+ \rightarrow e^+\nu_e$$

$$\rightarrow \mu^- \overline{\nu}_\mu \qquad \rightarrow \mu^+ \nu_\mu \tag{2.22}$$

led to successful *detection* of W^+ bosons [Arn83, Bag83, Ban83], the mass of which corresponded to the theoretically predicted value. In addition, the Z^0 boson was detected by observation of the decay

$$Z^0 \rightarrow e^+e^-,\ \mu^+\mu^-. \tag{2.23}$$

The generating reaction in the $p\overline{p}$ collision is

$$u\overline{u} \rightarrow Z^0 \qquad d\overline{d} \rightarrow Z^0. \tag{2.24}$$

This led to the spectacular confirmation of the theory of the electroweak interaction at the S$p\overline{p}$S collider at CERN by the detection of the exchange particles mediating the weak force.

This success ultimately led to the construction of the world's most powerful (in terms of centre-of-mass energy) hadron storage ring, the Tevatron at the Fermilab in Chicago. This is a 2×900 GeV $p\overline{p}$ collider with superconducting deflection magnets.

The large hadron ring currently planned, namely LHC at CERN, was discussed in the previous section. The achievable centre-of-mass energy is approximately 16–18 TeV.

In addition to the pure hadron colliders and to future electron linear colliders with energies beyond LEP, important information is expected from the hybrid electron–proton collider HERA at DESY in Hamburg. It is hoped that experiments on deep inelastic ep scattering will provide information about the internal structure, and particularly also the spin structure, of the nucleon. Together with other tests of the standard model, HERA is expected to lead to the discovery of new particles and perhaps phenomena which go beyond the standard model. These include the search for substructures of quarks and leptons, for leptoquarks [Ahm94], heavy neutrinos and heavy electrons. The latter three could be produced at HERA with masses up to 250 GeV/c^2 with a sensitivity far beyond that obtainable at other present-day accelerators [Buc92, Eis92].

The development of e^+e^- storage rings has proved particularly fruitful. In addition to the fact that all the collision energy after the annihilation reaction is available for the creation of new particles, e^+e^- rings have the major advantage over hadron machines of a small background and a simple initial state.

Particle storage rings were first developed for electrons. Electrons were first stored in 1960 in Frascati (Italy) in the ADA storage ring. The first e^+e^- collisions using a single magnetic ring were observed in 1964 in Orsay (France) with ACC and in 1969 in Frascati with ADONE.

Experiments at e^+e^- colliders have led, amongst other things, to the discovery of the τ lepton, mesons with charm (D mesons), the B meson, the

J/ψ particle and of hadron jets. The quarkonium states $J/\psi(c\bar{c})$ and $\Upsilon(b\bar{b})$ have been very thoroughly investigated. Because of the analogy with the H atom, the spectroscopy of the energy states of these resonances is particularly simple.

The e^+e^- storage rings SPEAR in Stanford and CESR at Cornell University played a major role in the advances in particle physics in the 1970s and 1980s. The SPEAR ring provided important data about ψ resonances, mesons with charm and the τ lepton.

The energy at the CESR ring permitted investigations of Υ resonances, B mesons and the properties of b quarks. Additional contributions to the physics of Υ resonances were provided by DORIS at DESY in Hamburg and the VEPP4 storage ring in Novosibirsk.

The e^+e^- storage rings PEP (Stanford) and PETRA (DESY, Hamburg) achieved energies of 20–30 GeV and 20–46 GeV, respectively. One important objective was to investigate the properties of c and b quarks. Successes using PETRA included the discovery of the carrier of the colour interaction (the gluon) and the observation of oscillations between neutral B mesons and their antiparticles (analogous to $K^0\overline{K^0}$ oscillations).

The next crucial step in particle physics was the introduction into operation of the largest e^+e^- storage ring ever built, LEP at CERN (1989) and the upgrading of the first e^+e^- linear collider SLC (Stanford Linear Collider) at the Stanford Linear Accelerator Center (SLAC) to energies of 50 GeV (1988). Since the recent results obtained using these two installations are very important in the next chapters, we shall discuss them in some detail in the following section.

2.2.1 Tests of the standard model by LEP and SLC

LEP, with a maximum energy of approximately 2×50 GeV in the first phase of its development, was designed to generate Z^0 bosons in large numbers to permit the study of the properties of these. Within the first 12 months of operation some 700 000 Z^0 bosons were detected, so that precise measurements were possible (see e.g. [Ste91] for a review). Very precise measurements have also been carried out at SLC in the area of the Z^0 resonance, even though the production rates are lower than for LEP (see section 2.1).

2.2.1.1 Z^0 physics, number of neutrino flavours, majorons

The Z^0 particles created in e^+e^- collisions

$$e^+e^- \rightarrow Z^0 \tag{2.25}$$

decay within 10^{-25} s, and thus cannot be directly observed but must instead be identified via the decay. Since the Z^0 with a rest mass of ~ 91 GeV/c^2 is

very heavy, the decay products have a high energy, so that large detectors are required to measure the energy. For example, the ALEPH detector weighs an impressive 3000 tonnes. The four detectors mentioned above (section 2.1) have a similar basic concept, although in each case particular value is placed on the detection of special particle types or decay channels.

Since the Z^0 is an important component of the standard model, determination of the properties of this boson is an important task. The cross section for the creation of a Z^0 particle in an e^+e^- annihilation has a maximum when the centre-of-mass energy corresponds exactly to the mass m_{Z^0} (see figure 2.17). As an extremely heavy particle, the Z^0 may decay into all lighter known, or possibly as yet undiscovered, lepton–antilepton pairs or quark–antiquark pairs. The Z^0 decay channels are

$$Z^0 \rightarrow e^+e^-,\ \mu^+\mu^-,\ \tau^+\tau^- \tag{2.26}$$

$$\rightarrow \nu_i\overline{\nu}_i \qquad i = e, \mu, \tau \tag{2.27}$$

$$\rightarrow q\overline{q} \qquad q = s, u, d, c, b$$

However, the quarks in (2.27) do not occur as free particles, but as jets of pions, kaons and other hadrons.

The reaction cross section for the reaction (2.27) may be determined from the number of jet-like events as a function of the collision energy. A typical resonance curve is obtained. The position of the maximum corresponds to the mass m_{Z^0}. According to the uncertainty principle, the width Γ of the curve is associated with the lifetime τ by

$$\tau = \frac{h}{\Gamma}. \tag{2.28}$$

The combined results from the four LEP experiments give a Z^0 mass [Bob91, Sch91, PDG92] (see also table 2.1)

$$m_{Z^0} = (91.174 \pm 0.005 \pm 0.020)\ \text{GeV}/c^2 \tag{2.29}$$

and a width

$$\Gamma = (2487 \pm 10)\ \text{MeV}/c^2. \tag{2.30}$$

The first error in m_{Z^0} is the experimental error, the second is a systematic error due to the energy calibration of the LEP beams.

The total width Γ is the sum of the contributions from the individual decay channels, i.e. the hadronic width Γ_h, the leptonic width $\Gamma_{lept} = \Gamma_{ee} + \Gamma_{\mu\mu} + \Gamma_{\tau\tau}$ and an invisible width Γ_{inv}. We have

$$\Gamma = \Gamma_h + \Gamma_{lept} + \Gamma_{inv}. \tag{2.31}$$

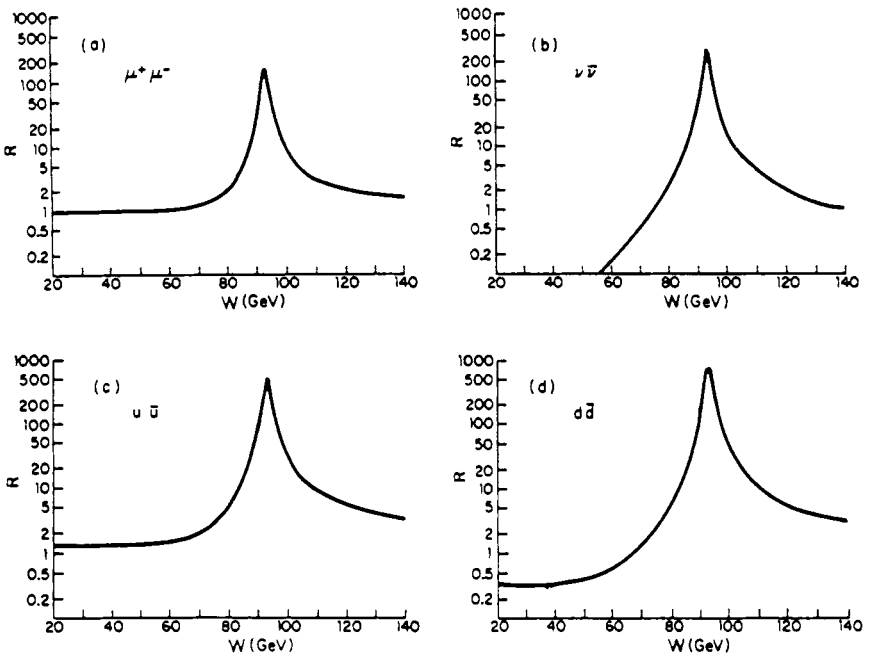


Figure 2.17. Dependence of the cross sections for the reactions $e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to f\overline{f}$ on the centre-of-mass energy W in the GWS model. The figure shows the energy dependence of the ratio $R \equiv \sigma(e^+e^- \to f\overline{f})/\sigma_{\mathrm{QED}}(e^+e^- \to \mu^+\mu^-)$ in the standard model for three generations of quarks and leptons. In (*c*) and (*d*), $u\overline{u}$ and $d\overline{d}$ stand for final states with quarks with charge 2/3 and $-1/3$, respectively. The resonance structure is due to the propagator of the Z^0 boson (from [Qui83]).

At this point, we recall briefly how the width Γ may be derived from a measurement of the cross section. The following formulae are given without proof in natural units ($\hbar = c = 1$). For their derivation, we refer the reader to, for example, [Nac86].

Let us consider the production of a Z^0 boson in the e^+e^- annihilation, with subsequent decay into an arbitrary final state

$$e^+e^- \to Z^0 \to f\overline{f}. \tag{2.32}$$

The scattering cross section for the reaction (2.32) is [Nac86]

$$\sigma_{f\overline{f}} = 12\pi \frac{\Gamma_{ee}\Gamma_{f\overline{f}}}{(s - m_{Z^0}^2)^2 + m_{Z^0}^2\Gamma^2} \tag{2.33}$$

where Γ_{ee} is the partial decay width for the channel $Z^0 \to e^+e^-$ and Γ is the total

Table 2.1. Data about the Z^0 boson obtained from four different LEP experiments; m_{Z^0} in GeV/c^2, decay widths in MeV (from [Scho91]).

	ALEPH	DELPHI	L3	OPAL
m_{Z^0}	91.182 ± 0.009	91.175 ± 0.010	91.181 ± 0.010	91.160 ± 0.009
Γ	2488 ± 17	2454 ± 20	2501 ± 17	2497 ± 17
Γ_{ee}	84.2 ± 0.9	81.6 ± 1.3	83.3 ± 1.0	83.5 ± 1.0
$\Gamma_{\mu\mu}$	80.9 ± 1.4	88.4 ± 2.4	84.5 ± 2.0	83.5 ± 1.5
$\Gamma_{\tau\tau}$	82.9 ± 1.6	84.9 ± 2.7	84.0 ± 2.7	83.1 ± 1.9
Γ_{lept}	83.3 ± 0.7	83.4 ± 1.0	83.6 ± 0.8	83.4 ± 0.7
Γ_{h}	1756 ± 15	1718 ± 22	1742 ± 19	1747 ± 19
Γ_{inv}	481 ± 14	486 ± 21	511 ± 18	499 ± 17
N_ν	2.90 ± 0.08	2.93 ± 0.13	3.05 ± 0.10	2.99 ± 0.10

	Average value
m_{Z^0}	$91.174 \pm 0.005 \pm 0.02$ (LEP)
Γ	2487 ± 9
Γ_{ee}	83.3 ± 0.5
$\Gamma_{\mu\mu}$	83.3 ± 0.9
$\Gamma_{\tau\tau}$	83.3 ± 1.0
Γ_{lept}	83.3 ± 0.4
Γ_{h}	1744 ± 10
Γ_{inv}	493 ± 9.5
N_ν	2.96 ± 0.06

decay width. s is the square of the centre-of-mass energy ($\sqrt{s} = E_{e^-} + E_{e^+}$). It is clear from this that the cross section $\sigma_{f\bar{f}}$ and Γ are interrelated. It follows from (2.33) that the production cross section for the Z^0 in the e^+e^- annihilation is

$$\sigma(e^+e^- \to Z^0) = 12\pi \frac{\Gamma_{ee}\Gamma}{(s - m_{Z^0}^2)^2 + m_{Z^0}^2\Gamma^2} \tag{2.34}$$

with a maximum value

$$\sigma_{\max} = 12\pi \frac{\Gamma_{ee}}{m_{Z^0}^2\Gamma} \tag{2.35}$$

for $\sqrt{s} = m_{Z^0}$. A typical value for $\sigma_{\max}$ is around 5×10^{-32} cm^2.

Using (2.33), for example, for the aforementioned hadronic channel we

have

$$\sigma_h = 12\pi \frac{\Gamma_{ee}\Gamma_h}{(s - m_{Z^0}^2)^2 + m_{Z^0}^2\Gamma^2}$$
$$= \sigma_0^h \frac{\Gamma^2 m_{Z^0}^2}{(s - m_{Z^0}^2)^2 + m_{Z^0}^2\Gamma^2}. \tag{2.36}$$

Here σ_0^h denotes the hadronic cross section at the maximum

$$\sigma_0^h = 12\pi \frac{\Gamma_{ee}\Gamma_h}{m_{Z^0}^2\Gamma^2}. \tag{2.37}$$

From (2.36) we see that the hadronic cross section may be described by a Lorentz curve with three parameters m_{Z^0}, Γ and σ_0^h. The three parameters may be determined by fitting the Lorentz curve to the measurement points.

The individual decay channels and their widths Γ_i were successfully measured at LEP by identifying the particles arising in the decay. The further breakdown of Γ_h into the components $\Gamma_{q\bar{q}}$ is highly complicated, so that in general only Γ_h is given. The results of the four experiments are summarized in table 2.1. The agreement between the individual measurement results is very good. The partial widths agree with the predictions of the standard model within the error limits. Moreover, the e–μ–τ universality is clearly fulfilled; in other words, the weak interaction does not distinguish between the different lepton types. Thus, we may write

$$\Gamma_{ee} = \Gamma_{\mu\mu} = \Gamma_{\tau\tau} = \Gamma_{\text{lept}}^* \tag{2.38}$$

and

$$\Gamma_{\text{lept}} = 3\Gamma_{\text{lept}}^*. \tag{2.39}$$

Measurements of the resonance curve (2.36) also provide fundamental evidence about the number of (light) neutrino types, when one compares the overall width with the predictions of the standard model.

The decay width Γ depends on the number and type of the possible decay channels. The more decay channels are open, the shorter the lifetime will be. From (2.28), this leads to a correspondingly greater decay width.

Through identification of the end products at LEP, it has been possible to measure the partial widths Γ_h and Γ_{lept} as well as Γ. The invisible width Γ_{inv} may be deduced directly from these results

$$\Gamma_{\text{inv}} = \Gamma - 3\Gamma_{\text{lept}}^* - \Gamma_h = \Gamma - \Gamma_{\text{lept}} - \Gamma_h. \tag{2.40}$$

In the framework of the standard model a contribution to the Z^0 width of

$$\Gamma_\nu \simeq 166\ \text{MeV} \tag{2.41}$$

is calculated for massless or light neutrinos which couple universally to the Z^0

boson. If Γ_{inv} is attributed solely to decay channels with $\nu_i\bar{\nu}_i$ pair formation, the number of neutrino types is then given by

$$N_\nu = \frac{\Gamma_{\text{inv}}}{\Gamma_\nu}. \tag{2.42}$$

N_ν may be measured with even greater sensitivity using the maximum of the hadronic cross section. Figure 2.18 shows the cross section for the creation of Z^0 particles in the e^+e^- annihilation with a subsequent hadronic decay channel for different numbers of neutrino flavours, $N_\nu = 2, 3, 4$.

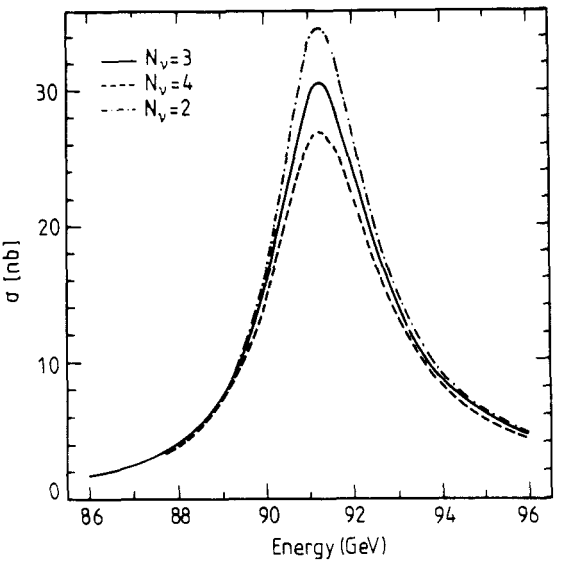


Figure 2.18. Cross section for the generation of the Z^0 boson from e^+e^- reactions and the decay into hadrons as a function of the collision energy (data from OPAL). The curves were calculated for different numbers of neutrino species (from [Scho91]).

The maximum cross section σ_0^{h} is very sensitive to N_ν. According to

$$\sigma_0^{\text{h}} = 12\pi \frac{\Gamma_{ee}\Gamma_{\text{h}}}{\Gamma^2 m_{Z^0}^2} \tag{2.43}$$

σ_0^{h} is inversely proportional to Γ^2. The existence of another ν flavour would lead to an increase of 7% in the total width; in other words, the maximum would decrease by 14%. Figure 2.18 shows that only $N_\nu = 3$ is consistent with the measurement. The LEP result for the number of light neutrinos is (see e.g. [PDG94])

$$N_\nu = 2.983 \pm 0.025 \quad \text{or} \quad N_\nu = 2.97 \pm 0.17. \tag{2.44}$$

Thus, there are three types of *light* neutrino, namely ν_e, ν_μ and ν_τ. The existence of a fourth type of neutrino with mass less than 45 GeV/c^2 can be ruled out. The restriction $m_\nu c^2 < 45$ GeV follows from the fact that only neutrinos with masses less than half the rest mass of the Z^0 can be created in the Z^0 decay

$$Z^0 \rightarrow \nu\overline{\nu}. \tag{2.45}$$

This result is of fundamental importance for the following chapters. According to our current knowledge, a charged lepton is assigned to each neutrino; in other words, we also expect only three lepton families. Moreover, from the symmetry between leptons and quarks, we deduce that there are also only three quark families. The number of neutrino flavours N_ν plays an important role, in particular in the discussion of the dark matter in the universe and the primordial nucleosynthesis (see chapters 3 and 9).

Here, we also note that, at LEP, N_ν is not necessarily an integer. In the present interpretation, the number of families N_{fam} is given by

$$N_{\text{fam}} = N_\nu = \frac{\Gamma_{\text{inv}}}{\Gamma_\nu}. \tag{2.46}$$

In theories which go beyond the standard model, the quantity

$$\tilde{N} = \frac{\Gamma_{\text{inv}}}{\Gamma_\nu} \tag{2.47}$$

cannot be equated automatically with N_ν or N_{fam}. In fact, if right-handed neutrinos (which are actually required by left–right symmetric models) are included in the standard model, it follows that [Jar90a]

$$\tilde{N} \leq N_{\text{fam}}. \tag{2.48}$$

$\tilde{N}$ can only be interpreted as the number of families if one assumes that the minimal standard model is valid. On the other hand, it follows from this discussion that a measured value of $\tilde{N} > 3$, i.e.

$$\tilde{N} = 3 + \varepsilon \qquad \varepsilon > 0 \tag{2.49}$$

excludes all models with three generations and arbitrarily many right-handed neutrinos. The measured value (2.44) is consistent with the assumption of only three left-handed neutrino types.

The LEP measurements of Z^0 also provide important information about the majoron (see chapters 1 and 6), the massless Nambu–Goldstone boson, which would result from a spontaneous breaking of a global $(B - L)$ symmetry.

From the various older type majoron models, triplet and doublet majorons are excluded by the measured Z^0 width. In fact, the contribution of the triplet majoron to the Z^0 width should correspond to that of two neutrino flavours [Geo81] and that of the doublet majoron to half a neutrino width. On the other hand, a singlet majoron [Chi80], which has recently attracted increasing interest in connection with a neutrino-mass hierarchy including a 17 keV neutrino [Gla91], or a mixture of singlet and doublet majorons, remain possible, and could occur in double-beta decay (see [Moh91a, Ber92b] and chapter 6).

2.2.1.2 Weinberg angle and top quark

LEP permits precise measurement of the Weinberg angle. A comparison with other data for $\sin^2\theta_W$ provides indirect upper bounds for m_t by virtue of the dependence of the radiative corrections on the mass of the top quark m_t.

The Weinberg angle links the electrical and weak charges e and g with one another. We have (see section 1.4.4)

$$\sin^2\theta_W = \left(\frac{e}{g}\right)^2 = 1 - \frac{m_W^2}{m_{Z^0}^2} \tag{2.50}$$

where m_{Z^0} is now known with great precision (see (2.29)), while estimates of m_W using colliders are associated with considerably greater errors. In addition to (2.50), we have the following relationship between m_{Z^0}, the fine structure constant α and the Fermi constant G_F (see (1.118))

$$m_{Z^0}^2 = \frac{\pi\alpha}{\sqrt{2}G_F}\frac{1}{\sin^2\theta_W\cos^2\theta_W}. \tag{2.51}$$

Since α and G_F are known very precisely from optical measurements and muon decay (respectively), it is possible to deduce $\sin^2\theta_W$ and ultimately to determine the mass m_W using (2.50).

Equations (2.50) and (2.51) are, strictly speaking, only valid if we neglect radiative corrections; in other words, they are only valid for Feynman diagrams of the lowest order (simple W, Z^0 or γ exchange). However, it is well known from quantum electrodynamics (QED) that higher-order effects such as the exchange of additional virtual field quanta or vacuum polarization play an important role. Examples of this include the Lamb shift in the atomic shell or the anomalous value of $g - 2$ for electrons and muons.

Radiative corrections may *in principle* be calculated without problems. In addition to a correction factor due to the strong interaction the terms due to the weak interaction are of particular interest. However, calculation of these requires a knowledge of all the masses, including those of the top quark (m_t) and the Higgs particle (m_{Higgs}).

Such radiative corrections are often taken account of by the introduction of 'running coupling constants' (see section 1.5.2), i.e. effective coupling constants which depend on the interaction energy. In this way, the fundamental equations between the coupling constants and the observable quantities retain their form.

The quantities measured at LEP, namely m_{Z^0}, Γ, Γ_{lept} and the angular distributions of the negatively charged leptons relative to the direction of incidence of the electron (forwards–backwards asymmetry), may each be used to derive a relationship between the Weinberg angle and the radiative corrections. On the other hand, radiative corrections can be calculated as a function of m_t and m_{Higgs} (see e.g. [Alt90]).

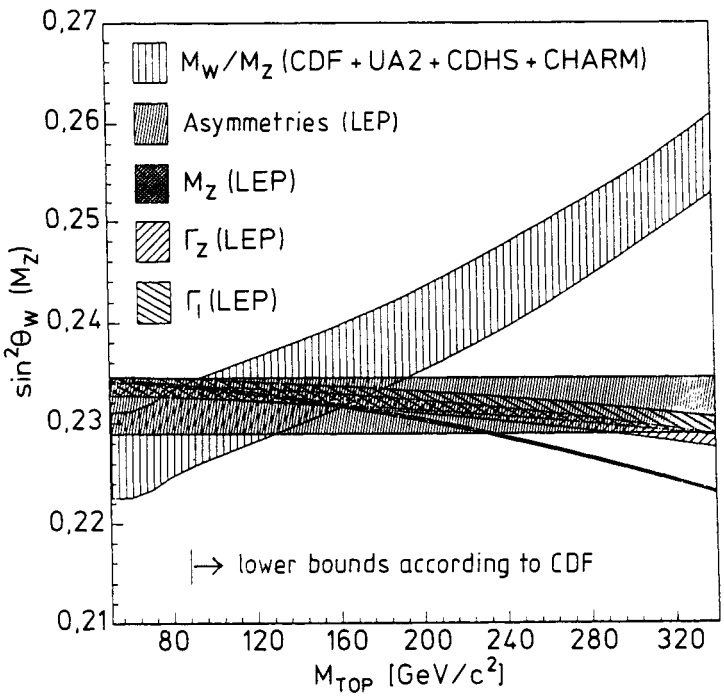


Figure 2.19. The mixing angle $\sin^2\theta_W$ as a function of the top mass m_t for a fixed Higgs mass $m_{\text{Higgs}} = 200$ GeV/c^2, derived from various measured quantities: CDF and UA2 $p\overline{p}$ annihilation, CDHS and CHARM neutrino scattering, LEP results averaged over four experiments (from [Scho91]).

Since m_{Higgs} only plays a minor role, $\sin^2\theta_W$ may be plotted as a function of m_t (see figure 2.19). Here, we note that at LEP, $\sin^2\theta_W$ was determined at an energy of 90 GeV. Taking into account a correction term Δr for the radiative

corrections [Sit80, Mar80], equation (2.51) becomes

$$\frac{G_F(1-\Delta r)m_{Z^0}^2}{8\sqrt{2}\pi\alpha} = \frac{1}{16\sin^2\theta_W\cos^2\theta_W} \tag{2.52}$$

where m_{Z^0}, α and G_F are known with great precision. Thus, we have a functional relationship between $\sin^2\theta_W$ and Δr.

The ratio m_W/m_{Z^0} is also known from experiments on $p\overline{p}$ colliders (CDF, UA2) [Abe90, Ali90] and from neutrino scattering experiments (CDHS, CHARM) [Abr86, All86, CHA89]. A combined analysis yields [Ama91]

$$\sin^2\theta_W = 1 - \left(\frac{m_W}{m_{Z^0}}\right)^2 = 0.2290 \pm 0.0035. \tag{2.53}$$

Using this result and m_{Z^0}, the radiative correction Δr and the mass m_t may be derived from (2.52). Figure 2.19, in which $\sin^2\theta_W$ is plotted against the top mass, illustrates the determination of m_t. The dependence of the radiative corrections on m_t is greater in $p\overline{p}$ and neutrino scattering experiments than in LEP experiments. Comparison of the self-explanatory data in figure 2.19 shows that a consistent picture is only obtained for a specific range of m_t.

The following limits for m_t were determined from the overlap area [Ama91]

$$m_t c^2 = \begin{cases} (116 \pm 38)\,\text{GeV} & m_{\text{Higgs}}c^2 = 45\,\text{GeV} \\ (144 \pm 37)\,\text{GeV} & m_{\text{Higgs}}c^2 = 1000\,\text{GeV}. \end{cases} \tag{2.54a}$$

Higgs masses less than 45 GeV/c^2 are ruled out by the latest LEP measurements (see e.g. [Akr91]). For the Weinberg angle this yields [Ama91, Scho91]

$$\sin^2\theta_W = \begin{cases} 0.2340 \pm 0.0014 & m_{\text{Higgs}}c^2 = 1\,\text{TeV} \\ 0.2331 \pm 0.0014 & m_{\text{Higgs}}c^2 = 45\,\text{GeV} \end{cases} . \tag{2.54b}$$

The most recent values are [PDG94]

$$m_t c^2 = (169^{+16+17}_{-18-20}\,\text{GeV}. \tag{2.55a}$$

and

$$\sin^2\theta_W(m_Z) = 0.2319 \pm 0.0005 \pm 0.0002 \tag{2.55b}$$

For a detailed actual discussion we refer to [Lan93, PDG94].

From (2.54*a*, 2.55*a*), we no longer need to rely on a blind search for the top quark. Instead, we know the range of masses in which the sixth quark is likely to be found. The recent indications on $m_t c^2 = (174 \pm 10^{+13}_{-12})$ GeV found at the $p\overline{p}$ Tevatron collider at the Fermilab [Abe94, Aba95] indeed lie in this range.

2.2.1.3 *Indications of supersymmetry: Weinberg angle, strong and electroweak coupling constants*

We would immediately add a note about the predictions of GUT models to the remarks about the Weinberg angle in the previous section. The new precise experimental results for $\sin^2\theta_W$ from LEP data and measurements of semileptonic neutrino reactions, like the lifetime of the nucleon, contradict the predictions of the minimal SU(5) model, which predicts the following value for $\sin^2\theta_W$ [Lan81]

$$\sin^2\theta_W = 0.214 \pm 0.004 \qquad (\mathrm{SU}(5)). \tag{2.56}$$

On the other hand, predictions based on minimal *supersymmetric* GUTs are *consistent* with the new LEP data within 2% [Ell90b] (see also [Cos88, Lan93]).

LEP may not only be used for accurate study of the electroweak interaction, but also provides important information about the strong interaction. Quantum chromodynamics (QCD) is a non-Abelian gauge theory. The exchange particles (gluons) themselves carry colour charges and therefore interact among themselves. Consequently, the coupling constant α_s decreases logarithmically as the interaction energy increases ('asymptotic freedom'). In order to verify these ideas about the colour force, it is therefore very important to measure α_s and its energy dependence as accurately as possible.

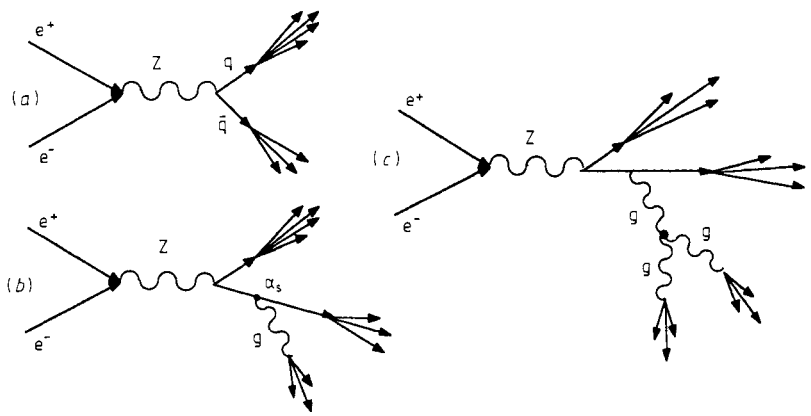


Figure 2.20. Diagrams for the decay of the Z^0: (*a*) decay into a quark–antiquark pair, which fragments into jets; (*b*) gluon bremsstrahlung with a three-jet fragmentation; (*c*) triple gluon vertex, whose fragmentation generates a four-jet event (from [Scho91]).

The coupling constant α_s is determined, for example, via the gluon bremsstrahlung (see figure 2.20(*b*)). The decay of a Z^0 particle may lead to the creation of a quark–antiquark pair (figure 2.20(*a*)). Such an event may be recognized from two hadron jets in opposite directions. By analogy with

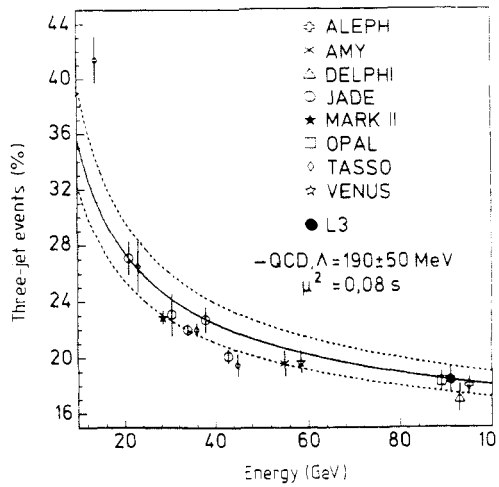


Figure 2.21. Percentage of the three-jet events as a function of the e^-e^- collision energy. This percentage is proportional to α_s (from [Scho91]).

bremsstrahlung, a quark may emit a gluon which also appears as a jet, so that we have a three-jet event (figure 2.20(*b*)). The probability of such an event depends on the coupling strength α_s.

The analysis yields [PDG94]

$$\alpha_s(m_{Z^0}) = 0.120 \pm 0.007 \pm 0.002. \tag{2.57}$$

This value is consistent with measurements of deep inelastic nucleon scattering.

Comparison of (2.57) with measurements at low energies clearly shows the dependence of α_s on energy (see figure 2.21) [L90], which agrees with that expected from QCD.

Another indication that the colour interaction is described by a non-Abelian gauge theory, follows from the existence of the triple gluon vertex (figure 2.20(*c*)), the fragmentation of which leads to a four-jet event (see, for example, [DEL91]).

Detailed analysis of LEP data has shown that a grand unification in the framework of the minimal *non*-supersymmetric standard model of particle physics is ruled out [Ama91, Ell90b]. On the other hand, a supersymmetric extension of the standard model appears to permit a unified description of the strong, weak and electromagnetic forces.

According to the renormalization group equation (1.128) the fact that different theories involve a different number of particles leads to different variations of the coupling constants as a function of the interaction energy.

The renormalization group equation may be used to extrapolate the functional relationships determined at low energies to higher energies.

If one extrapolates the very precise values for $\sin^2\theta_W$ and α_s obtained at LEP to higher energies, under the assumption of a *non*-supersymmetric standard model (one Higgs doublet), the three curves do *not* meet at one point [Ama91] (see figure 2.22). In addition to the discrepancy relating to the lifetime of the proton (see chapter 4), this is a further indication that the formulation of a GUT theory must go beyond the 'minimal' framework.

Although the α_i do not meet at a point in a normal GUT theory, SUSY–GUT models appear to permit a unification.

As previously mentioned, one consequence of supersymmetry is that there exists a fermionic partner for every boson, and conversely, but the masses of the new particles are undetermined. The search for supersymmetric particles at colliders has to date been unsuccessful for energies up to the 100 GeV area. On the other hand, if a reasonable high-energy behaviour of the theory is to be ensured (i.e. no violation of the unitarity limit), some of these masses should not be greatly in excess of 1 TeV.

The influence of the supersymmetric partner on the energy dependence of the coupling constants should be noticeable when the energy is greater than the rest energy of these new elementary particles. Thus, attempts have been made to adapt the experimental data assuming a common point of intersection, where the transition energy E_{SUSY} ($\approx$ average mass of the SUSY particles) and the unification energy E_{GUT} are chosen as free parameters. Analysis of the LEP data has given [Ama91] (see also [Ell90b])

$$E_{SUSY} = 10^{3.0\pm1.0}\ \text{GeV} \tag{2.58a}$$

$$E_{GUT} = 10^{16\pm0.3}\ \text{GeV} \tag{2.58b}$$

$$\alpha_{GUT}^{-1} = 25.7 \pm 1.7 \tag{2.58c}$$

where α_{GUT} denotes the coupling constant of the unified force. The result is shown in figure 2.23, from which it is clear that the gradients of the three lines change at around 1 TeV and there is a single point of intersection at around 10^{16} GeV. Above this energy the strong and the electroweak interaction become a unified force.

Based on the threshold behaviour, the mass of the X and Y bosons is given by

$$M_{X,Y} \simeq 0.3 M_{GUT} \simeq 3 \times 10^{15}\,\text{GeV}/c^2. \tag{2.59}$$

From equation (4.14), the lifetime of the proton τ_p may be estimated as [Ama91]

$$\tau_p \simeq \frac{m_X^2}{\alpha_{GUT}^2 m_p^5 c^2} \simeq 10^{33.2\pm1.2}\ \text{years}. \tag{2.60}$$

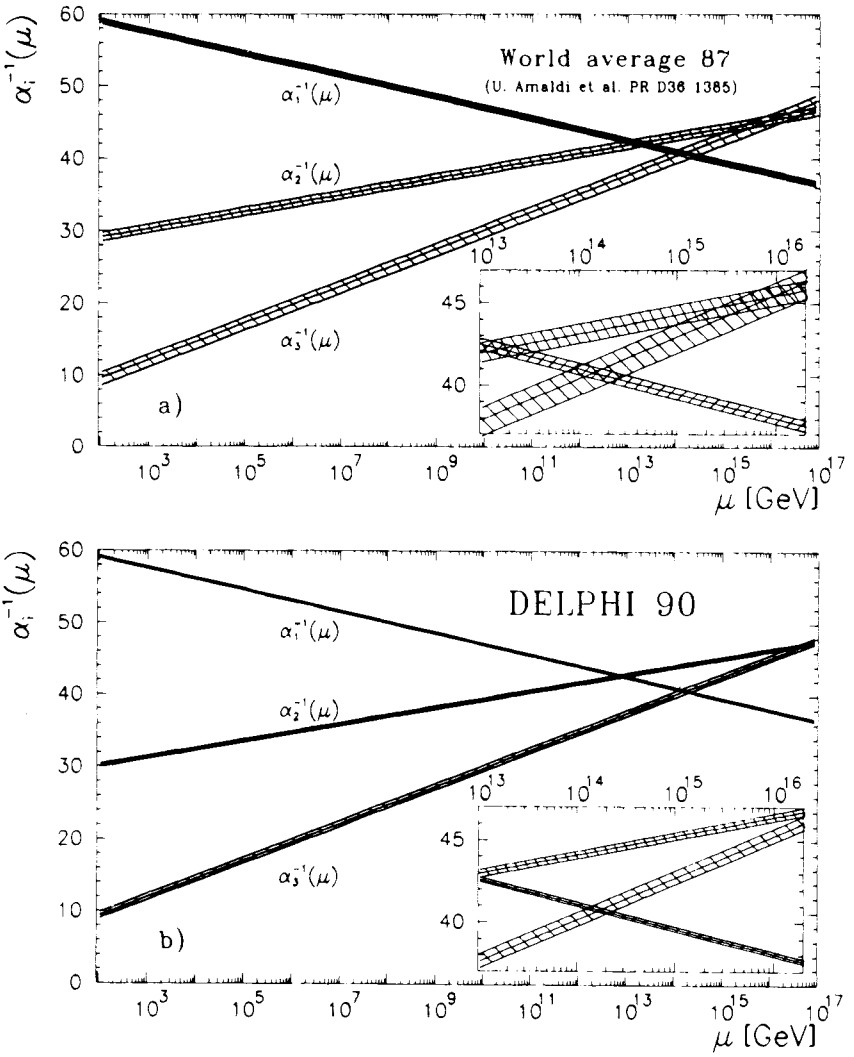


Figure 2.22. Extrapolation of the three coupling constants ($\alpha_1 = e^2/4\pi$, $\alpha_2 = g^2/4\pi$, $\alpha_3 = g_s^2/4\pi$) in the minimal standard model (average values 1987). The lower figure is based on the LEP data from DELPHI. The three coupling constants exhibit a deviation of more than seven standard deviations from a single unification point (from [Ama91]).

This result is consistent with the experimental lower bound for τ_p of 5.9×10^{32} years (see (4.26)), while, because of the relatively low GUT scale of 10^{15} GeV, the minimal SU(5) model has difficulty in explaining the stability of the proton.

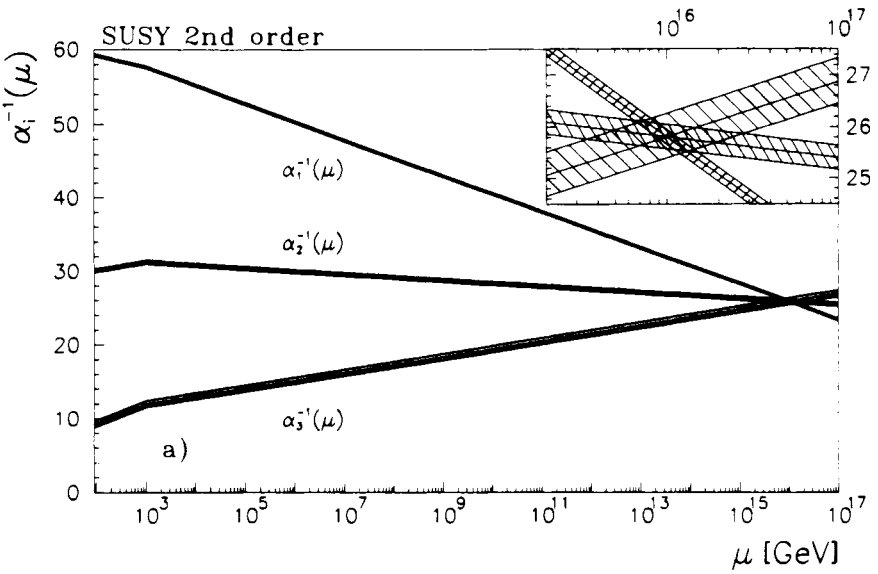


Figure 2.23. Evolution of the coupling constants in the minimal SUSY model. M_{SUSY} is fitted to a single point by the requirement that the three coupling constants should intersect.

However, we note that these results do not prove that the SUSY model is valid. Nonetheless, the consistency with the experimental data may be taken as an indication of a new physics beyond the standard model.

2.3 OUTLOOK: ACCELERATOR AND NON-ACCELERATOR EXPERIMENTS

We have tried in this chapter to give some insight into the possibilities of particle physics at accelerators. The presentation is not claimed to be unified or complete. Many important and interesting aspects could not be touched on. It is interesting that new phenomena and new particles may be expected in the energy area from 100 GeV to a few TeV (figure 2.24).

As far as the LEP experiments are concerned, we would add that, despite intense efforts, no new particles have been detected. In the framework of the minimal standard model, the existence of a Higgs boson can be ruled out at energies below 48 GeV/c^2 [Ade91c], so that

$$m_{\text{Higgs}} > 48\,\text{GeV}/c^2. \tag{2.61}$$

Similar limits for SUSY particles have been determined [PDG94].

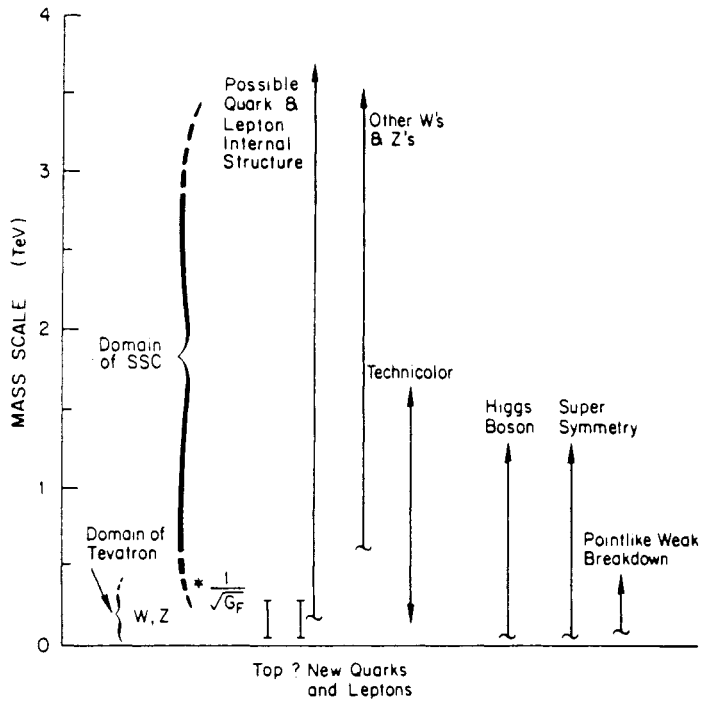


Figure 2.24. Mass scales at which the occurrence of new phenomena can be expected, which will be made accessible to research by the new colliders (from [Phy86]).

The energies of LHC will be sufficient to generate SUSY particles with masses up to the TeV area, including a heavy Higgs boson of the SUSY standard model [Ell91a]. Future e^+e^- colliders are ideal machines for searching for Higgs particles in the range up to 250 GeV/c^2. HERA will allow us to look for substructures of quarks and leptons, of leptoquarks, and for heavy electrons and neutrinos with masses up to 250 GeV, with outstanding sensitivity.

That there must be a physics beyond the standard model appears unavoidable, in view of the lepton–quark mass spectrum and the mixing behaviour of the quarks and in view of precision measurements of the Weinberg angle and of the energy dependence of the coupling constants at LEP. The mass problem for the leptons and the quarks is probably closely related to the physics beyond the standard model (see e.g. [Fri92]). The energy scale for the new physics, that is, for the break-up of the standard model or simple GUTs, cannot be predicted. Some arguments point to the 1 TeV energy scale, which would be reachable experimentally with LHC.

However, many conclusions as far as GUTs are concerned are based on

extrapolation further on across many orders of magnitudes in energy, which for the foreseeable future will not be accessible by accelerators. Thus, we require additional means of accessing the extreme energies of the GUT models. We shall present a number of prominent examples of how this is already possible with current techniques in non-accelerator experiments (see section 1.7) in the following chapters.

The two areas of particle physics, with and without accelerators, complement one another and together provide the tesserae which will be pieced together to form the new physical picture of the world.

Chapter 3

The Early Universe and Particle Physics

Elementary particle physics is concerned with the smallest building blocks of matter and their fundamental interactions, while cosmology deals with the development of the universe on very large space–time scales. However, since the advancement of the idea of grand unification, these two very different areas of physics have grown increasingly closer together (see e.g. [Oli85]). One important reason for this is that in the early stage of the development of the universe extremely large energies were available, at which the GUT symmetries could have been realized. Primordial nucleosynthesis, i.e. the generation of light nuclei during a relatively early phase in the formation of the universe, also represents a successful field for non-accelerator particle physics.

We shall now outline the cosmological standard model and indicate its close relationship to important open questions of particle physics, ranging from the structure of GUTs, through conservation theorems for B and CP to the neutrino mass and the quark–gluon plasma, to name but a few. We shall then discuss the early nucleosynthesis and the number of neutrino flavours derived from this.

3.1 THE COSMOLOGICAL STANDARD MODEL

Current ideas about the development of the universe are summarized in the big bang model (see e.g. [Wei72, 77, Bör88, Gro89, 90, 92, Kol90, Oli91, Cop95]). According to this model, the universe emerged from a singular configuration of space and time and developed from this initial state with an extremely high energy density in an explosive-like expansion. Despite this rapid expansion, thermal equilibrium may be assumed at any time, provided the particle interaction rates Γ_i are substantially greater than the rate of expansion of the universe H

$$\Gamma_i \gg H(t). \tag{3.1}$$

The expansion rate $H(t)$, which is also referred to as the Hubble constant, gives the relative change in size of the universe per unit time

$$H(t) = \frac{\dot{R}(t)}{R(t)}. \tag{3.2}$$

The scale factor $R(t)$ is a measure of the extent of the universe ('radius').

The cosmological standard model is essentially based on three experimental findings, namely: the recession of the galaxies measured from the red shift of spectral lines, the isotropic cosmic background radiation ('3 K radiation') and the abundances of light elements (in particular, of helium) in space.

Red shift. Until the 1920s, many cosmologists believed that the universe was stationary. It was in order to reproduce this that Einstein first introduced the famous cosmological constant Λ in his field equations. In 1929, the American astronomer E P Hubble discovered a linear relationship between the red shift $\Delta\lambda/\lambda_0$ of spectral lines in the spectra of extragalactic star systems and their distance away, s. The observed red shift was interpreted as a Doppler effect resulting from an expansion movement and was the first evidence for the big bang model. For small velocities v, we have

$$v = \frac{\Delta\lambda}{\lambda_0}c = Hs \tag{3.3}$$

where λ_0 is the wavelength of the spectral transition in the laboratory system and $\Delta\lambda$ is the shift in the wavelength ($\Delta\lambda = \lambda - \lambda_0$). The Hubble constant H now has a value of [Vau82, 86, Huc92]

$$H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \qquad \text{where } h = 0.4 \text{ to } 1. \tag{3.4}$$

The problem in determining H is that of independent measurement of the distance of galaxies, which is associated with large errors.

Cosmic background radiation. The second important discovery was the observation of an isotropic background microwave radiation ('3 K radiation') by A Penzias and R W Wilson in 1965 [Pen65]. This cosmic background radiation had already been predicted by G Gamow in 1948 and R Dicke in 1964 and represents important confirmation of the big bang theory. Assuming a big bang, after approximately 10^5 years the hot plasma should have become so optically thin that radiation and matter could become decoupled. This means that the temperatures had fallen to such an extent that electrons and nuclei were able to recombine into atoms. The photons would have too little energy to be absorbed. The radiation from the early phase of the cosmos should still be present today with an energy density corresponding to a temperature of 2.7 K. This was the microwave radiation detected by Penzias and Wilson. It is highly isotropic

and has a photon number density of approximately 400 cm^{-3}. Measurements of the cosmic background radiation by the COBE satellite (Cosmic Background Explorer) gave an almost perfect black-body spectrum with temperature 2.7 K [Mat90a]. The first extremely small anisotropies in the 3 K radiation have only recently been discovered [Ben92a, Sil92, Smo92, Wri92]. This indicates that at the time of the decoupling of the γ radiation from matter (i.e. at the time atoms were formed, approximately 10^5 years after the big bang) the universe was practically in a state of thermodynamic equilibrium. The observed tiny anisotropies may on the other hand be understood as a signature of the large-scale structure of the universe observed today (see e.g. [Schr85a, Sil93]).

The primordial nucleosynthesis. The third pillar supporting the big bang hypothesis, namely the primordial nucleosynthesis, will be discussed in detail in section 3.3.

In what follows, we shall describe the cosmological standard model in more detail. Until the time at which electrons and nuclei recombined into atoms the universe was dominated by radiation (energy density > density of matter). For a Euclidean universe this means that $t \propto T^{-2}$, i.e. the early period may be equally well discussed in terms of energy or time scales.

When the universe was born, thermal energies of 10^{14-16} GeV were present for a short time (up to $\approx 10^{-35}$ s). This explains the close relationship between cosmology and GUT theories. The standard model and the grand unification theories first enabled us to trace the development of the universe back to very early times and extremely high temperatures.

In current theories, the backwards extrapolation of the development of the cosmos comes to an end at the Planck time

$$t_{\mathrm{Pl}} = \sqrt{\frac{\hbar G}{c^5}} = 5.4 \times 10^{-44}\ \mathrm{s} \tag{3.5}$$

when the temperature corresponded to the Planck mass[1] $m_{\mathrm{Pl}} = 1.2 \times 10^{19}\ \mathrm{GeV}/c^2$ (see sections 1.5.5, 1.5.6).

At that time, the diameter of the universe observable today was only a few microns. In order to describe events prior to this time we require a quantum theory of gravitation, which is not yet available. One problem is that the space–time structure must have been subject to fluctuations so that concepts such as future, past and causality lose their meaning (see e.g. [Mis73]). Thus, it is possible that the state of the universe before t_{Pl} can no longer be described as a sequence of events in time. Currently, the state of the universe at time t_{Pl} must be assumed as an initial condition.

[1] The Planck mass is a characteristic mass of gravitation in that, expressed in units of m_{Pl}, Newton's gravitational constant is equal to one. For energies and momentum transfers of the order of m_{Pl} the gravitation dominates all other interactions (see e.g. [Oku87]).

The energy density at the Planck time (see [Lan79c])

$$\rho \simeq (kT)^4 \simeq (m_{\mathrm{Pl}}c^2)^4 \tag{3.6}$$

was so large that, according to GUT models, no symmetry breaking had yet occurred. The particles were all massless, or, at least, the masses at this early time could be neglected. Thus, we may assume an ideal gas of massless particles in thermodynamic equilibrium. Here, all the particle degrees of freedom were equally occupied. Up to a spin statistic factor the densities of the numbers of the various particle types were all the same. The subsequent development was then determined by the interplay between radiation pressure and gravitation.

In the standard model of cosmology it is assumed that the universe is isotropic and homogeneous in the space coordinates. The metric describing such a universe in the general theory of relativity is called the Robertson–Walker metric. In it, an infinitesimal line element ds of a space–time curve is given by

$$ds^2 = c^2dt^2 - R^2(t)\left(\frac{dr^2}{1-kr^2} + r^2(d\Theta^2 + \sin^2\Theta\, d\Phi^2)\right). \tag{3.7}$$

Here, r, Θ, and Φ are the polar coordinates of a point in space on the space–time curve. For a suitable choice of coordinates one can arrange for the metric parameter k to be either ± 1 or 0. These discrete values do not change with time.

We have

$$k = \begin{cases} +1 & \text{spherical metric} \\ 0 & \text{Euclidean metric} \\ -1 & \text{hyperbolic metric.} \end{cases} \tag{3.8}$$

The dynamic behaviour is completely contained in the scale factor $R(t)$ which describes the spatial separation between two neighbouring points of space with constant coordinates (r_1, Θ_1, Φ_1) and (r_2, Θ_2, Φ_2). The rate of expansion of the universe is defined using this parameter (see (3.2)). In the case of a spherical metric, $R(t)$ can be interpreted as the 'radius' of the universe.

$R(t)$ satisfies the Einstein–Friedmann–Lemaitre equations

$$\left(\frac{\dot{R}(t)}{R(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{R^2(t)} + \frac{1}{3}\Lambda c^2 \tag{3.9a}$$

and

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right) + \frac{1}{3}\Lambda c^2 \tag{3.9b}$$

where $p(t)$ denotes the pressure and $\rho(t)$ the density. Λ is the cosmological constant which is interpreted in modern quantum field theories as the energy

Table 3.1. GUT cosmology (from [Gro89, 90]).

	Time t [s]	Energy $E = kT$ [GeV]	Temp. T [K]	'Diameter' of Universe R [cm]
Planck-time t_{Pl}	10^{-44}	10^{19}	10^{32}	10^{-3}
GUT SU(5) breaking, m_X	10^{-36}	10^{15}	10^{28}	10
$SU(2)_L \otimes U(1)$ breaking, m_W	10^{-10}	10^{2}	10^{15}	10^{14}
Quark confinement, $p\overline{p}$-annihilation	10^{-6}	1	10^{13}	10^{16}
ν decouple, e^+e^--annihilation	1	10^{-3}	10^{10}	10^{19}
Formation of light nuclei	10^{2}	10^{-4}	10^{9}	10^{20}
γ decouple, transition from radiation-dominated universe to matter-dominated, formation of atoms, formation of stars and galaxies	10^{12} ($\approx 10^{5}$ a)	10^{-9}	10^{4}	10^{25}
Today, t_0	$\approx 5 \times 10^{17}$ ($\approx 2 \times 10^{10}$ a)	3×10^{-13}	3	10^{28}

density of the vacuum ρ_V [McC51, Zel68]

$$\Lambda = \frac{8\pi G \rho_V}{c^2}. \tag{3.10}$$

Λ is neglected in many discussions, although a rigorous derivation of Einstein's field equations shows that these *must* contain a Λ term [Lov72, Wei72] (for a discussion of the Λ problem see sections 3.2 and 9.2.1 and [Gro89, 90, Kol90, Wei89]).

The development of the universe may be roughly divided into two phases. In the early phase, when particle masses may be neglected, particles with and without mass have the same state equation. This is the so-called 'radiation-dominated' universe. The later phase in the development of the cosmos, after the thermal energy had become very much lower ($\geq 10^6$ years), was determined by cold, pressure-free matter ('matter-dominated phase') (see e.g. [Bör88, Gro89, 90, Kol90]). The most important stages in the development of the cosmos are described in table 3.1.

At the Planck time t_{Pl}, the temperature was approximately 10^{32} K. This corresponds to an average particle energy

$$E \simeq kT \tag{3.11}$$

of 10^{19} GeV. The original force had just divided into gravitation and a GUT interaction. This GUT interaction may, for example, be represented by an SU(5) (or higher) symmetry group. After approximately 10^{-36} s, the thermal energies reached a value of approximately 10^{15} GeV at which the GUT symmetry was broken. Through this symmetry breaking the X and Y bosons, which are responsible for the conversion from quarks to leptons, obtained their mass and froze out of the thermodynamic equilibrium; in other words, the existing X and Y bosons decayed, e.g. via

$$X \to u + u \quad X \to \bar{d} + e^+ \quad Y \to \bar{d} + \bar{\nu}_e \quad Y \to \bar{u} + e^+ \qquad (3.12a)$$

$$\bar{X} \to \bar{u} + \bar{u} \quad \bar{X} \to d + e^- \quad \bar{Y} \to d + \nu_e \quad \bar{Y} \to u + e^- \qquad (3.12b)$$

into quarks and leptons. In the period between $t_X \approx 10^{-36}$ s and $t_W \approx 10^{-10}$ s after the big bang, the universe consisted of a soup of massless leptons, quarks, W and Z bosons and photons, together with other hypothetical particles such as Higgs bosons etc. After 10^{-10} s, the universe had already grown to a size of 10^{14} cm and cooled to an energy of around 100 GeV, so that spontaneous breaking of the $SU(2)_L \otimes U(1)$ symmetry occurred. As a result of this, the massless particles mentioned above obtained their effective masses and the heavy W and Z bosons froze out of the thermodynamic equilibrium.

After approximately 10^{-6} s, the quark–gluon plasma which was present until then could no longer be supported by the thermal energy which was then $kT \approx 1$ GeV. A phase transition from a quark–gluon plasma to the quark confinement took place, i.e. mesons and baryons were formed. Attempts are currently being made to verify the existence of such a quark–gluon plasma using ultra-relativistic heavy-ion collisions in accelerator experiments. At this transition point, which is characterized by the scale factor Λ_{QCD} (see (1.131a), (4.15), (4.18)) in the renormalization group equation ($\Lambda \approx 300$ MeV), the cosmological axions also obtained their mass. Since subsequently no more hadrons could be created, the hadrons were annihilated with their antiparticles into photons. However, a small asymmetry between particles and antiparticles meant that the annihilation could not be total. The small amount of surplus matter forms the matter which is present in the universe today.

Comparison of the number of annihilation quanta N_γ which today form the 3 K background radiation with the number of baryons present in today's universe N_B shows that only a minute fraction of $\approx 10^{-9}$ of the nucleons originally present now remains.

$$N_B/N_\gamma \simeq 10^{-9\pm1}. \qquad (3.13)$$

No satisfactory quantitative explanation of the baryon asymmetry has yet been given. It is generally assumed that at least three conditions must be satisfied for such an asymmetry to occur if one starts from an initially symmetric state

and if it has not been present from the beginning ([Sac67a, b], see also [Wei79, Dol81, Kol90]):

(i) Both C and CP violation in one of the fundamental interactions (see chapter 1).
(ii) The baryon number B is not a conserved quantity.
(iii) Thermodynamic *non*-equilibrium.

It was recently shown, however, that *none* of these three conditions is really *necessary* for the baryogenesis and the formation of a matter–antimatter asymmetry (see [Dol92]). We note that the asymmetry in baryon number may be related to a Majorana mass of the neutrinos. It can be shown that a baryon–antibaryon asymmetry created in the early universe should be washed out by the so-called sphaleron effect [Kuz85] just below the electroweak phase transition, if there exist interactions of Majorana type, as is assumed for the explanation of small neutrino masses by the see-saw model (see sections 1.6.3, 1.6.4.2). The observed baryon asymmetry yields an upper bound for such interactions and thus a limit for Majorana masses of neutrinos of $m_\nu \leq 50$ keV/c^2 down to 1 eV/c^2 [Fuk90, Cam91, Gel92].

At a typical average energy of 1 MeV ($t \simeq 1$ s), the particle density had finally decreased to such an extent that the expansion rate H was greater than the interaction rate for neutrinos, so that the latter also became decoupled from the thermodynamic equilibrium. Apart from a possible neutrino decay, their numbers have scarcely altered since then. By analogy with the electromagnetic 3 K background radiation there should also be a neutrino background radiation with a temperature of 1.9 K (for the case of zero neutrino mass; for $m_\nu = 30$ eV, 0.005 K). Because of the very small interaction cross section it has not yet been possible to detect these background neutrinos (but see [Tup87]).

The contribution of the neutrinos to the density of matter in the universe is given by [Gel88, Kol90]

$$\frac{1}{2}\sum_i g_{\nu_i} m_{\nu_i} = 97 \text{ eV}(\Omega_\nu h^2) \tag{3.14}$$

where the sum is over all stable light neutrino types ($m_\nu < 1$ MeV), h determines the Hubble constant (see (3.4)), Ω_ν is the contribution of the neutrinos to the density of matter in the universe (see (3.35) and chapter 9) and g_ν is the statistical spin factor. In order not to contradict the observed densities ($\Omega_\nu \leq 1$), it follows that the sums of m_{ν_e}, m_{ν_μ} and m_{ν_τ} may not amount to more than 100 eV. This implies, for m_{ν_μ} and m_{ν_τ} in particular, a considerably sharper bound than the numbers obtained from laboratory experiments.

At somewhat lower energies, electrons and positrons annihilated one another and, after approximately 100 s, the light atomic nuclei D, ^{3}He, ^{4}He

and ^{7}Li were finally formed. At thermal energies in the eV range, after $\sim 10^5$ years, atomic systems had developed, since the photon energy was no longer sufficient to ionize these. This led to the decoupling of the photons. At this time, the universe changed from a radiation-dominated universe to one dominated by matter (see e.g. [Kol90]). At temperatures of approximately 3000 K, the gravitational energy exceeded the thermal energy of the molecules (Jeans criterion) and stars and galaxies began to form. In the next section, we shall briefly indicate the limitations of this standard model.

Before doing this it should at least be mentioned that there are also cosmological models which do not involve an initial singularity, e.g. steady-state models (see [Hoy75]).

3.2 INFLATIONARY MODELS

The cosmological standard model described above has a number of unsatisfactory aspects. It certainly provides for a reasonable linkage between cosmology and GUT models; however, a number of unresolved problems have also arisen (see e.g. [Bör88, Gro89, 90, Kol90]). We mention, in particular, the so-called flatness problem, the horizon problem and magnetic monopoles (GUT monopoles). The latter are required in GUT models, but have not yet been detected (see chapter 8). In addition, the fact that the evolution of the universe to its present state is critically dependent on the initial conditions at the Planck time is viewed as very unsatisfactory. A model in which the later development is essentially independent of the initial conditions at time t_{Pl} would be desirable so that our lack of knowledge of events prior to t_{Pl} would no longer be important.

Inflationary models [Gut81, Alb82, Lin82, 84, 90, Gro89, 90, Kol90, Ell94] provide one solution to the above problems. According to these models the universe passes through an inflationary phase at a certain time after t_{Pl}, during which all the previous conditions are, so to speak, equalized, and at the end of which, as a result of known physical laws, the conditions which led to the universe of today are established.

In the conventional big bang model discussed above, the cosmological constant Λ was assumed to be small and neglected. The case $\Lambda = 0$ corresponds to the assumption that the vacuum does not contribute to the energy density of the universe. However, in quantum field theory, the vacuum contains already different quantum fields. Although these are in a state of lowest energy, this energy is not necessarily equal to zero (cf. the zero-point energy of harmonic oscillators in quantum mechanics).

In inflationary models, it is assumed that $\rho_V \neq 0$. Using (3.10) we write

the Einstein–Friedmann–Lemaitre equations in the form

$$\left(\frac{\dot{R}(t)}{R(t)}\right)^2 = \frac{8\pi G}{3}(\rho(t) + \rho_V(t)) - \frac{kc^2}{R^2(t)} \tag{3.15a}$$

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{3}\left(\rho(t) - 2\rho_V(t) + \frac{3p(t)}{c^2}\right) \tag{3.15b}$$

$\rho(t)$ consists of the matter pressure ρ_M and the radiation pressure ρ_R. From (3.15*b*) it is clear that a positive vacuum density ρ_V corresponds to a negative pressure. If now ρ_V dominates the matter and curvature terms, i.e.

$$|\rho_V| \gg \rho,\ p/c^2 \qquad \left|\frac{8\pi G}{3}\rho_V\right| \gg \frac{|k|c^2}{R^2(t)} \tag{3.16}$$

it follows that

$$\frac{\dot{R}(t)}{R(t)} \simeq \sqrt{\frac{8\pi G}{3}\rho_V} = \sqrt{\frac{c^2\Lambda}{3}}. \tag{3.17}$$

This means that

$$R(t) \simeq R(0)\exp\sqrt{\frac{8\pi G}{3}\rho_V}\,t. \tag{3.18}$$

Thus, for $\rho_V > 0$, we obtain an exponentially expanding universe, a so-called de Sitter or inflationary universe. For such an exponential expansion to take place, the vacuum energy density must dominate the other terms (see (3.16)) for a certain period of time. One mechanism which could provide for this is the spontaneous symmetry breaking by Higgs fields.

During the inflationary phase, the matter becomes so extremely diluted that practically no particles from the time prior to the inflation are observed. All objects observable today stem from the time after this. For example, if magnetic monopoles developed before or in the initial phase of the inflation, we can no longer find them today. It can be shown easily that inflationary models would also solve the horizon and the flatness problems (see e.g. [Gro89, 90, Kol90]). The later development of the universe is as described in the standard model. However, we must stress that the idea of an inflationary universe is still only a hypothesis.

3.3 THE PRIMORDIAL NUCLEOSYNTHESIS

While cosmic background radiation (3 K radiation) provides information about the state of the universe some 10^5 years after the big bang (see chapter 9), the abundances of the lighter nuclei D, ^{3,4}He, and ^{7}Li may be used to obtain information about the universe at a very much earlier point in time. The

primordial nucleosynthesis, i.e. the formation of nuclei from free nucleons, took place approximately 100 s after the big bang. At that time, the temperature of the universe had fallen to approximately 10^9 K, so that the deuterons once formed could no longer be broken up by γ quanta. The existing nucleons mainly fused into ^{4}He, which since then forms about 25% of the mass of the universe.

All heavy elements in the universe, except Li, Be and B, which developed both primordially and by spallation in the cosmic radiation, were created in stars. Nuclei up to iron may be formed by fusion during the hydrostatic burning phases of heavier stars [Bur57]. Heavier elements are mainly formed by neutron-capture processes and subsequent β decays. β^- decay gives rise to a nucleus with atomic number one greater. In this context, we distinguish between the s and the r processes (cf. e.g. [Bur57, Gro89, 90, Käp89, Mat90b, Cow91, Kla91a]). In what follows, we shall only concern ourselves with the early synthesis of light nuclei.

3.3.1 Observed abundances of the primordial elements

The big bang model predicts abundances for the lighter nuclei which agree well with observations. This may be viewed as a great success, since the abundances range over 10 orders of magnitude. Before we discuss the course of the primordial nucleosynthesis, we give a brief summary of the existing experimental data about the cosmic abundances of the light nuclei (see e.g. [Boe85, Kol90]).

After hydrogen, ^{4}He is by far the most common light nuclide. The primordial origin of most ^{4}He is apparent from both its large amount ($Y_p \approx 25\%$) and its relatively uniform distribution throughout the whole universe. Y_p conventionally denotes the primordial mass fraction of ^{4}He in the overall mass of the cosmos. Since helium also developed later in stars, the present-day mass fraction Y_0 consists of the primordial component Y_p and a contribution ΔY from the development of stars.

$$Y_0 = Y_p + \Delta Y. \tag{3.19}$$

Since ^{4}He is very stable Y_0 gives an upper bound for the primordial mass fraction.

The abundance of ^{4}He in different astrophysical objects has been measured using various methods (see [Boe85, Ril91]). However, the derivation of the primordial component Y_p depends on certain model assumptions. Despite these problems, different observations give very similar results. As previously mentioned, the fact that all objects have a similar ^{4}He content points to the primordial origin. [Ril91] gives the following approximate average value from more recent measurements

$$Y_p = 0.230 \pm 0.010. \tag{3.20}$$

It is much more difficult to measure the mass fraction of other light nuclei, since the quantities are considerably smaller and thus scarcely any absorption lines can be observed. Often, it is even easier to detect the corresponding emission lines associated with recombinations, e.g. in H II regions.

In comparison with ^{4}He the deuteron is a weakly bound nucleus which is easily destroyed in the interior of a star when the temperature exceeds 6×10^5 K. The observed D abundances are thus generally regarded as a lower bound for the primordial abundances. To date, deuterium has been detected in our solar system and in the interstellar medium. The ratio of the number of deuterium nuclei to that of normal hydrogen is found to be [Bör88, Kol90, Ril91]

$$1 \times 10^{-5} < (\mathrm{D/H})_p < 2 \times 10^{-4}. \tag{3.21}$$

Only a fraction of the primordial ^{3}He will have survived until today, since ^{3}He is converted into ^{4}He in the fusion processes in stars. On the other hand, during the development of stars, deuterium is burnt into ^{3}He, so that the galactic abundance of ^{3}He is determined by two competing processes. It has been found that [Ril91]

$$\left(\frac{\mathrm{D} + {}^3\mathrm{He}}{\mathrm{H}}\right)_p < 10^{-4}. \tag{3.22}$$

The following has been obtained for ^{7}Li [Boe85]

$$10^{-10} < ({}^7\mathrm{Li/H})_p < 8 \times 10^{-10}. \tag{3.23}$$

3.3.2 The course of the nucleosynthesis

The course of primordial nucleosynthesis is described in detail e.g. in [Hay50, Wei72, Zel83, Bör88, Kol90, Wal91, Cop95]. The number of nuclei which may be formed is critically dependent on the ratio of protons to neutrons. For thermal energies $E > 1$ MeV ($T > 10^{10}$ K), leptons, hadrons and photons existed in thermodynamic equilibrium. The equilibrium between protons and the β-unstable neutrons was maintained by the following processes of the weak interaction

$$p + e^- \leftrightarrow n + \nu_e \qquad n + e^+ \leftrightarrow p + \overline{\nu}_e \tag{3.24}$$

The ratio of the neutron number N_n to the proton number N_p in thermodynamic equilibrium at temperatures $T > 10^{10}$ K is given by the Boltzmann factor

$$\frac{N_n}{N_p} \simeq \exp\left(\frac{-(m_n - m_p)c^2}{kT}\right) \tag{3.25}$$

where $(m_n - m_p)c^2 \simeq 1.3$ MeV. Thus, for a temperature $T = 3 \times 10^{10}$ K, it follows that

$$\frac{N_n}{N_n + N_p} \simeq 0.38. \tag{3.26}$$

As the average thermal energy fell below 1 MeV, the neutrinos were decoupled from the equilibrium, since the rate of the weak interaction Γ_w could no longer follow the rate of expansion of the universe H. The reactions (3.24) became too slow to maintain the equilibrium and the ratio N_n/N_p froze at the temperature $T_f \simeq 10^{10}$ K.

$$\left(\frac{N_n}{N_p}\right)_f = \exp\left(\frac{(m_p - m_n)c^2}{kT_f}\right). \tag{3.27}$$

The neutron fraction further only decreased slowly as a result of the neutron decay.

$$N_n(t) = (N_n)_f e^{-t/\tau_n} \tag{3.28}$$

where $t = 0$ for $T = T_f$. On freezing, the following held:

$$\frac{N_n}{N_n + N_p} \simeq \frac{1}{6} \qquad (T_f = 10^{10}\ \text{K}). \tag{3.29}$$

With the β decay of the neutron, this value decreased until the beginning of the primordial nucleosynthesis at $T \simeq 10^9$ K ($E = 100$ keV) to

$$\frac{N_n}{N_n + N_p} \simeq \frac{1}{7} \qquad (T = 10^9\ \text{K}). \tag{3.30}$$

The only way of forming complex nuclei involves a network of two-body reactions, the most important of which require deuterium

$$p + n \leftrightarrow \text{D} + \gamma \tag{3.31a}$$

$$\text{D} + \text{D} \leftrightarrow {}^3\text{He} + n \leftrightarrow {}^3\text{H} + p \tag{3.31b}$$

$${}^3\text{H} + \text{D} \leftrightarrow {}^4\text{He} + n. \tag{3.31c}$$

At temperatures of 10^{10} K, the deuterium which has developed is very quickly destroyed again by photodissociation since the binding energy of around 2.2 MeV is very small and the photons are approximately 10^9 times more common than nucleons. Heavier nuclei can only be formed by the reactions (3.31*b*, *c*) when the amount of D formed by the reaction (3.31*a*) is sufficiently high. This is the case for temperatures of 0.8×10^9 K [Bör88].

Under these conditions practically all the neutrons present fuse over the network (3.31) into ^{4}He. In addition, small quantities of D and ^{3}He are formed. Thus, the mass fraction of the primordial ^{4}He in the total baryonic matter

$$Y_p \simeq \frac{2N_n/N_p}{N_n/N_p + 1} \tag{3.32}$$

is expected to be approximately 25% (for $N_n/(N_n + N_p) = 1/7$ we have $Y_p = 2/7$).

Since there are no stable nuclei with mass numbers $A = 5$ and $A = 8$, practically no heavier elements are formed. Small amounts of ^{7}Li and ^{7}Be are formed via the reactions

$$^4\mathrm{He} + {}^3\mathrm{H} \rightarrow {}^7\mathrm{Li} + \gamma \tag{3.33a}$$

$$^4\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma \tag{3.33b}$$

$$^7\mathrm{Be} + e^- \rightarrow {}^7\mathrm{Li} + \nu_e. \tag{3.33c}$$

The calculated abundances of the primordial elements are critically dependent on the baryon density in the universe (cf. figures 3.1 and 3.2). Thus, the predicted abundances are plotted against the baryon density ρ_B or against

$$\eta = \frac{N_B}{N_\gamma}. \tag{3.34}$$

The ratio η determines when the photodissociation of the deuteron is overcome. A large value of η means fewer photons, in other words, the photodissociation of D decreases and the synthesis of helium may begin earlier, i.e. at higher temperatures. Since somewhat more neutrons are present then, correspondingly, more helium may be formed.

The calculated abundances in figures 3.1 and 3.2 are largely consistent with the observed values in the range $3 \times 10^{-10} < \eta < 5 \times 10^{-10}$ (however, see also [Ril91]). This is worth noting, insofar as the curves for the individual nuclides are very different. The ^{4}He fraction is relatively insensitive to η, while the other abundances, and in particular those of deuterium and lithium, vary considerably with η. In particular, the measured deuterium abundance may be used to obtain information about the baryonic density in the universe. Given the range $3 \times 10^{-10} < \eta < 5 \times 10^{-10}$ it follows that the ratio $\Omega_{0,\mathrm{bary}}$ of all present-day baryonic matter to the value of the critical density (for $\Lambda = 0$), which represents the boundary between an open and a closed universe, satisfies

$$0.01 < h^2\Omega_{0,\mathrm{bary}} < 0.018 \tag{3.35}$$

with h from (3.4) [Bör88]. If one considers only the bound derived from deuterium, it follows that

$$\Omega_{0,\mathrm{bary}} < 0.10\text{–}0.20. \tag{3.36}$$

The primordial nucleosynthesis determines an upper bound for the baryonic matter existing in the universe. According to (3.36), the latter is insufficient to close the universe. This is also not possible in scenarios of an *inhomogeneous*

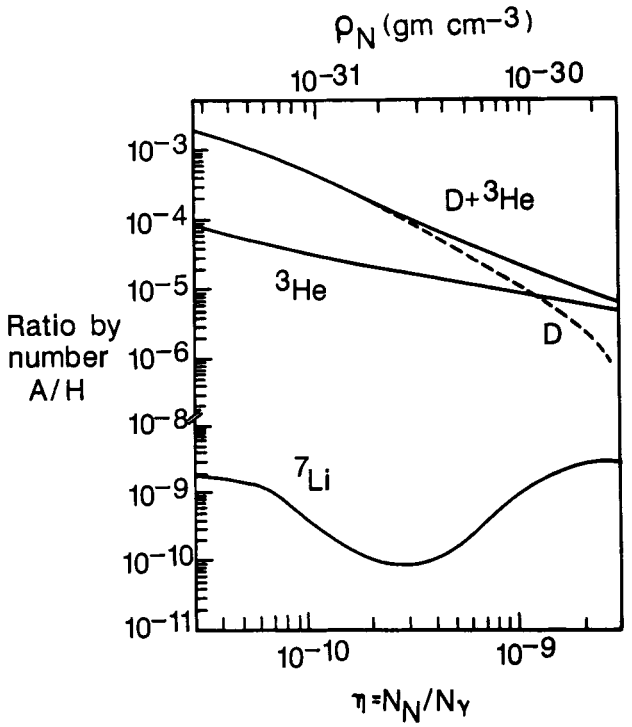


Figure 3.1. Calculated element abundances in the big bang nucleosynthesis as a function of the present-day baryon density ϱ_B (equals the nucleon density ϱ_N) (after [Boe85]). The half-life of the neutron is taken to be 10.6 min (from [Bör88]). For N_N ($= N_B$), N_γ see text.

big bang which were investigated in recent years (see e.g. [Mal91, 93, Oli91, Thi91]) and which are based on inhomogeneities in the baryon density that may originate in connection with the QCD phase transition. However, we note that the admission of a *time-dependent* gravitational constant could lead to values of $\Omega = 0.1$–1 in the early nucleosynthesis [Sta92c].

The theoretical predictions depend not only on η but also on the lifetime of the neutron τ_n. τ_n determines the rate at which the reactions (3.24) take place. The greater the lifetime is, the lower the reaction rate of the weak interaction is, i.e. the earlier the neutrinos freeze out of the equilibrium. This gives a higher temperature of freezing T_f, and thus also a larger value for $(N_n/N_p)_f$, so that more helium can be formed in the nucleosynthesis. In addition, a longer lifetime τ_n naturally ensures that more neutrons survive in the period between freezing and the beginning of the nucleosynthesis. However, the last effect only plays a subordinate role.

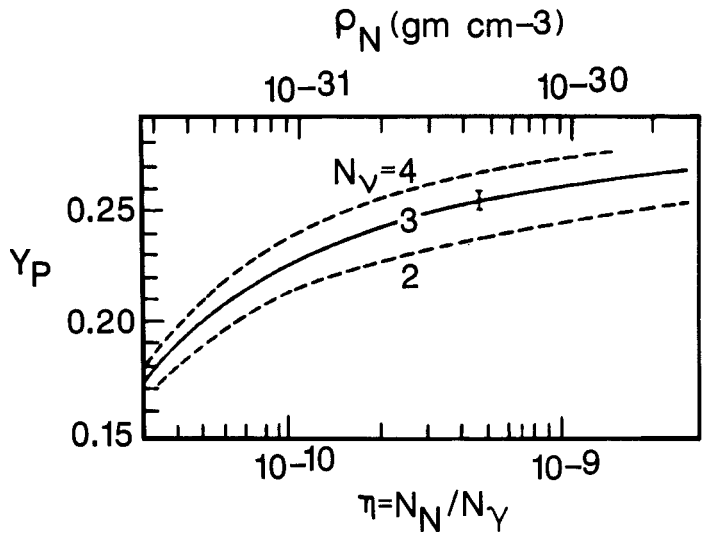


Figure 3.2. Dependence of the ^{4}He abundance on the number of neutrino generations N_ν (from [Bör88]). ϱ_N is the nucleon density ($\varrho_N = \varrho_B$).

The current value for the average lifetime of the neutron [Par92] is

$$\tau_n = (889.1 \pm 2.1)\ \text{s}. \tag{3.37}$$

The most accurate measurements to date were made using stored ultracold neutrons [Mam89].

$$\tau_n = (887.6 \pm 3.0)\ \text{s}. \tag{3.38}$$

The influence of τ_n on the calculated mass fraction of ^{4}He is shown in figure 3.3.

3.3.3 The number of neutrino flavours

The theoretically calculated ^{4}He abundances depend not only on η and τ_n, but also on the number of light neutrino flavours[2] ($m_\nu < 1\ \text{MeV}/c^2$). Since the fraction of primordial ^{4}He is supposedly very well known and the dependence of Y_p on the parameters η and τ_n is not very large, the number of neutrino families may be determined from measured abundances. We shall explain these relationships in what follows (cf. [Yan79a, 84, Oli81, Dol81, Blo84, Den90, Kol90]).

[2] More precisely, the quantity of helium depends on the number of types of light particles existing at the time of nucleosynthesis, or, more exactly, on the relativistic degrees of freedom. As yet unknown, exotic particles could also have contributed to this.

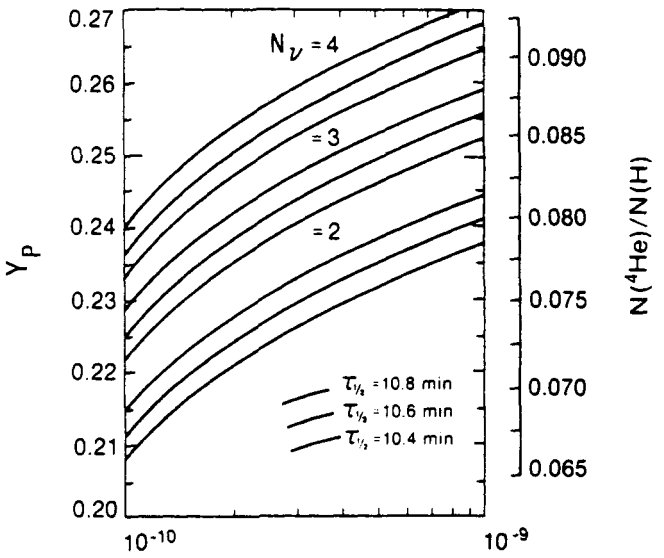


Figure 3.3. Effect of the neutron half-life on the calculated mass fraction Y_p of ^{4}He [Yan84].

The weak interaction rate Γ_w for processes which convert neutrons and protons into one another decreases as the temperature falls, and when

$$\Gamma_w < H(t) \tag{3.39}$$

the ratio N_n/N_p freezes. Practically all neutrons present at this time are fused into helium. Thus, the expansion rate at energies in the MeV region is a critical factor determining the neutron-to-proton ratio and thus the amount of ^{4}He. The expansion rate $H(t)$ is related to the density $\rho(t)$ in the universe. It follows from (3.15*a*), e.g. assuming a Euclidian metric ($k = 0$), that

$$H(t) = \sqrt{\frac{8\pi G}{3}(\rho(t) + \rho_V(t))}. \tag{3.40}$$

For $kT \simeq 1$ MeV the universe is still radiation dominated. All relativistic particles contribute to the total density $\rho(t)$ (in what follows, we shall neglect the vacuum energy density ρ_V), i.e. $\rho(t)$ is mainly composed from contributions of photons, electrons and light neutrinos

$$\rho(t) = \rho_\gamma + \rho_e + \rho_\nu. \tag{3.41}$$

Thus, the ^{4}He abundance depends on the number of types of relativistic particles on freezing. A larger number of neutrino flavours means a greater expansion rate and thus an earlier decoupling of the neutrinos. An earlier exclusion of the neutrinos from the equilibrium means that the ratio N_n/N_p is determined at a higher temperature $T_{f'}$ ($T_{f'} > T_f$). Thus,

$$(N_n/N_p)_{f'} > (N_n/N_p)_f \tag{3.42}$$

whence the fraction of primordial helium is greater the more neutrino flavours exist. In addition, a greater expansion rate means that the temperature decreases more rapidly, the nucleosynthesis begins earlier and the neutrons have less time to decay.

Corresponding calculations of Y_p as a function of η, τ_n and the number of neutrino flavours N_ν have been carried out by a number of authors (see e.g. [Blo84, Yan84]). The results of [Yan84] are illustrated in figures 3.2 and 3.3. For $Y_p < 0.25$ essentially only three light neutrino families are consistent with the assumptions of the primordial nucleosynthesis. However, four types of neutrino cannot be totally ruled out. Denegri *et al* [Den90] derive an upper bound of

$$N_\nu < 3.6 \tag{3.43}$$

with a confidence limit of 95%. This result agrees well with the most recent results from LEP (see chapter 2). Conversely, the LEP result $N_\nu = 3$ may be taken as a boundary condition for calculations relating to the primordial nucleosynthesis (see [Ril91]).

To end this section, we note that the above argument only holds for light neutrinos with $m_\nu < 1$ MeV/c^2. Heavier neutrinos are no longer relativistic at temperatures of 10^{10} K and their abundances are thermodynamically suppressed. Even if more than three families had been observed at LEP this would not necessarily have contradicted cosmology, since neutrinos with $m_\nu > 1$ MeV/c^2 play practically no role in the nucleosynthesis but would have been easily detectable via the decay width of the Z^0.

The cosmological abundance of heavier neutrinos is less than that of lighter neutrinos by a Boltzmann factor

$$a = \exp\left(-\frac{m_\nu c^2}{kT_f}\right) \tag{3.44}$$

where T_f denotes the temperature at which the heavy neutrino decouples. The factor (3.44) relates the abundance of a neutrino with mass m_ν to that of a massless neutrino at T_f.

The following formula has been obtained for the contribution of heavy stable neutrinos of mass m to the matter density of the universe [Kol90]:

$$\Omega_{\nu\bar{\nu}}h^2 = 3(mc^2/\mathrm{GeV})^{-2}\left(1 + \frac{3\ln(mc^2/\mathrm{GeV})}{15}\right) \tag{3.45}$$

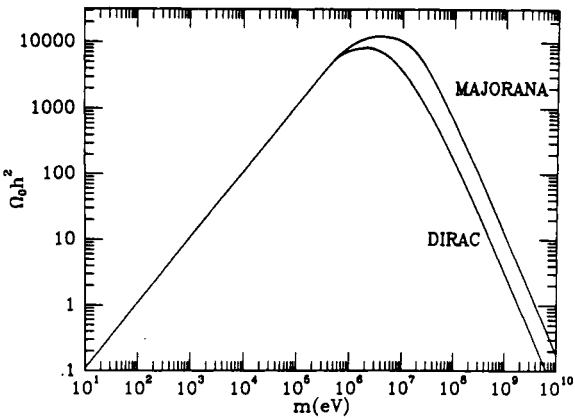


Figure 3.4. The contribution of a heavy, stable neutrino of mass m to the matter density of the universe (from [Kol90], $\Omega_0 = \Omega_\nu$). Only neutrino masses smaller than $92h^2$ eV or greater than 2 GeV (Dirac case) or 5 GeV (Majorana case) are cosmologically acceptable.

where the antineutrinos are assumed to have the same abundance as the neutrinos: $\Omega_{\nu\overline{\nu}} = 2\Omega_\nu$ (see figure 3.4).

So as not to contradict $\Omega \leq 1$, we must have $m \geq 2$ GeV/c^2 (see also [Kol86b]); this is often called the Lee–Weinberg bound. Thus, neutrinos with masses between 100 eV/c^2 and 2 GeV/c^2, including the 17 keV neutrino (chapter 6), must be unstable. However, masses greater than 2 GeV/c^2 are again cosmologically acceptable and such heavy neutrinos are good candidates for dark matter. Neutrino decay below the Lee–Weinberg bound cannot take place via $\nu_h \rightarrow \nu_l + \gamma$, since, in this case, the contribution to the cosmic γ background would be too great (see e.g. [Kol90]). However, a decay involving the emission of majorons would be allowed (see e.g. [Moh91a, Gel91]).

However, the Lee–Weinberg bound is only valid when the heavy neutrinos are stable. Since very heavy neutrinos are probably unstable, the decay mode and the lifetime of each hypothetical neutrino must be considered (see [Tur81, Kol84b, 86c]). A detailed overview of the cosmological and experimental bounds for unstable neutrinos is given in [Roo88].

Additional relativistic particles are not the only way of altering the expansion rate. A time-varying gravitational constant would also have affected the development of the cosmos. The temporal variation of G over cosmological time periods may be checked using the ^{4}He fraction. We shall return to this topic in chapter 12.

Chapter 4

Proton Decay

4.1 BARYON NUMBER

The baryon number B is an additive quantum number. No elementary interaction has yet been associated with the quantity B (see also chapter 11 on the fifth force). This new interaction should either have an infinite range, like the electromagnetic interaction, or be associated, under certain conditions, with the existence of a new particle, the majoron (see section 6.2.3). Although the conservation of the baryon number is not based on any fundamental symmetry principle, no process in which B changes has yet been observed. As far as is currently known, the decay of nucleon resonances and that of hyperons and their resonances leads, without exception, to a proton or a neutron, where the latter then decays via the weak interaction into a proton. Only transitions between different baryons occur; single baryons cannot be destroyed or created. If the baryons are assigned the baryon number $B = 1$, the antibaryons $B = -1$ and all other elementary particles $B = 0$, then experimental observation shows that the sum of the baryon numbers should be conserved for an arbitrary reaction. An example of an allowed reaction is $p\overline{p}$ pair creation in e^+e^- annihilation

$$e^+ + e^- \rightarrow p + \overline{p}. \tag{4.1}$$

If the baryon number is conserved exactly by all interactions, it follows that the proton is absolutely stable since it is the lightest baryon. The process

$$p \rightarrow e^+ + \gamma \tag{4.2}$$

which is hypothetically allowed by the laws of conservation of energy, angular momentum and electrical charge, would then be forbidden. Thus, the stability of the proton is a very sensitive test for the exact conservation of B.

In 1929, Weyl first attempted to explain the proton stability by postulating a new conserved quantity [Wey29]. This idea was taken up by Stückelberg

in 1939 [Stü39] and later by Wigner in 1949 [Wig49] and ultimately led to the above assignment of the baryon number. However, the explanation of the proton stability based on the introduction of a new quantum number unsupported by an underlying symmetry principle is unsatisfactory and is currently only viewed as a (very useful) 'book-keeping' idea.

As early as 1967, Sacharov described general reasons why an instability of the proton should be expected [Sac67a, b]. Based on the P and CP violation, and the lack of large amounts of antimatter in an expanding universe, he deduced the existence of nucleon decay. The excess of matter over antimatter in the universe of today contradicts the general assumption that a symmetry between particles and antiparticles should have arisen shortly after the big bang. The generation of this excess of matter requires a violation of the conservation of baryon number, CP violation and a thermal non-equilibrium (see chapter 3 and e.g. [Wei79a]).

The elementary interactions must be CP violating, since then processes which lead to baryons could occur with greater rates than the corresponding CP-conjugate processes which lead to antibaryons. The decays of the X and Y bosons

$$X \xrightarrow{r} u + u \qquad X \xrightarrow{1-r} \overline{d} + e^{+} \tag{4.3a}$$

$$\overline{X} \xrightarrow{\overline{r}} \overline{u} + \overline{u} \qquad \overline{X} \xrightarrow{1-\overline{r}} d + e^{-} \tag{4.3b}$$

could lead to an excess of u, d and e^{-} over $\overline{u}$, $\overline{d}$ and e^{+} when $r > \overline{r}$. On the other hand, the CP invariance would imply $r = \overline{r}$. If the baryon number is a conserved quantity, not all the decays in (4.3) are allowed. This means that the explanation of the matter–antimatter asymmetry necessarily requires a B violation if we assume an originally symmetrical initial state. Finally, the two conditions given above could only lead to an excess of baryons in thermal non-equilibrium. In thermal equilibrium no particular time direction is distinguished. Even when the reaction rates for creating baryons are greater than those for creating antibaryons, the same is also true for the inverse reactions. In equilibrium, the ratios of the particle numbers are independent of the dynamics of the reactions.

In fact, most grand unification theories predict a decay of the proton or the (bound) neutron and thus a violation of the conservation of the baryon number (see chapter 1). The replacement of global symmetries by local, possibly spontaneously broken, gauge invariances is an important feature in the development of modern field theories. The Lagrange density of the standard model of the electroweak interaction conserves the baryon number B and the lepton number L separately. On the other hand, the SU(5) gauge theory violates B and L conservation, but conserves the difference between the baryon number and the lepton number ($(B - L)$ conservation). In the extension to the SO(10)

gauge group, the latter is replaced by a spontaneously broken local gauge invariance [Mar80].

If the symmetries underlying the weak interaction originated from a substructure of quarks and leptons (and if the forces on the substructure level were similar to QCD), then $SU(2)_L \otimes SU(2)_R \otimes U(1)_{B-L}$ would occur as natural symmetry of the weak interaction instead of $SU(2)_L \otimes U(1)$ (see [Moh86a], p 116ff).

The search for proton decay, together with the $n\bar{n}$ oscillations to be discussed in the next chapter are of crucial importance as far as the development of GUT models is concerned, since they represent two of the few experimentally directly accessible consequences.

4.2 THEORETICAL PREDICTIONS OF THE LIFETIME OF THE PROTON

A lower bound for the lifetime of the proton may be deduced immediately from the fact we exist. This simple observation implies that the lifetime of the proton must be greater than the age of the universe of approximately 10^{10} years. This bound may be further increased by the following consideration. If the lifetime of the proton were less than 10^{16} years, then the approximately 10^{28} protons present in the human body would decay with an average rate of 10^{12} protons per year, i.e. approximately 30 000 decays per second. However, such a large number of decaying protons would be life threatening (see [Bet86a]). Current GUT models, on the other hand, predict very much smaller decay rates.

4.2.1 Proton decay in the SU(5) model

The minimal SU(5) model proposed by Georgi and Glashow in 1974 is the simplest gauge theory which contains the $SU(3)_c$ and the $SU(2)_L \otimes U(1)$ groups as subgroups of a *simple* gauge group. The 15 elementary left-handed fermion fields are arranged in a $\bar{5}$ multiplet and a decuplet (1.137).

The group SU(5) has 24 generators (see section 1.5), i.e. there are 24 gauge fields which mediate the SU(5) interactions. Twelve of these 24 gauge bosons are already known, namely, the eight gluons of the colour interaction, the $W^{\pm}$ and Z^0 bosons and the photon. In addition, 12 new gauge bosons, called X and Y bosons are introduced. The properties of these new gauge bosons were discussed in chapter 1 (see tables 1.3 and 1.4). The mass of these exchange particles corresponds approximately to the unification energy of the order of 10^{15} GeV.

The most important low-energy prediction of the SU(5) model is the nucleon decay. Since leptons and quarks belong to the same multiplet, the X and Y bosons may mediate transitions between both particle types. Such B violating

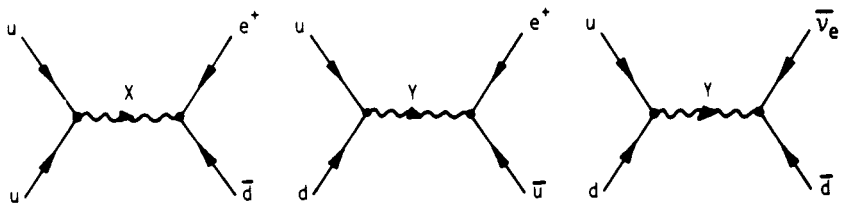


Figure 4.1. B violating transitions between quarks and leptons in the SU(5) model (from [Bec81]).

transitions are described by the diagrams in figure 4.1. At lower energies, these graphs may be replaced by an effective four-fermion interaction [Wei79b, Wil79, Bec81]

$$\mathcal{L}_{\mathrm{GUT}} = \frac{4}{\sqrt{2}} g_{\mathrm{GUT}} \left[(\overline{u}_L^c \gamma_\mu u_L)(2\overline{e}_L^+ \gamma^\mu d_L + \overline{e}_R^+ \gamma^\mu d_R) + (\overline{u}_L^c \gamma_\mu d_L)(\overline{\nu}_e \gamma^\mu d_R) + \mathrm{h.c.} \right]. \tag{4.4}$$

As in the transition from the GWS theory of the electroweak interaction to Fermi's four-fermion approach, the coupling constant of SU(5), g_5, translates into an effective coupling constant, g_{GUT}

$$\frac{g_{\mathrm{GUT}}}{\sqrt{2}} = \frac{g_5^2}{8m_X^2} = \frac{g_5^2}{8m_Y^2}. \tag{4.5}$$

Because of the large mass $m_X c^2 \sim 10^{15}$ GeV, the interaction mediated by the virtual X and Y bosons is very weak and has a very short range, $R \sim \hbar/m_X c$, which amounts to only approximately 10^{-29} cm. Since the extent of the nucleon of 10^{-13} cm is 16 orders of magnitude greater, quarks are extremely rarely found at a distance apart which permits a process such as that shown in figure 4.1.

Before we evaluate the proton lifetime we shall briefly discuss the different decay channels. In 1979, Weinberg described two general selection rules for baryon- and lepton-number violating processes, which are mediated by the exchange of a very heavy scalar or vector particle [Wei79b]

$$\frac{\Delta L}{\Delta B} = 1 \qquad \frac{\Delta S}{\Delta B} = 0 \quad \text{or} \quad -1 \tag{4.6}$$

where S denotes the strangeness quantum number. The first rule says that one of the decay products of the proton must be an antilepton (e^+, μ^+, $\overline{\nu}_i$; $i = e, \mu, \tau$). According to the second selection rule, either a meson with $S = 0$ (π, η, ρ, ω) or a meson with positive strangeness (K^+, K^0, $K^*(892)$) is generated so that the possible decay channels are very restricted. The allowed decay modes for

Table 4.1. Expected branching ratios for the proton decay in the SU(5) model (after [Luc86]).

Decay mode	Branching ratio [%]
$p \rightarrow e^+\pi^0$	31–46
$p \rightarrow e^+\eta$	0–8
$p \rightarrow e^+\rho^0$	2–18
$p \rightarrow e^+\omega$	15–29
$p \rightarrow \bar{\nu}_e\pi^+$	11–17
$p \rightarrow \bar{\nu}_e\rho^+$	1–7
$p \rightarrow \mu^+K^0$	1–20
$p \rightarrow \bar{\nu}_\mu K^+$	0–1

the proton and the neutron in the SU(5) model are [Moh86a]:

$$\begin{aligned} p &\rightarrow e^+\pi^0 \\ &\rightarrow e^+\rho^0 \\ &\rightarrow e^+\omega^0 \\ &\rightarrow e^+\eta \\ &\rightarrow \overline{\nu}\rho^+ \\ &\rightarrow \overline{\nu}\pi^+ \\ &\rightarrow \mu^+K^0 \\ &\rightarrow \overline{\nu}_\mu K^+ \end{aligned} \tag{4.7a}$$

and

$$\begin{aligned} n &\rightarrow \nu\omega \\ &\rightarrow \overline{\nu}\rho^0 \\ &\rightarrow \overline{\nu}\pi^0 \\ &\rightarrow e^+\rho^- \\ &\rightarrow e^+\pi^- \\ &\rightarrow \overline{\nu}_\mu K^0. \end{aligned} \tag{4.7b}$$

Additional symmetry considerations and phase-space arguments lead to predictions of the branching ratios which are summarized for the decay of the proton in table 4.1 [Luc86].

In the minimal SU(5) model the decay channel

$$p \rightarrow e^+\pi^0 \tag{4.8}$$

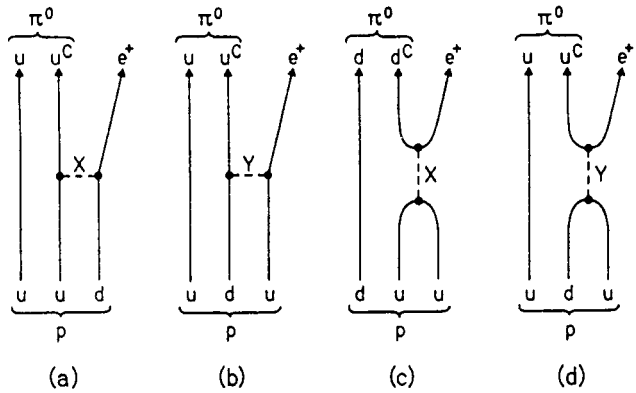


Figure 4.2. Diagrams contributing to the proton decay $p \rightarrow e^+ + \pi^0$ in the SU(5) model (from [Gro89, 90]).

is dominant. Figure 4.2 shows a number of diagrams which contribute to this decay channel. The decay rate is easy to evaluate from such graphs, as follows. The main contribution to the transition amplitude comes from the X-boson propagator function

$$A_p \sim \frac{1}{q^2c^2 - m_X^2c^4}. \tag{4.9}$$

Since the transfer of momentum is small in comparison with the mass of the gauge bosons, it follows that

$$A_p \sim \frac{1}{m_X^2c^4}. \tag{4.10}$$

The X and Y bosons couple with strength $\sqrt{\alpha_5}$ to the quarks and leptons, where

$$\alpha_5 = \frac{g_5^2}{4\pi\hbar c\epsilon_0} \tag{4.11}$$

(ϵ_0 is the permittivity of vacuum). Thus, for the second-order processes in figure 4.2, it follows that

$$A_p \sim \frac{\alpha_5}{m_X^2c^4}. \tag{4.12}$$

This essentially corresponds to the effective coupling constant introduced in (4.4). The following estimates for the decay width may be obtained using dimensional arguments

$$\Gamma_p \sim \alpha_5^2 \frac{m_p^5c^2}{m_X^4}. \tag{4.13}$$

Here, the rest energy of the decaying proton is used as energy scale. Thus, the lifetime of the proton

$$\tau_p = \frac{\hbar}{\Gamma_p} = D\frac{m_X^4}{\alpha_5^2 m_p^5 c^2} \tag{4.14}$$

increases with the fourth power of the mass of the exchange particle. The factor m_p^{-5} gives approximately the phase-space dependence. Since $m_p c^2 \gg m_\pi c^2 + m_e c^2$, the decay energy is approximately equal to the rest energy of the proton. The factor D takes account of the fact that the quarks between which bosons are exchanged are bound in a proton. However, these corrections are of order of magnitude 1, i.e. $D \approx \hbar$. If α_5 and m_X are known, a specific prediction for τ_p is obtained.

The GUT parameters m_X and α_5 may be obtained by extrapolating the U(1), SU(2) and SU(3) coupling constants to high energies up to the unification point using the renormalization group equation (section 1.5.2). The sensitivity to details of the model is low; however, because τ_p depends on m_X^4, the mass of the X bosons must be determined very accurately taking into account all possible corrections. Various investigations have given similar results for α_5 and m_X [Lan81, 86]

$$\alpha_5(m_X^2) = 0.0244 \pm 0.0002 \tag{4.15a}$$

$$m_X c^2 = 1.3 \times 10^{14}\,\text{GeV}\left(\frac{\Lambda}{100\,\text{MeV}}\right)(\pm 50\%). \tag{4.15b}$$

Λ is the scale factor of (1.131a). Overall in the SU(5) model the following lifetime for the dominant decay channel $p \to e^+ + \pi^0$ is obtained

$$\begin{aligned}\tau_p(p \to e^+ + \pi^0) &= 6.6 \times 10^{28\pm 0.7}\left(\frac{m_X c^2}{1.3 \times 10^{14}\,\text{GeV}}\right)^4 \text{years} \\ &= 6.6 \times 10^{28\pm 1.4}\left(\frac{\Lambda}{100\,\text{MeV}}\right)^4 \text{years}.\end{aligned} \tag{4.16}$$

For $\Lambda = 300$ MeV, it follows that

$$\tau_p(p \to e^+ + \pi^0) = 5.3 \times 10^{30\pm 1.4}\ \text{years} \tag{4.17}$$

i.e. in the SU(5) model the lifetime should be not more than about 10^{32} years.

4.2.2 Proton decay in supersymmetric GUT models

In addition to the minimal SU(5) model, there are a large number of other approaches to grand unification theories. However, most of these approaches

do not permit specific predictions about the lifetime of the proton. It is interesting that the Kaluza–Klein theories, described in somewhat more detail in section 12.2, give lifetimes of the order of 10^{45} years. Predictions of non-supersymmetric SO(10) models lie in the range from 10^{32} to 10^{38} years [Lee95].

We give now a brief discussion of proton decay in the framework of supersymmetric models. Since the lifetime of the proton is proportional to the fourth power of the mass of the X boson, the extension of the normal GUT models to supersymmetric GUT models has interesting consequences for the stability of the nucleon. The energy at which the GUT symmetry is broken may be extrapolated from the lower-energy behaviour of the coupling constants using the renormalization group equation (see section 1.5.2). This involves the number of participating particles. The introduction of supersymmetry, i.e. the assignment of a bosonic partner to each fermion and vice versa, necessarily leads to a doubling of the number of elementary particles. This enlargement of the particle spectrum means that the coupling constants only converge at approximately 10^{16} GeV [Lan86, 88]. Thus, in the minimal SUSY–GUT model, it follows that

$$m_X^{\mathrm{SUSY}} c^2 \simeq 4.8 \times 10^{15}\ \mathrm{GeV} \left(\frac{\Lambda}{100\ \mathrm{MeV}} \right) \tag{4.18}$$

whence the decay channel $p \rightarrow e^+ + \pi^0$ is strongly suppressed. The predicted dominant decay is now not that into a positron and a neutral pion, but the transitions $p \rightarrow K^+ + \overline{\nu}_\mu$ and $n \rightarrow K^0 + \overline{\nu}_\mu$. The latter are more difficult to access experimentally. Predictions of supersymmetric SO(10) models lead to far less straightforward predictions [Lee95a].

Observation of nucleon decay would be extraordinarily important as a test of grand unification theories, in particular, with respect to the extrapolation of the energy scale over 13 orders of magnitude.

4.3 PROTON DECAY EXPERIMENTS

The basic idea for the realization of experiments to search for the decay of the nucleon is very simple. However, theoretical predictions are extremely difficult to test. A half-life of 10^{30} years corresponds to the decay of approximately one nucleon per day in 1000 tonnes of matter ($\simeq 5 \times 10^{32}$ nucleons). Sensitivity to 10^{33} years requires the detection of one decay per year in 3000 tonnes of matter. Because of the enormously long half-lives expected for this process, proton decay experiments are only sensitive if many hundreds or thousands of tonnes of matter are observed for several years.

The decay $p \rightarrow e^+ + \pi^0$ with the subsequent decay of the pion into two γ quanta gives rise to two electromagnetic showers which could be detected using suitable counters. To ensure that interference effects due to the cosmic

background radiation are as small as possible, these experiments are carried out deep underground. In addition, the energies and particle trajectories must be measured in order to separate decay candidates from background events.

There are essentially two methods for searching for the decay of the nucleon:

(i) Indirect detection of the nucleus which results from the fact that a complex nucleus loses a nucleon. This method is advantageous when the daughter nucleus cannot be generated in any other way.
(ii) Direct detection of the particles emitted during the decay.

4.3.1 Indirect detection

The advantage of indirect detection (method (i)) is that it is practically independent of the decay channel. Even the disappearance without trace of a nucleon would be detectable from the fact that the nucleon leaves a hole in an occupied shell. The daughter nucleus is generated in an excited state, so that the nucleon decay could be detected via a chain of decays into an easily identifiable daughter nucleus.

The disappearance of a nucleon in a heavy nucleus could give rise, for example, to spontaneous fission if the excitation energy exceeds the height of the fission barrier. The half-life of ^{232}Th under spontaneous fission is greater than 10^{21} years [Fle58]. Since each of the 232 nucleons in the thorium nucleus could trigger a spontaneous fission, the lifetime of the nucleon must be greater than 10^{23} years. A similar bound is obtained in the search for the free neutron which remains after the decay of the proton in a deuterium nucleus.

Some daughter nuclei which may arise as a result of the decay of the proton are only generated in small quantities by other processes. This permits experiments based on the radiochemical analysis of geological samples. By determining the amount of an appropriate nuclide which has accumulated in mineral samples over geological time periods, it is possible to derive information about the lifetime of the proton. This method is associated with uncertainties resulting from limited knowledge of the past history of the mineral samples.

The decay of a nucleon bound in a ^{39}K nucleus generates either a ^{38}K or a ^{38}Ar nucleus. With a probability of approximately 21%, these emit a further nucleon and become ^{37}Ar. Since ^{37}Ar is radioactive, the concentration of these nuclei can be measured from their radiation. The greatest problem in this experiment is that of finding a rock sample which could not have reached the Earth's surface since its formation. The number of nucleons which have decayed during the geological period may be determined from the ratio of ^{37}Ar to ^{39}K (cf. [LoS85]).

Background effects may be very much better evaluated in radiochemical analyses which are independent of the uncertainty of the geological past. These

analyses use a pure sample of a suitable substance and try to chemically detect the nuclei generated over a certain period of time as a result of nucleon decay. In the most sensitive experiment of this type, two tonnes of potassium acetate were taken into the Homestake goldmine in South Dakota. It was possible to extract a few argon atoms from the two-tonne sample. The corresponding production rate of ^{37}Ar was less than one atom per day, corresponding to a lower bound of 2.2×10^{26} years [Ste77] for the lifetime of the nucleon.

Table 4.2 gives an overview of the bounds for the lifetime of the nucleon, which have been determined by indirect methods (from [Gol80]).

Table 4.2. Bounds for the lifetime τ_N of the nucleon, based on indirect methods (after [Gol80]).

Method	Bounds [years]
Spontaneous fission of ^{232}Th	$\tau_N > 2 \times 10^{23}$
Deuteron	$\tau_N > 3 \times 10^{23}$
$^{130}Te \rightarrow {}^{129}Xe$	$\tau_N > 1.6 \times 10^{25}$
$^{39}K \rightarrow {}^{37}Ar$	$\tau_N > 2.2 \times 10^{26}$

4.3.2 Direct detection

The second method involves identifying the proton decay by direct detection of the particles emitted during the decay. However, in general, such measurements are not sensitive to all conceivable decay channels.

In 1953, Reines, Cowan and Goldhaber first undertook a search for the proton decay using a massive detector [Rei54]. The detector consisted of a tank filled with 300 litres of scintillation fluid which was used to detect charged particles. The flashes of light were recorded by 90 photomultipliers. To screen against cosmic radiation the whole detector was installed at a depth of 30 m below ground. A lower bound of 10^{22} years for τ_p was determined.

A number of other experiments were carried out over the following years. Reines and his colleagues built a detector comprising 20 tonnes of CH_2 fluid scintillator, which was installed in a South African gold mine near Johannesburg at a depth of 3200 m [Rei74]. The scintillator was surrounded by approximately 84 000 photomultipliers. In particular, this was used to search for events in which a muon was stopped in the detector and then decayed. These muons may stem, for example, from the decays

$$p \rightarrow \mu^+ + \pi^0 \quad \text{or} \quad p \rightarrow \nu + \pi^+ \quad \pi^+ \rightarrow \mu^+ + \nu_\mu. \tag{4.19}$$

The depth of 3200 m guarantees that no muons from cosmic rays come into

contact with the detector. However, neutrinos cross the 3200 m thick screening practically unhindered. Muons may be produced by interactions in the detector and the surrounding rock. A total of six muons were observed during the period of measurement from 1965 to 1974. This corresponds to the expected neutrino-induced muon flux in the gold mine. Thus, there was no indication of a possible instability of the proton. From the six events a lower bound of

$$\tau_p > 3 \times 10^{30} \text{ years} \tag{4.20}$$

was derived for the decay of the proton in an end channel which contains a muon [Lea79].

Further experiments have been assembled and executed during the last 10 years; these may be divided into two classes, namely multilayer track detectors and Cerenkov water counters. These two types of detector will be described and compared in the following sections.

4.3.2.1 Multilayer track detectors

Multilayer track detectors are essentially giant iron calorimeters in which, for example, 3–12 mm thick steel or iron plates alternate with electronic counters which are used to detect charged particles passing through. The detectors include plastic flash tubes, proportional chambers, streamer tubes and Geiger counters. By way of example, figure 4.3 shows the Fréjus detector which was in operation until recently [Mey86]. The detector with a total mass of 912 tonnes and a fiducial mass of 750 tonnes was constructed in the Fréjus road tunnel on the border between France and Italy and screened against cosmic rays by 1550 m of rock. The counters between the 3 mm thick iron plates consisted of 930 000 plastic flash tubes and 40 000 Geiger tubes.

In this experiment the iron nuclei formed the supply of protons and neutrons. Unlike Cerenkov counters, track detectors have a very good resolution which permits a comparatively good reconstruction of the particle trajectories. As far as the resolution is concerned, thin plates are preferred; however, these also give rise to greater costs, since more particle counters are required.

Four large multilayer track detectors to search for the decay of the nucleon have been built: KGF [Kri81, 82], NUSEX [Bat82, 83], Fréjus [Bou88, Ber89, 91b] and Soudan II [All88, Thr93]. Table 4.3 summarizes the most important data about these experiments. One disadvantage of calorimetric methods of measurement is that no free nucleons are present and the protons and neutrons are bound in complex (iron) nuclei. This results in a reduced detection sensitivity for hadronic decay channels in comparison with the case of free nucleons, since the hadrons which are generated in the decay are absorbed with a probability of 30% in the nuclei in which the transition takes place [Ric87].

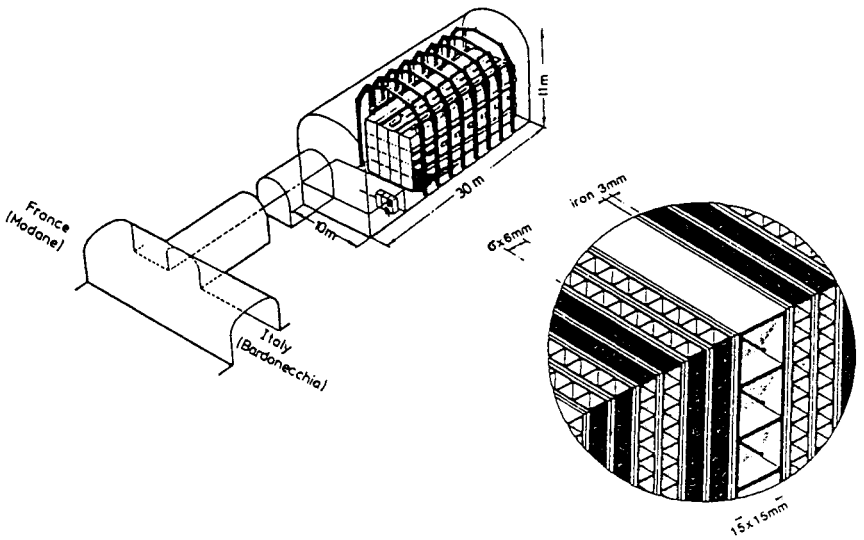


Figure 4.3. The Fréjus detector used to search for proton decay. The scale shows the colossal size of the detector. The material to be investigated comprises 750 tonnes of iron. The section on the right shows the fine structure of the detector. Since the potential charged decay products in iron have only a short range, the iron was arranged in the form of thin plates, with the strip-like scintillation detectors between them. The fact that these detectors were rotated by 90° every two positions made it possible to reconstruct the tracks of ionizing particles in three dimensions (after [Mey86]).

Table 4.3. Properties of proton decay experiments (iron calorimeters).

	KGF	NUSEX	Fréjus	Soudan II
M_{tot} [t]	140	150	912	1000
M_{eff} [t]	60	113	550	600
Depth [m]	2300	1850	1780	760
Water equivalent [m]	7600	5000	4850	1800
Vertex resolution [cm]	10	1	0.5	~ 0.5
Location	Kolar gold mine	Mont Blanc tunnel	Fréjus tunnel	Soudan ore mine

One advantage of the calorimeters is their modular structure. In particular, this allows one to study the neutrino-induced background by exposing individual modules to the neutrino radiation in an accelerator.

4.3.2.2 Cerenkov water counters

In the experiments with Cerenkov counters, a large volume of clear water is surrounded by photomultipliers. Here, the water is used both as a proton donor and as detector material. If a charged particle moves through a transparent medium with a velocity v which is greater than that of light in this medium (c/n), Cerenkov radiation is emitted. The angle α between the direction of flight of the charged particle and the direction in which the Cerenkov radiation is emitted depends on the ratio of the velocities.

$$\cos\alpha = \frac{1}{\beta n} \qquad \beta = \frac{v}{c}. \tag{4.21}$$

However, Cerenkov counters are only sensitive to charged particles with velocities greater than the threshold

$$\beta_T = \frac{1}{n} \tag{4.22}$$

where for water, $\beta_T = 0.75$. For a particle which moves through water at approximately the speed of light, the Cerenkov angle α is approximately 41°.

The basic information provided by a Cerenkov counter consists of Cerenkov rings for particles which are stopped or decay in the detector and Cerenkov circular surfaces for particles which leave the detector with a corresponding velocity. The arrangement of the photomultipliers and their operational time sequence may be used to reconstruct the particle trajectory and the direction. Pulse levels are a measure of the energy lost by a charged particle in the water. However, the resolution is worse than that of calorimeters.

Since only charged particles emit Cerenkov light, neutral particles cannot be directly detected. In particular neutrinos escape unhindered. Other electrically neutral elementary particles may be detected via secondary decay products. Neutral pions, for example, decay into two high-energy photons which are themselves electrically neutral but can generate observable signals by the formation of e^+e^- pairs.

A further advantage of Cerenkov water counters is that one-fifth of all protons in water are free protons. These produce a particularly clear trace, since there is no interaction with the other nucleons in a nucleus and since practically no smearing of energy occurs as a result of the Fermi momentum (see below).

Three large Cerenkov water counters have been built, namely Kamiokande I and II [Hir89], IMB I and III [Sei88] and HPW [Phi89]. The characteristics of these detectors are summarized in table 4.4. Figure 4.4 shows the IMB detector which was installed at a depth of 700 m in the Thiokol salt mine near Cleveland, Ohio. The overall mass was 8000 tonnes. The Cerenkov pulses are recorded by 2048 photomultipliers [Sei88]. For the Kamiokande detectors see figures 4.9–4.11.

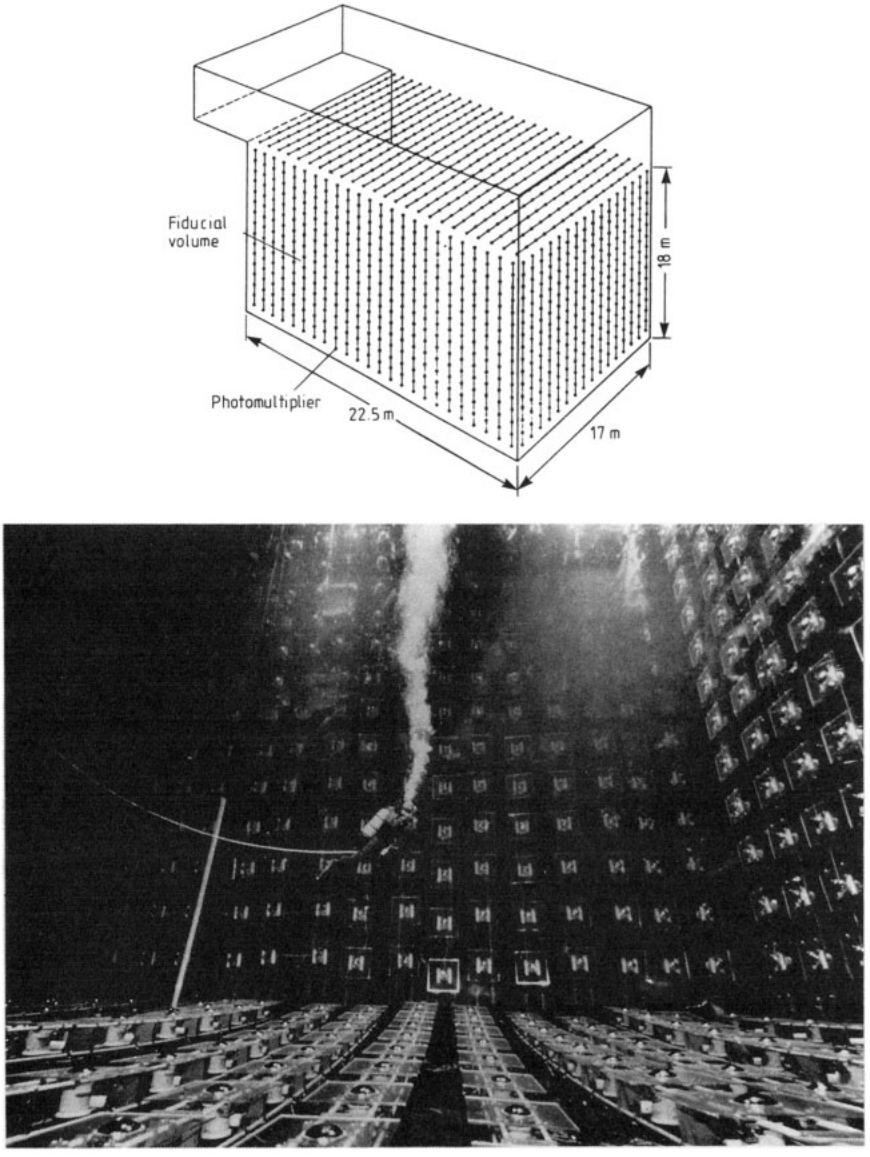


Figure 4.4. The IMB Cerenkov water counter with its 8000 tonnes of water and 2048 photomultipliers which was constructed in the Morton–Thiokol saltmine near Cleveland at a depth of approximately 700 m (1580 m water equivalent). Top: diagrammatic illustration (from [LoS85b]); bottom: view into the detector (photo: Joe Stancampiano and Karl Luttrell, © National Geographic Society).

Table 4.4. Properties of proton decay experiments (Cerenkov water counters).

	Kam I (II)	IMB I, III	HPW	Superkam
M_{tot} [t]	3000	8000	680	50 000
M_{eff} [t]	880 (1040)	3300	420	22 000
Depth [m]	825	600	525	825
Water equivalent [m]	2400	1600	1500	2400
Vertex resolution [cm]	100 (20)	100		10
Location	Kamioka ore mine	Thiokol salt mine	King silver mine	Kamioka ore mine

4.3.2.3 Background events

Since the expected decay rates for protons and bound neutrons are very small, a very precise knowledge of the background events is needed in order to identify true decay events. One source of background radiation is the natural radioactivity which cannot be fully suppressed since all screening material and the detector material itself contains certain radioactive impurities. However, the typical decay energies lie in the MeV area, i.e. they are only approximately 1% of the energy of the decay products freed in proton decay. Thus, by simple measurement of energy it is possible to suppress this background component.

It is much more difficult to identify perturbing background events from the cosmic radiation which covers all energy areas and a broad range of different particles on the Earth's surface. To decrease the flux of cosmic radiation, the large detectors are installed deep under the Earth's surface, for example in road tunnels or mines. The incident nucleons and pions may already be screened by a few metres of absorbing material. Muons, on the other hand, lose very little energy when crossing matter. Thus, effective screening must be several thousand metres thick. Even then high-energy muons are not completely held back by such enormous masses of rock. Figures 4.5 and 6.26 show the measured muon flux in various underground laboratories (cf. [Ern84]).

The penetrating muons may pass through the detector without reaction, decay in the detector or trigger reactions there. In addition, neutral hadronic particles are generated as a result of interactions in the surrounding rock; these may reach the detector practically unnoticed and trigger reactions there. This muon-induced background is reduced by choosing the fiducial detector volume to be less than the actual volume. Events which are outside the inner volume are rejected since these could be associated with particles penetrating from outside. Muons decaying in the fiducial detector area one tries to identify via their decay electrons.

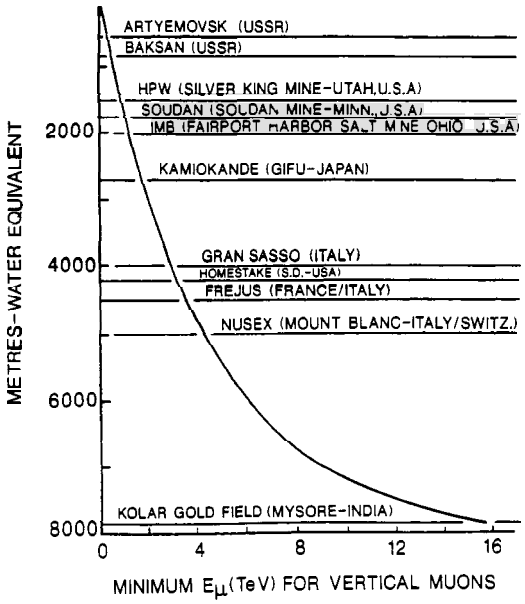


Figure 4.5. Muon screening of various underground laboratories (from [Bar88a]).

The sensitivity of a proton decay experiment is limited by neutrinos with energies in the range 0.5–2 GeV from the decay of cosmic muons, pions and kaons in the Earth's atmosphere. Figure 4.6 shows the flux of atmospheric neutrinos for various experiments. This amounts to approximately 130 neutrinos per 1000 tonnes per year (see [Los85a]). The reaction products from interactions of these neutrinos include muons, electrons and pions which have approximately the same energy as the expected nucleon decay products. To separate reactions induced by neutrinos from true candidates for nucleon decay one must consider the geometry of the events. When a neutrino interacts with a particle, the reaction products move in a forward direction. On the other hand, if a resting nucleon decays, the momenta of the emitted particles add to give an overall zero momentum. This means that true events may, in principle, be identified from momentum and energy conservation. For the decay of a free resting nucleon

$$\sum_i \boldsymbol{p}_i = 0 \tag{4.23a}$$

$$\sum E_i = m_N c^2 \qquad E_i = \sqrt{p_i^2 c^2 + m_i^2 c^4} \tag{4.23b}$$

where $\boldsymbol{p}_i$ and E_i denote, respectively, the momentum and the energy of the ith decay product and m_N is the rest mass of the nucleon. The decay channel

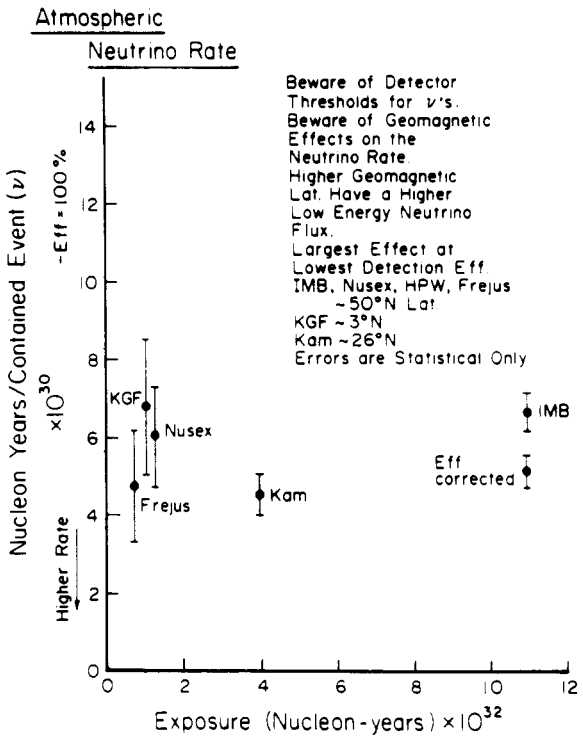


Figure 4.6. Inverse observed neutrino rate as a function of the exposure for various proton-decay detectors (after [LoS85a]).

$p \to e^+ + \pi^0$ which is predicted to be dominant in the minimal SU(5) model produces the characteristic signature shown in figure 4.7. Two electromagnetic showers with practically opposite directions are expected to develop in the detector. This example shows how the geometry of an event is used as a criterion in the evaluation. However, the limited position resolution imposes limits on this reconstruction.

Decay channels with invisible particles (neutrinos) cannot be detected in this way. Another problem arises in the case of nucleons bound in the nucleus. The geometry is smeared by the Fermi momentum p_F of the decaying nucleon. In this case (4.23) becomes

$$\sum_i p_i = p_F \qquad \sum_i E_i = E_N. \tag{4.24}$$

In addition, the hadronic decay products may interact in the nucleus and so reduce the efficiency of the detector.

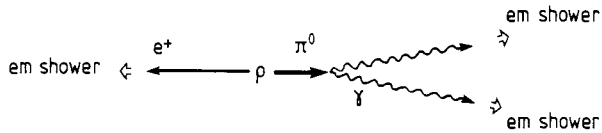


Figure 4.7. Characteristic signature of the dominant decay channel $p \rightarrow e^+ + \pi^0$ in the SU(5) model.

There are no such kinematic boundary conditions for the decay channels $p \rightarrow \overline{\nu} +$ mesons. The way in which the candidates are selected is heavily dependent on the particular channel. For example, the important decay in SUSY models $p \rightarrow K^+ + \overline{\nu}$ is characterized by a monoenergetic muon from the decay of the stopped kaon.

Monte Carlo simulations play an important role in the evaluation of the events and the analysis of the decay channels. The probability of a neutrino being mistaken for a proton decay is of the order of 1% in today's detectors, corresponding to a simulated proton decay rate of approximately 1 per 1000 tonnes per year. This limits the sensitivity of the detectors to proton lifetimes of approximately 10^{33} years.

4.3.2.4 Results

Practically all experiments have observed events which satisfy all the conditions for a nucleon decay. However, the corresponding rates are comparable with the expected background rate, so that as yet there is no evidence for an instability of the proton.

Only the Kolar group actually claims to have seen a proton decay [Kri81, 82]. The authors derive an average lifetime of approximately 7×10^{30} years from the six observed events. However, this result has not been confirmed by other, much more sensitive, experiments.

Currently, the best bounds come from the IMB, Kamiokande and, in particular, the Fréjus collaboration. From the number of background events for each decay channel it is possible to derive a lower bound for the partial lifetime using the following formula

$$\tau_{p/n}/B = \frac{S\epsilon N_{p/n}}{s_{90}}. \tag{4.25}$$

Here, B denotes the (unknown) branching ratio for the decay channel in question, $N_{p/n}$ is the number of protons/neutrons per kilotonne detector, S denotes the overall luminosity in kilotonne years), ϵ is the detector efficiency which is calculated using Monte Carlo simulations, and finally, s_{90} is the upper bound

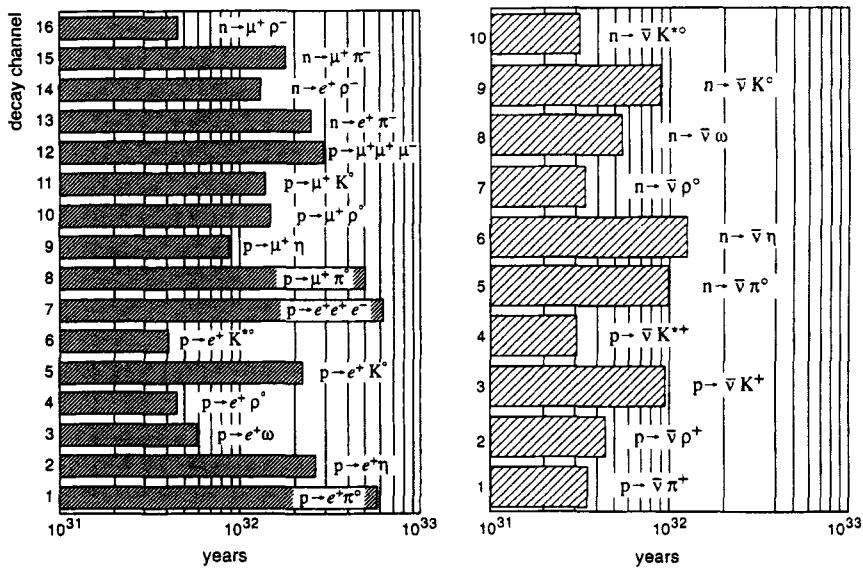


Figure 4.8. Lower bounds (90% confidence level) for the partial lifetimes of different decay channels for the nucleon decay. Left: decays of the form $N \to l^{\pm}x$. Right: decays of the form $N \to \overline{\nu}x$ (from [Rau89]).

for the number of decay candidates (the index 90 specifies a confidence level of 90%).

Bounds for the partial lifetimes are now available for a large number of decay channels ([Bou88, Sei88, Ber89, 91b, Hir89, Phi89]). Taking the results of the three largest experiments, IMB, Kamiokande and Fréjus together, it turns out that for the dominant SU(5) decay channel [Rau89]

$$\tau_p(p \to e^+\pi^0)/B > 5.9 \times 10^{32} \text{ years} \tag{4.26}$$

and for the modes preferred in the SUSY SU(5) model

$$\tau_p(p \to K^+\overline{\nu})/B > 1 \times 10^{32} \text{ years} \tag{4.27a}$$

$$\tau_n(n \to K^0\overline{\nu})/B > 0.9 \times 10^{32} \text{ years.} \tag{4.27b}$$

Figure 4.8 gives the lower bounds for a number of other decay channels. Comparison of (4.26) with the SU(5) predictions shows that the minimal SU(5) model can probably be ruled out since reasonable values of the parameter Λ give a maximum expected lifetime of about 10^{32} years.

4.3.2.5 *Future experiments*

No conclusive evidence for the instability of the nucleon has yet been obtained. The present lower bounds for the lifetime lie between 5×10^{31} and 5×10^{32} years, depending on the decay channel. With existing detectors, as mapped in figure 4.13 (see e.g. [Lea93, Arp94], IMB has since been shut down), only small improvements can be expected since the sensitivity of the experiments has practically reached the background level generated by atmospheric neutrinos.

Thus, real progress requires much larger and better detectors. Currently, there are no plans for an experiment which is solely concerned with the search for nucleon decay. Some detectors which are mainly intended to detect solar neutrinos or neutrinos from supernova explosions are also able to detect the nucleon decay. These include a new Cerenkov water counter with a fiducial mass of 22 000 tonnes (Superkamiokande in Japan, [Kaj89, Suz94]). Figure 4.9 shows the temporal development of the Kamiokande experiments and figure 4.10 the schematic structure of Superkamiokande, the 'big brother' of the Kamiokande detector (figure 4.11). Over a period of measurement of 10 years without a measured candidate, this project could increase the bound for the decay $p \rightarrow e^+ + \pi^0$ to approximately 1.6×10^{34} years [Rau89].

Two further experiments in Europe are at the construction stage. Both are multipurpose detectors (see [Bal92, Lea93]) and both are based in the underground laboratory of the Laboratorio Nazionale del Gran Sasso near Rome [Bel88], the largest and most modern underground laboratory in the world (see also chapters 6, 7 and 10). These comprise the LVD detector [Agl93, Bar88a] and the ICARUS project [Ben92b, 94], a giant time projection chamber (TPC) which will be filled with approximately 4000 tonnes of liquid argon (figure 4.12).

LVD stands for Large Volume Detector. It consists of 1800 tonnes of scintillator (2280 m^3 and 1800 tonnes steel). Over a measurement period of 15 years a small improvement of the bound for the preferred decay channel in SUSY models to approximately 3×10^{32} years is expected [Ber85, Bar88a]. The advantage of a scintillation counter over a Cerenkov water counter is that a scintillator which is equipped with faster electronics can detect both the K^+ and the μ^+ from the decay of the kaon. The K^+ from the proton decay $p \rightarrow K^+ + \overline{\nu}$ has a kinetic energy of 105 MeV, a total energy of 599 MeV and a velocity $\beta = 0.566$. Thus, the kaon lies below the Cerenkov threshold $\beta_T = 0.75$, so that a Cerenkov water counter only detects the muon. Consequently, it is impossible to distinguish between a proton decay and a neutrino-induced muon.

To date only a prototype of ICARUS of 3 tonnes is under operation (figure 4.12(*a*)). With the final setup (figure 4.12(*b*)) one expects to set limits for proton decay at around $(1–3) \times 10^{33}$ years [Ben94]. ICARUS is also planned to be used in long-baseline neutrino experiments with a CERN SPS neutrino beam directed to the Gran Sasso laboratory [Egg95] (see section 7.3.9).

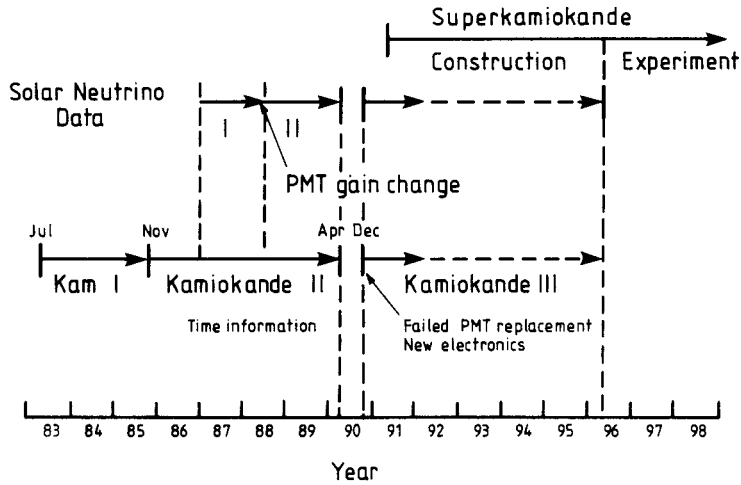


Figure 4.9. The temporal development of the Kamiokande and Superkamiokande experiments (from [Suz92]).

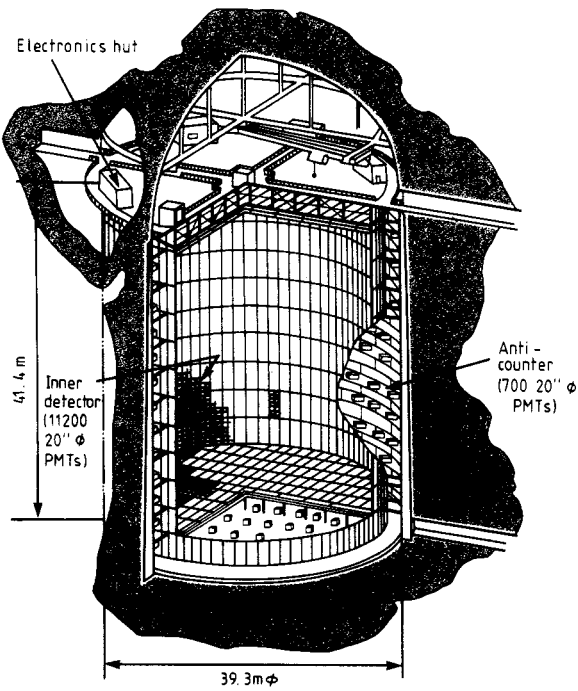
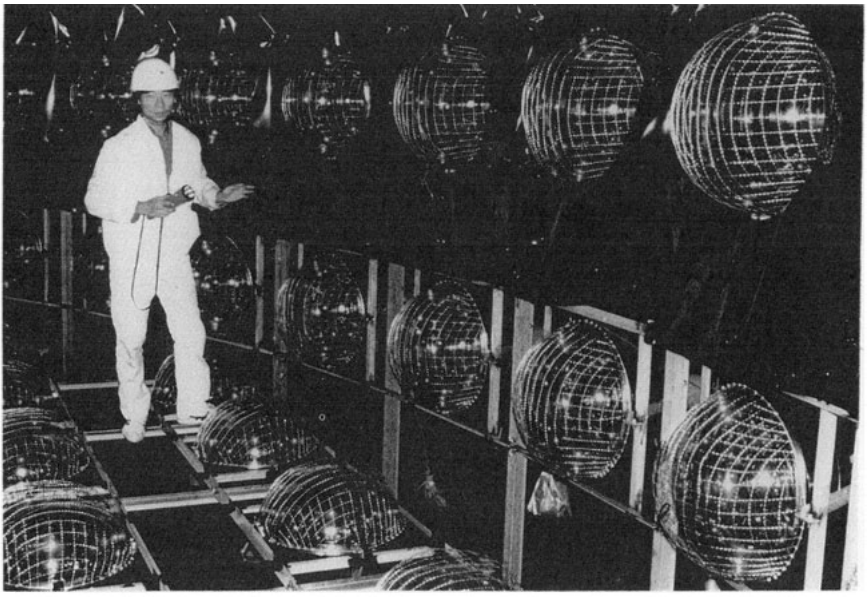


Figure 4.10. The schematic structure of the Superkamiokande detector. It has a height of 41 m and a diameter of 39 m and contains 50 000 tonnes of water (from [Sin91, Suz92]).



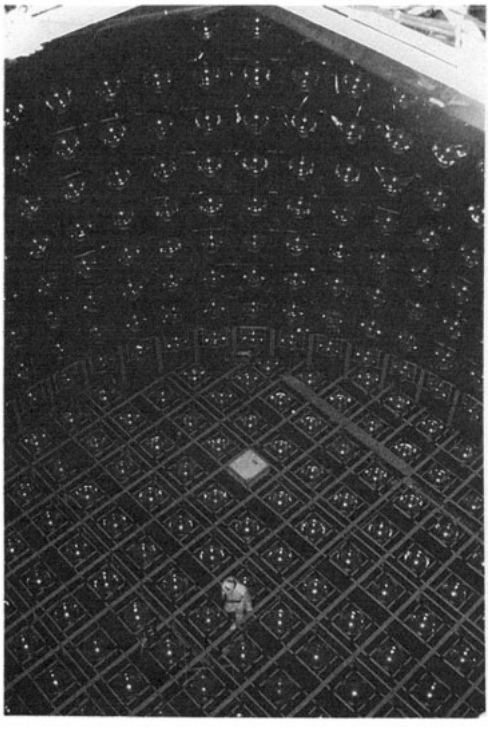


Figure 4.11. Two views into the Kamiokande II detector (courtesy of Y Totsuka).

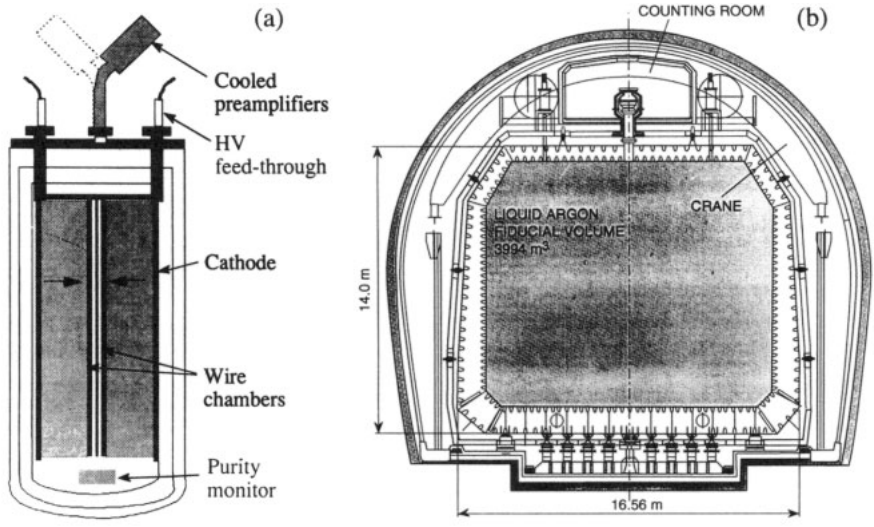


Figure 4.12. (*a*) Diagram of the prototype of the ICARUS detector (from [Ben93]). The active volume of the TPC, which is filled with 3 tonnes of liquid argon, is divided into two independent half cylinders. (*b*) Plan of the ICARUS project in the Gran Sasso laboratory (courtesy: Carlo Rubbia).

Figure 4.13. The most important underground laboratories in the world. Squares: underground laboratory. Circles: neutrino telescopes in lakes, in the sea or under ice.

An overview of the most important underground laboratories in the world, including sites of neutrino telescopes such as AMANDA, DUMAND, etc (see [Arp94, Kla95]), is given in figure 4.13 (cf. figure 2.13).

Chapter 5

Neutron–Antineutron Oscillations, and Electric Dipole Moment of the Neutron

The availability of cold neutrons in nuclear reactors has made it possible to undertake many experiments relating to fundamental questions of physics. These include precise measurements of the properties of the neutron, in particular of its mass, charge, lifetime and magnetic moment (see table 5.1). Studies of the electric dipole moment d_n of the neutron provide important boundary conditions for models of CP and T violation[1] (see section 1.3). Experiments on neutron–antineutron oscillations investigate a possible violation of the conservation of the baryon number B.

Table 5.1. Properties of the neutron [PDG92].

Mass m_n	$939.565\,63 \pm 0.000\,28$ MeV/c^2
Magnetic moment d_m	$(-1.913\,0427 \pm 0.000\,0005)\mu_N$
Lifetime τ	889.1 ± 2.1 s
Spin J^π	$1/2^+$

$\mu_N = 3.152\,451\,66(28) \times 10^{-14}$ MeV T^{-1} is the nuclear magneton.

In what follows, we shall discuss only two topics which are important in relation to GUT models, namely the electric dipole moment of the neutron and $n\bar{n}$ oscillations.

[1] C = charge conjugation, P = parity, T = time reversal.

5.1 ELECTRIC DIPOLE MOMENT OF ELEMENTARY PARTICLES

The electric dipole moment of a particle with angular momentum J is defined by

$$d = \int \rho_{JJ} z \, \mathrm{d}V. \tag{5.1}$$

The coordinate z is measured from the centre of the particle, ρ_{JJ} denotes the electric charge density inside the particle, where the orientation of the angular momentum J is given by $m = J$ relative to the z-axis. If the particle is charged, a non-zero d means that the mass centre of gravity and the charge centre of gravity do not coincide. On the other hand, for uncharged particles, a non-zero d would indicate an asymmetric charge distribution with vanishing net charge.

The permanent electric dipole moment of the neutron and, in general, that of every elementary particle, must vanish if there is to be T invariance. This can easily be seen from the following discussion [Pur50, Ram82]. In what follows, we shall talk about the electric dipole moment of the neutron, by way of a specific example, but the arguments hold for any particle. However, the neutrality of the neutron makes it a particularly suitable object for establishing the effects of a small electric dipole moment. The orientation of $\boldsymbol{d}_n$ is determined by the orientation of the spin $\boldsymbol{\sigma} = (2/\hbar)\boldsymbol{s}$, since the spin is the only preferred direction in space

$$\boldsymbol{d}_n \propto \boldsymbol{\sigma}. \tag{5.2}$$

If T invariance is required, the dipole moment $\boldsymbol{d}_n$ and the spin $\boldsymbol{\sigma}$ must change sign in the same way under T, if $d_n \neq 0$. When the time reversal transformation is applied all movement-related quantities change their sign; in particular, the spin changes sign, while the electric dipole moment is not subject to a change of sign. There would then be a preferred time direction. Thus, T invariance requires a vanishing electric dipole moment of the neutron (see also figure 5.1).

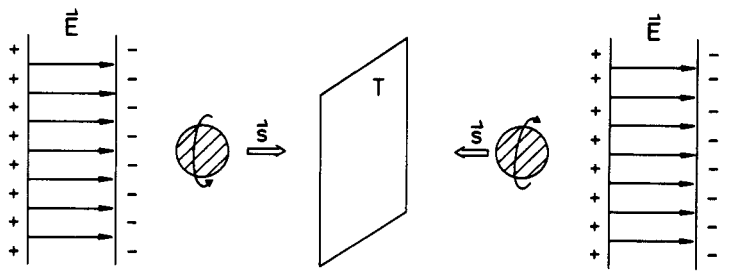


Figure 5.1. Electrical dipole moment of the neutron: behaviour of the electrical field $\boldsymbol{E}$ and the spin $\boldsymbol{s}$ under the time-reversal operation T.

We shall pursue this argument a little further. The interaction of an electric or magnetic dipole with the electromagnetic field is given in the classical approximation by

$$\text{magnetic} \qquad H_m = -d_m \boldsymbol{s} \cdot \boldsymbol{B} \tag{5.3a}$$

$$\text{electric} \qquad H_e = -d_e \boldsymbol{s} \cdot \boldsymbol{E}. \tag{5.3b}$$

The behaviour of the quantities $\boldsymbol{\sigma}$, $\boldsymbol{E}$ and $\boldsymbol{B}$ under T is summarized in table 1.5. We have

$$H_m = -d_m \boldsymbol{s} \cdot \boldsymbol{B} \xrightarrow{T} -d_m(-\boldsymbol{s}) \cdot (-\boldsymbol{B}) = H_m \tag{5.4a}$$

$$H_e = -d_e \boldsymbol{s} \cdot \boldsymbol{E} \xrightarrow{T} -d_e(-\boldsymbol{s}) \cdot \boldsymbol{E} = -H_e. \tag{5.4b}$$

Thus, a non-vanishing electric dipole moment violates time-reversal invariance (see figure 5.1). The parity transformation has a corresponding effect

$$H_e = -d_e \boldsymbol{s} \cdot \boldsymbol{E} \xrightarrow{P} -d_e \boldsymbol{s} \cdot (-\boldsymbol{E}) = -H_e. \tag{5.5}$$

In summary, an electric dipole moment of a particle violates both P and T invariance. At first sight this may appear somewhat surprising, since atoms and molecules have large electric dipole moments, while the electromagnetic interaction is invariant under P and T. However, these dipole moments may be explained by the occurrence of degenerate states. We shall discuss the hydrogen atom by way of example (cf. [Per82]). The first excited state of the H atom ($n = 2$) contains $s_{1/2}$ and $p_{1/2}$ levels which, apart from the Lamb shift, are energetically degenerate. The two levels have opposite parities. In the electrical field, a mixing now occurs due to the Stark effect. The new energy eigenstates are linear combinations of the s and p states and therefore have no definite parity. This is associated with an asymmetric distribution of the electron charge which results in a finite electric dipole moment.

If P and T are conserved, an electric dipole moment of the neutron could only arise if there was a second neutron state with opposite parity which was energetically degenerate with the first state. However, there is no evidence of two different states of the neutron. In particular, such a degeneracy would contradict the fact that neutrons satisfy the Pauli principle. This means that the measurement of an electric dipole moment is an important test for the time-reversal invariance, in particular, as far as Wolfenstein's superweak model and the strong CP problem of QCD are concerned (see sections 1.3 and 5.2.5).

We shall now estimate the possible order of magnitude of the electric dipole moment of the neutron d_n. We have

$$d_n \approx \text{charge} \times \text{length} \times g \tag{5.6}$$

where the parameter g describes the strength of the T violation. Since parity is also violated, the weak interaction must also be involved, so that d_n contains the Fermi coupling constant $G_F \approx 10^{-5}\hbar^3 c^{-1} m_p^{-2}$. The typical energy is given by the neutron mass m_n. Thus, the energy-independent coupling amounts to $G_F m_n^2 c^4/(\hbar c)^3 \approx 10^{-5}$. The neutron is uncharged, so the dipole moment must result from an asymmetry of the positive and negative charge clouds. The characteristic length scale for the neutron is its Compton wavelength

$$\lambda_n = \frac{\hbar}{m_n c} \approx 2 \times 10^{-14} \text{ cm}. \tag{5.7}$$

It seems reasonable to use the CP parameter $\epsilon \approx 2 \times 10^{-3}$ for the parameter g, so that

$$d_n \sim \epsilon e \lambda_n G_F m_n^2 c^4/(\hbar c)^3 \sim 4 \times 10^{-22} e \text{ cm}. \tag{5.8}$$

A very much smaller value of approximately

$$d_n \sim 10^{-32} e \text{ cm} \tag{5.9}$$

is expected in the standard model [Bar87b]. The Kobayashi–Maskawa model states that CP violation is associated with the heavy quarks, so that only diagrams which are of second order in the weak interaction (see figure 5.2) contribute to the electric dipole moment. This may explain, at least in part, the extremely small value in (5.9).

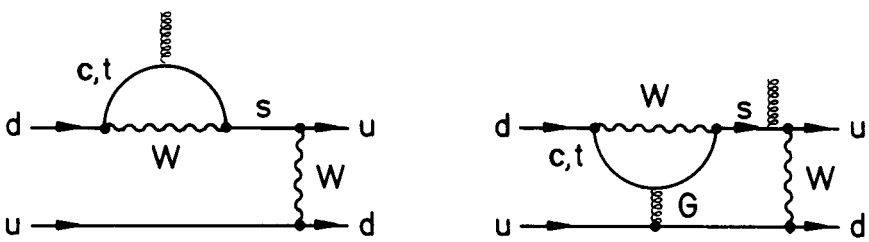


Figure 5.2. Contributions of second-order diagrams (diquark diagrams) to the electrical dipole moment of the neutron in the Kobayashi–Maskawa model (from [Ham85a]).

In right–left symmetric models and in supersymmetric theories the T violation may also occur in the light quark sector, so that terms of first order in the weak interaction may contribute to the electric dipole moment. An approximate value of

$$d_n \sim 10^{-26\pm1} e \text{ cm} \tag{5.10}$$

is expected for right–left symmetric models [Moh74, 86a, He89].

Wolfenstein's phenomenological model (see section 1.3.2) essentially gives a vanishing dipole moment. There also exist cosmological arguments relating to CP violation. If the baryon–antibaryon asymmetry observed in the universe is assumed to originate from the violation of CP, a lower bound of $3 \times 10^{-27} e$ cm may be obtained for the dipole moment [Ell81]. For a detailed overview we refer to [Tra90].

5.2 EXPERIMENTS TO MEASURE THE ELECTRIC DIPOLE MOMENT OF THE NEUTRON

5.2.1 The principle of the measurements

The principle of the experiments to search for an electric dipole moment of the neutron is based on the phenomenon of magnetic resonance. The experiments involve the determination of the precession frequency (Larmor frequency) of the neutron spin in a weak magnetic field

$$\nu_{\mathrm{L}} = -2d_m B/h \tag{5.11}$$

where d_m denotes the magnetic moment of the neutron (the minus sign is there because the magnetic moment is negative). A strong electrical field $\boldsymbol{E}$ is applied parallel and then antiparallel to the magnetic field $\boldsymbol{B}$. If the neutron has an electric dipole moment $\boldsymbol{d}_n$, an additional precession of the spin occurs and the resonance frequency is slightly shifted. From the point of view of quantum mechanics, the electrical field generates a splitting of the $(m_s = +1/2)$ and $(m_s = -1/2)$ states of the neutron. The frequency shift is

$$\Delta\nu_{\mathrm{L}} = 2d_n E/h. \tag{5.12}$$

When the direction of the $\boldsymbol{E}$ field changes, a shift of $2\Delta\nu_{\mathrm{L}}$ occurs. The resonance frequency of the neutron is

$$\nu_{\mathrm{R}} = -2d_m|\boldsymbol{B}|/h \mp 2d_n|\boldsymbol{E}|/h. \tag{5.13}$$

The minus sign applies for parallel fields, the positive sign for antiparallel positioning of the fields. The frequency shift $\Delta\nu_{\mathrm{L}}$ is extremely small. For an electrical field strength of 25 kV cm^{-1} and a hypothetical dipole moment of $10^{-25} e$ cm, it only amounts to 1.2×10^{-6} Hz.

Two classes of experiment may be distinguished:

(i) measurements with neutron beams;
(ii) measurements with stored neutrons.

The experiments of both classes use so-called neutron mirrors, the operation of which is described in the next section.

5.2.2 The neutron mirror

The passage of slow neutrons through matter may be described by a wave in a medium with a refractive index n [Fer50, Wla59]. If the wavelength of the neutron is large in comparison with the interatomic distances of a solid body, the interaction of the neutron with this body may be described by a neutron potential V averaged over the volume. The presence of this potential leads to the refraction of the neutron wave at the boundary of the body. The medium may be characterized by the specification of a refractive index for neutrons. The size of the refractive index is found by solving the Schrödinger wave equation for free neutrons in the vacuum and for neutrons in the medium. Assuming a small neutron absorption, one defines

$$n = \frac{k}{k_0} = \frac{\lambda_0}{\lambda} \tag{5.14}$$

where λ_0 is the neutron wavelength in vacuum and λ is the wavelength in the solid body. If the values of k and k_0 from the corresponding solutions of the Schrödinger equation are now inserted, it follows that

$$n^2 - 1 = -N\lambda^2 \frac{a}{\pi}. \tag{5.15}$$

a is the coherent scattering amplitude and the interaction cross section for neutron scattering is $\sigma = 4\pi a^2$. N denotes the number of nuclei per cm^3. Because $n^2 - 1 = (n+1)(n-1) \approx 2(n-1)$, the following approximation holds

$$n - 1 = -N\lambda^2 \frac{a}{2\pi}. \tag{5.16}$$

For thermal neutrons we have $\lambda \sim 10^{-8}$ cm and $a = \sqrt{\sigma/4\pi} \sim 10^{-13}$ cm. Taking N to be approximately 6×10^{23} particles per cm^3, it follows that

$$\tilde{\epsilon} = 1 - n \sim 10^{-6}. \tag{5.17}$$

Thus, the refractive index is of the same order of magnitude as for X-rays. Unlike X-rays, for which we always have $n < 1$, in the case of neutrons $\tilde{\epsilon}$ may be either greater than or less than zero. The sign depends on the sign of the scattering length a. For most materials, $a > 0$ and consequently the refractive index n is less than one. In these cases, the medium is optically thinner than the vacuum. On passing into an optically thinner material, i.e. when neutrons pass from the vacuum into a material with $n < 1$, total reflection may occur. If the angle of incidence θ_1 (measured from the perpendicular) is greater than a certain critical angle θ_c, the neutrons cannot pass through the boundary layer and are

totally reflected. The angle θ_c can be determined from Snell's law of refraction (figure 5.3) by requiring $\theta_2 = \pi/2$ for the angle of reflection. It follows that

$$\sin\theta_c = n. \tag{5.18}$$

Clearly, total reflection occurs for $\sin\theta_c > n$. This condition can only be satisfied for $n < 1$, in other words, by materials whose atoms have a positive scattering length for neutrons.

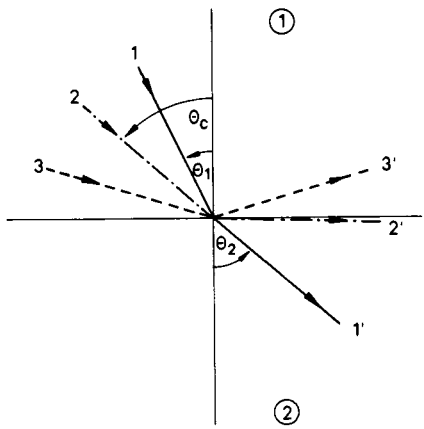


Figure 5.3. Total reflection of neutron rays. Ray 2 is incident below the critical angle of total reflection.

In the above discussion we have neglected the effect of the magnetic field $\boldsymbol{B}$. The correct expression for the refractive index is

$$n^2 = 1 - \frac{\lambda^2 Na}{\pi} \pm \frac{d_m B}{T}. \tag{5.19}$$

T denotes the kinetic energy of the neutron and the sign depends on the relative orientation of the neutron spin and the magnetic field. Cold neutrons with velocities of $v \approx 80$ m s^{-1} are totally reflected for angles of incidence $\theta > 85°$. This makes it possible to construct neutron conductors with a very small divergence of the neutron beam. Since the difference between the refractive index and 1 increases with λ^2, ultracold neutrons with $v < 6$ m s^{-1} ($\lambda \sim 670$ Å) may be totally reflected at all angles of incidence. Thus, it is possible to store such neutrons for periods of over 100 s in a so-called neutron bottle.

Since the refractive index depends on the orientation of the neutron spin, by an appropriate choice of the material and the magnetic field it is possible to ensure that neutrons with a certain spin orientation are totally reflected while those with the opposite polarization are not. This property is used, amongst other

things, to generate polarized neutrons and to analyse the degree of polarization. For typical velocities of $v = 100$ m s^{-1} and angle of incidence 88° from the perpendicular a beam polarization of 70% is obtained. Ultracold neutrons may be easily polarized by introducing a thin magnetic sheet into the beam. One spin direction will be reflected and the other allowed through.

5.2.3 Experiments on neutron beams

The first experiments relating to the electric dipole moment of the neutron were carried out using neutron beams. The apparatus is essentially similar to Rabi's molecular beam apparatus. A reactor forms the source of the neutron beam; the polarization of the neutrons and the subsequent analysis of the spin direction are implemented by transmission through a magnetized sheet of iron.

One of the first such experiments was carried out by Smith *et al* [Smi57]. The greatest sensitivity has been achieved by measurements at the research reactor of the Institut Laue-Langevin (ILL) in Grenoble [Dre77, 78, Ram82, 90]. The principle of these measurements is very simple. The neutrons are led from the moderator in the reactor via a neutron conductor to the polarizer. After passing through the spectrometer, the neutrons meet an analysing magnet and are finally detected in a neutron detector. A glass scintillation counter doped with ^{6}Li is a suitable detector. The transmitted intensity I is a maximum for neutrons which are not subject to depolarization in the spectrometer.

The beam from the reactor is polarized by transmission through magnetized iron. The beam prepared in this way enters a weak constant magnetic field, so that the spin precesses around the direction of the magnetic field. An alternating magnetic field induces transitions of the neutron from one spin direction to the opposite direction. By these spin flip transitions the beam is partially depolarized and the intensity let through by the analysing sheet changes. Typical resonance curves such as those shown in figure 5.4 are obtained.

Two coils are used to generate the alternating field, in order to generate interference phenomena. These lead to several minima and maxima of $I(\nu)$ and give a narrower full width of half maximum [Ram80]. If the frequency of the oscillating field is chosen so that the change in intensity for small frequency shifts is particularly large ($\mathrm{d}I/\mathrm{d}\nu$ as large as possible), a possible electric dipole moment may be detected by applying a strong electrical field. For a fixed oscillator frequency the reversal of the $\boldsymbol{E}$ field leads to a shift of $\Delta\nu_L = 4d_nE/h$ in the precession frequency, which should be apparent in a change in the intensity let through by the analyser. If the neutrons do not have an electric dipole moment, the electrical field does not affect the spin precession or the counting rate in the detector.

The strength of the electrical field in the direction of the constant $\boldsymbol{B}$ field amounted to 100 kV cm^{-1} in the most sensitive measurements, the static

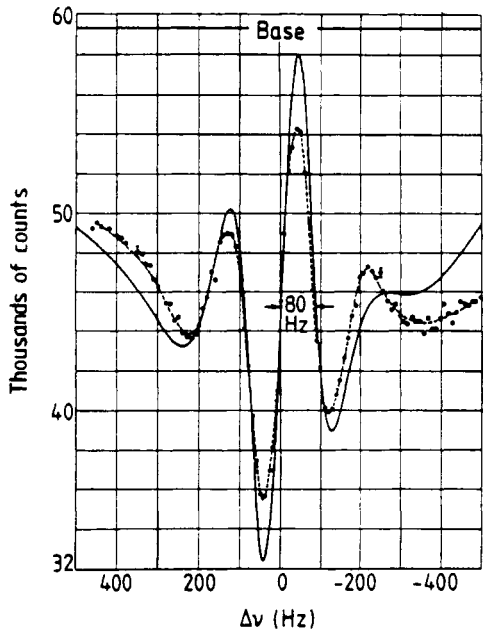


Figure 5.4. Typical magnetic resonance curve with a phase shift of $\pi/2$ between two oscillating fields. The full curve shows the calculated transition probability for a Maxwell–Boltzmann distribution at $T = 1$ K. The deviations of the experimental (broken) curve far from the resonance can be attributed to deviations of the beam velocity from the Maxwell–Boltzmann distribution (from [Ram82]).

magnetic field was of the order of 17 G and the polarization of the neutron beam reached 89%. No dependence of the counting rate on the relative orientation of the electrical and the static magnetic fields was detected. The result of the most sensitive measurement with a neutron beam is [Dre78]

$$d_n = +(0.4 \pm 1.5) \times 10^{-24} e \text{ cm} \tag{5.20a}$$

or

$$|d_n| < 3 \times 10^{-24} e \text{ cm}. \tag{5.20b}$$

In this method, a systematic error results from the difficulty of generating fields $\boldsymbol{E}$ and $\boldsymbol{B}$ which are exactly parallel. A particle moving with velocity $\boldsymbol{v}$ through an electrical field is subject to a magnetic field of magnitude $\boldsymbol{E} \times \boldsymbol{v}/c$. When $\boldsymbol{E}$ and $\boldsymbol{B}$ are exactly parallel the field $\boldsymbol{E} \times \boldsymbol{v}/c$ is perpendicular to $\boldsymbol{B}$ and does not alter the precession frequency, if the direction of $\boldsymbol{E}$ is altered, however, the absolute value is conserved. On the other hand, a small component $E_\perp$ perpendicular to $\boldsymbol{B}$ induces an additional magnetic field $\Delta B = vE_\perp/c$ in the

direction of $\boldsymbol{B}$. This effect on the Larmor frequency is scarcely distinguishable from the interaction of an electric dipole moment with the electrical field.

In addition, $\mathrm{d}I/\mathrm{d}\nu$ is proportional to the time the neutrons spend in the alternating field between the two RF coils. This leads to a requirement for the greatest possible distance between the coils and for very slow cold neutrons. For ultracold neutrons ($T = 0.002$ K, $v = 6$ m s^{-1}, $\lambda \sim 670$ Å) the critical angle for total reflection may become 0°. The neutrons are totally reflected for all angles of incidence and may be trapped in a bottle. Using these trapped neutrons one obtains very much longer observation times and thus an increased sensitivity of the experiment.

5.2.4 Experiments with stored neutrons

The current generation of experiments relating to the electric dipole moment of the neutron uses ultracold neutrons with velocities of approximately 5 m s^{-1}, which are stored in an evacuated bottle. This method has two essential advantages, namely a longer observation time and a smaller induced magnetic field $\boldsymbol{E} \times \boldsymbol{v}/c$.

Such experiments have been carried out at the research reactors of the ILL in Grenoble [Pen84, Smi90a] and the B P Konstantinov Institute for Nuclear Research in Leningrad [Alt81, 86, Lob84]. The early measurements with ultracold, stored neutrons [Alt81, Lob84, Pen84] gave an upper bound of

$$d_n < 6 \times 10^{-25} e \text{ cm} \tag{5.21}$$

for the electric dipole moment. The result of the Leningrad group provided the first indication of a finite value [Alt86]

$$d_n = (-14 \pm 6) \times 10^{-26} e \text{ cm} \tag{5.22a}$$

which, however, was interpreted as an upper bound

$$|d_n| < 26 \times 10^{-26} e \text{ cm} \quad (90\%). \tag{5.22b}$$

In what follows, we shall describe the most sensitive experiment to date, which was carried out at the ILL in Grenoble [Tho89a, Ram90, Smi90a] (see figure 5.5). The apparatus is largely similar to that of [Pen84] although it was possible to increase the neutron flux by more than two orders of magnitude. The neutrons from the moderator (liquid deuterium) of the ILL reactor reach a 1 μm thick polarization sheet consisting of an Fe–Co alloy via a nickel neutron conductor.

The polarized particles ultimately fill the neutron bottle, which may be used to store neutrons up to speeds of 6.9 m s^{-1}. The storage volume of 5 l is surrounded by a mu-metallic shield to screen the Earth's magnetic field

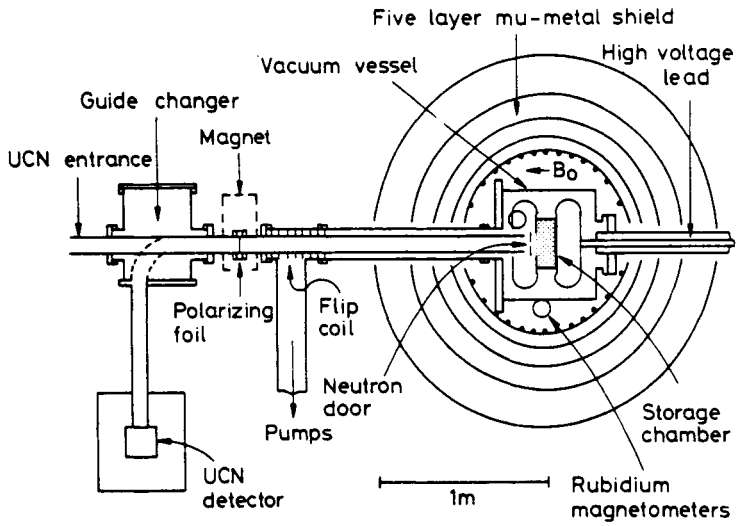


Figure 5.5. Experiment to measure the electrical dipole moment of the neutron at the Institut Laue-Langevin in Grenoble. Ultracold neutrons are stored for approximately 70 s in a magnetic field of 1 μT and an electrical field of order of magnitude 1 MV m^{-1} (from [Smi90a]).

(screening factor of $\sim 10^5$). There is a weak static magnetic field $B_0 = 1$ μT parallel to the axis of the bottle. The resonance frequency of the neutrons is 30 Hz. The storage volume is filled for approximately 100 s. The density of the polarized neutrons reaches 10 cm^{-3}. The neutron spins are oriented in the direction of the static magnetic field $\boldsymbol{B}_0$. After the neutrons have achieved an isotropic velocity distribution, the spin is rotated in the plane perpendicular to $\boldsymbol{B}_0$ by a magnetic field oriented perpendicular to $\boldsymbol{B}_0$ and oscillating at 30 Hz. This first so-called Ramsey pulse lasts approximately 4 s. After this, the neutrons precess around the static field for 70 s. Then, a 4 s long alternating field is applied (second Ramsey pulse) which is in phase with the first. The spin state depends on the extent to which the spin precession and the oscillating field have become out of phase (see [Ram80]).

If the angular momentum is initially parallel to the static magnetic field ($\phi = 0$) it is possible to select an alternating field perpendicular to this, which rotates the spin by 90° ($\phi = \pi/2$). In an area without an alternating field, the spin only performs a precession movement with the Larmor frequency of B_0. In the second alternating field, which is applied for the same length of time as the first, there results another angular momentum which attempts to change the angle ϕ. In what follows, we shall assume that the two oscillating fields are phase coherent.

When the frequency of the oscillating field coincides exactly with the Larmor frequency there is no phase shift between the spin and the second pulse. The second pulse, which is identical in length and strength to the first Ramsey pulse, has the same effect as the first; it rotates the spin by a further 90° so that $\phi = \pi$. This corresponds to a complete spin flip. However, if the field frequency and the Larmor frequency are slightly different, a complete spin reversal will not be achieved. If the relative phase angle between the second pulse and the spin is 180° the field rotates the spin to its original position ($\phi = 0$).

After the second Ramsey pulse the neutron bottle is opened and the neutrons in the correct spin state may pass unhindered through the polarization sheet which operates as analyser, and are detected in the detector. The spin of the remaining neutrons is then reversed using a spin flip coil so that these neutrons also pass through the sheet and can be detected. Figure 5.6 shows the resonance curve obtained by [Smi90a].

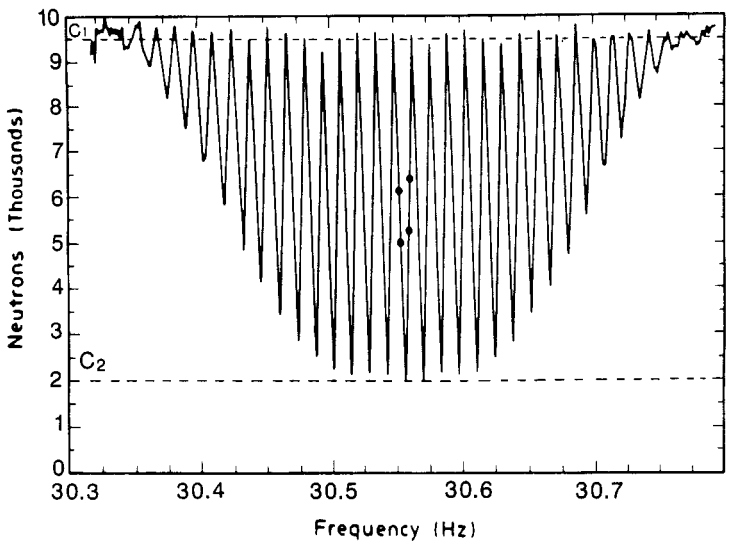


Figure 5.6. Measured magnetic resonance curve. The four points labelled were used for the evaluation (from [Smi90a]).

To measure the electric dipole moment, one now chooses a frequency with maximal gradient in figure 5.6, where an electrical field is applied parallel and then antiparallel to $\boldsymbol{B}_0$. In this experiment, the electrical field had strength up to 16 kV cm^{-1}. Overall, the reversal of the polarity of the electrical field shifts the resonance frequency by

$$2h\Delta\nu_{\mathrm{L}} = 4d_n E. \tag{5.23}$$

The dipole moment d_n is determined from the difference ΔN between the number of events with parallel and antiparallel fields

$$d_n = \frac{h\Delta N}{4ES} \tag{5.24}$$

where $S = \mathrm{d}N/\mathrm{d}\nu$ is the gradient of the resonance curve at the point of evaluation. The result of this experiment is [Smi90a]

$$d_n = -(3 \pm 5) \times 10^{-26} e \text{ cm}. \tag{5.25a}$$

This zero result implies an upper bound of

$$|d_n| < 12 \times 10^{-26} e \text{ cm} \quad (95\%). \tag{5.25b}$$

The bound of (5.25*b*) is in the area expected in right–left symmetric models (see section 1.3). Thus, an improvement in the experimental sensitivity could provide important information about the validity of these theories. Currently, the experiment at the ILL is being extended to a neutron bottle with a storage volume of approximately 60 l.

5.2.5 The θ problem

In section 1.3.2, we discussed the fact that there exists one reason for the existence of an electric dipole moment of the neutron inherent in QCD (see (1.64)), which gives

$$d_n \sim 4 \times 10^{-16} \theta e \text{ cm}. \tag{5.26}$$

From the current experimental bounds for d_n one may then derive

$$\theta < 3 \times 10^{-10}. \tag{5.27}$$

Thus, the angle θ, which *a priori* should be arbitrary must actually be very small; however, this is implausible if one assumes that θ is a purely random phase. This fact is called the θ *problem* or the *strong CP problem*. The solution of this puzzle is probably beyond the standard model (compare also section 12.2 on Kaluza–Klein theories). One possible explanation is provided by the introduction of additional Higgs fields and an additional chiral symmetry [Pec77], which make it possible to 'rotate θ away'. However, this Peccei–Quinn model requires the existence of a new pseudoscalar particle, the axion [Wei78, Wil78], which has not yet been detected despite an intensive search [Raf90, Tur90]. These axions are under discussion as possible candidates for dark matter (see chapter 9). For further interpretations of an electric dipole moment of the neutron in beyond-standard models, see [Gav90].

5.2.6 The electric dipole moment of other particles

Electric dipole moments of particles other than the neutron are also sought [Ram82]. However, because of its neutrality, the neutron is particularly suitable for such measurements. In addition, theoretical models generally give a larger dipole moment for the neutron than, for example, for the electron, for which the experiments have achieved a similar accuracy

$$d_e = (-1.5 \pm 5.5 \pm 1.5) \times 10^{-26} e \text{ cm} \qquad \text{[Mur89]} \qquad (5.28a)$$

and

$$d_e = (-2.7 \pm 8.3) \times 10^{-27} e \text{ cm} \qquad \text{[Abd90].} \qquad (5.28b)$$

According to the Kobayashi–Maskawa model, the electric dipole moment of the electron should not be greater than approximately $10^{-37} e$ cm. The most recent experiment carried out at Berkeley [Abd90] determined the energy of an electron in an electrical field. Valence electrons in heavy atoms (in this particular case ^{205}Tl) were used rather than free electrons. The measurements were carried out using an atomic beam resonance method with separate alternating fields. The atomic dipole moment of ^{205}Tl was determined to be

$$d(^{205}\text{Tl}) = (1.6 \pm 5.0) \times 10^{-24} e \text{ cm} \qquad (5.29a)$$

from which the value given for the electron in (5.28*b*) follows. Results of a similar sensitivity have been obtained for the electric dipole moments of other heavy atoms [Vol84, Lam87]

$$d(^{129}\text{Xe}) = -(0.3 \pm 1.1) \times 10^{-26} e \text{ cm} \qquad (5.29b)$$

$$d(^{199}\text{Hg}) = (0.7 \pm 1.5) \times 10^{-26} e \text{ cm.} \qquad (5.29c)$$

5.3 NEUTRON–ANTINEUTRON OSCILLATIONS

5.3.1 Introduction

As we saw in chapter 1, symmetries and conservation laws play an important role in modern gauge theories. The conservation theorems for the baryon and lepton number have a special position since they are not based on any known symmetry principle. These conservation theorems can only be justified phenomenologically and thus it is conjectured that they are only valid within the bounds of current experimental limits. There are a number of reasons for supposing that the baryon number symmetry is not an exact symmetry (see chapter 3). As early as 1967 Sacharov remarked that the origin of matter in the big bang model presupposed three important conditions [Sac67]: (i) violation of the conservation of baryon number; (ii) CP violation; and (iii) thermal non-equilibrium.

In addition, violations of the separate conservation theorems for B and L occur within the framework of the standard model. The underlying Lagrange density of the electroweak theory conserves both the baryon number and the lepton number, but the so-called triangle anomalies break the separate B and L symmetries, while the $(B–L)$ symmetry is conserved. However, the B violating amplitude introduced in this way is extremely small ($\sim e^{-4\pi/\alpha_w}$, where α_w is the coupling constant of the weak interaction).

In fact, practically all attempts to formulate a grand unification theory inevitably lead to a violation of the conservation of the lepton number and the baryon number and thus, in particular, permit the decay of the proton. A violation of the baryon number B might also manifest itself through the $(\Delta B = 2)$ process of neutron–antineutron oscillation. There are essentially two classes of B violating process:

(i) $\Delta B = 1$ (e.g. nucleon decay);
(ii) $\Delta B = 2$ (e.g. $n\bar{n}$ oscillations).

In the minimal SU(5) model, in which the baryon number and the lepton number are no longer conserved quantities, the difference $(B - L)$ is conserved, so that proton decay is permitted ($\Delta B = \Delta L = 1$), but the process of $n\bar{n}$ oscillation is forbidden since $\Delta B = 2$ and $\Delta L = 0$. Experiments on the stability of the nucleon and accurate measurements of the Weinberg angle, together with a number of other hints (see section 2.2.1), show that an extension of the minimal SU(5) model is probably required. In right–left symmetric GUT models such as the SO(10) model, processes with $\Delta B = 2$ and $\Delta L = 0$ are allowed. In the Pati–Salam model with a $SU(4)_{ec}\otimes SU(2)_L\otimes SU(2)_R$ symmetry, which is based on the SO(10) group with appropriate symmetry breaking, these processes should be much more frequent than the nucleon decay [Moh80]. The group $SU(4)_{ec}$ contains the $U(1)_{B-L}\otimes SU(3)_c$ as a subgroup. $B - L$ plays the role of a fourth colour. The $SU(4)_{ec}$ transformations include an extension of the strong colour interaction by this fourth colour ('extended colour', see [Moh86a]).

Nucleon decay was discussed in chapter 4. The Hamiltonian operator which is defined by the $SU(3)_c\otimes SU(2)_L\otimes U(1)_Y$ symmetry to describe $(\Delta B = 1)$ processes is given schematically by [Moh89]

$$H_{\Delta B=1} \sim \frac{1}{m_X^2} uude^- \tag{5.30}$$

where the mass of the X boson is of the order of magnitude 10^{15} GeV/c^2. $n\bar{n}$ oscillations, on the other hand, are described by a Hamiltonian operator with the following structure

$$H_{\Delta B=2} \sim \frac{1}{m_I^5} uddudd. \tag{5.31}$$

This operator not only induces $n\bar{n}$ oscillations, but is also responsible for the nuclear instability via nucleon decays of the form

$$p + n \rightarrow \text{pions}. \tag{5.32}$$

The experimental bound for the lifetime due to these non-leptonic decays is approximately 10^{31} years. This gives a mass scale of

$$m_I \sim m_X^{2/5} \sim 10^6 \text{ GeV}/c^2. \tag{5.33}$$

Unlike the decay of the proton with a very large mass scale in the area of the GUT scale, studies of ($\Delta B = 2$) processes thus yield information about new physics beyond the standard model in the area of an intermediate mass scale.

The various unification theories predict oscillation periods $\tau_{n\bar{n}} \sim 10^6$–$10^{10}$ s for the energy scale of $m_I \sim 10^2$–10^3 TeV/c^2. Above these energies, the existence of a local $(B - L)$ symmetry is expected. A detailed discussion of ($\Delta B = 2$) transitions in the context of GUT models may be found in [Moh89]. Table 5.2 summarizes the predictions of various models.

Table 5.2. GUT predictions for the period $\tau_{n\bar{n}}$ of the $n\bar{n}$-oscillations (from [Moh89]).

GUT model	$\tau_{n\bar{n}} = 10^6 - 10^{10}$ s
Standard model	No
SU(5)	No
$SU(2)_L \otimes SU(2)_R \otimes SU(4)_{ec}$	Yes
SO(10)	Yes
E_6	No
SUSY-$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$	Yes†
SUSY-right-left-symmetric with E_6-type spectrum	Yes

† Too fast, without fine tuning of the parameters.

Historically, $n\bar{n}$ oscillations were first discussed by Kuzmin in 1970 [Kuz70]. There are two experimental approaches to the search for $n\bar{n}$ transitions:

- Experiments on free neutrons. A beam of cold neutrons from a reactor is directed onto a target. The antineutrons which form along the flight path annihilate with neutrons or protons in the target into a number of pions.
- Experiments on bound neutrons. Underground experiments to search for the decay of the nucleon, as described in chapter 4, are also sensitive to $n\bar{n}$ oscillations. When a bound neutron in the nucleus turns into an antineutron, the latter annihilates with another bound nucleon into a number of pions. The only background sources in these experiments are interactions with atmospheric neutrinos in the detector.

Before we describe the experiments, we shall discuss the phenomenology of neutron–antineutron oscillations in more detail in the next section.

5.3.2 The phenomenology of the $n\overline{n}$ oscillations

As we have seen, ($\Delta B = 2$) transitions in the nucleus and, in particular, processes of the form $n + p \rightarrow$ pions, may occur in the framework of grand unification theories. According to this hypothesis, a mixing of neutron and antineutron states occurs, which is characterized by a mass splitting

$$\delta mc^2 = \langle \overline{n} | H_{n\overline{n}} | n \rangle. \tag{5.34}$$

The $n\overline{n}$ oscillations may be described phenomenologically by the following effective Hamiltonian

$$H_{n\overline{n}} = \delta mc^2 \int \mathrm{d}^3x \left(\overline{\psi}_n^c \psi_n + \overline{\psi}_n \psi_n^c \right). \tag{5.35}$$

Here the quark states of (5.31) have been replaced by the neutron field ($n = udd$). The description of this phenomenon corresponds completely to the strangeness oscillations ($\Delta S = 2$) in the $K^0\overline{K^0}$ system (see section 1.3.1).

Because of the mixing of states, the temporal development of this two-state system is described by the following equations

$$\mathrm{i}\hbar \frac{\partial}{\partial t} |n\rangle = m_n c^2 |n\rangle + \delta mc^2 |\overline{n}\rangle \tag{5.36a}$$

$$\mathrm{i}\hbar \frac{\partial}{\partial t} |\overline{n}\rangle = m_{\overline{n}} c^2 |\overline{n}\rangle + \delta mc^2 |n\rangle. \tag{5.36b}$$

In somewhat abbreviated notation this system of equation becomes

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \psi = \mathcal{M} \psi \tag{5.37}$$

where

$$\psi = \begin{pmatrix} |n\rangle \\ |\overline{n}\rangle \end{pmatrix} \quad \text{and} \quad \mathcal{M} = \begin{pmatrix} m_n c^2 & \delta mc^2 \\ \delta mc^2 & m_{\overline{n}} c^2 \end{pmatrix}. \tag{5.38}$$

The off-diagonal terms of $\mathcal{M}$ denote the transition energy between neutrons and antineutrons. Because of the CPT theorem, the diagonal elements must be equal ($m_n = m_{\overline{n}} = m$). In the following, we shall also assume CP invariance, so that the off-diagonal elements may also be assumed to be the same.

The eigenstates of the matrix $\mathcal{M}$ are

$$|n_{1/2}\rangle = \frac{1}{\sqrt{2}} (|n\rangle \pm |\overline{n}\rangle) \tag{5.39a}$$

with the mass eigenvalues

$$m_{1/2} = m \pm \delta m. \tag{5.39b}$$

In the interaction-free space, the temporal development of the eigenstates is described as follows

$$|n_1(t)\rangle = |n_1(0)\rangle \exp\left(-\mathrm{i}\frac{m_1c^2}{\hbar}t\right)\exp\left(-\frac{\Gamma}{2\hbar}t\right) \tag{5.40a}$$

$$|n_2(t)\rangle = |n_2(0)\rangle \exp\left(-\mathrm{i}\frac{m_2c^2}{\hbar}t\right)\exp\left(-\frac{\Gamma}{2\hbar}t\right). \tag{5.40b}$$

The exponential factor $\exp(-\Gamma t/2\hbar)$ describes the β decay of the neutron. It is assumed that the decay width

$$\Gamma = \hbar/\tau_n \tag{5.41}$$

is the same in both cases. Strictly speaking, the formula given in (5.40) is true for the rest system in which the energy corresponds to the rest mass $m_{1/2}c^2$. In this case, the time constant τ_n is the true lifetime of the neutron. The states $|n(t)\rangle$ and $|\overline{n}(t)\rangle$ are given by

$$|n(t)\rangle = \frac{1}{\sqrt{2}}(|n_1(t)\rangle + |n_2(t)\rangle) \tag{5.42a}$$

$$|\overline{n}(t)\rangle = \frac{1}{\sqrt{2}}(|n_1(t)\rangle - |n_2(t)\rangle). \tag{5.42b}$$

Let us assume that at time $t = 0$ there is a pure neutron beam from a reactor, whence $|n_1(0)\rangle = |n_2(0)\rangle = 1/\sqrt{2}$. At time t, the amplitude of the neutron beam is given by

$$|n(t)\rangle = \frac{1}{2}\exp\left(-\frac{\Gamma}{2\hbar}t\right)\left[\exp\left(-\mathrm{i}\frac{m_1c^2}{\hbar}t\right) + \exp\left(-\mathrm{i}\frac{m_2c^2}{\hbar}t\right)\right]. \tag{5.43}$$

Thus, the probability that an original neutron is still found to be a neutron at time t is

$$\begin{aligned} P_{nn}(t) &= \langle n(t)|n(t)\rangle \\ &= \frac{1}{4}\exp\left(-\frac{\Gamma}{\hbar}t\right)\left[2 + \exp\left(-\mathrm{i}\frac{(m_1 - m_2)c^2}{\hbar}t\right)\right. \\ &\quad \left. + \exp\left(\mathrm{i}\frac{(m_1 - m_2)c^2}{\hbar}t\right)\right] \\ &= \frac{1}{2}\exp\left(-\frac{\Gamma}{\hbar}t\right)[1 + \cos(2\delta mc^2t/\hbar)]. \end{aligned} \tag{5.44}$$

The probability that an original neutron is detected at time t as an antineutron is given by

$$\begin{aligned} P_{n\bar{n}}(t) &= \exp\left(-\frac{\Gamma}{\hbar}t\right) - P_{nn}(t) \\ &= \frac{1}{2}\exp\left(-\frac{\Gamma}{\hbar}t\right)[1-\cos(2\delta mc^2 t/\hbar)] \\ &= \exp\left(-\frac{\Gamma}{\hbar}t\right)\sin^2(\delta mc^2 t/\hbar). \end{aligned} \tag{5.45}$$

Similarly to the case of the $K^0\overline{K^0}$ system, the state mixing leads to state oscillations with time, and thus to transitions between neutrons and antineutrons. The oscillation period is

$$\tau_{n\bar{n}} = \frac{\hbar}{\delta mc^2}. \tag{5.46}$$

The number of antineutrons $\overline{N}$ arising from a beam of N original neutrons is

$$\overline{N} = N\exp\left(-\frac{\Gamma}{\hbar}t\right)\sin^2\left(\frac{t}{\tau_{n\bar{n}}}\right) \tag{5.47a}$$

$$\approx N\exp\left(-\frac{\Gamma}{\hbar}t\right)\left(\frac{t}{\tau_{n\bar{n}}}\right)^2. \tag{5.47b}$$

The approximation in (5.47*b*) is valid, since the observation time is always limited by the lifetime of the free neutron ($\tau_n \sim 10^3$ s) and very much larger oscillation periods are expected, so that $t \ll \tau_{n\bar{n}}$ is satisfied in a very good approximation. It is immediately clear that experiments to search for $n\bar{n}$ oscillations require very large neutron fluxes. For an oscillation time of 10^7 s, approximately 10^{16} neutrons are required in order to generate one neutron–antineutron transition in 0.1 s.

In nature there are no really free neutrons, since interactions with external fields such as the Earth's magnetic field always occur. In a magnetic field neutrons and antineutrons are no longer energetically degenerate since their magnetic moments have different signs, which leads to an energy splitting between n and $\bar{n}$ states. In the Earth's magnetic field the shift in the energy levels is $\Delta E = \pm 2d_m B_0 \simeq 10^{-12}$ eV. This leads to a strong suppression of $n\bar{n}$ oscillations, since, for example, an oscillation time of 10^7 s only corresponds to a transition energy of $\delta mc^2 \sim 10^{-22}$ eV [Moh89]. Thus, in what follows, we have to consider $n\bar{n}$ oscillations in external fields.

The equation of motion in an external field has the same form as (5.37), except that the matrix $\mathcal{M}$ has a different structure

$$\mathcal{M} = \begin{pmatrix} M_1c^2 & \delta mc^2 \\ \delta mc^2 & M_2c^2 \end{pmatrix}. \tag{5.48}$$

The requirement, based on the CPT theorem, that the diagonal elements should be the same is no longer valid, since the interaction with the external fields removes the degeneracy between neutrons and antineutrons. The effective n and $\overline{n}$ masses are

$$M_1c^2 = m_nc^2 + V_n \tag{5.49a}$$

$$M_2c^2 = m_nc^2 + V_{\overline{n}} \tag{5.49b}$$

where m_n denotes the (anti)neutron mass and V_n and $V_{\overline{n}}$ are external potentials for neutrons and antineutrons, respectively. These may be due to external electromagnetic fields or the nuclear potential for bound neutrons. In the case of neutrons bound in the nucleus, V_n and $V_{\overline{n}}$ are very different

$$V_n = U \qquad V_{\overline{n}} = \overline{U} - \mathrm{i}W. \tag{5.50}$$

The potential for bound neutrons is complex ('optical potential', see e.g. [Mar70]), since antineutrons may annihilate in the nucleus. This latter process is described by the imaginary component.

The eigenstates may now be described as follows

$$|n_1\rangle = |n\rangle \cos\theta + |\overline{n}\rangle \sin\theta \tag{5.51a}$$

$$|n_2\rangle = -|n\rangle \sin\theta + |\overline{n}\rangle \cos\theta. \tag{5.51b}$$

The mass eigenvalues are

$$m_{1/2}c^2 = \tfrac{1}{2}(M_1 - M_2)c^2 \pm \sqrt{\delta m^2c^4 + \Delta E^2} \tag{5.52}$$

where

$$\tan 2\theta = \frac{\delta mc^2}{\Delta E} \qquad \Delta E = \tfrac{1}{2}(V_n - V_{\overline{n}}). \tag{5.53}$$

The temporal development of the neutron state is described by the equation

$$|n(t)\rangle = |n_1(0)\rangle \exp\left(-\mathrm{i}\frac{m_1c^2}{\hbar}\right)\cos\theta + |n_2(0)\rangle \exp\left(-i\frac{m_2c^2}{\hbar}\right)\sin\theta. \tag{5.54}$$

Here, we have neglected the exponential factor for the β decay of the neutron. If one begins with a pure n beam at time $t = 0$, the transition probability at time t is

$$P_{n\overline{n}}(t) = \left(\frac{\delta m}{\Delta M}\right)^2 \sin^2(\Delta Mc^2t/\hbar). \tag{5.55}$$

The probability oscillates with the amplitude

$$A = \left(\frac{\delta m}{\Delta M}\right)^2 \tag{5.56}$$

and the period

$$\overline{\tau}_{n\bar{n}} = \frac{\hbar}{\Delta M c^2} \tag{5.57}$$

where

$$\Delta M c^2 = \tfrac{1}{2}(m_1 - m_2)c^2 = \sqrt{\delta m^2 c^4 + \Delta E^2}. \tag{5.58}$$

In the case of free neutrons, ΔE vanishes and (5.55) becomes (5.45). It is clear that $n\bar{n}$ oscillations are strongly suppressed, when ΔM is large in comparison with δm.

5.3.3 Experiments on $n\bar{n}$ oscillations

As mentioned in the introduction, there are two methods for searching for $n\bar{n}$ transitions. One approach uses neutrons bound in the nucleus, the second free neutrons.

5.3.3.1 Experiments with bound neutrons

Underground detectors to search for nucleon decay are also sensitive to $n\bar{n}$ oscillations [Dov83]. When a bound neutron in the nucleus changes into an antineutron the latter annihilates with another nucleon in the same nucleus

$$(A, Z) \rightarrow (A-1, Z, \bar{n}) \rightarrow (A-2, Z) + \text{ pions}. \tag{5.59}$$

This typically gives rise to five pions with a total energy of approximately 2 GeV. The $n\bar{n}$ annihilation may thus be detected via the following reactions

$$\bar{n} + n \rightarrow \text{ pions } \rightarrow \text{ muons} \tag{5.60a}$$

$$\bar{n} + p \rightarrow \text{ pions } + \text{ muons}. \tag{5.60b}$$

The Kamiokande collaboration's Cerenkov water counter [Tak86] and the Fréjus iron detector [Ber90a] have been used to search for such events in ^{16}O and ^{56}Fe, respectively.

The difficulty of this method is a result of the fact that the energy shift ΔE due to the nuclear potential has very large values of around $\Delta E = 100$–500 MeV [Moh89] and thus the oscillations are very strongly suppressed. For $\delta m c^2 \approx 10^{-22}$ eV, the amplitude of the oscillations is

$$A = \left(\frac{\delta m}{\Delta M}\right)^2 \sim \left(\frac{\delta m c^2}{\Delta E}\right)^2 \sim 10^{-60}. \tag{5.61}$$

Assuming that $\Delta E\, t \gg \hbar$, the function $\sin(\Delta M c^2 t/\hbar)$ oscillates very rapidly. Thus, the average probability of finding an antineutron is

$$\langle P_{n\bar{n}} \rangle = \frac{1}{2}\left(\frac{\delta m c^2}{\Delta E}\right)^2. \tag{5.62}$$

This gives an annihilation rate which is constant with time

$$T_{n\bar{n}}^{-1} \sim \delta m^2. \tag{5.63}$$

Comparison with (5.46) gives

$$\tau_{n\bar{n}} = \sqrt{T_R T_{n\bar{n}}}. \tag{5.64}$$

T_R is a typical period in nuclear physics, i.e. of the order of magnitude 10^{-23} s.

Thus, the measurement of nuclear stability makes it in principle possible to determine the $n\bar{n}$ oscillation time. Certain uncertainties arise from the fact that the parameter T_R must be determined by nuclear structure calculations.

No annihilation events have yet been detected. The Kamiokande collaboration [Tak86] gives a bound of

$$T_{n\bar{n}}(^{16}\mathrm{O}) > 4.3 \times 10^{31} \text{ years} \quad (90\% \text{ c.l.}) \tag{5.65}$$

for the lifetime of the neutron in ^{16}O, which corresponds to a lower bound of

$$\tau_{n\bar{n}} > 1.2 \times 10^8 \text{ s} \tag{5.66}$$

for the oscillation period. Measurements on the Fréjus iron detector [Ber90a] give a bound

$$T_{n\bar{n}}(^{56}\mathrm{Fe}) > 6.5 \times 10^{31} \text{ years} \quad (90\% \text{ c.l.}) \tag{5.67}$$

for the lifetime of the neutron in ^{56}Fe, which leads also to

$$\tau_{n\bar{n}} > 1.2 \times 10^8 \text{ s}. \tag{5.68}$$

A more recent calculation of the magnitude of T_R for the relevant nuclides using a realistic nuclear potential, the Paris potential, gives the following experimental bounds on the lifetimes [Alb91]

$$\tau_{n\bar{n}}(^{16}\mathrm{O}) > (0.7\text{–}0.9) \times 10^8 \text{ s} \tag{5.69a}$$

$$\tau_{n\bar{n}}(^{56}\mathrm{Fe}) > (0.8\text{–}1.0) \times 10^8 \text{ s}. \tag{5.69b}$$

The bounds which may be achieved in future with new underground detectors will be only slightly better than those of present experiments. A sensitivity to nuclear instabilities up to $T_{n\bar{n}} \sim 4 \times 10^{32}$ years, corresponding to an oscillation time of $(2\text{–}3) \times 10^8$ s, is expected from the Superkamiokande experiment (see chapter 4) using a 50 000-tonne Cerenkov water detector [Alb91]. The ICARUS project (see chapter 4) which is now being planned will achieve a corresponding sensitivity [Alb91].

5.3.3.2 Experiments with free neutrons

Experiments with free neutrons permit direct measurement of the magnitude of $\tau_{n\bar{n}}$. The interpretation does not depend on nuclear structure calculations.

As previously mentioned, there are no exactly free neutrons in nature, since external fields such as the Earth's magnetic field are always present. This means that ΔM is very much larger than δm, whence the oscillations are suppressed. According to (5.55) care must be taken to ensure that the energy gap between n and $\bar{n}$ states, $2\Delta E$, is as small as possible.

Neutrons essentially behave like free neutrons when the condition

$$\Delta M c^2 t \ll \hbar \tag{5.70}$$

is satisfied (quasi-free neutrons, [Moh80]), since then (5.47*b*) again follows from (5.55) for $t \ll \tau_{n\bar{n}}$, i.e. $P_{n\bar{n}}(t)$ increases quadratically with the time t. In experiments with neutron beams, the energy splitting between n and $\bar{n}$ states is determined by the interaction of the magnetic moments with the Earth's magnetic field $B_0 \approx 40\ \mu$T

$$\Delta E \approx 10^{-12}\ \text{eV}. \tag{5.71}$$

This would result in a suppression of the oscillation amplitude by a factor $(\delta m c^2/\Delta E)^2 \simeq 10^{-20}$. However, if the observation time is chosen to be sufficiently short, the condition (5.70) can be satisfied and the number of antineutrons will be independent of the strength of the magnetic interaction. In order to satisfy the condition for quasi-free neutrons with as large observation times as possible, the Earth's magnetic field must be well screened. The interaction energy with the residual field $\tilde{B}_0$ must be small in comparison with the energy uncertainty of the system

$$2d_m \tilde{B}_0 \ll \hbar/t. \tag{5.72}$$

For a given residual field the flight time must be set according to this requirement (for a discussion of the quasi-free condition see [Bit85]).

Experiments to search for $n\bar{n}$ oscillations must satisfy the following important requirements. First, good screening against the external fields is required in order to satisfy the quasi-free condition. The remaining requirements are easy to derive using (5.47*b*). Both a very large neutron flux and a long flight time t are required, at the same time satisfying inequality (5.70). Long flight times are obtained using cold neutrons with low velocities ($t = v/L$) over a long flight path L.

The principle of such measurements is very simple. A beam of cold, quasi-free neutrons from a reactor passes along a path L screened from external fields and then meets a carbon foil. The antineutrons formed during the flight time t annihilate in the target creating pions which are detected in a detector. The

neutrons pass through the foil largely unhindered. A typical $\bar{n}$ signal consists of five pions with a total energy of ~ 2 GeV and a vanishing total momentum. The detector must be shielded against cosmic rays. In comparison with the experiments with bound neutrons there is, however, the advantage that the background can be measured directly by applying an additional magnetic field which suppresses oscillation so that one detects only background events [Dub88].

The first experiment on free $n\bar{n}$ oscillations was carried out by a CERN–ILL–Padua–Rutherford–Sussex collaboration at the research reactor of the ILL in Grenoble in 1982–83 [Fid85]. The most important parameters of this measurement were:

- neutron flux: 1.5×10^9 s^{-1}
- average neutron velocity: $v = 161$ m s^{-1} (corresponding to a kinetic energy of 1.4×10^{-4} eV)
- flight path in the vacuum: $L = 2.7$ m
- flight time: $t = 26$ ms.

As a result the following bound was obtained

$$\tau_{n\bar{n}} > 1 \times 10^6 \text{ s} \quad (90\% \text{ c.l.}). \tag{5.73}$$

A further experiment was carried out at the Italian 250 kW Triga Mark II reactor in Pavia [Bre89]; the corresponding parameters were:

- neutron flux: 3.2×10^{10} s^{-1}
- average neutron velocity: $v = 2200$ m s^{-1}
- flight path in the vacuum: $L = 17.6$ m
- flight time: $t = 8$ ms.

The bound obtained for the oscillation time was

$$\tau_{n\bar{n}} > 4.7 \times 10^5 \text{ s} \quad (90\% \text{ c.l.}). \tag{5.74}$$

In what follows we shall discuss a new experiment at the ILL in somewhat more detail. This experiment aimed to raise the lower bound for the period to the 10^8 s range [Dub88, Pug89, Bal90]. The ILL's 57 MW high-flux reactor is used as the neutron source. Figure 5.7 outlines the experimental setup.

The reactor neutrons moderated in liquid deuterium ($T = 27$ K) reach the experiment through a 60 m long neutron conductor (H 53). The neutron conductor consists of a glass tube with a vapour-deposited nickel coating from the walls of which the neutrons are totally reflected. A slight curvature prevents γ quanta and fast neutrons from passing through directly from the reactor. The neutron intensity is approximately 2×10^{11} s^{-1} for an average wavelength of 6.5 Å, corresponding to a kinetic energy of 2×10^{-3} eV or an average velocity

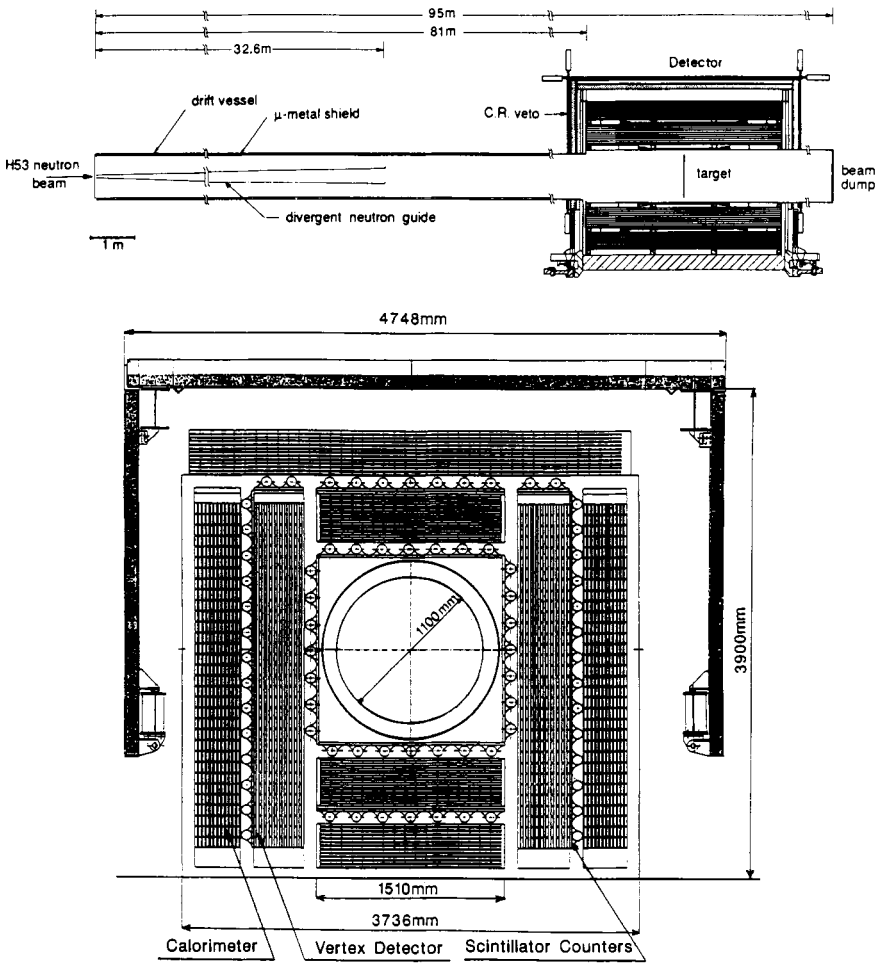


Figure 5.7. Experiment on $n\bar{n}$ oscillations of free neutrons at the ILL in Grenoble (from [Bal90]). Top: layout of the neutron guide, the target and the detector system; bottom: a cross section through the detector.

of 600 m s^{-1}. The measuring apparatus itself consists of an evacuated drift tube, the magnetic shielding and the detector system.

The drift tube. On leaving the H 53 neutron conductor the neutrons enter the 94.5 m long evacuated drift tube. The tube consists of an 81 m long flight path and another part which houses the target surrounded by detectors. The overall structure is evacuated to a residual pressure of 10^{-3} Pa to suppress

collisions between the neutrons and gas molecules, since these would decrease the probability of an $n\bar{n}$ transition. The large tube contains a 32 m long neutron conductor opening towards the front ('neutron trumpet') which is used to focus the beam.

The magnetic screening. As previously discussed, the neutrons must be very well screened against external magnetic fields to ensure that the quasi-free condition is satisfied in the area of the flight path. For this, a second mu-metallic tube is fitted inside the drift tube as passive screening, so that the transverse component of the Earth's magnetic field is well screened. The longitudinal component is actively reduced using a coil, the field of which compensates the Earth's magnetic field along the axis of the tube. The coil is wrapped around the mu-metallic screening [Bit85]. Using these measures the Earth's magnetic field may be reduced from 40 μT to a value of $\tilde{B}_0 < 10$ nT. Thus, the external interference energy ΔE amounts to less than 10^{-15} eV for all neutrons; whence the condition for quasi-free neutrons is satisfied for a flight time of 0.1 s.

The target and the detector. The target is a 200 μm thick foil of carbon. Antineutrons forming during the flight time of 0.1 s annihilate with a very high probability ($> 99\%$), while the foil is practically transparent to neutrons (transmission $> 95\%$). The area around the target is coated with ^{6}LiF to absorb scattered neutrons. The neutron beam is also stopped in ^{6}LiF 10.8 m after it passes through the target. ^{6}LiF is chosen because this substance absorbs neutrons without creating γ quanta.

The detector for detecting annihilation pions is subdivided into four quadrants and surrounds the target with a solid angle of $\Omega = 0.94 \times 4\pi$ sr. Ten layers of streamer tubes are used to detect the tracks of charged particles. In front of and behind these streamer layers are scintillation counters which function as triggers. These may be used to measure the flight time and the particle direction (from the interior outwards, or conversely). Finally, the internal detector is surrounded by a calorimeter to detect γ quanta from the decay of the π^0. In addition, the calorimeter may be used to measure the overall energy and the momentum of the events. The detector as a whole is surrounded by a veto system of scintillation counters which detects charged particles from cosmic rays. The neutral component of cosmic rays is screened by a 10 cm thick layer of lead.

No events which would indicate an $n\bar{n}$ transition were detected after a measurement time of 6.11×10^5 s. This gives a lower bound of

$$\tau_{n\bar{n}} > 10^7 \text{ s} \quad (90\% \text{ c.l.}) \tag{5.75}$$

for the oscillation period [Bal90].

The following value for the transition energy is obtained

$$\delta mc^2 < 6 \times 10^{-29} \text{ MeV}. \tag{5.76}$$

After a measuring time of 2.4×10^7 s the result is [Bald94]

$$\tau_{n\bar{n}} \geq 0.86 \times 10^8 \text{ s} \quad (90\% \text{ c.l.}). \tag{5.77}$$

This bound corresponds to those derived from the nuclear stability; however, it has the advantage of being independent of theoretical uncertainties of nuclear models.

Another experiment on $n\bar{n}$ oscillations was planned at the Moscow meson factory. A sensitivity of $\tau_{n\bar{n}} \sim 10^{10}$ s was aimed at [Ilj83].

Chapter 6

Experiments to Determine the Neutrino Mass

Since the time of Pauli and Fermi, neutrinos have played a central role in our understanding of the weak interaction. Today, they play a corresponding key role in the context of grand unification theories for which their type (Dirac or Majorana particles), their mass and their magnetic moment represent important boundary conditions. The embedding of massive neutrinos in the framework of GUT models has already been discussed in chapter 1 (see also [Gro89, 90, Moh91a]). On the other hand, neutrinos are the most important candidates for non-baryonic dark matter in the universe, and their mass could determine its large-scale structure and evolution. Moreover, neutrinos provide the only direct means of sampling processes in the central areas of collapsing stars. Neutrino properties are under discussion as a potential partial solution of a possible solar neutrino puzzle. In this and the following chapter we shall be concerned with various approaches to the experimental determination of a possible neutrino mass. Figure 6.1 gives an overview of these approaches.

The only method which permits a direct determination of the neutrino mass without additional assumptions involves purely kinematic considerations of processes of the weak interaction, including nuclear β decay (tritium) and π and τ decay. Internal bremsstrahlung after electron capture also provides information about neutrino properties. Another kinematic approach involves propagation-time measurements of neutrinos from the supernovae.

In addition to kinematic tests, other methods may also be used to obtain information about the neutrino mass. However, as well as a finite mass m_ν, these require the violation of lepton number conservation. This second group of experiments includes studies of neutrinoless double-beta decay. In addition, neutrino oscillations (see chapter 7) are sensitive to mass differences between different neutrino mass eigenstates. Finally, there is also the possibility of a

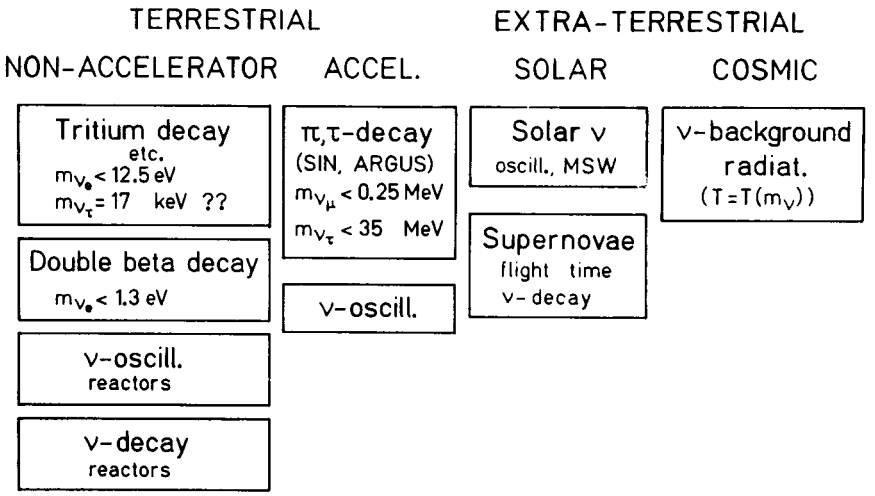


Figure 6.1. Overview of experimental methods for determining the neutrino mass.

neutrino decay, if these particles have a mass and the mass eigenstates are not identical with the interaction eigenstates. In principle it is possible to obtain information about the ν mass from (still missing) observation of the neutrino background radiation (more precisely, its temperature, see chapter 3) resulting from the big bang.

6.1 DIRECT DETERMINATION OF THE NEUTRINO MASS IN DECAY EXPERIMENTS

One of the most important methods of determining the neutrino mass involves studying the energy spectrum of nuclear β decays. The low-energy β decay of tritium is of particular importance, since it provides very sensitive information about the mass of the anti-electron neutrino. The bounds on the mass of μ and τ neutrinos are generally determined from the kinematics of pion and tauon decays.

6.1.1 The nuclear β decay and the mass of the electron neutrino

Studies of nuclear β decay essentially investigate the mass spectrum of the neutrino state which couples to the electron. The energy spectrum of the electrons or positrons emitted in β decay is modified by a finite rest mass of the neutrino. To understand this effect, we shall initially consider the β decay. For simplicity, we assume here that the interaction eigenstate is essentially given by

the mass eigenstate. Thus, the neutrino mass m_ν is given by the mass m_1 of the dominant mass eigenstate $|\nu_1\rangle$ (see, however, section 6.1.2 and chapter 7).

6.1.1.1 Introduction

The term β decay covers all nuclear decays in which the charge number of the nucleus Z for a constant mass number A changes by one unit. This includes the following decay processes (see figure 6.2)

$$ {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}X_{N-1} + e^- + \bar{\nu}_e \quad \beta^- \text{ decay} \tag{6.1a} $$

$$ {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}X_{N+1} + e^+ + \nu_e \quad \beta^+ \text{ decay} \tag{6.1b} $$

$$ e^- + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}X_{N+1} + \nu_e \quad \text{electron capture.} \tag{6.1c} $$

These transitions are closely related to the neutrino capture reactions

$$ \bar{\nu}_e + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}X_{N+1} + e^+ \tag{6.2a} $$

$$ \nu_e + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}X_{N-1} + e^-. \tag{6.2b} $$

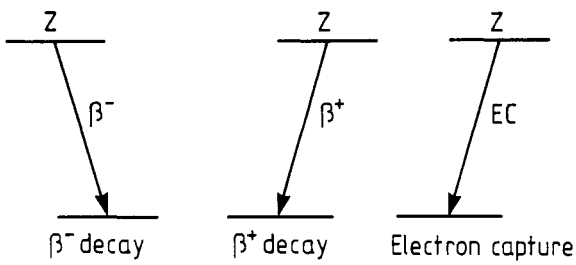


Figure 6.2. β decay and electron capture via the weak interaction.

If one does not initially consider the effects of the nuclear structure, the decay processes (6.1*a*) to (6.1*c*) are the following fundamental transitions (at the nucleon level; for a discussion on quark level see [Gro89, 90])

$$ n \rightarrow p + e^- + \bar{\nu}_e \tag{6.3a} $$

$$ p \rightarrow n + e^+ + \nu_e \tag{6.3b} $$

$$ e^- + p \rightarrow n + \nu_e. \tag{6.3c} $$

A β transition is only possible if the binding energy of the daughter nucleus is greater than that of the initial state. The energetic conditions for β^- and β^+ decay and electron capture are different. We note that in the following discussion of the energy balances $m(Z, A)$ denotes the mass of the neutral atom and not

that of the nucleus. We shall not initially consider a possible rest mass of the neutrino.

β^- decay. The decay energy Q_{β^-} in β^- decay is given by

$$\begin{aligned} Q_{\beta^-} &= [m(Z, A) - Zm_e]c^2 - [(m(Z+1, A) - (Z+1)m_e) + m_e]c^2 \\ &= [m(Z, A) - m(Z+1, A)]c^2. \end{aligned} \tag{6.4}$$

Thus, the so-called Q value corresponds in β^- decay to the mass difference between mother and daughter atom.

β^+ decay. Similarly, for β^+ decay, we have

$$\begin{aligned} Q_{\beta^+} &= [m(Z, A) - Zm_e]c^2 - [(m(Z-1, A) - (Z-1)m_e) + m_e]c^2 \\ &= [m(Z, A) - m(Z-1, A) - 2m_e]c^2. \end{aligned} \tag{6.5}$$

Since we are using atomic masses, the rest energy of two electrons is also involved.

Electron capture. For the electron capture we have

$$\begin{aligned} Q_{EC} &= [m(Z, A) - Zm_e]c^2 + m_e c^2 - [m(Z-1, A) - (Z-1)m_e]c^2 \\ &= [m(Z, A) - m(Z-1, A)]c^2. \end{aligned} \tag{6.6}$$

For a β transition to take place the corresponding Q value must satisfy the condition

$$Q_i > 0 \qquad i = \beta^-, \beta^+, EC. \tag{6.7}$$

It is clear that electron capture is energetically preferred in comparison with β^+ decay:

$$Q_{\beta^+} = Q_{EC} - 2m_e c^2. \tag{6.8}$$

For emission of a positron to take place the mass difference between mother and daughter atom must be at least $2m_e c^2$. Since both β^+ decay and electron capture lead to the same daughter nucleus, electron capture always occurs as a process in competition with the β^+ transition. If the mass difference Q_{EC} lies between 0 and $2m_e c^2$, only electron capture occurs.

In many cases, the β decay does not lead directly to the ground state but to an excited state of the daughter nucleus. This excited nucleus usually gives away its energy by emission of γ quanta or conversion electrons. If the excitation energy is greater than the neutron separation energy or the fission barrier, β delayed neutron (proton) emission and β delayed fission can occur. These β

delayed processes (see [Kla83, 86, Hir92b, Sta92a]) are of great importance in reactor physics and in the synthesis of heavy elements in the universe (r process) (see e.g. [Gro89, 90]). Since they only occur far away from the β stability line, they have no role to play in the following considerations.

6.1.1.2 The energy spectrum of the β decay

Since $\beta^{\pm}$ transitions involve many-body decays, the energy spectrum of the β particles (electrons or positrons) is continuous (see figure 6.3). This β continuum is overlaid by discrete lines which are caused by Auger electrons and conversion electrons.

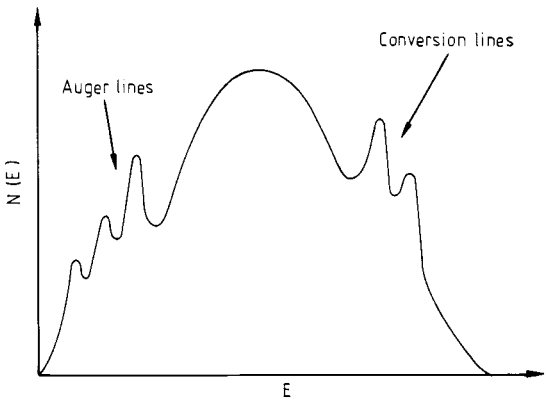


Figure 6.3. Continuous spectrum of the β decay overlaid with discrete lines of Auger and conversion electrons.

In what follows we shall give a simple derivation of the form of the β spectrum of so-called allowed transitions. Since the weak interaction is very weak, the β transition probability per unit time from a state $|i\rangle$ to the state $|f\rangle$ is calculated using time-dependent first-order perturbation theory (Fermi's golden rule, see [Lan79b])

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{2\pi}{\hbar}|\langle f|H_\beta|i\rangle|^2\frac{\mathrm{d}n}{\mathrm{d}E_0}. \tag{6.9}$$

The expression $\langle f|H_\beta|i\rangle$ denotes the matrix element of the Hamiltonian operator of the weak interaction between the initial state $|i\rangle$ and the final state $|f\rangle$. This quantity is initially unknown. However, experiments show that the *form* of most β spectra is determined solely by the factor $\mathrm{d}n/\mathrm{d}E_0$, which gives the density of the possible final states per energy interval $\mathrm{d}E_0$. The matrix element $\langle f|H_\beta|i\rangle$, which is discussed in more detail in section 6.1.1.3, may often be treated as

energy independent, to a good approximation. The form of the electron spectra is then determined solely by the final phase-space factor.

If the final state in the calculation of dn/dE_0 is characterized by the fact that an electron (positron) possesses a momentum between p and $p + dp$, the transition probability given in (6.9) is equal to the probability $N(p)dp$ that a β particle is emitted per unit time, with momentum in the given interval

$$N(p)dp = \frac{2\pi}{\hbar}|\langle f|H_\beta|i\rangle|^2 \frac{dn}{dE_0}. \tag{6.10}$$

$N(p)dp$ is the β (momentum) spectrum.

We require the number of states within a certain volume of the phase space. The phase space is generally understood to be the product of the coordinate space and the momentum space. Because of the Heisenberg uncertainty principle the minimum phase space to which an electron state may be narrowed down is

$$\Delta x \Delta p_x \simeq h. \tag{6.11}$$

For a six-dimensional phase space $(\Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z)$ it follows that there can only ever be one state in a phase-space volume of size h^3. In a system with independent particles with spin j, according to the Pauli principle, each of these states may be occupied by at most $\zeta_j = (2j + 1)$ particles.

Let us now consider the electron emitted in β decay, which is assumed to be located within a spatial volume V. According to our assumptions, the absolute value of the momentum lies in the interval between p_e and $p_e + dp_e$. In the momentum space, this momentum uncertainty is represented by a spherical shell with volume

$$d^3p_e = 4\pi p_e^2 dp_e. \tag{6.12}$$

Since the 'unit cell' in phase space has volume h^3, the number of possible states of the electron in the volume Vd^3p_e of phase space is

$$dn_e = \frac{V4\pi p_e^2 dp_e}{h^3}. \tag{6.13a}$$

There is a corresponding expression for the neutrino (provided it is impossible to confuse neutrinos and antineutrinos, we shall, for simplicity, refer in preference to a neutrino)

$$dn_\nu = \frac{V4\pi p_\nu^2 dp_\nu}{h^3}. \tag{6.13b}$$

Since a β transition is a three-particle decay, the momenta of the electron and the neutrino are not directly correlated. Thus, the probability that an electron is emitted in the momentum interval $[p_e, p_e + dp_e]$ and that at the same time a

neutrino is emitted in the interval $[p_\nu, p_\nu + \mathrm{d}p_\nu]$ is equal to the product of the individual probabilities. Consequently, we have

$$\mathrm{d}n = \mathrm{d}n_e \mathrm{d}n_\nu. \tag{6.14}$$

The density of final states for the energy interval $\mathrm{d}E_0$ may now be written as follows

$$\frac{\mathrm{d}n}{\mathrm{d}E_0} = \frac{16\pi^2 V^2}{h^6} p_e^2 p_\nu^2 \frac{\mathrm{d}p_e \mathrm{d}p_\nu}{\mathrm{d}E_0}. \tag{6.15}$$

Here, we have suppressed the spin factors ζ_s for the electron and the neutrino. The spins of the two leptons are not mutually independent. The leptons spins are coupled in different ways, depending on the angular momenta of the mother and daughter nuclei. For example, if both nuclei have angular momentum zero the spin of the electron and that of the neutrino are antiparallel as a result of the conservation of angular momentum. Consequently, one obtains only a factor two instead of a factor four for uncorrelated fermions (cf. [May84]). The coupling of angular momentum will be considered later in the calculation of the nuclear matrix elements.

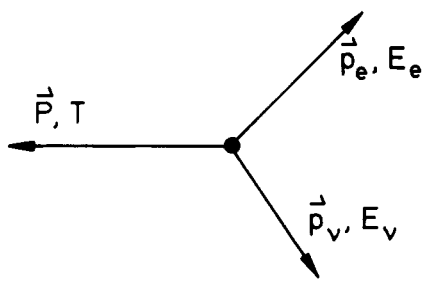


Figure 6.4. Momentum balance after the β decay of a nucleus ($\boldsymbol{P}$= momentum of the nucleus after the β transition, $\boldsymbol{p}_e$, $\boldsymbol{p}_\nu$ = electron and neutrino momentum, respectively).

Suppose that T and $\boldsymbol{P}$ are the energy and the momentum of the nucleus remaining after the β transition. From the conservation of energy and momentum, it follows immediately (see figure 6.4) that

$$\boldsymbol{P} + \boldsymbol{p}_\nu + \boldsymbol{p}_e = 0 \tag{6.16a}$$

$$T + E_\nu + E_e = E_0. \tag{6.16b}$$

Here E_0 denotes the Q value of the reaction, i.e. $E_0 = Q_{\beta^-}$ or $E_0 = Q_{\beta^+}$. Since the mass of the nucleus is very large in comparison with the lepton mass, the recoil energy transmitted to the nucleus may be neglected ($T \approx 0$) so that the total available decay energy E_0 divides into the kinetic energy of the electron and of the neutrino

$$E_e + E_\nu = E_0. \tag{6.17}$$

If we initially set the rest mass of the neutrino to zero, we have

$$p_\nu = \frac{E_\nu}{c} = \frac{E_0 - E_e}{c}. \tag{6.18}$$

Consequently, for a fixed energy of the electron we have

$$\frac{\mathrm{d}p_\nu}{\mathrm{d}E_0} = \frac{1}{c}. \tag{6.19}$$

Finally, by virtue of (6.18) and (6.19), (6.15) becomes

$$\frac{\mathrm{d}n}{\mathrm{d}E_0} = \frac{V^2}{c^3\hbar^6 4\pi^4} p_e^2 (E_0 - E_e)^2 \mathrm{d}p_e. \tag{6.20}$$

Thus, for the β spectrum we may write

$$N(p_e)\mathrm{d}p_e = \frac{V^2}{2\pi^3 c^3 \hbar^7} |\langle f|H_\beta|i\rangle|^2 p_e^2 (E_0 - E_e)^2 \mathrm{d}p_e. \tag{6.21}$$

6.1.1.3 The nuclear matrix elements for β decay

In this section, we shall briefly discuss the matrix element of the Hamiltonian operator of the weak interaction

$$\langle f|H_\beta|i\rangle = \int \mathrm{d}V \psi_f^* H_\beta \psi_i. \tag{6.22}$$

The wavefunction of the initial state ψ_i is determined by the wavefunction ϕ_i of the nucleons of the initial nucleus, while the wavefunction ψ_f is formed from the wavefunction ϕ_f of the nucleons in the daughter nucleus and the wavefunction of the electron–neutrino field $\phi_e\phi_\nu$. Since the electrons in the β decay are relativistic particles, strictly speaking the calculation of the matrix element (6.22) requires the solution of the Dirac equation. We shall not pursue this any further in this introduction (see e.g. [Gro89, 90, 92] for further details). However, we give some comments relating to the structure of this quantity.

The interaction between the nucleus and the leptons is very weak, so that, to a good approximation, the lepton wavefunctions may be described as plane waves (normalized to the volume V)

$$\phi_e(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \exp(\mathrm{i}\boldsymbol{k}_e \cdot \boldsymbol{r}) \tag{6.23a}$$

$$\phi_\nu(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \exp(\mathrm{i}\boldsymbol{k}_\nu \cdot \boldsymbol{r}). \tag{6.23b}$$

This ignores the deformation of the electron wavefunction by the Coulomb interaction with the nucleus. Expansion in a series of (6.23) about the origin $r = 0$ yields

$$\phi_j(r) = \frac{1}{\sqrt{V}}(1 + \mathrm{i}k_j \cdot r + \ldots) \qquad j = e, \nu. \tag{6.24}$$

The radius of the nucleus R may be estimated as

$$R = R_0 A^{1/3} \qquad R_0 \simeq 1.2 \text{ fm}. \tag{6.25}$$

The de Broglie wavelength $\bar{\lambda}$ of an electron with a typical energy of 2 MeV is

$$\bar{\lambda} = \frac{\hbar}{p} \simeq \frac{\hbar c}{E} = \frac{197 \text{ MeV fm}}{2 \text{ MeV}} \approx 10^{-11} \text{ cm}. \tag{6.26}$$

Thus, the wavevector $k = 1/\bar{\lambda}$ has order of magnitude 10^{-2} fm^{-1}, while the radius of the nucleus amounts to several fm; whence the product kr is small in comparison with 1. Thus, in a good approximation, we may truncate the series of (6.24) after the first step and obtain

$$\phi_j(r) \simeq \frac{1}{\sqrt{V}} \qquad j = e, \nu. \tag{6.27}$$

The lepton wavefunctions are practically constant over the volume of the nucleus. Thus, the expression $|\langle f|H_\beta|i\rangle|^2$ will contain a factor $|\phi_e(0)|^2|\phi_\nu(0)|^2 \simeq 1/V^2$, which gives the probability of finding both leptons on the site of the nucleus at their creation. If we introduce a coupling constant g as a measure of the strength of the interaction, it follows that

$$|\langle f|H_\beta|i\rangle|^2 = g^2|\phi_\nu(0)|^2|\phi_e(0)|^2|M_{fi}|^2 \simeq \frac{g^2}{V^2}|M_{fi}|^2 \tag{6.28}$$

where

$$M_{fi} = \int \mathrm{d}V \phi_f^* \mathcal{O} \phi_i \tag{6.29}$$

denotes the nuclear matrix element which describes the transition probability from the initial state of the nucleus ϕ_i to the final state ϕ_f. $\mathcal{O}$ is the corresponding transition operator.

For allowed β decays, for which the emitted leptons have no orbital angular momentum ($l = 0$), the matrix element M_{fi} is essentially determined by the nuclear structure. In addition, under this assumption M_{fi} is also energy independent, so that the form of the energy spectrum is determined solely by the phase-space factor $\mathrm{d}n/\mathrm{d}E_0$. Of course, there are also forbidden transitions, with $l \neq 0$, which we shall not discuss (see [Wu66, Gro89, 90]).

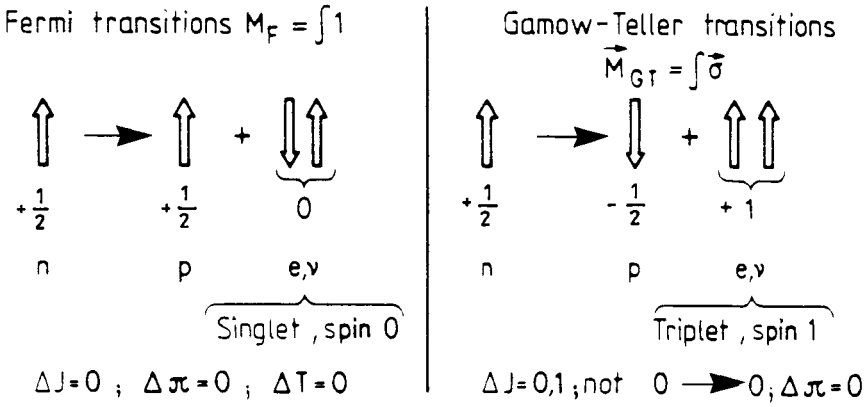


Figure 6.5. Selection rules for β decay. Fermi and Gamow–Teller transitions are distinguished according to the coupling of the lepton spins. The arrows denote the z component of the spin.

Two types of allowed β transition are distinguished, depending on whether the two leptons are emitted in a spin singlet or a spin triplet state (see figure 6.5). The two types of decay are associated with different nuclear matrix elements.

Fermi transitions. If the spins of the two leptons (e^- and $\bar{\nu}_e$ or e^+ and ν_e) are antiparallel, i.e. coupled to the overall spin 0 (singlet state), one speaks of a Fermi decay. The transition operator corresponds directly to the isospin ladder operator $\tau^-(\tau^+)$ which converts a neutron (proton) into a proton (neutron) without altering the state of motion. Thus, for the transition operator of the Fermi β^- decay, we write

$$\mathcal{O}_F = T^- = \sum_{i=1}^{A} \tau^-(i) \tag{6.30}$$

where we have summed over all nucleons in the nucleus. The associated Fermi nuclear matrix element M_F vanishes provided the nuclear wavefunctions do not belong to the same isospin multiplet. This is because the ladder operator in (6.30) only alters the z component of the isospin and not the isospin T itself[1]. This leads us to the following condition for Fermi decays

$$\Delta T = 0. \tag{6.31a}$$

Since the Fermi operator (6.30) does not alter either the spin J or the parity π of a wavefunction, there are two other selection rules

$$\Delta J = 0 \qquad \text{and} \qquad \Delta\pi = 0. \tag{6.31b}$$

[1] This is completely analogous to the behaviour of the angular-momentum ladder operators in quantum mechanics.

Gamow–Teller transitions. If, on the other hand, the lepton spins couple to the overall spin 1 (triplet state), one speaks of Gamow–Teller decays. In addition to the isospin operator, the transition operator also contains the spin operator which describes the spin flip of a nucleon

$$\mathcal{O}_{\mathrm{GT}} = \sum_{i=1}^{A} \boldsymbol{\sigma}(i)\tau^{-}(i). \tag{6.32}$$

The operator $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is defined by the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.33}$$

The corresponding selection rules for isospin, spin and parity in Gamow–Teller transitions are:

$$\Delta T = 0, 1 \tag{6.34a}$$

$$\Delta J = 0, 1 \quad \underline{\text{no}}\ 0 \rightarrow 0 \tag{6.34b}$$

$$\Delta \pi = 0. \tag{6.34c}$$

Since the operator $\boldsymbol{\sigma}$ alters the z component of the spin, the angular momentum is only conserved if the leptons have total spin 1. The case $\Delta J = 0$ is allowed, since a triplet state with z component 0 exists. However, no $0 \rightarrow 0$ transitions are allowed. For details of the explicit calculation of the nuclear matrix elements and the nuclear structure effects which must be taken into account, we refer readers to [Gro89, 90, 92].

We have seen that the nuclear matrix element (6.29) is composed of two contributions, i.e.

$$g^2|M_{fi}|^2 = g_{\mathrm{V}}^2|M_{\mathrm{F}}|^2 + g_{\mathrm{A}}^2|M_{\mathrm{GT}}|^2 \tag{6.35}$$

where we have taken account of the different coupling strengths of Fermi and Gamow–Teller decays by introducing the vector and axial coupling constants g_{V} and g_{A}. g_{V} is often also called G_β. It is

$$g_{\mathrm{V}} = G_\beta = G_{\mathrm{F}} \cos\theta_{\mathrm{C}} \tag{6.36}$$

where G_{F} denotes the universal coupling constant of the weak interaction (Fermi coupling constant). The latter determines, for example, the decay rate of the muon. θ_{C} is the Cabibbo angle ($\cos\theta_{\mathrm{C}} = 0.98$). The factor $\cos\theta_{\mathrm{C}}$ describes the modification of the coupling constant effective in the nucleus by the mixing of d and s quark states to a d' state (see sections 1.2.4 and 1.3.2, and [Gro89, 90]). Experimentally, one finds that [Par90]

$$\frac{g_{\mathrm{A}}}{g_{\mathrm{V}}} = -1.261 \pm 0.004 \tag{6.37a}$$

and

$$\frac{G_F}{(\hbar c)^3} = 1.6637(2) \times 10^{-5}\ \text{GeV}^{-2} = 1.02684 \times 10^{-5}\ m_p^{-2}. \qquad (6.37b)$$

Before we can finally write down the momentum spectrum of the β particles, we still need a correction for the Coulomb interaction between nucleus and emitted electron.

6.1.1.4 The Fermi function

The spectrum is distorted as a result of the Coulomb interaction between the emitted β particles and the nuclear charge and the shell electrons. The Coulomb field of the nucleus causes an acceleration of the emitted positrons and a deceleration of the electrons.

The decay probability $N(p_e)\mathrm{d}p_e$ is proportional to $p_e^2\mathrm{d}p_e$. Here, the momentum must be viewed as the asymptotic momentum of the electron at a large distance from the nucleus. However, the state of the β particle directly after its creation at the location of the nucleus is essential as far as the state density in the statistical factor (6.15) is concerned. Since an electron is subject to Coulomb attraction, the electron in β^- decay must be emitted with an energy greater than the asymptotic energy E_e. This leads to a phase space which is enlarged in comparison with $4\pi p_e^2\mathrm{d}p_e$. Thus, the Coulomb interaction leads to an increase in the statistical factor relevant for the decay probability in the case of β^- decay and, analogously, to a decrease in the corresponding kinematic factor for β^+ decay.

We note that the upper limit of the spectrum E_0 is not affected by the Coulomb interaction, since the Coulomb energy is already taken into account in the binding energies of the nuclei, from which the decay energy E_0 is calculated (see (6.4) and (6.5)).

The perturbation due to the Coulomb interaction may be described by a deformation of the wavefunction of the β particle, $\phi_e(r)$. The wavefunction of the electron at the site of the nucleus $|\phi_e(0)|^2$ enters into the matrix element $|\langle f|H_\beta|i\rangle|^2$. The effect of the Coulomb field may be taken into account using the following correction

$$F(Z, E_e) = \frac{|\phi_e(0)_{\text{Coul}}|^2}{|\phi_e(0)|^2}. \qquad (6.38)$$

Here, $\phi_e(0)_{\text{Coul}}$ denotes the electron wavefunction at the site of the nucleus, calculated taking into account the Coulomb force for an extended nucleus. This correction factor $F(Z, E_e)$ is called the *Fermi function*. Other small perturbations result from the screening of the Coulomb field by the shell

electrons. The Fermi function should actually be calculated relativistically (see [Gro89, 90]). However, we shall only give the non-relativistic approximation here, for which there exists an analytical expression

$$F(Z, E_e) = \frac{x}{1 - e^{-x}} \tag{6.39a}$$

where

$$x = \pm \frac{2\pi Z\alpha}{\beta} \quad \text{for } \beta^{\mp} \text{ decay.} \tag{6.39b}$$

We may now use (6.21), (6.28) and (6.38) to give the β spectrum in its final form

$$N(p_e)\mathrm{d}p_e = B^2 F(Z, E_e) p_e^2 (E_0 - E_e)^2 \mathrm{d}p_e \tag{6.40a}$$

where

$$B^2 = \frac{1}{2\pi^3 c^3 \hbar^7} (g_V^2 |M_F|^2 + g_A^2 |M_{GT}|^2). \tag{6.40b}$$

We briefly note that the decay constant λ for the β transition is easy to calculate using the expression

$$\lambda = \frac{\ln 2}{t_{1/2}} = \int_0^{p_0} N(p_e)\, \mathrm{d}p_e. \tag{6.41a}$$

This expression can be simplified if we measure energies and units in the natural units $m_e c^2$ and $m_e c$. For

$$\eta = \frac{p_e}{mc} \qquad \epsilon = \frac{W_e}{m_e c^2} = \frac{E_e + m_e c^2}{m_e c^2} \tag{6.41b}$$

it follows using $\eta \mathrm{d}\eta = \epsilon \mathrm{d}\epsilon$ that

$$\begin{aligned} N(\epsilon)\mathrm{d}\epsilon &= B^2 m_e^5 c^7 F(Z, \epsilon)\epsilon\sqrt{\epsilon^2 - 1}(\epsilon_0 - \epsilon)^2 \mathrm{d}\epsilon \\ &= \frac{m_e^5 c^4}{2\pi^3 \hbar^7} \left(g_V^2 |M_F|^2 + g_A^2 |M_{GT}|^2\right) \\ &\quad \times F(Z, \epsilon)\epsilon\sqrt{\epsilon^2 - 1}(\epsilon_0 - \epsilon)^2 \mathrm{d}\epsilon \end{aligned} \tag{6.41c}$$

from which it follows that

$$\lambda = \int_1^{\epsilon_0} N(\epsilon)\mathrm{d}\epsilon = \left(g_V^2 |M_F|^2 + g_A^2 |M_{GT}|^2\right) f(Z, \epsilon_0) \tag{6.41d}$$

where

$$f(Z, \epsilon_0) = \int_1^{\epsilon_0} F(Z, \epsilon)\epsilon\sqrt{\epsilon^2 - 1}(\epsilon_0 - \epsilon)^2 \mathrm{d}\epsilon \tag{6.41e}$$

is called the *Fermi integral*. Tabulated values for f can be found in [Gov71].

From (6.41*a*) we obtain

$$f\,t_{1/2} = \frac{D}{g_V^2|M_F|^2 + g_A^2|M_{GT}|^2} \tag{6.41f}$$

where

$$D = \frac{2\pi^3\hbar^7}{m_e^5 c^4 \ln 2}.$$

Here, $t_{1/2}$ is the *partial* half-life for a transition to a certain excited level E_f of the daughter nucleus. The *total* half-life $T_{1/2}$ for allowed β decay to the daughter nucleus is obtained by summing over all final states E_f ($T_{1/2}^{-1} \ln 2 = \lambda_{tot} = \sum_i \lambda_i$) populated in the β decay

$$T_{1/2} = \sum_{E_f \le E_0} \frac{g_V^2|M_F(E_f)|^2 + g_A^2|M_{GT}(E_f)|^2}{D} f(Z, E_0 - E_f) \tag{6.41g}$$

where $E_0 - E_f$ is the maximum kinetic energy of the electron on transition to the excited state E_f of the daughter nucleus.

6.1.1.5 The Kurie diagram

The electron spectrum (6.40*a*) is usually represented in a so-called Kurie diagram, which is obtained by plotting the expression $\sqrt{N(p)/[p^2F(Z, E)]}$ against the energy of the β particle. We have

$$\sqrt{\frac{N(p_e)}{p_e^2 F(Z, E_e)}} = B(E_0 - E_e) \tag{6.42}$$

where, for allowed transitions, B is an energy-independent constant. Thus, the graph shows a straight line, which intersects the energy axis at the end-point energy E_0 (see figure 6.6).

Corrections for a neutrino mass may be very precisely studied using the deviations from the linear relationship. The form of the spectrum has been calculated so far on the assumption that $m_\nu = 0$. In the case of a finite rest mass, we expect the spectrum to be affected in the area of the upper end of the electron spectrum, since we expect a small neutrino mass $m_\nu \ll m_e$. In particular, the end point of the β spectrum E_{max}, the maximum kinetic energy which an electron may receive, will be shifted by the rest energy of the neutrino

$$E_{max} = E_0 - m_\nu c^2. \tag{6.43}$$

The Kurie diagram deviates from the straight line in the neighbourhood of the end point (see figure 6.6). This distortion of the spectrum is easy to understand,

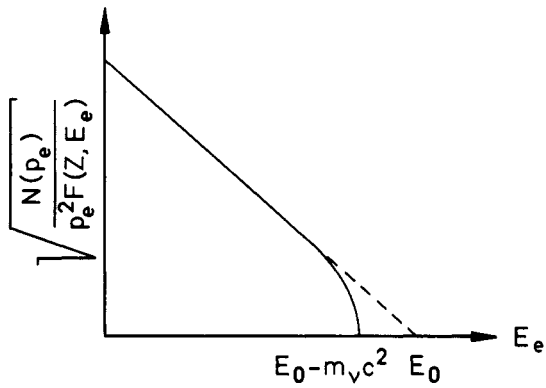


Figure 6.6. Kurie diagram of an allowed β transition. The broken line refers to the case of a massless neutrino, the continuous line to that of a neutrino with rest mass m_ν.

if we consider the rest mass of the neutrino explicitly in the relativistic expression for the neutrino energy when calculating the phase-space factor (6.15). We note that E_ν and E_e only denote the kinetic energy of the neutrino and the electron, respectively. The decay energy must now also include the rest energy of the neutrino

$$E_0 = E_e + E_\nu + m_\nu c^2. \tag{6.44}$$

For the kinetic energy of the neutrino, it follows from the relativistic equation for the total energy that

$$E_\nu = \sqrt{p_\nu^2 c^2 + m_\nu^2 c^4} - m_\nu c^2. \tag{6.45}$$

From this we obtain

$$p_\nu = \frac{1}{c}\sqrt{E_\nu(E_\nu + 2m_\nu c^2)}. \tag{6.46}$$

Substituting for E_ν according to (6.44) we have

$$p_\nu = \frac{1}{c}\sqrt{(E_0 - E_e - m_\nu c^2)^2 + 2m_\nu c^2(E_0 - E_e - m_\nu c^2)}. \tag{6.47}$$

Thus, $p_\nu^2\, \mathrm{d}p_\nu/\mathrm{d}E_0$ may be calculated as

$$\begin{aligned} p_\nu^2 \frac{\mathrm{d}p_\nu}{\mathrm{d}E_0} &= \frac{p_\nu}{c^2}(E_0 - E_e) \\ &= \frac{1}{c^3}(E_0 - E_e)^2\sqrt{1 - \left(\frac{m_\nu c^2}{E_0 - E_e}\right)^2}. \end{aligned} \tag{6.48}$$

The β spectrum is, with (6.10) and (6.15), then given by

$$N(p_e)\mathrm{d}p_e = B^2F(Z, E_e)p_e^2(E_0 - E_e)^2\sqrt{1 - \left(\frac{m_\nu c^2}{E_0 - E_e}\right)^2}\,\mathrm{d}p_e. \tag{6.49}$$

For the Kurie diagram we obtain

$$\sqrt{\frac{N(p_e)}{p_e^2F(Z, E_e)}} = B(E_0 - E_e)\left[1 - \left(\frac{m_\nu c^2}{E_0 - E_e}\right)^2\right]^{1/4} \tag{6.50}$$

instead of (6.42). Thus, it is possible to obtain information about the neutrino mass from the analysis of experimental decay spectra. One possible piece of circumstantial evidence would be the smaller maximum energy $E_{\max}$ of the electron for a finite neutrino mass. For this, one would have to compare the end-point energy determined from the Kurie diagram with the decay energy determined independently of $E_{\max}$. It is, however, not currently possible to measure the decay energy with sufficient accuracy. Instead, one tries to investigate possible alterations in the form of the spectrum in the end-point area. A downwards curvature in the Kurie diagram would point to $m_\nu \neq 0$.

Determination of the neutrino mass in this way requires careful recording of the Kurie diagram at the topmost end of the spectrum where the electron counting rate is particularly small. The momentum of the β particles is usually measured using a magnetic spectrometer, the momentum resolution Δp of which determines the smallest measurable neutrino mass. Since the resolution is in general proportional to the momentum itself

$$\Delta p \sim p \tag{6.51}$$

one investigates β emitters with smallest possible Q value E_0.

6.1.1.6 The tritium experiments

One particularly intensively studied transition is the β decay of tritium

$${}^3_1\mathrm{H} \rightarrow {}^3_2\mathrm{He} + e^- + \overline{\nu}_e. \tag{6.52}$$

This is an allowed β transition with an extremely small decay energy of only $E_0 = 18.594 \pm 0.008$ keV [Wap85]. Because of the small nuclear charge number Z, the distortion of the electron wavefunction is also small, so that this decay is very suitable for the search for deviations in the electron spectrum. It follows from (6.50) that the deformation of the Kurie diagram should be the more marked the greater the ratio $m_\nu c^2/E_0$ is.

Such experiments require great experimental skill. The counting rate of the electrons in the end area of the spectrum is extremely small and merges asymptotically with the background. Only a fraction of 10^{-9} of the total intensity is found in the last 20 eV of the tritium spectrum. Because of the relatively long half-life of tritium (12.33(6) years) the counting rates are very low so that experimental arrangements with a high luminosity and low background rate are required. Furthermore, the resolution of the spectrometer is required to be comparable with the mass to be determined.

Because any spectral deformation will be very small, effects resulting from the energy loss of the electrons in the tritium preparation and the apparatus must be taken into account. In particular, the tritium source takes on a critical importance. An extremely thin preparation of the source is chosen in order to keep the energy loss as small as possible.

Another effect concerns the atomic states of the $^3He^+$ ions populated in the β decay. In a free He^+ ion the 1s ground state is only reached in 70% of all β transitions, while approximately 25% of the decays populate the excited 2s state (40.5 eV); this leads to a spread of the line. The β spectrum derived in the last section is strictly speaking only valid for a free tritium atom. When molecular tritium is used molecule binding energies must be taken into account. When the tritium is built into carrier substances the situation may become even less clear as a result of lattice vibrations, etc (see e.g. [Boe87]). Finally the measured spectrum is the result of a convolution of the theoretical spectrum with the resolution function of the spectrometer. The latter may be decomposed into the geometrical resolution function and the energy loss function of the source.

A number of experiments involving tritium decay have been carried out. Table 6.1 gives a summary of the results derived for the neutrino mass. We shall describe some of these experiments below.

In a classical experiment, Bergkvist studied the tritium spectrum using a double-focus magnetic spectrometer [Ber72] with a geometrical resolution $\Delta E = 40$ eV ($\Delta p/p = 0.11\%$). The source used was a 0.2 mm thick aluminium foil in which the tritium molecules ($^3H_2$) were implanted. The average energy loss in the source was 20 eV, giving an overall resolution of 55 eV. The measurements gave an upper bound of

$$m_\nu c^2 < 60 \text{ eV} \tag{6.53}$$

for the mass of the neutrino.

The first evidence for a finite rest mass of the electron (anti)neutrino came from measurements carried out at the Institute for Theoretical and Experimental Physics (ITEP) in Moscow using an iron-free toroidal magnetic spectrometer [Lub80, Bor85, 87]. The energy resolution at the end point of the β spectrum was 45 eV [Lub80]. For the tritium source, the Moscow physicists used valine ($C_5H_{11}NO_2$) doped (18%) with tritium. This amino acid containing tritium was

Table 6.1. Results of experiments on tritium decay.

m_ν [eV/c^2]	Confidence level [%]	Ref.
< 250	—	[Lan52]
< 86	90	[Röd72]
< 60	90	[Ber72]
14–46	99	[Lub80]
< 65	95	[Sim81]
< 50	90	[Der83]
20–45	—	[Bor85]
< 18	95	[Fri86]
17–40	—	[Bor87]
< 32	95	[Kaw87]
< 27	95	[Wil87]
< 29	95	[Kaw88]
< 15.4	95	[Fri91]
< 13	95	[Kaw91]
< 9.3	95	[Rob91]
< 7.2	95	[Bac93]

vapour-deposited on a thin aluminium foil (2 μg cm^{-2} valine). The analysis of the data led to the deduction of a finite mass [Bor87]

$$17 \text{ eV} < m_\nu c^2 < 40 \text{ eV}. \tag{6.54}$$

However, there arose strong doubts as to the correctness of this result (see e.g. [Ber85c]). One of the greatest problems as far as the evaluation is concerned relates to the resolution function, the determination of which for the valine source is associated with large uncertainties. In the meantime, several other measurements now exist which contradict the result of the ITEP group (table 6.1).

An experiment with a similar magnetic spectrometer was carried out at the Physics Institute of the University of Zurich [Fri86]. Figure 6.7 gives a schematic illustration of the construction. A high voltage is applied between the cylinder-shaped source and the grid, which decelerates the electrons before they enter the analysing magnet, in order to improve the resolution according to (6.51). The electrons pass four times through a magnetic field of a few gauss generated by 36 rectangular current loops. They then enter a detector in which a high voltage is applied so that the electrons are again accelerated prior to detection. One major difference between the Zurich and the Moscow experiments concerns the tritium source. Fritschi *et al* [Fri86] implanted molecular tritium ($^3H_2$) in a carbon film which was deposited on an aluminium sheet. Unlike the vapour-deposited

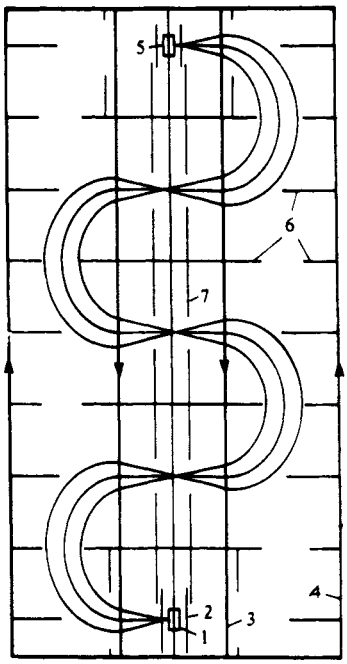


Figure 6.7. Diagram of the Zurich magnetic spectrometer (from [Boe92]). 1: source; 2: grid; 3, 4: coils, 5: detector, 6,7: apertures.

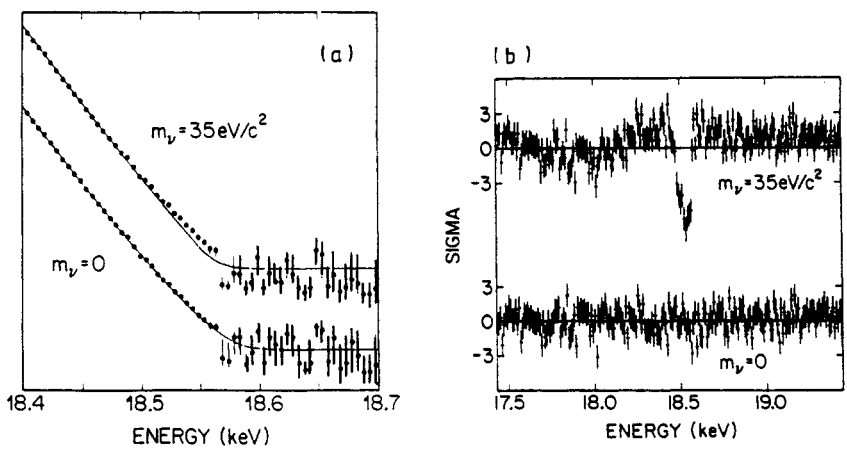


Figure 6.8. (*a*) Electron spectrum of the tritium decay from measurement at the Zurich spectrometer (from [Fri86]). The continuous curves fitted to the measurement points correspond to a neutrino mass of 0 and 35 eV/c^2, respectively. (*b*) Difference between the fitted function in (*a*) and the data, divided by the standard deviation for $m_\nu = 0$ and $m_\nu = 35$ eV/c^2 (from [Fri86]).

valine which tends to form islands, the implantation technique provides for a homogeneous distribution of the tritium. In addition, the carbon layer provides for a stable binding by means of C–H bonds.

The result of the measurements in Zurich is shown in figure 6.8. The data are consistent with a vanishing neutrino mass. The upper bound obtained for the neutrino mass is [Fri86]

$$m_\nu < 18 \text{ eV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.55a}$$

This value is compatible with (6.54) only in a narrow area. In the meantime the Zurich group has improved upon its result [Fri91]

$$m_\nu < 15.4 \text{ eV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.55b}$$

Since the determination of the spectrometer resolution and the energy loss function plays a central role when solid sources doped with tritium are used, a Japanese research group has developed a method for determining the response functions experimentally [Kaw87, 91]. This involves comparing the tritium source with a reference source with the same chemical structure. The β source used was the cadmium salt of a tritium doped organic acid ($C_{20}H_{40}O_2$). The reference source consisted of the same salt where, however, the tritium atoms were replaced by normal hydrogen atoms (^{1}H) and the natural cadmium by the radioactive isotope ^{109}Cd. The response function was evaluated using the spectrum of the observed monoenergetic KL_2L_3 Auger electrons from the ^{109}Cd decay, which have an energy around $E_0(^3\text{H})$. The result from this research is [Kaw91]

$$m_\nu < 13 \text{ eV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.56}$$

A somewhat sharper bound for the mass of the electron antineutrino has been obtained from measurements at the Los Alamos National Laboratory [Wil87, Rob91]. This group succeeded in constructing a source of gaseous molecular tritium, which represents a crucial step towards the reduction of systematic errors due to the energy loss in the source. In addition, the final states reached in ^{3}He in the β^- decay for the free $^3H_2$ molecule are very precisely known. While the first measurements only gave a bound of 27 eV/c^2, a number of improvements to the experimental arrangements, such as, for example, the replacement of the proportional counter by silicon semiconductor detectors, led to an upper bound of

$$m_\nu < 9.3 \text{ eV}/c^2 \quad (95\% \text{ c.l.}) \tag{6.57}$$

[Rob91]. Figure 6.9 shows a curve fitted to the data for $m_\nu = 0$ and $m_\nu = 30$ eV/c^2.

A research group in Mainz is currently working with a new type of β spectrometer. This Mainz solenoid retardation spectrometer is essentially an

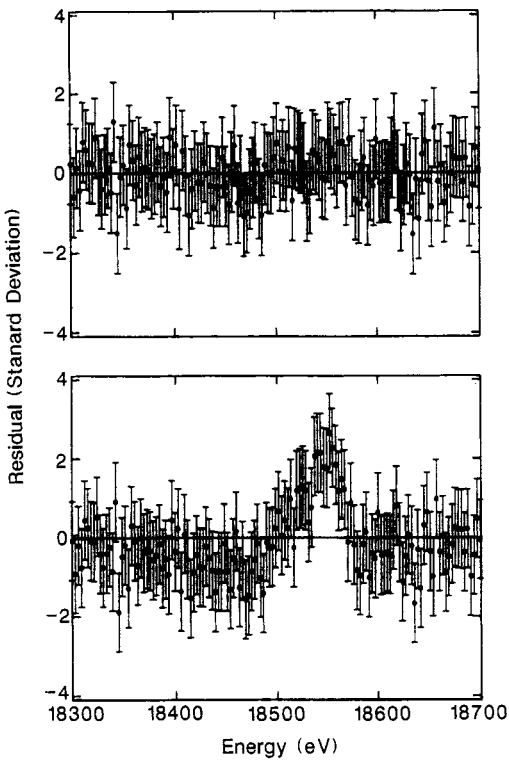


Figure 6.9. Result of the Los Alamos experiment on the β decay of free molecular tritium. The figure shows the difference between measurement data and fitted curves for $m_\nu = 0$ (top) and $m_\nu = 30$ eV/c^2 (bottom) (from [Rob91]).

electrostatic filter with a very high resolution; 13 monolayers of frozen $^3H_2$ are used as a source [Bac91, Pic92]. This spectrometer currently provides the sharpest bound [Bac92, Wei93]:

$$m_\nu < 7.2 \text{ eV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.58}$$

Thus, the more recent results contradict the ITEP measurements. A finite neutrino mass has not yet been detected in the tritium experiments. For a recent discussion of the various experiments we refer to [Kün94].

6.1.1.7 The internal bremsstrahlung

Strictly speaking the tritium experiments described above endeavoured to investigate the mass of the *anti*neutrino $\overline{\nu}_e$ emitted in the β^- decay of the tritium.

According to the CPT theorem, particles and antiparticles have the same mass, so that the above bounds on the mass apply directly to the neutrino ν_e (for simplicity, we have denoted this by m_ν). The validity of the CPT theorem has been established with an accuracy of $< 4 \times 10^{-18}$ using the difference in mass between K^0 and $\overline{K^0}$ [Par90]. However, it would be interesting to measure the mass of the electron neutrino independently.

In this respect, studies of the electron capture using the phenomenon of the internal bremsstrahlung seemed to be promising

$$(Z, A) + e^- \rightarrow (Z - 1, A) + \nu_e + \gamma. \tag{6.59}$$

Internal bremsstrahlung occurs both for $\beta^\pm$ transitions and electron capture. It is a continuous bremsstrahlung which results in β decay from the fact that the charged β particle detaches itself from the charge of the nucleus with a large acceleration, and in electron capture from the sudden transfer of the electron charge to the nucleus. Although the intensity of this emission is less in electron capture than in β transitions, the former has the advantage that the bremsstrahlung spectrum is not distorted (overlaid) by external bremsstrahlung, since no charges are emitted and consequently cannot be deflected (accelerated) by adjacent nuclei.

Table 6.2. Results from experiments on internal bremsstrahlung.

m_ν [eV/c^2]	Nucleus	Ref.
< 1250	^{163}Ho	[Yas83]
< 1300	^{163}Ho	[Jon83]
< 500	^{193}Pt	[Jon83]
< 550	^{163}Ho	[Yas86]
< 225	^{163}Ho	[Spr87]

Analogously to the nuclear β decay described above, a finite neutrino mass leads to a deformation of the bremsstrahlung spectrum in the neighbourhood of the end point. This type of experiment was first discussed by De Rujula [Ruj81]. Measurements were carried out on ^{193}Pt [Jon83] and ^{163}Ho [Yas83, 86, Spr87]. The bounds obtained for the mass of the electron neutrino are given in table 6.2. Compared with the values from tritium decay given above these results are somewhat disappointing. The experimental procedure used has a number of weaknesses which affect the sensitivity. These include the low yield of internal bremsstrahlung and the inadequate energy resolution of the available γ detectors. Springer *et al* [Spr87] showed that effects due to the atomic shell

limited the sensitivity to the given 225 eV. Thus, the accuracy achieved does not permit an interesting comparison with the results from tritium decay.

6.1.2 The 17 keV neutrino

In section 6.1.1, we neglected a possible mixing of the different mass eigenstates of the neutrino. Strictly speaking, the given bounds for the neutrino mass m_ν only relate to the mass m_1 of the dominant mass eigenstate $|\nu_1\rangle$.

We shall now consider the case in which the electron neutrino $|\nu_e\rangle$ is a mixed state formed from different mass eigenstates $|\nu_\alpha\rangle$. Thus, we shall write (see chapter 7 on neutrino oscillations)

$$|\nu_e\rangle = \sum_\alpha U_{e\alpha}|\nu_\alpha\rangle. \tag{6.60}$$

The β decay may be considered as a decay in various end channels, defined by the corresponding states $|\nu_\alpha\rangle$. The decay probability in the channel α is given by the square of the absolute value of the mixing amplitude $|U_{e\alpha}|^2$ and the phase-space factor (6.49). The total electron spectrum is a superposition of all individual spectra energetically possible ($m_\alpha c^2 \leq E_0$)

$$N(p_e) = \sum_\alpha |U_{e\alpha}|^2 N(p_e, m_\alpha) \tag{6.61}$$

where m_α denotes the mass of the eigenstate and $N(p_e, m_\alpha)$ is given by (6.49) with $m_\nu = m_\alpha$.

If we consider only a two-neutrino system ($m_2 > m_1$)

$$|\nu_e\rangle = |\nu_1\rangle \cos\theta + |\nu_2\rangle \sin\theta \tag{6.62}$$

then (6.61) may be written as follows

$$N(p_e) = N(p_e, m_1)\cos^2\theta + N(p_e, m_2)\sin^2\theta. \tag{6.63}$$

The energy spectrum of the β particles exhibits an additional deformation ('kink') at the threshold energy for the emission of the heavy neutrino mass eigenstate [Shr80, McK80, Shr81].

From the discussion of section 6.1.1, we deduce that the first mass eigenstate has a very small mass ($m_1 c^2 < 10$ eV). If we neglect m_1 ($E_0 - E_e \gg m_1 c^2$) it follows that for energies below the threshold for the emission of a $|\nu_2\rangle$ we have

$$N(p_e) = B^2 F(Z, E_e) p_e^2 (E_0 - E_e)^2 \cos^2\theta \tag{6.64a}$$

for $E_0 - E_e < m_2 c^2$. Otherwise, for $E_0 - E_e > m_2 c^2$, we have

$$N(p_e) = B^2 F(Z, E_e) p_e^2 (E_0 - E_e)^2 \cos^2\theta + B^2 F(Z, E_e) p_e^2 (E_0 - E_e)\sqrt{(E_0 - E_e)^2 - m_2^2 c^4}\sin^2\theta. \tag{6.64b}$$

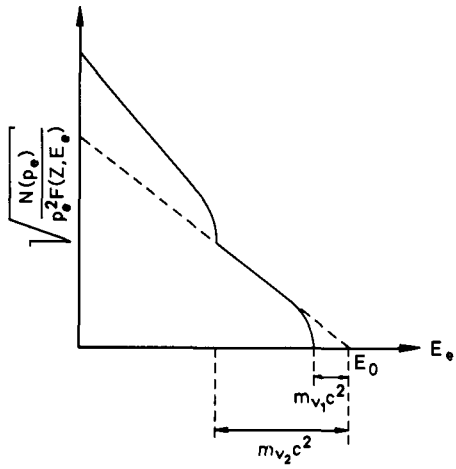


Figure 6.10. Kurie plot for two mass eigenstates m_{ν_1} and m_{ν_2}.

The Kurie diagram consists of two components, as schematically illustrated in figure 6.10. For electron energies E_e which are small in comparison with $E_0 - m_2c^2$, the Kurie diagram

$$K = \sqrt{\frac{N(p_e)}{p_e^2 F(Z, E_e)}} \tag{6.65}$$

is linear. Only when the energy E_e reaches the area of $E_0 - m_2c^2$, the line shows a kink. The absolute value of the gradient increases until the curve is perpendicular to the energy axis. For larger electron energies we are below the threshold for the emission of the heavy neutrino and we obtain the known graph of figure 6.6.

The energy region throughout which the Kurie diagram exhibits this nonlinearity (this kink) is very small. It depends on the mixing angle θ and the mass m_2 via the relationship

$$\frac{\Delta K}{K} \simeq \frac{\tan^2\theta}{2}\left(1 - \frac{m_2^2c^4}{(E_0 - E_e)^2}\right)^{1/2} \qquad \text{for } E_0 - E_e > m_2c^2 \tag{6.66a}$$

and

$$\frac{\Delta K}{K} = 0 \qquad \text{for } E_0 - E_e \leq m_2c^2. \tag{6.66b}$$

Here ΔK denotes the deviation of the Kurie diagram from a straight line. Thus, the position of the kink determines m_2 while the size of the kink is a direct measure of $\sin^2\theta$.

A number of experiments on neutrino mixing in the mass region of from 100 eV/c^2 to 500 keV/c^2 have been carried out based on these considerations (see e.g. [Sim81, Sch83]). First evidence of an effect associated with the existence of a heavy neutrino was claimed in 1985 in measurements carried out on tritium [Sim85]. Simpson used a lithium-drifted silicon detector (Si(Li)) implanted with tritium as β calorimeter. He observed the decay over a period of four years. The increase in the counting rate in comparison with (6.42) in the low-energy part of the electron spectrum (see figure 6.11(a)) was interpreted as a consequence of the emission of a heavy neutrino state with the mass

$$m_2 = 17.1 \text{ keV}/c^2 \tag{6.67a}$$

and a mixing coefficient

$$\sin^2\theta \simeq 0.03. \tag{6.67b}$$

The existence of such a heavy neutrino state (17 keV corresponds to approximately 1/30 of the rest mass of the electron) came as a great surprise to many people and led to a number of measurements with semiconductor counters and magnetic spectrometers to check this result using the β emitters ^{35}S [Alt85, Apa85, Dat85, Mar85, Ohi85] and ^{63}Ni [Het87]. However, none of these experiments has yet confirmed the existence of a 17 keV neutrino with such a large coupling to the electron (see, however, also [Sim86]). Hetherington *et*

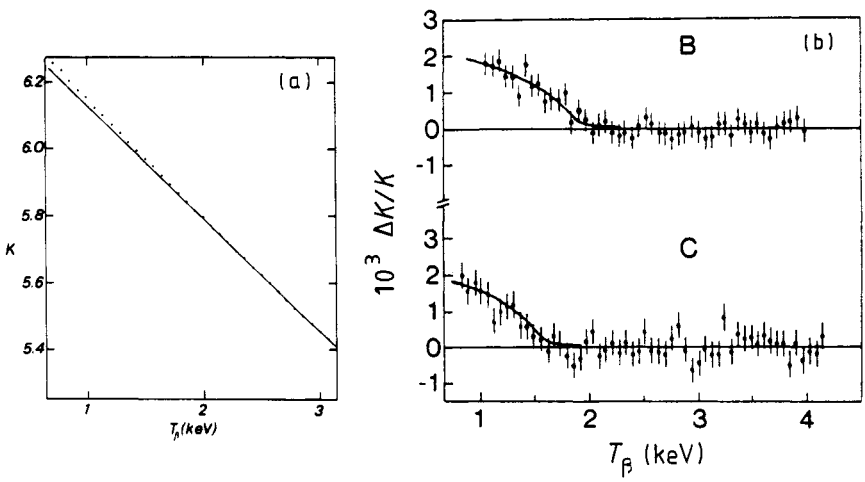


Figure 6.11. (a) Kurie plot of the tritium decay for the measurement of [Sim85] with a semiconductor counter; (b) deviation $\Delta K/K$ of the Kurie plot from a straight line in the low-energy part of the β spectrum of tritium (from [Him89]).

al studied the decay

$$^{63}\mathrm{Ni} \rightarrow {}^{63}\mathrm{Cu} + e^- + \bar{\nu}_e. \tag{6.68}$$

The Q value of 67 keV is less than the end point of the ^{35}S spectrum (167 keV) and thus provides for a better resolution of the spectrometer. On the other hand, the Q value is sufficiently high to avoid the inherent problems of tritium, that the neutrino mass sought for, m_2c^2, is only slightly smaller than the end-point energy itself. An upper bound for the mixing parameter of 0.3% for neutrinos with a rest energy of 17 keV was reported [Het87].

The tritium experiment was repeated with an improved apparatus [Him89]. In particular, the Si(Li) detector was replaced by a high-purity germanium crystal in which the tritium was implanted. The measurements essentially confirmed the result obtained in 1985, however, with a smaller mixing parameter than in the earlier measurements (see figure 6.11(*b*))

$$\sin^2\theta = 0.006 - 0.016. \tag{6.69}$$

The β spectrum of ^{35}S

$$^{35}\mathrm{S} \rightarrow {}^{35}\mathrm{Cl} + e^- + \bar{\nu}_e \quad E_0 = 167\ \mathrm{keV} \quad T_{1/2} = 87.5\ \mathrm{d} \tag{6.70}$$

was also studied using a windowless Si(Li) detector. Here, the source consisted of a 10 μm thin mylar substrate onto which the sulphur was adsorbed. The data again showed [Sim89] a threshold anomaly at an energy of $E_0 - 17$ keV, which was consistent with the emission of a neutrino with mass

$$m_2 = 16.9 \pm 0.4\ \mathrm{keV}/c^2 \tag{6.71}$$

and a mixing parameter

$$\sin^2\theta = 0.0073 \pm 0.0009 \pm 0006. \tag{6.72}$$

Later, several further experimental results were reported. Hime and Jelly have carried out another experiment on the β spectrum of ^{35}S in Oxford [Him91]. They also used an Si(Li) detector with open ^{35}S sources like [Sim89] but in an improved geometrical arrangement. The results again provided positive confirmation of the earlier measurements. Further evidence for the emission of a heavy neutrino in β decay came from investigations of the spectrum of ^{14}C ($E_0 = 156$ keV) which was embedded in a germanium crystal. This experiment, carried out in Berkeley, gave [Sur91]

$$m_2 = 17 \pm 2\ \mathrm{keV}/c^2 \tag{6.73}$$

$$\sin^2\theta = 0.0140 \pm 0.0045 \pm 0.0014. \tag{6.74}$$

The 17 keV neutrino hypothesis seemed to be confirmed by two other studies of the γ spectra of the internal bremsstrahlung of ^{55}Fe [Nor91] and ^{71}Ge [Zli91].

However, one recent experiment did not confirm the previous result. This was a study of the ^{35}S spectrum, carried out on the iron-free double-focus magnetic spectrometer of the California Institute of Technology (CalTech) in Pasadena [Bec91, Che92]. A 17 keV neutrino with an admixture of 0.8% was ruled out with more than 99% confidence. Conspicuously, the only more recent negative result is based on measurements using a magnetic spectrometer, while all the positive evidence was obtained using semiconductor detectors. Moreover, four of the six experiments carried out between 1985 and 1987 which failed to confirm the Simpson neutrino used magnetic spectrometers.

It has been conjectured that the deformation of the spectrum found by the silicon and germanium detectors could be the result of an unexpected effect of solid-state physics and not associated with the neutrino itself. This could in principle be checked by use of gas detectors. On the other hand, we note that the analysis of spectrometer data is critically dependent on the knowledge of the response function. The evaluation of the raw data requires a number of correction factors necessitated by the measuring instrument which could significantly reduce the sensitivity to a heavy neutrino.

The story and the excitement about the 17 keV neutrino seems to have come to an end by the recent announcement that a positive effect could be the result of scattering at the used diaphragms in the experimental setups [Him92]. This interpretation is convincingly confirmed by a recent work from Heidelberg [Abe93]. The above discussion thus may be considered as an example of the tortuous paths research sometimes has to follow[2].

The possible existence of a heavy neutrino would have far-reaching consequences for modern physics, in particular as far as the structure of grand unification theories is concerned (see e.g. [Cal91a, Gel91]). Naturally, there would also be implications for the problem of dark matter and the big bang theory. We shall return to such implications of a heavy neutrino later (section 6.2.6), after we have discussed the masses of the two other neutrino flavours and the bounds from investigations of neutrinoless double-beta decay.

6.1.3 The masses of the muon and the tau neutrino

Information about the masses of the muon and the tau neutrino is obtained from accelerator experiments which investigate the decays of pions and tauons in close analogy to the investigation of β decay. The neutrino mass is determined from the masses and the measured momenta of the other particles involved in the decay. Although this involves accelerator experiments, we shall discuss it here for the sake of completeness.

[2] For popular reviews to the 17 keV neutrino see [Sch91, Sel91, Zub93].

In particular, the weak semileptonic decay of the pion

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{6.75}$$

is used to determine the mass of the ν_μ, because of its simple kinematics. Let us consider a pion at rest. The muons and neutrinos generated in the reaction (6.75) are monochromatic. By virtue of the conservation of momentum and energy, it follows that

$$m_{\nu_\mu}^2 c^4 = m_\pi^2 c^4 + m_\mu^2 c^4 - 2m_\pi c^2 \sqrt{p_\mu^2 c^2 + m_\mu^2 c^4}. \tag{6.76}$$

Here, m_π and m_μ are the masses of the pion and the muon. Measurement of the muon momentum provides information about m_{ν_μ}. The uncertainties in the masses of the two charged particles, together with the measurement of the momentum, limit the sensitivity of this method.

The most accurate analysis to date of the decay of the positively charged pion was carried out at the Paul Scherrer Institute in Villigen, Switzerland [Abe84]. The π^+ beam generated in the proton cyclotron was stopped in a scintillation counter. The momentum of the emitted muons was measured using a magnetic spectrometer. The following upper bound for the rest mass of the ν_μ was obtained [Abe84]

$$m_{\nu_\mu} < 250 \text{ keV}/c^2 \quad (90\% \text{ c.l.}). \tag{6.77a}$$

Using more recent and more accurate measurements of the pion mass [Jec86], it follows from the measured momenta that

$$m_{\nu_\mu} < 270 \text{ keV}/c^2 \quad (90\% \text{ c.l.}). \tag{6.77b}$$

The sharpest bound for the mass of the ν_τ comes from studies of the τ decay at the electron–positron storage ring DORIS at DESY in Hamburg by the ARGUS Collaboration [Alb88, 92]. The $\tau^\pm$ are generated in e^+e^- collisions

$$e^+ + e^- \rightarrow \tau^+ + \tau^-. \tag{6.78}$$

The following decay channels were studied by the ARGUS group

$$\tau^\pm \rightarrow \pi^\pm + \pi^+ + \pi^- + \nu_\tau \tag{6.79a}$$

$$\tau^\pm \rightarrow \pi^\pm + 2\pi^+ + 2\pi^- + \nu_\tau. \tag{6.79b}$$

The decay into five pions gave the bound from only twelve detected events

$$m_{\nu_\tau} < 35 \text{ MeV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.80a}$$

The more recent experiment [Alb92] yields

$$m_{\nu_\tau} < 31 \text{ MeV}/c^2 \quad (95\% \text{ c.l.}). \tag{6.80b}$$

Since cosmological arguments exclude the mass range $0.5 < m_{\nu_\tau} < 35$ MeV [Tur92, Dol93], a considerably sharper limit may be deduced

$$m_{\nu_\tau} < 0.5 \text{ MeV}/c^2. \tag{6.80c}$$

Concluding, by far less sharp bounds are available for the two neutrino flavours ν_μ and ν_τ than for the electron neutrino ν_e.

6.2 DOUBLE-BETA DECAY

6.2.1 Introduction

In double-beta decay the atomic number or charge of a nucleus changes by two units. This process is one of the rarest decay modes in nature with typical half-lives of 10^{20} years or longer.

In section 6.1 we discussed simple nuclear β decay, which occurs as a first-order effect in the 'classical' theory of the weak interaction (the derivation of the β spectrum is based on Fermi's golden rule, which follows from first-order perturbation theory). In the modern electroweak theory (GWS theory) the point-like current–current interaction is replaced by a boson-exchange interaction so that single-beta decay is described as a second-order effect. Correspondingly, double-beta decay is a second-order process in the classical theory, or fourth order in the GWS model.

Since the coupling constant of the weak interaction is small and the $\beta\beta$ decay is a higher-order process, it can in general only be observed when the nucleus involved is stable against single β transitions, either for energy reasons or because of a strong suppression due to a large difference in the angular momentum of mother and daughter states. In heavy nuclei α decay also occurs as a competing process.

Whether or not a nucleus is stable against weak decay processes depends on the atomic mass $m(Z, A)$ (see equations (6.4)–(6.6)). A good approximation to the functional dependence of the atomic mass on the proton number Z and the mass number A is given by the semi-empirical mass formula due to von Weizsäcker [Wei35, Mar70] (m_H denotes the mass of the H atom)

$$m(Z, A) = Zm_H + (A - Z)m_n - a_V A + a_S A^{2/3} + a_C Z^2 A^{-1/3} + a_A \frac{(2Z - A)^2}{A} + \delta_P. \tag{6.81}$$

The main contribution to the binding energy is given by the volume energy ($a_V A$) which is proportional to the number of bound nucleons. This corresponds to the saturated nature of the nuclear forces, i.e. the fact that the binding energy per nucleon is approximately constant over a wide area of the chart of nuclides ($B/A = 7.5$–8.5 MeV for $A > 30$). Since the nucleons on the surface are less strongly bound than those in the interior there is a correction term proportional to the surface area $4\pi R^2$, the surface energy ($a_S A^{2/3}$). In addition, the binding energy is decreased by the Coulomb repulsion between the protons (Coulomb energy $\propto Z^2 A^{-1/3}$). The term proportional to $(2Z - A)^2 A^{-1}$ takes account of the asymmetry between the proton and the neutron numbers. An excess of neutrons gives rise to a decrease in the binding energy in comparison with a symmetric nucleus. The last term, δ_P, describes the effect of the pairing energy, which is important for the following discussion. For example, from a consideration of separation energies, it is known that paired nucleons are particularly strongly bound. Empirically, we find (see e.g. [Boh75])

$$\delta_P = \begin{cases} -a_p A^{-1/2} & \text{for even–even (ee) nuclei} \\ 0 & \text{for even–odd (eo) and odd–even (oe) nuclei} \\ +a_p A^{-1/2} & \text{for odd–odd (oo) nuclei} \end{cases} \tag{6.82}$$

where $a_p \approx 12$ MeV. Even–odd nuclei contain an even number of protons and an odd number of neutrons, etc. In β and $\beta\beta$ decay the mass number A is conserved. Thus, we are particularly interested in the behaviour of so-called isobars, i.e. nuclides with the same A but different Z. From (6.81) we infer that the atomic mass $m(Z, A)$ is quadratic in Z

$$m(Z, A = \text{constant}) \sim \text{constant} + \alpha Z + \beta Z^2 + \delta_P. \tag{6.83}$$

In order to represent the atomic mass as a function of Z for constant A, we must distinguish between two cases. For odd A (oe and eo nuclei) the pairing energy δ_P vanishes and one obtains a parabola (figure 6.12(a)). By β decay each nucleus changes into the corresponding adjacent energetically lower lying isobar, so that only one stable isobar is expected for nuclei with an odd mass number. Its atomic number Z_0 may be determined by calculating the minimum of (6.81).

For even A, as a result of the pairing energy, we obtain two parabolas a distance $2\delta_P$ apart, depending on whether an ee or an oo nucleus is present (see figure 6.12(b)).

As a result of a β decay an ee nucleus gives rise to an oo nucleus, and conversely. Thus, the decay steps oscillate between the two parabolas, provided this is energetically allowed. The decay chains end on the lower curve. Thus, we should not expect stable oo nuclei. In fact, only four stable oo nuclei are known ($^2_1$H, $^6_3$Li, $^{10}_5$B, $^{14}_7$N). These are very light and do not fall within the area of validity of the droplet model on which the mass formula (6.81) is based.

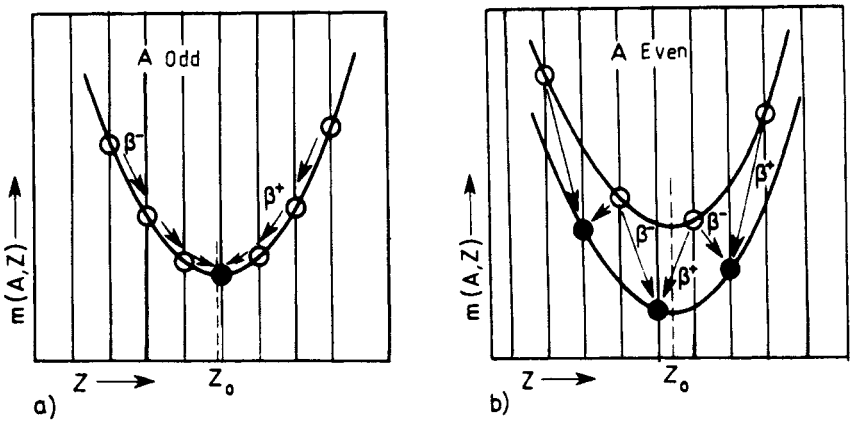


Figure 6.12. Dependence of energy on Z for nuclei with the same mass number A. Stable nuclei are denoted by bold circles; (*a*) nuclei with odd mass number A; (*b*) nuclei with even mass number A (from [May84]).

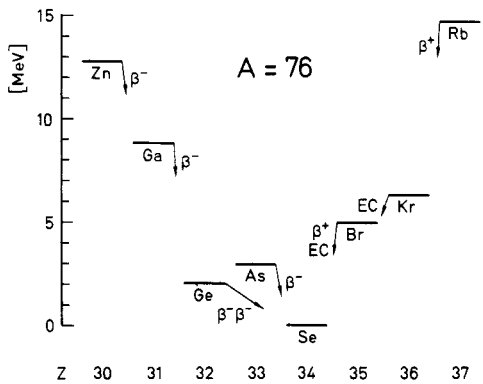


Figure 6.13. Mass spectrum of the ($A = 76$) isobars.

On the other hand, for a given even A, there may exist several β-stable isobars, since the adjacent ee nuclei on the lower parabola are separated from one another by two units of charge and thus cannot be changed into one another by β decay. However, since these β-stable nuclei do not generally have the same mass, the heavier may decay into the lighter by a second-order process. This process is known as double-beta decay. It may be thought of as simultaneous β decay of two neutrons or protons of the same nucleus.

Consequently, practically all potential double-beta emitters are ee nuclei which, as a result of the pairing interaction, have energetically lower ground

states than the adjacent oo nuclei. One of the few exceptions is ^{48}Ca, where while the single-beta decay is energetically allowed ($Q_{\beta^-} = 278 \pm 5$ keV) it is strongly suppressed as a result of a large difference in angular momentum ($0^+ \rightarrow 6^+$). Since the ground state of an ee nuclei has spin 0 and positive parity, the $\beta\beta$ decay usually involves ($0^+ \rightarrow 0^+$) transitions. In some cases population of the first excited 2^+ state and a number of other excited states in the daughter nucleus is energetically possible.

From the mass spectrum of the ($A = 76$) isobars in figure 6.13, it is clear that for ^{76}Ge both the $\beta^- EC$ and the $\beta^+ EC$ decay are energetically forbidden. The $\beta^-\beta^-$ decay to ^{76}Se is the only allowed decay mode. There is a total of approximately 36 potential $\beta\beta$ emitters (see [Gro89, 90, Sta90a]).

6.2.2 The various decay modes

We shall first consider the double-beta decay of type $\beta^-\beta^-$. We initially distinguish between two decay modes (figure 6.14)

$$(Z, A) \rightarrow (Z+2, A) + 2e^- + 2\overline{\nu}_e \qquad (2\nu\beta\beta) \qquad (6.84a)$$

$$(Z, A) \rightarrow (Z+2, A) + 2e^- \qquad (0\nu\beta\beta). \qquad (6.84b)$$

In the process (6.84a) two antineutrinos are emitted in addition to two electrons. The lepton number is conserved ($\Delta L = 0$). This decay mode is known as $2\nu\beta\beta$ decay. It is essentially similar to two consecutive simple β transitions, where the intermediate states are virtual (figure 6.15). This process is allowed in the standard model of the electroweak interaction, independently of the nature of the neutrinos. The $2\nu\beta\beta$ decay was first discussed by Maria Goeppert-Mayer at the suggestion of E Wigner [Goe35]. She calculated a half-life of approximately 10^{17} years.

While the existence of double-beta decay had been shown some time ago indirectly via geochemical experiments [Kir67] $2\nu\beta\beta$ decay was observed directly in counter experiments only recently (see section 6.2.5.3.1). Since it is a four-body decay the emitted electrons have a continuous energy spectrum (figure 6.16). Energetically, the process (6.84a) is possible if the atomic masses satisfy the following condition

$$m(Z, A) > m(Z+2, A). \qquad (6.85)$$

In addition, for practical reasons, the following requirement is imposed

$$m(Z, A) < m(Z+1, A). \qquad (6.86)$$

Of far greater interest is the neutrinoless double-beta decay ($0\nu\beta\beta$) (6.84b) first proposed by Furry [Fur39]. This represents a violation of the conservation of

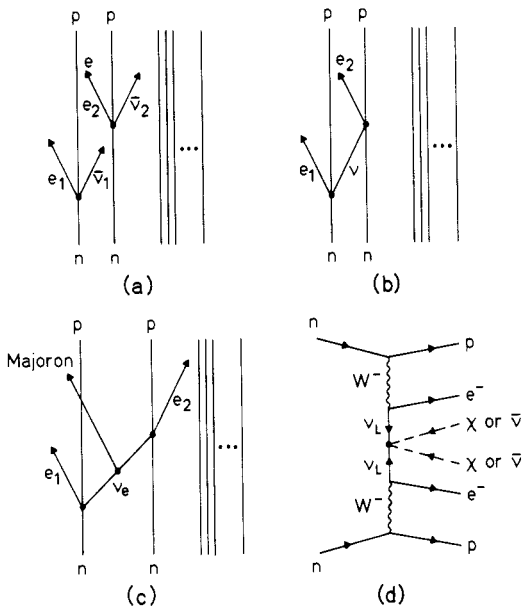


Figure 6.14. Feynman graphs for the $2\nu\beta\beta$ decay (*a*), the $0\nu\beta\beta$ decay (*b*), and the Majoron- and two-Majoron-accompanied decays, $0\nu\chi\beta\beta$ (*c*) and $0\nu 2\chi\beta\beta$ (*d*).

lepton number ($\Delta L = 2$) and is forbidden in the standard model. In this process a virtual neutrino is exchanged between two neutrons of the same atomic nucleus. The antineutrino emitted at the first vertex must be absorbed at the second vertex (figure 6.17). For this we consider the sequence

$$n \rightarrow p + e^- + \overline{\nu}_e \qquad \nu_e + n \rightarrow p + e^-. \tag{6.87}$$

Since a neutron can only absorb a neutrino, it is immediately clear that the $0\nu\beta\beta$ decay is in principle impossible for a Dirac neutrino for which the particle and antiparticle are different[3]. But this two-step process is also forbidden for a massless Majorana neutrino. We shall discuss the conditions for (6.84*b*) in somewhat more detail below.

Neutrinoless double-beta decay can only occur when two conditions are satisfied. Firstly, the neutrino must be a Majorana particle. This means that the neutrino and the antineutrino are identical particles, or more precisely, that the corresponding field is charge-conjugate to itself[4] ($\nu = \nu^c$). Secondly, the

[3] An antineutrino could be absorbed by a proton at the second vertex, but this would lead to a positron emission. This does not lead to the nucleus $(Z + 2, A)$.

[4] Note that the Dirac and Majorana descriptions for massless particles can only be distinguished if the weak interaction has a right-handed component [Gro89, 90, Kay89].

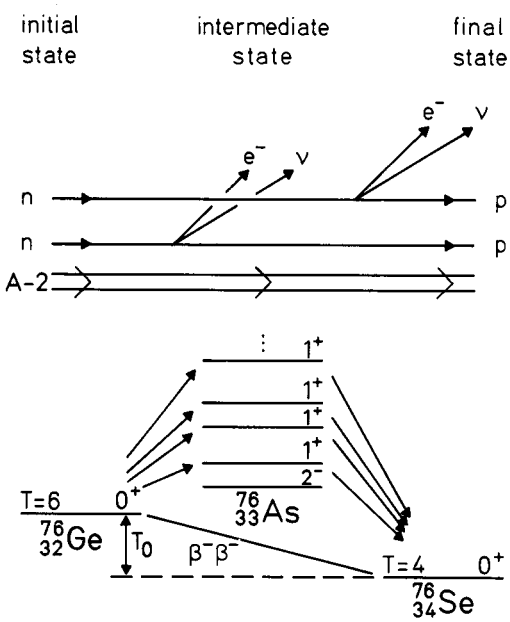


Figure 6.15. Illustration of the 2ν mode of double-beta decay.

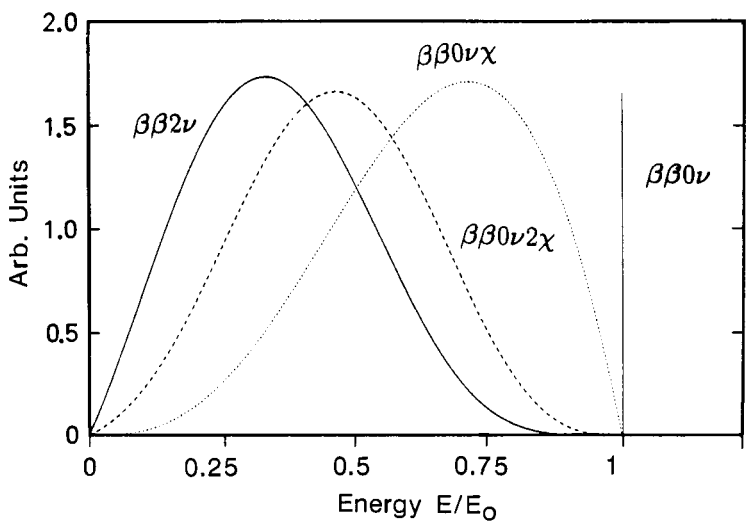


Figure 6.16. Total kinetic energy of the two emitted electrons for different decay modes ($2\nu\beta\beta$, $0\nu\beta\beta$, $0\nu\chi\beta\beta$ and $0\nu2\chi\beta\beta$); $E_0 = Q_{\beta\beta}$.

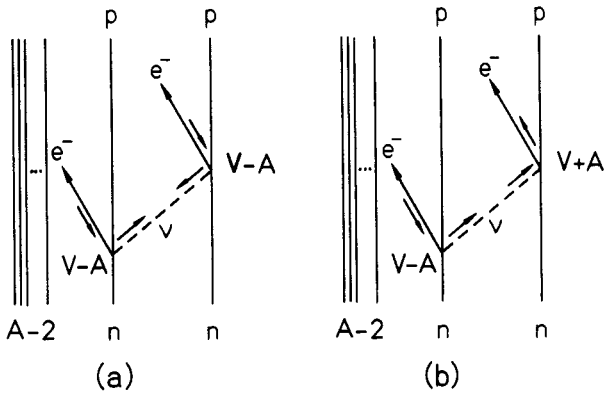


Figure 6.17. Diagrams for the 0ν mode in the two-nucleon mechanism. The arrows denote the dominant helicity components.

neutrino must have a non-zero rest mass and/or there must exist a right-handed component of the weak leptonic current. The second condition is necessary because of the helicity (see figure 6.17). In the case of a purely left-handed $(V - A)$ interaction the (massless) neutrino state emitted at the first vertex is purely right-handed, while only a left-handed neutrino can be absorbed. However, if the neutrino has a finite rest mass the helicity is no longer a good quantum number. A massive particle moves with a velocity which is less than the velocity of light in the vacuum, so that it is always possible to find a reference system, viewed from which the massive neutrino reverses its direction of motion. Since the spin vector remains unaffected by this, the projection of the spin with respect to the direction of the momentum is changed and the helicity is not conserved.

In addition to the dominant left-handed component, the wavefunction (spinor) which describes a neutrino mass eigenstate also has a right-handed admixture which is proportional to m_ν/E. This provides for the necessary helicity adjustment. Thus, the Majorana mass leads to induced $0\nu\beta\beta$ decay amplitudes proportional to m_ν. A right-handed neutrino can be absorbed at the second vertex if the charged weak currents contain a right-handed component $(V + A)$.

However, the two possibilities for the helicity adjustment cannot be considered separately from one another in GUT models (Grand Unified Theories). As long as the weak interaction is described by a gauge theory, experimental observation of neutrinoless double-beta decay *always* means that the neutrino is a *massive* Majorana particle [Sch82, Tak84, Moh91a], independently of the existence or non-existence of right-handed currents. A very clear derivation of this relationship can be found in [Kay89].

The $0\nu\beta\beta$ decay is a two-body decay. The two electrons in the final state share the available decay energy between themselves. Thus, the total (sum) energy spectrum has a peak at $Q_{\beta\beta}$, so that this process is easy to distinguish from the 2ν mode (see figure 6.16).

In $2\nu\beta\beta$ decay the kinetic energy is distributed across two electrons and two antineutrinos. In neutrinoless $\beta\beta$ decay only two electrons occur in the final state and the phase space is a factor of approximately 10^6 larger than in the case of 2ν decay since the number of the available final states is larger by this factor (see e.g. [Fae88]). The virtual neutrinos exchanged in the 0ν mode are spatially restricted to the volume of the nucleus, so that a large smearing of the momentum results. Thus, the energy region from 0 to $\sim$ 100 MeV is available to the neutrinos, while the emission of real neutrinos is limited by the $Q_{\beta\beta}$ value of a few MeV.

Here, we note that the detection of a $(0_i^+ \rightarrow 2_f^+)$ transition of the 0ν mode to the first excited 2^+ state in the daughter nucleus would enable us to make the important statement that there exists a right-handed admixture to the weak interaction. Because of the conservation of angular momentum and parity the contribution from the mass mechanism vanishes for this transition, in a first approximation (see e.g. [Tom88]). Because of the smaller decay energy, we expect longer half-lives in comparison with the ground state transitions; on the other hand, the additional detection of the γ quantum may help in reducing the background.

Other decay modes are discussed in the literature, including, for example, neutrinoless $\beta\beta$ decay with the emission of one or two majorons (χ) (see figures 6.14, 6.16) [Moh88a, 91a]

$$(Z, A) \rightarrow (Z+2, A) + 2e^- + \chi \tag{6.88}$$

$$(Z, A) \rightarrow (Z+2, A) + 2e^- + \chi + \chi \tag{6.89}$$

The introduction of a Majorana mass term implies a breaking of the $(B - L)$ symmetry at low energies. There are three different possibilities for this (see section 6.2.3). One of these, the spontaneous breaking of a global $(B - L)$ symmetry, is associated with the existence of a massless scalar Goldstone boson, which, in this case, is called a *majoron* (see also sections 6.2.3 and 6.2.4.3).

In *double-beta decay of type* $\beta^+\beta^+$ different decay possibilities exist for each of the two decay modes, since electron capture is possible in addition to positron emission:

$$(Z, A) \rightarrow (Z-2, A) + 2e^+(+2\nu_e) \tag{6.90a}$$

$$e_b^- + (Z, A) \rightarrow (Z-2, A) + e^+(+2\nu_e) \tag{6.90b}$$

$$e_b^- + e_b^- + (Z, A) \rightarrow (Z-2, A) + 2\nu_e \tag{6.90c}$$

$$e_b^- + e_b^- + (Z, A) \rightarrow (Z-2, A)^* \rightarrow (Z-2, A) + \gamma + 2X. \tag{6.90d}$$

In the reaction (6.90*d*) the daughter nucleus $(Z-2, A)$ is generated in an excited state. Thus, in addition to the two X-ray quanta (or Auger electrons) from the shell, there is a γ quantum from the nucleus. The electron capture is energetically favoured in comparison with the emission of a positron (see (6.8)). The mass difference between the (Z, A) and the $(Z-2, A)$ atoms corresponds to the Q value for the double electron capture (Q_0). The Q value for the emission of two positrons ($Q_{\beta^+\beta^+}$) is $4m_ec^2$ less than Q_0 since the masses of the two emitted positrons and the two excess shell electrons must also be taken into account[5]. Analogously the Q value for the decay (6.90*b*) is $Q_0 - 2m_ec^2$. Thus, the $\beta^+\beta^+$ decay is always accompanied by β^+EC and $ECEC$ decays. There are only seven nuclides for which the emission of two positrons is energetically allowed (see [Sta91, Hir94]).

6.2.3 Double-beta decay in the context of the grand unification theories[6]

In the GWS theory of the electroweak interaction, based on the gauge group $SU(2)_L \otimes U(1)$ the electromagnetic and the weak interaction are coupled. This model provides an excellent description of the experimental data. It contains no right-handed neutrinos and the left-handed neutrinos are massless. However, the GWS theory contains several free parameters including the Weinberg angle and the fermion masses. Moreover, the requirement for separate conservation of the baryon number B and the lepton number L cannot be considered to be fundamental since it is not based on any underlying symmetry principle.

Grand unification theories (GUTs) attempt to describe the known fundamental natural forces in a unified manner. The simplest GUT model is Georgi and Glashow's minimal SU(5) model [Geo74]. This contains only one independent neutrino field, which is left-handed (1.137). A Dirac mass is ruled out because there are no right-handed neutrinos, similarly a Majorana mass is also ruled out as a result of the exact $(B-L)$ symmetry; thus, the $0\nu\beta\beta$ decay is forbidden. A comparison of measured and predicted lifetimes of the proton (see chapter 4) together with other evidence (see chapter 2) suggests, however, that an extension of the SU(5) group is required.

One of the favourite such extensions, the SO(10) representation [Fri75], contains both left- and right-handed neutrinos. This is the simplest left–right symmetric GUT model. The left–right symmetry is broken at low energies. The left-handed $(W_L^{\pm}, Z_L^0)$ and right-handed vector bosons $(W_R^{\pm}, Z_R^0)$ obtain different masses through symmetry breaking, where the right-handed vector bosons must be considerably heavier, since, to date, no right-handed weak interaction has

[5] The atom (Z, A) has Z shell electrons, while only $(Z-2)$ shell electrons are included in the atomic mass for $(Z-2, A)$.

[6] For a detailed discussion see [Gro89, 90, Lan88, Moh86a, Val93], and in particular also [Moh91a, 94, Bur94, Lec94, Val94, Bam95a].

been found ($m_{W_R} > 1.6$ TeV/c^2 [Moh88b,c]). In many models based on the SO(10) group the neutrinos are Majorana particles. Since the $(B-L)$ symmetry is not necessarily conserved, Majorana mass terms are allowed.

If the neutrinos have a mass, mixing may occur, since the mass eigenstates $\nu_{iL/R}$ with the masses m_i are not necessarily identical to the eigenstates $\nu_{eL/R}$ of the weak interaction

$$\nu_{eL} = \sum_i U_{ei}\nu_{iL} \qquad \nu_{eR} = \sum_i V_{ei}\nu_{iR}. \tag{6.91a}$$

The indices L and R denote left- and right-handed particles, respectively. If more than one neutrino mass eigenstate couples to the electron, this leads to interference effects in $0\nu\beta\beta$ decay. Then an *effective mass* $\langle m_\nu \rangle$ will be observed

$$\langle m_\nu \rangle = \sum_j m_j U_{ej}^2 \tag{6.91b}$$

(see sections 6.2.4.2, 6.2.5.4 and 6.2.6). In general, in not too 'pathological' GUT models, however, $\langle m_{\nu_e} \rangle = m_{\nu_e}$ (see table 1.8).

In addition to left-handed leptonic (l) and hadronic ($\mathcal{L}$) currents, the Hamiltonian operator of the weak interaction also contains right-handed leptonic (r) and hadronic ($\mathcal{R}$) currents

$$H_W \sim (\mathcal{L}\cdot l + \kappa\mathcal{R}\cdot l + \eta\mathcal{L}\cdot r + \lambda\mathcal{R}\cdot r) \tag{6.92}$$

where $\eta, \kappa, \lambda \ll 1$. The mass eigenstates of the vector bosons $W_{1/2}^{\pm}$ (with masses M_1 and M_2) are mixtures of the left- and right-handed gauge bosons

$$W_1^{\pm} = W_L^{\pm}\cos\theta + W_R^{\pm}\sin\theta \tag{6.93a}$$

$$W_2^{\pm} = -W_L^{\pm}\sin\theta + W_R^{\pm}\cos\theta \tag{6.93b}$$

where $\theta \ll 1$ and $M_2 \gg M_1$. The parameters η, κ and λ in the Hamiltonian operator, which vanish in the GWS model, may be expressed in left–right symmetric GUT models in terms of the mixing angle θ and the masses M_1 and M_2 of the heavy vector bosons

$$\eta = \kappa \approx \tan\theta \qquad \lambda \approx (M_1/M_2)^2 + \tan^2\theta \tag{6.94}$$

(see also section 6.2.4.2). The predictions of the various GUT models for the Majorana mass of the neutrino range from 10^{-11} eV/c^2 to approximately 1 eV/c^2 [Lan88] (see table 1.8 in section 1.6). Currently, the neutrinoless $\beta\beta$ decay provides the most sensitive test for the Majorana nature of the neutrino and gives the sharpest bounds for the neutrino mass and the right-handed mixing parameter [Doi85, Mut88a, 89b, Tom91, Suh93, Kla94][7].

[7] For the latter a sharp limit was deduced also from analysis of data from the supernova SN1987A [Moh91b].

The creation of Majorana neutrino masses is associated with a violation of the lepton number by two units. Since in general gauge theories $B - L$ is the only gauge-anomaly free combination of quantum numbers, this also requires a breaking of the $B - L$ symmetry [Moh88b,c, 91a]. There are three different ways in which this may be done:

(i) Explicit $B - L$ breaking, i.e. the Lagrange density contains terms which break $B - L$.
(ii) Spontaneous breaking of a local $B - L$ symmetry.
(iii) Spontaneous breaking of a global $B - L$ symmetry.

The last possibility is associated with the existence of a massless Goldstone boson, the majoron. Depending on the model there are singlet [Chi80], doublet [Aul83] and triplet majorons [Gel81]. The singlet majoron arises in the extension of the standard model by right-handed neutrinos and a further Higgs singlet with lepton number $L = 2$. In the spontaneous symmetry breaking this obtains a vacuum expectation value v_{BL}. The majoron couples to the neutrino with a strength of $\approx m_{\nu L}/v_{BL}$. The coupling to the charged fermions takes place in first-order perturbation theory via W and Z boson exchange and results in a coupling strength of the order $G_F/\pi m_f m_\nu$, where m_f denotes the corresponding fermion masses. The couplings have order of magnitude 10^{-16}, which makes the singlet majoron difficult to detect. In the model of the triplet majoron the standard model is extended with the introduction of a new Higgs triplet with hypercharge $Y = 2$ and lepton number $L = 2$. Unlike the singlet model a Yukawa coupling of type $\nu_L^T \nu_L \Delta_L^0$ is now possible in this case. Here, the global $B - L$ symmetry is spontaneously broken by the expectation value $\langle \Delta_L^0 \rangle = v_T$, which leads to a neutrino mass $m_\nu = h v_T$, where h denotes the Yukawa coupling. Singlet and triplet majorons may be emitted as shown in figure 6.14.

Another decay possibility arises in supersymmetric models with R parity violation. Here, the lepton number may be spontaneously broken if a non-vanishing vacuum expectation value and a lepton number $L = 1$ are assigned to the sneutrino, the supersymmetric partner of the neutrino. This model leads to the doublet majoron and, according to figures 6.14 and 6.16 to a double-beta decay mode with emission of two majorons [Moh88a, 91a]. By virtue of this decay channel it may be possible to draw conclusions about the zino mass (see section 6.2.4.3).

However, the coupling of the triplet or doublet majoron to the Z^0 boson would give a contribution to the decay width of the Z^0 which corresponds to that of two (or 1/2) additional massless neutrino flavours [Geo81, Bar82, Des87]. The accurate measurement of the Z^0 width at LEP (see chapter 2; for an overview see [Jar90, Ste91]) gives the number of light neutrino flavours as three, so that both the triplet and the doublet majoron are ruled out. The singlet majoron (and a mixture of a singlet and a doublet) escapes this analysis since it is only coupled

extremely weakly to matter. However, a small gauge coupling does not rule out a significant Yukawa coupling to neutrinos, as recently discussed by [Ber92b]. Thus, singlet and doublet majorons could still contribute to $\beta\beta$ decay [Moh91a].

Particular interest has been shown in the model of the singlet majoron in connection with attempts to incorporate a 17 keV neutrino into the neutrino mass hierarchies [Gla91]. A short overview of the majoron can be found in [Kla92b]. For a discussion of other new classes of majoron models and multi-majoron modes in neutrinoless double-beta decay see [Bur93, 94, Car93, Bam95].

In addition to the masses M_{W_R} of right-handed W bosons of left–right symmetric models $0\nu\beta\beta$ decay also yields information about the mass M_N of the (super-) heavy neutrino occurring in such models (see section 1.6.4). Figure 6.18 shows the relationship between M_{W_R} and M_N deduced from $0\nu\beta\beta$ decay, together with a theoretical upper limit for M_N from considerations on the stability of the vacuum (from [Moh86c, 91a], see also [Kla98]).

Considering heavy composite neutrinos, $0\nu\beta\beta$ decay can also yield bounds on the compositeness scale [Bar81, Pan94, 95, Tak95], which are more restrictive than those from accelerators.

Also, exchange of Higgs particles could yield a non-negligible contribution to the $0\nu\beta\beta$ amplitude [Moh81, 91a]. Figure 6.19 shows the Feynman graph for the contribution of a doubly charged Higgs boson. Such Higgs bosons, which couple to two W_L bosons and two electrons, occur in numerous GUT models.

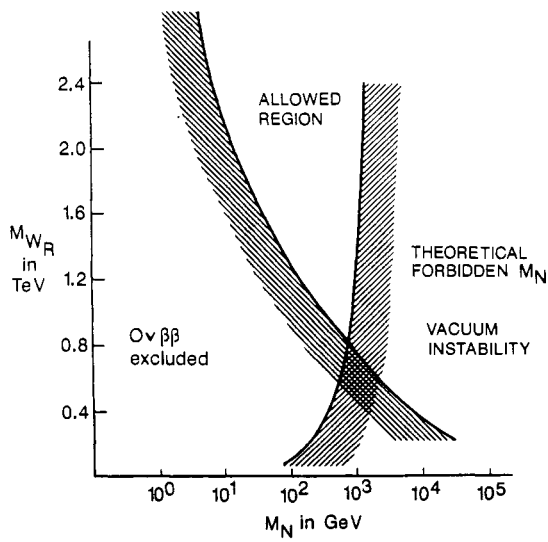


Figure 6.18. Allowed values of the parameters M_{W_R} and M_N of left–right symmetric models. The regions below the curves are excluded by $0\nu\beta\beta$ decay and the instability of the vacuum, respectively (from [Moh86c, 91a]).

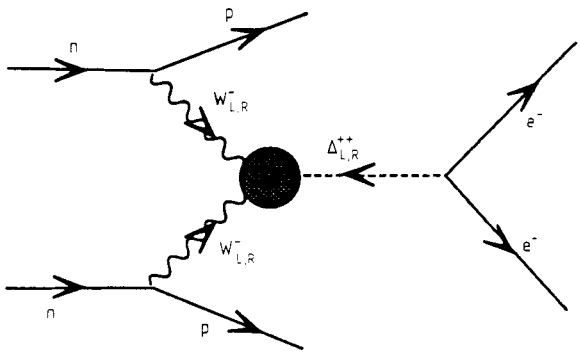


Figure 6.19. Contribution to the $0\nu\beta\beta$ decay induced by a doubly charged Higgs boson (from [Moh91a]).

In some supersymmetric theories with R parity breaking, a contribution to $0\nu\beta\beta$ decay is also possible via photino, gluino or zino exchange [Moh86b, 91a, Ver87, Hir95]. Figure 6.20 shows some of the corresponding Feynman graphs. Thus, for example, under certain assumptions, squark masses may be derived from $\beta\beta$ experiments as a function of the gaugino mass. More generally, very sharp bounds can (partly more restrictive than obtainable from high-energy accelerators) be deduced from $0\nu\beta\beta$ decay for the parameter space of R-parity violating SUSY models [Hir95] (see section 6.2.5.4).

Thus, even if the neutrino mass were too small to be observable in $0\nu\beta\beta$ decay, such experiments would still constitute a sensitive and very important test of physics beyond the standard model. Table 6.3 gives a list of some of the problems of particle physics to which double-beta decay may contribute.

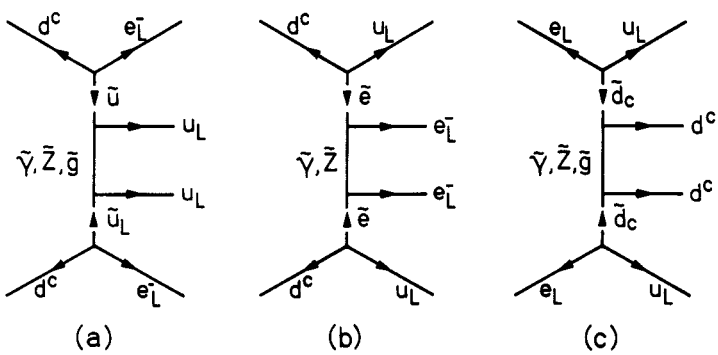


Figure 6.20. Contributions to the $0\nu\beta\beta$ decay in several SUSY models, due to photino, gluino or zino exchange (from [Moh86b]).

Table 6.3. $\beta\beta$ decay and particle physics.

Observable	Restrictions
0ν: via ν-exchange:	Beyond the standard model and SU(5) model;
Neutrino mass	early universe, matter–antimatter symmetry
Light neutrino ($\geq$ 0.1 eV)	Dark matter, L–R-symmetric models (e.g. SO(10)),
Heavy neutrino (GeV)	seesaw mechanism, compositeness
Right-handed weak currents	$V + A$ interaction, $W_R^{\pm}$ masses
via photino, gluino, zino (gaugino) exchange	SUSY models: Bounds for parameter space beyond the range of accelerators (R-parity breaking interaction, squark and slepton masses)
$0\nu\chi$: Existence of the majoron	Mechanism for the $(B-L)$ breaking: • explicit • spontaneous breaking of the local/global $B-L$ symmetry
$0\nu\chi\chi$: Majoron model	SUSY models, zino mass

6.2.4 The double-beta decay rates

In what follows, we shall only discuss ($0_i^+ \rightarrow 0_f^+$) transitions between the ground states since these are much more probable than transitions to the first excited 2^+ state in the daughter nucleus, some of which are energetically allowed. The $\beta\beta$ decay is a second-order process in the weak interaction. Mathematically, it is described using perturbation theory. The transition amplitude S_{fi} (scattering matrix) between the states $|i\rangle$ and $|f\rangle$ is represented in the time-dependent perturbation theory by an infinite series (see e.g. [Nac86, Gro89,90])

$$\begin{aligned} S_{fi} &= \sum_{n=1}^{\infty} \frac{(-\mathrm{i})^n}{n!} \langle f | \int_{-\infty}^{+\infty} \mathrm{d}^4x_1 \mathrm{d}^4x_2 \ldots \mathrm{d}^4x_n \\ &\quad \cdot \mathcal{T}[H_\beta(x_1), H_\beta(x_2), \ldots, H_\beta(x_n)] | i \rangle \\ &= \delta_{if} - \mathrm{i} \langle f | \int_{-\infty}^{+\infty} \mathrm{d}^4x H_\beta(x) | i \rangle \\ &\quad - \frac{1}{2} \langle f | \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{d}^4x_1 d^4x_2 \mathcal{T}[H_\beta(x_1), H_\beta(x_2)] | i \rangle + \ldots \end{aligned} \tag{6.95}$$

H_β denotes the Hamiltonian operator of the weak interaction and $\mathcal{T}$ is the time-ordered product

$$\mathcal{T}[H(x_1), H(x_2)] = \begin{cases} H(x_1)H(x_2) & \text{for } t_1 \geq t_2 \\ H(x_2)H(x_1) & \text{for } t_1 < t_2 \end{cases} \tag{6.96}$$

The term linear in H_β describes the single-beta decay. The double-beta decay is induced by the third term in (6.95). We shall not pursue this formal approach here and refer readers to [Gro89, 90, 92].

Typical $\beta\beta$ decay rates may be estimated using very simple considerations [Wu66]. According to (6.41*a*), the following holds for single-beta decay

$$\lambda_\beta = B^2 \int_0^{E_0} F(Z, E_e) p_e^2 (E_0 - E_e)^2 \mathrm{d}p_e. \tag{6.97}$$

For large decay energies E_0 with $E_e \simeq p_e c$, neglecting the distortion due to the Coulomb field ($F = 1$), we have

$$\lambda_\beta \simeq \frac{B^2}{c^3} \int_0^{E_0} E_e^2 (E_0 - E_e)^2 \mathrm{d}E_e = \frac{B^2}{c^3} \frac{E_0^5}{30}. \tag{6.98}$$

If we express E_0 in terms of the quantity $\epsilon_0 = E_0/m_e c^2$, we obtain

$$\lambda_\beta \simeq \frac{B^2}{c^3} \frac{\epsilon_0^5 m_e^5 c^{10}}{30}. \tag{6.99}$$

If we consider only Gamow–Teller transitions then, from (6.40*b*) with $M_\mathrm{F} = 0$, we obtain

$$\lambda_\beta \simeq \frac{m_e^5 c^4}{2\pi^3 \hbar^7} \frac{\epsilon_0^5}{30} g_A^2 |M_\mathrm{GT}|^2 \simeq \frac{m_e c^2}{\hbar} \frac{G^2}{2\pi^3} |M_\mathrm{GT}|^2 \frac{\epsilon_0^5}{30} \tag{6.100}$$

where

$$G = \frac{g_A}{m_e c^2} \left(\frac{m_e c}{\hbar} \right)^3. \tag{6.101}$$

The decay rate for double-beta decay (more precisely, $2\nu\beta\beta$ decay) in units of the characteristic rate $m_e c^2/\hbar$ follows easily from the square of the rate of the single-beta decay

$$\lambda_{\beta\beta} \sim \frac{m_e c^2}{\hbar} \left(\frac{G^2 |M_\mathrm{GT}|^2}{2\pi^3} \right)^2 \left(\frac{\epsilon_0^5}{30} \right)^2. \tag{6.102}$$

For typical values $G^2 |M_\mathrm{GT}|^2 / 2\pi^3 \sim 3 \times 10^{-27}$, as obtained from single-beta decays of light nuclei [Wu66], it follows that

$$T_{1/2}^{\beta\beta} = \frac{\ln 2}{\lambda_{\beta\beta}} \sim 3 \times 10^{27} \epsilon_0^{-10} \text{ years}. \tag{6.103}$$

Table 6.4 contains a number of examples of half-lives estimated in this way.

Table 6.4. Rough estimates of half-lives for the $2\nu\beta\beta$ decay (according to (6.103)).

Nuclide	$Q_{\beta\beta}$ [MeV]	ϵ_0 $[m_e c^2]$	$T_{1/2}^{\beta\beta}$ [years]
^{76}Ge	2.04	3.99	2.9×10^{21}
^{82}Se	3.01	5.89	6.0×10^{19}
^{100}Mo	3.03	5.93	5.6×10^{19}
^{130}Te	2.53	4.95	3.4×10^{20}
^{238}U	1.15	2.25	9.0×10^{23}

6.2.4.1 The decay rate for the $2\nu\beta\beta$ decay

The 2ν mode of $\beta\beta$ decay consists of two consecutive GT transitions. Since all potential $\beta\beta$ emitters are even–even nuclei with ground state spin 0^+, according to the selection rules for Gamow–Teller decays, the virtual intermediate states can only be 1^+ states. In addition, the isospin selection rule $\Delta T = 0$ prevents the occurrence of Fermi transitions since the isospins of the mother and daughter nuclei in the $\beta\beta$ decay differ by two units. As a simple rule we have: the ground state of a nucleus has the isospin $T = |T_z|$ where the z component T_z is defined to be half the neutron excess

$$T_z = \frac{N - Z}{2}. \tag{6.104}$$

Since the available decay energy $Q_{\beta\beta}$ is small (O (1 MeV)) and distributed across four leptons, we may assume that the electrons and antineutrinos are emitted as s waves. A good approximation is obtained if we neglect higher partial waves, nucleon recoil terms and contributions from right-handed weak currents (see e.g. [Doi85]).

In calculating the decay rate we may proceed analogously to the treatment of the single-beta decay in section 6.1, where the first-order expression (Fermi's golden rule) should now be replaced by the corresponding second-order term. For this, we present, without proof, a result of time-dependent perturbation theory (see e.g. [Gre79]). The transition probability per unit time for a transition from the state $|i\rangle$ to the state $|f\rangle$ is

$$\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)_{i\to f} = \frac{2\pi}{\hbar}|M_{fi}|^2\delta(E_f - E_i). \tag{6.105}$$

The δ function replaces the density of the final states in (6.9) for a defined level. In this case it just expresses the conservation of energy. The matrix element

M_{fi} has the form

$$M_{fi} = \langle f|H_W|i\rangle + \sum_m \frac{\langle f|H_W|m\rangle\langle m|H_W|i\rangle}{E_i - E_m} + \sum_{m,n} \frac{\langle f|H_W|m\rangle\langle m|H_W|n\rangle\langle n|H_W|i\rangle}{(E_i - E_m)(E_i - E_n)} + \dots . \tag{6.106}$$

The states $|m\rangle$, $|n\rangle$ are intermediate states which occur in the transitions. H_W is the associated interaction Hamiltonian operator. For the double-beta decay, we then have

$$\delta\lambda_{2\nu} = \frac{2\pi}{\hbar}\left|\sum_m \frac{\langle f|H_\beta|m\rangle\langle m|H_\beta|i\rangle}{E_i - E_m}\right|^2 \delta(E_i - E_f). \tag{6.107}$$

In contrast to section 6.1, the energies given in this section are total energies. The energy E_m of the intermediate state consists of the energy E_{N_m} of the intermediate nucleus together with the energy of the two leptons. Since we cannot distinguish which electron neutrino combination occurs in the first decay step and which in the second, the sum in (6.107) must be taken over all intermediate states with the energies

$$\begin{aligned} E_m &= E_{N_m} + E_{e_1} + E_{\nu_1} \quad E_m = E_{N_m} + E_{e_2} + E_{\nu_2} \\ E_m &= E_{N_m} + E_{e_1} + E_{\nu_2} \quad E_m = E_{N_m} + E_{e_2} + E_{\nu_1}. \end{aligned} \tag{6.108}$$

Without describing the derivation explicitly (see [Gro89, 90]) we give the result for the decay rate of the 2ν decay mode (in units $\hbar = c = 1$) (see [Boe87, Gro89, 90], see also [Kon66])

$$\begin{aligned} \lambda_{2\nu} = \frac{g_A^4}{32\pi^7} &\int_{m_e}^{Q_{\beta\beta}+m_e} F(Z, E_{e_1}) p_{e_1} E_{e_1} \mathrm{d}E_{e_1} \\ &\times \int_{m_e}^{Q_{\beta\beta}+2m_e-E_{e_1}} F(Z, E_{e_2}) p_{e_2} E_{e_2} \mathrm{d}E_{e_2} \\ &\times \int_0^{Q_{\beta\beta}+2m_e-E_{e_1}-E_{e_2}} E_{\nu_1}^2 E_{\nu_2}^2 \mathrm{d}E_{\nu_1} \sum_{a,a'} A_{aa'} \end{aligned} \tag{6.109}$$

where $Q_{\beta\beta}$ denotes the Q value of the transition, i.e. $Q_{\beta\beta}$ gives the available kinetic energy of the leptons which is calculated from the difference between the atomic masses. We have

$$Q_{\beta\beta} = E_{e_1} + E_{e_2} - 2m_e + E_{\nu_1} + E_{\nu_2}. \tag{6.110}$$

$p_e = \sqrt{E_e^2 - m_e^2}$ denotes the momentum of the electrons and $F(Z, E)$ is the Fermi function. E_e and E_ν are the lepton energies. The Cabibbo angle θ_C, which

describes a mixing in the quark sector, has value $\cos\theta_C = 0.9744 \pm 0.0010$ [Sir87]. The quantity $A_{aa'}$ contains the Gamow–Teller nuclear matrix elements and the typical energy denominators from the perturbation calculation

$$A_{aa'} = \langle 0_f^+ \| t_- \sigma \| 1_a^+ \rangle \langle 1_a^+ \| t_- \sigma \| 0_i^+ \rangle \langle 0_f^+ \| t_- \sigma \| 1_{a'}^+ \rangle \langle 1_{a'}^+ \| t_- \sigma \| 0_i^+ \rangle \times \tfrac{1}{3}(K_a K_{a'} + L_a L_{a'} + \tfrac{1}{2} K_a L_{a'} + \tfrac{1}{2} L_a K_{a'}) \tag{6.111}$$

where

$$K_a = \frac{1}{E_a + E_{e_1} + E_{\nu_1} - E_i} + \frac{1}{E_a + E_{e_2} + E_{\nu_2} - E_i} \tag{6.112a}$$

$$L_a = \frac{1}{E_a + E_{e_1} + E_{\nu_2} - E_i} + \frac{1}{E_a + E_{e_2} + E_{\nu_1} - E_i} \tag{6.112b}$$

E_a is an abbreviation for E_{N_a}. It is common to separate the nuclear physics component from the kinematic factors by replacing the sum of the lepton energies $E_e + E_\nu$ in the denominator of (6.112) by half of the available energy $Q_{\beta\beta}/2 + m_e$. The half-life may then be expressed as

$$T_{1/2}^{2\nu} = \left(G^{2\nu} |M_{\mathrm{GT}}^{2\nu}|^2 \right)^{-1}. \tag{6.113}$$

$G^{2\nu}$ is the phase-space integral $G^{2\nu} \sim Q_{\beta\beta}^{11}$, $M_{\mathrm{GT}}^{2\nu}$ denotes the nuclear matrix element

$$M_{\mathrm{GT}}^{2\nu} = \sum_a \frac{\langle 0_f^+ \| t_- \sigma \| 1_a^+ \rangle \langle 1_a^+ \| t_- \sigma \| 0_i^+ \rangle}{E_a + Q_{\beta\beta}/2 + m_e - E_i}. \tag{6.114}$$

A different approximation was often used in earlier calculations. If the energies of the virtual intermediate states E_a are replaced by an average energy $\langle E_a \rangle$ the sum may be taken over the intermediate states using $\sum_a |1_a^+\rangle\langle 1_a^+| = 1$ (closure approximation). This approximation has the advantage that only the wavefunctions of the initial state and the final state are required. The lengthy calculation of the intermediate states is avoided. However, since interferences between the individual terms of the matrix element $M_{\mathrm{GT}}^{2\nu}$ in (6.114) are important, the amplitudes must be weighted with the correct energies E_a. Thus, the closure approximation does not provide a reliable result for the 2ν mode. A much more successful method of avoiding the explicit calculation of the spectrum of the intermediate states is given by the so-called operator expansion method [Wu91, 92, Hir94a].

6.2.4.2 *The decay rate for the $0\nu\beta\beta$ decay*

As previously discussed, $0\nu\beta\beta$ decay is only possible if the neutrino is a Majorana particle and if either the mass of the neutrino is non-zero or the weak

charged current contains a right-handed admixture[8]. The last two possibilities are associated with different nuclear structure matrix elements [Doi85, Mut88, Tom91].

The following Hamiltonian density [Doi83, Tom86] (see also (6.92)), which contains both left- and right-handed currents, is generally used to calculate the $0\nu\beta\beta$ decay rates

$$H_W = \frac{G_F \cos\theta_C}{\sqrt{2}}(j_{L\mu}J_L^{\mu\dagger} + \kappa j_{L\mu}J_R^{\mu\dagger} + \eta j_{R\mu}J_L^{\mu\dagger} + \lambda j_{R\mu}J_R^{\mu\dagger}) + \text{h.c.} \tag{6.115}$$

The left- and right-handed leptonic currents are

$$j_L^\mu = \bar{e}\gamma^\mu(1-\gamma^5)\nu_{eL} \qquad j_R^\mu = \bar{e}\gamma^\mu(1+\gamma^5)\nu_{eR}. \tag{6.116}$$

The electron neutrinos are superpositions of mass eigenstates (see (6.91)). The hadronic currents $J_{L/R}^\mu$ could be expressed analogously in terms of quarks. Often, however, nucleon currents are used in a non-relativistic approximation. In this so-called impulse approximation, the nucleons in the nucleus are assumed to behave like free nucleons.

The $0\nu\beta\beta$ decay is given in a general form ($\hbar = c = 1$) by

$$\mathrm{d}\lambda_{0\nu} = 2\pi \sum_{\text{spin}} |R_{0\nu}|^2 \delta(E_{e_1} + E_{e_2} + E_f - M_i)\,\mathrm{d}^3\boldsymbol{p}_{e_1}\,\mathrm{d}^3\boldsymbol{p}_{e_2}. \tag{6.117a}$$

E_f is the energy of the daughter nucleus, $R_{0\nu}$ denotes the transition amplitude which contains both the lepton component and the nuclear matrix element [Doi85]. The leptonic component of the amplitude is a product of two left- or right-handed currents. For a purely left-handed (or right-handed) interaction only the mass term contributes. The case of a purely right-handed interaction is neglected, since this contribution depends quadratically on the small coupling constants η and λ. When left- and right-handed currents interfere the decay amplitude contains the component of the neutrino propagator proportional to the four-momentum q, which describes the exchange of the virtual neutrino between two nucleons. This neutrino propagator occurs as the so-called neutrino potential in the matrix elements.

The decay amplitude of the 0ν mode is calculated in second-order perturbation theory. The transition probability for the neutrinoless double-beta decay is obtained as the square of an amplitude with contributions proportional to the neutrino mass and to the parameters of the right-handed leptonic admixtures

[8] In the framework of gauge theories the two latter conditions cannot be considered as independent (as mentioned already in section 6.2.2). In this case a right-handed component is possible only in connection with a non-vanishing Majorana mass (see also [Moh86a, Ros88]).

to the weak interaction. The inverse of the half-life for $(0_i^+ \to 0_f^+)$ transitions is given by the expression [Doi85, Mut88a]

$$[T_{1/2}^{0\nu}(0_i^+ \to 0_f^+)]^{-1} = C_{mm}\left(\frac{\langle m_\nu\rangle}{m_e}\right)^2 + C_{\eta\eta}\langle\eta\rangle^2 + C_{\lambda\lambda}\langle\lambda\rangle^2 + C_{m\eta}\frac{\langle m_\nu\rangle}{m_e}\langle\eta\rangle + C_{m\lambda}\frac{\langle m_\nu\rangle}{m_e}\langle\lambda\rangle + C_{\eta\lambda}\langle\eta\rangle\langle\lambda\rangle. \quad (6.117b)$$

Here, CP conservation is assumed. The effective values of the neutrino mass and the right-handed parameters are defined by

$$\langle m_\nu\rangle = \sum_j{}' m_j U_{ej}^2 \quad \langle\eta\rangle = \eta\sum_j{}' U_{ej}V_{ej} \quad \langle\lambda\rangle = \lambda\sum_j{}' U_{ej}V_{ej}. \quad (6.118)$$

The sum $\sum_j{}'$ is taken over light neutrino states with $m_j < 10\ \mathrm{MeV}/c^2$. For heavier neutrinos, the effect of the mass m_j on the neutrino potential can no longer be neglected, so that the matrix elements become mass dependent[9] (see [Mut88a, 89b]). From the assumption of CP conservation it follows that η and λ are real. U_{ej} and V_{ej} are both either real or purely imaginary, depending on the CP eigenvalue of the Majorana mass eigenstate [Kay84]. Consequently, all three parameters, $\langle m_\nu\rangle$, $\langle\eta\rangle$ and $\langle\lambda\rangle$, which characterize the lepton number, are assumed to be real in what follows.

The ratio $R = \langle\lambda\rangle/\langle\eta\rangle$, which is independent of the amplitudes V_{ej}, is, under certain assumptions (see [Suh93]), a simple function of $K = (M_{W_L}/M_{W_R})^2$, where M_{W_L}, M_{W_R} are the masses of the left- and right-handed W bosons, and of the mixing angle θ between left- and right-handed W bosons (see (6.93a) and (6.93b)).

The coefficients C_{xy} consist of nuclear matrix elements and phase-space integrals. Readers are referred to [Mut88a, 89b, Tom91] for a detailed discussion of these quantities. When the right-handed currents are neglected ($\lambda = \eta = 0$) only the coefficient C_{mm} occurs. This is given explicitly by

$$C_{mm} = (M_{\mathrm{GT}} - M_{\mathrm{F}})^2 G_1. \quad (6.119)$$

The phase-space integral G_1 is given by

$$G_1 \sim \int_{m_e}^{Q_{\beta\beta}+m_e} F(Z, E_{e_1})p_{e_1}E_{e_1}F(Z, E_{e_2})p_{e_2}E_{e_2}\,\mathrm{d}E_{e_1} \quad (6.120)$$

[9] The effective mass then becomes $\langle m_\nu\rangle = \sum_j m_j U_{ej}^2 R(m_j)$, where for heavy neutrinos with masses beyond 1 GeV $R(m_\nu)$ decreases proportional to m_ν^{-2}. We have $R(m_\nu) = 3.2 \times (10^8\ \mathrm{eV}/m_\nu)^2$ with m_ν in eV, almost independent of the particular emitter.

where $Q_{\beta\beta} = E_{e_1} + E_{e_2} - 2m_e$. G_1 may be calculated using the following approximation for the Fermi function

$$F(Z, E) = \frac{E}{p} \frac{2\pi Z\alpha}{1 - \exp(-2\pi Z\alpha)} \tag{6.121}$$

which is obtained from (6.39) using $\beta = p/E$. It follows that

$$G_1 \sim \left(\frac{Q_{\beta\beta}^5}{30} - \frac{2Q_{\beta\beta}^2}{3} + Q_{\beta\beta} - \frac{2}{5} \right). \tag{6.122}$$

This should be compared with the dependence of the phase-space factor on $Q_{\beta\beta}^{11}$ for the 2ν mode.

The Gamow–Teller and Fermi matrix elements in (6.119) are

$$M_{\mathrm{GT}} = \sum_{m,n} \langle 0_f^+ \| t_{-m} t_{-n} \sigma_m \cdot \sigma_n H(r) \| 0_i^+ \rangle \tag{6.123}$$

$$M_{\mathrm{F}} = \sum_{m,n} \langle 0_f^+ \| t_{-m} t_{-n} H(r) \| 0_i^+ \rangle \left(\frac{g_{\mathrm{V}}}{g_{\mathrm{A}}} \right)^2 \tag{6.124}$$

where $r = |r_m - r_n|$. In addition to the known transition operators, these expressions also involve the neutrino potential $H(r)$ which describes the exchange of the virtual neutrino. For light neutrinos ($m_j < 10$ MeV/c^2) $H(r)$ behaves like a $1/r$ potential. For heavy neutrinos $H(r)$ is like a Yukawa potential $H(r) \sim e^{-mr}/r$ (see e.g. [Gro86c]). Unlike in the 2ν mode, because of the effect of the propagator, Fermi transitions also occur.

The matrix elements M_{GT} and M_{F} follow from the mass term, where the two electrons are emitted as s waves. Unlike in the $2\nu\beta\beta$ decay, the decaying nucleons in the $0\nu\beta\beta$ decay are spatially correlated, so that the transition is increasingly likely the closer the two nucleons are together. This means that long-range correlations only have a small effect. The 0ν mode is essentially determined by short-range correlations as produced by the pairing force. For very small distances $r < 1$ fm, the finite extent of the nucleons, which we have up to now assumed to be point-like, and their short-range mutual repulsion become noticeable. These effects must be taken account of in the nuclear structure calculations.

6.2.4.3 *The decay rates for the majoron-accompanied $0\nu\beta\beta$ decay ($0\nu\chi$ and $0\nu\chi\chi$ decay)*

The half-life for $0\nu\chi$ decay is given by [Doi88]

$$T_{1/2}^{-1} = |M_{\mathrm{GT}} - M_{\mathrm{F}}|^2 F^{0\nu\chi} |\langle g_{\nu\chi} \rangle|^2. \tag{6.125}$$

Here, the neutrino–majoron coupling constant $\langle g_{\nu\chi} \rangle$ is

$$\langle g_{\nu\chi} \rangle = \sum_{i,j} g_{\nu\chi} U_{ei} U_{ej} \tag{6.126}$$

and $F^{0\nu\chi}$ denotes the phase space (for the latter, see [Doi88]). Neglecting the exchange of heavy neutrinos, the nuclear matrix elements are the same as for 0ν decay (see previous section).

The half-life of the $0\nu\chi\chi$ decay is given by [Moh91a]

$$T_{1/2}^{-1} = (f_{\chi\chi} - m_e)^2 |M_{\mathrm{GT}} - M_{\mathrm{F}}|^2 F^{0\nu\chi\chi}. \tag{6.127a}$$

Here, the coupling constant $f_{\chi\chi}$ is related to the mass of the zino (the SUSY partner of the Z^0 boson) by

$$f_{\chi\chi} = \frac{g^2}{4 M_{\tilde{Z}} \cos\theta_{\mathrm{W}}}. \tag{6.127b}$$

The nuclear matrix elements again correspond to those of the 0ν decay. The phase-space factors $F^{0\nu\chi\chi}$ are given for various $\beta\beta$ emitters in [Moh88a].

6.2.4.4 *The decay rates for $0\nu\beta\beta$ decay induced by exchange of supersymmetric particles*

From the SUSY Feynman graphs as in figure 6.20 one derives [Hir95]

$$\left[T_{1/2}^{0\nu}(0^+ \to 0^+)\right]^{-1} \sim G_{01} \left(\frac{\lambda_{111}'^2}{m_{\tilde{q},\tilde{e}}^4 \, m_{\tilde{g}\chi}} M \right)^2 \tag{6.128}$$

where G_{01} is a phase-space factor, $m_{\tilde{q},\tilde{e},\tilde{g},\chi}$ are the masses of supersymmetric particles involved: squarks, selectrons, gluinos and neutralinos. λ'_{111} is the strength of an R-parity breaking interaction. M denotes a nuclear matrix element. Investigation of double-beta decay thus allows us to restrict the SUSY-model parameter space (see section 6.2.5.4).

6.2.4.5 *Effects of nuclear structure on the matrix elements of $\beta\beta$ decay*

In the framework of modern theories the neutrinoless double-beta decay provides sensitive information about the Majorana nature of the neutrino, about the neutrino mass and the right-handed admixtures in the charged weak currents, and also about parameters of more exotic models (see sections 6.2.3 and 6.2.4.3, and 6.2.4.4). However, in order that we may deduce the parameters $\langle m_\nu \rangle$, $\langle \eta \rangle$ and $\langle \lambda \rangle$ or neutrino–majoron coupling constants, etc from observed $0\nu\beta\beta$ half-lives, the corresponding nuclear matrix elements entering into equations

(6.117*b*), (6.125) and (6.127*a*), must be theoretically calculated. The significance of the information provided by experiments on $0\nu\beta\beta$ decay is thus crucially dependent on the reliability of the nuclear structure calculations.

In addition to the determination of the $0\nu\beta\beta$ matrix elements, the calculation of the $2\nu\beta\beta$ decay is also of interest. This is described by the conventional theory of the weak interaction. The $2\nu\beta\beta$ half-lives are directly related to the nuclear matrix elements by (6.113), and no free parameters are involved as far as particle physics is concerned. Comparison of experimentally measured and predicted $2\nu\beta\beta$ decay rates consequently represents a direct and sensitive test for the various nuclear structure models.

The problem that the predicted transition probabilities for the 2ν mode were a factor of 50 to 100 times too large, in comparison with the experimental values, was long standing. Studies of the dependence of the decay amplitudes on the complexity of the ground-state wavefunction showed that the detailed knowledge of the wavefunction of the ground state, in particular ground-state correlations, played a decisive role [Kla84, Gro85b,c, 86c]. Despite this realization which opened the way to the final solution of the problem, it was initially not possible to give a *complete* explanation of the experimentally observed suppression of the matrix elements. The breakthrough came in calculations using the so-called QRPA model (Quasiparticle Random Phase Approximation) taking into account the so-called particle–particle interaction which had previously been neglected. We refer readers who are interested to the original literature [Vog86, Civ87, Eng88, Mut88b, 89a,b, Sta90a,b, Wu91, Hir93a,b, 94, 94a].

In the $2\nu\beta\beta$ decay only virtual 1^+ states in the intermediate nuclei occur which are reached solely via Gamow–Teller transitions. Using a schematic force (δ force) it was shown [Vog86] that it is possible to reduce the $2\nu\beta\beta$ decay by adjusting two parameters. This suppression mechanism was confirmed in calculations with a realistic residual interaction derived from established nucleon–nucleon interactions, namely the Bonn and the Paris potentials [Civ87, Mut88b, 89a,b, Sta90a,b]. However, the results for the 2ν mode are very sensitive to the strength of the particle–particle interaction, which is not known very precisely. A complete calculation of the $2\nu\beta\beta$ and $0\nu\beta\beta$ half-lives of all potential double-beta emitters with $A \geq 70$ is given in [Gro86c, Sta90a, Hir94a]. The results of [Sta90a] are restated in table 6.5. These half-lives are systematically longer than those given by previous calculations [Hax82, Gro86c].

The calculated $2\nu\beta\beta$ half-lives are consistent with the known experimental data (see figure 6.21). It has been largely possible to reproduce the measured half-lives. Another advance in the calculation of the 2ν mode has recently been achieved with the application of the operator expansion method which seems to give some decrease in the dependence of the matrix elements on the strength of the particle–particle force [Wu91, 92, Hir93b, 94a] (see figure 6.22).

In the 0ν mode, virtual intermediate states with $J^\pi \neq 1^+$ must also be

Table 6.5. Predicted $2\nu\beta\beta$ and $0\nu\beta\beta$ half-lives for all $\beta\beta$ emitters with $A \geq 70$ (from [Sta90a]). $\overline{T}^{2\nu}_{1/2}$ is averaged over a physically reasonable range of the strength of the particle–particle component of the nuclear force (see [Sta90a]).

Nuclide	$\overline{T}^{2\nu}_{1/2}$ [years]	$T^{0\nu}_{1/2}\langle m_\nu\rangle^2$ [years eV^2]
^{70}Zn	3.99×10^{22}	9.83×10^{25}
^{76}Ge	2.99×10^{21}	2.33×10^{24}
^{80}Se	1.18×10^{30}	1.14×10^{27}
^{82}Se	1.09×10^{20}	6.03×10^{23}
^{86}Kr	6.93×10^{23}	2.78×10^{25}
^{94}Zr	6.93×10^{22}	3.97×10^{25}
^{96}Zr	1.08×10^{19}	5.30×10^{23}
^{98}Mo	2.96×10^{30}	1.05×10^{27}
^{100}Mo	1.13×10^{18}	1.27×10^{24}
^{104}Ru	6.29×10^{21}	4.24×10^{24}
^{110}Pd	1.16×10^{19}	1.96×10^{24}
^{114}Cd	1.78×10^{24}	5.07×10^{25}
^{116}Cd	6.31×10^{19}	4.87×10^{23}
^{122}Sn	5.44×10^{27}	1.27×10^{26}
^{124}Sn	5.25×10^{21}	1.36×10^{24}
^{128}Te	2.63×10^{24}	7.77×10^{24}
^{130}Te	1.84×10^{21}	4.89×10^{23}
^{134}Xe	6.09×10^{24}	1.69×10^{25}
^{136}Xe	4.64×10^{21}	2.21×10^{24}
^{142}Ce	1.58×10^{21}	2.77×10^{24}
^{146}Nd	—	4.36×10^{26}
^{148}Nd	5.17×10^{19}	1.36×10^{24}
^{150}Nd	7.37×10^{18}	3.37×10^{22}
^{154}Sm	1.80×10^{21}	1.39×10^{24}
^{160}Gd	4.21×10^{20}	8.56×10^{23}
^{170}Er	2.68×10^{25}	1.37×10^{25}
^{176}Yb	9.83×10^{21}	1.36×10^{24}
^{186}W	1.07×10^{25}	6.35×10^{24}
^{192}Os	1.94×10^{25}	4.08×10^{24}
^{198}Pt	3.53×10^{23}	4.70×10^{23}
^{204}Hg	1.87×10^{27}	8.22×10^{24}
^{232}Th	1.60×10^{23}	3.07×10^{23}
^{238}U	1.53×10^{23}	2.60×10^{23}
^{244}Pu	6.54×10^{22}	5.72×10^{23}

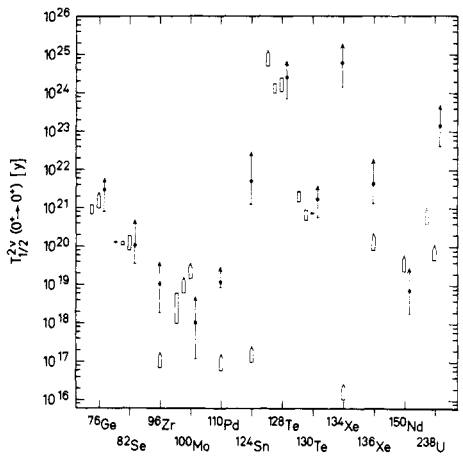


Figure 6.21. Comparison of $2\nu\beta\beta$ half-lives calculated using QRPA including particle–particle forces (described by g_{pp}) (thin arrows) with experimental values (thick arrows, stripes). The calculated arrows begin at the lower bounds for $T_{1/2}$, the circle denotes the average value of g_{pp} over the 1σ area. The upwards arrow indicates that the upper bound in the calculation is infinite (from [Mut88a, Gro89, 90]).

considered (see figure 6.23 and e.g. [Mut89b]). One consequence of the neutrino potential (see equations (6.123), (6.124)) is that long-range nucleon–nucleon correlations are almost ineffective in the $0\nu\beta\beta$ decay. Thus, unlike in the 2ν mode the nuclear matrix elements are only slightly reduced when, in addition to the pairing force, one takes account of longer-range components of the nuclear interaction [Kla84, Gro85b]. The particle–particle force leads to a reduction, which is, however, only small in comparison with that for the $2\nu\beta\beta$ decay [Tom87, Mut89b, Sta90a,b].

The nuclear matrix elements for the 0ν mode, unlike those for the 2ν mode, thus depend only very slightly on the details of the nuclear wavefunctions, which are of crucial importance for an understanding of the $2\nu\beta\beta$ decay.

More precisely, this is seen as follows: for light neutrinos, the following holds for the neutrino potential $H(r) \sim 1/r$. Using the expansion

$$\frac{1}{|\boldsymbol{r}_m - \boldsymbol{r}_n|} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\alpha) \tag{6.129}$$

one sees that not only are 1^+ states in the intermediate nucleus populated, but also different multipolarities J^π contribute to the transition amplitude (see (6.123), (6.124) and figure 6.23) ($r_<$ and $r_>$ denote the smaller and larger of

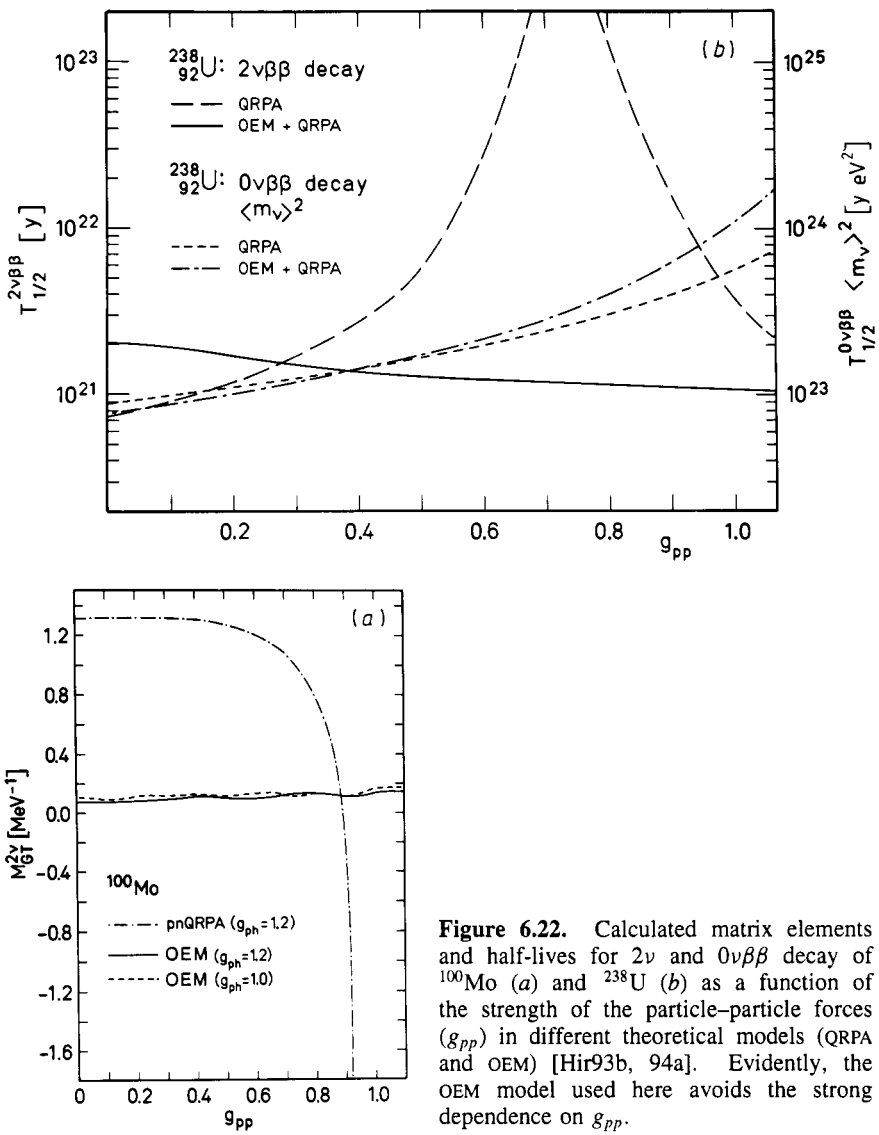


Figure 6.22. Calculated matrix elements and half-lives for 2ν and $0\nu\beta\beta$ decay of ^{100}Mo (*a*) and ^{238}U (*b*) as a function of the strength of the particle–particle forces (g_{pp}) in different theoretical models (QRPA and OEM) [Hir93b, 94a]. Evidently, the OEM model used here avoids the strong dependence on g_{pp}.

the distances r_m and r_n, respectively). As a consequence of the nature of the nucleon–nucleon interaction the influence of the particle–particle force for 1^+ intermediate states is particularly large, i.e. the 1^+ component is associated with particularly large uncertainties. The effect of the particle–particle force on the transition probability depends on the multipolarity. It is known that the

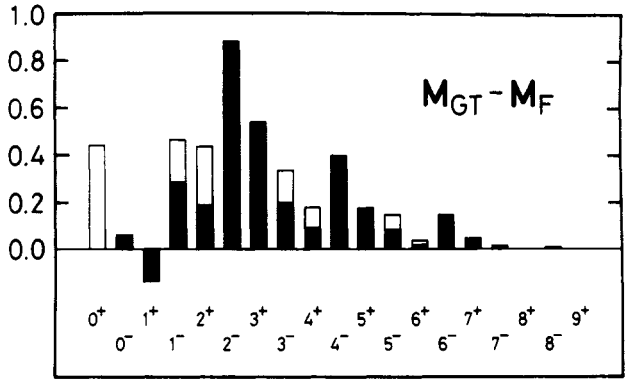


Figure 6.23. Decomposition of the nuclear matrix element $M_{GT} - M_F$ (see (6.123), (6.124)) into contributions of the intermediate states with spin and parity J^π for the $0\nu\beta\beta$ decay of ^{76} Ge. Open and filled histograms describe the contributions of $-M_F$ and M_{GT}, respectively (from [Mut89b]).

particle–particle interaction in the 1^+ channel is strongly attractive. This is seen, for example, from the fact that energetically low-lying 1^+ states are observed in odd–odd nuclei. For the higher multipoles the particle–particle force has a less attractive or slightly repulsive effect. This explains why the 1^+ component reacts in a particularly sensitive way to the introduction of this interaction component. Since $2\nu\beta\beta$ transitions only involve 1^+ states, the key role for the 2ν mode is apparent.

The products of the $0\nu\beta\beta$ half-lives and the effective neutrino mass, obtained by neglecting the right-handed currents (see (6.117*b*) and (6.118))

$$T_{1/2}^{0\nu}\langle m_\nu\rangle^2 = \frac{m_e^2}{(M_{GT} - M_F)^2 G_1} \tag{6.130}$$

are given in the right column of table 6.5 (from [Sta90a]). G_1 is the phase-space integral defined in (6.120). The relatively good agreement between theory and experiments for the neutrino-accompanied double-beta decay and the β^+ decay of proton-rich nuclei, in which the same suppression mechanism is effective [Suh88, Sta90c, Hir93a], as well as further investigations of the dependence on the choice of the nucleon–nucleon potential and of renormalization effects [Sta90b, 92b], gives one confidence that nuclear matrix elements for the 0ν mode are reliable (within a factor of 2–3). Thus reliable determinations of an (effective) neutrino mass and of right-handed coupling constants are possible from $0\nu\beta\beta$ decay. Figure 6.24 shows the estimated uncertainties of theoretically calculated half-lives of $0\nu\beta\beta$ decay.

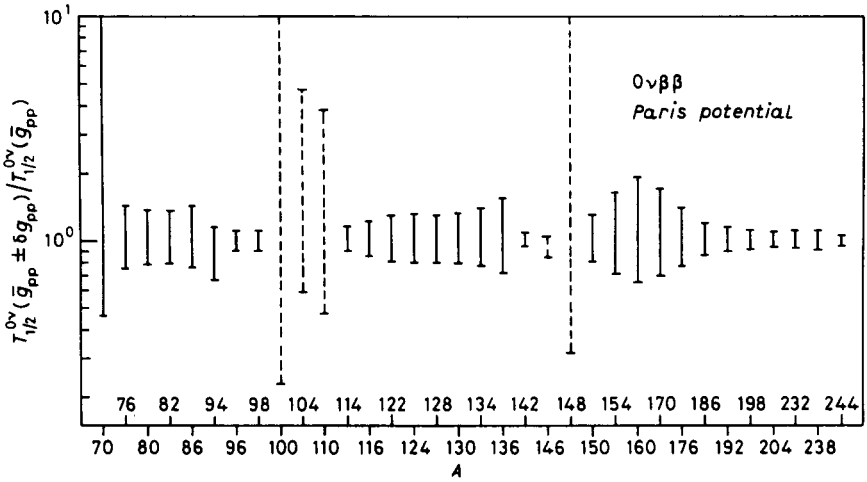


Figure 6.24. The uncertainty of $0\nu\beta\beta$ half-lives calculated using QRPA, resulting from the limited knowledge of the particle–particle force (g_{pp}) for potential double-beta emitters (from [Sta90a]).

Table 6.6. Calculated $2\nu\beta\beta$ and $0\nu\beta\beta$ half-lives for $\beta^+\beta^+$ emitters (from [Sta91]). For definition of $\overline{T}^{2\nu}_{1/2}$ see table 6.5.

Nuclide	$Q_{\beta^+\beta^+}$ [MeV]	$\overline{T}^{2\nu}_{1/2}$ [years]	$T^{0\nu}_{1/2}\langle m_\nu\rangle^2$ [years eV^2]
^{78}Kr	0.833	1.93×10^{26}	9.32×10^{27}
^{96}Ru	0.677	5.31×10^{26}	4.86×10^{28}
^{106}Cd	0.734	4.94×10^{25}	3.20×10^{28}
^{124}Xe	0.822	8.17×10^{25}	6.58×10^{27}
^{130}Ba	0.538	1.37×10^{29}	2.03×10^{29}
^{136}Ce	0.365	4.51×10^{31}	5.17×10^{30}
^{148}Gd	1.024	5.81×10^{26}	1.63×10^{28}

The matrix elements and the half-lives derived from these have recently also been calculated for the decay modes of $\beta^+\beta^+$ decay [Sta91, Hir94]. Table 6.6 summarizes the results of [Sta91] for the seven potential double-beta emitters. The decay rates are usually strongly suppressed in comparison with those of the $\beta^-\beta^-$ decay; exclusions are $2\nu\beta^+EC$ and $2\nu ECEC$ decays [Hir94]. This is mainly a result of the Coulomb barrier to the emission of positrons and the low Q values for these decays. For the potential importance of β^+EC decay for neutrino mass determination by $\beta\beta$ decay see section 6.2.5.5.

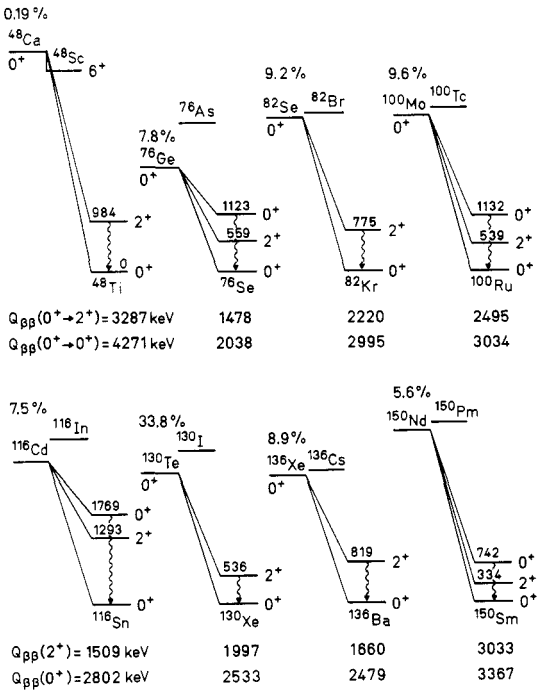


Figure 6.25. Some of the $\beta\beta$ emitters most studied at present.

6.2.5 Experiments on double-beta decay

A number of the $\beta^-\beta^-$ emitters which are currently subject to the most intensive study are shown in figure 6.25. Because of the typically extremely long half-life of 10^{20} years and more, detection of double-beta decay is associated with considerable experimental difficulties. In particular, the background of the cosmic radiation and radioactive impurities represent serious problems. Thus, most experiments are carried out in underground laboratories in order to reduce the muon flux from cosmic radiation (figure 6.26, see also figure 4.5). Table 6.7 gives the conditions for muon and neutron flux for a number of the most important laboratories.

In addition, there are limited possibilities for studying large amounts of source material since one must essentially rely on the nuclides listed in tables 6.5 and 6.6 (^{48}Ca is another important candidate). Whilst when studying proton decay it is possible to select favourable materials such as water from which to construct giant experiments with several thousand tonnes of source material, double-beta experiments have until now been restricted to a maximum of several kilograms. In addition, the decay energy of a few MeV is much lower than for

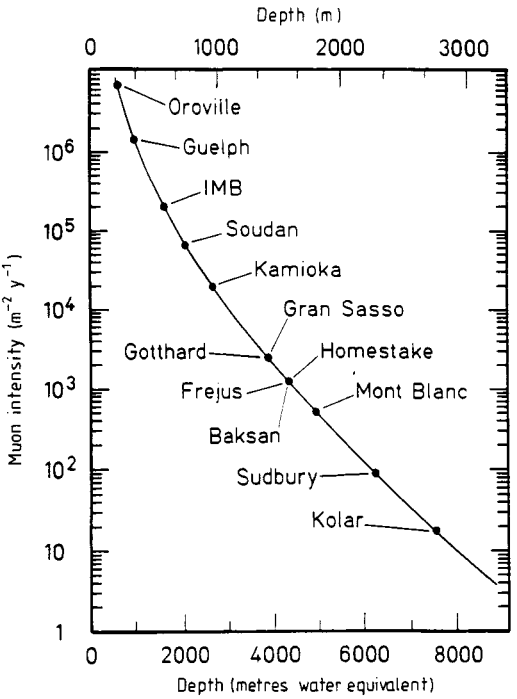


Figure 6.26. Some of the most important underground laboratories, in which $\beta\beta$ experiments are carried out, and their depth (in metres of water equivalent), together with the attenuation of the muon flux from the cosmic radiation.

Table 6.7. μ- and n-flux (unscreened) for a number of underground laboratories (after Yu Zdesenko).

Laboratory	Depth [m. w. e.]	μ flux [$m^{-2}\,d^{-1}$]	n flux [$cm^{-2}\,s^{-1}$]
Mont Blanc	5000	0.7	2.2×10^{-5}
Gran Sasso	3500	16	5.3×10^{-6}
Fréjus	4000	8	$< 3 \times 10^{-5}$
Broken Hill silver mine	3300	(20)	—
Solotvina salt mine	1000	1.5×10^{3}	$< 2.7 \times 10^{-6}$
Baksan	660	7×10^{3}	3×10^{-5}
Windsor salt mine	650 (350 m)	(7×10^{3})	—

Table 6.8. Half-lives from various experiments and the upper bounds for $\langle m_\nu \rangle$ calculated from them (the calculation used the matrix elements of [Sta90a] neglecting a right-handed weak interaction). The entries denoted by * refer to radiochemical experiments, those denoted by † refer to geochemical experiments. For a complete listing, see [Tre95].

Decay	Half-lives		Upper bound	Reference
	$T^{0\nu}_{1/2}$ [years]	$T^{2\nu}_{1/2}$ [years]	$\langle m_\nu \rangle$ [eV]	
$^{46}_{20}\mathrm{Ca}\rightarrow{}^{46}_{22}\mathrm{Ti}$	$>6.9\times10^{11}$(68%)			[Frem50]
$^{48}_{20}\mathrm{Ca}\rightarrow{}^{48}_{22}\mathrm{Ti}$	$>9.5\times10^{21}$(76%)	$>3.6\times10^{19}$(80%)	12.8[a](76%)	[Key91, Bar70]
$^{70}_{30}\mathrm{Zn}\rightarrow{}^{70}_{32}\mathrm{Ge}$	$>4.8\times10^{14}$(68%)			[Frem50]
$^{76}_{32}\mathrm{Ge}\rightarrow{}^{76}_{34}\mathrm{Se}$	$>1.1\times10^{25}$(90%)	$\left(1.77^{+0.01+0.13}_{-0.01-0.11}\right)\times10^{21}$	0.46(90%)	[Bau97, Kla97]
$^{82}_{34}\mathrm{Se}\rightarrow{}^{82}_{36}\mathrm{Kr}$	$>2.7\times10^{22}$(68%)	$\left(1.1^{+0.3}_{-0.1}\right)\times10^{20}$(68%)	4.7(68%)	[Ell92]
$^{94}_{40}\mathrm{Zr}\rightarrow{}^{94}_{42}\mathrm{Mo}$	$>1.5\times10^{16}$(68%)	$>6.0\times10^{15}$(68%)		[Bar90d]
$^{96}_{40}\mathrm{Zr}\rightarrow{}^{96}_{42}\mathrm{Mo}$	$>2.6\times10^{19}$(68%)	$>1.0\times10^{17}$(68%)	135(68%)	[Zde81, Bar90d]
$^{100}_{42}\mathrm{Mo}\rightarrow{}^{100}_{44}\mathrm{Ru}$	$>4.4\times10^{22}$(68%)	$(0.95\pm0.04\pm0.09)\times10^{19}$ (90%)	5.4(68%)	[Als93, Das95]
$^{110}_{46}\mathrm{Pd}\rightarrow{}^{110}_{48}\mathrm{Cd}$	$>6.0\times10^{16}$(68%)	$>6.0\times10^{16}$(68%) $>6.0\times10^{17}$(68%)		[Bar87] [Win52]
$^{114}_{48}\mathrm{Cd}\rightarrow{}^{114}_{50}\mathrm{Sn}$	$>2.0\times10^{20}$(90%)	$>9.2\times10^{16}$(99%)		[Dan95]
$^{116}_{48}\mathrm{Cd}\rightarrow{}^{116}_{50}\mathrm{Sn}$	$>2.9\times10^{22}$(90%)	$\left(2.7^{+0.5\,+0.9}_{-0.4\,-0.6}\right)\times10^{19}$(90%)	4.1(90%)	[Dan95, Arn95]
$^{128}_{52}\mathrm{Te}\rightarrow{}^{128}_{54}\mathrm{Xe}$	$>1.3\times10^{19}$(90%)	$\left(7.7^{+0.4}_{-0.4}\right)\times10^{24\dagger}$(68%)	1.1[b](68%)	[Mit88, Ber92a,93]
$^{130}_{52}\mathrm{Te}\rightarrow{}^{130}_{54}\mathrm{Xe}$	$>1.8\times10^{22}$(90%)	$\left(2.7^{+0.1}_{-0.1}\right)\times10^{21\dagger}$(68%)	5.3(90%)	[Ale94, Ber92a]
$^{134}_{54}\mathrm{Xe}\rightarrow{}^{134}_{56}\mathrm{Ba}$	$>8.2\times10^{19}$(68%)	$>1.1\times10^{16}$(68%)	454(68%)	[Bar89c]
$^{136}_{54}\mathrm{Xe}\rightarrow{}^{136}_{56}\mathrm{Ba}$	$>3.4\times10^{23}$(90%)	$>2.1\times10^{20}$(90%)	2.6(90%)	[Vui93]
$^{150}_{60}\mathrm{Nd}\rightarrow{}^{150}_{62}\mathrm{Sm}$	$>2.1\times10^{21}$(90%)	$\left(1.8^{+0.66}_{-0.39}\pm0.19\right)\times10^{19}$(90%)	4.1(90%)	[Moe95, Art95]
$^{186}_{74}\mathrm{W}\rightarrow{}^{186}_{76}\mathrm{Os}$	$>2.7\times10^{20}$(90%)	$>5.9\times10^{17}$ (90%)	150 (90%)	[Dan95]
$^{192}_{76}\mathrm{Os}\rightarrow{}^{192}_{78}\mathrm{Pt}$	$>9.8\times10^{12}$(68%)			[Frem50]
$^{198}_{78}\mathrm{Pt}\rightarrow{}^{198}_{80}\mathrm{Hg}$	$>3.2\times10^{14}$(68%)			[Frem50]
$^{238}_{92}\mathrm{U}\rightarrow{}^{238}_{94}\mathrm{Pu}$		$\left(2.0^{+0.6}_{-0.6}\right)\times10^{21}$ *(68%)	11.4	[Tur91]
$^{244}_{94}\mathrm{Pu}\rightarrow{}^{244}_{96}\mathrm{Cm}$		$>1.1\times10^{18}$ *(68%)		[Moo92]

[a] Calculated using [Mut91].
[b] Deduced from geochemical experiment.

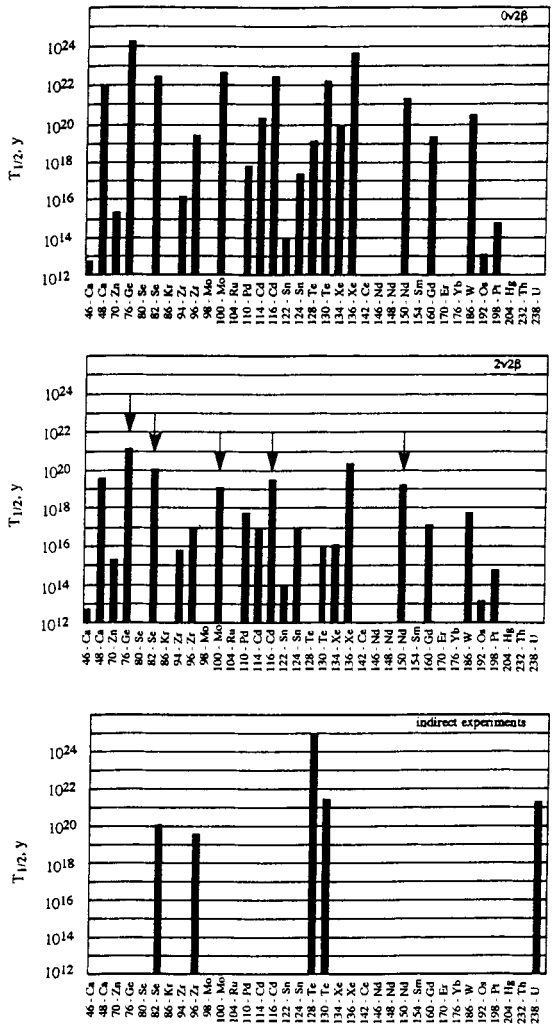


Figure 6.27. Measured 0ν and $2\nu\beta\beta$ half-life bounds (or half-lives as marked by arrows in middle part) for $\beta^-\beta^-$ emitters (g.s. $\rightarrow$ g.s. transitions). Top and middle: direct experiments. Bottom: indirect experiments (geochemical and radiochemical). (From [Tre95].)

proton decay, so that the self-absorption in the source plays a much greater role.

The double-beta decay was first detected indirectly by geochemical methods [Kir67, 86c, Man86, Ber92a]. These experiments, restricted to the selenium and tellurium isotopes ^{82}Se and ^{128,130}Te, were not able to distinguish between the different types of decay ($2\nu, 0\nu$). The $2\nu\beta\beta$ decay of ^{82}Se was finally first detected directly in 1987 in a counter experiment [Ell87]. The measured decay rate agreed well with that determined geochemically. In the meantime, the $2\nu\beta\beta$ decays of ^{76}Ge [Mil90, Vas90, Avi91, Bal94, Kla94] and ^{100}Mo [Eji91] and of some other isotopes have been observed.

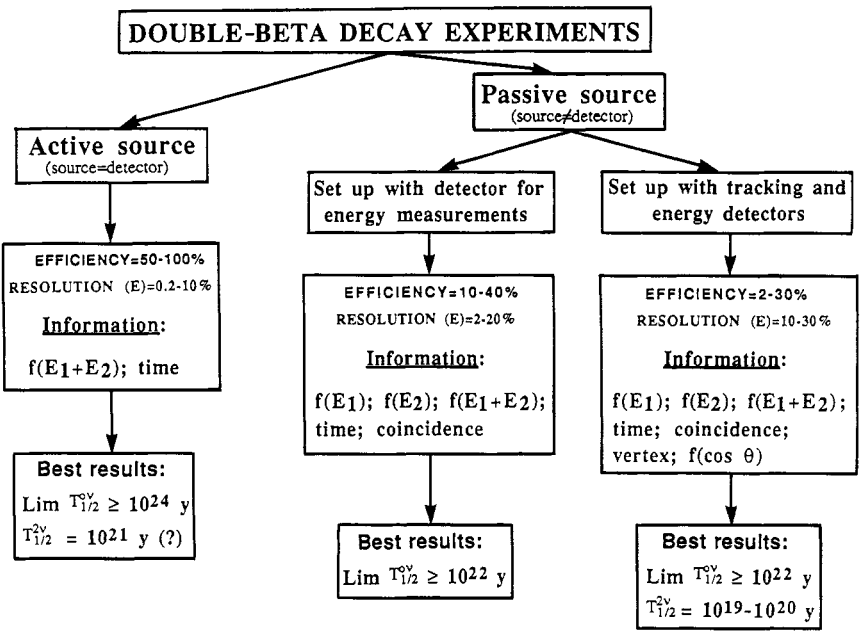


Figure 6.28. Classification of 'direct' $\beta\beta$ experiments (from [Kla92a, 94, Zde93]). There also exist geochemical and radiochemical experiments (see text).

Table 6.9. Some $\beta\beta$ emitters with their $Q_{\beta\beta}$ values and 0ν matrix elements (the latter from [Sta90a]).

Element	Energy $Q_{\beta\beta}$ [MeV]	$T^{0\nu}_{1/2} \times \langle m_\nu \rangle^2$ [years eV^2]
^{76}Ge	2.04	2.3×10^{24}
^{82}Se	3.0	6×10^{23}
^{96}Zr	3.3	5.3×10^{23}
^{100}Mo	3.03	1.3×10^{24}
^{116}Cd	2.8	4.9×10^{23}
^{130}Te	2.53	4.9×10^{23}
^{136}Xe	2.48	2.2×10^{24}
^{150}Nd	3.37	3.4×10^{22}

While the existence of the 2ν mode has been confirmed experimentally, the neutrinoless double-beta decay has not yet been detected. Until now, only lower bounds for the half-life have been given. Figure 6.27 and table 6.8 give

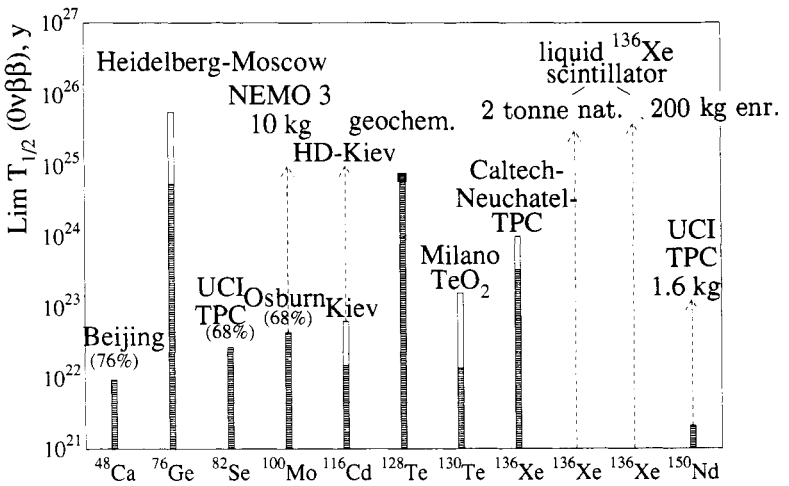


Figure 6.29. Present situation and prospects of the most promising $\beta\beta$ experiments for the attainable half-lives (for corresponding neutrino masses, see figure 6.40). Half-life bounds above 10^{21} years have only been obtained for the isotopes shown here. Shaded stripes correspond to the status in 1994, open stripes and broken lines refer to 'safe' and less 'safe' extrapolations to the year 2000 and beyond, respectively. For NEMO, liquid xenon and ^{150}Nd projects or plans see [Lal94, Rag94, Moe94a, 95].

overviews of the $0\nu\beta\beta$ and $2\nu\beta\beta$ half-lives or half-life limits obtained from measurements made to date.

In what follows, we shall discuss experimental methods for detecting these rare processes. As mentioned, there is an essential distinction between geochemical and radiochemical procedures and methods for direct observation of the decay in an appropriate detector. As far as the detector experiments are concerned, we further distinguish between experiments with active (source=detector) and passive sources (source≠detector) (figure 6.28). While the second class of experiments provides more complete information about the $\beta\beta$ events via measurements of coincidences, tracks and vertices of the electrons and their energy distribution, it is also clear (figure 6.27) that the greatest sensitivity is obtained for $0\nu\beta\beta$ decay with active sources, and in particular with enriched ^{76}Ge [Kla92a, 94] and ^{136}Xe [Won91, Vui93] (only the geochemical experiment with ^{128}Te achieves similar bounds). The main reason for this is that large source strengths can be used in conjunction with a much higher resolution. The best bounds on the $0\nu\beta\beta$ lifetime are greater than 5.6×10^{24} years. On the other hand, passive sources may be more promising as far as the detection of the 2ν decay is concerned.

Other criteria for the 'quality' of a $\beta\beta$ emitter include:

- a small product $T_{1/2}^{0\nu}\langle m_\nu\rangle^2$, i.e. a large matrix element $M^{0\nu}$ or phase space;
- a $Q_{\beta\beta}$ value above the energy limit for natural radioactivity (2.614 MeV).

Table 6.9 compares a number of important $\beta\beta$ emitters in terms of these criteria. The relatively small differences in the matrix elements should be noted (only ^{150}Nd is an exception, and thus a promising candidate for future application, see e.g. [Moe95]). Figure 6.29 shows the present situation and the perspectives of the most promising $\beta\beta$ experiments with respect to their sensitivity (see also figures 6.32, 6.40). Surveys of the experiments on double-beta decay can be found, e.g. in [Avi88, 89, Cal89, Moe91a, 93a,b, 94, 94a, 95, Kla92a, 93a–d, 94, 95a].

6.2.5.1 Geochemical experiments

The first experimental evidence for double-beta decay came from geochemical experiments. An enrichment of the daughter nuclei generated in $\beta\beta$ decay was noted in very old minerals containing a very large amount of the mother substance.

The geochemical method is based on the fact that the decay products from potential $\beta\beta$ emitters which are present in the mineral have accumulated over a very long time period, so that an enrichment of these nuclides in comparison with the normal isotope abundance is obtained. Because of the extremely long collection time (measurement time) such geochemical experiments are very sensitive. The decay rate $\lambda_{\beta\beta}$ is deduced from the age T of the rock sample and the measured abundances of the mother substance $N(Z, A)$ and the daughter substance $N(Z \pm 2, A)$ via the exponential decay law:

$$\lambda_{\beta\beta} \simeq \frac{N(Z \pm 2, A)}{N(Z, A)} \frac{1}{T}. \tag{6.131}$$

One major disadvantage is that only the decay product is detected, while the nature of its generation is not determined. Thus, it is impossible to distinguish between the 2ν and the 0ν mode. $\lambda_{\beta\beta}$ gives only the total decay rate, i.e. the sum of the decay rates of the two decay modes

$$\lambda_{\beta\beta} = \lambda_{2\nu} + \lambda_{0\nu}. \tag{6.132}$$

This means also that the sensitivity of such experiments for $0\nu\beta\beta$ decay cannot be increased beyond the half-life of $2\nu\beta\beta$ decay. The same is true for radiochemical experiments, to be discussed in the next section. In this sense these types of experiments have no future.

The samples used must satisfy certain geological and chemical conditions. The mineral must contain the nuclide under consideration in a high concentration.

In addition, one must make certain that a significant quantity of the daughter nuclide was not present before the rock was formed, independently of the double-beta decay, and that the concentration of the decay product has only altered as a result of $\beta\beta$ decay and not as a result of other effects such as the escape of volatile decay products. Furthermore, the age of the sample must be reliably determinable from the geological surroundings.

As a result of these requirements, the geochemical method is essentially restricted to selenium and tellurium ores. In both cases the daughter substances are volatile, chemically inert rare gases ($^{82}Se \rightarrow {}^{82}Kr$, $^{128,130}Te \rightarrow {}^{128,130}Xe$). Their concentration is heavily reduced during the crystallization of the minerals so that the initial abundances are very small. Because of the high sensitivity which may be achieved in studies of rare gases with a mass spectrometer, it is possible to detect a minute excess of daughter nuclei which has accumulated over geological time periods [Kir83b].

The first geochemical experiment was carried out in 1949 by Inghram and Reynolds [Ing49, 50]. The analysis of 1.5 billion-year-old tellurium ore showed an anomalous content of ^{130}Xe. The corresponding $\beta\beta$ half-life was determined as 1.4×10^{21} years [Ing50].

Table 6.10. Double-beta half-lives from geochemical experiments on selenium and tellurium ores.

$\beta\beta$ emitters	$T_{1/2}^{\beta\beta}$ [years]	Ref.
^{82}Se	$(1.30 \pm 0.05) \times 10^{20}$	[Kir86c]
	$(1.00 \pm 0.40) \times 10^{20}$	[Man86]
	$(1.2 \pm 0.1) \times 10^{20}$	[Lin88]
^{128}Te	$> 5 \times 10^{24}$	[Kir86c]
	$(7.7 \pm 0.4) \times 10^{24}$	[Ber92a, 93a]
^{130}Te	1.4×10^{21}	[Ing50]
	$(1.5 - 2.75) \times 10^{21}$	[Kir86c]
	$(2.7 \pm 0.1) \times 10^{21}$	[Ber92a]

The first convincing evidence of the existence of double-beta decay was given by Kirsten *et al* in studies of selenium and tellurium samples [Kir67, 68, 83, 86c]. The total $\beta\beta$ half-lives determined from geochemical experiments are summarized in table 6.10 (for an overview see also [Man91b]). The half-lives for ^{82}Se agree within the error limits and were recently confirmed by a direct counter experiment. The contradictions in the results for ^{128}Te have recently been resolved [Ber92a].

Although the geochemical experiments are not able to provide direct

information about the decay mode, a phase-space argument may be used to deduce from the ratio of the decay rates for the two tellurium isotopes that the observed half-life of ^{130}Te is determined by the 2ν mode [Pon68, Mut88a, b]. This reasoning is essentially based on the energy dependence of the phase-space factors ($G^{2\nu} \sim Q^{11}_{\beta\beta}$, $G^{0\nu}_1 \sim Q^5_{\beta\beta}$).

As far as the 0ν mode is concerned, from geochemical experiments it is only possible to obtain model-independent information in the form of an upper bound for the 0ν decay rate. It is practically impossible to obtain certain positive evidence for the *neutrinoless* double-beta decay.

6.2.5.2 *Radiochemical experiments*

Analogously to geochemical experiments, radiochemical experiments involve a search for an enrichment of the daughter nuclei in a sample. However, the radiochemical technique uses the advantage of radioactive daughter nuclides which are detectable in very much smaller quantities than stable rare gases. For this, one no longer requires geological integration periods and the measurements may be carried out over very much shorter timescales.

Following the purification of as large as possible a volume of the substance to be investigated, the decay products are accumulated over several years and their radioactive decay is detected. This method is independent of the uncertainties relating to the age of the rock sample, the original concentration of the daughter nuclide and possible diffusion effects of the rare gas over geological periods.

Two typical candidates for radiochemical experiments on double-beta decay are ^{232}Th and ^{238}U:

$$^{232}\mathrm{Th} \xrightarrow{\beta\beta} {}^{232}\mathrm{U}(Q_{\beta\beta} = 0.85\ \mathrm{MeV}) \qquad (6.133a)$$

$$^{238}\mathrm{U} \xrightarrow{\beta\beta} {}^{238}\mathrm{Pu}(Q_{\beta\beta} = 1.15\ \mathrm{MeV}). \qquad (6.133b)$$

Both daughter nuclei are α emitters with α half-lives of 70 years (^{232}U) and 87.7 years (^{238}Pu). One of the first experiments of this type was carried out in 1950 [Lev50]. Let us consider the case of ^{238}U in somewhat more detail.

^{238}U decays with a half-life of 4.5×10^9 years via α decay and also, with a much smaller probability, via spontaneous fission. The β decay to the neighbouring nucleus ^{238}Np is energetically forbidden. The double-beta decay leads to ^{238}Pu which again decays via α decay ($T_{1/2} = 87.74(9)$ years) and spontaneous fission (branching ratio $1.84(5) \times 10^{-9}$). As a result of the large kinetic energy of approximately 5.5 MeV of the α particle the α decay of ^{238}Pu has a unique signature. α detectors with an extremely low background are needed to detect this decay.

Levine *et al* [Lev50] carried out a corresponding investigation in which they extracted the plutonium from a six-year-old UO_3 sample. They could not find any 5.5 MeV α particles. This gave a lower bound of

$$T_{1/2}^{\beta\beta} > 6 \times 10^{18} \text{ years} \tag{6.134}$$

for the $\beta\beta$ half-life of ^{238}U. More recent measurements by Turkevich *et al* [Tur91] gave

$$T_{1/2}^{\beta\beta} = (2.0 \pm 0.6) \times 10^{21} \text{ years.} \tag{6.135}$$

This result agrees well with the theoretical expectation for $2\nu\beta\beta$ of ^{238}U [Wu92, Kla93b,d, 94, Hir94a] (see figure 6.22(*b*))

$$T_{1/2}^{\beta\beta} = 0.9 \times 10^{21} \text{ years} \tag{6.136}$$

so that the far-reaching conclusions drawn by [Tur91] from the result do not appear to be justified.

6.2.5.3 Counter experiments

The main advantage of counter experiments over geo- and radiochemical methods relates to the direct identification of the decay mode. The 2ν and 0ν decay are distinguished using the total energy spectrum of the two emitted electrons (positrons) (see figure 6.16). The $2\nu\beta\beta$ spectrum is continuous, since the energy is distributed across four leptons. For the $0\nu\beta\beta$ decay, we expect a sharp line at the upper end of the total energy spectrum at the maximum decay energy, since the sum of the kinetic energies of the two electrons corresponds to the Q value of the transition. The spectrum of the decay with the emission of one or two majorons would also be distinguishable from the spectrum of the $2\nu\beta\beta$ decay and that of the 0ν mode, as shown in figure 6.16.

The first counter experiment on double-beta decay was carried out in 1948 by Fireman [Fir48, 49]. The observation of two coincident electrons in a ^{124}Sn sample led to a half-life of $(4–9) \times 10^{15}$ years [Fir49] which was assigned to the 0ν mode. Later, however, more sensitive experiments showed that the recorded signal could be traced back to radioactive impurities.

We shall give a brief description of the most important experimental methods below.

Ionization chamber experiments. In this method the samples are placed in a gas detector in which the decay is observed. The major advantage is the ability to record the tracks of the two β particles involved. This results in a very effective suppression of background events. In addition, it is possible to determine the angular distribution of the electrons and the distribution of the individual and total energies, so that the double-beta decay has a unique signature.

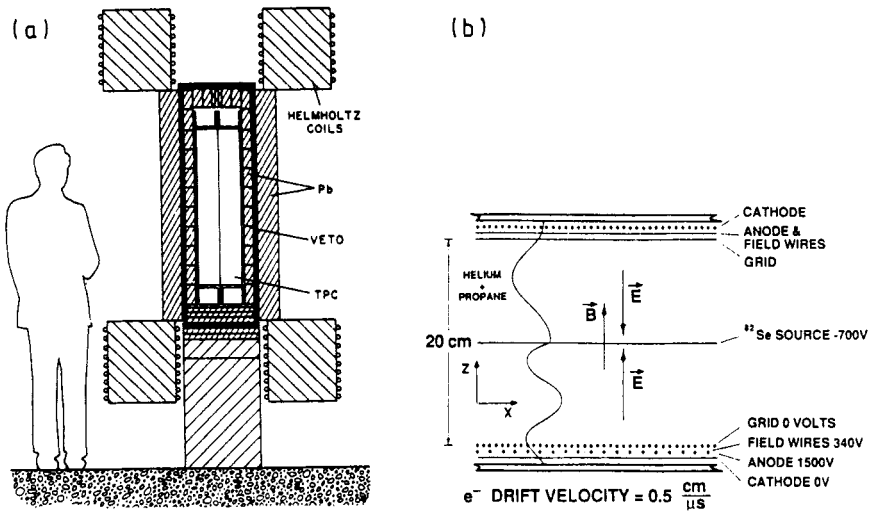


Figure 6.30. Schematic illustration of the $\beta\beta$ experiment at the University of Irvine, California [Ell87]. The right part shows a side view of the time projection chamber (TPC) (from [Avi88]).

This technology is limited by a relatively poor energy resolution of approximately 10% and a small probability of detection. Further, thin sources are needed to minimize the loss of energy of the β particles in the source. Consequently, in general, only small amounts of substances can be studied.

The first direct observation of a $\beta\beta$ event was achieved using a time projection chamber (TPC) [Ell87] (see figure 6.30). A sample foil with 14 g of isotope-enriched selenium (^{82}Se content 97%) was placed in the middle of the chamber which was screened against cosmic rays using lead and veto counters. The $\beta\beta$ electrons emitted from the foil in the gas-filled chamber (93% helium, 7% propane (C_3H_8)) leave an ionization track behind them, which is bent by an external magnetic field. The electrons released travel in the electrical field to the wire counters on the long sides, where they are recorded.

The energy and the starting point of the electron trajectories may thus be measured. The coincidence of two electrons in time and space is a condition for a true double-beta event. Other selection criteria exploit the energies of the electrons.

After a measurement time of 7960 hours, taking into account all background sources, 36 events had occurred which were assigned to the $2\nu\beta\beta$ decay of ^{82}Se. The half-life derived from this is [Ell87]

$$T_{1/2}^{2\nu}(^{82}\mathrm{Se}) = \left(1.1^{+0.8}_{-0.3}\right) \times 10^{20} \text{ years.} \tag{6.137}$$

(Earlier measurements with the same method had given a half-life which was shorter by a factor of 10 [Moe80]). This value agrees well with the results of the geochemical experiments. The bound for the 0ν mode is [Ell86]

$$T_{1/2}^{0\nu}(^{82}\mathrm{Se}) > 1.1 \times 10^{22} \text{ years.} \tag{6.138}$$

Consequently, the geochemically detected transition may essentially be traced back to the 2ν mode.

The same detector was also operated with a ^{100}Mo source. The following result was obtained [Ell91]

$$T_{1/2}^{2\nu}(^{100}\mathrm{Mo}) = (1.16^{+0.34}_{-0.08}) \times 10^{19} \text{ years.} \tag{6.139a}$$

A Japanese group from Osaka also studied the decay of ^{100}Mo with the ELEGANTS V detector (Electron gamma ray neutrino spectrometer V) which consists of a combination of drift chambers, plastic scintillators and NaI counters and is installed in the Kamioka mine (2700 m water equivalent). The source was a foil of enriched molybdenum (degree of enrichment of ^{100}Mo: 94.5%). The following results were obtained [Eji91, 92]

$$T_{1/2}^{2\nu}(^{100}\mathrm{Mo}) = (1.15^{+0.30}_{-0.20}) \times 10^{19} \text{ years} \tag{6.139b}$$

$$T_{1/2}^{0\nu}(^{100}\mathrm{Mo}) > 4.7 \times 10^{21} \text{ years.} \tag{6.139c}$$

For a more recent bound, see (6.157). A Soviet group gives a somewhat shorter half-life for the $2\nu\beta\beta$ decay of ^{100}Mo [Kli89]

$$T_{1/2}^{2\nu}(^{100}\mathrm{Mo}) = (3.3^{+2.0}_{-1.0}) \times 10^{18} \text{ years.} \tag{6.140}$$

We note that such experiments may also be used to study other substances such as ^{96}Zr, ^{130}Te and ^{150}Nd.

One particularly interesting variant of the gas detectors involves the use of xenon as a filling and counting gas in proportional counters, time projection chambers and drift chambers. Natural xenon contains both the $\beta\beta$ candidates ^{134}Xe ($Q_{\beta\beta} = 0.84$ MeV, natural isotopic abundance 10.4%) and ^{136}Xe ($Q_{\beta\beta} =$ 2.48 MeV, natural isotopic abundance 8.9%), and may be used both as source and detector at the same time. The idea of using the source as detector was first proposed in 1960 for studies of ^{136}Xe [Ant60].

The energy resolution is approximately an order of magnitude worse than in the germanium semiconductors to be discussed below; however, ^{136}Xe has the advantage of a somewhat larger $Q_{\beta\beta}$ value (see table 6.9). In addition, xenon can be relatively easily enriched in ultracentrifuges, so that experiments with large source strengths are possible.

The $\beta\beta$ decay of ^{136}Xe was investigated using scintillation counters [Bar86c] and ionization chambers [Bar89c]. A group in Milan installed a

multicell proportional counter in the Gran Sasso underground laboratory with an active volume of 79.4 l to study the double-beta decay of this nuclide. Samples with a natural isotopic composition and samples with an ^{136}Xe content of 64% were studied [Bel89, Bel91b]. The final evaluation of the data gave the following bound for $\beta\beta$ decay [Bel91b]

$$T_{1/2}^{2\nu}(^{136}\mathrm{Xe}) > 1.6 \times 10^{20} \text{ years (95\% c.l.).} \tag{6.141}$$

For the 0ν mode, Monte Carlo simulations of the angular distribution of the emitted electrons were used to derive separate bounds for the decay induced by the mass mechanism and right-handed currents, respectively

$$T_{1/2}^{0\nu}(^{136}\mathrm{Xe}) > 1.2 \times 10^{22} \text{ years (95\% c.l.) (mass mechanism)} \tag{6.142}$$

$$T_{1/2}^{0\nu}(^{136}\mathrm{Xe}) > 1.0 \times 10^{22} \text{ years (95\% c.l.) (r.h. currents).} \tag{6.143}$$

For the 0ν decay to the first excited 2^+ state of ^{136}Ba with an excitation energy of 818 keV, the measurement gave

$$T_{1/2}^{0\nu}(^{136}\mathrm{Xe}; 2^+) > 3.3 \times 10^{21} \text{ years (95\% c.l.).} \tag{6.144}$$

In the Swiss Gotthard underground laboratory (3000 m water equivalent screening thickness) a time projection chamber with an active volume of 180 l was built to study the $\beta\beta$ decay of ^{136}Xe [Won91, Tre91, Vui93]. The chamber was operated at a pressure of 5 atm and with a 62.5% level of enrichment. After a measurement time of 6830 hours no evidence for a neutrinoless double-beta decay was found. The bounds for the $(0^+ \rightarrow 0^+)$ transition are [Vui93]

$$T_{1/2}^{0\nu}(^{136}\mathrm{Xe}) > 3.4 \times 10^{23} \text{ years (90\% c.l.) (mass mechanism)} \tag{6.145}$$

$$T_{1/2}^{0\nu}(^{136}\mathrm{Xe}) > 2.6 \times 10^{23} \text{ years (90\% c.l.) (r.h. currents).} \tag{6.146}$$

A new proposal to detect the $\beta\beta$ decay of ^{136}Xe exploits the occurrence of the daughter nucleus as an additional signature. Although the stable daughter nucleus is generally difficult to detect, its atomic-physics related properties may be used. In particular, the ionized ^{136}Ba which is formed in the double-beta decay of ^{136}Xe could be detected via its laser fluorescence. The coincident detection of the daughter nuclide and the β particle would practically eliminate the background completely [Moe91b].

Larger drift chambers with highly enriched xenon were also planned for the Russian underground laboratory in Baksan (see e.g. [Kir88a]).

Semiconductor experiments. Because of their outstanding energy resolution semiconductors are excellently suited for study of $0\nu\beta\beta$ decay, which is manifest by a sharp line in the total energy spectrum. A particularly favourable

circumstance arises in the case of the $\beta\beta$ candidate ^{76}Ge. This germanium isotope occurs with an abundance of 7.8% in natural germanium, from which high-resolution detectors are manufactured. Thus, germanium can be used simultaneously as source and detector so that the problem of self-absorption does not arise.

This technique introduced by the Milan group in 1967 [Fio67] may be used to study a large number of ^{76}Ge nuclei with excellent efficiency. The resolution in the region of the decay energy of 2.04 MeV typically amounts to 3 keV.

On the other hand, for detecting the 2ν mode the extremely good energy resolution has less effect. Despite this the neutrino-accompanied double-beta decay of ^{76}Ge has now been detected [Mil90, Vas90, Avi91, Bal94, Kla94].

A Russian/Armenian group used three Ge(Li) detectors each of mass 0.5 kg which were introduced into a titanium cryostat system (the germanium crystals have to be cooled with liquid nitrogen) and surrounded by an NaI shield. Two of the crystals were enriched with the $\beta\beta$ emitter isotope ^{76}Ge to 85%. We shall discuss the importance of enrichment later in conjunction with the 0ν mode. The apparatus was operated in a salt mine in Yerevan at a depth of 645 m. To evaluate the data the difference spectrum between the natural and the enriched detectors, which contain a multiple of the ^{76}Ge $\beta\beta$ activity of the natural detector, was formed in order to eliminate the background. The spectrum obtained in this way is similar to the expected $2\nu\beta\beta$ spectrum. The half-life was given to be [Vas90]

$$T_{1/2}^{2\nu}(^{76}\mathrm{Ge}) = (9 \pm 1) \times 10^{20} \text{ years} \quad (68\% \text{ c.l.}). \tag{6.147}$$

An American collaboration (PNL–USC) studied the double-beta decay of germanium with two natural detectors in the Homestake gold mine. After a number of background corrections they also obtained positive evidence for the $2\nu\beta\beta$ decay of ^{76}Ge with a half-life of [Mil90]

$$T_{1/2}^{2\nu}(^{76}\mathrm{Ge}) = (1.12^{+0.48}_{-0.26}) \times 10^{21} \text{ years} \quad (95\% \text{ c.l.}). \tag{6.148}$$

A joint measurement by both groups [Avi91] used a Soviet enriched Ge crystal (enrichment 85%) of 0.25 kg in the cryostats of the PNL–USC collaboration. Initially 60 days were allowed to pass in order to permit the decay of possible short-lived cosmogenic radionuclides (i.e. radioactive impurities produced in the detectors and shielding material before putting them underground). The evaluation was carried out with the assumption that only $2\nu\beta\beta$ decay and a bremsstrahlung continuum were involved. The result was

$$T_{1/2}^{2\nu}(^{76}\mathrm{Ge}) = (9.2^{+0.7}_{-0.4}) \times 10^{20} \text{ years} \quad (95\% \text{ c.l.}). \tag{6.149}$$

Finally, the most significant result came from the measurements of the Heidelberg–Moscow collaboration [Bal94] using the first 2 kg of their total of 11.5 kg detector material enriched with ^{76}Ge up to a level of 86% (see below)

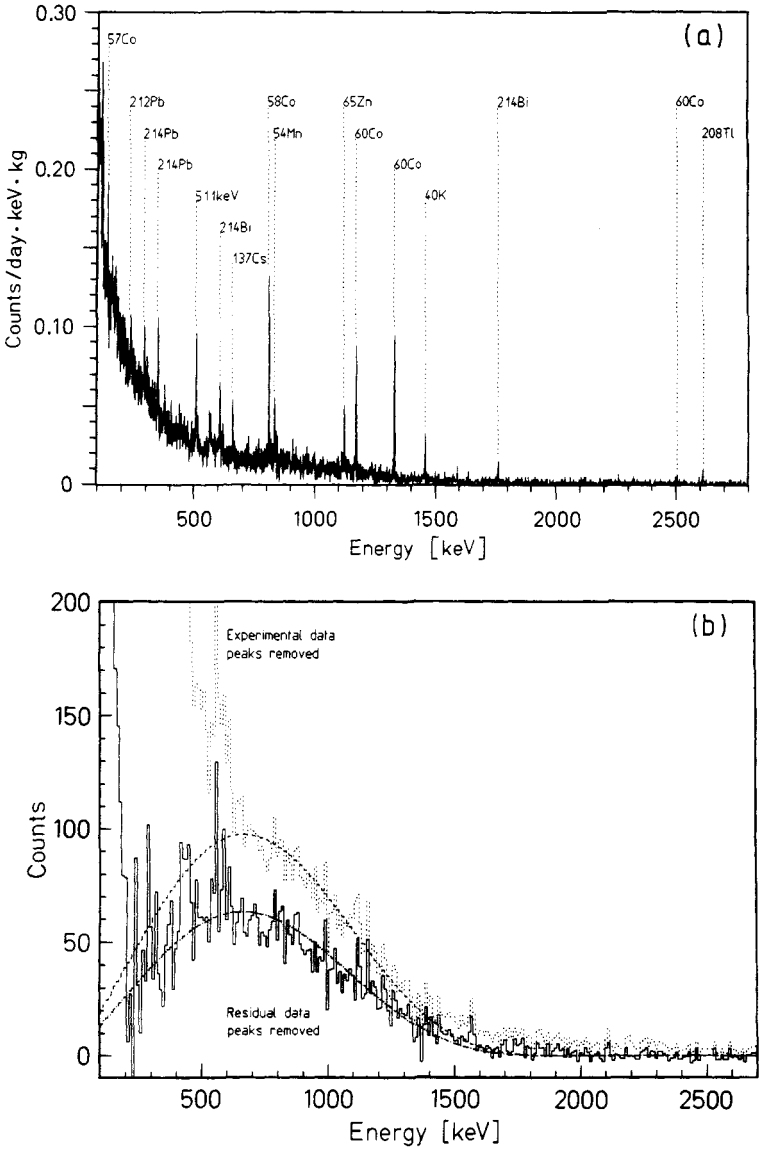


Figure 6.31. (*a*) Measured spectrum from the Heidelberg–Moscow $\beta\beta$ experiment after 1.68 kg years; (*b*) The $2\nu\beta\beta$ spectrum (continuous histogram) obtained from (*a*) after removal of the background. The dotted curves represent calculated $2\nu\beta\beta$ spectra with half-lives $T^{2\nu}_{1/2} = 0.92 \times 10^{21}$ and $T^{2\nu}_{1/2} = 1.42 \times 10^{21}$ years (from [Bec93, Bal94, Kla93b, 94]).

$$T_{1/2}^{2\nu}(^{76}\mathrm{Ge}) = (1.42 \pm 0.03\ (\mathrm{stat}) \pm 0.13\ (\mathrm{sys})) \times 10^{21}\ \mathrm{years} \quad (90\%\ \mathrm{c.l.}). \tag{6.150}$$

The measured spectrum is shown in figure 6.31. This result could even be the first undoubtable evidence for this nuclear decay mode.

Subsequent to the pioneering work of Fiorini *et al* [Fio67] a large number of experiments have been carried out to search for the $0\nu\beta\beta$ line at 2.04 MeV in the sum spectrum of the germanium detectors (see e.g. [Bel83, Lec83, Sim84, Avi87, Cal87, 90, Eji87, Fis89, Bal94a]. An overview is given in [Moe91a, 93a, 94a, Kla92a, 93b, 94]). The crucial problem is the suppression of background events, since the expected line still lies in the energy range of the natural radioactivity. Thus, in addition to the installation in underground laboratories, massive passive screens, for example, of copper and lead, and sometimes active screens of surrounding veto counters (often NaI counters) were used. Moreover, an extremely high requirement must be placed on the purity of the material used to construct the detector as far as radioactive impurities are concerned.

One important background source in experiments with natural germanium is the electron capture decay of ^{68}Ge ($Q_{EC} = 2.9$ MeV, $T_{1/2} = 270.8$ d). ^{68}Ge is formed cosmogenically as an activation product of ^{70}Ge (natural isotopic abundance 20.5%) via the reaction

$$^{70}\mathrm{Ge}(n, 3n)\,^{68}\mathrm{Ge}. \tag{6.151}$$

This means that part of the background comes from the detector crystal itself. This component may be reduced by decreasing the ^{70}Ge content (automatically associated with the enrichment of ^{76}Ge).

Decisive progress in research into the 0ν mode is expected from a new generation of experiments with isotopically enriched germanium or other enriched active sources. A strong enrichment of the detector material with ^{76}Ge nuclei increases the source strength without simultaneously increasing the sensitivity to background radiation. The latter would be an inevitable consequence of the use of a correspondingly large quantity of natural germanium. We shall discuss the importance of enrichment for the sensitivity of the experiments using the following expression which is approximately valid for germanium:

$$T_{1/2}^{0\nu}(^{76}\mathrm{Ge}) > (4.18 \times 10^{24}\ \mathrm{kg}^{-1}) a \sqrt{\frac{Mt}{B\Delta E}}. \tag{6.152}$$

Here, $T_{1/2}^{0\nu}$ is the half-life which can be extracted from a measurement after measurement time t (in years), if no line is visible. Further, a denotes the isotopic abundance of ^{76}Ge, M is the active mass of the detector in kg, B is the average background in the area of the expected line measured in events per keV,

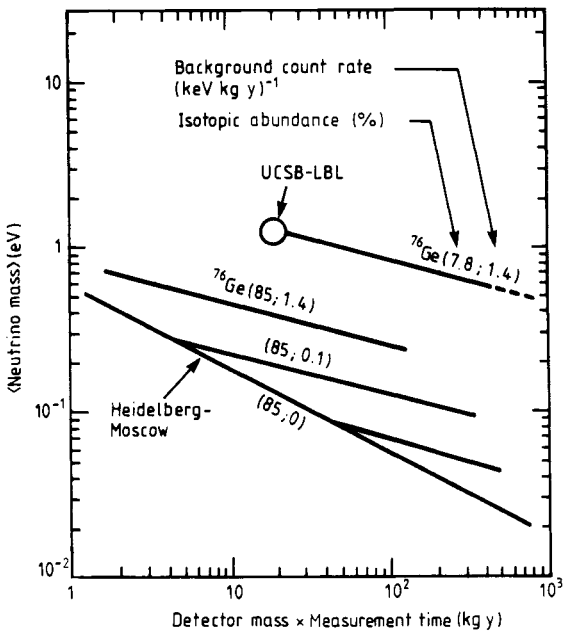


Figure 6.32. Regions for an effective neutrino mass $\langle m_{\nu_e} \rangle$ that can be investigated with Ge detectors with various levels of ^{76}Ge enrichment and of background reduction, as a function of the product of the detector mass times the measurement time (from [Moe91b, Kla91, 94]).

per year and per kilogram of detector and ΔE denotes the energy resolution in keV.

From (6.152), we see that the degree of enrichment a provides the most effective increase in the sensitivity, since it is the only parameter which does not appear under the square root. An experiment with about 10 kg of ^{76}Ge enriched to 86% (as used in the Heidelberg–Moscow experiment) corresponds in sensitivity to a natural Ge experiment with at least 1.2 t (in the most sensitive experiments with natural Ge to date around 8 kg of material were used). Over a measurement period of 5 years, the neutrino mass may be measured down to 0.2 eV (figures 6.32 and 6.40). As previously mentioned, the reduction of the background because of the tiny $\beta\beta$ decay rate is the greatest experimental requirement. The lowest background rates achieved in experiments of this type lie in the region $B \sim 0.1$–0.2 events/keV year kg [Mai94, Kla94].

The most informative germanium experiment to detect the $0\nu\beta\beta$ decay is currently being carried out by a Heidelberg–Moscow collaboration in the Gran Sasso underground laboratory near Rome (figures 6.33–6.35) [Kla87, 89, 94,

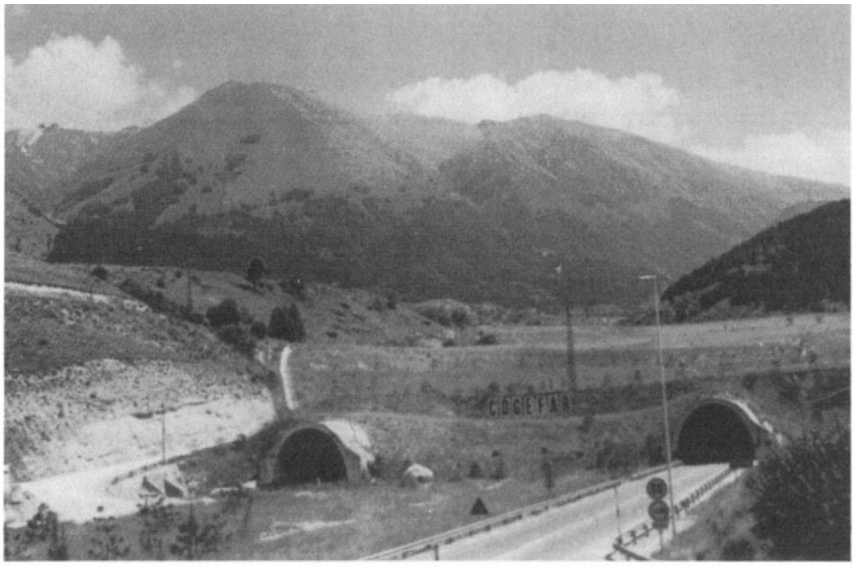


Figure 6.33. On the Heidelberg–Moscow double-beta experiment in the Gran Sasso underground laboratory near Rome. Entrance to the 11.4 km long road tunnel (photo: W Filser).

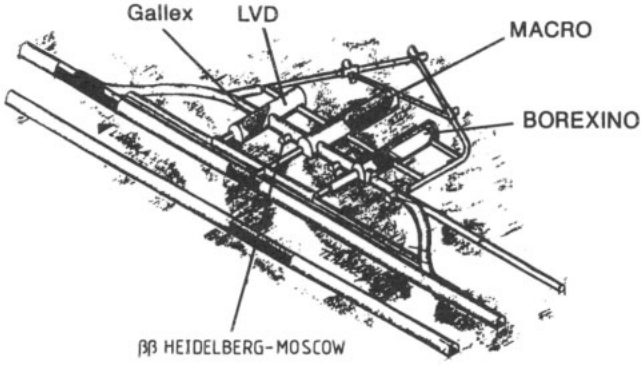


Figure 6.34. From the left tube of the road tunnel one branches into the three large halls of the Gran Sasso laboratory. The Heidelberg–Moscow experiment is located between Halls A and B, which house the GALLEX (see chapter 7) and LVD (see chapter 4) and MACRO (see chapter 8) experiments, respectively. Borexino is under development in Hall C (see chapter 7).

Bec93, 94, Bal93, 94, 94a, Mai94, Pet94]. A total of 19.2 kg of enriched germanium with a ^{76}Ge content of 86% is available.

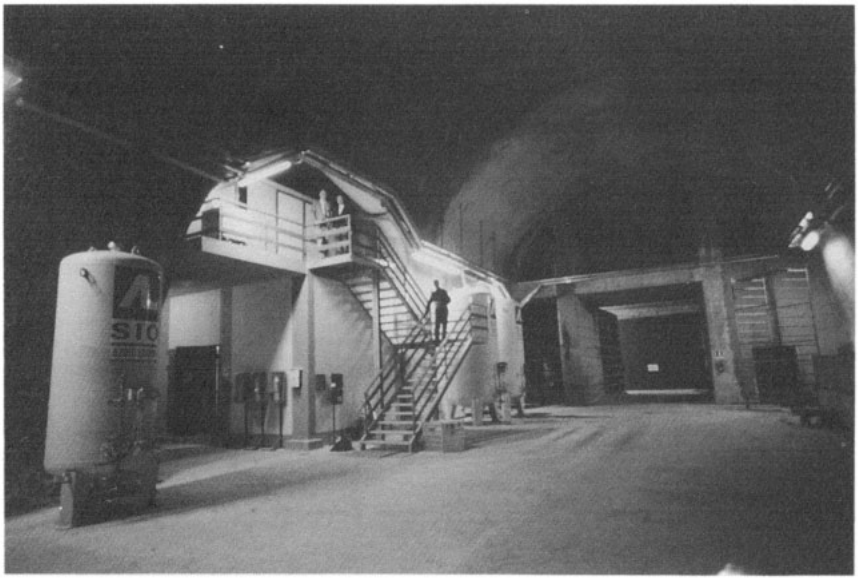


Figure 6.35. The experimental building of the Heidelberg–Moscow $\beta\beta$ experiment constructed in the Gran Sasso underground laboratory by the Italian National Institute for Nuclear Physics (INFN). The lower stories house the detectors, the upper stories the computers and the data processing. A liquid nitrogen tank can be seen in the foreground. The large door in the background leads to Hall A, where LVD and GALLEX are located (see chapters 4 and 7) (from [Kla91f]).

From July 1990 till February 1995 five high-purity enriched ^{76}Ge detectors with a total mass of around 11.5 kg have been installed in the Gran Sasso laboratory (figure 6.36) and measurements have been carried out on them. For screening, the detectors are surrounded by a total of 15 tonnes of extremely clean lead and electrolytic copper.

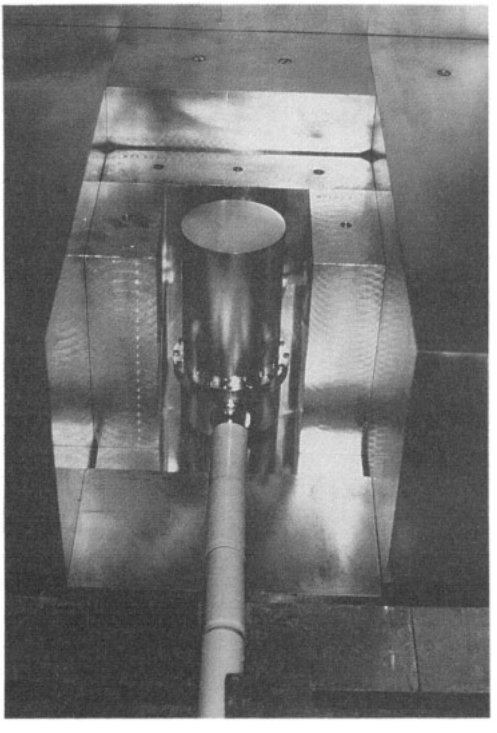
Figure 6.37 shows the total energy spectrum recorded after a measurement time of 10.23 kg years. A line has not yet been detected in the area of the $0\nu\beta\beta$ decay. The background in the area of the $0\nu\beta\beta$ line at 2038.6 keV corresponds to approximately 0.2 events.

A lower bound for the half-life of

$$T^{0\nu}_{1/2}(^{76}\mathrm{Ge}) > 5.6 \times 10^{24} \text{ years} \quad (90\% \text{ c.l.}) \tag{6.153}$$

is obtained [Bal95][10]. This means that the sensitivity has been increased now far beyond that obtained by the UCSB–LBL group [Cal90] from measurements

[10] The actual value at October 1997 is, after 31.8 kg years, $T^{0\nu}_{1/2}(^{76}\mathrm{Ge}) \geq 1.2 \times 10^{25}$ years [Kla98].



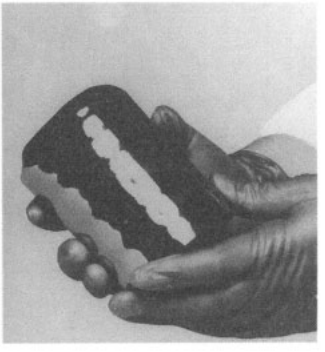


Figure 6.36. Top and lower left: the first five enriched 'high-purity' ^{76}Ge detectors in the world, in the Heidelberg–Moscow $\beta\beta$ setup (shielded by low-level lead and electrolytic copper) in the Gran Sasso (photos: B Maier). Lower right: an enriched 3.5 kg ^{76}Ge crystal before conversion into a 2.9 kg detector (photo: MPI Heidelberg).

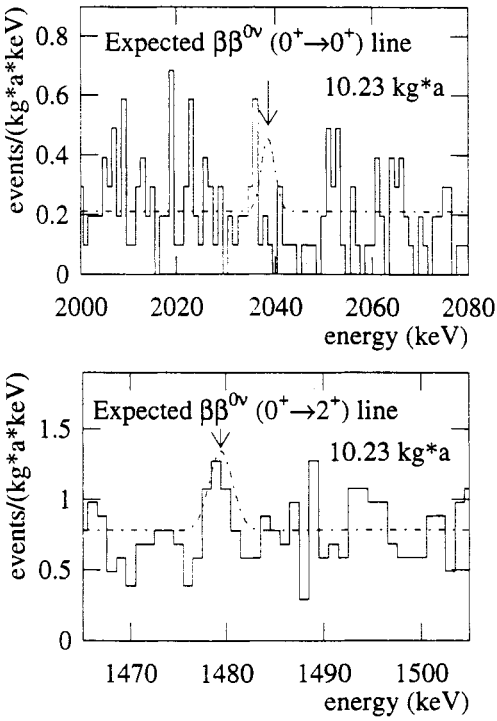


Figure 6.37. Spectrum of the Heidelberg–Moscow $\beta\beta$ experiment after a measurement time of 10.23 kg years in the region of the transition energies to the ground state and the first excited 2^+ state in ^{76}Se. The continuous lines are excluded with 90% confidence (from [Bal95, Hel95]).

with around 8 kg of natural germanium. The experiment also gives one of the sharpest bounds to date for the half-life of majoron-accompanied $0\nu\beta\beta$ decay modes together with the sharpest laboratory bound for the electron decay. The following limit was obtained after 615 kg days [Bec93, Kla94]

$$T_{1/2}^{0\nu\chi}(^{76}\mathrm{Ge}) > 1.66 \times 10^{22} \text{ years} \quad (90\% \text{ c.l.}) \tag{6.154}$$

(figure 6.38) and, based on this, one of the sharpest bounds for the majoron–neutrino coupling constant (see sections 6.2.4.3 and 6.2.5.4). For the process

$$e^- \rightarrow \gamma + \nu_e \tag{6.155a}$$

it was found that [Bal93]

$$\tau_e > 1.63 \times 10^{25} \text{ years} \quad (68\% \text{ c.l.}). \tag{6.155b}$$

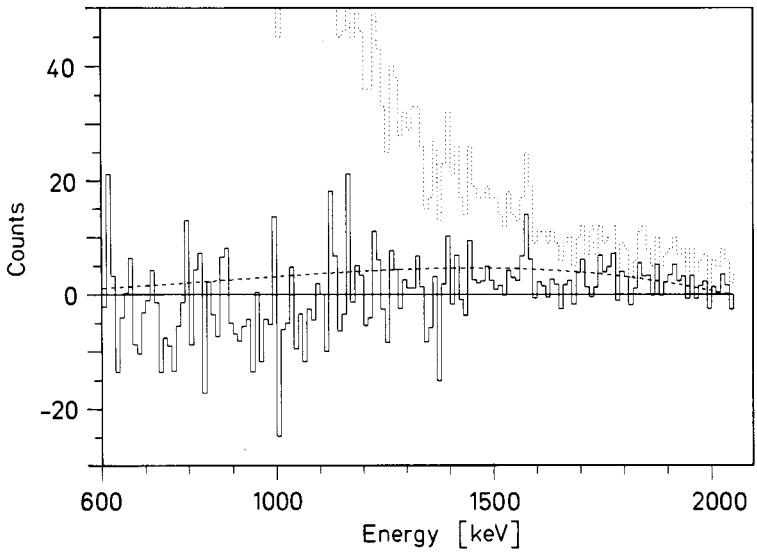


Figure 6.38. Residual spectrum obtained from the continuous histogram of figure 6.31(*b*) by subtracting the 2ν spectrum with $T_{1/2}^{2\nu} = 1.42 \times 10^{21}$ years, and calculated $0\nu\chi\beta\beta$ spectrum with $T_{1/2}^{2\nu} = 1.66 \times 10^{22}$ years (from [Bec93]). The broken histogram shows the uncorrected spectrum (only γ lines removed).

In the main phase of the experiment, which, with 11.5 kg of enriched detector material is intended to last at least until about the year 2000, the $0\nu\beta\beta$ mode is planned to be studied up to and beyond the region of 10^{25} years.

The $0\nu\beta\beta$ decay of ^{76}Ge to the first excited state in selenium at 559 keV (together with the decay to other excited states [Bec92a]) has also been studied. The measured lower bound for the half-life is [Hel95]

$$T_{1/2}^{0\nu}(^{76}\text{Ge}; 2^+) > 9.6 \times 10^{23} \text{ years} \quad (90\% \text{ c.l.}). \tag{6.156}$$

This rules out the result $T_{1/2}^{0\nu}(^{76}\text{Ge}; 2^+) = 1 \times 10^{22}$ years claimed by a French group [Bus90] looking for coincidences between a Ge detector and a surrounding counter detecting γ rays from the $2^+ \rightarrow 0^+$ transition in the $\beta\beta$ nucleus.

Semiconductor devices may also be used to study external probes. In this way, $\beta\beta$ emitters may be chosen, which have a larger $Q_{\beta\beta}$ value and thus the advantage of a larger phase space and smaller background. One example of such a setup is a sandwich detector consisting of silicon counters separated by molybdenum sheets (see e.g. [Avi88, Oka88, Als89, 93]). This setup yielded

$$T_{1/2}^{0\nu}(^{100}\text{Mo}) > 4.4 \times 10^{22} \text{ years} \quad (68\% \text{ c.l.}). \tag{6.157}$$

Scintillation counters. The approach proposed in 1960 involving simultaneous use of the source as detector [Ant60] was first implemented in Brookhaven to study the $\beta\beta$ decay of ^{48}Ca [Mat66]. ^{48}Ca benefits from a very large $Q_{\beta\beta}$ value of 4.271 MeV. However, this isotope is very rare (0.187%). This experiment used the fact that calcium in the form of CaF_2 may be used as a scintillation crystal.

The energy resolution of such scintillation detectors is lower than that of semiconductors. In addition, the photomultiplier is often a background source. The original experiment of 1966 had a source strength of 11.4 g ^{48}Ca. New measurements with 37.4 kg CaF_2 scintillation crystals (43 g ^{48}Ca) were carried out in a 512 m deep colliery shaft near Beijing (China). The bound obtained for the 0ν mode [You91] was

$$T_{1/2}^{0\nu}(^{48}\mathrm{Ca}) > 9.5 \times 10^{21} \text{ years} \quad (76\% \text{ c.l.}). \tag{6.158}$$

In an analogous way, ^{116}Cd ($Q_{\beta\beta} = 2.81$ MeV) can be built into scintillator material ($CdWO_4$). Three enriched $^{116}CdWO_4$ detectors (enriched in ^{116}Cd to 83%) are currently being operated in the Solotvina salt mine in the Ukraine. From these measurements it follows that [Dan89, 95, Zde91]

$$T_{1/2}^{0\nu}(^{116}\mathrm{Cd}) > 2.9 \times 10^{22} \text{ years}. \tag{6.159}$$

Cryodetectors. Cryodetectors are operated at very low temperatures and measure the overall energy deposited in the material, so that both ionizing and non-ionizing events are detectable. These calorimeters in principle have a very good energy resolution.

Some of the potential $\beta\beta$ candidates exhibit the phenomenon of superconductivity, but only at very low temperatures. For example, zirconium, molybdenum and cadmium have a transition (or critical) temperature T_c of less than 1 K. For tin, T_c is approximately 3.7 K. A possible detector would consist, for example, of a large number of small granules of one of these substances with a typical diameter in the μm area in an overheated superconducting state. The energy deposited in these granules as a result of the decay is sufficient to induce a transition from the superconducting state to a state of normal conductivity. This phase transition can be detected via the Josephson effect or via the change in the magnetic flux [Pre90, 93].

Other thermal low-temperature detectors, so-called *bolometers*, are also proposed for use in the search for the $\beta\beta$ decay [Fio84, 91]. They involve the use of a pure diamagnetic or dielectric crystal. At low temperatures, according to Debye, the thermal capacity is described by

$$C(T) \sim \left(\frac{T}{\Theta_D}\right)^3 \tag{6.160a}$$

where Θ_D denotes the Debye temperature. For $T \to 0$, the thermal capacity is so small that even the energy released in the double-beta decay causes a measurable increase in temperature

$$\Delta T \sim \frac{E}{C(T)} \qquad (6.160b)$$

where E denotes the energy deposited in the crystal. The first small bolometers have already been successfully tested. Four 334 g heavy TeO_2 crystals at 10 mK are currently being operated as bolometers in the Gran Sasso laboratory in the search for the decay of ^{130}Te [Fio91b, Giu91, Ale94].

6.2.5.4 *Bounds for the neutrino mass, the right-handed coupling constants, effective neutrino masses, majoron coupling constants and* SUSY *parameters*

Neutrino mass and right-handed currents. The general relationship between the rate of the neutrinoless $\beta\beta$ decay and the $(B - L)$ violating parameters $\langle m_\nu \rangle$, $\langle \eta \rangle$ and $\langle \lambda \rangle$ is given by (6.117*b*). To date the 0ν mode has not been observed experimentally and there exist only upper bounds for the decay rates $\omega_{0\nu}$. From these it is easy to derive upper bounds for the mass and the right-handed parameters, by setting the two other parameters in (6.117*b*) to be zero (projection onto one axis; this yields the values of $\langle m_\nu \rangle^*$, $\langle \eta \rangle^*$ and $\langle \lambda \rangle^*$ given in table 6.11).

However, for a more accurate analysis one must take into account the fact that the theoretical transition probability is a quadratic form in the parameter space of the neutrino mass and the two right-handed coupling constants. The inequality $\omega_{0\nu} < \omega_{\text{obs}}^{\text{lim}}$ spans an ellipsoid of the permissible neutrino masses and right-handed coupling constants in this three-dimensional space. The largest possible $\langle m_\nu \rangle$ occur, for example, for non-zero values of $\langle \lambda \rangle$ and $\langle \eta \rangle$ (see figure 6.39).

Tables 6.8 and 6.11 and figure 6.40 summarize the bounds which have been obtained with the matrix elements of [Mut89b, Sta90a]. The tightest bounds to date for the effective mass follow from the measurements of the $\beta\beta$ decay of ^{76}Ge and ^{128}Te. However, experiments with germanium detectors of enriched germanium may soon improve on these (see figure 6.32). The resulting bounds for the right-handed currents are several orders of magnitude smaller than those from measurements of angular correlations in the single-beta decay, from the μ decay or the K decay (compare [Doi85]). It should be kept in mind that $\beta\beta$ decay gives limits only on the *effective* values $\langle m_\nu \rangle$, $\langle \lambda \rangle$, $\langle \eta \rangle$ (see, however section 6.2.4.2 and [Suh93]). We recall further that the values given in tables 6.8 and 6.11 are only valid under the assumption that neutrinos are Majorana particles (see sections 1.6 and 6.2.3). We note again that the

Table 6.11. Upper bounds for the effective neutrino mass and the right-hand coupling constants, extracted from the given experimental bounds on the $0\nu\beta\beta$ half-life, using the matrix elements of [Sta90a]. Projection onto the axis: $\langle m_\nu\rangle^*$, $\langle\eta\rangle^*$, $\langle\lambda\rangle^*$. Evaluation of the quadratic form: $\langle m_\nu\rangle$, $\langle\eta\rangle$, $\langle\lambda\rangle$.

	^{76}Ge	^{82}Se	^{96}Zr	^{100}Mo
$\lvert\langle m_\nu\rangle^*\rvert$ [eV/c^2]	0.7	7.4	195	5.4
$\lvert\langle\eta\rangle^*\rvert$ [10^{-8}]	0.6	7.7	220	2.5
$\lvert\langle\lambda\rangle^*\rvert$ [10^{-6}]	1.1	9.5	210	4.7
$\lvert\langle m_\nu\rangle\rvert$ [eV/c^2]	0.8	8.1	213	5.9
$\lvert\langle\eta\rangle\rvert$ [10^{-8}]	0.7	8.4	230	2.8
$\lvert\langle\lambda\rangle\rvert$ [10^{-6}]	1.2	9.6	220	4.7
$T^{0\nu}_{1/2}$ [years]	$> 5.6 \times 10^{24}$	$> 1.1 \times 10^{22}$	$> 1.4 \times 10^{19}$	$> 4.4 \times 10^{22}$
Reference	[Bal95, Kla94]	[Ell86, 92]	[Zde81]	[Als93]
	^{114}Cd	^{116}Cd	^{128}Te	^{130}Te
$\lvert\langle m_\nu\rangle^*\rvert$ [eV/c^2]	2252	4.1	1.1	5.3
$\lvert\langle\eta\rangle^*\rvert$ [10^{-8}]	2900	5.4	0.9	5.0
$\lvert\langle\lambda\rangle^*\rvert$ [10^{-6}]	20 000	5.1	4.2	7.3
$\lvert\langle m_\nu\rangle\rvert$ [eV/c^2]	3818	4.6	1.4	5.8
$\lvert\langle\eta\rangle\rvert$ [10^{-8}]	4800	6.0	1.3	5.5
$\lvert\langle\lambda\rangle\rvert$ [10^{-6}]	20 000	5.2	4.3	7.4
$T^{0\nu}_{1/2}$ [years]	$> 1 \times 10^{19}$	$> 3 \times 10^{22}$	$> 7.7 \times 10^{24}$	$> 1.8 \times 10^{22}$
Reference	[Bar89e, Dan92]	[Dan95]	[Ber92a, 93a]	[Ale94]
	^{134}Xe	^{136}Xe	^{150}Nd	^{238}U
$\lvert\langle m_\nu\rangle^*\rvert$ [eV/c^2]	454	2.6	4.1	11.4
$\lvert\langle\eta\rangle^*\rvert$ [10^{-8}]	370	1.9	4.2	16
$\lvert\langle\lambda\rangle^*\rvert$ [10^{-6}]	2300	3.8	4.3	44
$\lvert\langle m_\nu\rangle\rvert$ [eV/c^2]	656	2.8	4.4	15.1
$\lvert\langle\eta\rangle\rvert$ [10^{-8}]	530	2.1	4.4	21
$\lvert\langle\lambda\rangle\rvert$ [10^{-6}]	2300	3.8	4.4	44
$T^{0\nu}_{1/2}$ [years]	$> 8.2 \times 10^{19}$	$> 3.4 \times 10^{23}$ [a]	$> 2.1 \times 10^{21}$	$> 2 \times 10^{21}$
Reference	[Bar89c]	[Vui93]	[Kli86, Moe95]	[Tur91]

[a] [Vui93] gives two half-lives. Using a Monte Carlo simulation they distinguish between the mass mechanism $T^{0\nu}_{1/2} > 3.4 \times 10^{23}$ years and the contribution due to right-handed currents $T^{0\nu}_{1/2} > 2.6 \times 10^{23}$ years.

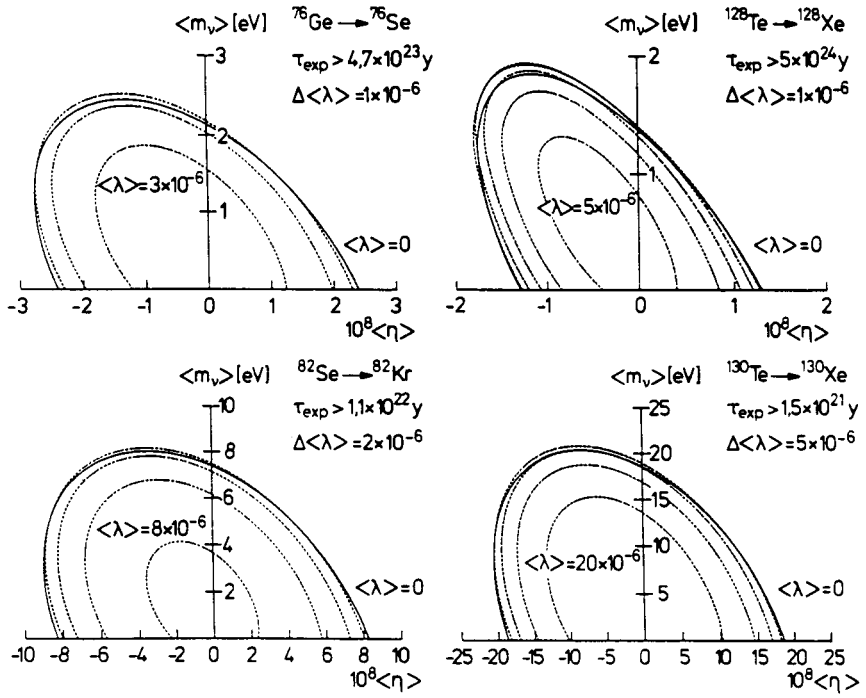


Figure 6.39. Allowed areas (inside the ellipses) for the parameters $\langle m_\nu \rangle$ and $\langle \eta \rangle$ for some fixed values of $\langle \lambda \rangle$ for a number of $\beta\beta$ emitters (from [Tom87]).

extraction of the bounds from the $0\nu\beta\beta$ decay was made possible by decisive progress in the area of nuclear structure physics, which permits the determination of the nuclear matrix elements with sufficient reliability.

Effective neutrino mass. In what follows we shall briefly discuss the importance of the *effective* neutrino mass determined from $0\nu\beta\beta$ decay (see sections 6.2.3, 6.2.6 and 6.2.4.2)

$$\langle m_\nu \rangle = |\sum_j m_j U_{ej}^2|. \tag{6.161}$$

This effective mass is not directly comparable with the neutrino mass determined from the single β decay of tritium or from propagation time measurements of the neutrino pulse from the supernova SN1987A. The different experiments measure different quantities.

If the electron neutrino is a Dirac particle the mass bounds derived from $\beta\beta$ decay are invalid, since in principle only Majorana neutrinos can induce the $0\nu\beta\beta$ decay. Majorana neutrinos may have masses m_j which are greater than the given bounds for the effective mass. In contrast to the case of neutrino oscillations

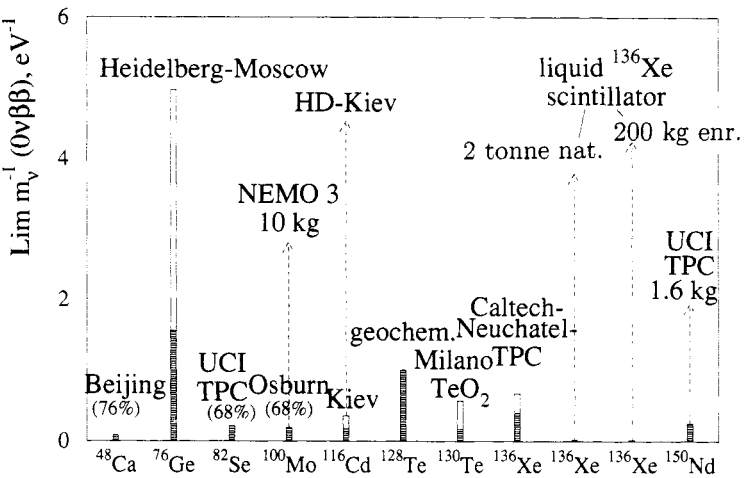


Figure 6.40. Bounds for an effective neutrino mass from $0\nu\beta\beta$ experiments (no geochemical experiments). Present status and future perspectives (after [Kla94]). Shaded stripes correspond to the status in 1994, open stripes and broken lines refer to 'safe' and less 'safe' extrapolations to the year 2000 and beyond, respectively. The mass limits shown correspond to the half-life limits shown in figure 6.29.

(see chapter 7) definition (6.161) involves not only the squares of the absolute values of the mixing parameters U_{ej}, but also the phase. Thus, the possibility of a destructive interference exists, so that $\langle m_\nu \rangle < m_j$ for all mass eigenstates ν_j is allowed [Wol81]. The $0\nu\beta\beta$ amplitude may become very small even when the 'true' masses of the neutrinos (the mass eigenvalues) are comparatively large. However, the possible mixing angles and mass differences $\Delta m^2 = (m_j^2 - m_k^2)$ are very restricted by oscillation experiments (see also section 6.2.6). For the dominant mass eigenstate ν_1 of the electron neutrino it follows that a mass substantially greater than $m_1 \simeq 1$ eV/c^2, while possible, cannot be considered to be very probable (see e.g. [Gro86c, 89, 90]).

Heavy and superheavy neutrinos. The observation of the neutrinoless double-beta decay would also lead to another interesting conclusion. In general, it is assumed that the $0\nu\beta\beta$ decay involves the emission of two W bosons. These exchange a neutrino mass eigenstate ν_j and thus create the two emitted electrons (figure 6.41). The W bosons in general could couple to the leptons via left- or right-handed currents. Under the additional assumption that the effective strength of a right-handed interaction G_r is not greater than G_F, many gauge theories give a *lower* bound for the mass of the *heaviest* neutrino [Kay89]. If the masses of the neutrinos exchanged are small in comparison with the typical momenta ($m_j < 10$ MeV/c^2), the effective mass is defined by (6.161). U_{ej}^2

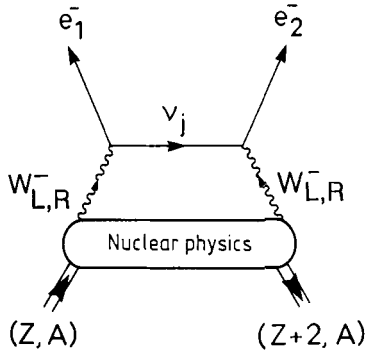


Figure 6.41. Neutrino-exchange diagram for $0\nu\beta\beta$ decay admitting the existence of right-handed W bosons. The indices a, b denote the handedness.

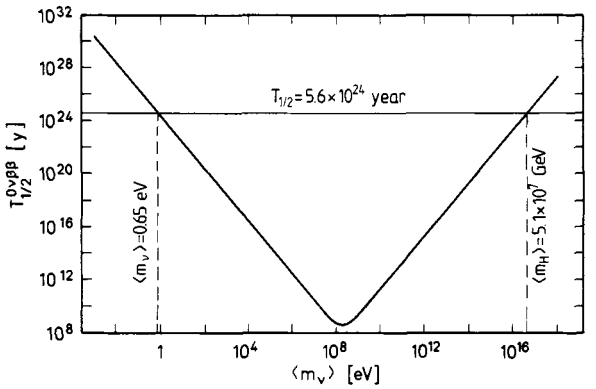


Figure 6.42. Calculated half-life of $0\nu\beta\beta$ decay of ^{76}Ge as function of neutrino mass and the resulting limits on the light and heavy neutrino mass from the Heidelberg–Moscow $\beta\beta$ experiment (from [Bal95]).

may be replaced by $|L_{ej}|^2$ and a phase factor ξ_j. It follows immediately from unitarity that $m_{\mathrm{H}} \geq \langle m_\nu \rangle$ where m_{H} is the biggest mass of the $\{m_j\}$. Also when right-handed currents are considered, the observation of $0\nu\beta\beta$ decay implies a lower bound for the mass of the heaviest neutrino. For ^{76}Ge, we have [Kay89]

$$m_{\mathrm{H}} > 1\ \mathrm{eV}/c^2 \left(\frac{10^{24}\ \mathrm{years}}{T^{0\nu}_{1/2}} \right)^{1/2}. \tag{6.162}$$

When including the dependence of the nuclear matrix elements in (6.117*b*) on the mass of the neutrino, according to [Mut89b] and section 6.2.4.2, we can deduce a lower limit on the mass of a superheavy neutrino $\langle m_{\mathrm{SH}} \rangle$ (see

section 1.6.3). The situation for ^{76}Ge is shown in figure 6.42. From the half-life limit of the Heidelberg–Moscow experiment a limit of

$$\langle m_{SH} \rangle > 5.1 \times 10^7 \text{ GeV} \tag{6.163}$$

is obtained [Bal95].

In conclusion, we note that the non-observation of neutrinoless double-beta decay at a certain level of measurement sensitivity does not give an upper bound for the neutrino mass itself but only for the effective neutrino mass; however, the observation of $0\nu\beta\beta$ decay defines a lower bound for the mass of the heaviest mass eigenstate. However, we also note that in most GUT models the effective neutrino mass agrees with the actual neutrino mass [Lan88] (see also table 1.8; for counterexamples, see e.g. [Cha82]).

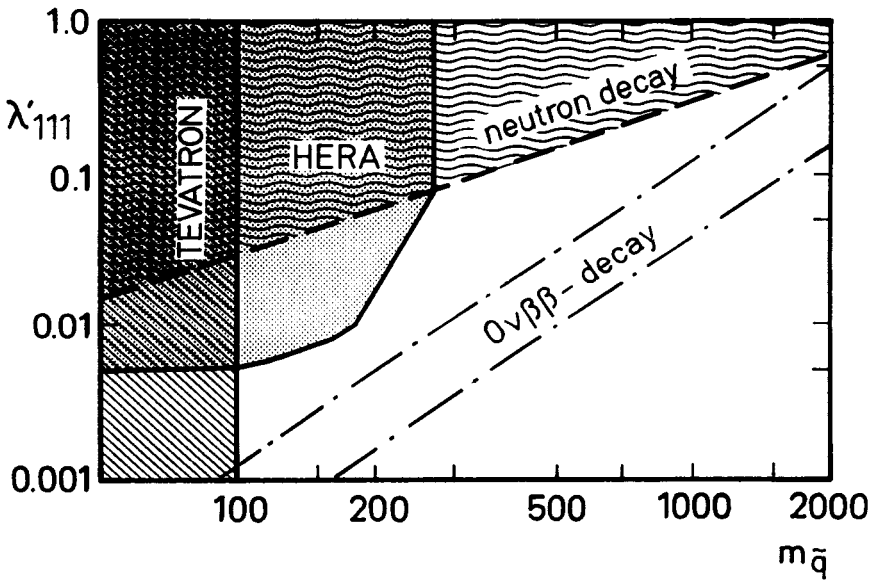


Figure 6.43. Limits on the strength of the R-parity violating interaction λ'_{111} from low-energy experiments and high-energy colliders. λ'_{111} is shown as a function of squark mass. Areas to the left of the curves are excluded. The vertical line is the lower limit for squark masses from Tevatron [Roi92], the bold line corresponds to the limit which might be accessible in future HERA experiments [But93]. The dashed line is the present best limit from low-energy experiments [Bar89], while the dashed-dotted lines correspond to the limits from $0\nu\beta\beta$ decay (the Heidelberg–Moscow experiment) – the upper and lower lines for a gluino mass of 1 TeV and 100 GeV, respectively (from [Hir95]).

Limits for SUSY model parameters. Double-beta decay induced by exchange of supersymmetric particles (e.g. figure 6.20) has the same experimental signature

as $0\nu\beta\beta$ decay induced by neutrino exchange. Therefore a $0\nu\beta\beta$ half-life limit allows restrictions of SUSY parameters (see (6.128)).

If we consider, for example, gluino exchange, one can show that the half-life limit of the Heidelberg–Moscow experiment yields the following limit on the strength λ'_{111} of an R_p-parity breaking interaction [Hir95]

$$\lambda'_{111} \leq 3.9 \times 10^{-4} \left(\frac{m_{\tilde{q}}}{100\,\mathrm{GeV}}\right)^2 \left(\frac{m_{\tilde{g}}}{100\,\mathrm{GeV}}\right)^{1/2}. \tag{6.164}$$

Figure 6.43 shows this result for two values of the gluino mass, $m_{\tilde{g}} = 100$ GeV and 1 TeV (the latter being some natural upper limit in SUSY models). For comparison also shown are the sharpest limits from low-energy experiments and high-energy colliders: from neutron decay [Bar89], from Tevatron dilepton data [Roi92], and from *future* HERA experiments [Dre94, But93, Ahm94]. For the calculation of the latter it must be assumed that the scalar quark can decay into a lighter supersymmetric particle (photino), i.e. that the photino mass is smaller than the squark mass. Neutrinoless double-beta decay thus yields the sharpest limits on R-parity violating SUSY models.

Table 6.12. Half-life bounds for the $0\nu\chi$ decay and the corresponding bounds for the neutrino–majoron coupling constant for various isotopes.

Isotope	Experiment	$T_{1/2}$ (10^{21} years)	$10^4\langle g_{\nu\chi}\rangle$	Ref.
^{76}Ge	MPIK–KIAE	16.6 (90%)	1.8	[Bec93]
^{76}Ge	ITEP	10 (68%)	2.2	[Vas90]
^{76}Ge	UCSB–LBL	1.4 (90%)	5.8	[Cal87]
^{76}Ge	PNL–USC	6.0	2.8	[Mil90]
^{76}Ge	Cal. PSI–New.	1.0 (90%)	6.9	[Fis89]
^{100}Mo	LBL–MHC–UNM	0.33 (90%)	6.2	[Als88, 93]
^{136}Xe	ITEP	0.19 (68%)	12.5	[Bar90c]
^{136}Xe	Cal. PSI–New.	4.9 (90%)	2.4	[Vui93]
^{82}Se	UCI	1.6 (68%)	2.0	[Ell87]
^{150}Nd	INR	0.17 (90%)	1.2	[Art95]
^{150}Nd	UCI	0.53 (90%)	0.7	[Moe94a]
^{48}Ca	ITEP	0.72	5.1	[Bar89d]
^{128}Te[a]	Wash. Univ.–Tata	7700	0.3	[Ber92a]

[a] Geochemical experiment.

Majoron–neutrino coupling. Finally, table 6.12 summarizes the bounds (applying (6.125)) for the majoron–neutrino coupling $\langle g_{\nu\chi}\rangle$ determined in various $0\nu\beta\beta$ experiments. The Heidelberg–Moscow experiment [Bec93] rules out recently presented evidence for a $0\nu\chi$ decay (e.g. [Moe92, 93c]).

6.2.5.5 Experiments on the $\beta^+\beta^+$ decay

Experimental and theoretical work on double-beta decay plainly focuses on the $\beta^-\beta^-$ decay. However, from the experimental point of view the two-positron emission (6.90*a*) and the electron–positron conversion (6.90*b*) have the advantage of a clear signature as a result of the emitted positrons. For example, if a $\beta^+\beta^+$ event occurs in a thick sample of material, the positrons will be stopped in the sample and annihilated there. The four emitted coincident 511 keV γ quanta would permit an unambiguous detection of a true event. This effect could compensate, at least in part, the kinematic suppression due to the Coulomb barrier and the low Q value,. However, the most recent measurements of $\beta^+\beta^+$ half-lives

$$T_{1/2}^{(2\nu+0\nu)}(^{96}\mathrm{Ru}) > 3.1 \times 10^{16} \text{ years} \quad [\mathrm{Nor85}] \qquad (6.165a)$$

$$T_{1/2}^{(2\nu+0\nu)}(^{106}\mathrm{Cd}) > 2.6 \times 10^{17} \text{ years} \quad [\mathrm{Nor84}] \qquad (6.165b)$$

$$T_{1/2}^{2\nu}(^{124}\mathrm{Xe}) > 2 \times 10^{14} \text{ years} \quad [\mathrm{Bar89c}] \qquad (6.165c)$$

$$T_{1/2}^{0\nu}(^{124}\mathrm{Xe}) > 4.2 \times 10^{17} \text{ years} \quad [\mathrm{Bar89c}] \qquad (6.165d)$$

are far below the theoretical predictions in table 6.6. The bounds for the neutrino mass derived from these (see table 6.13) are much less restrictive than those obtained from investigations of $\beta^-\beta^-$ decay.

Table 6.13. Upper bounds for the effective neutrino mass from $\beta^+\beta^+$-experiments.

Transition	$\langle m_\nu \rangle$ [eV/c^2]	Ref.
$^{96}\mathrm{Ru} \rightarrow {}^{96}\mathrm{Mo}$	$< 1.3 \times 10^6$	[Nar85]
$^{106}\mathrm{Cd} \rightarrow {}^{106}\mathrm{Pd}$	$< 3.5 \times 10^5$	[Nar84]
$^{124}\mathrm{Xe} \rightarrow {}^{124}\mathrm{Te}$	$< 1.3 \times 10^5$	[Bar89c]

It seems improbable that the sensitivity of present-day detectors will permit the detection of two-positron emission [Sta91]. The case of electron–positron conversion is somewhat more favourable, since the Q value is by around 1 MeV greater and there is still a clear signature due to the emitted positron. Expected half-lives for $2\nu\beta^+EC$ decay are partly more than four orders of magnitude lower than those for $2\nu\beta^+\beta^+$ decay (see table 6.6) [Hir94].

A β^+EC experiment could become of particular importance, if a $0\nu\beta^-\beta^-$ decay were to be observed. The additional information from the $0\nu\beta^+EC$ decay would allow us to deduce a *lower* limit of the neutrino mass or the existence of

right-handed weak currents. This point was discovered by Hirsch *et al* [Hir94], to which we refer for details.

6.2.6 The neutrino mixing in the $0\nu\beta\beta$ decay

In the previous section we have discussed the importance of the neutrino mass in double-beta decay. We shall now extend this discussion and place it in a broader context.

A large number of different measurements of the properties of the neutrino have now been taken (see this and the next chapter). The question arises as to the extent to which the results derived from these measurements are mutually compatible and what conclusions and restrictions have now become possible in the framework of existing models. In particular, since the discussion of a massive neutrino with mass 17 keV/c^2 a number of papers have attempted to place the collected experimental data in a common context and to interpret them (see e.g. [Bab91b, Ben91, Bil91, Cal91a, Gla91, Him91b, Kol91, Kra91, Man91a]). A number of interpretations and extensions of the standard model have been discussed. We cannot restate the interesting implications of the various experimental results in detail here; we shall merely point out a number of interconnections.

If we assume for the moment that there exists a mass eigenstate with a mass of 17 keV/c^2 and an admixture of around 1% to the electron neutrino, then the present non-observation of $0\nu\beta\beta$ decay already gives an apparent contradiction. The effective neutrino mass $\langle m_\nu \rangle$ is defined by (6.161)

$$\langle m_\nu \rangle = |\sum_j{}' m_j U_{ej}^2| = |\sum_j{}' m_j |L_{ej}|^2 \xi_j| \tag{6.166}$$

where the phase factor ξ_j takes the values ± 1. This corresponds to the case with CP conservation. For $m_1 < 10$ eV/c^2, $m_2 \simeq 17$ keV/c^2 and $U_{e2}^2 = 1\%$ there follows an effective mass $\langle m_\nu \rangle \simeq 170$ eV/c^2. Here, we have initially neglected other neutrino states with $i > 2$. However, from section 6.2.5.4 (see in particular table 6.11), we know that the experimental upper bound for $\langle m_\nu \rangle$ is already two orders of magnitude smaller.

A way out of this problem would be that a 17 keV neutrino would be a Dirac particle and thus cannot contribute to double-beta decay since (6.166) is only valid for Majorana neutrinos. An alternative interpretation asserts that the 17 keV neutrino would be a Majorana neutrino and at the same time requires that there should exist at least one other heavy Majorana neutrino with an appropriate phase factor ξ_j, so that a destructive interference in (6.166) is responsible for the small effective mass. Here, we note that single-beta decay is itself independent of the nature of the ν_e and does not enable us to draw any conclusions as far as discrimination between the two possibilities is concerned.

Other restrictions are possible by taking oscillation experiments into account. Let us consider the case of a two-neutrino mixing (let ν_2 be the 17 keV neutrino and ν_1 the dominant component of the electron neutrino), so that

$$\Delta m_{12}^2 = (17\ \text{keV}/c^2)^2 \tag{6.167a}$$

and

$$\begin{aligned} \sin^2 2\Theta_{12} &= 4 \sin^2 \Theta_{12} \cos^2 \Theta_{12} \\ &\simeq 4 \sin^2 \Theta_{12} \\ &= 4|U_{e2}|^2 \simeq 0.04. \end{aligned} \tag{6.167b}$$

From measurements of ν_e–ν_μ oscillations it follows for large mass parameters Δm_{12}^2 (see [Ahr85] and table 7.5) that

$$\sin^2 2\Theta_{12} < 3.4 \times 10^{-3}. \tag{6.167c}$$

Thus, a ν_e–ν_μ mixing in tritium decay is ruled out. Therefore, the heavy ν_2 could not be the dominant component of the ν_μ. However, the interpretation as a tau neutrino would be allowed. A more general consideration requires a consideration of three families with $m_1 \approx 0$, $m_2 \simeq 17$ keV/c^2 and $m_3 = ?$ (see e.g. [Cal91a]).

The possibility of harmonizing a hypothetical 17 keV neutrino with experiments on $0\nu\beta\beta$ decay by assuming it to be a Dirac neutrino causes cosmological problems. Stable neutrinos with such a large mass conflict with cosmology, since they would increase the existing mass density in the universe too greatly. Thus, the ν_2 would have to decay again sufficiently rapidly. If the decay products are only subject to the weak interaction cosmology requires a lifetime $\tau < 10^{14}$ s [Kol90]. The decay of heavy neutrinos is discussed in section 6.4. The decay rates given there for conventional Dirac neutrinos are far too small to be in agreement with the above cosmological requirement.

Glashow showed that a simple extension of the majoron model of [Chi81] discussed in section 6.2.4 could lead to the required short lifetime of the ν_2 [Gla91]. In this model the mass matrix involves both Dirac and Majorana terms. The Dirac masses are generated via the coupling to the 'usual' Higgs doublet, while the Majorana contributions come from the coupling which breaks the $(B - L)$ symmetry.

In this approach, the extended neutrino spectrum consists of two light Majorana neutrinos with masses around 10^{-3} eV/c^2 and two very heavy Majorana neutrinos with masses so large that they do not play a role in current experiments. Two other neutrino states are essentially degenerate and form the right- and left-handed components of a 17 keV Dirac neutrino, which could be

interpreted as a tau neutrino. The 17 keV neutrino would decay into one of the light neutrinos (ν_e or ν_μ) and a massless majoron.

However, as a result of the observation of electron neutrinos from the SN1987A the Dirac hypothesis appears very improbable. The results of the GALLEX collaboration on the solar neutrino flux [Ans92, 94] also give no indications of a magnetic moment of the neutrino and thus favour a Majorana neutrino (see chapter 7). A heavy neutrino produced in the supernova would be scattered by nucleons or electrons. As a massive Dirac neutrino, it would have a magnetic moment proportional to the mass, so that the original left-handed neutrino would eventually change into a right-handed state by a spin flip. A right-handed neutrino takes no more part in the weak interaction (*sterile neutrino*) and leaves the interior of the supernova unhindered. On the other hand, as a result of their interaction, left-handed neutrinos are trapped in the supernova for a certain time.

The emitted sterile neutrinos would result in an additional cooling of the supernova. This would decrease the overall energy emitted in the form of electron neutrinos and shorten the length of the detected neutrino pulses [Gan90, Gri90]. The time length of the pulse for the IMB detector should then have been only 1.4 s, and 3.3 s for Kamiokande. In fact, neutrinos were recorded over an interval of from 6 s to 12 s. Since Majorana neutrinos cannot have a magnetic moment, the above problem of a conversion into a sterile neutrino with associated excessive cooling of the supernova does not apply. Thus, this argument favours the Majorana hypothesis (with another heavy neutrino state). To prevent the 'overclosing' of the universe, a neutrino decay as discussed above had again to be assumed. A very detailed discussion of this area is given, for example, in [Cal91a] (see also [Zub93]). The authors conclude that if a 17 keV neutrino was actually a physical reality the current bounds from laboratory measurements and astrophysical and cosmological observations would give the following picture [Cal91]:

(i) The 17 keV neutrino (ν_2) is a Majorana particle and the dominant component of the ν_τ.

(ii) The ν_2 does *not* form the dark matter in the universe.

(iii) The ν_μ is a massive Majorana particle with a mass of 17 keV/c^2 or a mass in the region from 170–270 keV/c^2 with a CP eigenvalue opposite to ν_2, to suppress the neutrinoless double-beta decay.

(iv) The MSW solution (see chapter 7) of the solar neutrino problem requires the conversion (oscillation) of ν_e into a light, sterile neutrino which forms an SU(2) singlet.

Also, while the above discussion may have lost part of its relevance because of the experimental development relating to the 17 keV neutrino, it remains

instructive and gives insights into the flexibility of the theory and into typical methods of reasoning.

6.3 THE SUPERNOVA SN1987A

Another spectacular way of determining neutrino properties involves the observation of neutrinos from supernovae. The measurement of differences in the propagation times of neutrinos with different energies from supernova explosions yields direct information on the neutrino mass. It follows from $E = mc^2$ and $p = mv$ that the relativistic velocity of a particle is given by

$$v = \frac{pc^2}{E}. \tag{6.168}$$

The propagation time of the neutrino between the time t_{em} of its emission and the time t_{obs} of its detection is a function of the neutrino mass m_ν and the neutrino energy E_ν given by

$$\begin{aligned} t_{\mathrm{obs}} - t_{\mathrm{em}} = \frac{l}{v} = \frac{lE_\nu}{p_\nu c^2} &= t_0 \frac{1}{\sqrt{1 - \left(\frac{m_\nu c^2}{E_\nu}\right)^2}} \\ &\simeq t_0 \left(1 + \frac{m_\nu^2 c^4}{2E_\nu^2}\right). \end{aligned} \tag{6.169}$$

Here l denotes the propagation distance, and $t_0 = l/c$ is the corresponding propagation time of light.

If neutrinos with different energies are emitted, for a finite rest mass, differences in the propagation time result. We shall assume in what follows initially that all neutrinos are emitted at the same time and denote the minimum and maximum energy of their energy spectrum by E_{min} and E_{max}, respectively; thus neutrinos are observed in the detector over a time interval

$$\Delta T = t_0 m_\nu^2 c^4 \left(\frac{1}{2E_{\mathrm{min}}^2} - \frac{1}{2E_{\mathrm{max}}^2}\right). \tag{6.170}$$

The neutrino mass can be determined from the time interval ΔT and the energies using (6.170) (see figure 6.44).

If the emission takes place over a time interval Δt, the possible information which can be gleaned about the neutrino mass is weaker. The minimum (maximum) time interval within which the neutrinos may reach the detector is now

$$\Delta T(\Delta t) = t_0 m_\nu^2 c^4 \left(\frac{1}{2E_{\mathrm{min}}^2} - \frac{1}{2E_{\mathrm{max}}^2}\right) \pm \Delta t \tag{6.171}$$

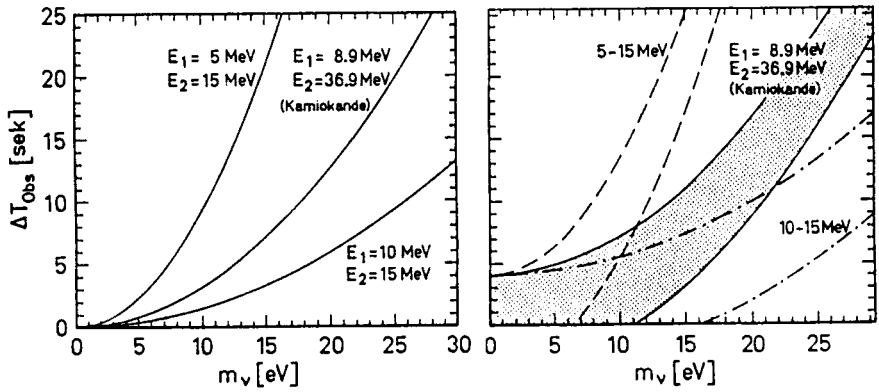


Figure 6.44. On supernova SN1987A: (*a*) neutrino mass as a function of the time interval ΔT between the observation of *simultaneously* emitted neutrinos with different energies E_1 and E_2 on the Earth. Within $\Delta T = 12.439$ s (1.915 s), 11 (8) neutrinos were observed in the Kamiokande experiment with energies between 8.9 and 36.9 MeV; (*b*) neutrino mass as a function of ΔT assuming emission of the neutrinos within a time interval $\Delta t = 4$ s. According to (6.171), for each value of ΔT, we now have a *region* of possible m_ν between two associated curves. For example, for $\Delta T = 1.915$ s, we have $m_\nu < 13.5$ eV (from [Gro89,90]).

where the upper bound corresponds to the situation in which the neutrino with the greatest energy is emitted first and that with the least energy is emitted last (see e.g. [Gro89, 90]).

In a supernova explosion immense numbers of neutrinos are released within a very brief period of time (several seconds). It is assumed, that stars with masses greater than approximately eight solar masses end their life in supernova explosions (of type II) forming neutron stars or black holes.

The gravitational collapse of a massive star occurs at the end of the various hydrostatic burning phases in which successive, increasingly heavier nuclei up to iron are formed. If the mass of the iron core passes the so-called Chandrasekhar limit [Cha39], the pressure of the relativistically degenerate electron gas can no longer resist gravity. The core of the star becomes unstable and a dynamic collapse begins. The total energy released during the gravitational collapse corresponds to the gravitational energy of a neutron star of radius R

$$E \simeq \frac{GM^2}{R} \tag{6.172}$$

and amounts to approximately 3×10^{46} J. Most of this enormous amount of energy is carried away by neutrinos.

The gravitational collapse is triggered by the photodisintegration of nuclei

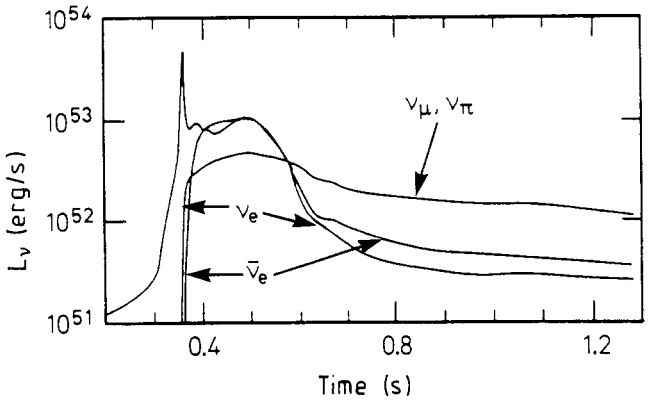


Figure 6.45. Luminosity of a $2M_\odot$ Fe core of a $\approx 25M_\odot$ main sequence star as a function of the time after the start of the collapse for the different neutrino flavours (from [Bru87]).

of the iron group and the onset of electron capture by protons and nuclei, which gives rise to a short neutrino pulse lasting approximately 10 ms. This pulse contains approximately 5% of all emitted neutrinos, and according to (6.1*c*) consists largely of electron neutrinos (ν_e). During the cooling phase of the star core, thermal processes such as pair generation

$$e^+ + e^- \rightarrow \nu_i + \overline{\nu}_i \quad i = e, \mu, \tau \tag{6.173}$$

give rise to neutrinos of all flavours (the ν_μ and ν_τ production only takes place via neutral weak currents (Z^0 exchange)). Roughly speaking, one obtains a statistical mixing of all three neutrino flavours with the corresponding particles and antiparticles. The typical energy is approximately 10 MeV. Figure 6.45 shows the neutrino luminosity of a main sequence star with mass $M = 25M_\odot$. The process (6.173) is of practically no importance during the collapse [Bru85]. Thus, we expect a short, sharp ν_e pulse from the electron capture in the collapsing core, followed by the emission of the main part of the neutrino flux from the pair-generation reaction in the radiating core, which is determined by the diffusion time.

On 23rd February 1987 a supernova explosion at a distance accessible to measurement is known to have been observed (SN1987A). This rare event occurred in the Large Magellanic Cloud some 170 000 light years distant. According to the spectral characteristics this was a type-II supernova. The neutrino signal was detected by various neutrino detectors in Japan (Kamiokande [Hir87, Kosh92]), the USA (IMB [Bio87]), France (Mont Blanc [Agl87]) and the

USSR (Baksan [Ale88]), before the explosion became visible[11]. This was the first time neutrinos created outside our solar system had been unambiguously detected.

Since the interaction cross section for the reaction

$$\nu_e + e^- \rightarrow \nu_e + e^- \tag{6.174}$$

in the energy region around 10 MeV is approximately two orders of magnitude smaller than that for the reaction

$$\overline{\nu}_e + p \rightarrow e^+ + n \tag{6.175}$$

the Cerenkov water counters and scintillation detectors are for the most part sensitive to $\overline{\nu}_e$ so that it is likely that no ν_e from the first short pulse were seen [Bru87, Sat87]. According to (6.173) the $\overline{\nu}_e$ form around one-sixth of the total neutrino flux from a supernova.

Table 6.14 summarizes the neutrino events and their energies recorded by the Cerenkov water counters (Kamiokande II and IMB) and the scintillation detectors (Mont Blanc and Baksan). The fact that the neutrino signal in the Mont Blanc detector was recorded almost five hours before the other events casts doubt on its interpretation as a supernova neutrino signal. The Mont Blanc detector and the Baksan detector each recorded five neutrinos over periods of 7 s and 9.1 s, respectively. IMB detected eight events within 6 s. The greatest number of events, namely eleven with energies above a threshold of 7.5 MeV, was recorded by the Kamiokande II detector within an interval of 12 s. However, it is apparent that the last three of these eleven events are very late since eight neutrinos were seen within 1.915 s.

Based on their angular distribution, the first two Kamiokande events may possibly be attributable to the neutrino–electron scattering rather than to (6.175) [Hir87]. Assuming nine antineutrino events, from the sensitivity of the Cerenkov counter and the interaction cross section for the reaction (6.175) we deduce an overall $\overline{\nu}_e$ flux of 1.0×10^{10} cm^{-2}. For an average neutrino energy of 15 MeV this implies that the energy released by the supernova SN1987A in the form of electron antineutrinos amounted to 8×10^{45} J [Hir87]. For monoenergetic 32 MeV $\overline{\nu}_e$ the IMB data give a $\overline{\nu}_e$ flux of 8×10^8 cm^{-2}; based on this the $\overline{\nu}_e$ luminosity of SN1987A is calculated to be 1×10^{45} J [Bio87].

In the analysis of the data, the crucial question arises as to whether the temporal scattering of the events is caused by the timescale of the neutrino emission by the supernova or by the dispersion resulting from a finite neutrino mass. Unfortunately, the interpretation of the data is fraught with problems

[11] Some of these detectors were used to search for the decay of the proton and have already been described in chapter 4. The origin of the Mont Blanc and the Baksan signal is disputed.

Table 6.14. Table of the neutrino events recorded by the four neutrino detectors Kamiokande II [Hir87], IMB [Bio87], Mont Blanc [Agl87] and Baksan [Ale88]. T denotes the time of the event, E gives the energy of the electron (positron) visible in the detector.

Detector	Event no.	T [UT]	E [MeV]
Kamiokande	1	7 : 35 : 35.000	20 ± 2.9
	2	7 : 35 : 35.107	13.5 ± 3.2
	3	7 : 35 : 35.303	7.5 ± 2.0
	4	7 : 35 : 35.324	9.2 ± 2.7
	5	7 : 35 : 35.507	12.8 ± 2.9
	(6)	7 : 35 : 35.686	6.3 ± 1.7
	7	7 : 35 : 36.541	35.4 ± 8.0
	8	7 : 35 : 36.728	21.0 ± 4.2
	9	7 : 35 : 36.915	19.8 ± 3.2
	10	7 : 35 : 44.219	8.6 ± 2.7
	11	7 : 35 : 45.433	13.0 ± 2.6
	12	7 : 35 : 47.439	8.9 ± 1.9
IMB	1	7 : 35 : 41.37	38 ± 9.5
	2	7 : 35 : 41.79	37 ± 9.3
	3	7 : 35 : 42.02	40 ± 10
	4	7 : 35 : 42.52	35 ± 8.8
	5	7 : 35 : 42.94	29 ± 7.3
	6	7 : 35 : 44.06	37 ± 9.3
	7	7 : 35 : 46.38	20 ± 5.0
	8	7 : 35 : 46.96	24 ± 6.0
Baksan	1	7 : 36 : 11.818	12 ± 2.4
	2	7 : 36 : 12.253	18 ± 3.6
	3	7 : 36 : 13.528	23.3 ± 4.7
	4	7 : 36 : 19.505	17 ± 3.4
	5	7 : 36 : 20.917	20.1 ± 4.0
Mont Blanc	1	2 : 52 : 36.79	7 ± 1.4
	2	2 : 52 : 40.65	8 ± 1.6
	3	2 : 52 : 41.01	11 ± 2.2
	4	2 : 52 : 42.70	7 ± 1.4
	5	2 : 52 : 43.80	9 ± 1.8

resulting from the inaccurate time measurements and the different energy thresholds of the various detectors.

Many authors have published analyses of the data listed in table 6.14 relating to the neutrino mass. These involve more or less well-founded model

assumptions. By way of example, we shall discuss the data from the Kamiokande collaboration in more detail. Eleven events above the energy threshold of 7.5 MeV were detected. The neutrino energy is calculated from the electron energy E_e given in table 6.14 using

$$E_{\bar{\nu}_e} = E_e + m_e c^2 + (m_n - m_p)c^2 = E_e + 1.3 \text{ MeV}. \tag{6.176}$$

Thus, in a time interval of $\Delta t = 12.439$ s eleven events with a minimum energy $E_{\min} = 8.8$ MeV and a maximum energy $E_{\max} = 36.7$ MeV were recorded. Let us initially assume that all neutrinos were emitted at the same time ($\Delta t = 0$), then from (6.170), based on the distance of 170 000 light years ($t_0 \simeq 5.3 \times 10^{12}$ s), the neutrino mass is given by

$$m_\nu = 19.6 \text{ eV}/c^2. \tag{6.177a}$$

Neglecting the last three events, for $\Delta T = 1.915$ s, we obtain a mass of

$$m_\nu = 7.7 \text{ eV}/c^2. \tag{6.177b}$$

The emission intervals Δt for the emission of the neutrinos from a supernova vary according to current models in a range of several seconds; this considerably weakens the information which may be obtained about the neutrino mass. Assuming that $\Delta t = 4$ s, (6.171) yields only an upper bound for the neutrino mass (see figure 6.44)

$$m_\nu < 22.6 \text{ eV}/c^2 \qquad (\Delta T = 12.439 \text{ s}) \tag{6.178a}$$

$$m_\nu < 13.5 \text{ eV}/c^2 \qquad (\Delta T = 1.915 \text{ s}) \tag{6.178b}$$

from the Kamiokande data. Of the many published bounds on the mass derived from observation of the supernova neutrinos, we mention only $m_\nu c^2 < 27$ eV [Arn87] and $m_\nu c^2 < 11$ eV [Bah87]. In a detailed analysis, taking into account model dependences, Kolb *et al* obtained an upper bound

$$m_\nu < 20 \text{ eV}/c^2. \tag{6.179}$$

It seems almost impossible to make a more far-reaching, reliable statement. However, we note that, despite the small number of events recorded, the sensitivity of this propagation-time method to the neutrino mass is very close to that of the tritium experiment.

For further details of the physics of SN1987A we refer readers to [Sch90b], for example.

6.3.1 The neutrino decay

If neutrinos have a finite rest mass and the mass eigenstates are not identical with the interaction states, then the possibility of neutrino decay exists. A

heavy neutrino state which is admixed to the neutrino wavefunction is generated in every neutrino source, provided the kinematic conditions are satisfied. For example, in nuclear β decay, the branching ratio for the emission of a heavy neutrino ν_2 with an energy E_{ν_2} (see (6.49) and (6.61)) is given by

$$B(E_{\nu_2}) = |U_{e2}|^2 \sqrt{1 - \frac{m_2^2 c^4}{(E_{\nu_2} + m_2 c^2)^2}}\,\Theta(E_{\nu_2}) \tag{6.180}$$

where U_{e2} denotes the mixing angle. The Θ function (step function) takes account of the fact that a neutrino state of mass m_2 can only be emitted when sufficient energy is available.

Various decay channels are conceivable, according to the energy conditions. For example, the following transitions are possible (see figures 6.46 and 6.47)

$$\nu_2 \to \nu_1 + l_i^- + l_j^+ \quad l_{i,j} = e, \mu, \ldots \quad \text{for } m_2 > m_1 + m_{l_i} + m_{l_j} \tag{6.181a}$$

$$\nu_2 \to \nu_1 + \nu_1 + \overline{\nu}_1 \quad \text{for } m_2 > 3m_1 \tag{6.181b}$$

$$\nu_2 \to \nu_1 + \gamma \quad \text{for } m_2 > m_1. \tag{6.181c}$$

Let us consider the decay of a heavy neutrino into an electron–positron pair and a light neutrino

$$\nu_2 \to \nu_1 + e^- + e^+. \tag{6.182}$$

For $m_2 \gg 2m_e$ the inverse of the lifetime of the heavy neutrino in its rest system is given by [Shr81]

$$\frac{1}{\tau} = \frac{G_F^2 m_2^5 c^4}{192\pi^3 \hbar^7} |U_{e2}|^2. \tag{6.183}$$

This expression is completely analogous to the expression for the decay rate of the muon for the process $\mu^- \to e^- + \overline{\nu}_e + \nu_\mu$. The Feynman diagrams

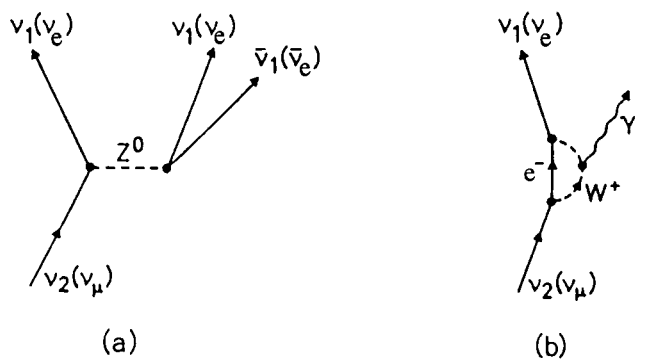


Figure 6.46. Graphs for the neutrino decay: (*a*) $\nu_2 \to \nu_1 + \nu_1 + \overline{\nu_1}$; (*b*) $\nu_2 \to \nu_1 + \gamma$.

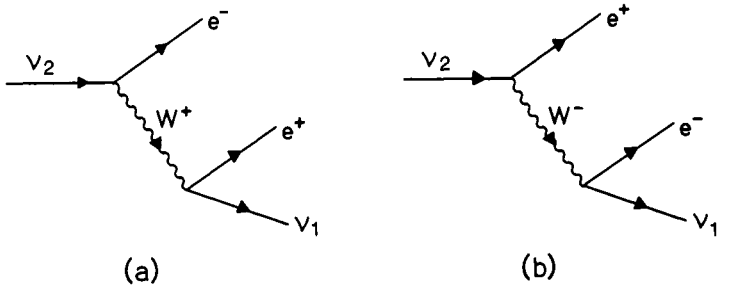


Figure 6.47. Graphs for the neutrino decay $\nu_2 \rightarrow \nu_1 + e^- + e^+$. Diagram (*a*) is valid for Dirac neutrinos. For Majorana neutrinos, the graphs (*a*) and (*b*) should be added (after [Boe92]).

which describe the decay (6.182) are given in figure 6.47. If the heavy neutrino is a Dirac particle, only the diagram shown in figure 6.47(*a*) must be taken into account, while in the case of a Majorana particle the two graphs must be added. Since the neutrino ν_1 in the process (*a*) is left-handed, in the process (*b*), however, right-handed, the terms do not interfere so that the decay rates for Dirac and Majorana neutrinos are identical (see also [Boe87, 92]). However, the angular distribution of the electrons could be used to distinguish between Dirac and Majorana neutrinos [Li82].

According to (6.58) and (6.77*b*) the masses of the dominant mass eigenstates of the ν_e and the ν_μ are already below the threshold for the decay into an e^+e^- pair. While it appears difficult to reduce the upper bounds on the mass of the ν_τ further using direct kinematic measurements, the search for a decay of the ν_τ provides a unique possibility of testing its mass in the region far below 31 MeV/c^2.

The decay of neutrinos into charged leptons has been studied at high-energy accelerators by various groups (see e.g. [Ber86] and the references given in [Fei88a]). Figure 6.48 lists the results of these measurements for the mixing parameters $|U_{e2}|^2$ and $|U_{\mu 2}|^2$ as a function of the mass of the heavy neutrino component.

It is known (see chapter 7) that a reactor represents an intense antineutrino source with $\overline{\nu}_e$ energies up to around 8 MeV. Thus, reactors are suitable for experiments on the neutrino decay into e^+e^- in the energy region[12]

$$1 \text{ MeV} \leq m_2 c^2 \leq 8 \text{ MeV}. \tag{6.184}$$

The formal theory for the decay of reactor neutrinos is described in detail in [Vog84]. An experiment was carried out at the Swiss nuclear power reactor in

[12] The threshold for the decay (6.182) is about 1 MeV.

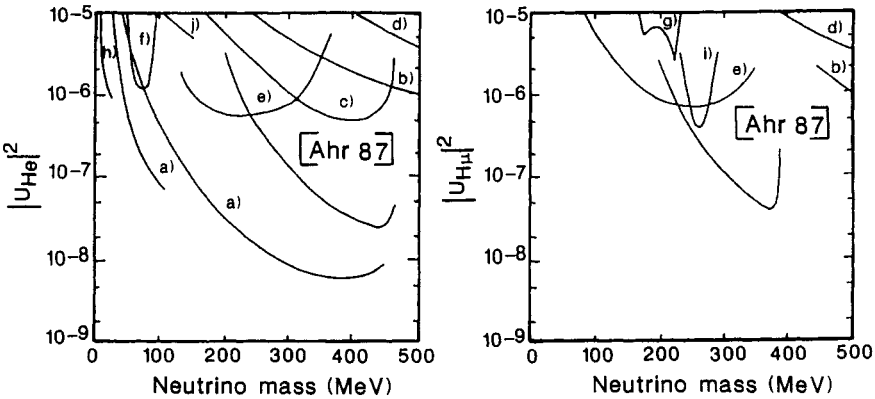


Figure 6.48. Bounds (90% confidence level) for the mixing matrix elements $|U_{e2}|^2 = |U_{He}|^2$ and $|U_{\mu 2}|^2 = |U_{H\mu}|^2$ as a function of the mass of the heavy neutrino component from accelerator experiments. The regions above the curves are excluded (after [Ahr87], from [Fei88a]).

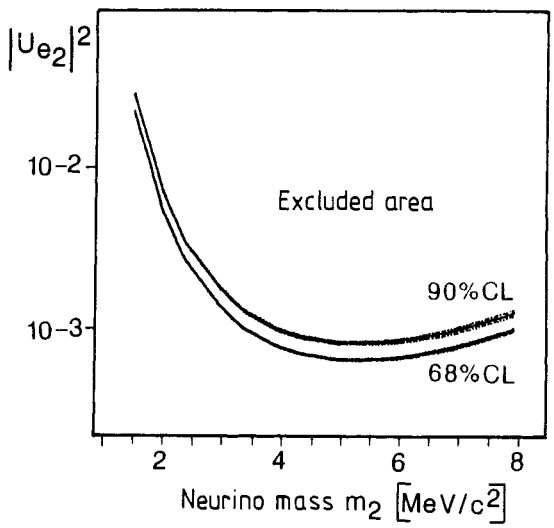


Figure 6.49. Bounds for $|U_{e2}|^2$ derived from measurement of the decay of $\overline{\nu}_e$ into e^+e^- pairs on the Gösgen reactor [Obe87], as a function of the mass m_2 (from [Fei88a]).

Gösgen [Obe87]. Electron–positron pairs created by the possible decay of the neutrino ν_2 were searched for by comparing the counting rates during reactor operation and with the reactor switched off. The detector comprised 375 litres of liquid scintillator. The bounds derived from these measurements are shown

in figure 6.49 in the $|U_{e2}|^2$–m_2 plane.

For heavy neutrinos with masses below the threshold of 1 MeV/c^2 the decay into a light neutrino and a γ quantum is the only directly observable process. Lower bounds for the lifetime were also derived in the reactor experiment mentioned above [Obe87].

Very sensitive information was obtained from the observation of the neutrino flux from the supernova SN1987A and recordings of the γ flux taken during this period by the Solar Maximum Mission satellite (see [Obe88]). No significant increase of the γ rate was observed in the energy interval 4.1 MeV $< E_\gamma <$ 6.4 MeV. The following bound for the stability of the electron antineutrino was derived from these data [Fei88b]

$$\frac{\tau_{\bar{\nu}_e}}{m_{\bar{\nu}_e}c^2} > 8.3 \times 10^{14} \text{ s/eV}. \tag{6.185}$$

Although no supernova neutrinos of other flavours have been detected here on Earth, the following bound for the dominant mass eigenstate ν_α may be calculated

$$\frac{\tau_\alpha}{m_\alpha c^2} > 3.3 \times 10^{14} \text{ s/eV}. \tag{6.186}$$

A neutrino decay has not yet been experimentally detected. Of course, under certain circumstances, such a process would have serious cosmological consequences, which we shall not go into here; readers are referred to the corresponding sections of [Gel88, Gro89, 90, Kol90].

Chapter 7

Neutrino Oscillations

7.1 INTRODUCTION TO NEUTRINO OSCILLATIONS AND THEIR PHENOMENOLOGY

We have already indicated in chapters 1 and 6 that the neutrino has a special place among the known elementary particles (see also [Gro89, 90, Lan88, Moh91a]). Experimental questions relating to the neutrino mass and the neutrino mixing are of particular interest. The question of the conservation of lepton number is closely related to these. If, as a result of mass terms in the Lagrange density, the different neutrino flavours mix, the lepton numbers relating to the flavours L_i ($i = e, \mu, \tau$) cannot be conserved, since a mixing permits conversion between the different neutrinos. If the neutrinos are Dirac particles the total lepton number $L = \sum_i L_i$ may be conserved. L cannot be a conserved quantity for Majorana particles since in this case particles and antiparticles are identical and the usual assignment of this additive quantum number no longer makes sense.

One important experimental approach to the determination of properties of neutrinos, including, in particular, their mass and mixing parameters, involves the search for neutrino oscillations. This phenomenon was first proposed by Pontecorvo [Pon57, 58] and is under discussion as a solution of the so-called solar neutrino problem (see below).

Neutrino oscillations involve the conversion of a neutrino of a certain flavour into a neutrino of a different flavour (see figure 7.1). Neutrino–antineutrino oscillations are also conceivable. In what follows, we shall initially only discuss flavour oscillations. This phenomenon means that, for example, given initially a pure ν_μ beam, then after a distance of flight x, there is a non-zero probability of detecting a ν_e.

Three different types of neutrino are currently known, namely ν_e, ν_μ and ν_τ (to date the existence of the latter has only been inferred indirectly). These flavours are eigenstates of the weak interaction, under which the neutrinos are

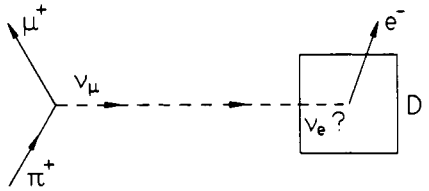


Figure 7.1. The principle of neutrino oscillation experiments (from [Kay89]). D stands for detector.

generated and under which the following decays occur

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{7.1a}$$

$$\pi^+ \rightarrow e^+ + \nu_e \tag{7.1b}$$

$$\tau^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\tau. \tag{7.1c}$$

It is possible to introduce a lepton number L_i $(i = e, \mu, \tau, \ldots)$ which takes the value $+1$ for a particle and -1 for an antiparticle. This lepton number is defined to be the eigenvalue of the lepton number operator $\mathcal{L}_i$

$$\mathcal{L}_i|\nu_j\rangle = \delta_{ij}|\nu_j\rangle \tag{7.2a}$$

$$\mathcal{L}_i|\overline{\nu}_j\rangle = -\delta_{ij}|\overline{\nu}_j\rangle. \tag{7.2b}$$

Experiments show that, to a good approximation, this lepton number is a conserved quantity. Since this conservation principle is not based on any known symmetry of the Lagrange density, it is possible that it is not a fundamental physical law.

The occurrence of neutrino oscillations

$$\nu_i \leftrightarrow \nu_j \quad i \neq j \tag{7.3}$$

would violate the principle of conservation of the lepton number, although the total lepton number L would still be conserved, as the following example shows

$$\begin{array}{cc} \nu_\mu \rightarrow \nu_e & \\ \delta L_\mu = -1 & \delta L_e = +1 \end{array} \tag{7.4}$$

implies

$$\delta L = \delta L_\mu + \delta L_e = 0.$$

Neutrino oscillations are only possible if at least one neutrino has a non-zero rest mass. If the eigenstates $|\nu_i\rangle$ of the lepton-number operator or the Hamiltonian of the weak interaction are not eigenstates of the mass operator a neutrino mixing occurs. This phenomenon is known from the quark sector. The quark states $|s\rangle$ and $|d\rangle$ are not eigenstates of the weak interaction but are rotated from these by the Cabibbo angle θ_C (see section 1.3.2).

7.2 THE FORMAL THEORY

7.2.1 The mass matrix and particle mixing

In this section we shall use the example of two lepton generations (e, μ) to show how neutrino mixing comes about as a result of the mass matrix. If the neutrinos are assigned a rest mass the Lagrangian then includes a neutrino mass term $\mathcal{L}_m$. For simplicity, we shall assume that neutrinos are Dirac particles and that CP is conserved. Under these conditions, the most general mass term is

$$-\mathcal{L}_m = \bar{\nu}\mathcal{M}_\nu \nu \tag{7.5a}$$

where

$$\nu = \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} \tag{7.5b}$$

and

$$\mathcal{M}_\nu = \begin{pmatrix} m_{\nu_e\nu_e}c^2 & m_{\nu_e\nu_\mu}c^2 \\ m_{\nu_e\nu_\mu}c^2 & m_{\nu_\mu\nu_\mu}c^2 \end{pmatrix}. \tag{7.5c}$$

In order to formulate the equations in a clear way, we shall not use the SI unit system in what follows, but the notation often used in the literature, $\hbar = c = 1$. In the basis of the interaction eigenstates (ν_e, ν_μ), the Lagrange density $\mathcal{L}_m$ generally has non-vanishing off-diagonal elements. However, because of its symmetry, the mass matrix may be diagonalized by a unitary transformation. This is achieved by the transformation

$$\nu = U^\dagger \nu' \tag{7.6a}$$

i.e. in our particular case

$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}. \tag{7.6b}$$

Since the physical properties are independent of the choice of basis, we have

$$\bar{\nu}\mathcal{M}_\nu \nu = \bar{\nu}'\mathcal{M}_\nu^{\text{Diag}}\nu' \tag{7.7a}$$

where

$$\mathcal{M}_\nu^{\text{Diag}} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}. \tag{7.7b}$$

Whence it follows that

$$\mathcal{M}_\nu^{\text{Diag}} = U\mathcal{M}_\nu U^\dagger. \tag{7.8}$$

For the mass eigenvalues m_1 and m_2 we have

$$m_{1/2} = \frac{1}{2}\left[m_{\nu_e\nu_e} + m_{\nu_\mu\nu_\mu} \pm \sqrt{(m_{\nu_e\nu_e} - m_{\nu_\mu\nu_\mu})^2 + 4m_{\nu_e\nu_\mu}^2}\right]. \tag{7.9}$$

The mixing angle is given by

$$\tan 2\theta = \frac{2m_{\nu_e\nu_\mu}}{m_{\nu_\mu\nu_\mu} - m_{\nu_e\nu_e}}. \tag{7.10}$$

The particles ν_1 and ν_2 have the masses m_1 and m_2, respectively. The matrix elements of $\mathcal{M}_\nu$ may be expressed in terms of the eigenvalues m_1 and m_2 and the angle θ

$$m_{\nu_e\nu_e} \equiv m_{\nu_e} = m_1 \cos^2\theta + m_2 \sin^2\theta \tag{7.11a}$$

$$m_{\nu_\mu\nu_\mu} \equiv m_{\nu_\mu} = m_1 \sin^2\theta + m_2 \cos^2\theta \tag{7.11b}$$

$$m_{\nu_e\nu_\mu} = (m_2 - m_1)\sin\theta\cos\theta. \tag{7.11c}$$

To represent a neutrino state in Hilbert space we require a basis. The interaction eigenstates $|\nu_i\rangle$ of the lepton-number operator $\mathcal{L}_i$ form a possible system of basis vectors. Since the lepton number is an observable, $\mathcal{L}_i$ is a self-adjoint operator so that the $|\nu_i\rangle$ form an orthonormal basis

$$\langle\nu_i|\nu_j\rangle = \delta_{ij}. \tag{7.12}$$

In the example at the beginning of this section we defined another basis in which the mass operator is diagonal

$$\mathcal{M}_\nu|\nu_\alpha\rangle = m_\alpha|\nu_\alpha\rangle \quad \alpha = 1, 2, 3, \ldots \tag{7.13}$$

where the m_α denote the individual particle masses. The $|\nu_\alpha\rangle$ are generally referred to as mass eigenstates. The temporal evolution of these states in the free space without interaction is given by the Dirac equation

$$|\nu_\alpha(t)\rangle = |\nu_\alpha\rangle \exp[\mathrm{i}(p_\alpha x - E_\alpha t)] \tag{7.14}$$

where

$$E_\alpha = \sqrt{p_\alpha^2 + m_\alpha^2}. \tag{7.15}$$

7.2.2 Flavour oscillations

In what follows (see the detailed description in [Gro89, 90, 92]) we shall assume, that the neutrinos are either of a purely Dirac or a purely Majorana nature, which does not lead to any measurable differences as far as flavour oscillations are concerned (see [Boe92]). In addition, we shall assume that the neutrinos are stable, ultrarelativistic particles located in the vacuum. For unstable neutrinos the decay width would have to be taken into account in (7.14) by a factor $\exp(-\Gamma t/2\hbar)$. The important phenomenon of neutrino oscillations in matter is discussed in section 7.3.5.

If neutrinos have a mass, the mass and flavour eigenstates are generally not identical, as we saw in section 7.2.1. A condition for this non-identity is that the eigenvalues of $\mathcal{M}_\nu$ should not all be degenerate; otherwise, $\mathcal{M}_\nu$ would simply be a multiple of the unit operator, i.e. the eigenstates of the mass operator would also be eigenstates of the lepton-number operator.

Thus, the flavour eigenstates $|\nu_i\rangle$ contain several mass components $|\nu_\alpha\rangle$. If the mass differences $m_\alpha - m_\beta$ are small enough (i.e. strictly speaking, smaller than the resolution of the experiment, so that ν_α and ν_β are indistinguishable [Nus76, Kay81]), then the $|\nu_i\rangle$ may be described by a coherent, quantum-mechanical superposition of mass eigenstates

$$|\nu_i\rangle = \sum_\alpha U_{i\alpha}|\nu_\alpha\rangle. \tag{7.16}$$

Here, U denotes a unitary mixing matrix. The unitarity follows from the fact that U associates two orthonormal bases with one another

$$U^\dagger = U^{-1} \quad \text{and} \quad \sum_i U_{\alpha i} U^\dagger_{i\beta} = \delta_{\alpha\beta}. \tag{7.17}$$

In the case of CP conservation, U is orthogonal and all elements $U_{\alpha i}$ are real.

Let us now consider the case in which a source at time $t = 0$ (at the point $\boldsymbol{x}$) emits neutrinos of a certain flavour l with fixed energy E and fixed momentum $\boldsymbol{p}$[1]. At that time, the neutrino state generated is represented by a superposition of plane waves corresponding to the different mass components

$$\begin{aligned} |\nu(\boldsymbol{x}, 0)\rangle = |\nu_l\rangle &= \sum_\alpha U_{l\alpha}|\nu_\alpha(\boldsymbol{x}, 0)\rangle \\ &= \sum_\alpha U_{l\alpha} \exp(-\mathrm{i}\boldsymbol{p}_\alpha \boldsymbol{x})|\nu_\alpha\rangle. \end{aligned} \tag{7.18}$$

At a later time t, we have

$$|\nu(\boldsymbol{x}, t)\rangle = \sum_\alpha U_{l\alpha} \exp(\mathrm{i}\boldsymbol{p}_\alpha \boldsymbol{x}) \exp(-\mathrm{i}E_\alpha t)|\nu_\alpha\rangle. \tag{7.19}$$

Since we are dealing with ultrarelativistic neutrinos, the following approximation holds

$$E_\alpha = \sqrt{m_\alpha^2 + p_\alpha^2} \simeq p_\alpha \left(1 + \frac{m_\alpha^2}{2p_\alpha^2}\right) = p_\alpha + \frac{m_\alpha^2}{2p_\alpha}. \tag{7.20}$$

If $m_\alpha \ll p_\alpha$ is satisfied for all mass eigenstates the neutrino moves at practically the speed of light, i.e. at time t it is to be found at the point $x = t$. Thus, for

[1] By definition neutrinos are always produced with a definite flavour.

points in the direction of momentum $\boldsymbol{p}$ we have

$$|\nu(\boldsymbol{x},t)\rangle = \sum_{\alpha} U_{l\alpha} \exp(\mathrm{i}\boldsymbol{p}_{\alpha}\boldsymbol{x}) \exp\left[-\mathrm{i}\left(p_{\alpha} + \frac{m_{\alpha}^2}{2p_{\alpha}}t\right)\right] |\nu_{\alpha}\rangle$$

$$= \sum_{\alpha} U_{l\alpha} \exp\left(-\mathrm{i}\frac{m_{\alpha}^2}{2p_{\alpha}}t\right) |\nu_{\alpha}\rangle. \tag{7.21}$$

The mass eigenstates $|\nu_{\alpha}\rangle$ may also be expressed in terms of interaction eigenstates

$$|\nu_{\alpha}\rangle = \sum_{k} U_{k\alpha}^{*} |\nu_{k}\rangle. \tag{7.22}$$

Substitution in (7.21) gives

$$|\nu(\boldsymbol{x},t)\rangle = \sum_{k} \left[\sum_{\alpha} U_{k\alpha}^{*} U_{l\alpha} \exp\left(-\mathrm{i}\frac{m_{\alpha}^2}{2p_{\alpha}}t\right)\right] |\nu_{k}\rangle. \tag{7.23}$$

It is clear that the wavefunction $\nu(\boldsymbol{x},t)$ which at time $t = 0$ describes a neutrino of flavour l now represents a superposition of all flavours.

We shall now calculate the probability of a transition from ν_l to ν_k $(k \neq l)$. For the probability amplitude we have

$$A_{l\to k}(t) = \langle \nu_k | \nu(\boldsymbol{x},t)\rangle = \sum_{k'} \sum_{\alpha} U_{k'\alpha}^{*} U_{l\alpha} \exp\left(-\mathrm{i}\frac{m_{\alpha}^2}{2p_{\alpha}}t\right) \langle \nu_k | \nu_{k'}\rangle \tag{7.24a}$$

$$= \sum_{\alpha} U_{k\alpha}^{*} U_{l\alpha} \exp\left(-\mathrm{i}\frac{m_{\alpha}^2}{2p_{\alpha}}t\right). \tag{7.24b}$$

For antineutrinos, we have an analogous equation

$$A_{\bar{l}\to\bar{k}}(t) = \sum_{\alpha} U_{k\alpha} U_{l\alpha}^{*} \exp\left(-\mathrm{i}\frac{m_{\alpha}^2}{2p_{\alpha}}t\right). \tag{7.24c}$$

Assuming CP conservation, it follows that

$$A_{l\to k}(t) = A_{\bar{l}\to\bar{k}}(t) \tag{7.25}$$

since the matrices U are real. The probability that after time t the neutrino is in state $|\nu_k\rangle$ is given by the square of the absolute value of the amplitude (7.24*b*)

$$P_{l\to k}(t) = |A_{l\to k}(t)|^2$$

$$= \sum_{\alpha} \sum_{\beta} U_{k\alpha}^{*} U_{l\alpha} U_{k\beta} U_{l\beta}^{*} \exp\left[-\mathrm{i}\left(\frac{m_{\alpha}^2}{2p_{\alpha}} - \frac{m_{\beta}^2}{2p_{\beta}}\right)t\right]. \tag{7.26}$$

By virtue of the assumption $m_\alpha \ll p_\alpha$, we have $p_\alpha = p_\beta \equiv p$. Thus, (7.26) becomes

$$P_{l\to k}(t) = \sum_\alpha \sum_\beta U_{l\alpha} U^*_{l\beta} U^*_{k\alpha} U_{k\beta} \exp\left(-\mathrm{i}\frac{\Delta m^2_{\alpha\beta}}{2p} t\right) \tag{7.27a}$$

where

$$\Delta m^2_{\alpha\beta} = m^2_\alpha - m^2_\beta. \tag{7.27b}$$

For increased clarity (7.27*a*) may also be written as follows

$$P_{l\to k}(t) = \sum_\alpha |U_{l\alpha}|^2 |U_{k\alpha}|^2 + \sum_{\alpha\neq\beta} U_{l\alpha} U^*_{l\beta} U^*_{k\alpha} U_{k\beta} \exp\left(-\mathrm{i}\frac{\Delta m^2_{\alpha\beta}}{2p} t\right). \tag{7.28}$$

The diagonal term is not dependent on time; it represents an average transition probability which is modulated by the time-dependent second term. The probability that there is no transition into a flavour different from l is

$$P_{l\to l}(t) = 1 - \sum_{l\neq k} P_{l\to k}(t). \tag{7.29}$$

Equation (7.28) exhibits a nice oscillatory behaviour as a function of time and position (since $v \approx c = 1$, we have $x = t$). This can be particularly clearly seen from the decomposition into real and imaginary parts

$$\begin{aligned} P_{l\to k}(t) = {} & \sum_\alpha |U_{l\alpha}|^2 |U_{k\alpha}|^2 \\ & + \sum_{\alpha\neq\beta} \mathrm{Re}\,(U_{l\alpha} U^*_{l\beta} U^*_{k\alpha} U_{k\beta}) \cos\left(\frac{\Delta m^2_{\alpha\beta}}{2p} t\right) \\ & + \sum_{\alpha\neq\beta} \mathrm{Im}\,(U_{l\alpha} U^*_{l\beta} U^*_{k\alpha} U_{k\beta}) \sin\left(\frac{\Delta m^2_{\alpha\beta}}{2p} t\right). \end{aligned} \tag{7.30}$$

If we assume CP conservation, the last term in (7.30) vanishes. The probability $P_{l\to k}(x)$ varies periodically with time and distance x from the source. The periodicity is characterized by the oscillation length

$$L_{\alpha\beta} = \frac{4\pi p}{\Delta m^2_{\alpha\beta}} = \frac{4\pi E}{\Delta m^2_{\alpha\beta}}. \tag{7.31}$$

We note that for a relativistic particle, $p = E$. These oscillations represent an interference effect, which ultimately is a result of the non-diagonal mass terms in (7.5*c*). If all masses are identical, in particular, if all masses vanish, no oscillations occur ($L_{\alpha\beta} \to \infty$). If the neutrino is already in a mass eigenstate ($|\nu_l\rangle = |\nu_\alpha\rangle$) at time $t = 0$, then no oscillations take place into a flavour different from l. This again shows that the occurrence of neutrino oscillations requires both a mass, and a neutrino mixing.

7.2.3 Average values over time

Until now, we have implicitly assumed a point-like neutrino source, where the neutrino was emitted at time $t = 0$ at the point x. However, in general, the neutrinos in experiments come from extended sources (e.g. a reactor or the Sun), so that it is impossible to determine the exact flight time between the point of creation and the source. Consequently, measurement of the fluxes of the different neutrino flavours only provides information about the average values of the transition probabilities over time

$$\langle P_{l\to k}(t)\rangle_T = \frac{1}{T}\int_0^T P_{l\to k}(t)\,\mathrm{d}t \tag{7.32a}$$

$$= \sum_\alpha |U_{l\alpha}|^2|U_{k\alpha}|^2 + \frac{1}{T}\int_0^T \sum_{\alpha\neq\beta} U_{l\alpha}U_{l\beta}^*U_{k\alpha}^*U_{k\beta}\exp\left(-\mathrm{i}\frac{\Delta m_{\alpha\beta}^2}{2p}t\right)\mathrm{d}t. \tag{7.32b}$$

Since the integral in (7.32*b*) is finite, for large T, the second term in the sum may be neglected in comparison with the first. Thus, as $T \to \infty$ we have

$$\langle P_{l\to k}(t)\rangle_T = \sum_\alpha |U_{l\alpha}|^2|U_{k\alpha}|^2. \tag{7.33}$$

In this case, the transition probability no longer depends on the mass difference but only on the mixing matrix. Thus, an experiment designed to measure the average value over time is only sensitive to the mixing matrix.

7.2.4 Neutrino oscillations and the principles of quantum mechanics

Oscillations between different observable states constitute a fundamental characteristic of a quantum-mechanical system. An analogous phenomenon is found in particle physics for the neutral kaons. The following discussion is oriented towards the corresponding sections of [Kay89] and [Boe92].

To remain as specific as possible, we introduce a simple 'gedanken' experiment to investigate the transition from muon neutrinos to electron neutrinos (figure 7.1). Since this relates to the search for a flavour which was not previously present, such an experiment is also called an 'appearance' experiment. The ν_μ should come from the reaction

$$\pi^+ \to \mu^+ + \nu_\mu. \tag{7.34}$$

An electron neutrino detector is now placed at various distances from the decay region. In addition, detectors for determining the energies and the momenta

of the pion and the muon are also assumed to be present. By virtue of the kinematics, it is therefore possible to determine the energy and the momentum of the emitted neutrinos, whence also their mass. We now suppose that the muon neutrino generated in the decay (7.34) is a superposition of two mass eigenstates with the masses m_1 and m_2; thus, the wavefunction consists of two components with different energies and momenta, where

$$E_1 - E_2 = \frac{m_1^2 - m_2^2}{2E_\pi} \tag{7.35a}$$

$$p_1 - p_2 = \frac{m_1^2 - m_2^2}{p_1 + p_2} \frac{2E_\pi - E_1 - E_2}{2E_\pi}. \tag{7.35b}$$

Here, E_π denotes the energy of the decaying pion. The phase difference between the two components of the neutrino wavefunction is given by

$$\begin{aligned} \Phi(x,t) &= (E_1 - E_2)t - (p_1 - p_2)x \\ &= \frac{m_1^2 - m_2^2}{E_1 + E_2} t + (p_1 - p_2)\left(\frac{p_1 + p_2}{E_1 + E_2} t - x\right). \end{aligned} \tag{7.36}$$

The first term leads directly to the definition of the oscillation length (see (7.31))

$$L_{12} = 2\pi \frac{E_1 + E_2}{|m_1^2 - m_2^2|} \simeq 2\pi \frac{2\overline{E}}{|m_1^2 - m_2^2|}. \tag{7.37}$$

The second term is usually neglected, which represents a very good approximation for highly relativistic neutrinos. It vanishes for

$$x = v_0 t = \frac{p_1 + p_2}{E_1 + E_2} t \tag{7.38}$$

where v_0 denotes the average velocity of the two components. In this case, the two components of the neutrino wavefunction remain coherent and the oscillation phenomenon may occur.

One may also enquire about the coherence length, i.e. at what distance from the source is there still coherence. Suppose that the pulse width of the neutrino source is $\delta p = 1/\delta x$ where δx is a measure of the extent of the source. Using (7.36) the phase uncertainty is determined to be

$$\delta\Phi(x) \simeq \frac{\Delta m_{12}^2}{2p^2} \delta p x \tag{7.39}$$

where $\Delta m_{12}^2 = |m_1^2 - m_2^2|$. Coherence is guaranteed for $\delta\Phi < 1$. If we define the distance x to be the coherence length L_{coh} at which $\delta\Phi(x) = 1$, the maximum number of observable oscillations is then given by

$$N = \frac{L_{\text{coh}}}{L_{12}} \simeq \frac{p}{\delta p} = p\delta x. \tag{7.40}$$

In standard experimental arrangements N is a very large number. In general, the oscillations will vanish earlier than one might expect from (7.40), since the measurements will not be carried out under ideal conditions (see [Boe87, 92]).

At this point, we return to the gedanken experiment mentioned above. Let us suppose that we detect the occurrence of oscillations, by observing electrons in our detector which are generated by an interaction with a ν_e. Suppose that the transition probability $P_{\nu_\mu \to \nu_e}(t)$ is given by (7.28). If one were now to measure the momentum and the energy of the pions very accurately, one could determine which neutrino mass eigenstate was actually emitted (this is not currently experimentally possible). The neutrino wavefunction would then no longer be a superposition of mass eigenstates, but a specific individual component would be singled out. However, in this case, oscillations could no longer occur since these are the result of interference between different mass eigenstates.

On the other hand, according to the Heisenberg uncertainty principle, very accurate determination of the pion momentum would mean that information about the location of the decay would be largely lost. Calculating the energy and the momentum of the emitted neutrino kinematically, the mass is given by

$$m_\nu^2 = E_\nu^2 - p_\nu^2. \tag{7.41}$$

The error, for uncorrelated p_ν and E_ν, is given by

$$\Delta(m_\nu^2) = \sqrt{(2E_\nu)^2(\Delta E_\nu)^2 + (2p_\nu)^2(\Delta p_\nu)^2}. \tag{7.42}$$

Now, only if $\Delta(m_\nu^2)$ is smaller than $|m_1^2 - m_2^2|$ is it possible to distinguish between the two eigenstates experimentally, i.e. the condition

$$2p_\nu \Delta p_\nu < |m_1^2 - m_2^2| \tag{7.43}$$

must be satisfied. Thus, the uncertainty in the position of the source (up to constant factors) amounts to

$$\delta_x > L_{12} \tag{7.44}$$

and is greater than the oscillation length. Consequently, the oscillations are smeared out. When one determines the emitted mass eigenstate it is impossible to determine the location of the emission sufficiently accurately that oscillations may be seen. A more adequate treatment of this problem using wave packets leads to the same results [Kay81].

7.2.5 Mixing of two neutrino flavours

When analysing experimental data, a two-state mixing, in general the $\nu_e \nu_\mu$ mixing, is often assumed.

The mixing matrix U then is a unitary 2×2 matrix of the form (see (7.6*b*))

$$U = \begin{pmatrix} \cos\theta & e^{i\phi}\sin\theta \\ -e^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}. \tag{7.45}$$

The phase $e^{\pm i\phi}$ is responsible for the CP violation. This phase factor has no part to play as far as the phenomenon of neutrino oscillations is concerned, but is relevant to the discussion of double-beta decay (see chapter 6). In what follows, we shall assume CP conservation, i.e. the phase factor may only take the values 1 or i [Wol81]. The expression (7.45) then becomes (7.6*b*).

In the mass basis the states ν_e and ν_μ are explicitly given by

$$|\nu_e\rangle = \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \tag{7.46a}$$

$$|\nu_\mu\rangle = -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle. \tag{7.46b}$$

The mixing angle is limited to the interval $0 \leq \theta \leq \pi/4$. Until now the transition probabilities have always been expressed as a function of time t. In experiments, one does not usually know the flight time of the neutrino, but only the distance of the detector from the source. Since neutrinos move at practically the speed of light ($v = c = 1$) it is easy to rewrite the formulae using $x = t$. From (7.30) it follows that the probability of finding a ν_e emitted at $x = 0$ with energy E_ν at distance x as a ν_μ is

$$\begin{aligned} P_{\nu_e \to \nu_\mu}(x) &= 2\cos^2\theta\sin^2\theta - 2\sin^2\theta\cos^2\theta\cos\frac{\Delta m^2}{2p_\nu}x \\ &= \tfrac{1}{2}\sin^2(2\theta)\left(1 - \cos\frac{\Delta m^2}{2p_\nu}x\right) \\ &= \tfrac{1}{2}\sin^2(2\theta)\left(1 - \cos\frac{2\pi x}{L}\right) \\ &= \sin^2(2\theta)\sin^2\frac{\pi x}{L} \end{aligned} \tag{7.47}$$

where $\Delta m^2 = |m_1^2 - m_2^2|$ and the oscillation length is

$$L = \frac{4\pi p_\nu}{\Delta m^2} = \frac{4\pi E_\nu}{\Delta m^2} \tag{7.48}$$

(defined by $(\Delta m^2/2E_\nu)L = 2\pi$). Correspondingly, we have

$$P_{\nu_e \to \nu_e}(x) = 1 - \sin^2(2\theta)\sin^2\frac{\pi x}{L}. \tag{7.49}$$

The probabilities oscillate with the characteristic length L (see figure 7.2). The amplitude of the oscillations depends on the mixing angle θ and is largest

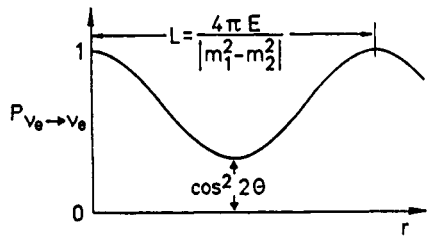


Figure 7.2. Neutrino oscillation in the ν_e–ν_μ system. The probability that an electron neutrino emitted at $r = 0$ is found after a distance r as an electron neutrino oscillates with the characteristic length $L = 4\pi E/|m_1^2 - m_2^2|$ (from [Gro89, 90]).

for $\theta = 45°$ (maximum mixing), when the electron neutrino at the points $x = L(n + 1/2)$ for integers n is fully converted into a muon neutrino.

The search for neutrino oscillations is one of the most important approaches to the experimental determination of a non-zero neutrino mass. More precisely, however, only mass differences are measured since only the quantity Δm^2 is included as a parameter. We note that, strictly speaking, in the case $\theta \neq 0$ one should no longer speak of a mass of the ν_e or the ν_μ since these states do not have a defined mass. However, if the mixing angle is very small there is a still a large overlap of flavour and mass eigenstates (i.e. $|\nu_e\rangle \approx |\nu_1\rangle$ etc), so that there is some justification for speaking of the mass of the ν_e.

Before we discuss the experimental methods, we shall briefly consider the conditions under which we may hope to observe neutrino oscillations at all. There are two approaches to the detection. Either one determines the oscillating $\sin^2$ terms in the neutrino intensity or one finds a constant transition probability $P_{\nu_e \to \nu_\mu} \neq 0$ or $P_{\nu_e \to \nu_e} \neq 1$.

If the distance x of the detector from the source is very much smaller than the oscillation length L, the neutrino will remain in its original flavour

$$P_{\nu_e \to \nu_e}(x) \simeq 1 \quad \text{for } x \ll L. \tag{7.50}$$

For large x, as a result of the momentum uncertainty the oscillations will be washed out; one finds

$$P_{\nu_e \to \nu_\mu}(x) \to \tfrac{1}{2} \sin^2(2\theta) \quad \text{for } x \gg L. \tag{7.51}$$

Then, it is in principle possible to find a ν_μ neutrino in a ν_e beam; however, the probability no longer varies as a function of distance.

From these considerations we deduce, that an oscillatory behaviour is only found if the distance between the source and the detector is of the order of magnitude of the oscillation length L. In addition the dimensions of the source

and the detector must be less than L, otherwise it is only possible to detect the average value

$$\langle P_{\nu_e \to \nu_\mu}(t) \rangle_T = \tfrac{1}{2} \sin^2(2\theta). \tag{7.52}$$

At this point we shall briefly discuss the distinction between Dirac and Majorana neutrinos in the context of neutrino oscillations. Until now, we have discussed flavour oscillations in which the individual lepton numbers (L_e, L_μ, L_τ) are no longer conserved, unlike the total lepton number. Oscillations only take place between states with the same helicity. In this case, it is not possible to distinguish between Dirac and Majorana particles. However, flavour oscillations are not the most general oscillations possible. In the general case, transitions between states with the opposite helicity are also possible, i.e. $\nu \to \bar{\nu}$ where the lepton number is altered by two units ($|\Delta L| = 2$). A neutrino beam ν_i may create a positively charged antilepton $\bar{l}_i^+$ at a distance x from the source with a certain probability. However, in the case of a purely left-handed weak interaction, the corresponding amplitude contains a suppression factor of the form m_ν/E [Bah78]. This helicity factor does not arise in the presence of right-handed charged currents. For a further discussion readers are referred to [Boe87, 92, Gro89, 90].

7.3 EXPERIMENTS ON NEUTRINO OSCILLATIONS

7.3.1 The sensitivity of different experimental arrangements

In order to investigate the sensitivity of different experiments, we substitute numerical values for $\hbar$ and c in the expression for the oscillation length and obtain

$$L(E) = \frac{4\pi E}{\Delta m^2} = 2.5 \frac{E[\mathrm{MeV}]}{\Delta m^2[\mathrm{eV}^2]}[\mathrm{m}]. \tag{7.53}$$

In practice, reactors, the Sun and high- and low-energy accelerators are available as neutrino sources. Table 7.1 gives an overview of the associated oscillation lengths for a given mass difference. Solar neutrinos have a comparable energy to reactor neutrinos.

The neutrino sources referred to here do not in general give rise to a monoenergetic neutrino beam. Figure 7.3 shows measured antineutrino spectra for the thermal fission of ^{235}U and ^{239}Pu [Fei82, Sch85a], i.e. typical spectra as they occur in nuclear reactors. Since the oscillation length is energy dependent, a measurement of the energy would be desirable when detecting neutrinos, since otherwise it is only possible to measure the transition probability averaged over the energy. As in the case of the average value over time, the information about the mass difference is lost.

Since when detecting neutrino oscillations, the distance between the source and the detector x is required to be not too small in comparison with the

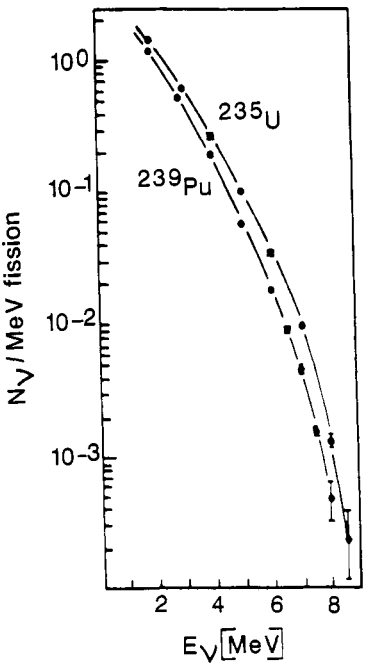


Figure 7.3. Experimental neutrino spectra of ^{235}U and ^{239}Pu from the experiments of [Fei82] and [Sch85a] (from [Fei88a]).

Table 7.1. Neutrino sources and typical energies, together with the oscillation lengths which result for given mass parameters.

Source	Energy	$\Delta m^2 = 1$ eV^2 L [m]	10^{-6} eV^2 L [km]	10^{-11} eV^2 L [km]
CERN SPS	100 GeV	250 000	2.5×10^8	2.5×10^{13}
CERN PS, BNL AGS	5 GeV	12 500	1.25×10^7	1.25×10^{12}
LAMPF	30 MeV	75	75 000	7.5×10^9
Reactor	4 MeV	10	10 000	10^9
Sun	0.2–10 MeV			1.5×10^8

oscillation length, in the case of small mass differences, low-energy experiments are much more sensitive than measurements at high-energy accelerators, as table 7.1 shows. The oscillatory behaviour is determined by the ratio of x to the neutrino energy. Figure 7.4 illustrates the areas which are accessible

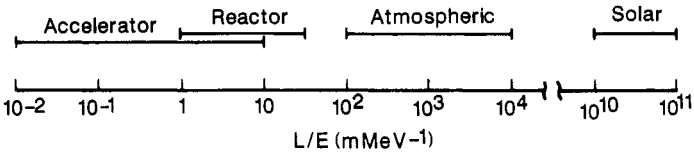


Figure 7.4. Overview of the accessible parameter area L/E in various experiments (after [Boe92]).

in the experiments. The different methods are complementary as far as their sensitivity is concerned (see also figure 7.5).

There are two main types of experiment:

- Experiments which search for the occurrence of a neutrino flavour which was not originally present in the neutrino beam ('appearance' experiments).
- Experiments which search for a decrease of the flavour which was originally present in the beam (so-called 'disappearance' experiments).

In appearance experiments one begins with a beam which ideally consists solely of neutrinos of flavour l (or a known fraction of other flavours) and searches for the occurrence of a flavour l' at distance x from the source. This requires neutrino energies which are sufficient for detection of the neutrino with the flavour l' in the detector (e.g. by the reaction (7.56)). The advantage of this type of measurement is its high sensitivity to small mixing angles, since even the detection of a few neutrinos of the 'wrong' flavour is sufficient. Usually only one specific channel (e.g. $\nu_\mu \rightarrow \nu_e$) is measured.

The second method is less sensitive to the mixing amplitudes. The sensitivity is limited by uncertainties in the neutrino flux, in the neutrino spectra and the detection probabilities of the detectors. On the other hand, disappearance experiments allow one to detect transitions in all possible final channels, including transitions into sterile neutrinos (see section 6.2.6), i.e. neutrinos which are not observed in nature (e.g. right-handed neutrinos and left-handed antineutrinos), the existence of which, however, is required by certain theories. In addition, when low-energy neutrinos are used, these experiments are very sensitive to small mass parameters $\Delta m^2_{\alpha\beta}$. Figure 7.5 summarizes the sensitivity of the different experiments as far as the mass and the mixing angle are concerned.

7.3.2 Reactor experiments

A number of experiments with low-energy antineutrino beams were carried out using nuclear reactors. In these, the disappearance channel $\overline{\nu}_e \rightarrow \nu_x$ was studied, where the antineutrino flux was determined over typical distances of 10–100 m. Because of the low neutrino energy, these measurements are of great interest,

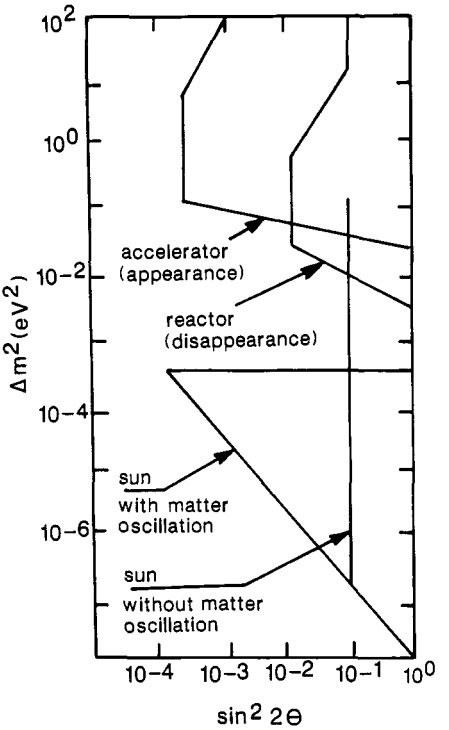


Figure 7.5. Sensitivity of different experiments to neutrino oscillations in the Δm^2–$\sin^2 2\theta$ plane (schematic). The area to the right of each curve is experimentally accessible (from [Fei88a]).

particularly for determination of small mass parameters. Since only the decrease in the $\overline{\nu}_e$ flux and not the occurrence of a $\overline{\nu}_\mu$ flux is experimentally accessible, reactor experiments are not well suited to small mixing angles.

Nuclear reactors form the most intensive terrestrial $\overline{\nu}_e$ source. On average, six antineutrinos are generated per fission event as the result of the β decay of unstable fission products. The energy released in the fission of a ^{235}U nucleus is approximately 200 MeV. This gives rise to an antineutrino production of 1.9×10^{20} per second per GW of thermal power. This flux is emitted isotropically over the whole solid angle. Figure 7.3 shows the typical energy spectrum of the neutrinos. It extends up to approximately 10 MeV, but the flux falls off practically exponentially at large energies. Figure 7.4 shows that reactor experiments with neutrino energies in the range from 2–10 MeV with typical distances of from 10 to 100 m between source and detector are able to investigate mass parameters from ~ 0.02 eV^2 to 10 eV^2.

Since the fission products are very neutron rich, practically only β^- transitions occur, which lead to an emission of antineutrinos. (Because of the low β decay energies, no other flavours occur apart from the electronic antineutrino.) The fraction of ν_e from β^+ decays and electron capture of the rare, neutron-deficient nuclei only amounts to from 10^{-5} to 10^{-8}, depending on the energy [Sch84]. Thus, a nuclear reactor is a very pure $\overline{\nu}_e$ source.

We now note one difficulty in the evaluation and interpretation of the results of measurements on reactor neutrinos. The determination of the oscillation parameters Δm^2 and $\sin^2 2\theta$ requires a knowledge of the fission yields and of the neutrino spectrum formed from the contributions of all fission products. To reduce the uncertainties associated with this the spectra measured at different distances may be compared.

A number of experiments to search for neutrino oscillations have been carried out on various reactors. Table 7.2 gives an overview of the measurements to date.

Table 7.2. List of reactor experiments to date.

Reactor	Thermal capacity [MW]	Distance [m]	Reference
ILL-Grenoble (France)	57	8.75	[Kwo81]
Bugey (France)	2800	13.6, 18.3	[Cav84]
Rovno (USSR)	1400	18.0, 25.0	[Afo85]
Savannah River (USA)	2300	18.5, 23.8	[Bau86]
Gösgen (Switzerland)	2800	37.9, 45.9, 64.7	[Zac86]
Krasnoyarsk (USSR)	?	57.0, 57.6, 231.4	[Vid94]
Bugey III (France)	2800	15, 40, 95	[Dec95]

Oscillations with large mass differences have been sought, in particular at the 57 MW research reactor of the Institut Laue-Langevin in Grenoble. Because of the associated small oscillation length, the research reactor in Grenoble, which has a small reactor core in comparison with a normal power reactor, is particularly appropriate, despite the lower neutrino production. The nuclear fuel was enriched uranium with a 97% enrichment of the isotope ^{235}U. The neutrino detector was installed at a distance 8.76 m from the 'point-like' neutrino source. There, the $\overline{\nu}_e$ flux was 9.8×10^{11} cm^{-2} s^{-1}. In this experiment 4890 ± 180 neutrino events were recorded (with a count rate of 1.58 per hour). No evidence of the occurrence of neutrino oscillations was found in the parameter area investigated [Kwo81].

In contrast, an experiment on the French pressurized water reactor at

Bugey, which has a thermal power of 2800 MW, initially provided evidence of neutrino oscillations [Cav84]. Antineutrinos were detected at distances of 13.6 m and 18.3 m from the reactor core. The flux in the detector at 13.6 m was 2×10^{13} cm^{-2} s^{-1}. While the individual spectra at both positions were consistent with those expected without oscillations the ratio of the two spectra provided clear indications of an effect. The best fit to the data was given by $\Delta m^2 = 0.2$ eV^2, $\sin^2 2\theta = 0.25$. However, this result, which has meanwhile been withdrawn (see [Pes88]), was not confirmed by other experiments and contradicts more recent measurements (see below).

Other experiments on a Soviet reactor in Rovno [Afo83, 85] and on the reactor in Savannah River in the USA [Bau86, Sob86] also produced no evidence for the occurrence of neutrino oscillations.

In what follows, as a representative case we shall discuss in somewhat more detail what was, until recently, the most sensitive reactor experiment, performed on the Swiss power reactor in Gösgen (see [Vui82, Gab84, Zac85, 86a–c]). This pressurized water reactor has a thermal power of 2800 MW, corresponding to an antineutrino flux of approximately 5×10^{20} per second. We note that these neutrinos carry away around 140 MW irretrievably. The working cycle of this reactor is approximately 11 months; in the following month a third of all fuel rods are replaced before the reactor is brought into operation again.

The experiment in Gösgen was based on the measurement of the energy spectrum and the flux of the antineutrinos at three different distances from the reactor core (37.9 m, 45.9 m, 64.7 m). Approximately 10 000 antineutrinos were recorded at each of these positions. Because the flight path of the antineutrinos between the source and the detector is relatively long in comparison with other experiments, it was possible to measure smaller mass parameters. To ensure that the effect of variations in the composition of the nuclear fuel was small the individual measurements were carried out during equivalent periods of the reactor cycle.

If there are neutrino oscillations, we expect a deformation of the measured antineutrino spectra as a function of the distance (see figure 7.6). While a monochromatic neutrino beam of a specific flavour is subject to intensity modulation in the direction of propagation, a continuous energy spectrum at a fixed distance from the source is deformed according to the energy dependence of the expression

$$P_{\nu_e \to \nu_\mu} = \sin^2 2\theta \, \sin^2 \frac{\Delta m^2 x}{4E_\nu}. \tag{7.54}$$

The maxima of the transition probability occur at

$$\frac{\Delta m^2 x}{4E_\nu} = (2n-1)\frac{\pi}{2} \qquad n \in N \tag{7.55}$$

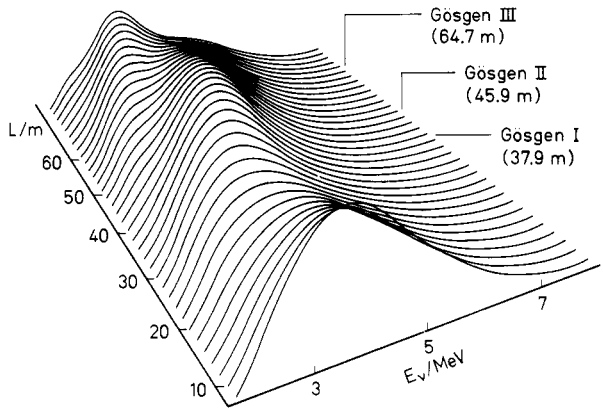


Figure 7.6. Modulation of the neutrino spectrum as a function of the distance from the reactor core in the presence of neutrino oscillations, based here, as an example, on the oscillation parameters $\Delta m^2 = 0.2$ eV^2 and $\sin^2 2\theta = 0.4$ (from [Zac86b]).

i.e. they lie on bands of straight lines $x \sim E_\nu$. These measured spectra may be used to analyse forbidden and allowed parameter combinations. A considerably more restrictive analysis of the data is possible, if one knows the absolute spectral $\overline{\nu}_e$ flux of the source. The reason for this is that because of the final extension of the source and the limited energy resolution of the detector, neutrino oscillations with a small oscillation length are only manifest through a decrease in the flux of the detectable flavour.

We shall now briefly consider the neutrino spectrum of a reactor. In a pressurized water reactor, such as the Gösgen power reactor, the overall power consists of contributions from four fissile isotopes, namely ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu. At the beginning of the reactor cycle (Gösgen) the composition of the fissile material leads to the following percentage fractions of the total number of fission events: ^{235}U (69%), ^{238}U (7%), ^{239}Pu (21%) and ^{241}Pu (3%). Other fissile isotopes such as ^{236}U and ^{240,242}Pu make only a very small contribution ($< 0.1\%$) to the neutrino spectrum and may be neglected.

We note that the relative fraction of the fissile isotopes in the overall fission rate changes over the period of operation as the uranium isotopes are burnt up while plutonium is bred (see figure 7.7). Since the average numbers of neutrinos released per fission and the energy distributions of these are different for the different isotopes the time dependence of the composite spectrum over the whole period must be explicitly taken into consideration. The calculation of the theoretically expected neutrino spectrum thus requires a knowledge of the operational parameters of the reactor and of the individual spectra. Figure 7.8

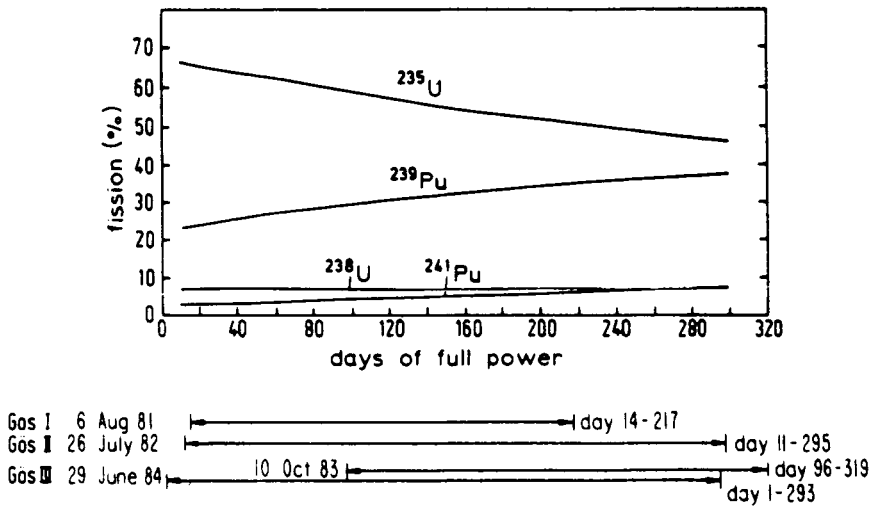


Figure 7.7. Relative contributions of the four relevant nuclides ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu to the overall number of fission events, as a function of days of reactor operation at full power. The lower part of the figure shows the measurement periods and the start of measurements for experiments I–III of the Gösgen experiment (from [Zac86a,b]).

shows the differences between the reactor neutrino spectra for various core compositions. Table 7.3 gives the contributions of the individual isotopes to the reactor power, averaged over the time of operation for various experiments.

The neutrino spectra of ^{235}U, ^{239}Pu and ^{241}Pu are now essentially known from β-spectroscopic measurements of the energy distributions of the simultaneously emitted electrons, although inconsistencies remain in the high-energy range [Fei82, Sch85a, 86, Key85]. However, as far as the contribution from fission of ^{238}U induced by fast neutrons is concerned, it is still necessary to rely on theoretical estimates (see e.g. [Kla81, 82a,b, Vog81]).

The detection of neutrinos in such reactor experiments is based on the reaction

$$\overline{\nu}_e + p \rightarrow e^+ + n. \tag{7.56}$$

Because of the low neutrino energies, corresponding reactions of the other flavours cannot occur in the detector since the threshold energies are far too high (table 7.4). Thus, the presence of neutrino oscillations can only manifest itself through a deficit in the measured $\overline{\nu}_e$ rate or a deformation of the spectrum.

For shielding against cosmic radiation the neutrino detector system is located in a concrete bunker. The neutrino detector consists of two systems, which are arranged in a cube with side length of about 1 m (see figure 7.9) and permit separate detection of the two reaction products. A total of 30 liquid

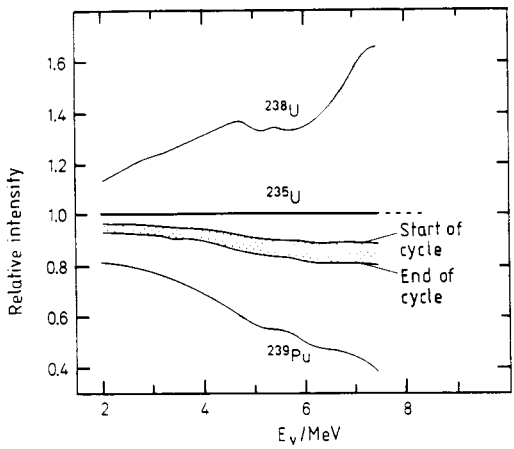


Figure 7.8. Differences in the neutrino spectra from different fissile materials relative to pure ^{235}U. The two curves, cycle start and cycle end, demonstrate the temporal evolution of a reactor neutrino spectrum during a full reactor cycle for a mixed reactor core of ^{235}U, ^{239}Pu, ^{238}U and ^{241}Pu, such as that typically found in a nuclear power plant (from [Zac86b]).

Table 7.3. Average contributions of the individual isotopes to the reactor output over the period of operation.

Isotope	^{235}U	^{238}U	^{239}Pu	^{241}Pu
Energy/fission [MeV]	201.7 ± 0.6	205.0 ± 0.9	210 ± 0.9	212.4 ± 1.0
Gösgen 37.9 m	61.9%	6.7%	27.2%	4.2%
Gösgen 45.9 m	58.4%	6.8%	29.8%	5.0%
Gösgen 64.7 m	54.3%	7.0%	32.9%	5.8%
Rovno 18 m	60.6%	7.4%	27.7%	4.3%
Bugey 13.6 m	62.1%	7.6%	26.4%	3.9%
Bugey 18.3 m	47.9%	8.2%	36.9%	7.0%

scintillator cells are used to determine the energy of the positrons from the reaction (7.56); they also form the target for the antineutrinos. The scintillator liquid consists of 377 litres of mineral oil. This provides the antineutrinos with a target of 2.4×10^{28} protons. At both ends of the cells are photomultipliers to detect the scintillation light, the intensity of which, for full absorption of the positron, is proportional to its kinetic energy. The antineutrino spectrum may be determined directly from the positron energy E_{e^+}, since E_{e^+} only differs from

Table 7.4. Energy thresholds for the reactions of type $\overline{\nu}_l + p \rightarrow n + l^+$.

Reaction	Threshold energy [MeV]
$\overline{\nu}_e + p \rightarrow n + e^+$	1.804
$\overline{\nu}_\mu + p \rightarrow n + \mu^+$	100
$\overline{\nu}_\tau + p \rightarrow n + \tau^+$	3600

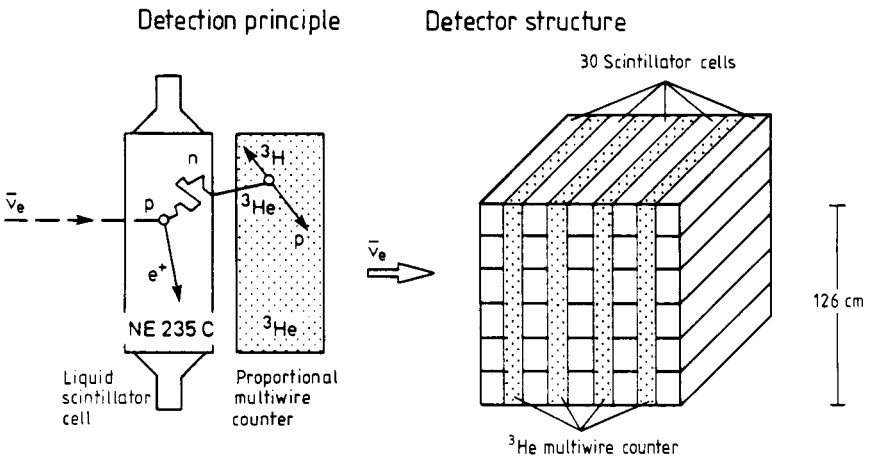


Figure 7.9. The neutrino detector in the Gösgen experiment. It consists of two different types of detector for separate detection of the end products of the reaction $\overline{\nu}_e + p \rightarrow n + e^+$. Positrons are recorded in the scintillation counter, which also supplies the protons for the neutrino reaction. The neutrons produced at the same time are detected in a ^{3}He multiwire counter with the formation of a tritium nucleus and a proton (from [Zac86b]).

the energy $E_{\overline{\nu}}$ of the incident neutrino by the threshold energy of 1.804 MeV

$$E_{\overline{\nu}} = E_{e^+} + 1.804 \text{ MeV} + \mathrm{O}(E_{\overline{\nu}}/M_n). \tag{7.57}$$

Because of its comparatively large mass, the neutron only absorbs a very small fraction of the energy. These cells are stacked together in groups of six in five layers, alternating with four neutrino counters.

The detection of the neutrons involves two steps. The neutrons, with an energy of a few keV, are decelerated by collisions with the protons of the scintillator within 10 μs to thermal energies ($\sim$ 10 meV), and diffuse into an adjacent ^{3}He multiwire proportional counter (typical diffusion times have order

of magnitude 100 μs). There, they are detected via the reaction

$$n + {}^3\mathrm{He} \rightarrow p + {}^3\mathrm{H} + 765\ \mathrm{keV}. \tag{7.58}$$

The cross section for this (n,p) reaction is around 5500 b. Ultimately, the two charged reaction products, the proton and the triton, generate an ionization in the multiwire chamber. A $\bar{\nu}$-induced event is manifest via the signal of a positron in a scintillation cell, followed by a signal delayed by the diffusion time in one of the adjacent ^{3}He counters. In addition, to suppress the background, one may also require spatial coincidence of the positron and the neutron, since these two events are separated by a distance corresponding to at most the neutron diffusion time ($\sim$ 24 cm).

The background induced by cosmic radiation represents one of the greatest problems in these experiments. The concrete bunker essentially serves to attenuate the nucleonic component. A 15 cm thick steel screening suppresses the ^{40}K γ-radiation coming from the concrete. A 20 cm thick water layer decelerates fast neutrons and a subsequent boron carbide plate (0.5 cm) absorbs the thermal neutrons[2].

In addition, the detector is surrounded by a system of liquid scintillators used as veto counters, to reduce the muon background. However, the neutron background induced by cosmic muons in the shielding is not recognized by the veto system. Neutrons may be scattered in the detector by protons and so simulate a neutrino event. While the recoil proton generates a signal in the scintillator the decelerated neutron is detected in a ^{3}He counter. However, e^+- and p-induced signals may be distinguished according to their pulse shape, as a result of their different ionization densities. The remaining background is measured when the reactor is switched off annually to change the fuel rods.

Figure 7.10 shows the recorded positron spectra, after removal of the background. The measured values may now be evaluated either using measured and calculated reactor spectra, or independently of the reactor spectra. Neither form of analysis provides evidence of the existence of neutrinos.

An oscillation hypothesis is characterized by the mass parameter Δm^2 and the mixing parameter $\sin^2 2\theta$. Thus, results of oscillation experiments (without a positive outcome) are represented in the form of contour plots, excluding areas in a plane spanned by these two parameters. Figure 7.11 shows the results obtained by the Gösgen experiment. The contour lines pass through pairs of values which can be ruled out with 90% (or 68%) confidence. The area to the right of this is the excluded area, the parameter combinations to the left of the curves are still allowed. Figure 7.11(a) is based on the relative analysis of the data for the different detector distances without using the neutrino source spectrum. Neglect of the additional information about the

[2] Because of its strong shielding, the reactor itself is not a background source.

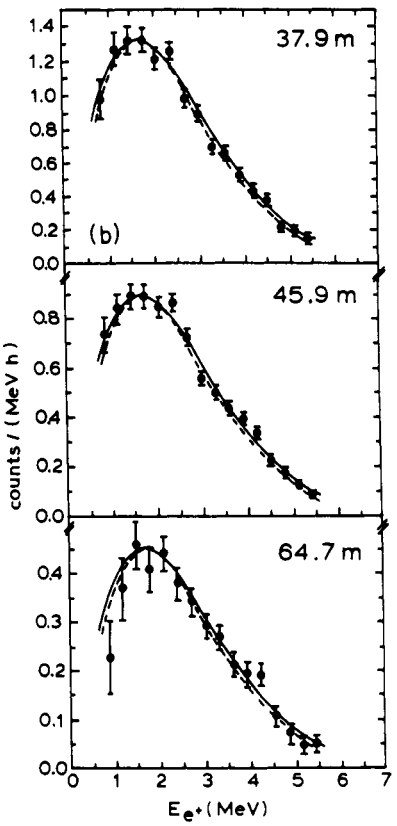


Figure 7.10. Measured positron spectra after subtraction of the background in comparison with predicted spectra. The error bands shown are of statistical nature. The continuous lines represent the positron spectra calculated using the fitted antineutrino spectrum (that is, independently of additional assumptions about the form of the reactor neutrino spectrum); the prediction shown by a broken line was derived using theoretical calculations about the reactor neutrino spectrum. In both cases, it was assumed that no oscillations occur (from [Zac86a]).

absolute reactor-neutrino flux and consideration of only the spectral information from the three measurements results in a smaller excluded area; however, the evaluation is essentially independent of uncertainties in the expected spectrum and of the detector detection probability. Figure 7.11(*b*) is obtained from the same measured values, making use of a theoretical neutrino spectrum for the reactor core.

Oscillations with very small mixing and/or mass parameters are still possible

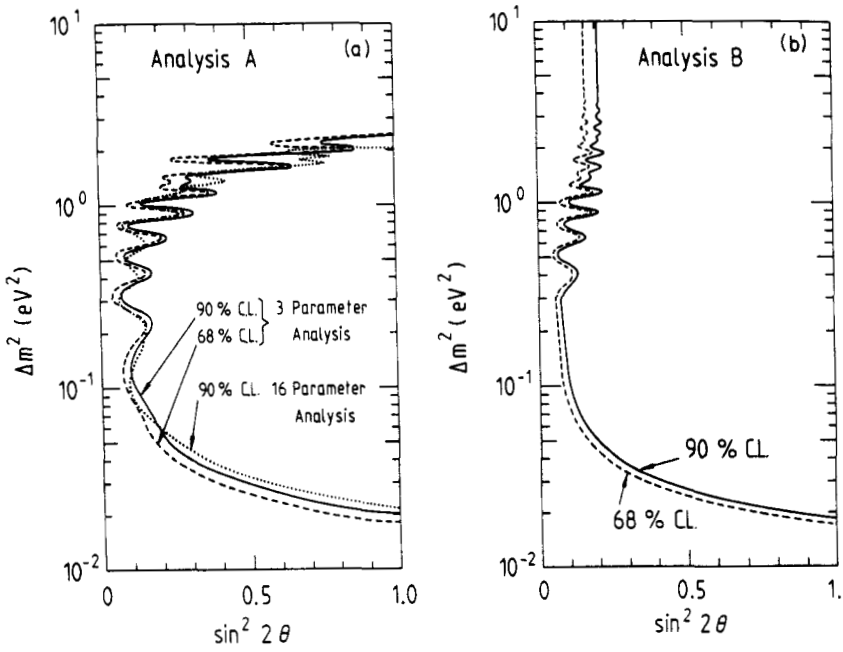


Figure 7.11. Exclusion region for the oscillation parameters Δm^2 and $\sin^2 2\theta$ for the Gösgen experiment [Zac86]. The parameter combinations to the right of the contour lines are excluded by the experiment. (*a*) Analysis A is based solely on data from the reactor experiment; (*b*) analysis B uses additional information about the neutrino spectrum of a reactor.

according to these results. For a mass parameter $\Delta m^2 < 0.01$ eV^2, the oscillations are too slow, even in the case of maximum mixing, to have a perceptible effect on the intensity. For large Δm^2, the boundary of the excluded area in the case of using a theoretical neutrino-core spectrum runs along a line $\sin^2 2\theta$ = constant, since the modulation of the intensity is washed out because of the limited energy resolution of the detector. Transitions to other neutrino flavours can then only manifest themselves through an overall decrease in the flux which is determined solely by $\sin^2 2\theta$. As previously mentioned, when average values are measured the information about the mass is lost.

Figure 7.12 shows exclusion diagrams in the (Δm^2–$\sin^2\theta$) plane for more recent reactor experiments (from [Dec95]). For earlier experiments see [Zac86c, Gro89, 90]. The Bugey III experiment excludes to a large extent the range of possible ν_e–ν_μ oscillations deduced from the Kamioka experiment on atmospheric neutrinos [Hir92a, Fuk94] (see section 7.3.9).

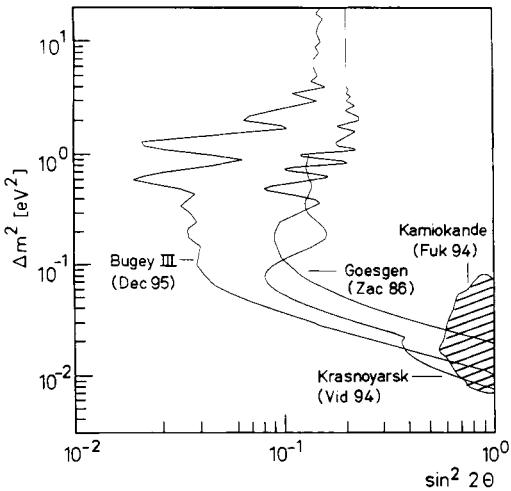


Figure 7.12. Exclusion diagram in the (Δm^2–$\sin^2 2\theta$) plane for the most recent reactor experiments carried out (after [Dec95]). Excluded are areas to the right of the full curves. The shaded area corresponds to an allowed range for ν_e–ν_μ oscillations, according to the Kamiokande experiment [Fuk94] (see section 7.3.9).

There is also another way of studying the phenomenon of neutrino oscillations using reactor neutrinos. If deuterium nuclei are exposed to the $\overline{\nu}_e$ flux of a nuclear reactor, two processes are possible

$$\overline{\nu}_e + d \rightarrow n + n + e^- \qquad \text{'charged current (cc)'} \qquad (7.59a)$$

$$\overline{\nu}_e + d \rightarrow n + p + \overline{\nu}_e \qquad \text{'neutral current (nc)'} \qquad (7.59b)$$

The reaction (7.59*a*) corresponds to the antineutrino capture by a proton (7.56) and is sensitive to oscillations. However, the threshold energy exceeds that for the reaction (7.56) by the binding energy of the deuteron of 2.226 MeV, i.e. $E_s = 4.03$ MeV. The process (7.59*b*), which is mediated by Z^0 exchange, occurs with the same probability for all neutrino flavours, and is therefore not sensitive to oscillations. In this case, the threshold corresponds to the binding energy of the deuteron.

For reactor neutrinos the expected reaction cross section for (7.59) averaged over the energy spectrum is (see e.g. [Boe87, 92])

$$\overline{\sigma}_{\overline{\nu}}(\text{cc}) = 1.2 \times 10^{-44}\ \text{cm}^2/\text{fission} \qquad (7.60a)$$

$$\overline{\sigma}_{\overline{\nu}}(\text{nc}) = 2.9 \times 10^{-44}\ \text{cm}^2/\text{fission}. \qquad (7.60b)$$

The antineutrino–deuterium scattering was studied at the reactor in Savannah River [Pas79, Rei80, 83]. A comparison of the expected ratio of cc to nc

reactions in the case in which there are no oscillations with the measured ratio gave a value of [Rei83]

$$R = 0.74 \pm 0.23. \tag{7.61}$$

This result is consistent with 1 to within one standard deviation. More recent measurements on the reactor in Rovno agree well with the theoretical predictions [Fay91] for R, for the case in which there are no oscillations.

Finally, it should be noted that there are plans for new long-baseline reactor neutrino oscillation experiments (San Onofre and Chooz) aiming at sensitivities down to $\Delta m^2 \approx 10^{-3}$ eV^2 and $\sin^2 2\theta \approx 0.1$ (see e.g. [Che94]).

7.3.3 Accelerator experiments

This book is primarily concerned with non-accelerator particle physics. Despite this, we now include a brief discussion of oscillation experiments on high-energy accelerators, also as example that accelerator and non-accelerator particle physics can be very nicely complementary.

Intensive ν_μ and $\overline{\nu}_\mu$ beams are available at high-energy accelerators. These are mainly generated from the decays of pions and kaons

$$\pi^+ \to \mu^+ + \nu_\mu \qquad K^+ \to \mu^+ + \nu_\mu \tag{7.62a}$$

$$\pi^- \to \mu^- + \overline{\nu}_\mu \qquad K^- \to \mu^- + \overline{\nu}_\mu \tag{7.62b}$$

where the lifetimes are $\tau_\pi = 2.6 \times 10^{-8}$ s and $\tau_K = 1.2 \times 10^{-8}$ s. These beams may be used in appearance experiments to detect the occurrence of a different neutrino flavour. If the energy of the primary beam is above the threshold for μ or τ production, transitions from ν_e to ν_μ or ν_τ may be studied with high sensitivity to the mixing angle. In contrast to disappearance experiments, this does not require a very precise knowledge of the source spectrum.

Figure 7.13 shows the typical structure of an accelerator experiment. A proton beam from a high-energy accelerator hits a target. The reaction products are primarily pions and kaons. These are then focused and enter an evacuated decay channel. There, the mesons decay, essentially according to (7.62). Since the mesons are highly relativistic, they transfer the focusing onto the emitted

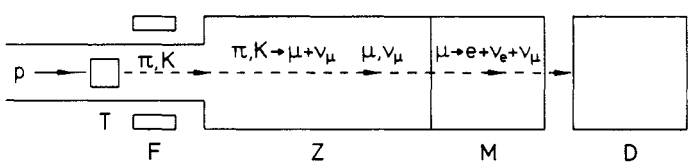


Figure 7.13. Typical accelerator experiment on neutrino oscillations: T = target, F = focusing, Z = decay tunnel, M = muon shield, D = detector.

neutrinos. The muons which are also created are stopped in the muon shield consisting of iron and concrete, where they decay via

$$\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu \qquad \mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu. \tag{7.63}$$

The lifetime for this process is $\tau_\mu = 2.2 \times 10^{-6}$ s. Since the velocity of the decaying muons is very small, the neutrinos from (7.63) are emitted practically isotropically in space. Thus, the impurity of the beam due to ν_e or $\overline{\nu}_e$ from (7.63) is greatly reduced. The beam purity is essentially limited by the following processes

$$K^+ \rightarrow e^+ + \nu_e + \pi^0 \qquad (B = 4.82 \pm 0.06\%) \tag{7.64a}$$

$$\pi^+ \rightarrow e^+ + \nu_e \qquad (B = (1.218 \pm 0.014) \times 10^{-4}) \tag{7.64b}$$

which occur with the given branching ratios B [Par90].

For high-energy accelerators, the decay (7.64*a*) (K_{e3} decay) is dominant. If one stays with the proton energies below the production threshold for K mesons, a very much greater beam purity is achievable. This is used on low-energy accelerators such as the LAMPF in Los Alamos (Los Alamos Meson Physics Facility) with a proton energy of 800 MeV (see [Wil80, Nem81, Wan84, Dom87]). The selection between muonic neutrinos and antineutrinos is made by a charge separation of the mesons before they enter the decay tubes.

So-called 'beam dump' experiments constitute another possible experimental possibility. Here, high-energy protons ($\sim$ 400 GeV) at CERN or Fermilab are stopped in a thick target (the beam dump). The secondary particles from the proton–nucleon collisions, also including hadrons with charm, may promptly decay into charged leptons and high-energy neutrinos. The semileptonic decay of the very short-lived hadrons with charm (D^0, $\overline{D^0}$, Λ_c, ...) also gives rise to electron neutrinos or antineutrinos. An example is the decay of the D^0 ($c\overline{u}$), with a rest mass of 1864.5 ± 0.5 MeV and an average lifetime of $\tau_D = (4.20 \pm 0.08) \times 10^{-13}$ s

$$D^0 \rightarrow K^- + e^+ + \nu_e \qquad (B = 3.4 \pm 0.4\%). \tag{7.65}$$

ν_τ and $\overline{\nu}_\tau$ may also be produced by semileptonic processes such as

$$D_s \rightarrow \tau + \nu_\tau. \tag{7.66}$$

Because of the high energies the phase spaces for the decays into muons and electrons are essentially equally large; thus, we expect approximately the same numbers of ν_e and ν_μ or $\overline{\nu}_e$ or $\overline{\nu}_\mu$ in these beam dump experiments. The neutrino components from the decays of the pions and kaons are also present. Because of the long lifetime of the light mesons, the latter lose a large part of

their energy as a result of the interaction in the target before the decay occurs; thus, the corresponding neutrinos have a comparatively small energy (< 1 GeV) and are easy to distinguish from the prompt high-energy neutrinos.

Beam dump experiments involve the measurement of the ratio of the ν_e rate to the ν_μ rate at distance x from the detector, where it is assumed that this ratio was equal to unity at the time of generation.

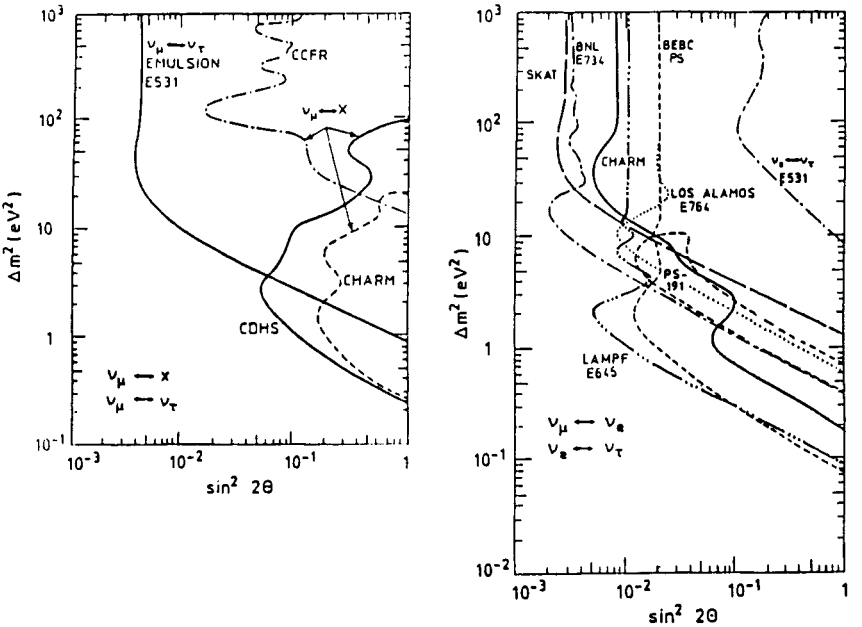


Figure 7.14. Exclusion regions for oscillation parameters from various accelerator experiments (from [Pes88]); left: ν_μ disappearance experiments and the ($\nu_\mu \leftrightarrow \nu_\tau$) channel; right: ($\nu_\mu \leftrightarrow \nu_e$) and ($\nu_e \leftrightarrow \nu_\tau$) channels. For more recent results see [Bor92, Win95].

Table 7.5 gives an overview of various appearance experiments on accelerators. A summary and short description can be found, for example, in [Pes88, Win95]. These measurements also establish very restrictive bounds for possible oscillation parameters (see figure 7.14). An experiment (experiment PS191) carried out at CERN's proton synchrotron (proton energy 600 MeV) gave an excessively large number of electron events in the reaction with a ν_μ beam, which was interpreted as neutrino oscillation with the parameters $\Delta m^2 = 5$ eV^2, $\sin^2 2\theta = 0.03 \pm 0.01$ [Ber86]. However, this result contradicts the measurement E734 on the AGS accelerator at the Brookhaven National Laboratory (BNL) [Ahr85].

Table 7.5. Results of oscillation experiments at accelerators.

Channel	Experiment	$(\Delta m^2)^a$ [eV^2]	$(\sin^2 2\theta)^b$
$\nu_\mu \rightarrow \nu_e$[d]	COL-BNL [Bak84]	< 0.6	$< 6 \times 10^{-3}$
	BNL-E734 [Ahr85]	< 0.43	$< 3.4 \times 10^{-3}$
	BEBC/PS [Ang86]	< 0.09	$< 1.3 \times 10^{-2}$
	PS191 [Ber86][c]	$= 5$	$= 0.03 \pm 0.01$
	LAMPF-E764 [Dom87]	< 0.67	$< 8 \times 10^{-3}$
$\overline{\nu}_\mu \rightarrow \overline{\nu}_e$[d]	FNAL [Tay83]	< 2.4	$< 1.3 \times 10^{-2}$
$\nu_\mu \rightarrow \nu_\tau$	FNAL [Ush81]	< 3	$< 1.33 \times 10^{-2}$
	FNAL-E531 [Gau86]	< 0.9	$< 4 \times 10^{-3}$
$\overline{\nu}_\mu \rightarrow \overline{\nu}_\tau$	FNAL [Asr81]	< 2.2	$< 4.4 \times 10^{-2}$
$\nu_e \rightarrow \nu_\tau$	COL-BNL [Bak84]	< 8	< 0.6
	FNAL-E531 [Gau86]	< 9	< 0.12

[a] For maximal mixing ($\sin^2 2\theta = 1$).

[b] For large Δm^2.

[c] Consistent with [Ast89].

[d] BNL-E776 [Bor92] yields $(\Delta m^2)^a \leq 0.075$ eV^2, $(\sin^2 2\theta)^b \leq 0.003$.

The CERN PS191 experiment was repeated at BNL with an improved apparatus and a larger beam intensity. This again gave an excess of electron neutrinos which is consistent with the original result [Ast89], but the statistical significance is far too small to interpret the result unambiguously as neutrino oscillations. Other possible explanations for the ν_e excess are discussed, for example, in [Ast89]. Some indications for $\nu_\mu \rightarrow \nu_e$ oscillations have recently been discussed from an experiment at LAMPF (Los Alamos) using a Liquid Scintillator Neutrino Detector (LSND) [Lou95, Smi95]. The experiments CHORUS and NOMAD being carried out at CERN at present will look for ν_μ–ν_τ oscillations at considerably smaller mixing angles than earlier experiments (figure 7.14) (see [Win95, Lee95a]). They are motivated mainly by a hypothetical MSW solution to the solar neutrino problem (see section 7.3.8).

In summary, it can be said, that no unambiguous evidence of the occurrence of neutrino oscillations has yet been found in reactor and accelerator experiments.

7.3.4 Experiments with solar neutrinos

The Sun forms another extremely interesting source of neutrinos. Because of the large distance between the Sun and the Earth, solar neutrinos are suitable for research into oscillations with mass parameters Δm^2 down to the region of

10^{-12} eV^2 (see table 7.1). In addition, the neutrinos generated by thermonuclear reactions in the interior of the Sun provide important information about the Sun itself, and, in particular, about the core. As a result of multiple scattering and absorption processes, photons take approximately 10^4 years to reach the surface. Consequently, on the Earth we only see photons which are essentially emitted on, or ultimately scattered by, the surface. Thus, these only provide information about the luminosity and the chemical composition in the Sun's shell. Neutrinos, on the other hand, are only subject to the weak interaction and leave the Sun practically unhindered after their generation. The observation of neutrinos thus constitutes a unique way of obtaining information about the interior of the Sun and in particular about the processes involved in the generation of energy. Experiments to detect the flux of solar neutrinos have been carried out on the Earth since 1965.

The next section is concerned with the generation of energy and the production of neutrinos in the Sun; we restrict the discussion here to the so-called Standard Solar Model (SSM).

7.3.4.1 Generation of energy and neutrino production in the Sun

The Standard Solar Model (SSM) [Bah82, 88, 89, 92, Ber93, Tur88, 93a,b] is based on the assumption that the structure and the temporal evolution of a star are unambiguously determined by its mass and chemical composition. The chemical composition is usually given in mass percentages (X for hydrogen, Y for helium and Z for the so-called metals, i.e. all elements heavier than helium). The concepts underlying this model, which may be traced back to Eddington [Edd26], may be summarized as follows:

- Hydrostatic equilibrium. At every point of the star the pressure gradient and the gravitational pressure are balanced. In most stars the pressure is essentially determined by the gas pressure; the radiation pressure need only be considered in very hot and massive objects.
- The pressure, density and temperature are related by the state equation for the ideal gas.
- There is thermal equilibrium, i.e. the generation of energy is equal to the emission.
- The energy transfer in the interior of the star takes place mainly via radiation. Convection plays a role in the outer zones.
- The energy is generated via fusion processes.

The energy which is released as a result of continuous thermonuclear reactions acts against the increasing compression due to gravitation so that an equilibrium between generated and emitted energy is established at certain density and temperature values. It is assumed that the Sun had a homogeneous structure

shortly after its generation, i.e. the chemical composition found in the Sun's atmosphere today should correspond to the abundances of elements in the Sun at its birth around 4.6 billion years ago. The principles underlying this model are highly plausible; for details of the calculation, we refer readers to the specialist literature [Bah82, 88, 89, 92, Tur93a,b]. Table 7.6 summarizes the most important properties of the Sun, according to the Standard Solar Model.

Table 7.6. Properties of the Sun, according to the Standard Solar Model [Bah88].

	$t = 4.6 \times 10^9$ y (today)	$t = 0$
Luminosity $L_\odot$	$\equiv 1$	0.71
Radius $R_\odot$	696000 km	605500 km
Surface temperature T_S	5773 K	5665 K
Central temperature T_c	15.6×10^6 K	–
Central density	148 g cm^{-3}	–
X(H)	34.1%	71%
Y(He)	63.9%	27.1%
Z	1.96%	1.96%

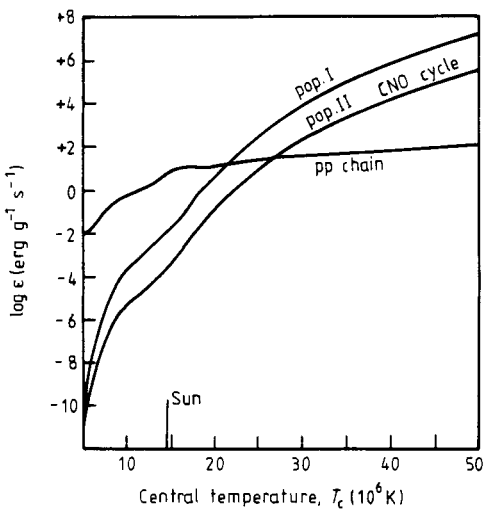


Figure 7.15. Evolution of energy ε for stars with various central temperatures. The central density is uniformly assumed to be $\varrho_c = 100$ g cm^{-3}. In cooler stars the *pp* chain dominates, in hotter stars the CNO cycle dominates (from [Uns81]).

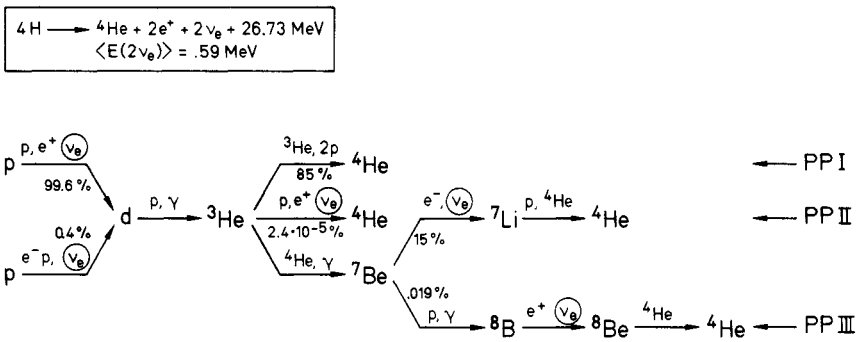


Figure 7.16. The *pp* cycle.

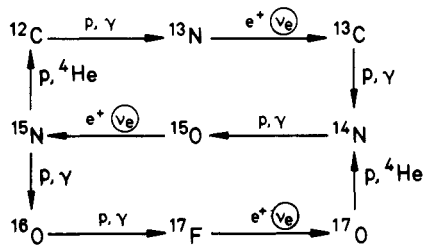


Figure 7.17. The CNO cycle.

The energy emitted from the Sun's surface is generated in the interior of the Sun by fusion of light atomic nuclei. With a central density of 148 g cm^{-3} and a temperature of 1.56×10^7 K, the light atoms are completely ionized and form a plasma. Relatively cold stars such as our Sun obtain their energy mainly from the so-called *pp* cycle, while the CNO cycle (also called the Bethe–Weizsäcker cycle) plays a relatively subordinate role (see figures 7.15–7.17).

The *pp* cycle begins with hydrogen nuclei which are gradually fused via deuterium nuclei into ^{4}He. The following reaction is by far the most important[3]

$$p + p \rightarrow d + e^+ + \nu_e + 0.42 \text{ MeV}. \tag{7.67}$$

Since the diproton is not stable, the transition into the deuteron occurs via a β decay process. This is a weak interaction process, which has a timescale of around 10^{10} years. Thus, the reaction (7.67) determines the temporal evolution of the subsequent (faster) fusion processes. This *pp* process generates neutrinos with a continuous spectrum up to 0.42 MeV.

The average lifetime of 10^{10} years of a proton in the Sun explains why the energy output by the Sun has remained constant for billions of years. Despite

[3] The energies given in these equations relate to nuclear masses, i.e. they are the difference in rest energies of the nucleons and nuclei on the two sides.

this extremely tiny reaction rate (10^{-10} per year corresponds to a reaction cross section of approximately 10^{-23} b at 1 MeV) approximately 10^9 reactions occur per second per cm^3, since the proton density is approximately 10^{26} cm^{-3}. In addition to the dominant reaction (7.67) with a branching ratio of 99.6% there is another process that leads to deuterium

$$p + p + e^- \rightarrow d + \nu_e + 1.44 \text{ MeV } (0.4\%). \tag{7.68a}$$

Here neutrinos are emitted with the fixed energy of 1.44 MeV. The deuterium formed fuses practically immediately into ^{3}He

$$d + p \rightarrow {}^3\text{He} + \gamma + 5.49 \text{ MeV}. \tag{7.68b}$$

The ^{3}He thus created passes through further reaction steps involving the formation of the particularly stable ^{4}He:

$$\text{(I)} \qquad {}^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He} + 2p + 12.86 \text{ MeV } (85\%) \tag{7.69}$$

$$\text{(II)} \qquad {}^3\text{He} + p \rightarrow {}^4\text{He} + e^+ + \nu_e + 18.77 \text{ MeV } (2.4 \times 10^{-5}\%) \tag{7.70}$$

$$\text{(III)} \qquad {}^3\text{He} + {}^4\text{He} \rightarrow {}^7\text{Be} + \gamma + 1.59 \text{ MeV}. \tag{7.71a}$$

The ^{7}Be is also converted into ^{4}He by two different routes:

$$\begin{aligned} \text{(IIIb1)} \qquad {}^7\text{Be} + e^- &\rightarrow {}^7\text{Li} + \nu_e + 0.8617 \text{ MeV } (15\%) \\ {}^7\text{Li} + p &\rightarrow {}^4\text{He} + {}^4\text{He} + 17.35 \text{ MeV} \end{aligned} \tag{7.71b}$$

or

$$\begin{aligned} \text{(IIIb2)} \qquad {}^7\text{Be} + p &\rightarrow {}^8\text{B} + \gamma + 0.14 \text{ MeV } (0.019\%) \\ {}^8\text{B} &\rightarrow {}^8\text{Be} + e^+ + \nu_e + 14.6 \text{ MeV} \\ {}^8\text{Be} &\rightarrow {}^4\text{He} + {}^4\text{He} + 3 \text{ MeV}. \end{aligned} \tag{7.71c}$$

The *pp* cycle is illustrated graphically in figure 7.16. The net effect is the conversion of four protons into an α particle

$$2e^- + 4p \rightarrow {}^4\text{He} + 2\nu_e + 26.7 \text{ MeV}. \tag{7.72}$$

The gain in energy is 26.7 MeV, the average energy carried away by the neutrinos amounts to a few per cent of this ($\langle E(2\nu) \rangle \sim 0.6$ MeV). The expected neutrino flux may be estimated using (7.72). On average, one neutrino is generated per 13 MeV thermal energy generated in the Sun. Based on the assumption of the Standard Solar Model that the Sun has been in thermal equilibrium for a long time, the energy production is equivalent to the energy emitted. The solar constant, i.e. the flux of energy which reaches the Earth, is approximately $S = 0.13$ J cm^{-2} s^{-1} $\approx 8 \times 10^{11}$ MeV cm^{-2} s^{-1}. Thus, the neutrino flux is calculated to be

$$\Phi_\nu = \frac{S}{13 \text{ MeV}} \sim 6 \times 10^{10} \text{ cm}^{-2} \text{ s}^{-1}. \tag{7.73}$$

Table 7.7. SSM predictions for the flux of solar neutrinos on the Earth [Bah88, 89].

Source	Φ_ν [10^{10} cm^{-2} s^{-1}]
pp	$6.0 \pm 2\%$
pep	$0.014 \pm 5\%$
^{7}Be	$0.47 \pm 0.15\%$
^{8}B	$5.8 \times 10^{-4} \pm 37\%$
^{13}N	$0.06 \pm 50\%$
^{15}O	$0.05 \pm 58\%$
^{17}F	$5.2 \times 10^{-4} \pm 46\%$

In addition to the *pp* cycle, there is another process which contributes approximately 1.6% of the Sun's energy production, namely the CNO (or Bethe–Weizsäcker) cycle. This plays a subordinate role at the low temperature of our Sun, but constitutes the most important energy source in heavy hot stars (see figure 7.15). In the CNO cycle ^{12}C acts as a catalyst for the burning of hydrogen into helium (figure 7.17).

Both cycles include various reactions involving the emission of neutrinos with either continuous or discrete energy distributions. We summarize these below. The most important ν generating reactions of the *pp* chain are

$$\begin{aligned}
&p + p \rightarrow d + e^+ + \nu_e && (E_\nu \leq 0.420 \text{ MeV}) \\
&p + e^- + p \rightarrow d + \nu_e && (E_\nu = 1.442 \text{ MeV}) \\
&^3\text{He} + p \rightarrow {}^4\text{He} + e^+ + \nu_e && (E_\nu \leq 18.77 \text{ MeV}) \\
&^7\text{Be} + e^- \rightarrow {}^7\text{Li} + \nu_e && (E_\nu = 0.862 \text{ MeV (90\%)}, \ 0.384 \text{ MeV (10\%)}) \\
&^8\text{B} \rightarrow {}^7\text{Be}^* + e^+ + \nu_e && (E_\nu \leq 14.6 \text{ MeV}).
\end{aligned} \tag{7.74}$$

The important reactions in the CNO cycle, as far as the production of solar neutrinos is concerned, are

$$\begin{aligned}
&^{13}\text{N} \rightarrow {}^{13}\text{C} + e^+ + \nu_e \quad (E_\nu \leq 1.199 \text{ MeV}) \\
&^{15}\text{O} \rightarrow {}^{15}\text{N} + e^+ + \nu_e \quad (E_\nu \leq 1.732 \text{ MeV}) \\
&^{17}\text{F} \rightarrow {}^{17}\text{O} + e^+ + \nu_e \quad (E_\nu \leq 1.740 \text{ MeV}).
\end{aligned} \tag{7.75}$$

Measurements of the flux of the ^{17}F neutrinos would be particularly interesting, since this is a measure of the oxygen content of the interior of the Sun. However, because of the very low reaction rates, the reactions (7.75) are difficult to detect.

The fundamental process (7.67) generates the most neutrinos, although these have relatively low energies; on the other hand, the high-energy ^{8}B neutrinos

are only formed in a rare side branch of the *pp* cycle. The calculated fluxes of solar neutrinos on the Earth are summarized in table 7.7.

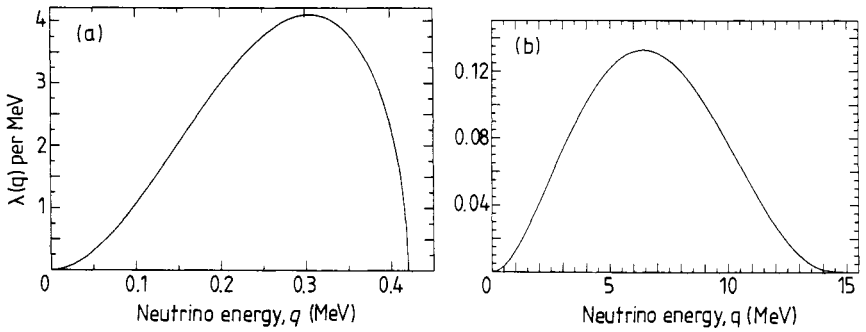


Figure 7.18. Energy spectra of (*a*) the solar *pp* and (*b*) ^{8}B neutrinos (from [Bah89]).

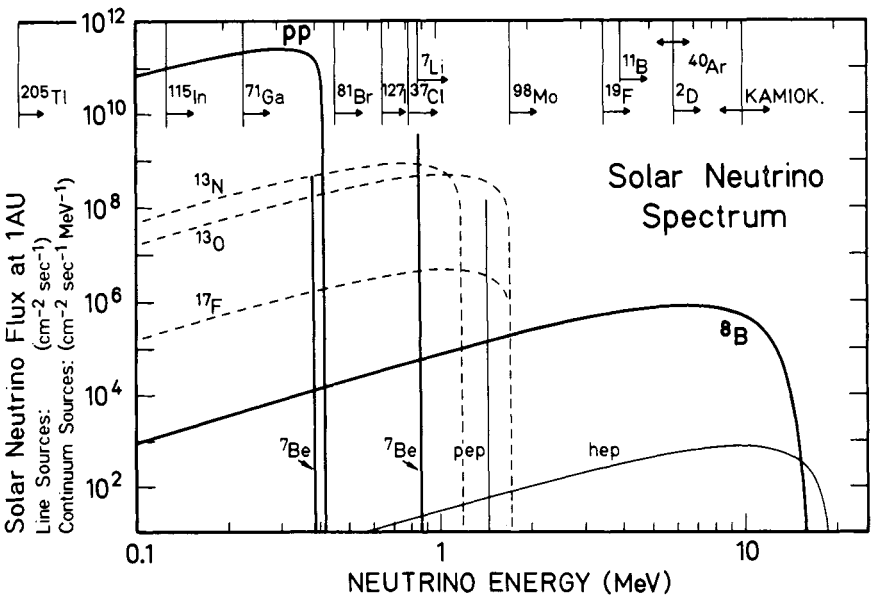


Figure 7.19. The solar neutrino spectrum according to the Standard Solar Model. The figure also shows the response regions of different detectors (after [Ham93, 94]).

The cross sections of the individual reactions in the Sun are energy dependent, so that the production rates are determined by the ambient temperature in the interior of the Sun. The flux of the *pp* neutrinos is largely determined by the luminosity of the Sun, which is well known from

measurements; the dependence of the production rate on the central temperature of the Sun is relatively weak ($\sim T^4$). The ^{8}B neutrinos come from a side branch which is less important for the energy production. The formation of the ^{8}B involves several fusion processes in which nuclei with a higher nuclear charge ($Z \geq 2$) also participate. Thus, the production rate for these neutrinos exhibits a very sensitive dependence on the temperature ($\sim T^{19}$). Measurements of the relative abundances of the individual neutrino components would thus provide a sensitive thermometer for the Sun's interior.

Here, we note that the statement that the *pp* neutrino flux is independent of the Sun model is only true under certain conditions, since, for example, at a somewhat higher temperature, the majority of the neutrinos would come from the CNO cycle rather than from the *pp* cycle (see the discussion in [Bah89]).

In addition to the absolute neutrino flux, a knowledge of the neutrino spectrum, i.e. the flux as function of the energy, is of great interest, since different energy ranges of the spectrum correspond to different neutrino sources. Because of their different energy thresholds, different detectors are sensitive to different energy regions and thus detect different components. Figure 7.18 shows the continuous neutrino spectra of the two most important reactions, $pp \rightarrow de^+\nu_e$ and $^8\mathrm{B} \rightarrow {}^8\mathrm{Be}^*e^+\nu_e$. The solar neutrino spectrum resulting from the SSM is shown in figure 7.19.

7.3.4.2 The detection of solar neutrinos, the chlorine experiment and the Kamioka experiment

An ideal detector for solar neutrinos should be able to determine the time of the interaction together with the energy and the direction of flight of the incident neutrino. However, these requirements can only be partially satisfied.

One way of detecting neutrinos on the Earth involves the use of radiochemical detectors. These are based on the capture of neutrinos in individual atomic nuclei of a suitable element. The radioactive daughter nuclei formed in this way are detected via their decay. For this reason, for a certain period of time, which depends on the half-life of the reaction product, the detector material is exposed to solar neutrinos. The measurement period is determined by the onset of equilibrium between production and decay of the radioactive isotope. The few nuclei which are converted by neutrino capture must then be extracted from a very large reservoir of target nuclei and detected.

No individual events are detected in this type of detector, which instead involves integration over the exposure time. In addition, only neutrinos with energy above a certain threshold are recorded. The measured signal is a measure of the neutrino spectrum integrated over the whole energy region from the threshold upwards, weighted with the energy-dependent reaction cross section for the capture in the target nucleus.

R Davis's famous chlorine experiment in operation since 1968 [Dav68, 84, 87a, 88, 94, 94a] uses such a radiochemical detector. The neutrino is detected via the neutrino capture in ^{37}Cl nuclei with the formation of radioactive ^{37}Ar

$$^{37}\mathrm{Cl} + \nu_e \rightarrow {}^{37}\mathrm{Ar} + e^- \qquad E_t = 814\ \mathrm{keV}. \tag{7.76}$$

This reaction has a threshold energy of 814 keV. This means that only neutrinos with relatively high energies can be detected; in particular, the chlorine detector does not respond to the *pp* neutrinos (see figure 7.19). The argon formed is radioactive. ^{37}Ar decays back to ^{37}Cl with a half-life of 35.04 days via electron capture. This transition can be detected using the 2.82 keV Auger electrons which follow the electron capture. The neutrino flux to which the target was exposed is then determined from the amount of argon generated.

The detector consists of a large tank containing 380 000 litres of perchloroethylene (C_2Cl_4) weighing 615 t. The tank was installed in the Homestake gold mine in Lead, South Dakota, at a depth of around 1400 m below the Earth's surface (figure 7.20). The shielding thickness corresponds to that of 4100 m of water. The nuclide ^{37}Cl occurs in the natural isotopic distribution with a relative abundance of 24.23%. Thus, the tank contains 2.2×10^{30} ^{37}Cl nuclei as a target for the neutrinos ($\simeq$ 133 tonnes).

Chlorine was used since the combination of its chemical and physical properties is well suited for the construction of a large radiochemical neutrino detector. The threshold energy is relatively low and the transition to the ground state of ^{37}Ar has a favourable reaction cross section, i.e. the $ft_{1/2}$ value (see below) is small ($\log ft = 5$). In addition, transitions to excited states are also possible, which again increases the rate considerably.

Moreover, chlorine is comparatively inexpensive, so that it is possible to build a detector with a mass of several hundred tonnes (the liquid perchloroethylene is used in chemical dry cleaning). The reaction product ^{37}Ar is a rare gas, and thus chemically inert, so that it is easy to flush out. The nucleus ^{37}Ar generated by neutrino capture has a recoil energy which is sufficient for it to leave the molecular bond. Argon dissolves without entering into complex chemical relationships, so that the few atoms can be driven out by rinsing with helium. One other important point is that ^{37}Ar has a reasonable half-life of 35 days.

Every few months (corresponding to an exposure time of several half-lives) the argon is rinsed out and detected in a proportional counter. The electron capture leads to an excited ^{37}Cl atom which emits either X-rays or Auger electrons. In approximately 90% of all decays, electrons are released from the K shell with an energy of 2.82 keV; these are counted in the detector. Since only a few decays are expected per measurement period, very high requirements are placed on the extraction and detection techniques. In addition, the background has to be reduced as far as possible. For this, in addition to the usual shielding

measures and an NaI veto counter, the rise times of the pulses are also measured in counting tubes. We refer readers to the literature for further details of the experiments [Dav68, 84, 94, Row85, Bah89].

What neutrino capture rates should we expect for the chlorine detector, according to the Standard Solar Model? To answer this question, we need to know the energy-dependent reaction cross section for the reaction (7.76) in addition to the solar neutrino spectrum discussed above. For the transition between the ground states, the matrix element can be determined from the ft value of the β decay of the daughter nucleus. In addition, there are contributions from excited states which give rise to some uncertainty. These matrix elements

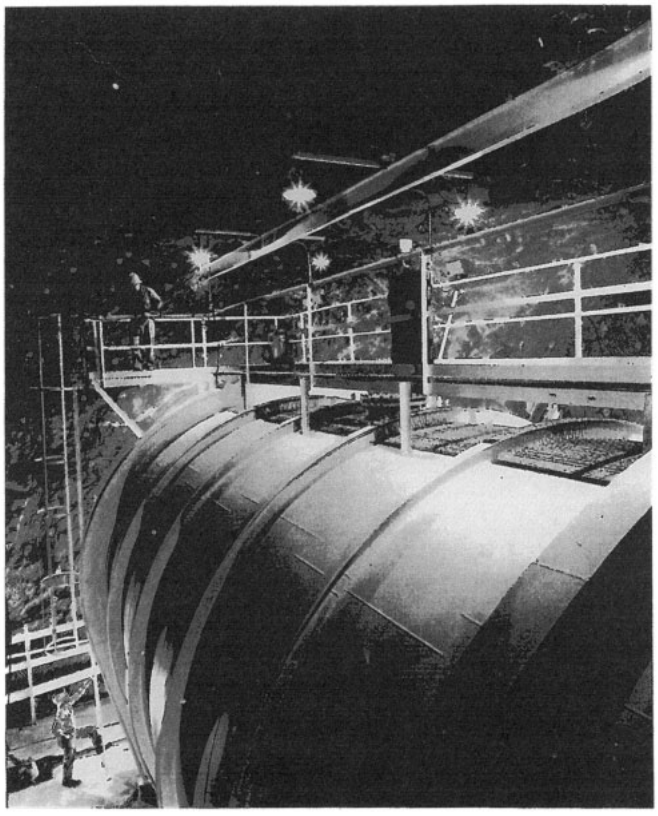


Figure 7.20. The detector of Davis's chlorine experiment to detect solar neutrinos in the approximately 1400 m deep Homestake mine in Lead, South Dakota, USA, around 1967. The tank filled with 380 000 litres of perchloroethylene is shown. R Davis can be seen at the top (courtesy of Brookhaven National Laboratory).

may be obtained from measurements of (p, n) reactions; alternatively, they may be determined in nuclear structure calculations (theoretical approaches to the calculation of the reaction cross section may be found e.g. in [Gro86a, 89, 90, Sta89]). In the case of ^{37}Cl, this does not lead to substantial uncertainties in the predictions since the corresponding matrix elements may be deduced from the β^+ decay of ^{37}Ca [Sex74]. Larger uncertainties originate certainly from the opacity of the Sun [Ber93], which also enters into the calculation.

The standard model gives a capture rate of

$$\sum_i \Phi_i \sigma_i = 7.9 \pm 2.6 \text{ SNU} \quad \text{[Bah92]} \quad (7.77a)$$

$$= 6.4 \pm 1.4 \text{ SNU} \quad \text{[Tur93b]} \quad (7.77b)$$

$$= 7.43 \pm 2.7 \text{ SNU} \quad \text{[Ber93]} \quad (7.77c)$$

for the chlorine detector. Since the expected rates are very small, the solar neutrino unit (SNU) was introduced. 1 SNU means one neutrino capture per 10^{36} target atoms per second. The capture rates for the individual neutrino sources are given in table 7.8.

Table 7.8. SSM predictions for the capture rates in the chlorine detector [Bah88].

Source	$\Phi_i \sigma_i$ [SNU]
pp	0
pep	0.2
^{3}He *p*	0.03
^{7}Be	1.1
^{8}B	6.1
^{13}N	0.1
^{15}O	0.3
^{17}F	0.003
$\sum_i$	7.9

The chlorine experiment is mainly sensitive to ^{8}B neutrinos which contribute 77% of the event rate. This is essentially because the energy of the ^{8}B neutrino is sufficient to excite the isobaric analogue state in ^{37}Ar ($\sim$ 5 MeV). The ^{7}Be neutrinos provide a further contribution of 14%.

From 1970 to 1984, 61 measurement runs were undertaken [Row85] (see also [Bah89]). The production rate amounted to 0.462 ± 0.040 ^{37}Ar atoms per day. The background due to cosmic radiation amounted to 0.08 ± 0.03 per day. From the number of ^{37}Cl atoms in the tank, it follows that a production rate of

one atom per day corresponds to 5.35 SNU. This gives a measured capture rate of

$$\sum_i \Phi_i \sigma_i = 2.1 \pm 0.3 \text{ SNU } (3\sigma) \text{ (1970–1984).} \tag{7.78}$$

The discrepancy between the measured value and that expected according to the standard model is generally referred to as the solar neutrino problem. One should, however, keep in mind, that the *predicted* capture rates underwent a considerable development with time (see figure 7.21).

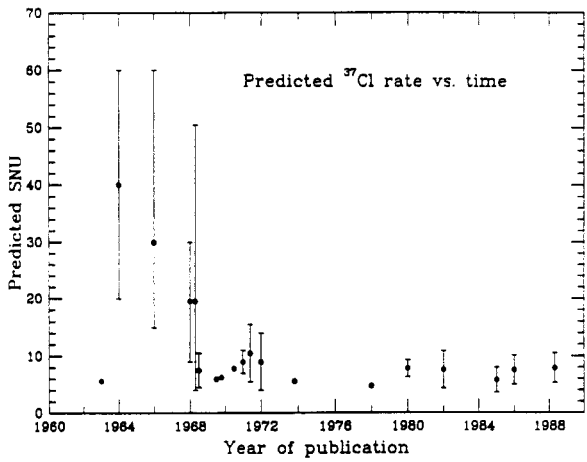


Figure 7.21. Variation of the predicted value of the neutrino flux from the Sun from the decay of ^{8}B over the last 30 years (error bands correspond to 1σ) (from [Bah89, Dav92b]).

Measurements were broken off in 1985 and started again in 1986. Surprisingly, the first 10 new measurement runs (#90–99, October 1986–April 1988) gave an average capture rate of [Bah89]

$$\sum_i \Phi_i \sigma_i = 3.6 \pm 0.7 \text{ SNU } (1986\text{–}1988) \tag{7.79}$$

which is two standard deviations larger than the average rate up to 1984. We note that the new measurements were carried out during a period of little solar activity. The average value for all the data up to April 1988 is

$$\sum_i \Phi_i \sigma_i = 2.2 \pm 0.3 \text{ SNU } (1970\text{–}1988). \tag{7.80}$$

The measurements were continued [Dav94, 94a]. Figure 7.22 shows the ^{37}Ar production rate over the measurement period from March 1970 to April 1988. Table 7.9 summarizes the resulting data.

Table 7.9. Summary of the ^{37}Ar production rates in the Davis experiment. The table also shows the results over the period during which data was recorded by the Kamiokande II detector (from [Bei91]). The rates are given in atoms per day and 615 tonnes of C_2Cl_4.

Period of observation	March 1970–April 1989	August 1986–April 1989
^{37}Ar rate	0.518 ± 0.036	0.87 ± 0.13
Background	0.08 ± 0.03	0.08 ± 0.03
Corrected ^{37}Ar rate	0.438 ± 0.047	0.79 ± 0.13
^{37}Ar rate [SNU]	2.33 ± 0.25	4.2 ± 0.7

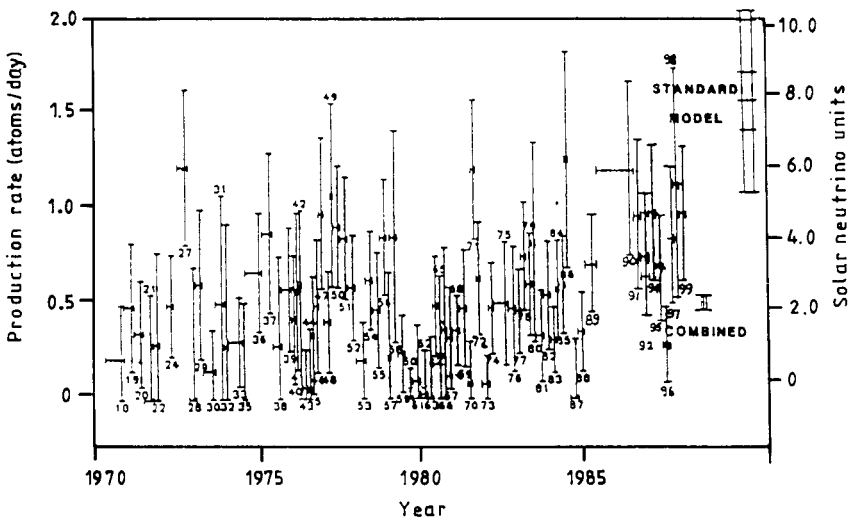


Figure 7.22. Event rate for the Homestake ^{37}Cl detector since 1970 (from [Dav88]). For further data see [Dav94, 94a].

The data in figure 7.22 are compatible with the assumption of a time-independent flux. On the other hand, it is clear that the rates in the measurement periods 1979–1980 and 1988–1989 are systematically smaller than the average value, while each of the previous periods gave relatively large neutrino fluxes. Consequently, an anticorrelation between these temporal variations and the 11-year cycle of solar activity has been discussed (see figure 7.32, for a possible interpretation see section 7.3.6). On the other hand it was pointed out [Baz84, Ale87, Dav87a, 94, 94a], that some of the measurement periods with particularly high neutrino flux (runs 27, 51, 71, 86, 117) coincide in time with the occurrence

of large solar flares (see figure 2 in [Dav87a]). Neutrinos might be produced during the solar flares by pp interactions of protons (accelerated to > 10 GeV in the flares) in the exterior parts of the Sun, and the decay of pions and muons produced in such reactions. Another possibility could be thermonuclear reactions in the interior of the Sun, whose external phenomena might be the solar flares. For discussions and attempts to test these hypotheses by other underground detectors, see [Ale87, Agl91, Kri94].

Before we discuss possible solutions to the solar neutrino problem arising from the Davis experiment, we shall describe another experiment which has confirmed the neutrino deficit, namely the Cerenkov water counter of the Kamiokande collaboration, which was originally constructed to detect proton decay [Hir87, 90a,b, Kos92]. Unlike radiochemical detectors, this is a real-time experiment. The detection of neutrinos is based on the neutrino–electron scattering

$$\nu + e^- \rightarrow \nu' + e' \tag{7.81}$$

in water as target and detector material. The recoil electrons are detected via the Cerenkov light to which they give rise, using photomultipliers. The Kamiokande detector is described in chapter 4, so we shall only summarize the most important data here. It consists of 3000 tonnes of water, of which only the inner 680 tonnes are used as a sensitive volume (fiducial volume) for measurement (number of electrons = 2.27×10^{32}). The remaining water is used to suppress the background. The giant steel tank is installed at a depth of 1000 m (2700 m water equivalent) in the Kamioka mine 300 km to the west of Tokyo.

Neutrino–electron scattering experiments have several experimental advantages over radiochemical detectors:

- It is possible to determine a direction. The recoil electrons are preferentially scattered in the forward direction, i.e. in the direction of flight of the neutrinos. It is thus possible to determine whether the neutrinos come from the direction of the Sun from the reconstructed electron track.
- Scattering experiments give the arrival time of individual events. There is no integration over long exposure times, so that temporal variations of the neutrino flux may be studied.
- While radiochemical detectors can only detect ν_e, the νe^- scattering process covers all flavours. However, the reaction cross sections for ν_x, $x \neq e$ are much smaller than for ν_e. This is because both Z^0 and $W^{\pm}$ exchange play a role in $\nu_e e$ scattering, while only the Z^0 term contributes to $\nu_\mu e$ or $\nu_\tau e$ (see figure 7.24(a)). The differential reaction cross section for the creation of a recoil electron with kinetic energy T via scattering with a neutrino of

energy E_ν is given by [t'Ho71]

$$\frac{d\sigma}{dT} = \sigma_e \left[g_L^2 + g_R^2 \left(1 - \frac{T}{E_\nu}\right)^2 - g_L g_R \left(\frac{m_e c^2 T}{E_\nu^2}\right) \right] \tag{7.82}$$

where

$$\sigma = \frac{2G_F^2 m_e}{\pi \hbar^4}. \tag{7.83}$$

According to the standard model, the coupling constants are given by

$$g_R = \sin^2 \theta_W \qquad g_L = \sin^2 \theta_W \pm \tfrac{1}{2} \tag{7.84}$$

where the positive sign applies for the electron neutrino and the negative sign for the muon and tau neutrino. The total reaction cross sections $\sigma_{\nu_e e}$, $\sigma_{\nu_\mu e}$, ... are obtained by integration over the energy spectrum of the solar neutrinos. A comparison of absorption and scattering experiments could in principle provide information about the flavour of the neutrinos involved.

- The energy distribution of the electrons to a certain extent reflects the energy spectrum of the neutrinos.

One large disadvantage of the Cerenkov counter is the high energy threshold. This is necessary because the background rate at low energies is too large. Typically, the threshold amounts to 8 MeV. This means that the Kamiokande detector too is essentially only sensitive to the flux of ^{8}B neutrinos. At this threshold energy, the ratio of $\sigma_{\nu_e e}$ to $\sigma_{\nu_\mu e}$ is approximately $\sim$ 6–7.

The expected ^{8}B neutrino flux according to the standard model amounts to $\Phi(^8\text{B}) = 6.0 \times 10^6$ cm^{-2} s^{-1} on the Earth's surface. The data obtained after a measurement period of 1040 days are shown in figure 7.23(*a*) (450 days with a threshold energy $E_t = 9.3$ MeV and 590 days with $E_t = 7.5$ MeV) [Hir90a]. The measured flux normalized to the expectation from the SSM over the period from January 1987 to April 1990 amounts to

$$\frac{\Phi(^8\text{B})}{\Phi(^8\text{B})_{\text{SSM}}} = 0.46 \pm 0.05 \text{ (stat)} \pm 0.06 \text{ (syst)}. \tag{7.85}$$

These data thus also seem to show a clear deviation from the value expected according to the SSM and qualitatively confirm the deficit of solar neutrinos found in the chlorine experiment. However, we stress that both experiments essentially measure only high-energy ^{8}B neutrinos which stem from a rare side branch of the *pp* cycle.

Since the direction of the neutrinos is correlated with that of the electrons recorded in the Cerenkov counter, it was possible to measure the angular distribution of the events in relation to the direction of the Sun (see

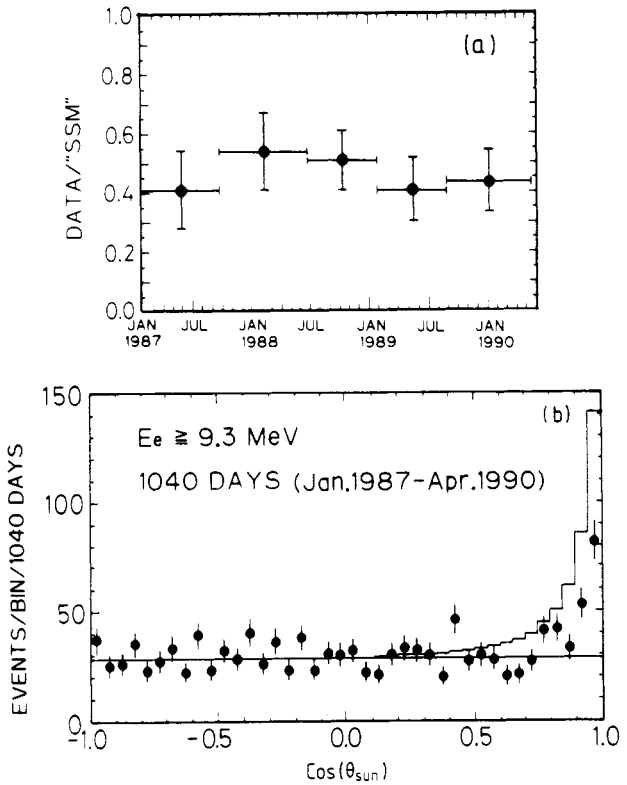


Figure 7.23. (*a*) Time variation of the ^{8}B neutrino signal in the Kamiokande II detector. The measurements are normalized to the Standard Solar Model (SSM). For the first two measurements the threshold was $E_e \geq 9.3$ MeV, for the last three $E_e \geq 7.5$ MeV; (*b*) angular distribution of the events in the Kamiokande II detector, relative to the direction of the Sun (from [Hir90a]).

figure 7.23(*b*)). The given histogram shows the expected distribution according to the SSM. The increase in the forward direction indicates that the Sun is in fact the source of the neutrinos.

Figure 7.23(*a*) does not show any statistically significant temporal variation of the flux over a period of approximately three years, although the solar activity over this period varied by approximately an order of magnitude.

7.3.4.3 Attempts to understand the solar neutrino problem

The deficit of solar neutrinos observed in the chlorine experiment and confirmed by Kamiokande has preoccupied astro-, nuclear and particle physicists for a

number of years. There are essentially two approaches to solving this problem: either the Solar Model is in some way false ($\Rightarrow$ Non-standard Solar Model) or our knowledge of the properties of the neutrino is still incomplete ($\Rightarrow$ MSW effect, magnetic moment of the neutrino).

Non-standard solar models As long as the solar *pp* neutrinos, which represent the greatest part of the solar ν flux and are essentially linked to the known luminosity of the Sun, were undetected, there was no secure basis on which the SSM could be founded. Attempts were then made to alter the SSM so that the ^{8}B flux was reduced. As previously discussed in section 7.3.4.1, the flux of ^{8}B neutrinos depends in a very sensitive manner on the central temperature of the Sun. A fall of only 6% from 15.6 to 14.6 million degrees would already be sufficient to explain the given discrepancy. Thus, most non-standard solar models are based on a reduction of the central temperature and include the luminosity of the Sun as a boundary condition [Bah71, 82b, Boy85, Fau85, Sch85b]. Other sharp boundary conditions for non-standard models come from helioseismology [Bah88, 93, Chr94, Els90, Lei85, Tur93a,b].

There are many ways in which a temperature reduction could be achieved. We shall not discuss these individually here, but list only a selection of the many ideas (see e.g. [Kir86]). High magnetic fields or centrifugal pressure due to a rapidly rotating solar core could reduce the thermal pressure which is needed to balance the force of gravity. Another proposal involves the decrease of the central temperature gradient by turbulent mixings or a decrease in the opacity. The latter would be achieved by a smaller metal content z in the interior of the Sun. This would decrease the absorption of the photons and thus also the temperature gradient, which would result in a lower central temperature (see e.g. [Ber93]).

The existence of new particles, so-called cosmions (solar WIMPs; weakly interacting massive particles, see also chapter 9) has also been discussed. Cosmions trapped in the centre of the Sun could affect the energy transfer by radiation, by transporting energy away from the core, and again decrease the temperature (see e.g. [Fau85, Spe85, Gil86, Gri87, Pri88]). Particle masses of from 2 to 10 GeV are needed to explain the solar neutrino problem. The required density amounts to around 10^{-11} times the proton density. This new exotic particle would also play a role in the much discussed problem of dark matter. Recent experiments [Cal91c], however, have ruled out this particle and thus this solution (see chapter 9).

The proposals listed above represent only a selection. A more detailed discussion is given in [Bah89]. We note that most schemes concern only the ^{8}B neutrino flux, while the *pp* neutrino flux remains practically unaltered. A totally different resolution of the problem is based on particle physics.

7.3.5 The Mikheyev–Smirnov–Wolfenstein effect

7.3.5.1 Vacuum oscillations

Neutrino oscillations in the vacuum as discussed in section 7.2 were a possible explanation for the solar neutrino problem, i.e. the fact that only about one-third or half of the flux expected according to the Standard Solar Model is observed. Assuming that neutrinos have a mass and that there is a mixing of different flavour states, electron neutrinos produced in the Sun could change into muon or tau neutrinos on their way. Thus, in the case of complete mixing, only one-third of the original ν_e flux would be observable on the Earth. The other two-thirds would reach the Earth as muon or tau neutrinos (one-third each), which cannot be detected in the chlorine detector.

The probability of a transition from the state $|\nu_e\rangle$ to $|\nu_k\rangle$ is given by (7.27*a*). In typical solar neutrino experiments, the averaging over the neutrino energy leads to cancellation of the oscillating terms for $\alpha \neq \beta$ [Bah69], if the distance between the Sun and the Earth is very much greater than the oscillation length. Neglecting the oscillating terms, the probability that at time t (or, after a flight distance $x \approx t$) a ν_e remains in the original state is given by

$$P_{e\to e}(t) = \sum_{\alpha} |U_{e\alpha}|^4. \tag{7.86}$$

Starting from the highly likely hypothesis that there is a mass eigenstate $|\nu_1\rangle$ which forms the dominant fraction of the electron neutrino, we would then have $U_{e1} \approx 1$. The probability that an electron neutrino does not undergo a transition would then be practically equal to one, so that vacuum oscillations would appear to be a very unlikely explanation of the solar neutrino problem. However, they cannot be ruled out [Bar90, Ack91].

In what follows, we shall consider which reduction factors due to vacuum oscillations are possible. We assume N mass eigenstates, so that

$$P_{e\to e}(t) = \sum_{\alpha=1}^{N} |U_{e\alpha}|^4. \tag{7.87}$$

Since the overall probability is conserved, the matrix U must satisfy the following boundary condition

$$\sum_{\alpha=1}^{N} |U_{e\alpha}|^2 = 1. \tag{7.88}$$

We now seek the smallest value of (7.87), taking into account the conservation of the overall probability. Thus, we seek the minimum of the function

$$F = \sum_{\alpha=1}^{N} |U_{e\alpha}|^4 + \lambda \left(\sum_{\alpha=1}^{N} |U_{e\alpha}|^2 - 1 \right) \tag{7.89}$$

where the additional condition (7.88) is taken account of by the Lagrange multiplier λ. From the requirement

$$\frac{\partial F}{\partial U_{e\alpha}} = 0 \tag{7.90}$$

it follows that

$$|U_{e\alpha}|^2 = -\frac{\lambda}{2} \qquad \alpha = 1, 2, \ldots, N. \tag{7.91}$$

When all mass eigenstates contribute with the same amplitude to the state $|\nu_e\rangle$, the flux of electron neutrinos has its lowest value (this situation is referred to as *maximal mixing*). By virtue of (7.88), it follows in particular that

$$|U_{e\alpha}|^2 = \frac{1}{N}. \tag{7.92}$$

Thus, we finally obtain the result

$$[P_{e\to e}]_{\min} = \left(\sum_{\alpha=1}^{N} |U_{e\alpha}|^4\right)_{\min} = N\,\frac{1}{N^2} = \frac{1}{N}. \tag{7.93}$$

For N neutrino flavours, the flux of the electron neutrinos can be reduced by a factor of at most N. Assuming three light flavours, a decrease in the solar neutrino flux by a factor of three would be conceivable in the case of maximal mixing.

Thus, the solar neutrino problem could be explained by neutrino oscillations in vacuum, provided the mixing is sufficiently large. However, a large mixing appears improbable (but not impossible). This follows from a comparison with the quark sector, where the analogous Cabibbo angle $\theta_C \approx 13°$ is rather small. In summary, the explanation of the solar neutrino problem by vacuum oscillations would require maximal mixing of three neutrino eigenstates for all $\Delta m^2_{\alpha\beta} > 3 \times 10^{-11}\ \mathrm{eV}^2$.

7.3.5.2 Oscillations in matter

Based on fundamental considerations by Wolfenstein [Wol78, 79a], Mikheyev and Smirnov pointed out that the presence of matter may have a strong effect on the phenomenon of neutrino oscillations [Mik86a–c].

Wolfenstein recognized that matter affects the propagation of neutrinos through coherent, elastic forward scattering. Until now we have only considered oscillations and neutrino propagation in the vacuum, where the following equation holds for the propagation of a relativistic neutrino of mass m (see (7.14) and (7.20))

$$\nu(x, t) = \nu(0)\exp[\mathrm{i}(px - Et)] \simeq \nu(0)\exp\left(-\mathrm{i}t\frac{m^2}{2p}\right). \tag{7.94}$$

In matter, the phase in (7.94) is altered from ipx to $ipxn$, where n denotes the refractive index. Because of the weak interaction, n is different from the vacuum value $n_0 = 1$. The classical expression for the index n is [Wol78]

$$n_i = 1 + \frac{2\pi N}{p^2} f_i(0). \tag{7.95}$$

N is the density of the scattering centres, p is the neutrino momentum and $f(0)$ is the scattering amplitude for forward scattering. The index i is intended to denote the different flavour states. If the refractive index were the same for all i, nothing in the previously discussed picture of neutrino oscillations would change, since the phase of all components would be modified in the same way. Only a relative change of the phases amongst themselves would have physical consequences. Wolfenstein showed that matter affects the propagation of the individual neutrino flavours in different ways and thus has the effect of a relative phaseshift, which is equivalent to a change in the mass matrix (see section 7.2.1).

The physical reason for this is easy to see. Matter consists of quarks and electrons. The contribution of the quarks to the scattering amplitude $f(0)$ is the same for all neutrino flavours, provided there are no weak neutral currents which change the flavour quantum number ($f_e^q(0) = f_\mu^q(0)$). The neutrino–quark scattering process only involves Z^0 exchange (figure 7.24(a)). This contribution results in a minor, since for all neutrinos the same, change in the amplitude (7.94), and thus is of no importance to the following discussion.

However, there are differences as far as the scattering by electrons is concerned. While the neutral weak current again gives identical scattering amplitudes for all neutrino flavours, there are differences for the charged weak

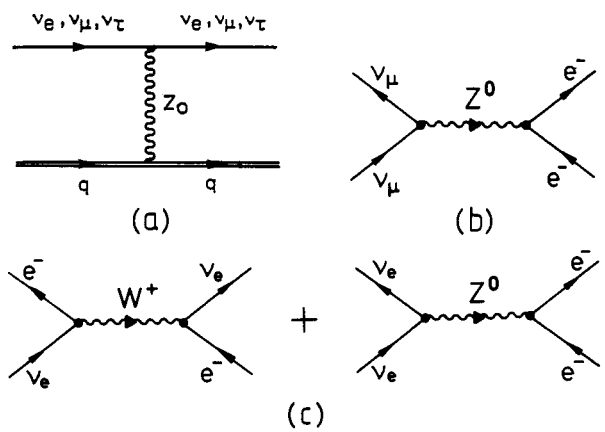


Figure 7.24. Feynman diagrams for (a) neutrino–quark scattering, (b) $\nu_\mu e^-$ scattering, and (c) $\nu_e e^-$ scattering.

current. The situation is illustrated in figures 7.24(b) and 7.24(c). The process $\nu_e e^- \rightarrow \nu_e e^-$ may take place via Z^0 exchange and also via $W^{\pm}$ exchange. A ν_e emits an electron and a positively charged vector boson which is absorbed at the second vertex by an electron, forming a ν_e. There are no corresponding Feynman diagrams for neutrinos of the other flavours.

The interaction via the charged weak currents thus results in a difference in the propagation of ν_e and $\nu_{\mu/\tau}$. Calculation of the Feynman diagram gives (the formula includes a factor $-\sqrt{2}$ [Lew80, Lan83] which was not present in the original work [Wol78])

$$\Delta f(0) = f_e(0) - f_i(0) = -\sqrt{2}\frac{G_F p}{2\pi}. \tag{7.96}$$

The sign must be changed for $\overline{\nu}_e$.

We shall now discuss the effect of the additional term in figure 7.24(c) in more detail and return to the mass matrix introduced in section 7.2.1. In order to follow the basic concepts, in what follows, we shall only consider two neutrino flavours (ν_e, ν_μ). A simple presentation of the MSW effect was given by Bethe [Bet86b]. We shall consider this in what follows, before we discuss a somewhat more formal approach.

We assume that neutrinos have a mass and that the mass matrix $\mathcal{M}_0$ in the basis of the flavour eigenstates is not diagonal (the index 0 stands for the vacuum). According to (7.6b), the flavour and mass eigenstates are related as follows

$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = U^\dagger \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}. \tag{7.97}$$

If the mass matrix $\mathcal{M}_0$ has the eigenvalues m_1 for ν_1 and m_2 for ν_2, the following holds for the matrix of the squares of the masses in the basis (ν_1, ν_2)

$$(\nu_1^\dagger, \nu_2^\dagger) \begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = (\nu_e^\dagger, \nu_\mu^\dagger) U^\dagger \begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix} U \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix}. \tag{7.98}$$

We shall transform this a little further

$$\begin{aligned} U^\dagger \begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix} U &= \frac{1}{2}(m_1^2 + m_2^2) U^\dagger \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} U \\ &\quad + \frac{1}{2}(m_1^2 - m_2^2) U^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} U \\ &= \frac{1}{2}(m_1^2 + m_2^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \frac{1}{2}(m_1^2 - m_2^2) \begin{pmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ -2\sin\theta\cos\theta & -\cos^2\theta + \sin^2\theta \end{pmatrix} \end{aligned}$$

$$= \frac{1}{2}(m_1^2 + m_2^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}(m_2^2 - m_1^2)\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \tag{7.99}$$

Thus, in the vacuum we obtain

$$\mathcal{M}_0^2 = \frac{1}{2}(m_1^2 + m_2^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}(m_2^2 - m_1^2)\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \tag{7.100}$$

The interaction of the electron neutrinos with the electrons of the matter via $W^{\pm}$ exchange now leads to a change in the square of the mass matrix to

$$\mathcal{M}^2 = \mathcal{M}_0^2 + \mathcal{M}_{\text{Mat}}^2. \tag{7.101}$$

The interaction is described by the Hamiltonian

$$H_{WW} = \frac{G_F}{\sqrt{2}}\left[\bar{e}\gamma_\mu(1-\gamma_5)\nu_e\right]\left[\bar{\nu}_e\gamma^\mu(1-\gamma_5)e\right]. \tag{7.102}$$

This may be brought to the following form by a Fierz transformation (see [Gre86b])

$$H_{WW} = \frac{G_F}{\sqrt{2}}\left[\bar{\nu}_e\gamma_\mu(1-\gamma_5)\nu_e\right]\left[\bar{e}\gamma^\mu(1-\gamma_5)e\right]. \tag{7.103}$$

G_F denotes the Fermi coupling constant of the weak interaction. In addition to (7.20), the energy of a neutrino of mass m with momentum $p \gg m$ also contains a contribution due to H_{WW}. The effective energy in matter is

$$E_{\text{eff}} = p + \frac{m^2}{2p} + \langle e\nu|H_{WW}|e\nu\rangle. \tag{7.104}$$

We shall now evaluate the four-current density

$$j^\mu = \bar{e}\gamma^\mu(1-\gamma_5)e \tag{7.105}$$

of the electrons in the rest system of the Sun. By virtue of the statistical distribution of the electrons, the current density vanishes and the only contribution comes from the ($\mu = 0$) component, i.e.

$$j^\mu = N_e\delta^{\mu 0} \tag{7.106}$$

where N_e denotes the average number of electrons per unit volume. For left-handed neutrinos ν_L, since $\nu_L = (1/2)(1 - \gamma_5)\nu$, we may replace the matrix $(1 - \gamma_5)$ by a factor 2 and obtain

$$H_{WW} = \sqrt{2} G_F N_e \bar{\nu}_e \gamma_0 \nu_e. \tag{7.107}$$

The electrons contribute an additional potential for the electron neutrinos

$$V = \sqrt{2} G_F N_e. \tag{7.108}$$

Thus, the momentum and the energy of the neutrino in matter are related by the following equation

$$p^2 + m^2 = (E - V)^2 \simeq E^2 - 2EV \tag{7.109}$$

in which the quadratic term V^2 is neglected. The potential V is equivalent to a change in the square of the mass m^2 to

$$m^2 \to m^2 + A \tag{7.110}$$

where

$$A = 2EV = 2\sqrt{2} G_F N_e E. \tag{7.111}$$

Thus, the matrix of the squares of the masses of ν_e and ν_μ in matter is given by

$$\begin{aligned} \mathcal{M}^2 &= \mathcal{M}_0^2 + \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \\ &= \frac{1}{2}(m_1^2 + m_2^2 + A) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \frac{1}{2} \begin{pmatrix} A - \Delta \cos 2\theta & \Delta \sin 2\theta \\ \Delta \sin 2\theta & -A + \Delta \cos 2\theta \end{pmatrix} \end{aligned} \tag{7.112}$$

where $\Delta = m_2^2 - m_1^2$. The eigenvalues of $\mathcal{M}^2$ are

$$m^2_{\nu_{1,2}} = \frac{1}{2}(m_1^2 + m_2^2 + A) \pm \frac{1}{2}\sqrt{(\Delta \cos 2\theta - A)^2 + \Delta^2 \sin^2 2\theta}. \tag{7.113}$$

Thus, the mass depends on A, i.e. on the electron density (see figure 7.25). For $\Delta = m_2^2 - m_1^2 > 0$, the splitting of the square of the mass has a minimum as a function of A, since A and $\cos 2\theta$ are positive.

If ν_e and ν_μ are not mixed with one another ($\theta = 0$) the eigenvalues lie on the broken curves. The two levels cross at $A = \Delta$. If the mixing angle is non-vanishing, the eigenvalues lie on the full curves. At smaller density of matter the electron neutrino has the smaller mass. When A attains the value

$$A_0 = 2\sqrt{2} G_F N_e E_0 = \Delta \cos 2\theta \tag{7.114}$$

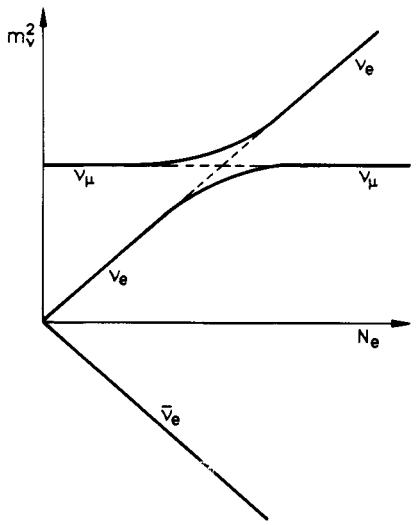


Figure 7.25. The dependence of the masses of ν_e, ν_μ and $\bar{\nu}_e$ on the electron density N_e of the surrounding matter (after [Bet86b]).

the splitting has a minimum. The two curves would cross for vanishing $\Delta \sin 2\theta$. If the density increases further, the electron neutrino then has a larger mass than the muon neutrino. The solar neutrino problem is now explained as follows. An electron neutrino is generated in the interior of the Sun at a sufficiently large electron density ($A > \Delta \cos 2\theta$). This lies on the upper curve in figure 7.27. If the ν_e moves outwards, the electron density decreases and ultimately reaches the resonance density (7.114). If the density of matter decreases slowly (adiabatic approximation) the neutrino moves along the continuous line and finally leaves the Sun as a muon neutrino which cannot be detected in the chlorine detector. In this adiabatic approximation the state vector is rotated from the direction of the $|\nu_e\rangle$ to that of the $|\nu_\mu\rangle$. Because of the slow change in density the neutrino remains in the state with the greater mass. If muon neutrinos are heavier than electron neutrinos, essentially only muon neutrinos will occur on the edge of the Sun.

Equation (7.114) defines the condition for a resonant conversion from ν_e to ν_μ. The resonance condition is satisfied by a certain electron density for each neutrino energy. A critical energy E_c is defined in terms of the density at the centre of the Sun N_e^c

$$E_c = \frac{m_2^2 - m_1^2}{2\sqrt{2} G_F N_e^c}. \tag{7.115}$$

All neutrinos with energies $E > E_c$ pass through the resonance and appear as

muon neutrinos, while electron neutrinos with $E < E_c$ cannot meet the resonance condition and leave the Sun unchanged and can be detected. Applying the MSW effect to the chlorine experiment, estimation of the order of magnitude of N_e^c gives a mass parameter $m_2^2 - m_1^2 \simeq 6 \times 10^{-5}$ eV^2, i.e. $m_2 \simeq 0.008$ eV for $m_1 \ll m_2$ [Bet86b]. Such a sensitivity is very difficult to achieve in laboratory experiments.

We shall now discuss the problem of neutrino oscillations in matter in somewhat more detail, since it provides a very elegant solution of the solar neutrino problem. A phase change in the neutrino wavefunction may be caused by two things:

(i) mass difference between ν_1 and ν_2;
(ii) different interaction of ν_e and ν_μ.

Both effects may be described by differential equations [Wol78, 79, Mik88]. We saw in section 7.2 that the temporal evolution of a mass eigenstate $|\nu_\alpha\rangle$ in vacuum is described by the following expression

$$|\nu_\alpha(t)\rangle = |\nu_\alpha\rangle \exp\left(\mathrm{i}\frac{m_\alpha^2}{2p}t\right). \tag{7.116}$$

In vacuum, states with a defined mass develop independently of one another and the equation of motion (Schrödinger equation) is

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\nu_\alpha = E_\alpha \nu_\alpha = \left(p_\alpha + \frac{m_\alpha^2}{2p_\alpha}\right)\nu_\alpha. \tag{7.117}$$

Since $p \simeq p_\alpha$, the linear term in p_α affects the phase of all $|\nu_\alpha\rangle$ in the same way and thus is unimportant as far as the following discussion is concerned. Thus, the equation of motion is

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\nu_1\\ \nu_2\end{pmatrix} = \begin{pmatrix} m_1^2/2p & 0 \\ 0 & m_2^2/2p \end{pmatrix}\begin{pmatrix}\nu_1\\ \nu_2\end{pmatrix}. \tag{7.118}$$

It is easy to check by substitution that equation (7.116) is a solution of this Schrödinger equation. The corresponding expressions for the flavour eigenstates are obtained using the transformation (7.97) (compare with the derivation in (7.98)–(7.100)).

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\nu_e\\ \nu_\mu\end{pmatrix} = \tilde{H}_f\begin{pmatrix}\nu_e\\ \nu_\mu\end{pmatrix} \tag{7.119a}$$

with

$$\tilde{H}_f = H_f + H_1 \tag{7.119b}$$

where

$$H_f = \frac{m_2^2 - m_1^2}{4p} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \tag{7.119c}$$

$$H_1 = \frac{m_1^2 + m_2^2}{4p} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.119d}$$

The term H_1 is a multiple of the unit matrix. This constant again alters the phase of the states ν_e and ν_μ in the same way, and thus is not taken into account in the derivation of the probability amplitudes for neutrino oscillations. It is customary to set

$$\tilde{H}_f \equiv H_f. \tag{7.120}$$

However, we shall still retain the term H_1 for a while.

The effect of the matter is manifest through an additional potential $V = \sqrt{2} G_F N_e$ for the electron neutrinos. Using the abbreviations

$$a = \frac{m_1^2 + m_2^2}{4p} \qquad b = \frac{m_2^2 - m_1^2}{4p} \tag{7.121a}$$

the Schrödinger equation in matter may be written in the form

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \left[\begin{pmatrix} a + V & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} -b\cos 2\theta + V & b \sin 2\theta \\ b \sin 2\theta & b \cos 2\theta \end{pmatrix} \right] \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix}. \tag{7.121b}$$

Conversion to the system of mass eigenstates using the transformation (7.97) gives

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \begin{pmatrix} (m_1^2/2p) + \sqrt{2} G_F N_e \cos^2\theta & \sqrt{2} G_F N_e \sin\theta \cos\theta \\ \sqrt{2} G_F N_e \sin\theta \cos\theta & (m_2^2/2p) + \sqrt{2} G_F N_e \sin^2\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}. \tag{7.122}$$

The terms containing the electron density N_e relate to the forward scattering of the ν_e in four channels [Ric87]

$$\nu_1 e^- \to \nu_1 e^- \quad (\sim G_F \cos^2\theta) \tag{7.123a}$$

$$\nu_1 e^- \to \nu_2 e^- \quad (\sim G_F \sin\theta \cos\theta) \tag{7.123b}$$

$$\nu_2 e^- \to \nu_1 e^- \quad (\sim G_F \sin\theta \cos\theta) \tag{7.123c}$$

$$\nu_2 e^- \to \nu_2 e^- \quad (\sim G_F \sin^2\theta). \tag{7.123d}$$

The new Hamiltonian (7.122) may be brought to diagonal form using the usual methods of quantum mechanics. The eigenstates for the propagation in matter

are density dependent and differ from those in the vacuum (ν_1, ν_2)

$$\begin{aligned} |\nu_{1m}\rangle &= |\nu_e\rangle \cos\theta_m - |\nu_\mu\rangle \sin\theta_m \\ &= |\nu_1\rangle \cos(\theta_m - \theta) - |\nu_2\rangle \sin(\theta_m - \theta) \qquad (7.124a) \\ |\nu_{2m}\rangle &= |\nu_e\rangle \sin\theta_m + |\nu_\mu\rangle \cos\theta_m \\ &= |\nu_1\rangle \sin(\theta_m - \theta) + |\nu_2\rangle \cos(\theta_m - \theta). \qquad (7.124b) \end{aligned}$$

The new mixing angle θ_m is related to the vacuum mixing angle by the following equation

$$\tan 2\theta_m = \frac{\sin 2\theta}{\cos 2\theta + (L/L_e)}. \qquad (7.125)$$

L is the vacuum oscillation length defined in (7.48). The so-called neutrino–electron interaction length L_e is given by

$$L_e = \frac{2\pi}{\sqrt{2} G_F N_e}. \qquad (7.126)$$

The effective oscillation length in matter is given by

$$L_m = L \frac{\sin 2\theta_m}{\sin 2\theta} = L \left[1 + \left(\frac{L}{L_e} \right)^2 + \frac{2L}{L_e} \cos 2\theta \right]^{-1/2}. \qquad (7.127)$$

If the electron density N_e vanishes, then $\theta_m = \theta$ and $\nu_{\alpha m}$ becomes the mass eigenstate in the vacuum ν_α. For an infinite density N_e, on the other hand, we have

$$\theta_m = \begin{cases} 0 & \text{if } m_1 > m_2 \\ \pi/2 & \text{if } m_2 > m_1. \end{cases} \qquad (7.128)$$

In the case of an infinite density the matter eigenstates are given by the flavour eigenstates. The eigenvalues of the Hamiltonian are shown in figure 7.26 as a function of the electron density (for $m_2 > m_1$). For large densities ($\theta_m \simeq \pi/2$), according to (7.124), the electron neutrino is the state with the larger mass ($|\nu_e\rangle \simeq |\nu_{2m}\rangle$).

Assuming a constant density N_e, (7.122) is easy to solve. The probability $P_{\nu_e \to \nu_e}$ that an electron neutrino is found to be unchanged after crossing a distance x in matter is given by

$$\begin{aligned} P_{\nu_e \to \nu_e}(x) &= 1 - \frac{1}{2} \sin^2 2\theta_m \left(1 - \cos \frac{2\pi x}{L_m} \right) \\ &= 1 - \sin^2 2\theta_m \sin^2 \frac{\pi x}{L_m}. \qquad (7.129a) \end{aligned}$$

Correspondingly, we have

$$P_{\nu_e \to \nu_\mu}(x) = \sin^2 2\theta_m \sin^2 \frac{\pi x}{L_m}. \qquad (7.129b)$$

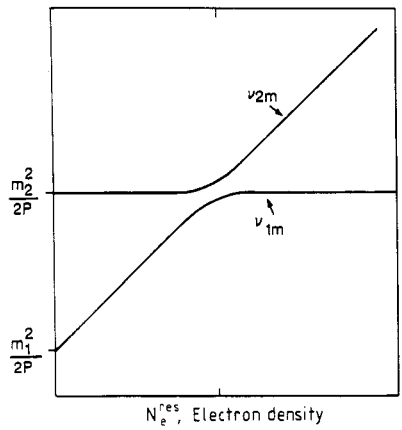


Figure 7.26. The eigenvalues of the matrix of equation (7.122) as a function of the electron density. The curves were calculated for the case $m_2 > m_1$ and $\sin\theta \ll 1$. The upper curve corresponds to the eigenvalue of $|\nu_{2m}\rangle$, the lower to that of $|\nu_{1m}\rangle$ (from [Ric87]).

Matter oscillations are not a new physical phenomenon, but are closely associated with the vacuum oscillations. They can only occur if neutrinos have a mass and mix (this changes if there exist neutral weak currents which alter the flavour quantum number; however, these currents are not contained in the standard model of the electroweak theory and have not yet been found experimentally). As $N_e \to 0$, (7.129) turns into the expression (7.47).

Mikheyev and Smirnov [Mik86a,b] found an interesting resonance phenomenon, which may occur for $m_2 > m_1$. The resonant nature of matter oscillations is clear from (7.125). We shall consider three special cases as follows:

- $(|L|/L_e) \ll \cos 2\theta$ According to (7.126), this case corresponds to a small electron density. The matter has only a small effect on the oscillatory phenomenon

$$P_{\nu_e \to \nu_\mu}(x) = \sin^2 2\theta \sin^2 \frac{\pi x}{L}. \tag{7.130}$$

- $(|L|/L_e) \gg \cos 2\theta$ This corresponds to the case of a very large electron density. We have $\theta_m \approx \pi/2$. We note that the mass hierarchy is reversed; the electron neutrino which is lighter in the vacuum now corresponds to the heavier mass eigenstate ($\nu_e = \nu_{2m}$ and $\nu_\mu = -\nu_{1m}$). In the limiting case of an infinite electron density neutrinos no longer oscillate

$$P_{\nu_e \to \nu_\mu}(x) = \left(\frac{L_e}{L}\right)^2 \sin^2 2\theta \sin^2 \frac{\pi x}{L_e}. \tag{7.131}$$

- $|L/L_e| \approx \cos 2\theta$ In this case, the oscillations may be increased by a resonance effect. For $\theta_m = \pi/4$, we have

$$P_{\nu_e \to \nu_\mu}(x) = \sin^2(\pi x \sin 2\theta / L). \tag{7.132}$$

We shall discuss this resonance pattern in detail below.

From (7.125), it is clear that the mixing angle attains a maximum ($\theta_m = \pi/4$) when the electron density satisfies the resonance condition

$$\frac{L}{L_e} = -\cos 2\theta. \tag{7.133}$$

For $m_2 < m_1$, this condition cannot be satisfied for electron neutrinos, since the oscillation length L is positive. We would then expect resonance to occur for $\bar{\nu}_e$. In what follows we shall only consider the (natural) case in which $m_2 > m_1$.

Even for very small vacuum mixing angles θ, the parameter θ_m may have value $\pi/4$ in matter (maximal mixing), provided the electron density satisfies the condition

$$N_e^{\text{res}} = \frac{(m_2^2 - m_1^2)\cos 2\theta}{2\sqrt{2}G_F E} \qquad \text{(MSW resonance density)}. \tag{7.134}$$

At this resonance density the two diagonal elements of the mass matrix are of equal size. The mixing between electron and muon neutrinos is increased, so that the size of the vacuum mixing angle is no longer important. A typical resonance curve (figure 7.27) is obtained when plotting the oscillation amplitude $\sin^2 2\theta_m$ as a function of the electron density. The amplitude takes its maximum value at a certain density N_e^{res}. The resonance density depends on the energy of the neutrinos.

In order to apply the MSW effect to the solar neutrino problem, we must also take into account the fact that the density along the flight path of the neutrino is not constant, but decreases from a large value at the point of creation to zero at the edge of the Sun. This means that the resonance condition may be satisfied for a wide range of values of $E/(m_2^2 - m_1^2)$. From a certain critical energy, the neutrinos on their way from the centre to the outer edge of the Sun enter an area in which the electron density satisfies condition (7.133) or (7.134), so that a considerable conversion $\nu_e \to \nu_\mu$ may occur. An electron neutrino always passes through the resonance region, provided its energy exceeds the value which would be needed to satisfy the resonance condition at the density of the centre of the Sun, i.e.

$$E_{\min} = \frac{(m_2^2 - m_1^2)\cos 2\theta}{2\sqrt{2}G_F N_e(0)}. \tag{7.135}$$

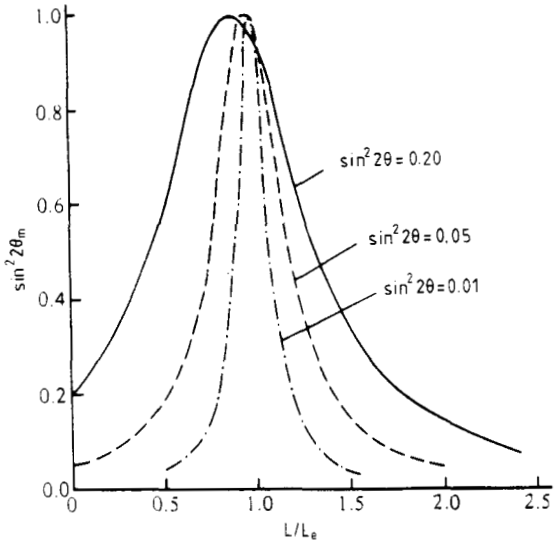


Figure 7.27. Dependence of the effective mixing angle θ_m in matter on the ratio L/L_e. The curves are calculated for various vacuum mixing angles θ (from [Boe92]).

Substituting numerical values, we obtain

$$E_{\min} = 6.6 \cos 2\theta \frac{\Delta m^2}{10^{-4}\ \text{eV}^2}\ [\text{MeV}]. \tag{7.136}$$

This means many of the ^{8}B neutrinos pass through a resonance area, if $\Delta m^2 < 10^{-4}$ eV^2.

For a variable density N_e, the neutrino state at each point of the path may be written as a linear combination of the local eigenstates $|\nu_{1m}\rangle$ and $|\nu_{2m}\rangle$

$$|\nu_e(t)\rangle = a(t)|\nu_{1m}\rangle + b(t)|\nu_{2m}\rangle. \tag{7.137}$$

Suppose that the electron neutrino is generated at time $t = 0$, so that

$$a(0) = \cos\theta_m \qquad b(0) = \sin\theta_m \tag{7.138}$$

where θ_m is the mixing angle at the point of generation. The parameters $a(t)$ and $b(t)$ are solutions of the Schrödinger equation (7.122). The neutrino leaves the Sun as a certain combination of $|\nu_e\rangle$ and $\nu_\mu\rangle$. The description is particularly simple in the case of the adiabatic approximation. It follows from the adiabatic theorem of quantum mechanics [Mes76, 86] that, for a slow variation of the Hamiltonian, the basis states change, but no transitions between

the individual states are induced. The system remains in the original (time-dependent) eigenstate. Thus, the absolute values of the coefficients $a(t)$ and $b(t)$ are time independent

$$|a(t)| = |a(0)| \qquad |b(t)| = |b(0)|. \tag{7.139}$$

We have a slow variation of the density, if N_e may be viewed as approximately constant over the effective oscillation length L_m.

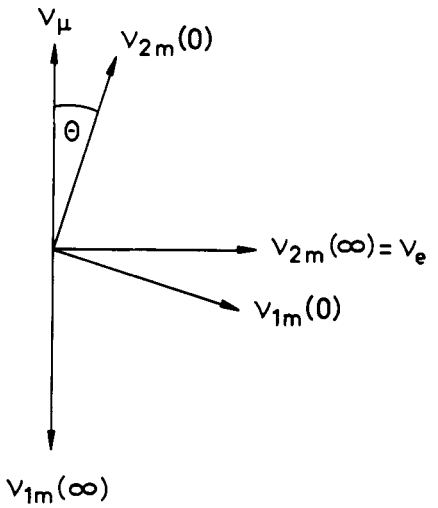


Figure 7.28. Illustration of the MSW effect in the adiabatic approximation (see text).

In the case of the adiabatic approximation, the MSW effect may be understood as follows (see figure 7.28). We assume that $m_2 > m_1$. Electron neutrinos ν_e are generated in the interior of the Sun. At large densities, the effect of forward scattering is dominant compared with the mass splitting of the mass eigenstates. The electron neutrino essentially behaves like the heavier effective mass eigenstate ν_{2m}; in the limit $N_e \to \infty$, we have $|\nu_e\rangle = |\nu_{2m}(\infty)\rangle$. On its outward path the neutrino enters regions with lower electron density. The vectors $|\nu_{1m}\rangle$ and $|\nu_{2m}\rangle$ are rotated relative to the states $|\nu_e\rangle$ and $|\nu_\mu\rangle$, since the mixing angle θ_m depends on N_e. In our case, the neutrino state is rotated in the direction of ν_μ. According to the adiabatic theorem, the neutrino remains in the state $|\nu_{2m}\rangle$ and only its relative orientation changes. For an average density the mixing angle has value $\theta_m = \pi/4$, i.e. the resonance case occurs. If the vacuum is attained, the neutrino originally generated as a ν_e is identical with the mass eigenstate $\nu_{2m}(0)$. The crucial point is that the electron neutrino generated in the state $|\nu_{2m}\rangle$ remains in this time-varying state and essentially leaves the Sun as a muon neutrino.

Since the neutrino is in a mass eigenstate on leaving the Sun, no more oscillations occur. The probability of detecting it on Earth as a muon neutrino is given by the projection of $|\nu_{2m}(0)\rangle$ onto $|\nu_\mu\rangle$, i.e. we obtain

$$P(\nu_\mu) = \cos^2\theta \tag{7.140a}$$

$$P(\nu_e) = \sin^2\theta. \tag{7.140b}$$

Thus, the MSW effect is able to provide a very elegant explanation of why the original electron neutrino is highly likely to reach the Earth as a muon neutrino, even though the mixing angle in vacuum, θ, is very small. However, we note that this conversion is energy dependent. The energy must be sufficiently large that the resonance condition is satisfied. We also note that oscillations in matter, unlike those in the vacuum, depend explicitly on the signs of Δm^2 and $\Delta f_e(0)$. For example, if one were to change the sign of the interaction ($\Delta f_e(0)$) by using $\overline{\nu}_e$, or if the dominant mass eigenstate of the ν_e in the vacuum were the heavier ($m_1 > m_2$), no resonances would occur. There would be no decrease in the solar neutrino flux.

We have described the effect for the adiabatic approximation. If we are no longer able to neglect the change in the electron density over an oscillation length, the situation becomes very much more complicated. In this case, one has to calculate the probability of a jump from one adiabatic mass eigenstate to another. We shall not discuss this further here and refer readers to the specialist literature [Bil87, Mik88, Bah89]. Theoretical studies of the MSW effect in a non-homogeneous medium can also be found, for example, in [Hal86].

The MSW effect may be taken to be an explanation of the solar neutrino problem for the parameter region $\Delta m^2 = 10^{-4} - 10^{-8}$ eV^2 and $\sin^2\theta > 10^{-4}$. This solution is particularly attractive, especially since such small masses and mixing angles are very easy to incorporate into the standard model of the electroweak theory in a natural way using the 'see-saw' mechanism (see [Lan81]).

Assuming the MSW effect, the results of the chlorine and the Kamiokande experiments may be used to deduce allowed parameter combinations in the Δm^2–$\sin^2 2\theta$ plane, under the SSM (figure 7.29). Bahcall and Bethe claimed that the non-adiabatic solution of the MSW equations with the parameter combination

$$\sin^2\theta\,\Delta m^2 \approx 10^{-8}\ \text{eV}^2 \tag{7.141}$$

may provide a consistent description of the data from the chlorine and from the Kamiokande experiment [Bah90]. This would mean that the gallium detectors to be discussed below should detect approximately one order of magnitude fewer neutrinos than one would expect from the SSM (but, see also [Bal91] which shows that the proposed solution is only one of many possible solutions).

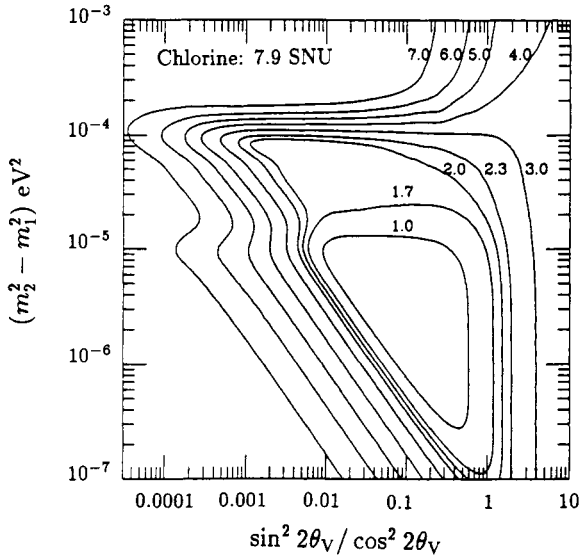


Figure 7.29. MSW diagram for the chlorine experiment. Each contour line is labelled by the corresponding value of the neutrino flux in SNU (from [Bah89]). Here, θ_V is the vacuum mixing angle.

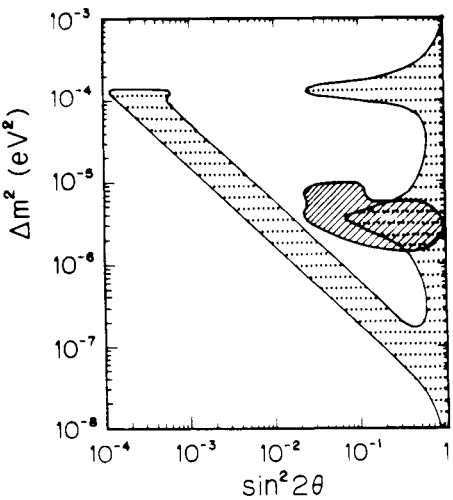


Figure 7.30. The dashed region in the Δm^2–$\sin^2 2\theta$ plane is excluded by virtue of the absence of a day–night effect (confidence level 90%). The dotted area corresponds to the allowed parameter region according to the measurement of the total neutrino flux and the energy spectrum of the recoil electrons by Kamiokande II (from [Hir91]).

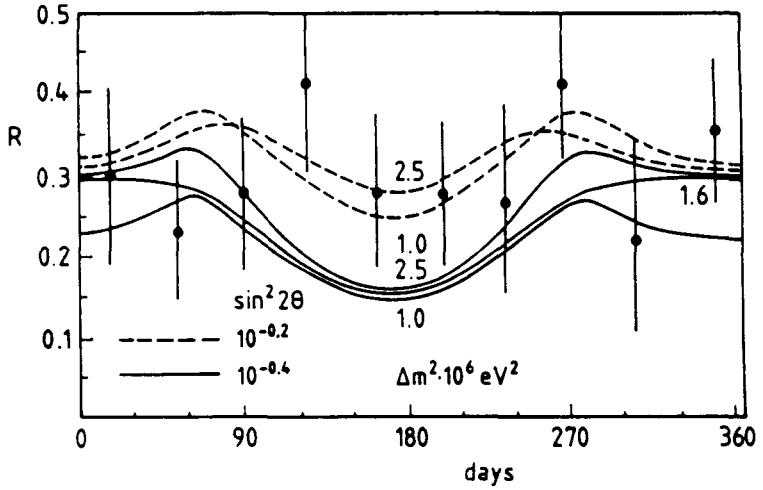


Figure 7.31. Seasonal fluctuations of the ^{37}Ar production in the chlorine detector. The curves shown correspond to different parameter combinations (from [Mik88]).

The MSW effect leads to an interesting regeneration effect, which should be observable on Earth, analogous to the phenomenon in the K^0 system [Bou86, Cri86, Mik86c, Bal87, 88]. Electron neutrinos which are converted into muon neutrinos as they fly through the Sun may be in part reconverted into electron neutrinos as they pass through the Earth, since the Earth also has a non-zero electron density. Consequently, a *day–night effect* should be observable. At night the solar neutrinos should pass through the Earth, so that, as in the Sun, a resonant transition may occur which would lead to a recovery of electron neutrinos. However, such a day–night modulation of the solar neutrino flux cannot be detected with a radiochemical detector since the latter integrates the flux over several weeks. However, it should be possible to detect such modulations using the Kamiokande Cerenkov counter. However, the analysis of measurements taken over 1040 days does not provide evidence for a temporal variation, to within the statistical error [Hir91]. In particular, this analysis was also used to further restrict the allowed parameters in the MSW Δm^2–$\sin^2 2\theta$ plane (see figure 7.30).

In addition to a day–night modulation, the Earth also induces a seasonal modulation, since, as a result of the variation of the relative orientation of the Earth and the Sun, the average density of the matter which a neutrino must penetrate depends on the season. Thus, we expect maxima of the ν_e flux in spring and autumn. These fluctuations should also be detectable with a radiochemical detector. Figure 7.31 illustrates the data from the chlorine experiment. However,

analysis has not produced an unambiguous result. The curves shown correspond to various combinations of parameters. The data from Kamiokande II do not exhibit any significant daily or seasonal variation [Hir91].

In describing the MSW effect, we considered only the case of two flavours. The more general case with three neutrino flavours is discussed in [Kuo86, Pet87, Mik88b]. Three crossings of levels are obtained. In an analogous way to the case discussed above, the neutrino leaves the Sun in the heaviest mass eigenstate if the density at the point of generation is sufficiently large and the momentum lies in a specific interval. For smaller neutrino momentum the neutrino leaves the Sun in the mass eigenstate with the second heaviest mass. Below a certain threshold the neutrino occurs in the lowest mass eigenstate.

To end this section on the MSW effect, we note that [Ber91a] recently discovered a possible relationship between the strong CP problem (see chapter 1) and the solar neutrino problem. The authors proposed a solution of the strong CP problem using the Peccei–Quinn mechanism, which automatically leads to a small Majorana mass for the neutrino ($< 10^{-1}$ eV) – without using the 'see-saw' mechanism. This mass lies in precisely the area required to explain the neutrino enigma in terms of the MSW effect. The axion necessarily generated by breaking of the introduced chiral U(1) symmetry only couples weakly to neutrinos.

7.3.6 The magnetic moment of the neutrino

The data from the chlorine experiment appear to exhibit another interesting characteristic. If one plots the time-varying flux against the 11-year solar cycle the neutrino flux and the solar activity (number of sunspots) appear to be anticorrelated [Row85, Dav88, 94, Bah89] (see figure 7.32). It is noticeable that the capture rate in the period around 1980 practically sank to the background rate. A discussion of the statistical significance of this apparent anticorrelation may be found in [Dav87b, 94, Bah89]. As previously mentioned, the Kamiokande measurements provide no evidence that the flux is time dependent.

If we take this anticorrelation seriously, we encounter the problem that the previous proposals for a solution of the solar neutrino problem do not provide for a relationship between the solar activity and the neutrino flux. One possible explanation involves a magnetic moment of the neutrino [Cis71]. Okun later noted that a magnetic moment and an electric dipole moment are practically indistinguishable for a relativistic, light neutrino [Oku86]. If the neutrino had a large magnetic (and/or electric) dipole moment, the neutrino spin could be flipped by the interaction with the solar magnetic field ('spin flip'). This effect would transform a left-handed electron neutrino into a right-handed electron neutrino, which cannot be detected in the detectors, since right-handed neutrinos do not participate in the weak interaction (*sterile neutrinos*) [Oku86, Vol86]. The flux

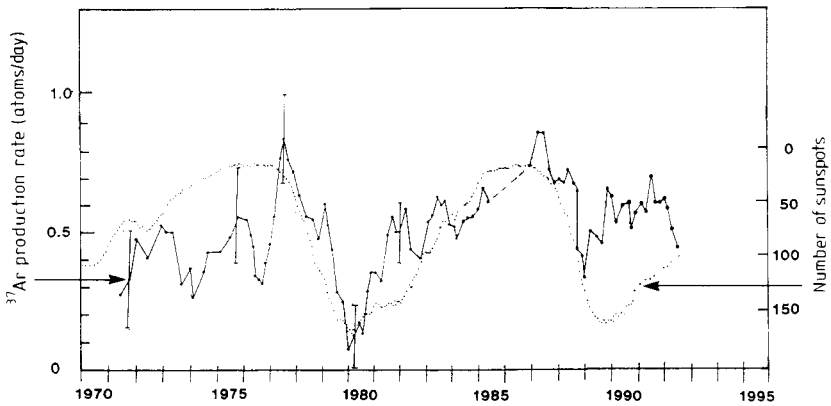


Figure 7.32. On the correlation between the neutrino rate of the chlorine detector and the solar activity. The dotted curve represents the solar activity (number of sunspots per month) (inverse scale), the black points (full curve) the observed neutrino flux, averaged over five measurement periods in each case (from [Dav94, 94a]).

of neutrinos which are not subject to a spin reversal, depends on the magnetic moment μ_m, and on the strength and extent of the magnetic field. We have

$$\Phi_B(E) = \Phi_0(E)\cos^2\left(\mu_m \int B_\perp \mathrm{d}l\right). \tag{7.142}$$

Here $B_\perp$ denotes the component of the B field perpendicular to the direction of propagation of the neutrino. $\Phi_0(E)$ is the flux of the left-handed neutrinos produced in the Sun. Spin oscillations between non-degenerate mass eigenstates then play an important role, if the magnetic energy is greater than the mass difference

$$|\mu_m B_\perp| > |\Delta m^2|. \tag{7.143}$$

It follows from this that a magnetic moment in the region

$$\mu_m = (10^{-11} - 10^{-10})\,\mu_B \qquad \mu_B = \frac{e}{2m_e} \tag{7.144}$$

is needed in order to obtain a sufficiently high probability for a spin flip. In this situation, it would be perfectly possible to explain an anticorrelation with the solar activity.

The solar activity and the solar magnetic field are closely interrelated. An increased number of sunspots means a greater magnetic field and thus a greater probability of conversion into sterile, right-handed neutrinos. Thus, when the solar activity was a maximum, a minimum neutrino flux would be measured on

the Earth. In addition, one might expect a half-yearly variation, which, however, has not been seen by Kamiokande [Hir91].

It is generally assumed that neutrinos do not take part in the electromagnetic interaction. As we saw in chapter 1, Majorana neutrinos cannot have an electric or magnetic dipole moment

$$\mu_m = 0 \qquad \mu_e = 0 \quad \text{(Majorana neutrinos).} \tag{7.145a}$$

Here, we note that, strictly speaking, only the diagonal magnetic moment disappears, while non-zero off-diagonal elements, which describe transitions between the different flavours, may occur.

For Dirac neutrinos, on the other hand, we expect a (diagonal) magnetic moment which is proportional to the mass [Lee77, Fuj80, Liu87]

$$\mu_m \sim 10^{-19} \frac{m_\nu}{1\ \text{eV}} \mu_B \qquad \text{(Dirac neutrinos).} \tag{7.145b}$$

This magnetic moment is approximately eight orders of magnitude smaller than that required to explain the solar neutrino problem. There are, however, also models which give a value of μ_m in the range from 10^{-10} to 10^{-11} μ_B (see e.g. [Vol86, Fuk87, Ste88]).

The current experimental bounds for the magnetic moment are still very far from the value given in (7.145*b*). A bound

$$\mu_m < 2 \times 10^{-10}\ \mu_B \tag{7.146}$$

has been derived from $\bar{\nu}_e$ scattering experiments at reactors [Lim88]. Astrophysical observations are more restrictive, but also more uncertain, and give [Mar86]

$$\mu_m < 10^{-10}\ \mu_B. \tag{7.147}$$

A value

$$\mu_m < 10^{-12}\ \mu_B \tag{7.148}$$

has been derived from observations of the most recent supernova SN1987A [Bar88b, Gol88a, Lat88]. However, this presupposes that the dynamics of the star collapse are sufficiently well understood.

Laboratory detection of a magnetic moment would be most convincing. Possible approaches might include [Cli87]:

- Deviations from the predictions of the standard model in high-energy $\nu_e e^-$ and $\bar{\nu}_e e^-$ scattering experiments.
- Direct detection of the $\nu_e e^-$ scattering via the magnetic moment for small energies. Since the reaction cross section for an interaction due to the magnetic moment increases with $\ln E_\nu$ and that for weak interaction increases as E_ν^2 for energies of a few MeV and less, it is in principle possible to observe the magnetic moment of the neutrino at very low energies.

- Observation of the coherent interaction of neutrinos in matter via the magnetic moment.

Analogous to the case of the MSW effect, the probability of a spin reversal due to the magnetic moment may be increased in matter (see e.g. [Akh88, Lim88]).

7.3.7 Neutrino decay

In principle, an instability of the electron neutrino could also provide an explanation for the solar neutrino problem. However, over timescales of a few hundred seconds, corresponding to the neutrino propagation time from the Sun to the Earth, no significant decrease in the flux due to a decay is expected (see the discussion in chapter 6).

7.3.8 More recent experiments to detect solar neutrinos

Further measurements are needed to understand the underlying cause of the solar neutrino problem. Currently a number of experiments are being designed, in preparation or under way. An overview is given in [Kir88a,b, Man88, Bah89, Sin91, Bei91, Bor91, Arp94, McD94]. Here, we shall describe three experiments, two of which, the gallium experiments GALLEX and SAGE, have provided first results, while the third, the Sudbury Neutrino Observatory (SNO), is under construction. For further projects like Superkamiokande (see also chapter 4, figure 4.10), Borexino and others, we refer to [Suz92, 94, Gia94, McD94].

7.3.8.1 The gallium experiments

The use of gallium as a detector material for solar neutrinos was first proposed by Kuzmin [Kuz66]. Neutrinos are detected via the reaction

$$^{71}\mathrm{Ga} + \nu_e \rightarrow {}^{71}\mathrm{Ge} + e^-. \tag{7.149}$$

The isotope ^{71}Ga has a natural isotopic abundance of 39.9%. The daughter nuclide ^{71}Ge is radioactive and decays via electron capture ($T_{1/2} = 11.4$ days) back into gallium. This decay is recorded via the emitted Auger electrons (L peak: 1.2 keV; K peak: 10.4 keV). So, this is a radiochemical experiment. The decisive advantage over the chlorine experiment is the lower energy threshold of only $E_t = 0.233$ MeV for neutrinos; thus, a gallium detector is sensitive to most of the solar neutrino flux (see figure 7.19). In particular, this now makes it possible to measure the flux of *pp* neutrinos for the first time. This is very important when one comes to decide between the various possible explanations for the deficit of neutrinos, since the generation rate for *pp* neutrinos depends

only slightly on the central temperature of the Sun and cannot be reduced by a decrease in the temperature to the same extent as the rate for ^{8}B neutrinos.

If the neutrino flux found in the gallium detector were equally strongly suppressed, this would be evidence for the existence of neutrino oscillations with large mixing, i.e. for the MSW effect. Figure 7.33 shows an MSW diagram for the ^{71}Ga detector. Because of the extremely weak interaction of neutrinos with matter, large amounts of gallium are required. Only one capture per day is expected in 30 tonnes. Thus, the greatest experimental problems relate to the suppression of the background and the extraction of the few germanium isotopes formed during the exposure time.

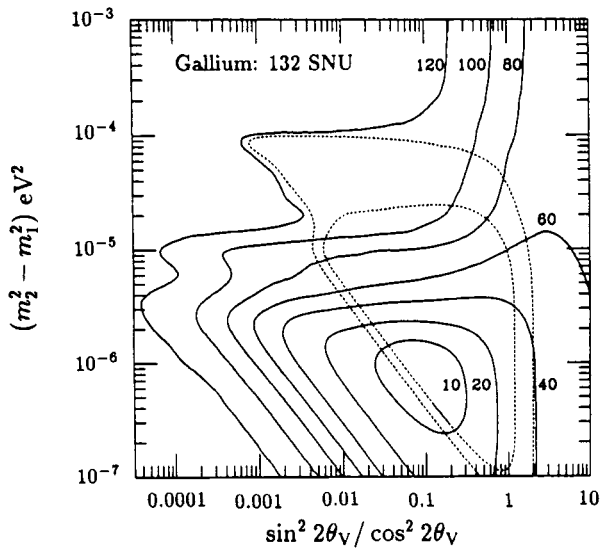


Figure 7.33. MSW diagram for ^{71}Ga experiments. The contour lines are calculated for various SNU values. The band between the dotted lines is defined by the lower and upper bounds from the chlorine experiment (2.1 ± 0.3 SNU) (from [Bah89]).

Precise knowledge of the neutrino-capture cross section for the reaction (7.149) is needed, in order to obtain an informative result. The capture of the low energy *pp* neutrinos leads mainly to the ground state in ^{71}Ga. The corresponding matrix element is known from the β^+ decay of ^{71}Ge. On the other hand, the capture rate at higher energies is decisively determined by the distribution of the Gamow–Teller strength in ^{71}Ge. This may be determined from (p, n) reactions [Kro85] or via nuclear structure calculations [Gro86, Kla86, Sta89]. Table 7.10 shows the expected capture rates for the ^{71}Ga detector, based on the SSM. The

Table 7.10. Capture rates for the gallium detector, according to the Standard Solar Model [Bah88].

Source	Capture rate [SNU]
pp	70.8
pep	3.0
^{3}He p	0.06
^{7}Be	34.3
^{8}B	14.0
^{13}N	3.8
^{15}O	6.1
^{17}F	0.06
$\sum$	132

total expected capture rate

$$\sum_i \Phi_i \sigma_i = 132^{+20}_{-17}\ \text{SNU} \tag{7.150}$$

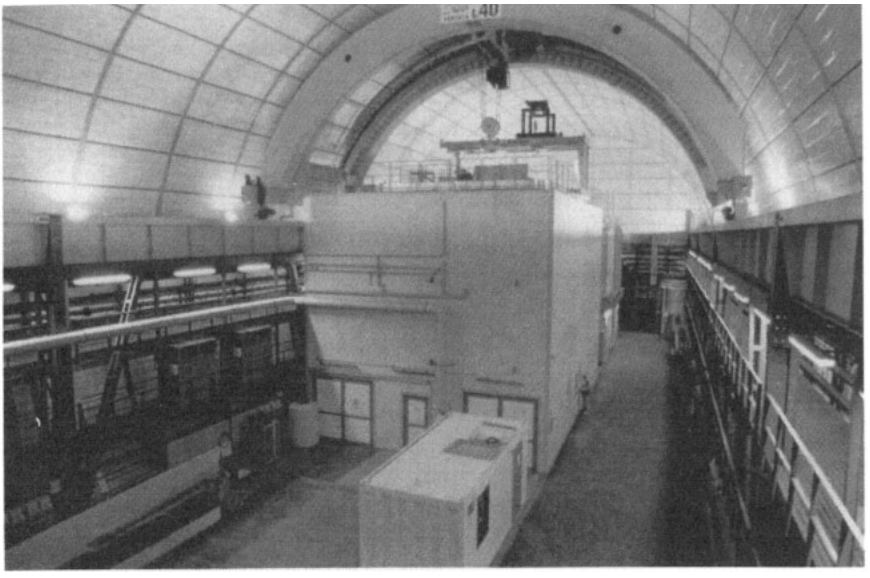
is very much higher than in the chlorine experiment.

Two gallium experiments have been designed and measurements started. These are the GALLEX experiment [Kir84, Ham85, 86a, b, 88a,b, Kir86b, Ans92, 93, 94, Kir95] and the SAGE experiment [Bar85, Aba88, 91, Ano92, Abd95]. Both are described in detail in [Bah89].

GALLEX This gallium experiment was essentially designed by a European collaboration. The detector is installed in the underground laboratory of the LNGS (Laboratorio Nazionale del Gran Sasso [Bel91]) in the Gran Sasso tunnel in the Abruzzi mountains 150 km to the east of Rome (figure 7.34). The shielding thickness of 1400 m of rock corresponds to that of approximately 3400 m of water.

The GALLEX collaboration uses 30 tonnes of gallium ($\simeq 1.03 \times 10^{29}$ ^{71}Ga atoms) in the form of an 8N aqueous gallium chloride solution ($GaCl_3$). This solution weighs 105 tonnes. The reaction product ^{71}Ge forms the slightly volatile germanium tetrachloride ($GeCl_4$) which is rinsed out of the tank after the exposure period using nitrogen. The $GeCl_4$ extracted is chemically converted into the gas GeH_4. The latter is finally passed with xenon as counting gas into a small proportional counter, in which the number of germanium atoms is detected via their radioactive decay.

The results after approximately one year of measurement (see figure 7.35)



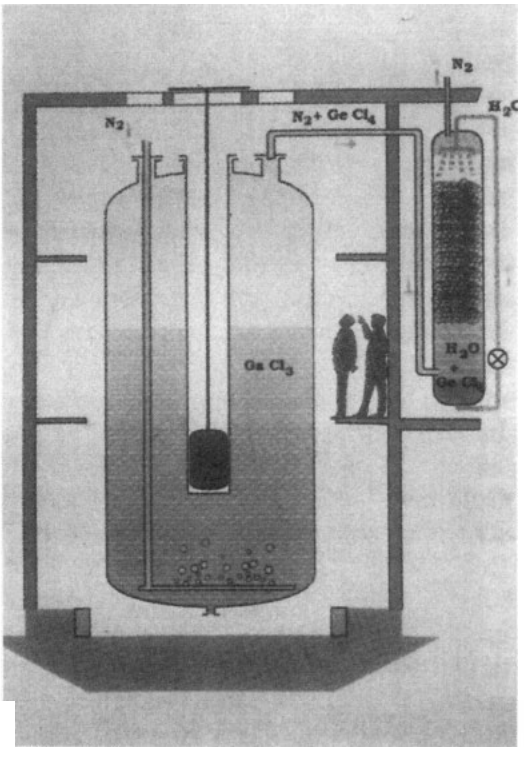


Figure 7.34. The GALLEX experiment in Hall A of the Gran Sasso underground laboratory (see figure 6.33). (*a*) In the foreground, the main building with the gallium tank and the extraction facilities, and, in the background, the counting-tube building; (*b*) section through the GALLEX tank (from [Kir93]).

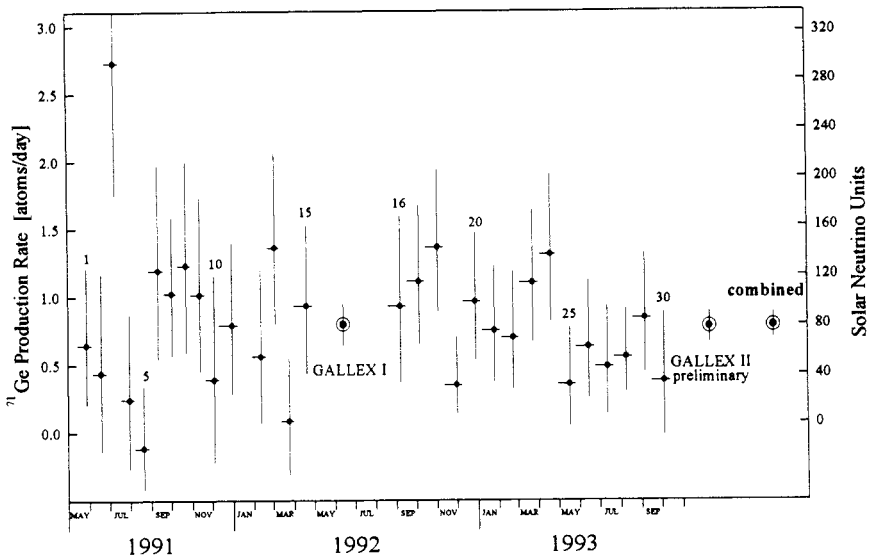


Figure 7.35. End result of all 15 solar neutrino measurements by GALLEX I (stripes to the left of May 1992) and result of the first 15 GALLEX II measurements (stripes to the right of August 1992). The left ordinate gives the production rate of the measured ^{71}Ge, the right, the nett Solar production rate (in SNU) after subtraction of contributions from side reactions. Error bars correspond to $\pm 1\sigma$. The point marked 'combined' corresponds to the global average of the maximum likelihood analysis of all 30 measurements. Horizontal bands show the duration of the individual measurements (from [Ans94]).

give a flux of

$$\sum_i \Phi_i \sigma_i = 83 \pm 19 \text{ (stat.)} \pm 8 \text{ (sys.) SNU.} \qquad (1\sigma) \qquad (7.151a)$$

The (total) result after two further years of measurement was [Ans93, 94]

$$\sum_i \Phi_i \sigma_i = 79 \pm 10 \text{ (stat.)} \pm 6 \text{ (sys.) SNU} \qquad (1\sigma). \qquad (7.151b)$$

This would imply the first observation of solar *pp* neutrinos. The result is consistent with the occurrence of the full expected *pp* flux (table 7.10) together with a reduced flux of ^{8}B and ^{7}Be neutrinos, according to the observations in the Homestake and Kamiokande experiments. It is only two standard deviations below the predictions of the SSM (124–132 SNU). Thus, this does not require neutrino oscillations. On the other hand, an interpretation within the framework of the MSW mechanism together with the chlorine and Kamiokande experiments would restrict the parameters Δm^2 and $\sin^2 2\theta$ to two narrow areas around

$\Delta m^2 = 6 \times 10^{-6}$ eV^2, $\sin^2 2\theta = 7 \times 10^{-3}$ and $\Delta m^2 = 8 \times 10^{-6}$ eV^2, $\sin^2 2\theta = 0.6$ (figure 7.36). The first of these solutions would lead, *assuming* a mass hierarchy of neutrinos as in some see-saw models, to $m_{\nu_\mu} \sim 3$ meV and $m_{\nu_\tau} \sim 10$ eV, which would make the ν_τ a candidate for dark matter (see chapter 9). This is one of the main motivations for the CHORUS and NOMAD experiments (see section 7.3.3 and [Win95]). However, it has already been stressed in section 1.6.4.2, that there are also models other than the see-saw mechanism which lead to degenerate neutrino masses and would make the solar neutrino results consistent with a ν_e mass in the 1 eV region (see, for example, [Lee94, Moh94, Pet94, Ioa94, Raf95, Cal95]).

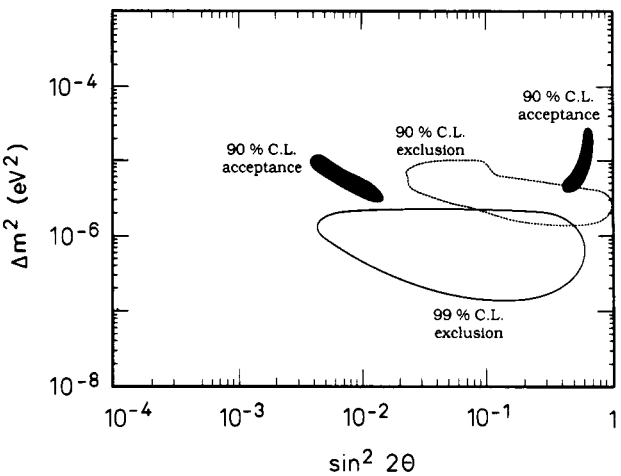


Figure 7.36. Diagram of Δm^2 against $\sin^2 2\theta$ for solar neutrino experiments. Parameters within the black areas enable us to explain the results of the ^{37}Cl, Kamiokande and GALLEX experiments within the Standard Solar Model (within 90% confidence limits). Values within the dotted line are excluded with 90% confidence by the Kamiokande experiment by studying the day–night effect. Values within the continuous curve are excluded with 99% confidence by the GALLEX I experiment (from [Ans92]).

The mechanism for neutrino decay proposed by [Fri88] would lead one to expect a value less than 45 SNU.

Neither does the GALLEX experiment point to magnetic interaction of the neutrino with the solar magnetic field, such as a spin flip due to an interaction with a magnetic moment (in the case of a Dirac neutrino) or a spin flip plus a flavour change in the case of a magnetic dipole interaction (in the case of a Majorana neutrino). Under both of the scenarios described in [Bab91a, Ono91], the result from the gallium detector should fluctuate between 75 and 80 SNU

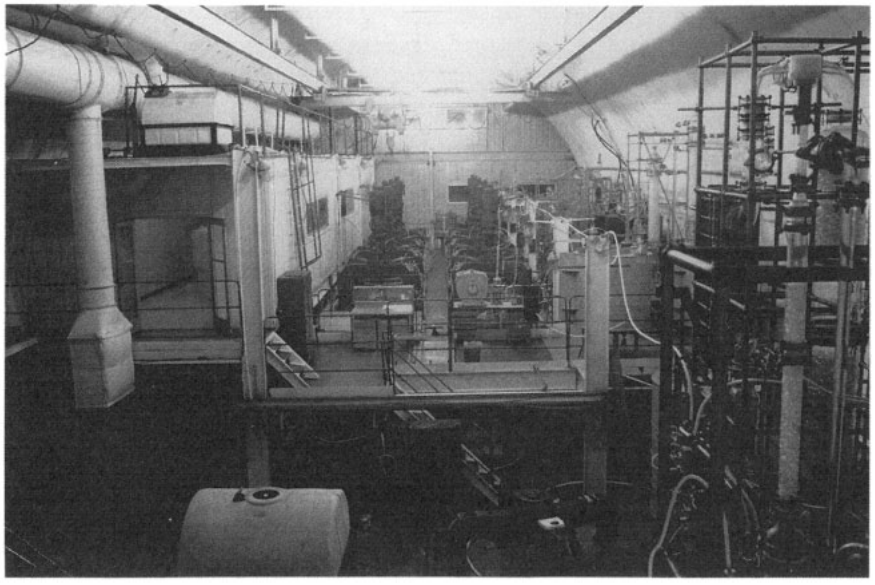


Figure 7.37. The SAGE experiment in the Baksan neutrino observatory (Caucasus). The figure shows the ten so-called reactors. Eight of these contain a total of 57 tonnes of metallic gallium. (Photo: Tom Bowles, Los Alamos, with permission of the SAGE collaboration.)

at the minimum and approximately 25 SNU near the maximum of the sunspot cycle. The latter value is scarcely consistent with the measured one.

SAGE. The other gallium experiment, SAGE (Soviet–American gallium experiment), is being carried out in the Baksan neutrino observatory in the northern Caucasus (shielding thickness: 4700 m water equivalent) (figure 7.37). In the end phase, the detector includes 57 tonnes of metallic gallium. Metallic gallium has the advantage of a small detector volume, which contributes to a decrease in the background. One large disadvantage is that the extraction of germanium is more difficult than in the case of a solution.

Gallium has a melting point of 30 °C. To extract the germanium, liquid gallium metal is mixed with thinned hydrochloric acid (HCl). Hydrogen peroxide (H_2O_2) is then added and the whole is well stirred. Unlike gallium, the GeH_4 formed dissolves. The solution is distilled and finally mixed with concentrated hydrochloric acid so that germanium tetrachloride is formed. The extraction and the detection then continue analogously to the GALLEX experiment.

The first (preliminary) data from the SAGE detector caused a great sensation [Bei91, Gav91, Aba91] (see also [Sch90a]). The recorded event rate was

consistent with the expected background. In addition, the decay of the ^{71}Ge should have exhibited an exponential dependence corresponding to a half-life of 11 days, while background events should be practically uniformly distributed. The SAGE data did not contain any component which could be said to be associated with a radioactive decay with a half-life of 11 days. As a result an event rate of

$$\sum_i \Phi_i \sigma_i = 20^{+15}_{-20} \text{ (stat.)} \pm 32 \text{ (sys.) SNU} \qquad (7.152a)$$

and an upper bound of

$$\sum_i \Phi_i \sigma_i < 79 \text{ SNU} \qquad (90\% \text{ c.l.}) \qquad (7.152b)$$

were initially obtained [Aba91, Ano92]. Meanwhile the published event rate is [Gav94, Abd95]

$$\sum_i \Phi_i \sigma_i = 74 \pm 13 \text{ (stat.)} \pm 7 \text{ (sys.) SNU} \qquad (90\% \text{ c.l.}). \qquad (7.152c)$$

In this context, the fact that the non-adiabatic MSW solution due to Bahcall and Bethe predicted a heavily reduced flux for gallium detectors may be commented upon. Only approximately 5 SNU were expected due to [Bah90]. This is because most low-energy electron neutrinos are converted into ν_x with $x \neq e$. A slightly larger value is obtained if one takes into account the regeneration of electron neutrinos due to oscillations in the vacuum and passage of the neutrinos through the Earth at night time. Baltz and Weneser [Bal91] pointed out, however, that [Bah90] only picked out the minimum solution. A more accurate consideration of the MSW effect, using the chloride and Kamiokande data allows capture rates of around 8–90 SNU in the gallium detector [Bal91].

7.3.8.2 The Sudbury Neutrino Observatory (SNO)

The Sudbury experiment [Ewa87, Aar87, Sin87, Bei91, McD94] uses a real-time Cerenkov counter which will begin to record data about the end of 1996. The detector consists of 1000 tonnes of highly pure heavy water (D_2O) in a transparent acrylic tank which is surrounded by 9600 photomultipliers. Around the acrylic tank there are another 7300 tonnes of water (H_2O). This detector is constructed at a depth of 2070 m (5900 m water equivalent) in the Creighton mine near Sudbury (Ontario) in Canada (figure 7.38). The muon intensity is a factor 200 less than that in the Kamioka mine in Japan. The most important parameters of the detector are summarized in table 7.11.

Table 7.11. The SNO detector.

Location	Creighton mine
Depth	2070 m
Water equivalent	5900 m
Detector material	1000 tonnes D_2O
Threshold	~ 5 MeV
Number of D atoms	6.02×10^{31}

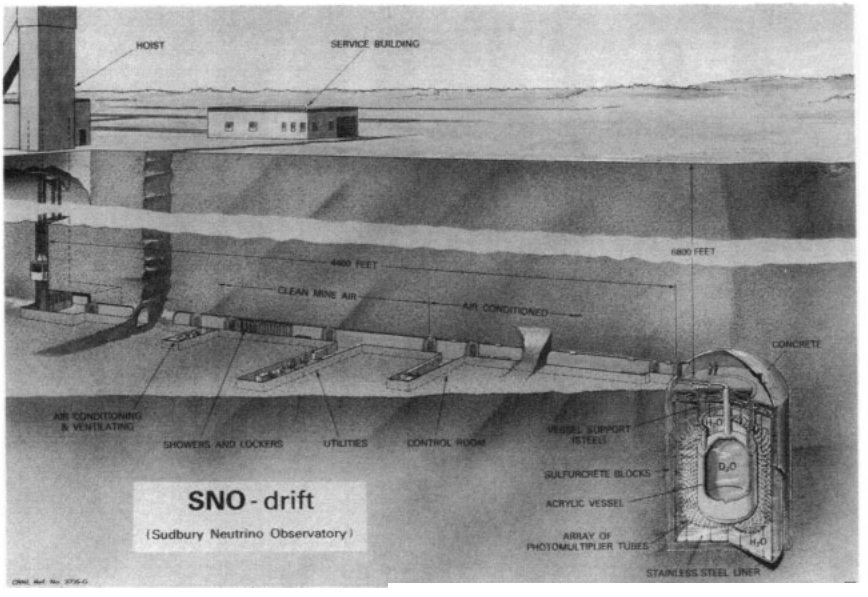


Figure 7.38. Early artist's conception of the Sudbury Neutrino Observatory (SNO) (status 1987). The acrylic tank filled with 1000 tonnes of heavy water (diameter $\sim$ 12 m) is surrounded by 7300 tonnes of water. The diameter of the cavity driven into the cliff is 20 m (from [Ewa87], with permission of the SNO collaboration).

The following three reactions will be used to detect neutrinos

$$\nu_e + d \rightarrow p + p + e^- \tag{7.153a}$$

$$\nu_x + e^- \rightarrow \nu_x + e^- \tag{7.153b}$$

$$\nu_x + d \rightarrow \nu_x + p + n. \tag{7.153c}$$

The reaction (7.153a) involves the charged weak current. Only electron neutrinos

may participate in it. The neutrino energy is related to the electron energy by

$$E_\nu = E_e + 1.442\ \text{MeV}. \tag{7.154}$$

The elastic scattering of neutrinos of all flavours by electrons has already been discussed in connection with the Kamiokande II detector. The event rate amounts to around 1/10 of the event rate due to absorption (7.153*a*). However, since the electrons are emitted in forward direction, the elastic scattering may be detected from the angular distribution.

Independently of the neutrino flavour, the total neutrino flux may be determined from the reaction (7.153*c*), which only involves the neutral weak current (threshold energy: 2.225 MeV). Such events are detected via the capture of the neutron in an appropriate nucleus and the subsequent γ emission. Comparison of the rates of the reactions (7.153*a*) and (7.153*c*) provides information about whether neutrino oscillations are responsible for the solar neutrino problem, independently of the Solar Model.

The SNO experiment offers for the first time the possibility of measuring the background directly – by replacing the heavy water by normal water so that the reactions (7.153*a*) and (7.153*c*) cease. The reaction (7.153*b*) may in principle be distinguished from the background by its directional dependence. The expected counting rate for the SNO detector of 10 events per day for each of reactions (7.153*a*) and (7.153*c*) is large in comparison with detectors currently in operation. As a result of the large statistics, one may hope for an explanation of the time dependence of the flux found in the chlorine experiment.

While the mechanism proposed by Bahcall and Bethe [Bah90] could have drastic consequences for the gallium experiments (see above), the effect on the SNO experiment would be comparatively small. The rates for reactions which are mediated by neutral weak currents would remain unchanged. The neutrino absorption and the neutrino–electron scattering are decreased by a factor of approximately three, depending on the threshold energy.

7.3.9 Atmospheric neutrinos

Atmospheric neutrinos are also used in oscillation experiments. The primary cosmic radiation interacts with nuclei (N) in the Earth's atmosphere, and thus generates complex hadronic showers. The atmospheric neutrinos occur as a result of decays of the unstable secondary particles, in particular of pions, kaons and muons. The production scheme is as follows

$$\begin{aligned} p + N &\rightarrow n + \pi/K + \ldots \\ \pi/K &\rightarrow \mu^+(\mu^-) + \nu_\mu(\overline{\nu}_\mu) \\ \mu^+(\mu^-) &\rightarrow e^+(e^-) + \nu_e(\overline{\nu}_e) + \overline{\nu}_\mu(\nu_\mu). \end{aligned} \tag{7.155}$$

The energy spectra of the neutrinos show a very rapid decrease

$$\frac{dN}{dE} \sim E^{-3.7}. \tag{7.156}$$

The ratios of the different neutrino flavours to one another may be estimated using the decay chain (7.155). We expect about twice as many muon neutrinos as electron neutrinos and an approximately equal number of neutrinos and antineutrinos. A number of calculations of the atmospheric neutrino flux are carried out in the energy region around 1 GeV (see e.g. [Bar89a, Gai94]).

Atmospheric neutrinos may be observed in well-shielded underground detectors, although the typical rate of one event per day per 3000 tonnes of detector material is very small. Because of their long flight paths, atmospheric neutrinos are very good as far as the search for oscillations in the region $\Delta m^2 > 10^{-4}$ eV is concerned, although because of the poor statistics this is true only for relatively large mixing angles. The large proton decay detectors, Kamiokande II, Fréjus and NUSEX, are (or were) able to distinguish between ν_e and ν_μ induced reactions, so that they could look for ν_e–ν_μ and ν_μ–ν_τ transitions.

What effects do we expect in the presence of oscillations? The measured atmospheric neutrino flux may differ in three ways from the calculated flux:

- Oscillations would alter the flavour composition of the flux.
- If the oscillation length is much larger than the height of the atmosphere, but smaller than the diameter of the Earth, then only neutrinos which pass through the Earth may have a measurable oscillation probability. This would be apparent in a distortion of the angular distribution.
- If the oscillation length of some of the neutrinos is of the order of magnitude of the flight path, we would be able to see a modulation of the energy distribution.

Let us consider, for example, the composition of the flux as determined by evaluation of the Fréjus experiment [Ber90b]. The measured quantity is the ratio of electron to muon events (e/μ). A direct observation of the transition into a tau neutrino is impossible, since approximately 90% of the atmospheric neutrino flux lies below the threshold for τ production. If the measured ratio (e/μ) is smaller than expected, this may be interpreted as ν_e–ν_τ or $\overline{\nu}_e$–$\overline{\nu}_\tau$ oscillation. An excessively large measured value may be explained by ν_μ–ν_e or ν_μ–ν_τ oscillations (or oscillations of the corresponding antineutrinos).

The Kamiokande collaboration [Hir88, 92a, Suz94] found a deficit of muon neutrinos in the low-energy region (< 1 GeV) and also in the multi-GeV area [Fuk94], which was interpreted as evidence for neutrino oscillations [Bar88c, Hid88, Lea88, Bug89] (see figure 7.39). The Fréjus collaboration, on the other hand, has found no such effect [Ber90b]. Figure 7.39 also shows the exclusion diagram obtained from the 200 or so events in the Fréjus detector. The bounds

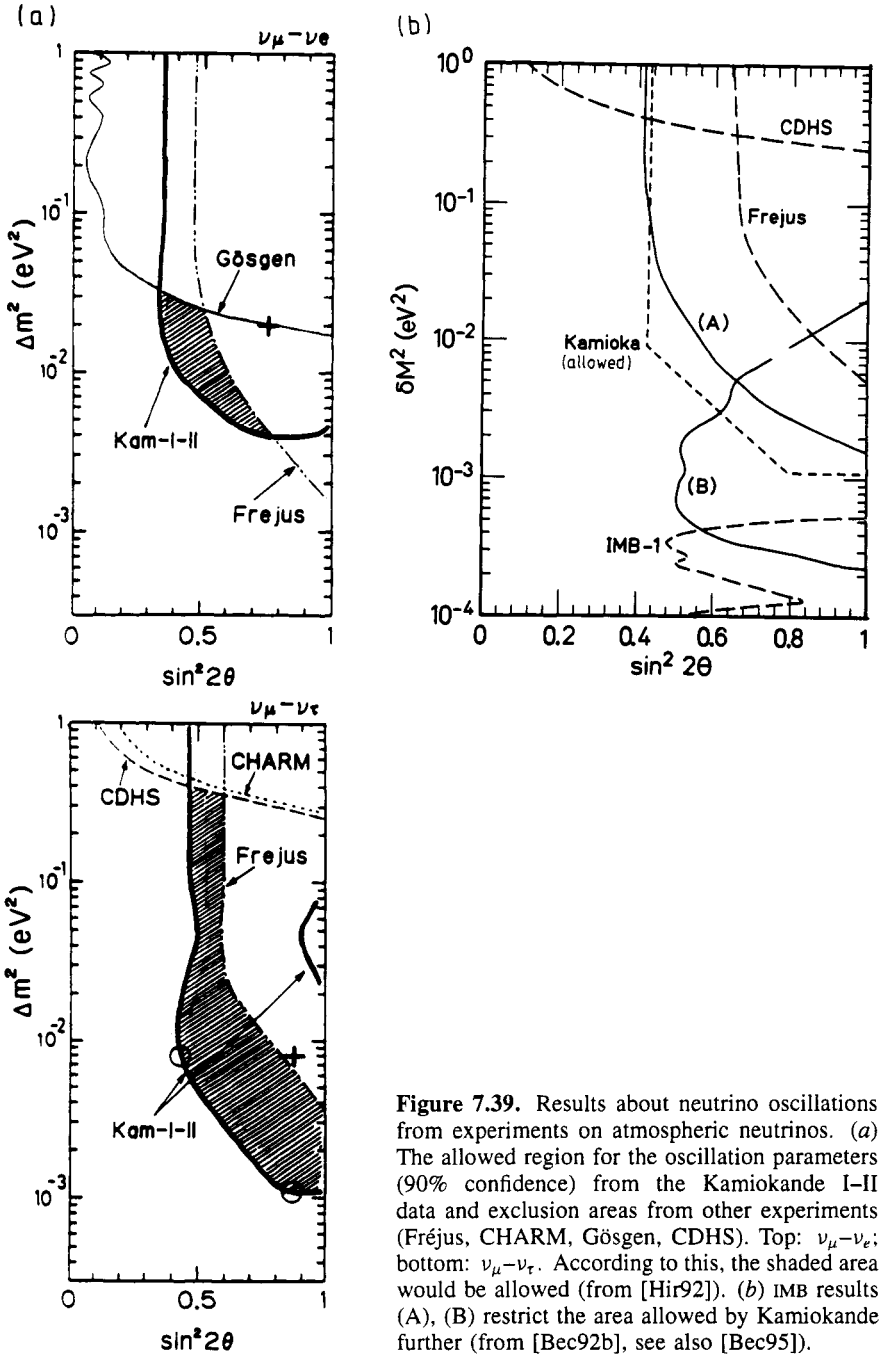


Figure 7.39. Results about neutrino oscillations from experiments on atmospheric neutrinos. (*a*) The allowed region for the oscillation parameters (90% confidence) from the Kamiokande I–II data and exclusion areas from other experiments (Fréjus, CHARM, Gösgen, CDHS). Top: ν_μ–ν_e; bottom: ν_μ–ν_τ. According to this, the shaded area would be allowed (from [Hir92]). (*b*) IMB results (A), (B) restrict the area allowed by Kamiokande further (from [Bec92b], see also [Bec95]).

obtained from the latter may be summarized as follows. For ν_e–ν_μ or $\overline{\nu}_e$– $\overline{\nu}_\mu$ oscillations

$$\begin{aligned} \Delta m^2 &\leq 1.5 \times 10^{-3}\ \text{eV}^2 \quad \text{(maximum mixing)} \\ \sin^2 2\theta &\leq 0.47 \qquad\qquad (\Delta m^2 \geq 1\ \text{eV}^2). \end{aligned} \tag{7.157a}$$

The corresponding values for the flavours μ and τ are

$$\begin{aligned} \Delta m^2 &\leq 3.5 \times 10^{-3}\ \text{eV}^2 \quad \text{(maximum mixing)} \\ \sin^2 2\theta &\leq 0.60 \qquad\qquad (\Delta m^2 \geq 1\ \text{eV}^2). \end{aligned} \tag{7.157b}$$

The data are not sufficient to provide information about the ν_e–ν_τ channel.

Despite the small number of events it is possible to sharpen the bounds for the mass parameters by a factor of 10 in comparison with accelerator experiments. The Fréjus result contradicts the Kamiokande measurement, but is consistent with NUSEX [Agl89].

IMB also shows no indication of ν_μ–ν_τ oscillations [Bec92b] and excludes the range 'allowed' by Kamiokande to a large extent (figure 7.39). The rest of the 'allowed' range of Kamiokande seems to be excluded by the Baksan experiments (see [Tot92]). These discrepancies will be solved only at a future date, when more data or a new generation of detectors (see e.g. [Lea93, Ben94]) will be available. Possibly this question will be decided only by terrestrial long-baseline neutrino oscillation experiments, for example, with ICARUS in the Gran Sasso (section 4.3.2.5) and a CERN SPS or LHC neutrino beam, or in reactor experiments like Chooz or San Onofre (see section 7.3.2 and [Vog95]).

Chapter 8

Magnetic Monopoles

8.1 INTRODUCTION, HISTORICAL OVERVIEW

In 1269, the French weapon-maker Petrus Peregrinus (Pierre de Maricourt) discovered that iron filings around lodestone were arranged along certain lines. The lines of force emanating from a magnetic body ended in the same body on the opposite side. This and all later observations confirmed the bipolarity of magnetism, i.e. that a magnet always has a south and a north pole. It seems impossible to separate the two poles. This property of magnetism contrasts with the existence of isolated electrical charges.

According to our current understanding of electromagnetism, the field of an electric dipole is generated by electrically charged particles of opposite polarity, while a magnetic dipole field is based on circular electrical currents. If we consider magnetism at the atomic level, the electrons circulating around an atomic nucleus have the effect of a circular current which can give rise to a magnetic moment. If the atoms are randomly oriented in a solid body, these atomic magnetic moments cancel one another out; the material is not magnetic. In a magnet, these elementary magnets are arranged in a preferred direction. However, since each atom is a magnetic dipole in its own right, it is impossible, according to this picture, to separate south and north poles from one another.

While the field equations of electrodynamics are symmetric as far as electrical and magnetic forces are concerned, the symmetry between electricity and magnetism is destroyed by the fact that isolated electrical charges exist but no analogous magnetic charges. Dirac first introduced magnetic monopoles into modern physics in 1931 [Dir31] in order to guarantee a far-reaching symmetry of the Maxwell equations. In particular, he was able to show that the existence of a particle with a magnetic charge g would automatically imply the quantization of the electrical charge according to the condition

$$eg = \tfrac{1}{2}n\hbar c \tag{8.1}$$

where n is an integer ($n \in \mathbb{Z}$). As far as quantum mechanics is concerned, the existence of magnetic monopoles is permissible, the charge of such a magnetic pole μ is an integer multiple of the elementary magnetic charge g

$$\mu = ng = n\frac{\hbar c}{2e} \simeq n\frac{137}{2}e. \tag{8.2}$$

A further step in the history of magnetic monopoles occurred in 1974, when the work of 't Hooft ['tHo74] and Polyakov [Pol74] pointed to a close relationship with modern elementary particle theories. Magnetic monopoles arise as stable solutions in spontaneously broken, non-Abelian gauge field theories as topological defects (see section 9.2.3.5). The breaking of the SU(5) group in SU(3) ⊗ SU(2) ⊗ U(1) requires, for example, the existence of magnetic monopoles with masses in the region of 10^{16} GeV/c^2. Magnetic charges also occur naturally [Per84] in the context of Kaluza–Klein theories, from which one hopes that they may permit a unification of the elementary forces including gravity (see chapter 12).

After this brief historical overview of the development of the concept of magnetic monopoles, we shall now discuss the underlying physical approaches in somewhat more detail. For overviews we refer also to [Car83, Sto84].

8.2 THEORETICAL CONCEPTS OF MAGNETIC MONOPOLES

8.2.1 The symmetry of the Maxwell equations

Initially, the question arises as to whether it is at all possible to determine unambiguously whether particles may have a magnetic charge in addition to an electric charge.

In what follows, we postulate the existence of a magnetic charge density ρ_m and a corresponding magnetic current density $\boldsymbol{j}_m$. These may be used to bring the Maxwell equations to a symmetric form (see the discussion in [Jac75])

$$\operatorname{div} \boldsymbol{D} = 4\pi\rho_e \tag{8.3a}$$

$$\operatorname{div} \boldsymbol{B} = 4\pi\rho_m \tag{8.3b}$$

$$\operatorname{curl} \boldsymbol{H} = \frac{1}{c}\frac{\partial \boldsymbol{D}}{\partial t} + \frac{4\pi}{c}\boldsymbol{j}_e \tag{8.3c}$$

$$-\operatorname{curl} \boldsymbol{E} = \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} + \frac{4\pi}{c}\boldsymbol{j}_m. \tag{8.3d}$$

The continuity equation should apply to both types of charge; thus, charge conservation should also apply

$$\frac{\partial \rho_e}{\partial t} + \operatorname{div} \boldsymbol{j}_e = 0 \qquad \frac{\partial \rho_m}{\partial t} + \operatorname{div} \boldsymbol{j}_m = 0. \tag{8.4}$$

It now appears that the existence of magnetic charges should lead to measurable physical effects. We see that this is not automatically the case when we carry out the following dual transformation

$$\boldsymbol{E} = \boldsymbol{E}' \cos\phi + \boldsymbol{H}' \sin\phi \qquad \boldsymbol{D} = \boldsymbol{D}' \cos\phi + \boldsymbol{B}' \sin\phi \tag{8.5a}$$

$$\boldsymbol{H} = -\boldsymbol{E}' \sin\phi + \boldsymbol{H}' \cos\phi \qquad \boldsymbol{B} = -\boldsymbol{D}' \sin\phi + \boldsymbol{B}' \cos\phi. \tag{8.5b}$$

For real mixing angles, this transformation leaves quadratic forms such as $\boldsymbol{E} \times \boldsymbol{H}$ and $(\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{B} \cdot \boldsymbol{H})$ unchanged. In addition, the components of the Maxwell stress tensor $T_{\mu\nu}$ are invariant under (8.5).

Simple calculations show that the extended Maxwell equations (8.3) are form-invariant under the transformation (8.5) if the sources are transformed analogously.

$$\rho_e = \rho_e' \cos\phi + \rho_m' \sin\phi \qquad j_e = j_e' \cos\phi + j_m' \sin\phi \tag{8.6a}$$

$$\rho_m = -\rho_e' \sin\phi + \rho_m' \cos\phi \qquad j_m = -j_e' \sin\phi + j_m' \cos\phi. \tag{8.6b}$$

The equations of (8.3) again hold for the field quantities with the prime sign, where the sources should also be replaced by the quantities with the prime sign.

Because the fundamental equations of electrodynamics are invariant under these dual transformations, it is largely a question of convention as to whether one assigns a particle a magnetic in addition to an electrical charge. Provided the ratio of the magnetic charge to the electrical charge is the same for all particles, the angle ϕ which defines the transformation may be chosen so that ρ_m and j_m vanish. For this, one defines ϕ_0 such that

$$\rho_m = \rho_e' \left(-\sin\phi_0 + \frac{\rho_m'}{\rho_e'} \cos\phi_0 \right) = 0. \tag{8.7}$$

Whence the components of the current density satisfy

$$\begin{aligned} j_{m_i} &= j_{e_i}' \left(-\sin\phi_0 + \frac{j_{m_i}'}{j_{e_i}'} \cos\phi_0 \right) \\ &= j_{e_i}' \left(-\sin\phi_0 + \frac{\rho_m'}{\rho_e'} \cos\phi_0 \right) = 0. \end{aligned} \tag{8.8}$$

For this particular choice of the angle, the extended Maxwell equations become the well-known Maxwell equations. According to the normal convention, the charges of the electron are chosen such that

$$q_e^e = -e \qquad q_m^e = 0. \tag{8.9}$$

The electrical charge of the proton is then given, for example, relative to that of the electron by $q_e^p = +e$ with an experimental uncertainty of

$|q_e^e + q_e^p| < 1 \times 10^{-21} e$ [Dyl73]. For the corresponding magnetic charges, we find $||q_m^p| - |q_m^e|| \leq 10^{-26} g$ [Van68]. The magnetic charge of the electron itself was determined to be $|q_m^e| < 4 \times 10^{-24} g$ in this experiment.

The question then arises as to whether the ratio of the magnetic to the electrical charge is the same for all particles. The Maxwell equations in the generalized form need only be discussed if there exist particles with different electric to magnetic charge ratios.

8.2.2 The Dirac quantization condition

Dirac investigated the behaviour of an electron in the presence of a magnetic monopole [Dir31, 48]. He was able to show that it is possible to quantize the equations of motion only under the assumption that the charges are multiples of the elementary charges e and g, where condition (8.1) must be satisfied. Thus, the discrete nature of the electrical charge follows from the existence of a magnetic monopole. The quantity e is not determined directly, but expressed in terms of the (unknown) magnetic charge g.

On the other hand, since the fine structure constant $\alpha = e^2/\hbar c \simeq 1/137$ is known, we may deduce the existence of a magnetic monopole, the charge of which should be

$$g = \frac{n}{2}\frac{\hbar c}{e} = \frac{n}{2}\frac{e}{\alpha} \simeq n\frac{137}{2}e. \tag{8.10}$$

According to equation (8.10), the magnetic fine structure constant α_m has the value

$$\alpha_m = \frac{g^2}{\hbar c} = \frac{n^2}{4}\left(\frac{\hbar c}{e^2}\right) \simeq \frac{137}{4}n^2. \tag{8.11}$$

It follows from (8.10) that magnetic monopoles with charge g should exercise approximately the same force as an electrically charged body with charge $137n/2$. This means, in particular, that because of the very large coupling constants, magnetic monopoles could experience a considerable acceleration even in weak fields. Thus, they should be relatively easy to detect.

We shall now attempt to understand the relationship between charge quantization and the existence of magnetic monopoles. The following discussion is oriented towards the semiclassical derivation in [Gol65]. We shall consider the deflection of a charged particle with charge e and mass m in the field of a monopole. The collision parameter b is assumed to be sufficiently large that there exists a good straight line approximation to the path of the particle (see figure 8.1). The magnetic monopole is assumed to be located at the origin of the coordinate system. The radial magnetic field at the point $\boldsymbol{r}$ is given by

$$\boldsymbol{B}(\boldsymbol{r}) = g\frac{\boldsymbol{r}}{r^3}. \tag{8.12}$$

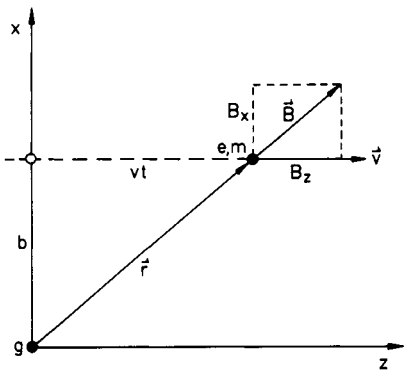


Figure 8.1. Flight of a charged particle with charge e and mass m past a magnetic monopole with charge g with a large collision parameter p (after [Jac75]).

The charge, which moves with velocity $v = v\hat{e}_z$ in the z direction is subject to a Lorentz force in the y direction

$$F = F_y\hat{e}_y = \frac{ev}{c}B_x\hat{e}_y. \tag{8.13}$$

It is easy to deduce the component B_x of the magnetic field using figure 8.1; it is given by

$$B_x = \frac{bg}{(v^2t^2 + b^2)^{3/2}}. \tag{8.14}$$

Thus, the y component of the Lorentz force is

$$F_y = \frac{eg}{c}\frac{vb}{(v^2t^2 + b^2)^{3/2}}. \tag{8.15}$$

The momentum transmitted to the particle is given by

$$\begin{aligned}\Delta p_y &= \int_{-\infty}^{\infty} F_y \mathrm{d}t \\ &= \frac{egvb}{c}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{(b^2 + v^2t^2)^{3/2}} \\ &= \frac{2eg}{cb}.\end{aligned} \tag{8.16}$$

This change in the momentum means that the particle is deflected out of the plane of figure 8.1. There is an associated change in the angular momentum

$$\Delta L_z = b\Delta p_y = \frac{2eg}{c}. \tag{8.17}$$

ΔL_z is independent of the collision parameter and of the particle velocity. It is a universal value for a charged particle flying past a magnetic monopole at an arbitrary distance.

From quantum mechanics, we know that the orbital angular momentum is quantized

$$L_z = n\hbar \quad n \in \mathbb{Z}. \tag{8.18}$$

This immediately leads to the condition

$$\frac{eg}{\hbar c} = \frac{n}{2} \tag{8.19a}$$

which implies a quantization of the electrical charge

$$e = \frac{n}{2}\left(\frac{\hbar c}{g}\right). \tag{8.19b}$$

The existence of a single magnetic monopole is sufficient as far as these considerations are concerned.

In what follows, we shall discuss a somewhat different approach to the derivation of equation (8.1). The magnetic field of a monopole of strength g is given by

$$\boldsymbol{B} = g\frac{\boldsymbol{r}}{r^3} = -g\,\mathrm{grad}\left(\frac{1}{r}\right). \tag{8.20}$$

Since $\nabla^2 1/r = -4\pi\delta^3(r)$, it follows that

$$\mathrm{div}\,\boldsymbol{B} = 4\pi g\delta^3(r). \tag{8.21}$$

Because the magnetic field is radially symmetric, the flux through the surface of a sphere enveloping the monopole is calculated as

$$\Phi = 4\pi r^2 B = 4\pi g. \tag{8.22}$$

We now consider an electrical charge e in this field. A free particle is described by the wavefunction

$$\psi = \psi_0 \exp\left(\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}\boldsymbol{r} - Et)\right). \tag{8.23}$$

In the presence of the electromagnetic field, according to the usual rules, the momentum $\boldsymbol{p}$ is replaced by

$$\boldsymbol{p} \to \boldsymbol{p} - \frac{e}{c}\boldsymbol{A} \quad \text{(minimum coupling)} \tag{8.24}$$

where $\boldsymbol{A}$ denotes the vector potential of the magnetic field. We obtain

$$\psi \to \psi \exp\left(-\frac{\mathrm{i}e}{\hbar c}\boldsymbol{A}\boldsymbol{r}\right) \tag{8.25}$$

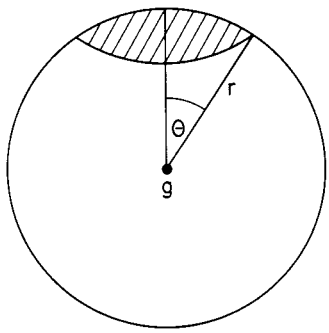


Figure 8.2. On the quantization of the electrical charge (see text).

i.e. the phase of the wavefunction changes

$$\alpha \to \alpha - \frac{e}{\hbar c} \boldsymbol{A}\boldsymbol{r}. \tag{8.26}$$

For a closed path with fixed coordinates r, θ and ϕ, the phase change amounts to

$$\begin{aligned} \Delta\alpha &= \frac{e}{\hbar c} \oint \boldsymbol{A}\,\mathrm{d}\boldsymbol{l} = \frac{e}{\hbar c} \int (\operatorname{curl} \boldsymbol{A})\,\mathrm{d}\boldsymbol{f} = \frac{e}{\hbar c} \int \boldsymbol{B}\,\mathrm{d}\boldsymbol{f} \\ &= \frac{e}{\hbar c} \Phi(r, \theta). \end{aligned} \tag{8.27}$$

In figure 8.2, $\Phi(\theta, r)$ is the flux through the hatched surface. As $\theta \to 0$, the surface shrinks to a point and the flux through the 'sphere cap' vanishes ($\Phi(r, 0) = 0$). As $\theta \to \pi$, according to (8.22)

$$\Phi(r, \pi) = 4\pi g \tag{8.28}$$

since the hatched surface now covers the whole of the surface of the sphere. On the other hand, as $\theta \to \pi$, the closed curve again contracts to a point, at which the vector potential has a singularity. The phase is not defined there. To ensure that the wavefunction is well-defined, the condition

$$\Delta\alpha = n2\pi \quad n \in \mathbb{Z} \tag{8.29}$$

must be satisfied, i.e. we again obtain the known result

$$n2\pi = \frac{e}{\hbar c} 4\pi g \to eg = \frac{1}{2} n\hbar c. \tag{8.30}$$

The Dirac considerations have nothing further to say about the properties of the magnetic monopole. However, we may give a rough argument of plausibility

about the mass we may expect it to have. Suppose that r_e is the classical electron radius $r_e = e^2/m_e c^2 = 3 \times 10^{-15}$ m. Assuming that the corresponding monopole radius is of the same order of magnitude, it follows that

$$\frac{g^2}{m_M} \approx \frac{e^2}{m_e} \rightarrow m_M \approx \left(\frac{g}{e}\right)^2 m_e \approx 4700 m_e. \tag{8.31}$$

Thus, the mass of the particle introduced by Dirac might be in the region of a few GeV/c^2.

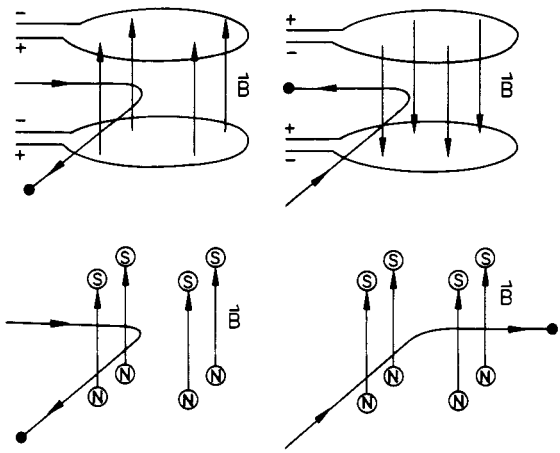


Figure 8.3. Violation of time-reversal invariance by magnetic monopoles (see text) (after [Car82]).

For a long time there was a serious objection to the possible existence of magnetic monopoles, since the latter violate the time-reversal invariance. We shall explain this using figure 8.3. The upper part of the figure shows the situation in which a proton enters the magnetic field between two loops with current passing through, and is consequently deflected perpendicularly to the field direction. If the time direction is reversed the proton moves backwards. However, since the current in the loops and thus the induced magnetic field also reverse their direction, the proton follows the original path exactly. The reversal of the temporal evolution again describes a process which is physically possible. This invariance is no longer satisfied for a field which is generated by isolated magnetic monopoles (see figure 8.3(*b*)). Since the polarity of a magnetic monopole is not reversed with the time direction, the magnetic field remains unchanged. The path of the backwards-moving proton is no longer the original path. Thus, there is a violation of time-reversal invariance. However, since the discovery of CP violation in the K^0 system, which, according to the

CPT theorem, is equivalent to an indirect T violation, this objection to the possible existence of monopoles can no longer be justified.

Until 1975 experimental efforts essentially concentrated on light relativistic magnetic monopoles ($mc^2 < 10$ GeV) (see [Cra86]); however, the search has been unsuccessful. The monopole hypothesis received a new impetus through the work of 't Hooft and Polyakov [Pol74, 'tHo74] in 1974, since the authors were able to show that magnetic monopoles appear to be a natural consequence in the framework of GUT models.

8.2.3 GUT monopoles

8.2.3.1 GUT *monopoles as topological defects*

't Hooft and Polyakov ['tHo74, Pol74] found that gauge theories predict the existence of isolated magnetic poles as a physical necessity. However, these objects must have a far greater mass than had previously been supposed. This could explain why the intensive search for magnetic monopoles had failed until then.

We saw that Maxwell's electrodynamics, which is based on the Abelian gauge group U(1), may be extended by the introduction of an isolated magnetic charge so that far-reaching symmetry between electricity and magnetism can be achieved. But there is no need to introduce magnetic monopoles. However, if one extends the gauge symmetry to a non-Abelian group such as SU(5), the spontaneous symmetry breaking leads to a stable solution of the field equations, which exhibits the properties of a magnetic monopole. This means that the existence of this particle is required in the framework of current model schemes.

According to these modern theories, magnetic monopoles arise as topological defects of space–time, precisely when a large local gauge group is spontaneously broken and the Abelian group U(1) occurs for the first time as a subgroup. In particular, this phenomenon is found in connection with the phase transition

$$\mathrm{SU}(5) \rightarrow \mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1) \quad (E \sim 10^{15}\ \mathrm{GeV}). \tag{8.32}$$

Here, magnetic monopoles occur as zero-dimensional defects of the Higgs field. The mass is closely related to the breaking energy (see below). Further breaking of $\mathrm{SU}(2) \otimes \mathrm{U}(1)$ into U(1) does not lead to further monopoles associated with the electroweak energy scale of 100 GeV [Jaf80].

The general picture is as follows. The SU(5) symmetry breaking began at typical temperatures of 10^{28} K, when the 24-dimensional Higgs field in the SU(5) space oriented itself analogously to the orientation of a Weiss domain of a ferromagnet when the temperature decreases below the Curie temperature (see chapter 1). Since the orientations are not specified for such a phase transition,

spatial domains with various orientations must have developed. On the bounding surfaces of areas with different SU(5) breakings, topologically stable defect points developed which carry a magnetic charge corresponding to that of the Dirac monopole.

8.2.3.2 *The mass of* GUT *monopoles*

The mass of the GUT monopole is correlated with the temperature of the phase transition

$$m_M \simeq \frac{m_X}{\alpha_5} \simeq 10^{16}\ \mathrm{GeV}/c^2. \tag{8.33}$$

Here, m_X is the mass of the X bosons and α_5 denotes the strong coupling constant (analogous to the electromagnetic fine structure constant) in the region of the unification energy. The mass m_M is enormously large for an '*elementary particle*', 10^{16} GeV/c^2 corresponds to 10^{-8} g, i.e. a monopole would be approximately as heavy as a bacterium. Because of this enormous mass, it will hardly be possible to generate such objects using particle accelerators.

What velocity are GUT monopoles expected to have [Gro86b, Ric87]? A gas of monopoles, without interaction with external fields would have cooled to around 10^{-5} mK during the evolution of the universe, i.e. the thermal velocity would be less than $10^{-21}\,c$. It is conjectured that the current movement of superheavy monopoles is determined by galactic gravitational and magnetic fields. For masses above 10^{17} GeV/c^2, gravity dominates and the velocities should be close to the galactic virial velocity of 200 km s^{-1} ($\simeq 10^{-3}\,c$). For lighter monopoles, the effect of the magnetic fields dominates and the velocities are somewhat higher, with

$$\beta = \frac{v}{c} \sim 10^{-3}\left(\frac{10^{17}\ \mathrm{GeV}}{m_M c^2}\right)^{1/2}. \tag{8.34}$$

In gauge groups larger than the SU(5), other possibilities for the charge and the mass of magnetic monopoles exist. In particular, lighter monopoles with $m_M \sim 10^4$–10^{10} GeV/c^2 may occur. m_M is larger in supersymmetric GUT models than in the corresponding simple theories since the energy scale of the spontaneous symmetry breaking in (8.33) is higher as a result of the greater number of fundamental particles. Finally, monopole masses in the region of the Planck mass of 10^{19} GeV/c^2 and over are expected in Kaluza–Klein theories (see [Wit81]).

8.2.3.3 *The structure of the* GUT *monopoles*

GUT monopoles are not point-like but have a relatively complicated structure which is shown schematically in figure 8.4. They have an onion-like structure

and contain practically the whole particle spectrum of GUT models. Near the centre ($r_c \sim 1/m_X \sim 10^{-29}$ cm) there is a GUT-symmetric vacuum. Outside this is the region of the electroweak unification. The next shell is called the confinement shell and contains photons and gluons. On the outer edge there is a region with fermion–antifermion pairs. At large distances (r greater than several fm) this construction behaves like a Dirac-like monopole.

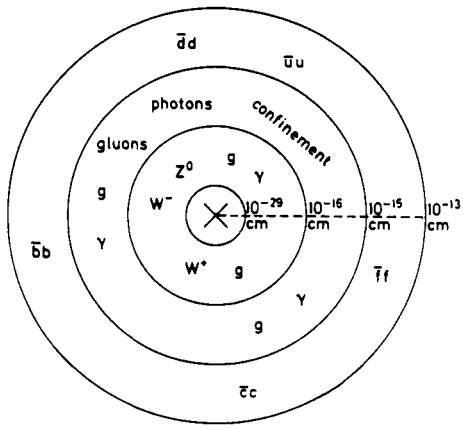


Figure 8.4. Onion structure of a GUT monopole. The inner GUT core is surrounded by other shells: 1 the region of the electroweak unification, 2 the confinement region, 3 the fermion–antifermion condensate (from [Bör88]).

Because of their complicated structure, GUT monopoles may catalyse a decay of the proton[1] (Rubakov–Callan effect) [Rub81, Cal82, Tro83]. A collision between a proton and a magnetic monopole may lead to an interaction of the quarks of the proton with the inner GUT zone of the monopole. The presence of the X and Y bosons causes a rapid decay of the corresponding proton. One possible reaction would be, for example, (M = monopole)

$$M + p \rightarrow M + e^+ + \pi^0. \tag{8.35}$$

The reaction cross section for a catalysed proton decay is [Ell82]

$$\sigma = \frac{c}{v}\left(\frac{\hbar c}{1\ \text{GeV}}\right)^2 \sigma_0 = 0.4\frac{c}{v}\sigma_0\ [\text{mb}]. \tag{8.36}$$

Here, v is the relative velocity between the proton and the monopole. Unfortunately, the factor σ_0 is only known very inaccurately. According to the original work it should have order of magnitude O(1) to O(100). Thus,

[1] Proton decay is discussed in detail in chapter 4.

the reaction cross section for such a process would typically be in the region of reactions of the strong interaction. However, other authors give very much smaller values for σ_0 (see e.g. [Gro86b]). It also appears possible that the Rubakov–Callan effect may vanish completely, at least for the SU(5) gauge group [Wal84]. Clearly, the theoretical situation is fraught with considerable uncertainty.

A magnetic monopole which passes through a proton decay detector may (for sufficiently large σ_0) induce around 10 decays. Since no such events have yet been observed it is in principle possible to derive an upper bound for the flux of monopoles (an overview is given in [Err84]). The IMB Cerenkov water counter has been used to study reactions of the form M + nucleon $\rightarrow M + e^+ +$ mesons. The bound for the flux, derived from the non-observation is [Err83]

$$F_M < 7.2 \times 10^{-15}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1}. \tag{8.37a}$$

However, astronomical objects such as neutron stars are more suitable for this. Magnetic monopoles are captured by these provided they have non-relativistic velocities, and their masses are less than about 10^{21} GeV/c^2 [Dim82, Kol82, 84a, Fre83]. They act as catalysts for the decay of the neutrons which would provide a considerable contribution to the luminosity of the star, particularly in the X-ray region. The measured values of the X-ray luminosity of neutron stars give upper bounds of

$$F_M f_0 < (10^{-22}\text{–}10^{-24})\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \tag{8.37b}$$

for the monopole flux [Gro86b], where the factor f_0 reflects the theoretical uncertainties in σ_0 from equation (8.36) ($f_0 = 4\sigma_0$). For a detailed discussion of the problem of catalysed proton decay and the uncertainties in the bounds derived for F_M as a result of non-observation, readers are referred to [Kol84, Gro86b].

8.2.4 The abundance of magnetic monopoles in the universe

Because of their extremely large mass, magnetic monopoles can only have been formed at a very early stage after the big bang, when the temperature was still sufficient to create such particles. The original considerations, which were based on a combination of the unified theory of the weak, electromagnetic and strong forces and the cosmological standard model, gave unbelievably high numerical densities n_M of magnetic monopoles in the universe. According to these considerations n_M should have been of the same order of magnitude as the baryon density. However, this contradicts the following consideration.

A simple estimate of the abundance of GUT monopoles may be obtained if we assume that these massive objects are responsible for the dark matter in the

universe. This is conceivable, since monopoles can emit very little light. Let us assume that such a monopole is 10^{16} times as heavy as a nucleon. From the condition that the monopole mass density $\rho_M = n_M m_M$ should not be greater than the critical mass density ρ_c of the universe, we deduce an approximate frequency of one monopole to 10^{15} protons. This may be reformulated to give a bound for the flux [Bör88]

$$F_M < 5 \times 10^{-15}(\beta \times 10^3)\left(\frac{10^{16}\ \text{GeV}}{m_M c^2}\right)\text{cm}^{-2}\ \text{s}^{-1}. \tag{8.38}$$

where $\beta = v/c$ is the typical monopole velocity. For heavy monopoles $\beta \sim 10^{-3}$.

The monopole density n_M is naturally related to the coherence of the Higgs field since monopoles are defined to be defects in this field. It is possible to give a simple lower bound for n_M if we assume that there exists at least one monopole per spatial domain [Kib76, Gut81, Gro89, 90]. If l denotes the diameter of a typical spatial domain then

$$n_M \propto l^{-3}. \tag{8.39}$$

The contribution to the mass density of the universe is

$$\rho_M = m_M n_M \propto m_M l^{-3}. \tag{8.40}$$

According to the experimental findings the current density of matter ρ_0 of the universe (for $\Lambda = 0$) lies in the region [Wei72, Mis73, Blo84]

$$\frac{1}{10}\rho_c \leq \rho_0 \leq 10\rho_c. \tag{8.41}$$

Thus, the monopole mass density must also satisfy the condition

$$\rho_M \leq \rho_c. \tag{8.42}$$

On the other hand, l cannot be larger than the event horizon at the time of the symmetry breaking, since each spatial domain must be causally related. This leads to an upper bound for l which depends on m_M. Conversely, by virtue of equations (8.40) and (8.42) and the causality requirement, we have

$$m_M \leq 10^{10}\ \text{GeV}/c^2 \tag{8.43}$$

i.e. a phase transition at the typical energy for the SU(5) breaking (10^{15} GeV) would result in an excessively large monopole density according to this rough estimate.

Inflationary models avoid the large number of monopoles, since the concentration is greatly decreased during the exponential expansion phase. It is

conceivable that the current monopole density is so small that we could never observe it (see [Pre84]).

We have already seen that the catalytic proton decay provides another way of obtaining information about the flux F_M. Another important bound for the flux (the *Parker limit*) follows from astrophysical observations. Parker used the existence of the galactic magnetic field to derive a largely model-independent bound for the flux of slow monopoles [Par70, 84]. The galactic B field of approximately 2–3 μG may be detected, for example, via the Zeeman splitting of the 21 cm line of the hydrogen atom or by observation of synchrotron radiation. A background flux of magnetic monopoles would attenuate this magnetic field. If there were too many monopoles, they would gradually destroy it. However, since the field exists, the total number of monopoles must be less than a certain value. A bound for the flux F_M may be derived (see below) by comparing the decay time of the B field and the regeneration time $\tau_R \simeq 10^8$ years (τ_R corresponds approximately to the period of the galaxy). An analogous effect is responsible for the fact that the galactic electrical field is extremely small. It is greatly attenuated by the work expended to accelerate charged particles.

A free magnetic monopole with charge $g \simeq 137/2e$ would be strongly accelerated in the galactic magnetic field. The current density which essentially runs in the direction of $\boldsymbol{B}$ may be described as follows

$$j_M = g n_M v. \tag{8.44}$$

The work which the field expends on the monopoles leads to a decrease in the energy density of the field $\omega = B^2/8\pi$. This decrease amounts to

$$P_- = \boldsymbol{B} \cdot j_M = B n_M g v = B g F_M \tag{8.45}$$

per unit time.

As a result of the movement of the interstellar gas, the galactic magnetic field may be renewed on a timescale of $\tau_R = 10^8$ years by virtue of an effect similar to the dynamo principle. The consequent increase in the energy density of the field amounts to

$$P_+ = \frac{B^2}{8\pi\tau_R}. \tag{8.46}$$

per unit time. The existence of the galactic field now implies that the outgoing energy should not be less than the incoming energy. Consequently, the flux of the magnetic monopoles F_M is restricted by the condition

$$F_M g B \leq \frac{B^2}{8\pi\tau_R}. \tag{8.47}$$

This leads to the *Parker limit*[2]

$$F_M \leq \frac{B}{8\pi g \tau_R} \simeq 1.3 \times 10^{-15}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}. \tag{8.48}$$

This limiting value corresponds to a very small flux of approximately one monopole per year and surface area of 2500 m^2.

These bounds hold for monopoles with a mass of 10^{16} GeV/c^2 which are relatively strongly accelerated in the galactic field. For larger masses in the range from 10^{19} to 10^{20} GeV/c^2 the bounds are a factor of 100 larger. If no regeneration mechanism for the galactic magnetic field is present, the time factor in equation (8.46) then amounts to 10^{10} years and a bound of

$$F_M \leq 10^{-17}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1} \tag{8.49}$$

may be derived.

8.3 PRINCIPLES FOR THE DETECTION OF MAGNETIC MONOPOLES

The direct detection of these exotic particles attracts particular attention. The experimental search for heavy magnetic monopoles in the cosmic radiation involves a number of different techniques including normal and superconducting coils, scintillation counters, counting tubes and 'track-etch' procedures. These detectors may be divided into two classes:

(i) Induction detectors which are based on the electromagnetic induction due to a monopole flying through an electrical loop.
(ii) Detectors which are based on the interaction of a monopole with matter as the former passes through the detector.

In what follows we shall briefly describe the individual techniques before discussing the experimental results.

8.3.1 Induction techniques

This method is based solely and exclusively on Faraday's law of induction and measures the magnetic charge directly. A very nice overview may be found, for example, in [Fri84]. The use of the magnetic induction in a superconducting ring to detect magnetic monopoles was first proposed by Alvarez [Alv63] and Tassie [Tas65]. Figure 8.5 is a schematic illustration of such a monopole detector. The idea was to pass various materials several times through the ring and to measure

[2] A stricter limit of $F_M \leq 10^{-16}$ (m/10^{17} GeV/c^{-2}) $\mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1}$ is given by [Ada93] (so-called *extended Parker limit*).

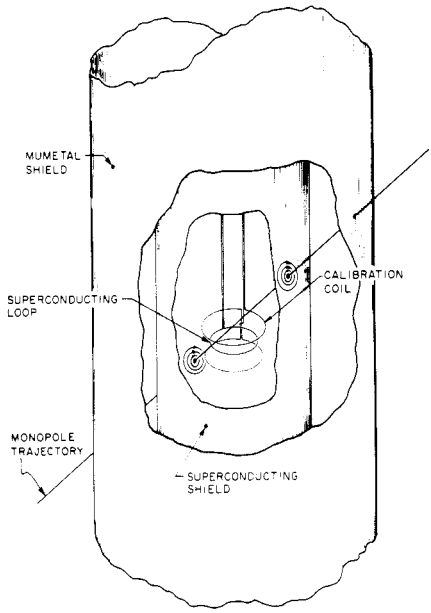


Figure 8.5. Cabrera's superconducting induction detector for detecting magnetic monopoles (from [Gro86b]).

the induced current. A number of experiments of this type were carried out shortly afterwards [Van68, Alv71, Kol71]. However, no evidence for monopoles was found.

The passage of a magnetic monopole through a superconducting ring alters the magnetic flux by

$$\Delta\Phi = 4\pi q_m. \tag{8.50}$$

In the case of a simple magnetic charge ($q_m = g$), it follows that $\Delta\Phi = 4.14 \times 10^{-7}$ G cm^2. On the other hand, the flux through such a superconducting ring is known to be quantized, with flux quantum (the charge $2e$ in the denominator is based on the fact that in the superconductor so-called Cooper pairs are responsible for the charge transportation)

$$\Phi_0 = \frac{h}{2e} = 2.07 \times 10^{-15} \text{ Tm}^2 = 2.07 \times 10^{-7} \text{ G cm}^2. \tag{8.51}$$

The change in the flux due to a simply charged monopole

$$\Delta\Phi = 4\pi g = 4\pi \frac{1}{2\alpha} e = \frac{hc}{e} = 2\Phi_0 \tag{8.52}$$

is twice the elementary flux quantum.

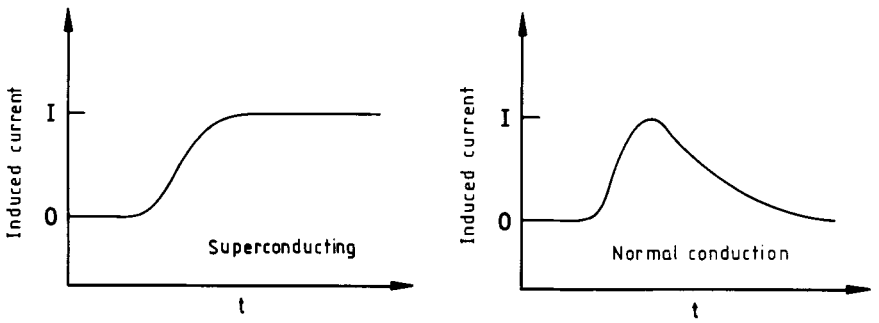


Figure 8.6. The induced current resulting from the passage of a magnetic monopole through a ring for a superconducting coil (left) and a coil with normal conductivity (right).

The change in the flux is measured via the induced current I

$$I = \frac{\Delta\Phi}{L} = \frac{4\pi q_m}{L} \tag{8.53}$$

where L denotes the inductance. Figure 8.6 shows the typical course of such an induction signal for both a superconducting coil and a normal coil. The main advantage of this method is that the signal is independent of the mass m_M, the velocity β and a possible electrical charge of the magnetic monopole. For a circular ring with diameter D and wire thickness d, the inductance is calculated as [Gro46]

$$L = 0.4\pi \frac{D}{1\,\mathrm{m}} \left(\ln \frac{8D}{d} - 1.75 \right) \mu\mathrm{H}. \tag{8.54}$$

Typical values for such experiments are $D = 1$ m and $d = 250$ μm. This gives an inductance of $L = 10$ μH. Thus, the passage of the monopole induces a current of 0.4 nA.

Signals of this order of magnitude may be measured without difficulty using a SQUID (superconducting quantum interference device). Thus, it appears in principle possible to detect individual monopoles. However, the external magnetic fields must be well screened or stabilized, since the signal is very much weaker than the noise due to the normal fluctuations of the Earth's magnetic field. For the above dimensions a comparable signal would be induced already by a change of only 10^{-11} G in the field. Thus, the change in the flux due to the monopole corresponds to extremely small changes in the field (see [Fri84]). In addition, currents which are induced by the passage of a monopole in the magnetic screening may result in a decrease in the current signal.

The disadvantages of this method include the extreme sensitivity to

thermal and mechanical interference effects and the limitation to relatively small dimensions of ≤ 1 m^2 as a result of the extensive cryosystems.

8.3.2 Interaction of monopoles with matter

The detectors based on the interaction of the monopoles with matter may be divided into two subgroups. We distinguish between detectors which use the ionization or the excitation of the matter directly and so-called 'track-etch' detectors which are based on the fact that the passage of a charged particle destroys the crystal structure along the path. This track can be made visible later using chemical methods. Another method, the search for proton decay catalysed by monopoles, has already been discussed.

8.3.2.1 Ionization detectors

This detection technique involves the use of conventional particle detectors such as proportional counting tubes or scintillation counters to identify the particles passing through from the loss of energy due to ionization or excitation. These detectors may be very much larger than induction detectors. However, the interpretation of the measurement results depends on calculations of the detector response to the passage of a monopole.

At this point, we must consider the energy lost by a monopole in matter $(\mathrm{d}E/\mathrm{d}x)_m$. This quantity is well known for relativistic particles, whereas the energy-loss mechanisms and detector signals are not yet fully understood for small projectile velocities (see [Gro86b, Ric87]). Detailed calculations of atomic collisions and excitations are required for this.

A magnetic monopole moving with velocity β produces an electrical field, with lines of force perpendicular to the path of the particle. This field causes an excitation or ionization of the neighbouring atoms or molecules. The energy loss due to ionization by a moving magnetic charge $q_m = ng$ is given by [Gia83]

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_m = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_e \left(\frac{g}{e}\right)^2 (n\beta)^2 \tag{8.55}$$

where $(g/e)^2 = 4700$. $(\mathrm{d}E/\mathrm{d}x)_e$ is given by the Bethe–Bloch formula provided the velocity is not too small [Per82]. It is immediately clear that the highly relativistic Dirac monopoles have an enormous ionization capability in comparison with electrically charged particles. On the other hand, because of their large mass, GUT monopoles have very much smaller velocities ($\beta \sim 10^{-3}$). Figure 8.7 shows the typical energy-loss curves for electrically and magnetically charged particles. However, the behaviour in the region of small velocities must be viewed as relatively uncertain (see also [Gia83]), since a number of

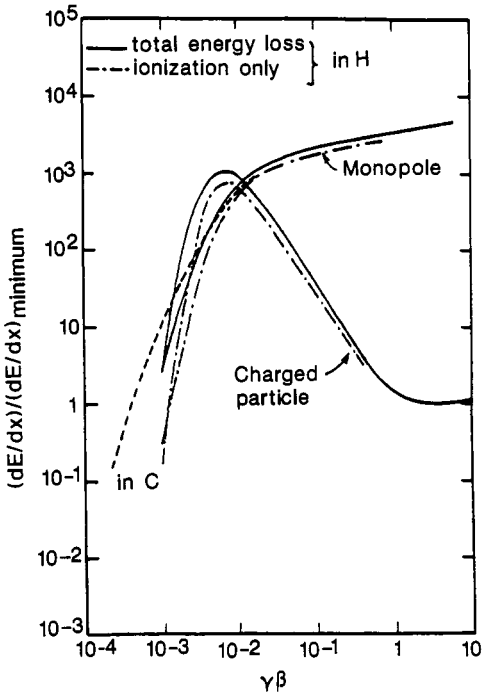


Figure 8.7. Total energy loss (full curves) and energy loss due to ionization (chain curves) for magnetic and electrical elementary charge in atomic hydrogen as a function of $\beta\gamma$. For $\beta\gamma < 10^{-2}$, the energy loss due to ionization in carbon is also shown (from [Gia83], see also [Ahl83]).

other effects must be taken into account in that region (these are discussed, for example, in [Gro86b]).

Two effects in particular play an important role in gas counters, namely the Drell and the Penning effects [Dre83, Kaj84, Gro86b]. When a monopole passes near an atom, it generates crossings of levels and transitions in the atomic shell via the Zeeman effect so that the atom can remain in an excited state. In helium, this results, for example, in an excited metastable state with an excitation energy of around 20 eV. The corresponding energy loss by the monopole is approximately 10 times greater than for normal ionization (*Drell effect*).

If the helium is mixed with a second gas with a lower ionization potential (e.g. *n*-pentane with an ionization energy of 10 eV) the excitation energy of the helium may be transmitted to a pentane molecule by a collision which is ionized in this way. This *Penning effect* forms the basis for all more recent ionization counters for the detection of slow monopoles [Kaj85].

8.3.2.2 Track-etch detectors

When particles carrying electric or magnetic charges pass through matter, they leave a track behind them as a result of the local destruction of the material. In many dielectrics, the damaged regions are chemically very much more reactive than the surrounding material. These tracks may be made visible by treatment (etching) with appropriate chemicals [You58, Fle75, Car78, Pri83, Ahl84]. The advantage of this method is that it is possible to build large detectors and carry out observations over long periods.

8.4 EXPERIMENTAL RESULTS

8.4.1 The search for Dirac monopoles

Since the formulation of the Dirac hypothesis in 1931, the search for magnetic monopoles, and initially, naturally, for the classical light monopoles, has been on. For example, magnetic monopoles have been routinely sought on various accelerator installations, in particular when new energy regions became available. It is conceivable that monopoles (M) might be generated in high-energy reactions such as

$$e^+ + e^- \rightarrow M + \overline{M} \tag{8.56a}$$

$$p + p \rightarrow p + p + M + \overline{M} \tag{8.56b}$$

$$p + \overline{p} \rightarrow M + \overline{M}. \tag{8.56c}$$

They must be produced in pairs M, $\overline{M}$, also because of the assumed conservation of magnetic charge. Since the Dirac monopoles would be highly relativistic, they could be easily detected via their enormous ionization potential. Corresponding experiments were carried out on both electron and proton accelerators. All measurements to detect a monopole have ultimately proved unsuccessful.

Attempts were made to detect monopoles either directly after their generation in high-energy collisions or by indirect methods long after their production. In typical indirect measurements ferromagnetic materials such as iron or manganese are bombarded with highly energetic particles (e.g. protons from the CERN SPS). The monopoles generated in proton–proton or proton–neutron collisions rapidly lose their energy in the solid and remain bound there since they induce opposite charges. Using a strong electromagnet, it should ultimately be possible to extract them from the iron and then accelerate them. These accelerated monopoles should be detectable in ionization counters. Ancient iron ores may also contain magnetic monopoles which could be detected by the same method. However, to date, all experimental searches for classical Dirac monopoles have been unsuccessful. Table 8.1 summarizes some of the experimental results (see also [Gia88]). A full overview is given in [PDG90].

Table 8.1. Search for monopoles at accelerators. The table gives the bounds for the monopole production cross section together with the monopole mass, for various energy regions.

σ_M [cm^2]	m_M [GeV]	Beam	$\sqrt{s}$ [GeV]	Events	Ref.
$< 2 \times 10^{-35}$	< 1	p	6	0	[Bra59]
$< 1 \times 10^{-35}$	< 3	p	28	0	[Fid61]
$< 2 \times 10^{-40}$	< 3	p	30	0	[Pur63]
$< 5 \times 10^{-42}$	< 13	p	400	0	[Car74]
$< 5 \times 10^{-43}$	< 12	p	400	0	[Ebe75]
$< 4 \times 10^{-38}$	< 10	e^+e^-	34	0	[Mus83]
$< 3 \times 10^{-32}$	< 800	$p\overline{p}$	1800	0	[Pri87b]
$< 1 \times 10^{-38}$	< 17	e^+e^-	35	0	[Bra88]
$< 1 \times 10^{-37}$	< 29	e^+e^-	50–61	0	[Kin89]
$< 2 \times 10^{-34}$	< 850	$p\overline{p}$	1800	0	[Ber90d]

8.4.2 The search for GUT monopoles

None of the experiments mentioned in section 8.4.1 were sensitive to GUT monopoles, since the available energies were not sufficient to produce such heavy particles or the magnetic fields used could only accelerate light Dirac monopoles to velocities which would have led to a measurable ionization. In what follows, we shall give a brief survey of attempts to detect heavy GUT monopoles. These examples are a nice further illustration that it is possible to advance into energy regions which are not accessible using current accelerator technologies.

A number of experiments used the induction in superconducting coils resulting from the passage of a magnetic charge. The use of several loops allows for coincidence measurements which permit a reduction of the background. Figure 8.5 shows the first of these induction detectors due to Blas Cabrera in Stanford [Cab82]. The coil included four loops and its diameter was only 5 cm. A possible candidate for a magnetic monopole was observed using this detector (see figure 8.8). A conclusive explanation of this event has not yet been given. If it were due to a magnetic monopole, the measurement time of 151 days would imply a flux of

$$F_M \approx 6 \times 10^{-10}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1}. \tag{8.57a}$$

However, subsequent experiments give very much sharper bounds. Cabrera himself interpreted his result as a bound [Cab82]

$$F_M < 1.4 \times 10^{-9}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \quad (90\%\ \text{c.l.}). \tag{8.57b}$$

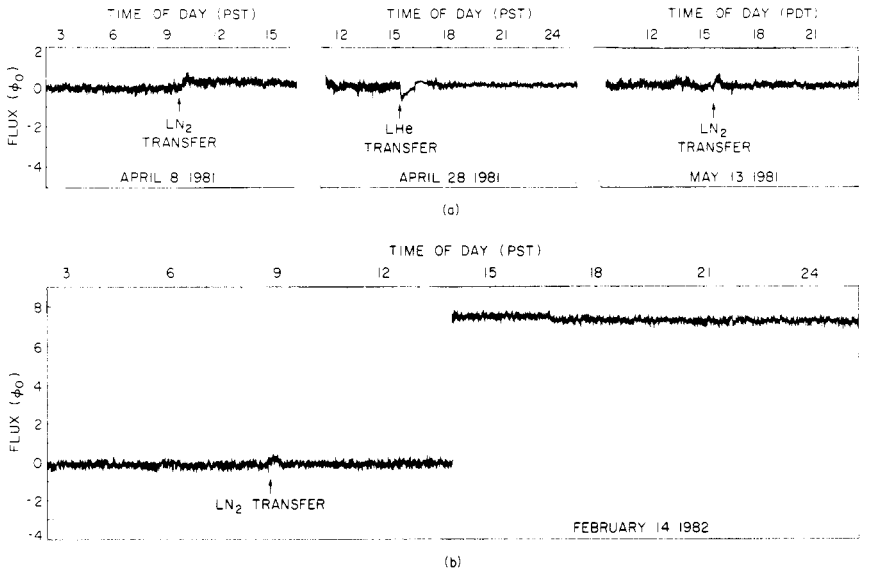


Figure 8.8. Measurement data from Cabrera's original apparatus acquired on 14 February 1982. The lower figure shows the jump in the flux by 8 units, which one would expect for Dirac monopoles. The upper part of the figure shows the effect of typical perturbing effects on the measured flux (from [Gro86b]).

Table 8.2. Flux bounds for magnetic monopoles from induction detectors.

F_M [$cm^{-2}\,sr^{-1}\,s^{-1}$]	Ref.
$< 1.4 \times 10^{-9}$	[Cab82]
$< 3.7 \times 10^{-11}$	[Cab83]
$< 5.9 \times 10^{-10}$	[Ebi84]
$< 6.7 \times 10^{-12}$	[Inc84]
$< 5.5 \times 10^{-12}$	[Ber85b]
$< 6.0 \times 10^{-12}$	[Cap85, 86]
$< 5.0 \times 10^{-12}$	[Cro86]
$< 3.8 \times 10^{-13}$	[Ber90c]
$< 7.2 \times 10^{-13}$	[Hub90]
$< 4.4 \times 10^{-12}$	[Gar91]

Following this measurement a number of new experiments were carried out using increasingly large detectors. Table 8.2 summarizes the bounds derived for the flux. Another experiment [Cap85] observed an unexplainable induced

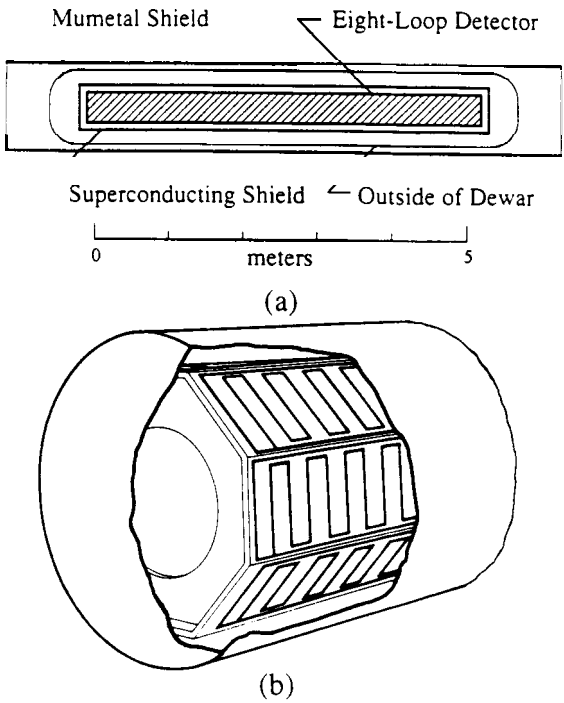


Figure 8.9. (*a*) Schematic structure of the Stanford monopole detector comprising eight independent coils with the superconducting shielding, the Dewar vessel and the mu-metal shielding; (*b*) arrangement of the coils in the superconducting shielding (from [Hub90]).

current which could point to the passage of a magnetic charge. However, this candidate also cannot be unambiguously interpreted as a monopole.

The sharpest bounds come from three new large detectors. The IMB detector consists of six mutually independent planar detector coils which are arranged on the surface of an rectangular parallelepiped. Thus, a monopole passing through will induce a signal in exactly two of these coils. For an isotropic flux, the effective surface area of the detector averaged over the total solid angle of 4π is approximately 1 m^2. No events were detected after a measurement time of 13 410 hours, giving a bound of [Ber90c]

$$F_M < 3.8 \times 10^{-13}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1} \tag{8.58}$$

for the flux. The second experiment, in Stanford, used a detector consisting of eight independent coils with an effective area of 1.1 m^2 [Hub90, 91] (figure 8.9). No monopoles were recorded after a measurement time of 6482 hours, giving a

bound of

$$F_M < 7.2 \times 10^{-13}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1} \tag{8.59}$$

for the flux. A further measurement at Stanford University using three superconducting detector coils with a total sensitive surface area of 476 cm^2 gave an upper bound of [Gar91]

$$F_M < 4.4 \times 10^{-12}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1} \tag{8.60}$$

after a measurement period of 24 190 hours. At this point we again note that the values obtained by the induction method are independent of the mass and the velocity of the monopole. However, the bounds are still approximately three orders of magnitude larger than the value $F_M < 1 \times 10^{-16}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1}$ derived by Parker [Par70].

Experiments using organic scintillators and gas-filled wire chambers have the advantage that, unlike superconducting coils, they may incorporate large

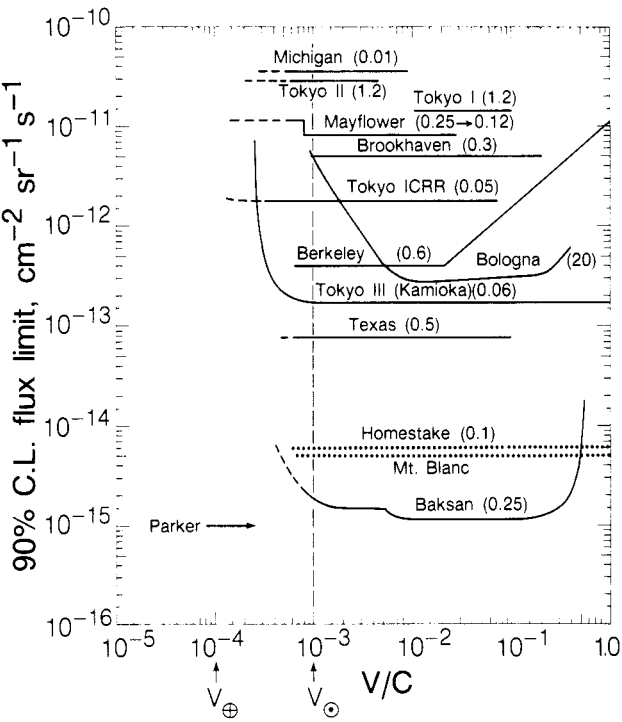


Figure 8.10. Bounds for the flux of magnetic monopoles from ionization experiments with scintillation counters (from [Gro86b]).

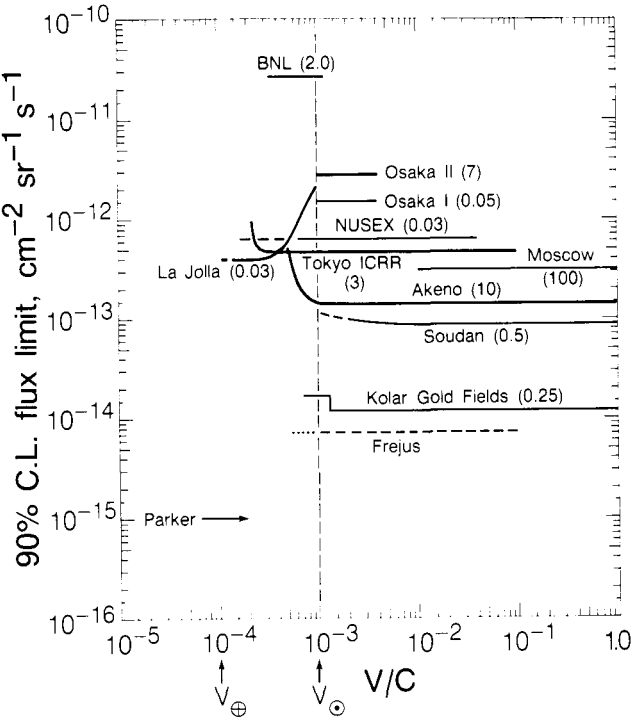


Figure 8.11. Bounds for the flux of magnetic monopoles from wire chamber experiments (from [Gro86b]).

surface areas. However, the measurement results depend on the velocity of the monopoles. The lower bound as far as the sensitivity of scintillation counters is concerned is approximately $\beta \sim 10^{-4}$. The bounds on the flux from ionization experiments are summarized for scintillation counters in figure 8.10 and for wire chambers in figure 8.11. A summary is given in [Gro86b]. The most sensitive experiment of this type to date was carried out in the Baksan underground laboratory [Ale82, 85]. The bounds for the monopole flux are already quite close to the Parker limit.

The largest detector to date in the search for magnetic monopoles, the MACRO detector (monopole and cosmic ray observatory) is almost completed in the Gran Sasso underground laboratory (see chapters 4, 6 and 7) at a depth of 3600 m water equivalent [Bat88, Cal88b, Bel90, Bar92, Ahl93, Hon94] (figure 8.12). When it is completed the MACRO detector, which measures $72 \times 72 \times 10$ m^3, will have an acceptance of $S\Omega \sim 1000$ m^2 sr for an isotropic particle flux. It is intended to measure particle traces, velocities and ionization

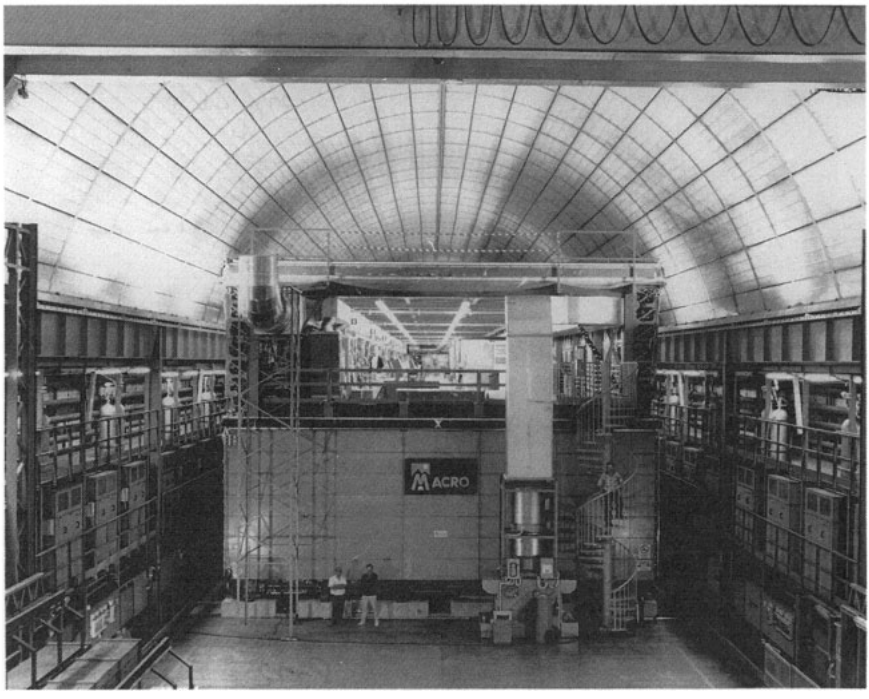


Figure 8.12. The MACRO detector in Hall B of the Gran Sasso underground laboratory near Rome (status beginning of 1993). (Courtesy of Gran Sasso Laboratory.)

capabilities. The detector consists of twelve independent modules ($12 \times 12 \times 4.8$ m^3). Figure 8.13(*a*) shows one of these twelve building blocks. Each module consists of two liquid scintillators and ten streamer chambers (see the cross section in figure 8.13(*b*)). The distance between the streamer chambers is 32 cm. This space is filled with absorbent material (rock). The middle layer of the modules also includes a track-etch detector. The four vertical sides are each enclosed by a scintillation detector and six layers of streamer chambers.

This gigantic detector will be used to search for monopoles in the 10^{16} GeV/c^2 mass region below the Parker limit (see figure 8.15). The first results with one module have already given a bound of

$$F_M < 5.6 \times 10^{-15} \ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \quad (10^{-4} \lesssim \beta < 4 \times 10^{-3}) \tag{8.61}$$

for the flux of simply charged GUT monopoles [Bel90, Hon94, Ahl94]. However, the measurements are not restricted solely to magnetic monopoles. The MACRO detector is also suitable for detecting cosmic radiation and other charged superheavy particles or supernova neutrinos [Cal88b, Bar92].

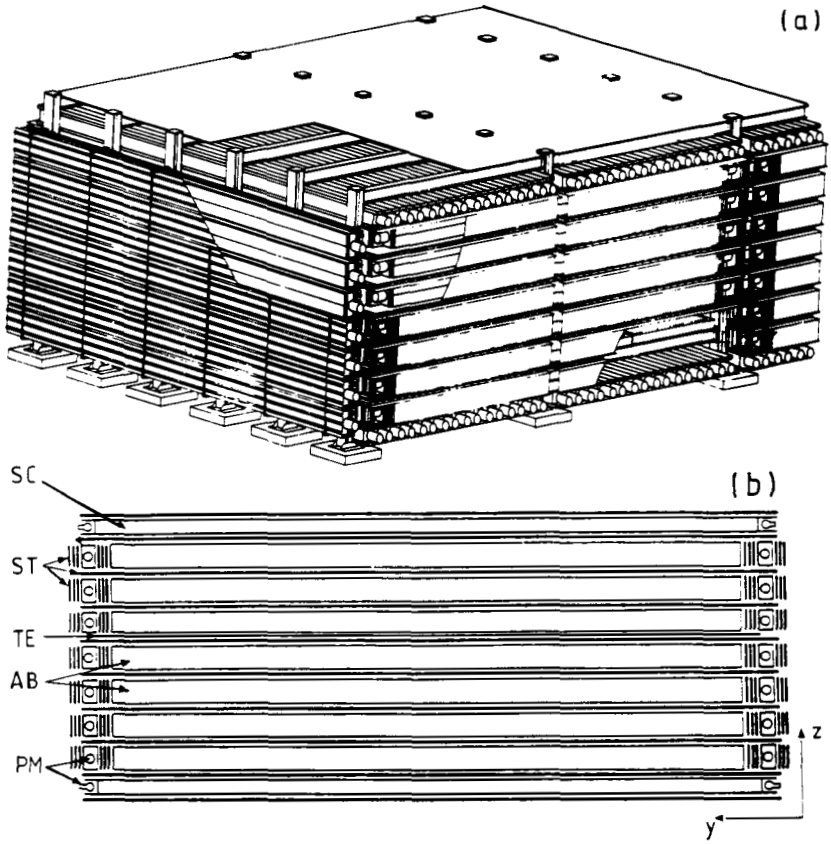


Figure 8.13. The MACRO detector in the Gran Sasso underground laboratory: (*a*) structure of one of the 12 modules (12 m × 12 m × 5 m) of the MACRO detector (from [Cal88b]); (*b*) cross section of a detector module: SC = scintillator, ST = streamer chamber, TE = 'track–etch' detector, AB = absorber, PM = photomultiplier (from [Ahl90]).

Bounds for the flux F_M have also been determined using track-etch detectors.

$$F_M < 1.6 \times 10^{-13}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \text{ for } \beta > 0.02 \quad \text{[Kin81]} \qquad (8.62a)$$

$$F_M < 10^{-12}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \text{ for } \beta > 0.007 \quad \text{[Bar83]} \qquad (8.62b)$$

$$F_M < 5 \times 10^{-15}\ \text{cm}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1} \text{ for } \beta > 0.04 \quad \text{[Dok83].} \qquad (8.62c)$$

[Pri84] points out that this method may be used down to velocities of $\beta \sim 10^{-4}$. A new track-etch experiment with a detector surface area of 2000 m^2 was recently carried out. The material (plastic discs) was exposed to a possible

monopole flux over a period of 2.1 years [Ori91]. A new upper bound for F_M has been given, depending on the velocity and charge of the monopole. The sensitivity ranges down to

$$F_M < 3.2 \times 10^{-16}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1}. \tag{8.63}$$

Figure 8.14 shows the results of a precise analysis of the data.

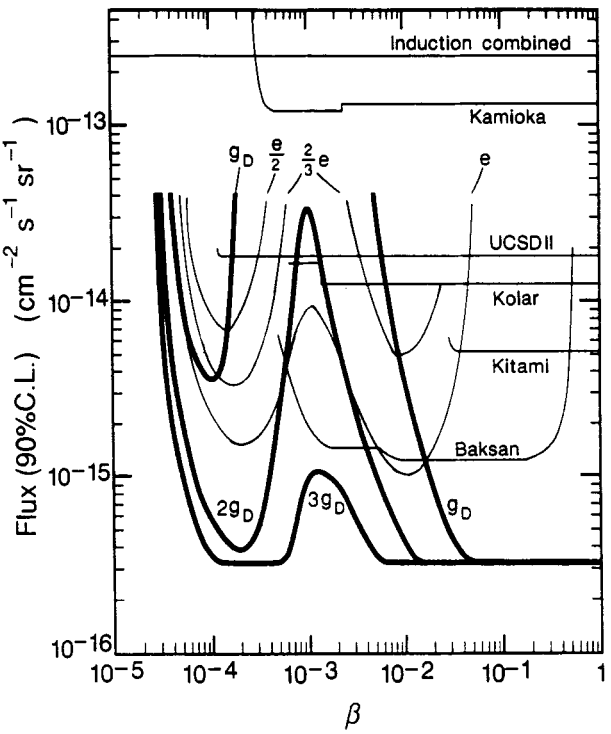


Figure 8.14. Flux bounds (90% confidence level) for magnetic monopoles (bold curves) (and particles with a fractional electrical charge) as a function of β from the track-etch experiment of Orito *et al* (from [Ori91]). Bounds from other experiments are also shown.

The indirect method, which uses old muscovite (potash mica ($K_2Al_4[Si_6Al_2O_{20}](OH,F)_4$)) as a track-sensitive detector, is of particular interest. This involves the use of the long exposure times of the mica samples. As early as 1969 studies of 2×10^8 year-old mica samples gave a value of [Fle69]

$$F_M < 1 \times 10^{-19}\ \mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{s}^{-1}. \tag{8.64}$$

More recent studies of samples with an age of 4.6×10^8 years were carried out by [Pri86]. It is conjectured that the track is formed by the following

mechanism. A monopole which penetrates the rock captures a nucleus with a large magnetic moment (in general, this is assumed to be a ^{27}Al nucleus). This bound object should leave visible tracks behind it in the mica. As a result of the non-observation of such tracks, bounds in the region between 5×10^{-19} and 5×10^{-17} cm^{-2} sr^{-1} s^{-1} were derived (for $\beta \sim 10^{-3}$). However, this method involves a number of problems, so that the results are only valid under certain conditions (see [Gro86b]).

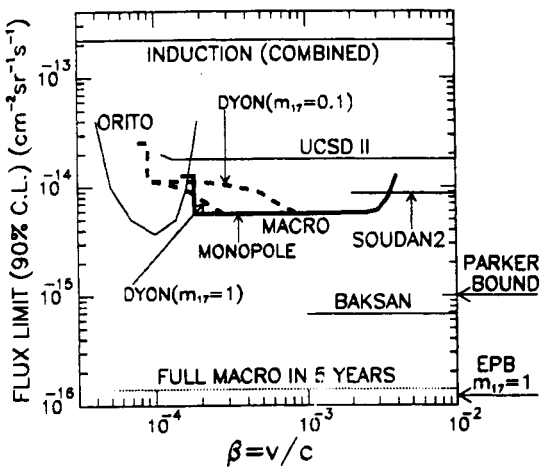


Figure 8.15. Upper bounds on the flux (90% confidence level) for monopoles and dyons (bound monopole–proton states, see [Bra84]) from various experiments (UCS-DII [Buc90], Orito [Ori91], SOUDAN2 [Thr92], Baksan [Ale82], MACRO [Hon94]). The figure also shows the Parker bound, the extended Parker bound (EPB) for monopoles with 10^{17} GeV mass ($m_{17} = 1$) and the expected value from five-year measurements with the completed MACRO detector (from [Hon94]).

In this section we have given a brief survey of the search for magnetic monopoles. Figure 8.15 summarizes the most important results. A full evaluation of the ideas and experiments is beyond the scope of this book, and we have only been able to mention a certain selection of these. In summary, modern theories appear to require the existence of isolated magnetic charges which have not yet been unequivocally detected.

Chapter 9

The Search for Dark Matter in the Universe

There is evidence that a large part of the matter in the universe does not emit radiation. However, this invisible matter can be recognized from its gravitational interaction with the emitting matter [Fab79, Sch89]. Studies of galaxy clusters and galaxy rotation curves provide evidence for the existence of this so-called dark matter.

The density of matter in the universe ρ may be estimated from the observation of the movement of individual galaxies. Usually ρ is given in units of the so-called critical density ρ_c

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{9.1}$$

ρ_c denotes the critical density such that for $\rho > \rho_c$ the universe is closed, i.e. the gravitational interaction is sufficient to reverse the expansion of the universe into a contraction (see chapter 3). G denotes the gravitational constant, H is the Hubble constant. The latter is only known up to a factor of two

$$H = 100h \text{ km s}^{-1} \text{ Mpc}^{-1} \qquad 0.4 < h < 1. \tag{9.2}$$

Thus, the critical density is given by

$$\rho_c = 2 \times 10^{-29} h^2 \text{ g cm}^{-3}. \tag{9.3}$$

The cosmological density $\Omega \equiv \rho/\rho_c$ (see below) determined from the dynamics of clusters and superclusters of galaxies is [Blo84, Sch86b, Ton93]

$$0.1 \lesssim \Omega \lesssim 0.3. \tag{9.4a}$$

The observation of recession velocity fields of large-scale areas via the InfraRed Astronomy Satellite (IRAS) yields [Dek93a,b, Ton93]

$$0.25 < \Omega < 2. \tag{9.4b}$$

On the other hand, estimation of the baryonic density Ω_b from the luminosity of galaxies gives a considerably smaller mass density [Wal91]

$$\Omega_b < 0.02. \tag{9.5}$$

This discrepancy is generally interpreted as an indication of the existence of invisible matter (see e.g. [San83, Ton93]).

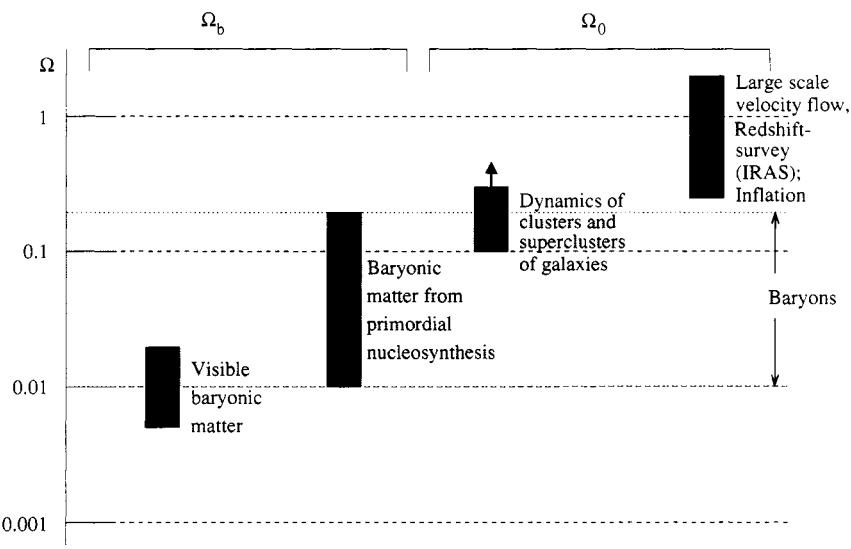


Figure 9.1. Estimates of the total density parameter Ω_0 and the baryonic component Ω_b (see the following chapters and section 3.3 and, in particular, [Ton93, Dek93, 93a]).

The search for dark matter has recently become a very active research area. That dark matter which reveals itself through its gravitational effect but emits no electromagnetic radiation must by necessity exist is undisputed. Even when all baryonic forms such as dust, brown and white dwarfs, neutron stars and black holes are taken into account, it still appears that a considerable fraction of *non-baryonic* matter is needed to explain all observations (figure 9.1). This remains true also after taking into account recent indications of so-called MACHOs (Massive Compact Halo Objects) (see [Sad94]) by gravitational lens (see [Pac86]) effects. In what follows, we shall firstly investigate observable

events which call for the existence of dark matter. We shall then discuss the possible composition of this matter. A final section is devoted to the detection of possible candidates.

A short historical comment is appropriate at this point. The detection of 'dark matter' dates back to the last century [Schr89]. In a letter to the mathematician Carl Friedrich Gauss in 1844, the physicist Friedrich Bessel, wrote that the hitherto unexplained fluctuation of Sirius was a result of the gravitational interaction with a neighbouring body, where the latter must have a large mass in order to be responsible for the conspicuous effect. In Bessel's day this body was invisible; however, this dark companion of Sirius was first detected optically in 1862. In this example 'invisible' matter was first detected by its gravitational effect.

9.1 EVIDENCE FOR THE EXISTENCE OF DARK MATTER

9.1.1 Galaxy rotation curves

We shall initially consider so-called spiral galaxies, i.e. collections of large numbers of stars in the shape of a flattened rotating disc (figure 9.2). The rotational velocity of the individual stars around the centre of the galaxy is determined from the condition for a stable orbit. By virtue of the equality of the centrifugal and the gravitational force

$$\frac{GmM_r}{r^2} = \frac{mv^2}{r} \tag{9.6}$$

the rotational velocity is given by

$$v(r) = \sqrt{\frac{GM_r}{r}} \tag{9.7}$$

where M_r denotes the total mass within the radius r. The effects of the mass elements outside the radius r cancel each other out in the case of ideal spherical or cylindrical symmetry. To a first approximation, the central bulge of the galaxy may be assumed to be spherical, i.e.

$$M_r = \overline{\rho}\frac{4}{3}\pi r^3 \tag{9.8}$$

where $\overline{\rho}$ denotes the average density. Thus, we obtain the following expression for the rotational velocity

$$v(r) = \sqrt{\frac{4}{3}\pi G\overline{\rho}r^2} \sim r. \tag{9.9}$$

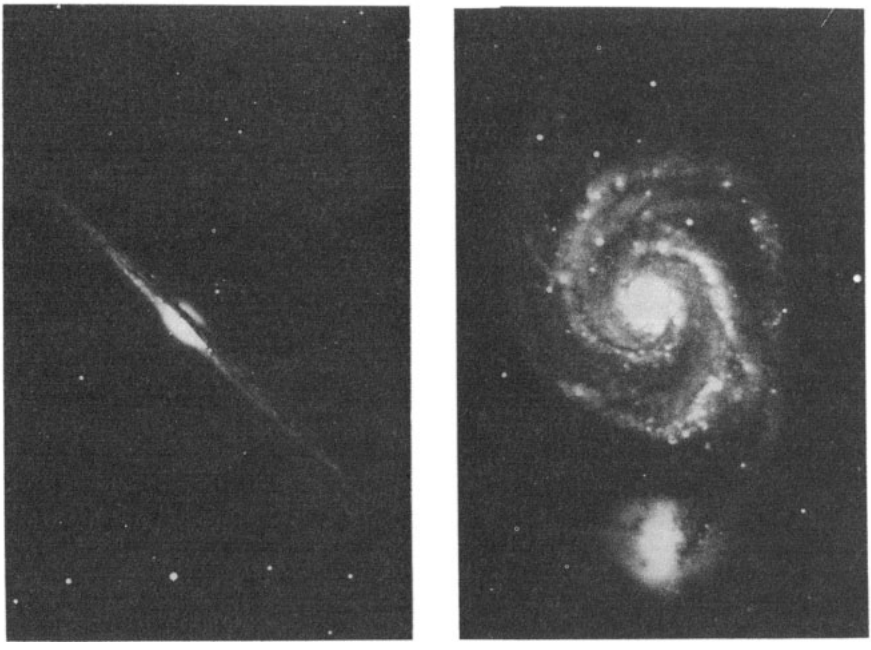


Figure 9.2. Spiral galaxies NGC 4565 (left) and M51 (with its companion NGC 5195 in the Canes Venatici constellation) (right). The central bulge and the thin disc in which the spiral arms are embedded are clearly recognizable in NGC 4565 (from [Sch89] and [Rie87], respectively).

In the inner part of the galaxy we expect a linear increase in the rotational velocity as the distance from the centre increases. In the outer area of the galaxy the mass M_r is essentially constant. The behaviour is similar to that of a point mass in the centre of the galaxy: we have

$$v(r) \sim 1/\sqrt{r}. \tag{9.10}$$

Figure 9.3(a) shows the expected rotation curve for a spiral galaxy.

The rotational velocity $v(r)$ is determined experimentally, for example by measuring the Doppler shift in the emission spectrum of the HII regions around so-called O stars. The experimental rotation curves of spiral galaxies do not show a decrease in $v(r)$ for large radii (see figure 9.3(b)–(d)). Studies of the 21 cm line (hyperfine structure transition in atomic hydrogen) emitted by interstellar matter give the same result. The fact that $v(r)$ is constant for large radii means that the mass M_r increases as the radius increases

$$M_r \sim r. \tag{9.11}$$

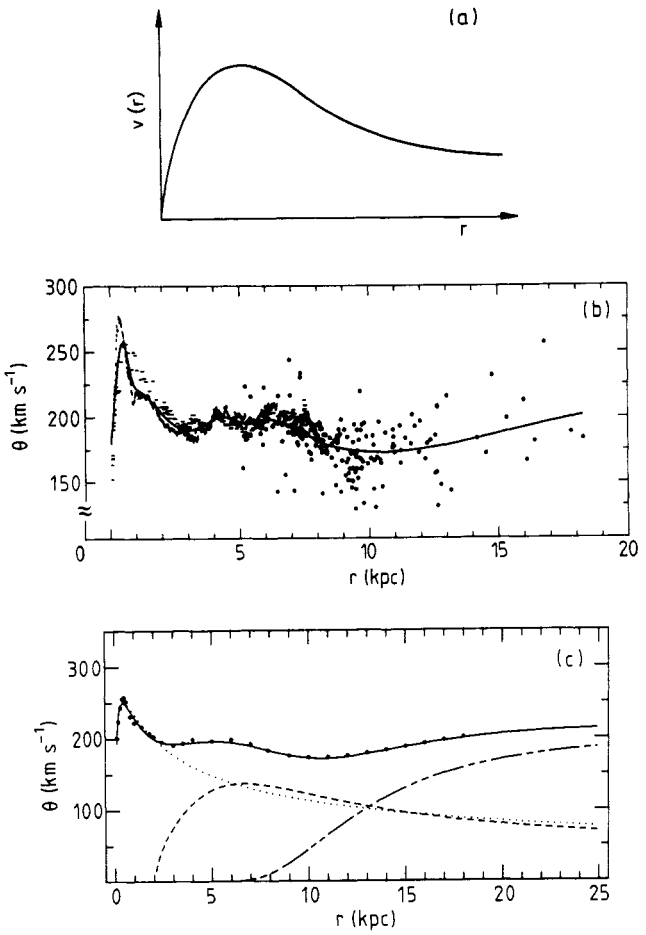


Figure 9.3. (*a*) Expected shape of the rotation curve of a galaxy. (*b*) Orbital velocity θ of objects in the disc of the Milky Way (rotational velocity of the Galaxy) as a function of the distance r to the centre of the Galaxy. (*c*) Example of the fitting of a model to the rotation curve (cf. (*b*)) of the Milky Way. The model comprises three components: bulge–disc–halo, whose contributions, plotted as θ versus r, are shown by $\cdots$, – – – and — · · —, respectively.

This points to the presence of unseen matter. The stars move faster than one would expect from the amount of visible matter.

This fact led to the postulation of a spherical *halo of dark matter* surrounding the galaxy which is responsible for the flat rotation curve. In addition, a

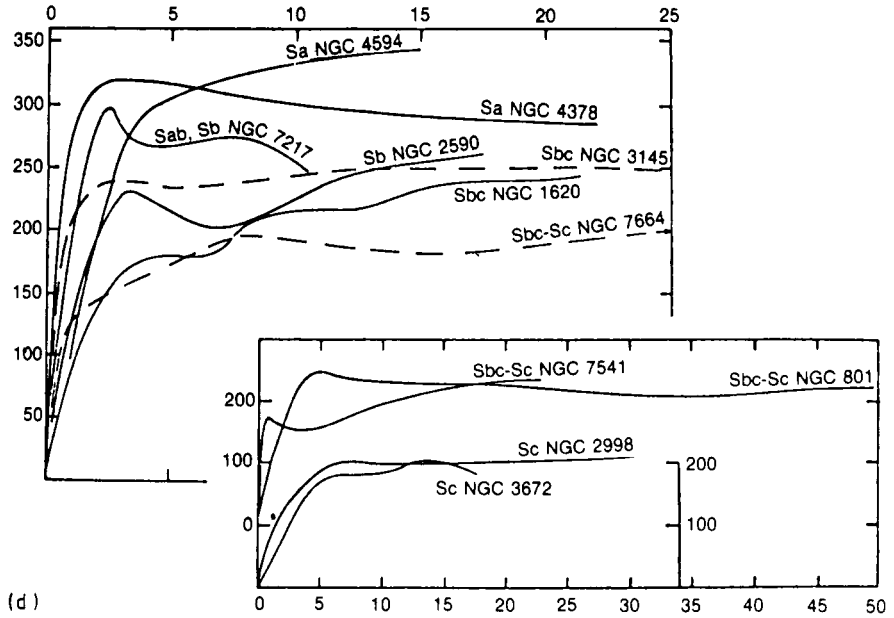


Figure 9.3. (continued) (*d*) Observed rotation curves, θ versus r, of Sb spiral galaxies (from [Sch89]).

spherical halo could also contribute to the stability of the disc shape of the galaxies. It would also support the hypothesis that galaxies have developed from a spherical protogalaxy. Model calculations for our Milky Way, which are able to reproduce the rotation curves taking into account the halo, point to the fact that a considerable part of the mass is to be found in the halo [Roh86, Fic91]. Evidence for spherical halos is also provided by the globular clusters, spherical star clusters, which are the oldest objects in the galaxy and are spherically distributed.

However, recent research into the transparency of galaxies cast doubt on this picture [Val90]. By considering the opacity of spiral galaxies as a function of the angle of inclination, it is possible to draw conclusions about the transparency of these objects. If a galaxy were completely transparent the total luminosity would be independent of the angle of inclination at which it is observed, since all stars would always be visible (if one neglects star diameters). On the other hand, a constant surface brightness means that the galaxy is opaque. One would then only ever be able to see the outer stars, i.e. always the same number per surface area, independently of the viewing angle. Figure 9.4 shows the results of studies of several thousand galaxies. It is clear that the surface brightness remains

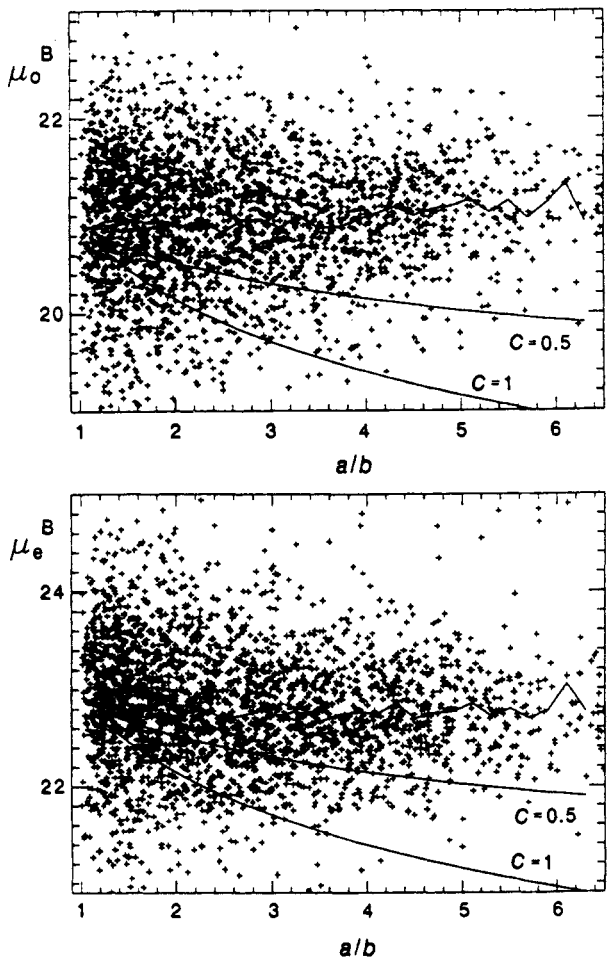


Figure 9.4. Opacity of spiral galaxies. The figure shows the distribution of 2639 type 2.5–5.0 ($\sim$ Sb) galaxies in a diagram of the local surface brightness (μ^B) as a function of the angle of inclination (corresponding to the observed ratio of the axes a/b). The two curves ($C = 1$ and $C = 0.5$) show the expected trend for transparent discs ($C = 1$ corresponds to full transparency) (from [Val90]).

on average constant, which would indicate that spiral galaxies are practically opaque. Under this assumption optical studies to determine the mass density in the universe would not be very appropriate.

A more accurate analysis of the measurement results points to molecular clouds as absorbent material (diameter $\sim$50 pc, temperature $\sim$20 K). According

to Wien's displacement law, these clouds should emit radiation in the submillimetric area. This result could lead to an explanation of the flat rotation curves, even without the assumption of additional exotic dark matter [Val90]. This argument has been questioned elsewhere [Ces90]. A more recent investigation is given by [Jam93] (see also [Bot91]).

Evidence for dark matter has also been found in elliptical galaxies. Gas halos with temperatures of around 10^7 K have been detected by their X-ray absorption. The velocities of the gas molecules were far greater than the recession velocity

$$v_r = \sqrt{\frac{2GM}{r}} \tag{9.12}$$

if the mass is assumed to correspond to the luminosity. The ratio of mass to luminosity for elliptical galaxies is approximately two orders of magnitude greater than that of the Sun, which is a typical example of an average star. This large value is generally attributed to the existence of dark matter.

9.1.2 The dynamics of galaxy clusters

The dynamics of galaxy clusters also provides evidence for dark matter. If the movement of a system, the potential energy of which is a homogeneous function of the coordinates, takes place in a limited spatial region, then the average values over time of the kinetic and potential energies are related by the virial theorem of Clausius (see e.g. [Lan79]). This theorem may be used to estimate the matter density of clusters of large numbers of galaxies.

If the potential energy U is a homogeneous function of degree k of the radius vectors r_i, then U and the kinetic energy T are related as follows (see [Lan79])

$$2\overline{T} = k\overline{U}. \tag{9.13}$$

Since $\overline{T} + \overline{U} = \overline{E} = E$, equation (9.13) may also be represented by the formulae

$$\overline{U} = \frac{2}{k+2}E \qquad \overline{T} = \frac{k}{k+2}E \tag{9.14}$$

which express the average values in terms of the total energy E. For the gravitational interaction ($U \sim 1/r$) with $k = -1$ we have

$$2\overline{T} = -\overline{U}. \tag{9.15}$$

For a cluster of N galaxies the average kinetic energy is given by

$$\overline{T} = \frac{1}{2}N\overline{mv^2}. \tag{9.16}$$

The N galaxies may interact with one another in pairs. Thus, there are $\frac{1}{2}N(N-1)$ independent pairs of galaxies, the overall average potential energy of which is given by

$$\overline{U} = -\frac{1}{2}GN(N-1)\frac{\overline{m}^2}{r}. \tag{9.17}$$

For $N\overline{m} = M$ and $(N-1) \approx N$ this gives a dynamic mass of

$$M \approx \frac{2\overline{r}\,\overline{v^2}}{G}. \tag{9.18}$$

Measurement of $\overline{r}$ and $\overline{v}$ gives a dynamic mass which is approximately two orders of magnitude greater than that derived from the average ratio of luminosity to mass. This may be interpreted as further evidence for the existence of dark matter.

However, this argument also has weak points. The virial equation is only valid on average over a long period for closed systems which are in equilibrium. But the measurements of galaxy clusters are only instantaneous snapshots. Moreover, galaxy clusters are not closed systems but are linked together. Finally, it is unclear whether they have reached equilibrium since their creation.

9.1.3 Evidence from cosmology

We have already defined the critical density ρ_c in the introduction to this chapter. The expression (9.1) may be derived formally from Newtonian dynamics by calculating the critical recession velocity for a spherical universe

$$E = T + U = \frac{1}{2}mv^2 - \frac{Gm}{r}\rho\frac{4}{3}\pi r^3 = 0. \tag{9.19a}$$

Equation (9.1) now follows with

$$H = \frac{\dot{r}}{r} = \frac{v}{r} \tag{9.19b}$$

For a better understanding of this critical density we shall consider the dynamics of the universe in the framework of the theory of general relativity.

The description of the dynamics of the universe is based on Einstein's field equations. These may be simplified if one assumes a homogeneous, isotropic space. In the associated metric, the Robertson–Walker metric, an infinitesimal line element is given by

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R^2(t)\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right). \tag{9.20}$$

Here, r, θ and ϕ are the polar coordinates of a point in space. The degrees of freedom of the metric are included in the parameter k and the scale factor R. k can take only discrete values (see section 3.1) and is time invariant. The initial value assumed for k is a characteristic of each model. We have

$$k = \begin{cases} -1 & \text{hyperbolic metric (open universe)} \\ 0 & \text{Euclidean metric (flat universe)} \\ +1 & \text{spherical metric (closed universe).} \end{cases} \tag{9.21}$$

The dynamics is completely contained in the scale factor $R(t)$ (the distance between two neighbouring points in space with constant coordinates r, θ and ϕ, scales with time by $R(t)$). In the case of the spherical metric $R(t)$ is also a graphic representation of the 'radius' of the universe. The scale factor satisfies the Einstein–Friedmann–Lemaitre equations

$$\left(\frac{\dot{R}(t)}{R(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{R^2(t)} + \frac{\Lambda c^2}{3} \tag{9.22a}$$

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3}{c^2}p(t)\right) + \frac{\Lambda c^2}{3}. \tag{9.22b}$$

Here, $p(t)$ denotes the overall pressure and Λ is the so-called cosmological constant, which is interpreted as the vacuum energy density in the framework of modern quantum field theories (see below).

In what follows, we shall initially assume, as is often done (but see [Gro89, 90]) that the cosmological constant Λ vanishes. The quotient $\dot{R}_0/R_0$ gives the Hubble constant H_0, where the subscript 0 stands for the current value of the corresponding quantity. It follows then from (9.22*a*) that, for the curvature parameter $k = 0$, the current critical density of the universe (see (9.1)) is given by

$$\rho_{c0} = \frac{3H_0^2}{8\pi G}. \tag{9.23}$$

Thus, this density represents the boundary between an open and a closed universe. It determines the boundary between perpetual expansion and a collapse after a finite period of expansion.

The so-called density parameter Ω_0 is often used

$$\Omega_0 = \frac{\rho_0}{\rho_{c0}} = \frac{8\pi G\rho_0}{3H_0^2} = 1 + \frac{kc^2}{R_0^2H_0^2} = 2q_0 \tag{9.24}$$

where q_0 is the so-called deceleration parameter ($q(t) = -R(t)\ddot{R}(t)/\dot{R}^2(t)$). Thus, the following holds

$$\begin{aligned} &\Omega_0 < 1 \Rightarrow \text{ open universe} \\ &\Omega_0 = 1 \Rightarrow \text{ flat universe} \\ &\Omega_0 > 1 \Rightarrow \text{ closed universe.} \end{aligned} \tag{9.25}$$

Measurements of the density parameter give a value of approximately (see (9.4))

$$\Omega_0 \approx 0.2 \tag{9.26}$$

which would point to an open universe. However, certain theoretical considerations are difficult to reconcile with an open universe, namely the so-called flatness problem and the genesis of the galaxies.

The flatness problem. It is astonishing that the universe comes at all so close to the critical density. From the Einstein–Friedmann–Lemaitre equations it follows (again with $\Lambda = 0$) that

$$\frac{\dot{R}^2(t)}{kc^2} = \frac{8\pi G\rho(t)R^2(t)}{3kc^2} - 1. \tag{9.27}$$

Since the density $\rho(t)$ is proportional to $1/R(t)^3$, by virtue of (9.24) for $k \neq 0$ we have

$$\Omega(t) = 1 + \frac{1}{(a/kR(t)) - 1} \tag{9.28}$$

where $a = 8\pi G/3c^2$. Thus, the value $\Omega \approx 1$ is very unstable. Any deviation from the exactly flat value increases with the expansion of the universe. This means that at the time of the primordial nucleosynthesis the universe must have been a *very* great deal flatter than it is today.

One possible answer to this problem is provided by so-called inflationary models [Gut81, Alb82, Lin82, 84, 90, Kol90]. These assume that the expansion of the early universe (between 10^{-34} and 10^{-31} s after the big bang) took place exponentially in an inflationary phase. These models generally imply a time-independent density parameter

$$\Omega = 1 \qquad \text{(in inflationary models)}. \tag{9.29}$$

However, some theoretical work also points to the fact that density parameters of $0.01 < \Omega_0 < 2$ could be consistent with the inflationary model [Hoe90].

Genesis of galaxies. The genesis of galaxies requires density inhomogeneities. Galaxies must have originated in spatial regions in which the densities were greater than in the surroundings so that, as a result of the gravitational interaction, these regions collapsed before the general expansion was able to separate this matter.

However, such an accumulation of matter could only begin after the formation of atoms from nuclei and electrons, i.e. approximately 150 000 years after the big bang, at temperatures of around 3000 K. Noticeable density fluctuations of ordinary matter at that time have been ruled out down to a very low level by the isotropy of the cosmic background radiation, which has now

been well measured. By the formation of neutral atoms the radiation could decouple, i.e. from that time it was no longer in thermal equilibrium with the matter; thus, density fluctuations occurring then would no longer be reflected in it.

However, if one calculates the temporal evolution of the compression of matter which was beginning at that time, it turns out that the time until today is not sufficient to permit the development of such large structures as galaxies or galaxy clusters. Thus, it appears necessary to require the existence of massive particles which decoupled from the thermal equilibrium at an early time so that they were able to act as seeds for the condensation of the ordinary matter (see e.g. [Kol90]). Candidates for these would be the so-called WIMPs (see section 9.2.3). This must have been subject to the boundary condition that the isotropy of the background radiation [Mat90] was essentially unaffected. Tiny anisotropies in the 3 K radiation were discovered only recently using the COBE satellite [Ben92a, Sil92, Smo92, Wri92].

In summary, we note that there is a range of evidence for the possible existence of baryonic as well as non-baryonic dark matter, although this is–in particular for the latter–not mandatory and other, less exotic explanations might be possible. The most difficult problem at the moment is that of the genesis of the galaxies and the associated large-scale structure of the universe (see [Gel89, Huc90, Sch90c, Schr85a, Sil93]). In addition to the genesis of individual galaxies, the formation of clusters and superclusters remains to be explained. Since the distribution of matter in the universe on scales of ≈ 100 Mpc also appears to be extremely inhomogeneous, this must be brought into harmony with the extraordinary homogeneity of the 3 K radiation.

On the other hand, various theories beyond the standard model predict a range of new particle types such as supersymmetric partners to the particles known today, which could provide an explanation for the problems described above. In the next section we shall discuss possible candidates for dark matter.

9.2 CANDIDATES FOR DARK MATTER

9.2.1 The cosmological constant, MOND theory, time-dependent gravitational constant

The cosmological constant Λ was originally introduced by Einstein in the field equations of the theory of general relativity, to guarantee a static universe, corresponding to the ideas of that time. However, after Hubble's discovery in the late 1920s that the universe is expanding, the cosmological constant appeared to be superfluous. Consequently, from then on Λ was often set to $\Lambda = 0$. However, in the framework of modern field theories the cosmological constant

is interpreted as the vacuum energy density ρ_V. The following equation holds [McC51, Zel68]

$$\Lambda = \frac{8\pi G}{c^2}\rho_V. \tag{9.30}$$

The case $\Lambda = 0$ corresponds to the assumption that the vacuum does not contribute to the energy density of the universe. This picture corresponds to the ideas of classical physics, and must be modified in the framework of quantum theory, as, for example, the typical zero-point energy of the harmonic oscillator shows. In quantum field theory, the vacuum contains various quantum fields. These are in the state of lowest energy, which is not necessarily equal to zero. Thus, many field-theoretical approaches require a non-vanishing zero-point vacuum energy (see e.g. [Gro89, 90, Kol90, Wei89]).

Taking into account a non-zero cosmological constant, by virtue of

$$\rho_{c0} = \frac{3H_0^2 - \Lambda c^2}{8\pi G} \qquad \Omega_0 = \frac{8\pi G\rho_0}{3H_0^2 - \Lambda c^2} \tag{9.31}$$

we obtain a lower critical density and a larger density parameter than expected from (9.23) and (9.24), respectively. Astronomical observations based on counts of galaxies give an upper bound of $\Lambda \leq 3 \times 10^{-56}$ cm^{-2} for the current value of the cosmological constant [Abb88]. It is also easy to estimate an upper bound, since the critical density ρ_{c0} cannot be negative. Thus, the factor Λc^2 in (9.31) can be at most as large as $3H_0^2$. This leads to:

$$\Lambda \leq \frac{3H_{0,\max}^2}{c^2} \approx 3.5 \times 10^{-56}\ \text{cm}^{-2} \tag{9.32}$$

where we have used a value of 100 km s^{-1} Mpc^{-1} for $H_{0,\max}$. While a non-vanishing cosmological constant appears necessary for the interpretation of the early phase of evolution (see e.g. [Gro89, 90, Kol90, Lin90]), a number of authors have come to the conclusion that $\Lambda \neq 0$ could also have played a role in the later development of the universe [Blo84, Pee84, Tur84, Kla86b, Tay86, Pri87, Chu88, Gro89, 90].

A cosmological constant of size

$$\Lambda = \frac{3H_0^2}{c^2}\left(1 - \frac{\Omega(\Lambda = 0)}{\Omega(\Lambda \neq 0)}\right) \tag{9.33}$$

could simulate a $\Omega(\Lambda = 0)$ although a $\Omega(\Lambda \neq 0)$ is in force. An apparent $\Omega(\Lambda = 0)$ determined from ρ_0 would be sufficient to generate a $\Omega = 1$ as required in inflationary models, provided the cosmological constant had a value

$$\Lambda = \frac{3H_0^2}{c^2}(1 - \Omega(\Lambda = 0)). \tag{9.34}$$

Use of the numerical values $H_0 = 75 \pm 25$ km s^{-1} Mpc^{-1} and $\Omega_{0,\mathrm{obs}} = 0.2 \pm 0.1$ gives

$$\Lambda = (1.6 \pm 1.1) \times 10^{-56}\ \mathrm{cm}^{-2}. \tag{9.35}$$

A vacuum energy density corresponding to this value could resolve the 'contradiction' between the observed density parameter and the density parameter $\Omega = 1$ required by various theories.

In addition to the introduction of the cosmological constant, other models exist which manage, at least for part of the problem, without postulating dark matter. We mention two such approaches here.

MOND theory. M Milgrom's MOND theory (modified Newton dynamics) (see the survey in [San90]) assumes that the law of gravity is different from the normal Newtonian form and looks as follows

$$a_{\mathrm{G}} = \frac{GM}{r^2} + \frac{\sqrt{GMa_0}}{r}. \tag{9.36}$$

According to this, the force of attraction would be larger and would have to be compensated by a faster periodic movement, which could explain the flat rotation curves. Such deviations from the known law of gravity are discussed in detail in chapter 11.

Time-dependent gravitational constant. A time-dependent gravitational constant $G(t)$ would have a large effect on the genesis of galaxies [Sta92c]. However, to date, precise measurements have not shown a variation of G with time. The time-dependence of natural constants and the closely related non-conservation of energy are discussed in chapter 12.

9.2.2 Baryonic dark matter

The most obvious candidate for dark matter would be ordinary baryonic matter, which would have to be both abundant and dark. One possibility would be interstellar or intergalactic gas. However, this should give rise to typical emission or absorption lines, which have not been observed.

Another candidate would be so-called *brown dwarfs*, i.e. bodies with mass very much smaller than the mass of the Sun ($M < 0.08 M_\odot$). The gravitational pressure in the interior of these objects is not sufficient to generate temperatures which would permit a fusing of protons into helium. Since no nuclear fusion is involved, brown dwarfs only radiate weakly, apart from during the initial phase of their development. The planets could also be counted among this group. However, because of our lack of knowledge of the genesis of stars and planets and because of the constraint of photometric detectability to distances of only a few light years, it is extremely difficult to estimate the number of these objects.

Very compact objects from the end stages of the development of stars, such as white dwarfs, neutron stars or black holes, might also constitute dark matter. Since, in the course of its lifetime, practically every star reaches one of the three end stages, a large part of the mass of earlier, heavier stars must be present in dark form as white dwarfs, neutron stars or black holes. Part of this material is returned to the interstellar space via supernova explosions and other processes and plays a role in the formation of other new stars. In this discussion, stars with mass $M < 0.9M_{\odot}$ must be ruled out, since these have a lifetime which is greater than the age of the universe, so that these objects have not yet reached the end stage.

Upper bounds for the maximum possible density of baryonic matter in the universe may be derived from studies of the primordial nucleosynthesis (see chapter 3) which began approximately 3 min after the big bang (see [Yan79a, 84, Blo84, Kol90]). Measurements of the current abundance of deuterium are particularly important

$$\left(\frac{D}{H}\right)_0 \approx 10^{-5} \tag{9.37}$$

since deuterium is essentially only generated during the primordial nucleosynthesis. While deuterium occurs later also as an intermediate product in fusion reactions, the overall amount is not increased further. Calculations of the early nucleosynthesis give an upper bound of

$$\Omega_{0,\text{bary}} < 0.1\text{–}0.2 \tag{9.38}$$

for the density of possible baryonic matter in the universe. This includes all matter which was formed during the nucleosynthesis in the early universe. This value agrees well with the value obtained from the rotation pattern of galaxies (9.4*a*).

On the other hand, it is clearly recognized that baryonic dark matter alone cannot satisfy the requirement $\Omega = 1$ derived from inflationary models. The problem of the genesis of the galaxies is also unresolved. This raises the question of the existence of *non-baryonic* dark matter. In particular, this appears to be necessary when one requires $\Omega = 1$ for a vanishing cosmological constant (see section 9.2.1).

9.2.3 Non-baryonic dark matter

Theoretical models provide a large selection of possible candidates for additional non-baryonic invisible matter, including: light and heavy neutrinos, supersymmetric particles from SUSY models, axions, cosmions, magnetic monopoles, Higgs particles and many others (table 9.1), some of which we shall discuss below. The term WIMP (= weakly interacting massive particles)

is used to describe particles with mass > GeV/c^2 which only take part in the weak interaction. The large-scale distribution of galaxies and the development of the large-scale structure of the Galaxy has recently been investigated [Tag92, Dar92a] in the light of the new COBE results on the cosmic microwave radiation (3 K radiation) [Wri92, Smo92] (see section 3.1) and of the QDOT IRAS red-shift survey [Fish93]. The authors conclude on the basis of the analysis of various structure formation models, that there is *only one* satisfactory model of the universe with $\Omega = 1$, namely a model with mixed dark matter: 70% in the form of cold dark matter and 30% in the form of hot dark matter (see below), the latter in the form of two massless neutrinos and one neutrino with a mass of 7.2 ± 2 eV (1σ). This means a revival of the earlier disqualified (see e.g. [Schra85a]) mixed dark matter model (for a detailed discussion see [Sil93]).

9.2.3.1 Light neutrinos

Unlike all subsequent candidates for dark matter, neutrinos have the advantage that they are known to exist. Even their abundance in the universe is approximately known. For neutrinos to be a candidate for dark matter, they must clearly have a mass. To achieve the critical density of the universe the neutrino masses m_ν must lie either in the region of a few GeV/c^2 or in the region from 10 to 100 eV/c^2 [Kol86b, 90] (see also chapter 3). The heavy neutrinos are allowed, since the cosmologically relevant product $m_\nu \exp(-m_\nu / kT_f)$ becomes small for large masses. Here T_f denotes the temperature at which the heavy neutrinos decouple from the thermal equilibrium. The Boltzmann factor describes the abundance of a neutrino of mass m_ν relative to that of a massless neutrino. A more detailed treatment of this topic, taking into account a possible instability of the neutrino, may be found in [Gro89, 90] and references therein.

We shall initially consider light neutrinos; heavy neutrinos will be discussed in brief in the next section. The neutrino density n_ν for each individual type of neutrino in the universe is related to the photon density n_γ by the equation [Boe87, 92, Gro89, 90]

$$n_\nu = \frac{3}{11} n_\gamma. \tag{9.39}$$

Strictly speaking, this expression is only valid for light Majorana neutrinos (under certain circumstances a further statistical factor of two must be introduced for Dirac neutrinos). The photon density may be determined from the 3 K background radiation and amounts to $n_\gamma \approx 400$ cm^{-3}. It turns out (see e.g. [Ell88b, Kol90]) that a neutrino mass density in the region of the critical density is obtained if the following condition is satisfied

$$\sum_\nu \frac{g_\nu}{2} m_\nu \approx 100 \text{ eV } c^{-2} \left(\frac{\rho_\nu}{\rho_c}\right) h_0^2 \tag{9.40}$$

Table 9.1. Summary of non-baryonic candidates for dark matter (DM) (from [PRI88]).

Candidate particle	Approx. mass	Predicted by	Astrophysical effect
$G(R)$	-	Non-Newtonian gravity	Apparent DM on large scales
Λ (cosmological const.)	-	General relativity	$\Omega = 1$ without DM
Axion, majoron, Goldstone boson	10^{-5} eV	QCD; Peccei–Quinn symmetry breaking	Cold DM
Conventional neutrino	10–100 eV	GUTs	Hot DM
Light Higgsino, photino, gravitino axino, sneutrino[a]	10–100 eV	SUSY/DM	
Para-photon	20–400 eV	Modified QED	Hot, warm DM
Right-handed neutrino	500 eV	Superweak interaction	Warm DM
Gravitino, etc[a]	500 eV	SUSY/SUGRA	Warm DM
Photino, gravitino,	keV	SUSY/SUGRA	Warm/cold DM
Axino, mirror particles, Simpson neutrino[a]			
Photino, sneutrino,	MeV	SUSY/SUGRA	Cold DM
Higgsino, gluino heavy neutrino[a]			
Shadow matter	MeV	SUSY/SUGRA	Hot/cold (like baryons)
Preon	20–200 TeV	Composite models	Cold DM
Monopoles	10^{16} GeV	GUTs	Cold DM
Pyrgon, maximon, Perry pole, newtorite, Schwarzschild	10^{19} GeV	Higher-dimensional theories	Cold DM
Supersymmetric strings	10^{19} GeV	SUSY/SUGRA	Cold DM
Quark nuggets, nuclearites	10^{15} g	QCD, GUTs	Cold DM
Primordial black holes	10^{15}–10^{30} g	General relativity	Cold DM
Cosmic strings, domain walls	$(10^{8}$–$10^{10})M_{\odot}$	GUTs	Supported formation of galaxies, may not provide a large contribution to Ω

[a] Of these various supersymmetric particles, predicted by different supersymmetric (SUSY) theories and supergravity (SUGRA), only one, the lightest, can be stable and contribute to Ω; however, at the current time, the theories say very little about the nature and the expected mass of the particles.

where g_ν is a statistical factor which gives the number of helicity states per neutrino flavour. For Majorana neutrinos, this factor has value two. Dirac neutrinos could have a statistical factor of four. However, it is generally assumed that the right-handed component decouples considerably earlier from the equilibrium so that we may also suppose that $g_\nu = 2$ in the Dirac case.

Since the neutrino density is of the same order of magnitude as the photon density there are around 10^9 times more neutrinos than baryons; thus, even a small neutrino mass could dominate the dynamics of the universe. Neutrino masses of [Har89]

$$m_\nu c^2 \approx 15\text{–}65 \text{ eV}/N_\nu \tag{9.41}$$

where N_ν denotes the number of light neutrino flavours, are needed to achieve $\Omega = \rho_\nu/\rho_c = 1$. Experimental upper bounds for the masses of the three known neutrino flavours are (see the discussion in chapter 6)

$$m_{\nu_e} < 7.2 \text{ eV}/c^2 \qquad m_{\nu_\mu} < 250 \text{ keV}/c^2 \qquad m_{\nu_\tau} < 31 \text{ MeV}/c^2. \tag{9.42}$$

Thus, the electron neutrino is practically ruled out as a candidate for the dominant fraction. The experimental data for the other two flavours are less restrictive, so that muon and tau neutrinos remain in question.

Neutrinos decouple in the early universe after approximately 1 s at a temperature of 10^{10} K (corresponding to an energy of 1 MeV). At that time they possessed relativistic energies and are thus referred to as *hot dark matter*.

Neutrinos may also have contributed to the genesis of galaxies. In an expanding universe dominated by particles of mass m_i, according to the Jeans criterion, the mass which may collapse under gravity is [Tri87]

$$M_J \approx 3 \times 10^{18} \frac{M_\odot}{m_i \text{ [eV]}}. \tag{9.43}$$

In a neutrino-dominated universe the necessary clumping may have set in at a relatively late stage; the first structures would have corresponded in size to superclusters. Thus, galaxy clusters and galaxies would have developed by fragmentation of these original structures (*top-down model* [Oor83]). However, there are problems as far as the development of very small structures such as dwarf galaxies is concerned. The generation of relatively very massive compression also requires one to take into account the Pauli principle for fermions. For details, we refer readers to [Bör88, Kol90, Schra85a, Sil93].

9.2.3.2 Heavy neutrinos

According to the results of LEP and SLC relating to precision measurement of the decay width of the Z^0 there are only three light neutrino flavours and the

existence of heavier neutrinos may be ruled out up to the kinematic limit of approximately 45 GeV/c^2.

When neutrinos with such a large mass decoupled they would already have attained non-relativistic velocities; thus they are referred to as *cold dark matter*. Heavy neutrinos could have allowed an early gravitational compression of matter. According to (9.43) smaller structures were formed first. Galaxy clusters and superclusters were formed later by accumulation of individual groups of galaxies (*bottom-up model* [Efs90, Schra85a]). For a detailed discussion see [Sil93].

9.2.3.3 Axions

Axions are hypothetical particles which occur in connection with the strong CP problem (θ-problem). The existence of this pseudoscalar particle follows from the breaking of the chiral Peccei–Quinn symmetry [Pec77]. This is described in some more detail in chapters 1 and 5. For a detailed discussion see [Raf90, Sik90, Tur90].

This hypothetical particle has mass

$$m_a \simeq 0.62 \text{ eV} \frac{10^7 \text{ GeV}}{f_a}. \tag{9.44}$$

The interactions with fermions and with the gauge bosons are described by the following coupling constants, respectively

$$g_f \sim \frac{m_f}{f_a} \qquad g_b \sim \left(\frac{\alpha}{4\pi}\right) \frac{1}{f_a}. \tag{9.45}$$

f_a is the decay constant of the axion which is given by the vacuum expectation value of a Higgs field. Since f_a is a free constant between the electroweak and the Planck scale, this gives a range of approximately 18 orders of magnitude for possible axion masses. We distinguish between DFSZ axions [Zhi80, Din81] and hadronic axions [Kim79], depending on whether axions couple to electrons directly or in first order perturbation theory. Axions are generally counted as cold dark matter. For their density to be smaller than the critical density, we require [Abb83, Din83, Pre83]

$$f_a \leq 10^{12} \text{ GeV}. \tag{9.46}$$

The standard Peccei–Quinn axion with $f_a \approx 250$ GeV has been experimentally ruled out; other variants with smaller masses and correspondingly larger coupling parameters are considerably restricted by a variety of data, including primarily astrophysical data [PDG90, Jon90, Raf90, Tur90] (see section 9.3.1).

9.2.3.4 Supersymmetric particles

The idea of supersymmetry was discussed in chapter 1. Most of the supersymmetric theories contain one *stable* particle, which is a new candidate for dark matter. The existence of a stable supersymmetric partner particle results from the fact that these models include a conserved multiplicative quantum number, the so-called R parity, which takes value $+1$ for particles and -1 for the corresponding supersymmetric partner. According to this conservation theorem SUSY particles can only be generated in pairs. SUSY particles may only decay into an odd number of SUSY particles. Consequently, the lightest supersymmetric partner particle must be stable.

There is a possibility to violate the conservation rule for R parity. The quantum number R is related to the baryon number B and the lepton number L by

$$R = (-1)^{3B+L+2S} \tag{9.47}$$

where S stands for the particle spin. In other words, a violation of B and/or L may lead to non-conservation of the R parity. However, there exist very sharp bounds for an R violation [Hal84, Lee84b, Hir95].

It is conjectured that the lightest supersymmetric particle (LSP) does not take part in either the electromagnetic or the strong interaction. Otherwise, it would have condensed with normal matter and would be detectable today as an unusual heavy particle. For the calculated abundance of such LSP, normalized by the abundance of the proton, we have [Dov79, Wol79b]

$$\frac{n(\text{LSP})}{n(p)} \sim \begin{cases} 10^{-10} & \text{strong interaction} \\ 10^{-6} & \text{electromagnetic interaction.} \end{cases} \tag{9.48}$$

However, these values contradict the experimental upper bounds [Ell88a]

$$\frac{n(\text{LSP})}{n(p)} < 10^{-15}\text{–}10^{-30}. \tag{9.49}$$

These values are mass dependent and relate to the region 1 GeV $< m_{\text{LSP}}c^2 <$ 10^7 GeV. Thus, it is likely that the lightest SUSY particle only takes part in the weak interaction in addition to gravity.

Possible candidates for the neutral lightest supersymmetric particle include the photino ($S = 1/2$), the higgsino ($S = 1/2$), the zino ($S = 1/2$), the sneutrino ($S = 0$) and the gravitino ($S = 3/2$). In most theories, it is likely that the LSP is a mixture of the above SUSY particles with half-integral spin, which are generally referred to as *gauginos*. The mass of this so-called *neutralino* should preferably be greater than 10 GeV/c^2. SUSY particles as dark matter are of interest, since they occur in a totally different context and are not specially introduced to solve the problem of (non-baryonic) dark matter. For detailed discussions we refer to [Ell93, 94, Bed94a,b, Bot94].

9.2.3.5 *Cosmions*

Cosmions were primarily introduced to solve the solar neutrino problem [Fau85, Sper85]. Because of their high velocity, these particles penetrate the surface of a star practically unhindered. In the centre, they collide with atomic nuclei. If the energy loss is sufficiently large they may not leave the star again and accumulate there over the course of time. In the interior of the Sun trapped cosmions affect the energy transport in the Sun and thus contribute to a decrease in the central temperature. This would result in a lower production rate for ^{8}B neutrinos and would explain why the neutrino flux measured on the Earth is less than expected (see chapter 7). To solve the neutrino problem the mass must lie in the range from 4 to 11 GeV/c^2 and the reaction cross-section for the interaction between cosmions and matter must have order of magnitude 10^{-36} cm^2 [Bou89]. However, experimental data appear to rule this solution out (see section 9.3 and [Cal90a, 91c, 92]).

9.2.3.6 *Topological defects of space–time*

In addition to the 'true' particles discussed hitherto, so-called topological defects may also contribute to the dark matter. It is conjectured that at the end of the GUT symmetry which was present in the early universe at

$$t \approx 10^{-36}\ \text{s} \qquad E \approx 10^{15}\ \text{GeV} \qquad T \approx 10^{28}\ \text{K} \tag{9.50}$$

a symmetry breaking occurred which led to the separation of the interactions described by the groups SU(3) and SU(2) $\otimes$ U(1). The 24-dimensional Higgs field aligned itself (where the orientation phase angles for the spontaneous symmetry breaking are random) as described in chapter 1. As a consequence of this phase transition spatial domains with different alignments must have been formed. These domains grew with time and ultimately touched one another.

According to current ideas topologically stable defect points developed on the bounding surfaces where different orientations met (figure 9.5). These could have dimension from zero to three. They would consist only of vacuum, but of the vacuum from the previously unbroken symmetry. After the symmetry breaking this original vacuum has a very large energy and matter density.

The most important defects are point-like defects. These should carry an isolated magnetic charge, i.e. they are magnetic monopoles ['tHo74]. Their mass is related to the temperature of the phase transition and is in the region around 10^{16} GeV/c^2. Magnetic monopoles are discussed in detail in a different chapter. Until now, the existence of these objects has not been detected despite an intensive search.

Analogously to the magnetic monopoles, linear defects may also have formed; these are known as *cosmic strings*. These thread-like objects have a

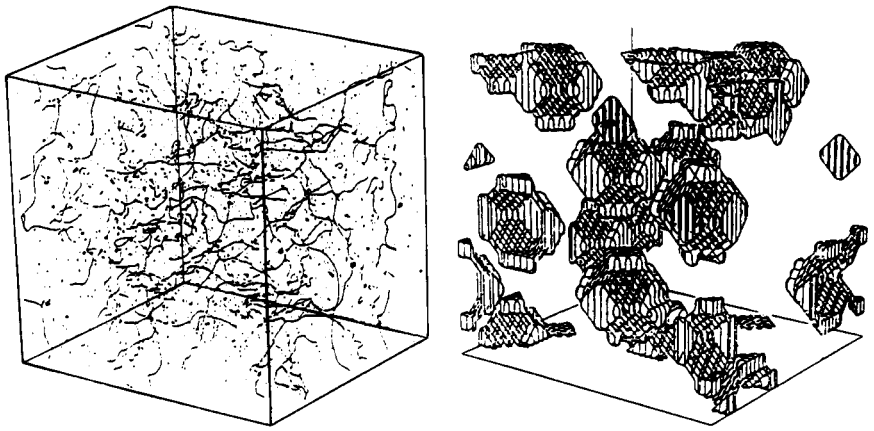


Figure 9.5. Topological defects: cosmic-string and domain-wall simulation (see [All90, Pre89]).

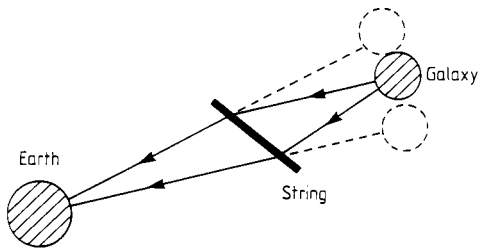


Figure 9.6. Gravitational lensing effect. Cosmic strings would bend the space surrounding them in such a way that objects behind them would appear as double images.

typical linear mass density of 10^{22} g cm^{-1} and may be either open or closed. As a result of their gravitational effect they may have served as condensation seeds for the genesis of galaxies [Alb85, Vil87, 88].

As a result of their large mass, it may be possible to detect such strings through the gravitational lens effect. Strings would curve the space around them in such a way that the objects behind them would appear as double images (see figure 9.6). The light from very remote galaxies could be deflected by the string according to the laws of the general theory of relativity. The observer on the Earth would see two adjacent mirror-image galaxies with identical spectral properties. This gravitational lens effect has already been found for remote quasars, for which a galaxy in the foreground serves as a gravitational lens (see [Kay92]).

The possibility that cosmic strings are superconducting is also under

discussion. Electrically charged particles such as electrons would be massless in the symmetric vacuum of the string, since they obtain a mass only as a result of symmetry breaking via the Higgs mechanism. Thus, particle–antiparticle pairs, which move with the velocity of light, may be created for a very low expenditure of energy. This would result in a flow of current with vanishing resistance. Superconducting strings could be excited by interaction with charged particles to emit radio waves [Vil87].

Higher-dimensional defects, including two-dimensional 'domain walls' and, in particular, three-dimensional defects or 'textures', are also under discussion [Tur89a].

9.2.3.7 Other exotica

There are practically no limits to imagination for ideas for new candidates for dark matter. In what follows we shall briefly mention just two exotic sounding approaches.

Shadow matter. By assuming that fermions are one-dimensional extended objects, superstring theories attempt to apply the success of the supersymmetric models in eliminating the divergences in quantum field theory of the non-gravitational forces to the area of gravitation, and to penetrate into energy regions above the Planck mass. Mathematically, anomaly-free superstring theories can only be given for the gauge groups SO(32) and $E_8 \otimes E_{8'}$. The latter splits into two sectors, one of which describes the normal matter while the other describes the shadow matter ($E_{8'}$). The two sectors may only interact with one another via gravity [Kol85, Gre86a].

Quark nuggets. Quark nuggets were proposed in 1984 [Wit84]. These are stable macroscopic objects of quark matter, consisting of u, d and s quarks, with densities in the area of the nuclear density of 10^{15} g cm^{-3}, and with masses ranging from several GeV/c^2 to the mass of a neutron star. They are formed in a hypothetical QCD phase transition, but are generally considered to be very unlikely.

9.3 EXPERIMENTS TO DETECT DARK MATTER

To end this chapter on dark matter we shall give a number of examples of the wide range of methods used to detect the above candidates. We shall not discuss the determination of a possible neutrino mass for the light neutrino types here, since this has been done in detail in previous chapters. The detection of magnetic monopoles has also been discussed. A survey of the detection of dark matter may be found in [Prim88, Smi90b, Cal91c, 92, 93].

The general difficulty as far as the detection of one of the above particles is concerned arises from the fact that they are electrically neutral and only take part in the weak interaction. There are two main detection methods (see figure 9.7), namely direct and indirect detection. In direct detection the consequences of the interaction with electrons or atomic nuclei are studied using Earth-bound apparatus. In indirect methods, an attempt is made to detect the flux of secondary particles which arises, for example, due to the annihilation of solar or galactic dark matter.

9.3.1 Experiments to detect the axion

Because of the small coupling constant and the correspondingly small number of events to be expected, detection of the axion is very difficult. An overview of experiments relating to the axion may be found in [Raf90, Tur90]. Despite great efforts these hypothetical particles have not been detected; only a few

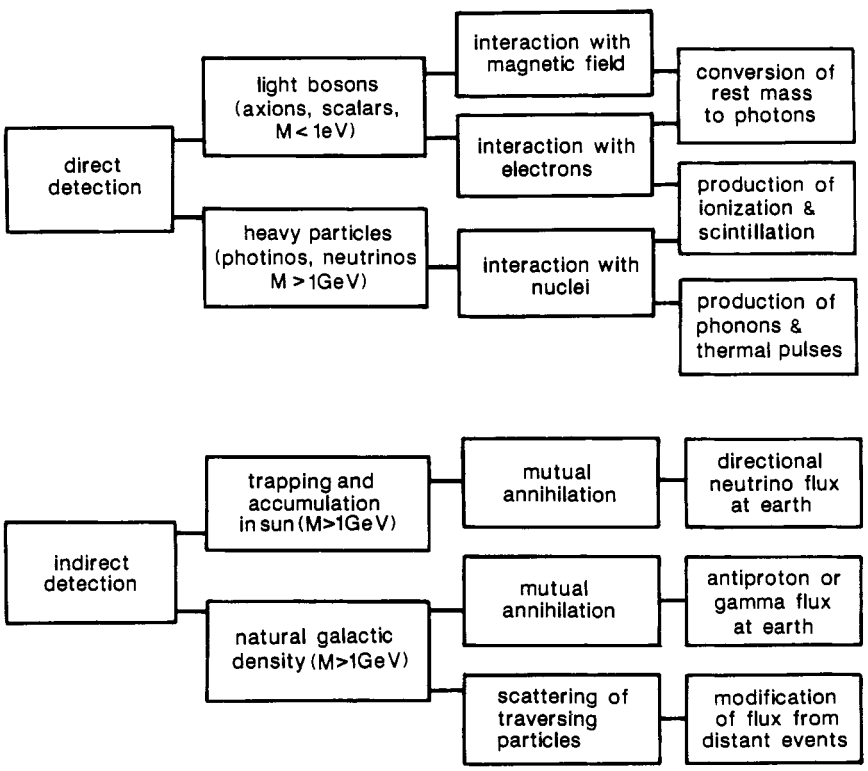


Figure 9.7. Overview of the various methods of detecting dark matter (after [Smi90b]).

parameter areas (axion mass, coupling constant) have not yet been ruled out by experiments. Here, we shall only sketch two possible methods of detection and include some remarks about astrophysical restrictions.

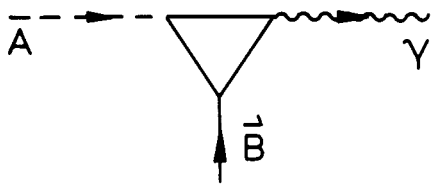


Figure 9.8. Interaction of an axion (A) with a magnetic field.

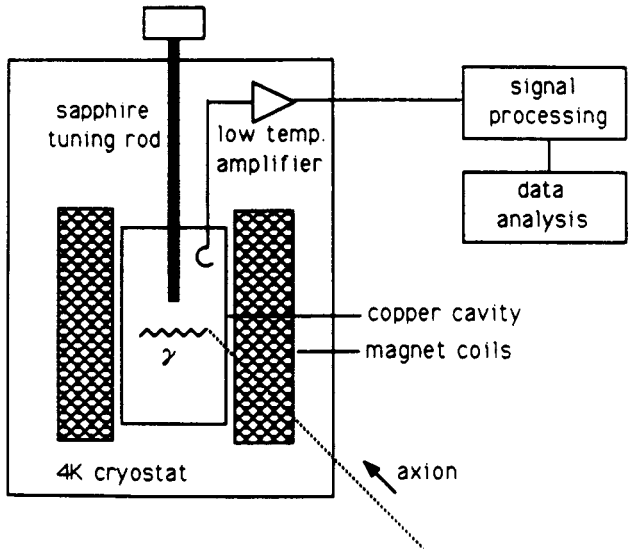


Figure 9.9. Experiment to detect axions based on the reaction shown in figure 9.8. A copper cavity resonator located in a magnetic field is introduced into a 4 K cryostat (from [Smi90b]).

[Sik83, 85] pointed out that the mass of an axion may be converted into an observable photon by interaction with a magnetic field (figure 9.8). The experiment shown schematically in figure 9.9 was based around this fact. A copper cavity resonator, which is located in a magnetic field, is placed in a 4 K cryostat. If the frequency of the photons resulting from the interaction between axions and the magnetic field corresponds to the resonator frequency, these photons gather together in the resonator and yield a detectable signal. An axion mass in the region from 10^{-5} to 10^{-3} eV/c^2 would correspond to a

frequency of 2–200 GHz in the microwave area. In this experiment, the whole of the accessible frequency range has to be sampled systematically and accurately, since the expected line is extremely sharp ($\Delta E/E \approx 10^{-6}$). A magnetic field strength of 5 T and a resonator volume of 0.01 m^3 are used. No positive signal has yet been detected [Wue89, Smi90b].

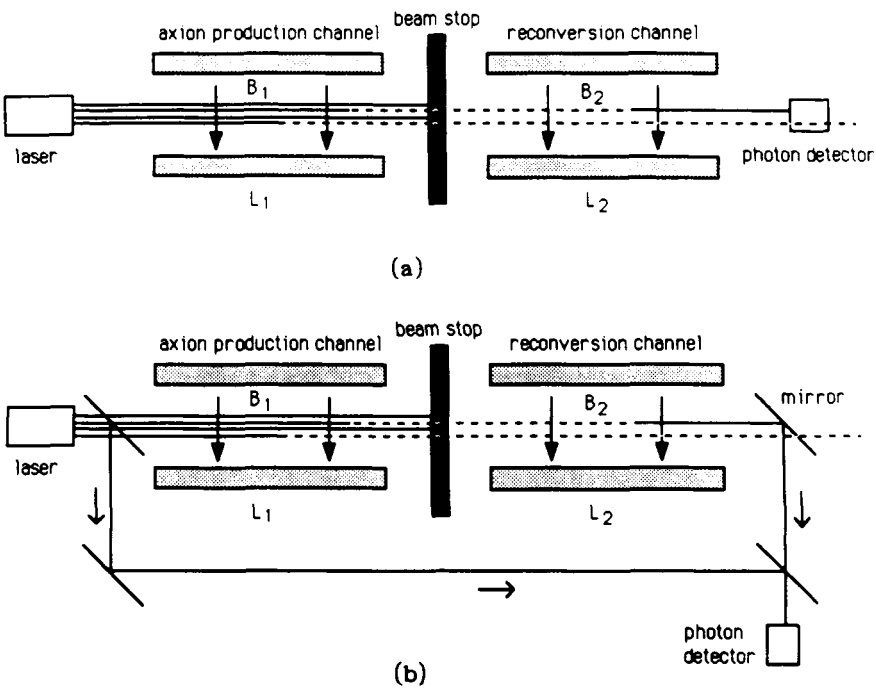


Figure 9.10. Generation and detection of axions with a high-performance laser and using the process shown in figure 9.8: (*a*) direct detection of the regenerated photons; (*b*) interference between regenerated and original photons (from [Smi90b]).

Another very interesting detection method is illustrated in figure 9.10 [Bib87, Smi90b]. A high-energy laser beam passes through a transverse magnetic field. A number of photons may be converted into axions there. A beam stopper, which is opaque to the laser light, may let the axions through unhindered. The latter reach a second transverse magnetic field in which some of the axions are converted into photons with the original energy. The regenerated photons could be detected using a photomultiplier. The sensitivity may be increased by making the original laser beam interfere with the regenerated photons.

Axions with masses greater than 1 eV/c^2 are ruled out primarily by star-

evolution calculations, since such axions would form an additional energy-loss mechanism and lead to a more rapid evolution than that observed. Strictly speaking, this is only true for the DFSZ axion; a mass in the eV/c^2 region remains possible for hadronic axions. The duration of the neutrino pulse from the supernova SN1987A also appears to rule out masses greater than 10^{-3} eV/c^2. Cosmological arguments rule out masses less than 10^{-5} eV/c^2, since otherwise too much matter in the form of axions would be present. Thus, the region from 10^{-3} to 10^{-5} eV/c^2 remains together with the region of a few eV/c^2 for hadronic axions (see [Kol90, Raf90]). Experimental results relating to solar axions have been obtained from double-beta decay experiments with Ge detectors; these rule out masses greater than 14.4 eV/c^2 [Avi88]. Recently, it has been suggested that Bragg scattering be used for the investigation of solar axions [Pas94].

9.3.2 The detection of WIMPs (Weakly Interacting Massive Particles)

WIMPs could be detected from the recoil energy which is transmitted, for example to an atomic nucleus, by the weak interaction. The collision of such a particle of mass m_d gives a recoil spectrum for the nucleus which typically lies in the region of $10^{-6}m_dc^2$. For the calculation of cross sections for elastic scattering of neutralinos on nuclei and for the possible restriction of the parameter space of supersymmetric GUT models by search for these particles we refer to [Ell93, 94, Bed94a,b, Bot94].

To detect WIMP candidates with mass from 1–10 GeV/c^2 requires the measurement of recoil energies in the region of a few keV. This can be done using ionization or the thermal effects. We now give a brief discussion of low-temperature methods and the measurement of ionization in semiconductors (for details see [Prim88, Smi90b, Cal91b,c, Boo92, Fif93, Pre93, Sad94]).

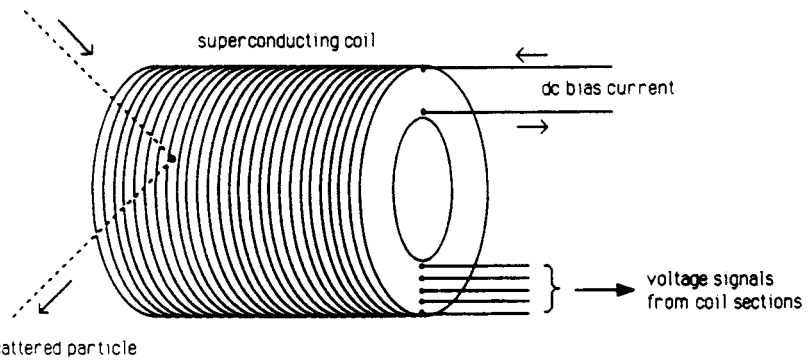


Figure 9.11. On the detection of WIMPs: detection of the nuclear recoil energy with superconducting coils.

9.3.2.1 *Low-temperature methods*

One of these low-temperature methods uses superconducting coils (see figure 9.11). A WIMP meets a coil, which is operated at a temperature slightly below the critical temperature. The recoil energy is sufficient to destroy the local Cooper pairs in the conductor, which should generate a measurable voltage signal.

One possible implementation of this technique (see [Pre87, 90, 93, Mös91]) involves a detector consisting of superconducting granules (type I superconductor) with a diameter of 5–10 μm which are embedded in a dielectric carrier material. The whole arrangement is located in a magnetic field. According to the Meissner–Ochsenfeld effect, below the critical temperature the magnetic lines of force are forced out from the granule. The temperature and the outer magnetic field are adapted so that the superconductor is in a metastable state. If an interaction takes place the recoil energy is sufficient to return the corresponding granule to a state of normal conduction. The resulting change in the flux can be detected.

It is also possible to attempt to determine the recoil energy E_R calorimetrically, i.e. to detect the warming of a crystal due to E_R. For this, the crystal must be cooled to very low temperatures (in the region of a few mK). Because of the dependence of the thermal capacity on T^3

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V \tag{9.51a}$$

$$\sim \left(\frac{T}{\Theta_D}\right)^3 \qquad (T \ll \Theta_D) \tag{9.51b}$$

below the Debye temperature, the recoil energy $E_R = \Delta Q$ would give rise to a relatively large increase in the temperature. To increase the sensitivity, the target should be divided into a large number of regions with very small volumes, since the temperature increase is greater the smaller the volume is. In addition, it is customary to use a material with the lowest possible thermal capacity and to work in the mK region. The first such detectors, including a 334 g TeO_2 crystal, are now in use in double-beta decay experiments [Giu91, Ale94].

9.3.2.2 *Ionization in a semiconductor*

When the interaction (elastic scattering) of a WIMP takes place in a semiconductor, the recoil nucleus may generate a number of electron–hole pairs. As a result of the applied field, these free charge carriers move in opposite directions out of the depletion region and generate a current pulse. Since only very few events are expected, the detector used must be well screened against cosmic rays and other external effects. Germanium semiconductor

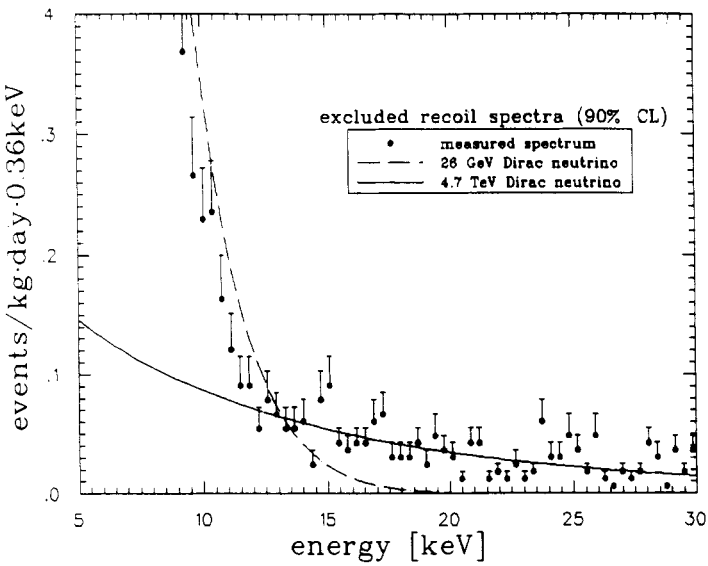


Figure 9.12. Lower-energy part of the spectrum recorded by a detector enriched with ^{76}Ge in 166 kg days. Error bars denote 90% confidence level. The figure also shows the expected recoil spectra for Dirac neutrinos with masses of 26 GeV/c^2 and 4.7 TeV/c^2 (from [Bec94, Kla94]).

counters of this type are used to study the double-beta decay of ^{76}Ge (see chapter 6). Figure 9.12 shows the low-energy spectrum of one of these experiments (Heidelberg–Moscow experiment) which is being carried out in the Gran Sasso underground laboratory [Bec94, Kla93c,d, 94]. The figure also shows the expected recoil spectra for Dirac neutrinos with masses of 26 and 4700 GeV/c^2. Such evaluations give upper bounds for the masses and coupling strengths of weakly interacting massive particles. The detector noise forms a limitation of sensitivity for lower masses. Figure 9.13 shows exclusion diagrams for WIMPs (masses and cross sections) which were obtained from such germanium experiments [Bec94, Kla94]. Figure 9.14(*a*) (from [Ell90a, Schr90d]) shows corresponding results of another ^{76}Ge experiment [Cal88a, 91c]. The sensitivity of these experiments clearly exceeds that of the most recent LEP results [ALE92], which are given for comparison. More precisely, they complement the LEP experiments, which are more sensitive to lighter particles, while the sensitivity of the semiconductor detector extends to very large masses.

We note that these semiconductor experiments are based on the *coherent* scattering of WIMPs, for which

$$\sigma^{\mathrm{coh}} \simeq N^2 \tag{9.52}$$

where N is the number of neutrons in the scattering nucleus. In addition to such a coherent interaction a *spin-dependent* interaction is also possible.

A germanium detector in which the isotope ^{73}Ge is enriched would be of particular interest here since this isotope has a nuclear spin of $J = 9/2$, while the other stable Ge isotopes have a vanishing nuclear spin. Such a detector could be used to detect spin-dependent interactions. The class of candidates for dark matter which interact largely spin dependently includes Majorana neutrinos and the LSP (light supersymmetric particle). Figure 9.14(*b*) shows the possibilities for detecting WIMPs using a 2 kg detector of ^{73}Ge enriched to $\approx 85\%$ (natural abundance 7.8%) (see [Kla91a,d]) in comparison with the existing LEP data. The evaluation of these experiments required not only knowledge of the nuclear matrix elements but also of the spin structure of the proton (see [Ell88, Prim88, Iac91, Nik93, Bed94]).

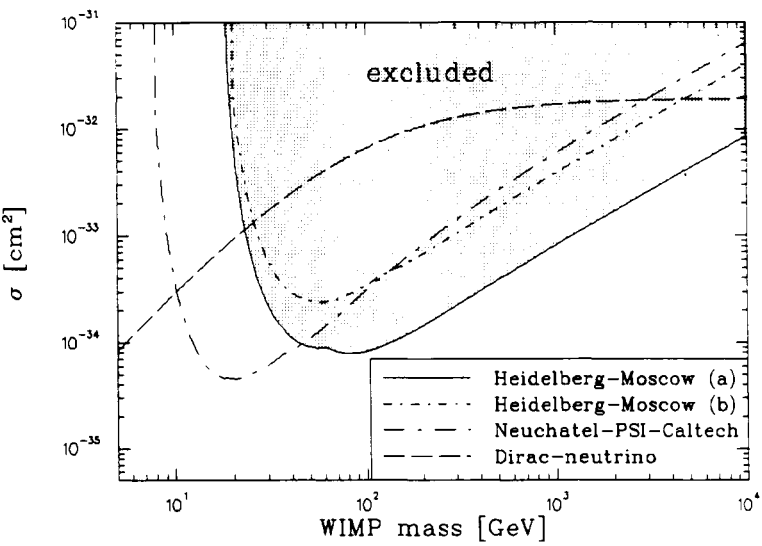


Figure 9.13. Exclusion diagram for WIMPs from the Heidelberg–Moscow experiment [Bec94, Kla93b, 94]: WIMPs with masses and scattering cross sections σ in the shaded area are excluded, e.g. Dirac neutrinos with masses between 26 GeV/c^2 and 4.7 TeV/c^2 and with standard coupling. For coherently interacting WIMPs the cross section has been corrected for loss of coherence (line (*b*)). Since the cross section for sneutrino–nucleus interaction is just four times the cross section for Dirac neutrino–nucleus interaction, the Heidelberg–Moscow experiment also excludes heavy sneutrinos of the minimal supersymmetric standard model as candidates for dark matter [Fal94]. The figure also shows results of the Gotthard experiment [Reu91, Tre91].

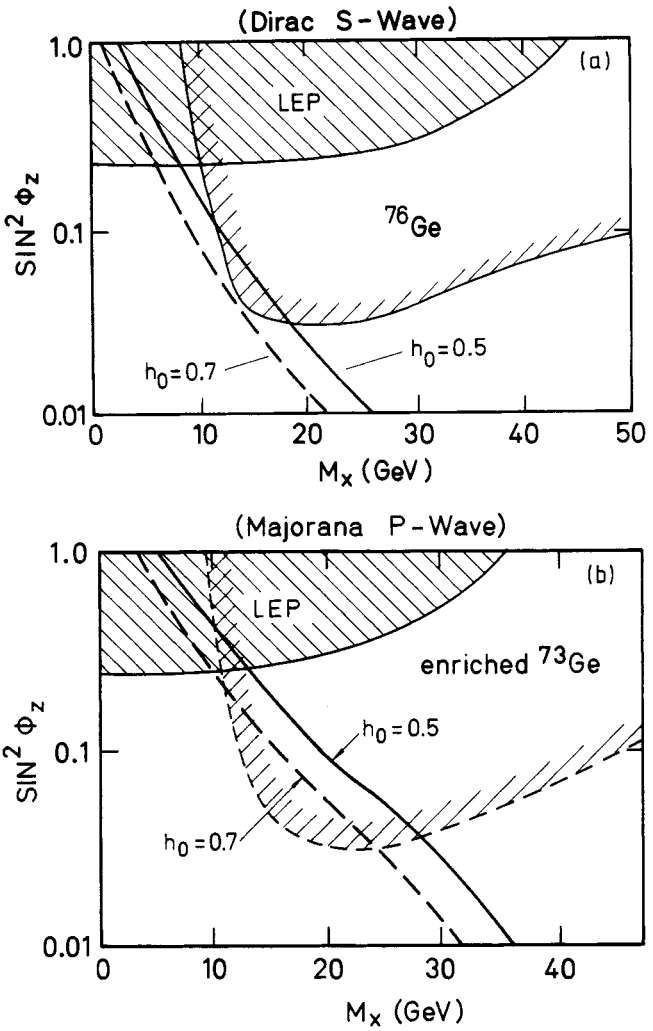


Figure 9.14. Exclusion diagrams for WIMPs with mass M_X and coupling strength $\sin^2\theta_Z$ (relative to the coupling of the neutrino to the Z^0) for Dirac (*a*) and Majorana particles (*b*). In (*a*) the region above the shaded curve is excluded by experiment [Cal91c], while the exclusion boundary in (*b*) corresponds to a planned experiment (after [Ell90a, Schr90d, Kla91c–e]).

The *simultaneous* detection of the ionization and the phonons at low temperatures has recently been demonstrated in Berkeley using a 60 g Ge detector (see [Cal91c, 94, Sad94]).

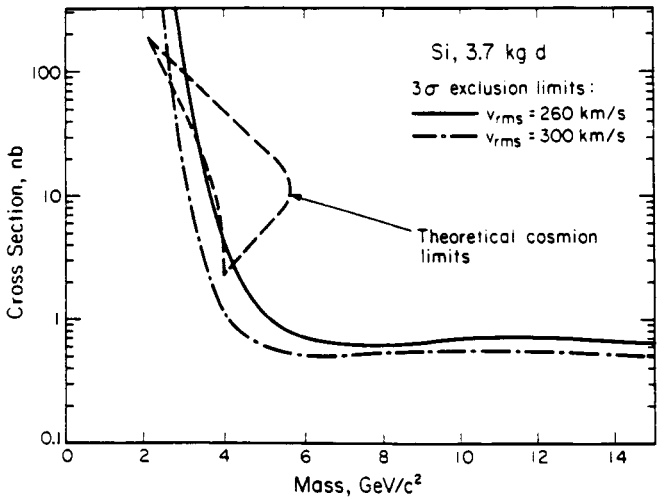


Figure 9.15. Exclusion diagram for dark matter particles (various velocities) in an Si experiment. The regions above the curves are excluded, among them cosmions whose theoretically allowed region lies between the broken curves (from [Cal91c, 92]).

Experiments with silicon semiconductors are also of interest. The advantage of silicon is its smaller mass, which results in a larger recoil energy and thus, in principle, a greater sensitivity. Figure 9.15 shows an exclusion diagram for cosmions measured using Si detectors [Cal90b, 91c, 92]. The region required (see chapter 7) for the solution of the solar neutrino problem (within the dashed curves) is practically completely excluded. Figure 9.16 shows the present experimental status of some leading dark matter experiments, perspectives for the future with ^{73}Ge cryodetectors, and theoretical expectations.

9.3.2.3 Accelerator experiments

We conclude with some remarks about recent accelerator results. Great efforts in the search for Higgs and SUSY particles have been made at LEP [ALE92]. The four LEP experiments have searched the whole of the kinematically accessible mass region. No events have been found which could be ascribed to a Higgs particle. As for the Higgs particle which was ruled out in the mass range $m_{\text{Higgs}} < 48$ GeV/c^2 [Ade91c], also the search for the SUSY particle was unsuccessful in practically the whole of the kinematically accessible mass region. Lower bounds of 40–45 GeV/c^2 on the mass of squarks and sleptons were determined.

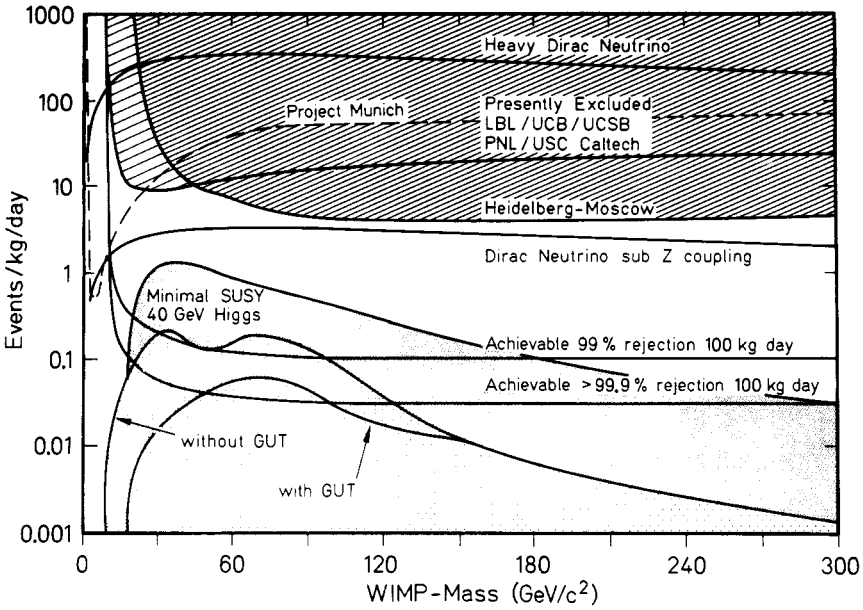


Figure 9.16. Current detection bounds for WIMPs with Ge detectors (shaded regions) and anticipated increase in the sensitivity with enriched ^{73}Ge cryodetectors (curves 99 and 99.9%). The figure also shows expected values for *spin-dependently* interacting WIMPs in various GUT models (after [Cal94, Sad94, Kla94]). Also shown is the sensitivity range of the Munich cryodetector project [Coo93, Sei93, Fer94].

9.3.3 Search for quark nuggets (nuclearites)

Part of the MACRO detector (see chapter 8) was used for the search for quark nuggets (called also nuclearites, see section 9.2.3.7) in the cosmic radiation [Ahl92]. The observed flux limits are 1.1×10^{-14} cm^{-2} sr^{-1} s^{-1} for 10^{-10} g $< m < 0.1$ g and 5.5×10^{-15} cm^{-2} sr^{-1} s^{-1} for $m > 0.1$ g.

We have only discussed a few results from the many experiments. Despite all the efforts no unambiguous evidence for the existence of a candidate for dark matter has yet been found.

Chapter 10

Particles with a Fractional Charge

10.1 THE QUARK CONFINEMENT

Experiments to search for particles with electrical charges q_e less than the elementary charge [Coh87]

$$e = 1.60217733(49) \times 10^{-19}\ \mathrm{C} \tag{10.1}$$

began with the postulation of the quark model by Gell-Mann and Zweig [Gel64, Zwe64]. Gell-Mann showed that the observed particle multiplets could be easily understood by thinking of the hadrons built-up from constituents with charges a multiple of a third ($+2/3e$ and $-1/3e$). He called these constituents quarks. Although the quarks were initially viewed as a mathematical way of expressing an underlying symmetry, the experimental search for them started soon after their postulation. Gell-Mann himself proposed such experiments to prove that quarks did not exist. As discussed in chapter 1, the physical reality of the quarks is no longer in question, even though no free quarks have yet been identified.

The fact that most experiments gave a negative result ultimately led to the concept of confinement, which we have already described in chapter 1. The current opinion is that, as a consequence of the special properties of the colour interaction, all physical systems such as mesons and baryons appear neutral in colour ('white') from the outside, i.e. they form a singlet for the colour interaction. Thus, quarks can only exist in bound systems and not as free particles (quark confinement). The smallest colour neutral quark systems consist of three quarks of different colours (qqq = baryons) or of a quark–antiquark pair ($q\overline{q}$ = mesons).

The search for free quarks remains an important test of the confinement hypothesis. In particular, the question arises as to whether the confinement of the quarks is exact and whether it is valid for arbitrary energies, or whether the confinement breaks down above a certain energy. In addition, some grand

unification theories require the existence of other particles with a fractional charge [Ing86, Wet86]. Certain higher-dimensional models introduce exotic quarks with electrical charges $q_e = +1/6e$ and $q_e = -5/6e$. Together with the standard quarks, these particles form baryons and mesons with electrical charges a multiple of a half. The lightest of these new hadrons should be stable, provided charge conservation applies. While four-dimensional GUT models such as SU(5), SO(10) and E_6 require that the charge of all colour singlet states should be an integer multiple of the elementary charge e, higher-dimensional models also allow colourless particles with charges a multiple of a half. By way of example, we mention superstring theories based on $E_8 \otimes E_{8'}$ or SO(32) symmetry groups. Since the masses of the new exotic quarks are associated with the $SU(2) \otimes U(1)$ symmetry breaking, values in the region of the mass of the W boson or below are expected [Ing86].

Because of the relatively large number of quarks and leptons which are now known, theories in which these 'elementary particles' are built from even more elementary components (subquarks or preons) have been developed (see e.g. [Ter80, Schr85b, Moh86a] and references therein). According to [Ter80] these subquarks have electrical charges

$$q_e/e = \pm 1/2,\ 0,\ +1/6. \tag{10.2}$$

There is also speculation about heavy leptons with a fractional charge [Bar83b]. These new leptons should be stable and thus could occur in the cosmic radiation. For a general discussion of the origin of fractional charges we refer to [Ber85d].

10.2 EXPERIMENTS TO SEARCH FOR FREE QUARKS

10.2.1 Millikan's experiment

Millikan was the first to prove the existence of the quantization of the electrical charge and the first to determine the exact size of the elementary charge, in 1910 [Mil10]. Millikan's experimental method was based on the direct measurement of the charge of very small oil droplets. Small droplets of liquid were introduced between the plates of a horizontally charged condenser using an atomizer. Some of these droplets were given a charge q either via the atomization process itself or by subsequent irradiation with ionizing X-rays. In the field-free space ($\boldsymbol{E} = 0$) a droplet is acted on by the force of gravity, corrected by the upthrust, acting downwards

$$F_G = \frac{4\pi}{3} r^3 (\rho - \rho_A) g \tag{10.3}$$

and the force of friction opposite to the direction of motion

$$F_F = 6\pi \eta r v_g \qquad \text{Stokes' law.} \tag{10.4}$$

Here, ρ denotes the density of the oil, and ρ_A is the density of the air. The radius of the droplet may be determined from measurements of the free-fall velocity v_g by virtue of the equation

$$\frac{4\pi}{3}r^3(\rho - \rho_A)g = 6\pi\eta r v_g \tag{10.5}$$

whence

$$r = 3\sqrt{\frac{\eta v_g}{2(\rho - \rho_A)g}}. \tag{10.6}$$

The movement may be observed using an optical microscope. A charged droplet with charge q may be held in suspension by applying an appropriate voltage U to the condenser: i.e. we require

$$F_G = qE = q\frac{U}{d}. \tag{10.7}$$

Thus, the charge may be easily determined using equation (10.6). In suspension, as a result of the Brownian motion, only very imprecise results may be expected. However, by reversing the polarity of the electrical field, the droplet may be made to fall and rise again in alternation and the velocities v_1 and v_2 in the two directions of motion (v_1 and v_2 are constant because of the friction) may be measured. Two simple equations of motion are obtained

$$\frac{4\pi}{3}r^3(\rho - \rho_A)g + qE - 6\pi\eta r v_1 = 0 \quad \text{downwards direction} \tag{10.8a}$$

$$\frac{4\pi}{3}r^3(\rho - \rho_A)g - qE + 6\pi\eta r v_2 = 0 \quad \text{upwards direction.} \tag{10.8b}$$

Finally, the charge of the droplet is obtained from precision measurements of the velocities v_1 and v_2

$$q = \frac{9\pi}{2E}\sqrt{\frac{\eta^3(v_1 - v_2)}{g(\rho - \rho_A)}}(v_1 + v_2). \tag{10.9}$$

Millikan's measurements yielded the important result that the electrical charge of each droplet is an integer multiple of the elementary charge e

$$q = ze, \quad z \in \mathbb{Z}. \tag{10.10}$$

We have described this fundamental experiment here because similar methods are still used today in the search for free quarks. In his experiment, Millikan also observed a candidate which did not satisfy condition (10.10) but had a 30% lower charge (see also [McC83]).

There are various techniques for searching for free quarks (see figure 10.1). Attempts may be made to generate them in interaction processes, or existing quarks in matter may be sought. We shall describe these two possibilities below.

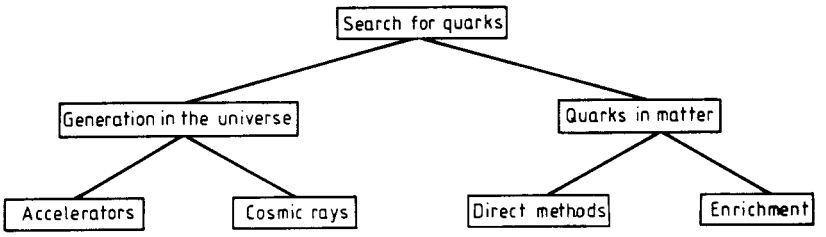


Figure 10.1. Various techniques for searching for free quarks (after [Lyo85]).

10.2.2 The search for free quarks at accelerators

Since the generation of particles with a fractional charge requires a certain centre-of-mass energy depending on the mass of the new particle, these studies involve either accelerator experiments or measurements of cosmic radiation.

The detection of particles with non-integral charges is based on the fact that these have a lower ionization capability than particles with an integral charge at the same velocity. Thus, information about the charge may be obtained from simultaneous measurement of the velocity and the ionization. The energy loss by a charged particle in matter is proportional to the square of the electrical charge

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{q_e^2}{\beta^2} f(\beta). \tag{10.11}$$

The function $f(\beta)$ depends only on the particle velocity. Thus, the passage of a quark through a scintillation counter could be identified from the unusually low ionization.

The search for quarks at accelerators began in 1964 at CERN in Geneva and at the Brookhaven National Laboratory (BNL) in New York, with typical proton energies of 30 GeV. This was followed by a variety of other experiments to search for free quarks and other particles with unusual charges in hadronic collisions, in deep inelastic lepton scattering and in e^+e^- storage rings. Despite the increasing centre-of-mass energies which have become available over the years, no evidence for the existence of free quarks has yet been obtained (a survey of accelerator experiments can be found in [Lyo85, PDG90, 94]).

10.2.3 The search in the cosmic radiation

In comparison with experiments at accelerators, cosmic radiation has the advantage that the studies may extend to higher energies. However, the results of these measurements may be interpreted in two ways. Either the observed candidates for particles with a fractional charge are primary components of the cosmic radiation or they are formed as secondary particles in reactions triggered

by the cosmic radiation. Since the spectrum of the cosmic radiation falls off sharply at high energies, a conversion of the flux (or of a flux limit) to a production cross section is associated with great uncertainties, since it involves assumptions about the energy dependence of the reaction cross section.

The cosmic radiation, which was discovered in 1912, was for a long time the most fruitful field for experiments on particle physics. Accelerator experiments only came to dominate later. The cosmic radiation is still associated with many enigmas. For example, the sources of these particles and the acceleration mechanisms are largely unknown. Figure 10.2 shows the spectrum of the cosmic radiation, which extends to energies of at least 100 EeV = 10^8 TeV. This spectrum follows a power law

$$-\frac{\mathrm{d}N}{\mathrm{d}E} \sim E^{-\gamma} \tag{10.12}$$

over a large energy region. At $E \simeq 1$ PeV = 10^3 TeV, the curve has a slight kink, where the index γ changes from $\gamma \simeq 2.7$ to $\gamma \simeq 3.1$. This kink and the shape of the spectrum at very high energies are currently the subject of intensive research.

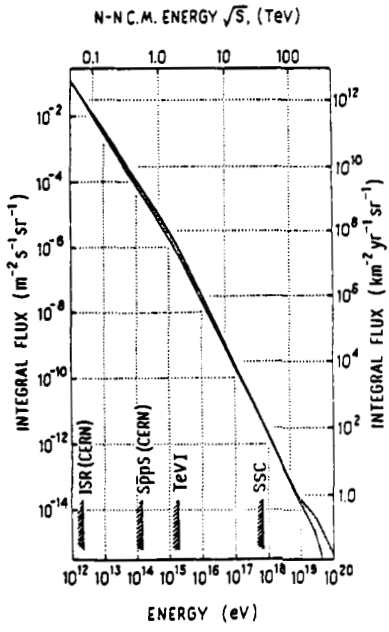


Figure 10.2. Energy spectrum of the primary cosmic radiation. For comparison, the equivalent centre-of-mass energy in the nucleon–nucleon system is also given (from [Ric87]).

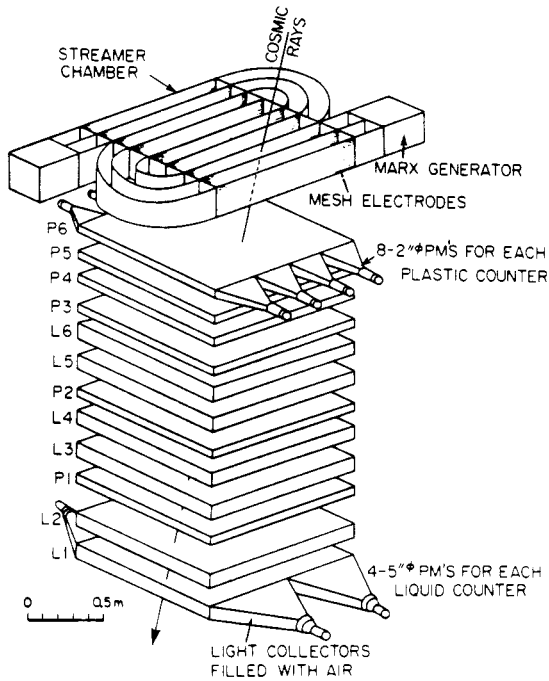


Figure 10.3. 'Quark telescope' for detecting quarks in cosmic radiation (from [Fuk89]) (see text).

The primary flux consists mainly of light nuclei. However, it contains practically all other stable nuclei, many long-lived radionuclides, electrons and antiprotons. In the lower energy region protons are overwhelmingly dominant. Typical detectors for detecting quarks in the cosmic radiation are designed as telescope detectors. Figure 10.3 shows the quark telescope of Fukushima *et al* [Fuk89]. A streamer chamber on the upper side is used to identify tracks, the 12 attached scintillation counters measure the ionization of the particles passing. These and other detectors have been used to carry out a number of experiments at various heights. Table 10.1 shows a selection of the results for the flux of particles with a fractional charge in the cosmic radiation. A complete list may be found in [PDG90]. Until recently, the upper bounds for the flux were in the range from 10^{-10} cm^{-2} s^{-1} sr^{-1} to 10^{-12} cm^{-2} s^{-1} sr^{-1}. Possible candidates have also been observed in some measurements.

An underground experiment, which is primarily intended to detect magnetic monopoles, provides very sensitive measurements of the flux of particles with a fractional charge [Mas83, Kaw84]. The detector was installed at a depth of

Table 10.1. The quark flux in the cosmic radiation.

F_{q_e} [cm^{-2} s^{-1} sr^{-1}]	q_e [e]	Events	Ref.
$< 5 \times 10^{-10}$	$+4/3$	0	[Bea72]
$< 2 \times 10^{-12}$	$\pm 1/3, \pm 2/3, \pm 1$	0	[Mas83]
$< 1 \times 10^{-12}$	$\pm 2/3, \pm 1/2$	0	[Kaw84]
$< 9 \times 10^{-10}$	$\pm 1/3, \pm 2/3$	0	[Wad84]
4×10^{-9}	$\pm 4/3$	7	[Wad84]
$< 2 \times 10^{-10}$	$\pm 1/3, \pm 2/3$	0	[Wad88]
	$\pm 4/3$	12	[Wad88]
$< 2.1 \times 10^{-15}$	$\pm 1/3$	0	[Mor91]
$< 2.3 \times 10^{-15}$	$\pm 2/3$	0	[Mor91]

250 m in the Kamioka mine, 300 km to the west of Tokyo. It consisted of a total of 60 plastic scintillators, arranged in six layers. The energy loss and the flight time of the particles passing through were measured. No unusual events were detected for velocities in the region $3 \times 10^{-4} < \beta < 1$. Analysis of the data gave the following results for the flux of magnetic monopoles (see chapter 8) [Mas83]

$$F_M < 1.5 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad 6 \times 10^{-4} < \beta < 1 \tag{10.13a}$$
$$F_M < 1.8 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad \beta = 5 \times 10^{-4} \tag{10.13b}$$
$$F_M < 2.5 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad \beta = 4 \times 10^{-4} \tag{10.13c}$$
$$F_M < 9.2 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad \beta = 3 \times 10^{-4}. \tag{10.13d}$$

The bounds on the flux of particles with different fractional charges with $\beta < 0.4$ are

$$F_{q_e} < 1.5 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad 3.5 \times 10^{-4} < \beta \quad |q_e| = 1e \tag{10.14a}$$
$$F_{q_e} < 1.8 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad 4 \times 10^{-4} < \beta \quad |q_e| = 2/3e \tag{10.14b}$$
$$F_{q_e} < 2.5 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad 4.5 \times 10^{-4} < \beta \quad |q_e| = 1/2e \tag{10.14c}$$
$$F_{q_e} < 9.2 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad 6 \times 10^{-4} < \beta \quad |q_e| = 1/3e. \tag{10.14d}$$

Bounds for the flux of leptons with $|q_e| = 2/3e$ and $|q_e| = 1/2e$ with relativistic energies ($\beta \simeq 1$) have also been obtained

$$F_{q_e} < 2.1 \times 10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad |q_e| = 2/3e \quad \text{(leptons)} \tag{10.15a}$$
$$F_{q_e} < 1.6 \times 10^{-14} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad |q_e| = 1/2e \quad \text{(leptons).} \tag{10.15b}$$

Currently, the sharpest bounds on the flux of particles with charge a multiple of one third in the cosmic radiation come from measurements with the Cerenkov

water counter of the Kamiokande II collaboration. This detector is described in more detail in chapter 4 on proton decay. Particles with a fractional charge may be distinguished from those with an integral charge by the intensity of the Cerenkov radiation emitted. The number of photons per path length in water and per unit wavelength emitted by a particle with charge q_e and velocity $\beta > 1/n$ is given by [Mor91]

$$\frac{\mathrm{d}^2N}{\mathrm{d}x\mathrm{d}\lambda} = 2\pi\left|\frac{q_e}{e}\right|^2\alpha\left(1-\frac{1}{(n\beta)^2}\right)\frac{1}{\lambda^2} \tag{10.16}$$

where n is the refractive index of water and α is the fine structure constant. Consequently, the number of Cerenkov quanta depends on the square of the charge. Thus, the intensity of the radiation emitted by quarks amounts to only 1/9 or 4/9 of the muon intensity. Evaluation after 1009 days of measurement gave [Mor91]

$$F_{q_e} < 2.1 \times 10^{-15}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1} \quad |q_e| = 1/3e \tag{10.17a}$$

$$F_{q_e} < 2.3 \times 10^{-15}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1} \quad |q_e| = 2/3e. \tag{10.17b}$$

10.2.4 The search for particles with a fractional charge in matter

If free quarks exist around us, they must be bound in matter, without forming a colour-neutral state. It is conjectured that a positively charged quark together with an electron will form a hydrogen-like state with charge $-2/3e$ or $-1/3e$. Negatively charged quarks, on the other hand, could be bound in the inner shell of an atom or even in the nucleus. The chemical properties of such atoms with a fractional charge are discussed in [Lac82, 83]. However, experiments are designed so that the results are largely independent of assumptions about the quark distribution and the chemical properties of the resulting exotic atoms.

A simple estimate of the concentration of particles with a fractional charge in matter is obtained under the assumption that all particles with non-integral charges present on the Earth originally stem from the cosmic radiation (see [Ric87]). We may then use the bounds on the flux obtained in the previous section to determine the maximum number of free quarks per nucleon. The free quarks may be distributed down to a depth of about 3 km in the Earth's crust (a geological mixing takes place down to this depth). Assuming a constant flux of $10^{-11}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1}$ over the age of the Earth, we expect a maximum of one free quark per 200 mg of material. This corresponds to a concentration of 10^{-23} free quarks per nucleon. However, the new measurements by the Kamiokande collaboration show [Mor91] that the actual flux is at least four orders of magnitude smaller, i.e. the concentration is correspondingly smaller.

There are essentially two experimental methods which have been used to study samples of material. These involve measurements with ion beams and

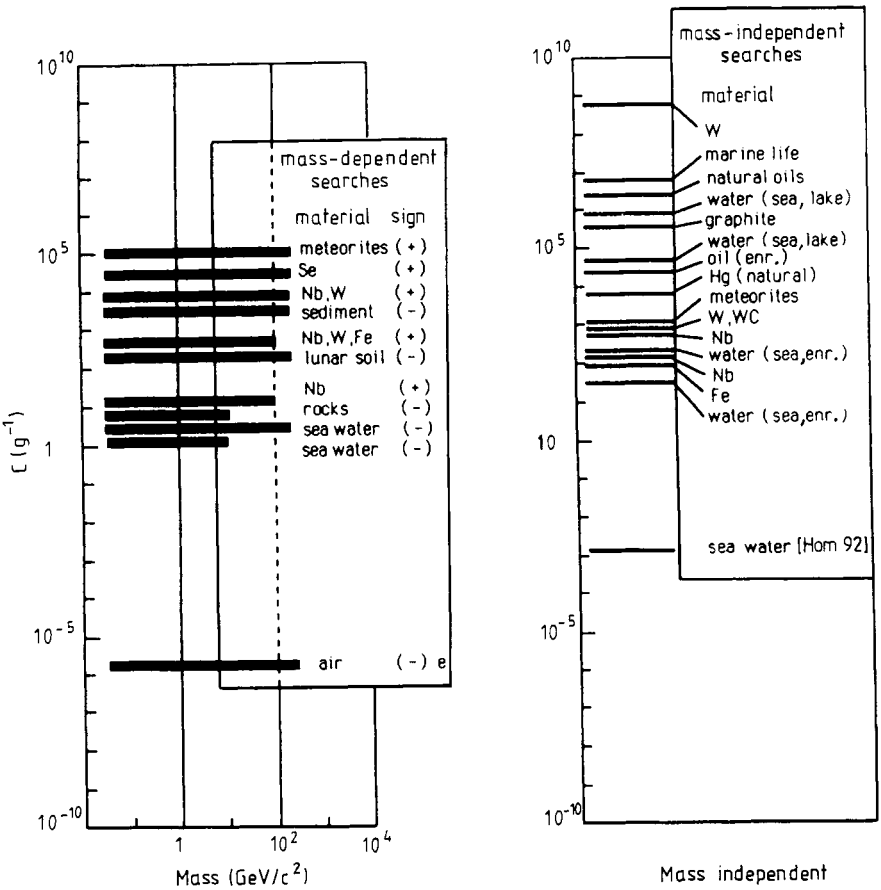


Figure 10.4. Bounds for the concentration of fractionally charged particles in different materials. The left part of the figure summarizes the results of experiments with ion beams. The black stripes show the region of mass sensitivity for the measurements, which only search for a certain sign of the charge. The right part summarizes the results of direct charge measurements. These results are independent of the particle number and the sign of the charge (from [Smi89], completed by [Hom92]).

suspension experiments similar to that due to Millikan. The latter have the advantage that they are independent of the particle mass. Figure 10.4 gives an overview of the bounds on the concentration from such experiments. A very good summary of the experimental search for fractional charges in matter may be found in [Smi89].

10.2.4.1 Ion-beam experiments

Figure 10.5(*a*) illustrates the principle of an ion-beam experiment. The sample of material to be studied is vaporized and ionized. Then the ions of charge q_e are accelerated to the kinetic energy $T = q_e V$ by the potential V. The energy, and thus also the electrical charge q_e may be determined using a silicon barrier detector or an electron multiplier. To decrease the background, the ions are often 'stripped', i.e. the electrons are removed from them. In addition, different T/q_e values can be selected by additional electrostatic fields [Mil85, 87].

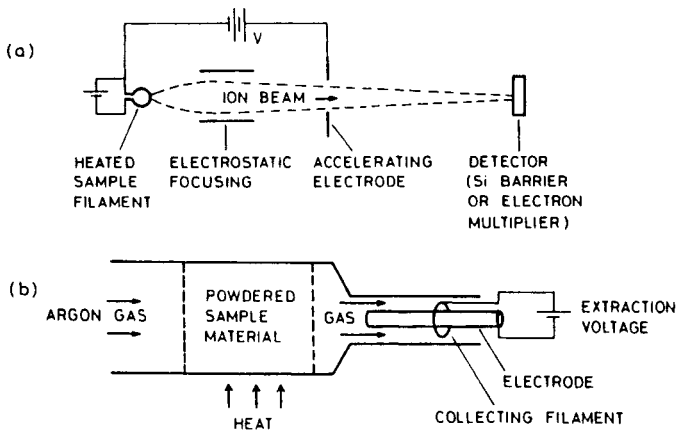


Figure 10.5. (*a*) Principle behind an ion beam experiment to search for fractionally charged particles; (*b*) procedure to extract fractionally charged particles from a sample and concentration on a small string (after [Smi89]).

Since the ion flux is limited, this method is only suitable for direct study of small amounts of material of around 10^{-4} g. However, higher sensitivities may be achieved by enriching the sample with free quarks or other particles. For this, air or water vapour is allowed to flow slowly through an electrical field so that all charged ions may be extracted and collected on a small thread which is used as source for the ion-beam experiment. This method is possible since the charge of a particle with a fractional charge cannot in general be completely neutralized. The procedure has also been extended to other materials (see figure 10.5(*b*)). The substance to be studied is heated in a stream of inert gas (e.g. argon). The carrier gas takes the charges released with it and passes into an electrical field for extracting these ions. The probability of extraction achieved in this way is, however, largely unknown.

Ion-beam experiments have been used to study various samples such as air, seawater, rock and metal (Fe, Se, Nb, W). Quantities up to $\sim 10^{-4}$ g have been studied using direct measurements and up to ~ 1 g using the extraction method.

In the case of air, the amount used has even reached 10^6 g. Because free charges are taken up by water droplets, only very few particles with non-integral charges are expected in air, so that the corresponding bounds for the concentrations should be not very significant (see [Smi89]). Table 10.2 summarizes a selection of the results obtained in ion-beam experiments.

Table 10.2. The quark concentration c in matter from ion-beam experiments.

c [quarks/nucleon]	q_e [$e/3$]	Events	Ref.
$< 1 \times 10^{-17}$	+1.2	0	[Chu66]
$< 5 \times 10^{-23}$		0	[Coo69]
$< 1 \times 10^{-21}$		0	[Ste76]
$< 1 \times 10^{-22}$		0	[Sch78]
$< 4 \times 10^{-28}$		0	[Ogo79]
$< 2 \times 10^{-20}$	$\pm > 1$	0	[Mil85]
$< 1 \times 10^{-19}$	$\pm 1, 2$	0	[Mil87]

One major disadvantage of the ion-beam method is the restriction to a mass region up to around 100 proton masses (see figure 10.4). For larger masses, the accelerating potential V is no longer sufficient to achieve the velocity of at least $\sim 10^7$ cm s^{-1} needed for a detection via ionization [Lew85]. In addition, the sensitivity of an experiment also depends on the type of ion and the charge. Furthermore, the upper bound of approximately 100 GeV/c^2 on the mass is a strong restriction, since the mass scales of current theories lie, in part, considerably above this. Thus, it appears to be necessary to carry out experiments, the interpretation of which is independent of the mass of a possible particle with a fractional charge. We shall discuss such a method in the next section.

10.2.4.2 Suspension experiments

As previously mentioned, suspension experiments provide for mass-independent detection of particles with a fractional charge. This direct determination of the electrical charge is based on two important principles. The sample is isolated from the surroundings to avoid fluctuations in the charge (a sample which is in contact with the surrounding material is subject to charge fluctuations of size $\Delta q_e^2 \simeq akT$ [Smi89], where a denotes the radius of the sample). The movement of the body in an electrical field $\boldsymbol{E}$ is then observed and from this the charge is determined.

One important implementation of these principles is the classical experiment

Table 10.3. Bounds for the concentration of fractionally charged particles in Millikan-type experiments (see also [PDG90]).

c [quarks/nucleon]	Material	Events	Ref.
$< 2 \times 10^{-20}$	175 μg Hg	0	[Hod81]
$< 9 \times 10^{-20}$	50 μg seawater	0	[Joy83]
$< 3 \times 10^{-21}$	2 mg Hg	0	[Sav86]

due to Millikan, which we have already discussed in detail. Here, the movement of a charged droplet subjected to gravitational and electrical fields in a viscous medium is measured. This technique has also been used in the search for particles with a fractional charge. While Millikan in his original experiment used masses of approximately 10^{-11} g, more recent automatic measurements involve masses in the region of 10^{-9} g and much shorter measurement times [Hod81, Joy83, Sav86]. Automation has made it possible to achieve repetition rates of 1 s^{-1} and thus to measure a total of $\sim 10^{-4}$ g per day [Hod81, Joy83]. This method may be used to study practically any liquid. Table 10.3 shows the bounds obtained for the free quark concentrations. The search in mercury exposed before to a beam of heavy ions has also been unsuccessful to date [Lin83].

Most mass-independent measurements are carried out in an inhomogeneous magnetic field rather than in an electrostatic field. For this, a ferromagnetic or diamagnetic sample body is brought into an equilibrium situation, generated by an inhomogeneous magnetic field and the Earth's gravitational field (see figure 10.6). The electrical charge may be determined from the motion induced by an alternating electrical field [Mar82].

Ferromagnetic implementations involve the use of small steel balls (or non-ferromagnetic samples with a steel jacket) with typical diameter 0.25 mm as the sample body in a vacuum chamber at room temperature. The almost perfect diamagnetism of a small superconducting sphere at very low temperatures ($\sim$ 4 K) may also be used to stabilize the sample in the magnetic field. In both cases, two disc conductors generate an oscillating electrical field which excites the spheres into a damped oscillation with amplitude proportional to the charge.

To search for particles with a fractional charge, one begins with a negatively charged sample body and removes successively electrons by irradiation with UV light or using a radioactive source. When only a few surplus charges are still present, these are removed individually until the sample has a few positive charge units. If no particles with a fractional charge are present, the amplitude of the oscillations should show a zero passage. However, the presence of a particle

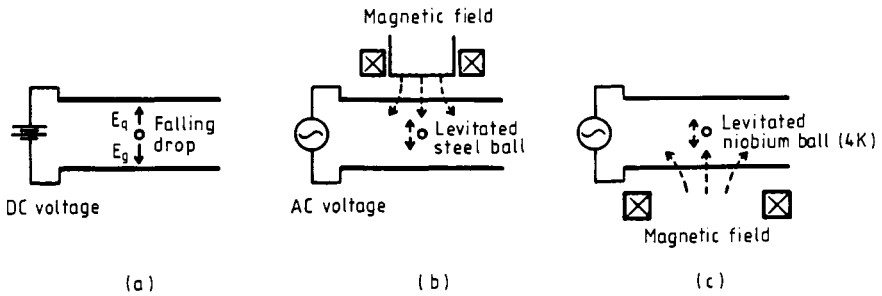


Figure 10.6. Suspension experiments to measure the electrical charge: (*a*) Millikan's falling-drop technique; (*b*) suspension experiment with a ferromagnetic sphere in a magnetic field; (*c*) suspension experiment with a superconducting sphere in a magnetic field (after [Smi89]).

with charge $1/3e$ or $2/3e$ results in a shift from the zero line corresponding to a charge of $\pm 1/3e$. Various sources of systematic error are discussed in [Mar82, 84b]. Apart from one exception, all ion-beam and suspension experiments are consistent with a vanishing rest charge in all samples.

A group from Stanford has observed evidence for the existence of free quarks [LaR77, 79, 81]. LaRue *et al* studied superconducting niobium spherules at the temperature of liquid helium. Figure 10.7 is a schematic illustration of the Stanford experiment. A niobium spherule is suspended in a magnetic field generated by the superconducting coil M. It is located between two horizontal plates of a condenser, by which an oscillating electrical field may be applied. The whole apparatus is cooled to a temperature $T = 4.2$ K. The position of the spherules is measured using an extremely sensitive SQUID (superconducting

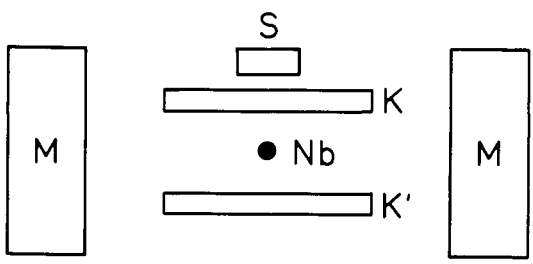


Figure 10.7. Schematic structure of the Stanford experiment to investigate superconducting niobium spheres at the temperature of liquid helium: K, K′ = flat condenser plates between which an alternating electrical field can be generated; M = superconducting magnetic coils; S = SQUID to determine the position of the Nb sphere.

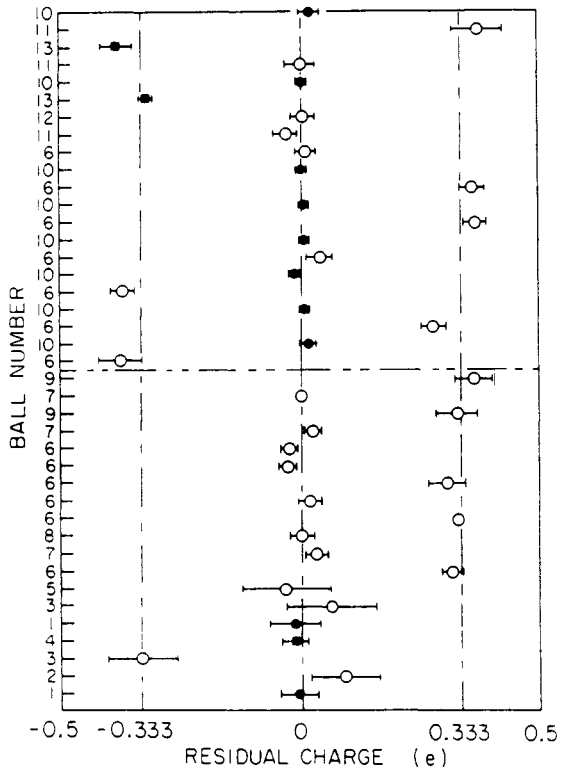


Figure 10.8. Results of the Stanford experiment of La Rue *et al* (from [LaR81]). Fourteen measurements showed residual charges of size $\pm(1/3)e$ (see text).

quantum interference device) magnetometer.

A total of 40 measurements with 13 niobium spherules, each of mass 0.1 mg, were taken. Fourteen measurements (5 spherules) indicated residual charges of size $\pm 1/3\,e$. The results are summarized in figure 10.8. The authors give the following values for the concentration c ((10.18*a,b*) [LaR77], (10.18*c*) [LaR79], (10.18*d*) [LaR81]):

$$c = 4 \times 10^{-21} \text{ quarks/nucleon} \quad q_e = +1/3\,e \qquad (10.18a)$$

$$c = 2 \times 10^{-21} \text{ quarks/nucleon} \quad q_e = -1/3\,e \qquad (10.18b)$$

$$c = 1 \times 10^{-20} \text{ quarks/nucleon} \quad q_e = +1/3\,e \qquad (10.18c)$$

$$c = 1 \times 10^{-20} \text{ quarks/nucleon} \quad q_e = \pm 1/3\,e. \qquad (10.18d)$$

However, these values contradict the results of other subsequent experiments (see table 10.4). Since we do not have a precise knowledge of

Table 10.4. Bounds for the concentration c of fractionally charged particles from suspension experiments (see also [PDG90]).

c [quarks/nucleon]	Charge [e]	Material	Events	Ref.
4×10^{-21}	$+1/3$	Nb	2	[LaR77]
2×10^{-21}	$-1/3$	Nb	1	[LaR77]
$< 3 \times 10^{-21}$		Fe	0	[Gal77]
$< 5 \times 10^{-15}$		W	0	[Bla77]
$< 6 \times 10^{-15}$	$> 1/6$	W	0	[Put78]
1×10^{-20}	$+1/3$	Nb	2	[LaR79]
$< 1 \times 10^{-21}$		Fe	0	[Mar80b]
1×10^{-20}	$+1/3$	Nb	4	[LaR81]
1×10^{-20}	$-1/3$	Nb	4	[LaR81]
$< 2 \times 10^{-21}$	$\pm > 1/2$	Fe	0	[Lie83]
$< 5 \times 10^{-22}$		Fe	0	[Mar84b]
$< 1 \times 10^{-21}$	$\pm 1/3$	Nb	0	[Smi85]
$< 3 \times 10^{-22}$	$\pm 1/3, \pm 2/3$	Nb	0	[Smi86]
$< 5 \times 10^{-22}$	$\pm 1/3, \pm 2/3$	W	0	[Smi87]
$< 4 \times 10^{-20}$	$\pm 1/3, \pm 2/3$	Meteorite	0	[Jon89]

the chemical behaviour of atoms with a fractional charge, differences between various materials (and, under certain circumstances, between samples of the same material, depending on the pre-processing) may arise as a result of possible enrichment processes. In these experiments, events which simulate the signals from possible charges must be thoroughly analysed and investigated. These include, amongst other things, interference from residual magnetic or electrical fields. These effects are discussed in detail in [Mar82, 84b], also in relation to the measurements at Stanford.

As a result of unresolved technical problems, no more reliable measurements have been carried out by the Stanford group since 1982. In the meantime, the authors themselves have conjectured that a small magnetic effect might provide a possible explanation for the rest charges which were apparently multiples of a third [Phi88].

The existence of particles with a fractional charge, and, in particular, of free quarks, has not yet been confirmed beyond all doubt. Thus, it remains unclear whether quark confinement corresponds to an exact symmetry or whether it is only valid at today accessible energies. For a detailed discussion of the experimental problems and a survey of future possibilities for the search for free quarks, we refer interested readers to [Smi89].

Chapter 11

Fifth Force: Theoretical Expectations and Experimental Status

11.1 INTRODUCTION

One of the underlying aims of physics is to understand the fundamental forces (interactions) in nature. At the present time, we know of four basic forces which rule the world, namely gravity, the weak interaction, the electromagnetic interaction and the strong interaction (colour force). Of these, Newtonian gravity has been known for the longest time. This force, with which two bodies attract one another depends only on their mass and obeys a $1/r^2$ law. Newtonian gravity is independent of the chemical composition of the interacting objects. In the last century, another fundamental force, the electromagnetic interaction, was added to this. The two remaining forces, which only operate at very small distances, were finally discovered during this century.

Naturally, the question immediately arises as to whether only these four basic forces exist or whether there are other interactions which we cannot yet see because they are too weak. Within the framework of the accuracy of modern measurement techniques, it is possible to define certain ranges of distances in which there is still space for a fifth force. In addition, there exist indications from theoretical models which point to further interactions.

11.2 THEORETICAL EXPECTATIONS

In the framework of attempts to unify the four interactions known today, i.e. to trace these back to a common cause, quantum theories of gravity have been proposed which contain further attractive and repulsive gravity-like forces which decrease exponentially as the distance increases. We shall discuss these ideas in some detail. However, we cannot discuss the mathematical structures of these

theories, which are in part very complicated, in this book, and restrict ourselves to essentially only describing the relevant results and predictions.

The objective of deriving all interactions from one underlying force has been largely achieved in the case of the weak and the electromagnetic interaction within the framework of the standard model. GUT models aim to unify the electroweak theory with the strong interaction (see chapter 1). Inclusion of gravity is currently still the greatest problem. However, there are also very promising approaches to this (supergravity, and, in particular, superstring models).

11.2.1 The equivalence principle

The difficulties in formulating a quantum field theory of gravity based on the tried and tested pattern of local gauge symmetries, as for example in the electroweak theory or quantum chromodynamics, stem from the incompatibility of the Heisenberg uncertainty principle and the so-called *weak equivalence principle* which requires the equivalence of inertial and gravitational mass. Before we discuss this contradiction to the general theory of relativity we shall first explain the terms gravitational and inertial mass.

The mass of a body may be determined by measuring the acceleration a for a known force F from the equation

$$m_t a = F. \tag{11.1}$$

This mass is called the inertial mass. The so-called gravitational mass m_s may be determined by determining the gravitational force F which another body of mass m_0 exercises on the sample body. According to Newton's law of gravity

$$F = G\frac{m_0 m_s}{r^2}. \tag{11.2}$$

The gravitational mass is given by

$$m_s = \frac{Fr^2}{Gm_0}. \tag{11.3}$$

It is a remarkable fact that the inertial and the gravitational mass of all bodies are proportional to one another, to within the accuracy of measurement. If the units are suitably chosen, the numerical values of m_s and m_t become equal.

The equivalence principle states that the effect due to an accelerated movement and that due to gravitational force are indistinguishable. From this it follows, in particular, that the inertial and the gravitational mass are equal.

One very interesting experiment concerns the measurement of the gravitational mass of the photon, which has a vanishing rest mass. The energy

of a photon with frequency ν is given by

$$E = h\nu. \tag{11.4}$$

Thus, by virtue of the equation $E = mc^2$, it must have an inertial mass m_t of size

$$m_t = \frac{h\nu}{c^2}. \tag{11.5}$$

Pound and Rebka [Pou60] attempted to detect the gravitational mass of the photon directly via the red shift due to gravity. They studied photons in the Earth's gravitational field. According to the equivalence principle, a photon with frequency ν, located at height H above the Earth's surface, has a potential energy

$$V = m_s Hg = \frac{h\nu}{c^2} Hg. \tag{11.6}$$

If the photon falls from the height H towards the centre of the Earth, its energy increases by this amount and the frequency is shifted. The new energy is

$$h\nu' \simeq h\nu + \frac{h\nu}{c^2} Hg. \tag{11.7}$$

Here, we have assumed that the mass $h\nu/c^2$ is essentially unchanged. It follows from (11.7) that

$$\nu' = \nu \left(1 + \frac{Hg}{c^2}\right). \tag{11.8}$$

Pound and Rebka used a distance of fall of some 20 m only. This gives a frequency shift of

$$\frac{\Delta\nu}{\nu} \approx 2 \times 10^{-15}. \tag{11.9}$$

This extremely subtle effect was detected using the Mössbauer effect on the 14.4 keV line of ^{57}Fe. The experimental arrangement provided for measurements of energy shifts $\Delta\nu/\nu$ down to 5×10^{-16} [Pou60]. The measured relative energy shift of the photons agreed with expectation from equal inertial and gravitational mass:

$$\Delta(h\nu)^{\text{exp}} = +(1.05 \pm 0.10)\, \Delta(h\nu)^{\text{meas}}. \tag{11.10}$$

However, one should bear in mind, that it is extremely difficult to exclude systematic errors in this experiment. In particular, even tiny temperature differences between source and absorber could produce effects of comparable size.

After this small excursion, we return to the problem of the fifth force. The general theory of relativity is based on the weak equivalence principle, one particular consequence of which is the $1/r^2$ dependence of the Newtonian

force of gravity. However, this postulate means that the world line of an object is exactly defined. This requirement is incompatible with the Heisenberg uncertainty principle. One possible way of removing this contradiction in the framework of a quantum theory of gravity involves the violation of the equivalence principle by further gravitational forces, which would modify the $1/r^2$ law. If there is a unified theory which includes gravity, then the general theory of relativity can no longer hold in its existing form as a non-renormalizable theory. For example, supersymmetric theories contain elementary particles, the existence of which could imply a violation of the equivalence principle, at least for certain ranges of distances between the interacting objects [Gol86].

11.2.2 The Yukawa potential in boson-exchange models

According to our present understanding, all elementary interactions are mediated by the exchange of a virtual field quantum. These field quanta are bosons; they have spin 1 for all interactions except gravity. The as yet hypothetical graviton which mediates the gravitational interaction would have to have spin 2 (see figure 11.1).

Massless exchange particles give a force law with a radial dependence of the form $1/r^2$, as in the example of electromagnetism (because of the self-coupling of the gluons, the strong interaction is an exception). Using the example of electromagnetism, we shall now consider a plausibility argument which shows how the exchange of a massless virtual photon between two charges q_1 and q_2 is related to the $1/r^2$ dependence of Coulomb's law (see figure 11.2). The photon exchanged is said to be virtual, since it can only exist for a limited time defined by the uncertainty relation

$$\Delta E \Delta t \geq \hbar. \tag{11.11}$$

The momentum of the photon and the spatial distance between the two charges

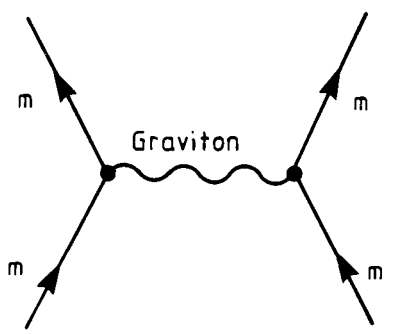


Figure 11.1. Elementary vertex for gravity.

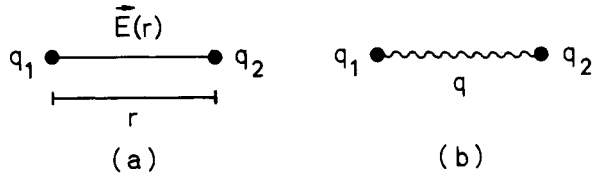


Figure 11.2. The electromagnetic interaction: (*a*) classical electrical field $\boldsymbol{E}(r)$; (*b*) exchange of a virtual boson with momentum q.

are related by the equation

$$qr = \hbar. \tag{11.12}$$

Every photon exchanged transmits a momentum $q = \hbar/r$. The propagation time for a massless particle is given by

$$t = \frac{r}{c}. \tag{11.13}$$

The effect of the force now follows easily from the equation

$$F = \frac{\mathrm{d}q}{\mathrm{d}t}. \tag{11.14a}$$

The absolute value of the force is given by

$$F = \frac{\hbar c}{r^2}. \tag{11.14b}$$

Let us now assume that the number of emitted and absorbed photons is proportional to the product of the couplings $q_1/\sqrt{\hbar c}$ and $q_2/\sqrt{\hbar c}$; then, we have the well-known formula for Coulomb's law

$$F = \text{constant} \times \frac{q_1 q_2}{r^2}. \tag{11.15}$$

From the theoretical point of view, gravitational forces are also caused by boson exchange. If their rest mass is zero, they move with the speed of light and generate Newton's law of gravity (11.2) with an infinite range. In other words, the interaction potential is of the form

$$V(r) \sim \frac{1}{r}. \tag{11.16}$$

However, if the exchange bosons have a finite rest mass m_B, then they generate a potential of the Yukawa form

$$V(r) \sim \frac{1}{r} \exp(-r/\lambda_B). \tag{11.17}$$

The range of this force is given by the Compton wavelength of the boson

$$\lambda_B = \frac{\hbar}{m_B c}. \tag{11.18}$$

We have already come across this equation relating the finite range to the mass of the exchange quanta in chapter 1 in the discussion of the weak interaction.

Since, here, we wish to discuss forces in the macroscopic region, we are interested in bosons with mass less than $\sim 10^{-4}$ eV/c^2, corresponding to a range in the millimetric region and above. The existence of such light particles has been discussed in a number of theoretical works (see e.g. [Gib81, Moo84, Cha85, Fay86, Bar86a, b, Gol86, Pec87] and section 11.2.4). In the simplest case, the Yukawa potential (11.17) may be added directly to the classical gravitational potential. Thus, the overall gravitational interaction potential between two masses m_1 and m_2 separated by a distance r may be written as follows

$$V(r) = -G_\infty \frac{m_1 m_2}{r}\left[1 + \alpha \exp(-r/\lambda)\right]. \tag{11.19}$$

Here G_∞ is the 'true' Newtonian constant of gravity which is measured for an infinite separation, α denotes the coupling constant of the new field and λ stands for the range of the fifth force. Such a modification of the law of gravity was already proposed in the 1970s by Fujii [Fuj72]. As a general result of quantum field theory, the exchange of bosons with an even (odd) spin results in an attractive (repulsive) force between two fermions of the same charge. For large ($r \gg \lambda$) and very small ($r \ll \lambda$) distances, (11.19) becomes Newton's law. However, for small distances we have now an effective constant of gravity

$$G_0 = G_\infty(1 + \alpha). \tag{11.20}$$

Such a hypothesis does not contradict astronomical observations, since only the product Gm is derived from these and the masses of the corresponding astronomical objects are only known very imprecisely [Spe88].

In general, scalar exchange particles mediate an attractive interaction. [Pec87] includes a theoretical study of possible effects of scalar pseudo-Goldstone bosons. These particles are a consequence of spontaneously broken global symmetries. In classical theories without a mass scale, the so-called dilation symmetry ('length rescaling') must be spontaneously broken. The resulting Goldstone boson is called a dilaton. This particle acquires a small mass, since the quantization requires the introduction of an effective mass scale. It is discussed [Pec87] that the mass m_D has order of magnitude $\Lambda^2/m_{\mathrm{Pl}}$, where Λ is a typical mass scale of the known interactions and m_{Pl} denotes the Planck mass of $\sim 10^{19}$ GeV/c^2. Thus, for $\Lambda \simeq 1$ GeV/c^2, we obtain a range of λ_D for the interaction mediated by dilaton exchange of from several hundred metres to

a few kilometres. For reasons of symmetry, the coupling of this force to leptons, quarks and gauge bosons is very weak, and we expect $0.01 < \alpha < 1/3$.

To conclude, we note that the Yukawa potential in (11.19) is not the most general potential possible, which is generated by the exchange of massive bosons [Moo84]. Additional spin-dependent terms may arise. We shall not discuss these modifications here.

11.2.3 Baryon-number-dependent fifth force

In addition to obeying a law other than $1/r^2$ radial dependence, a new force could manifest itself via a dependence on the composition of the interacting bodies. The 'charge' of an atom, as far as the hypothetical new force is concerned, or the coupling constant of the new force, need not necessarily be proportional to the mass of the atom. For example, a new interaction has been proposed which is mediated by the exchange of vector bosons which couple to the baryon number B [Fis86], i.e. for this force B plays the role of a 'charge'. If the baryon number could be associated with an elementary interaction in this way, this would provide a theoretical explanation for the conservation of this quantity which has been confirmed, without exception, in experiments to date (see also chapter 4 on the decay of the proton).

Suppose that Q is the 'charge' to which the fifth force is coupled. We write the interaction potential generated in the form

$$V_Q(r) = \frac{G_\infty m_1 m_2}{r} \alpha_Q \left(\frac{Q}{\mu}\right)_1 \left(\frac{Q}{\mu}\right)_2 \exp(-r/\lambda_Q) \tag{11.21}$$

where α_Q denotes the strength of the new interaction. The numbers μ give the masses in units of the mass of atomic hydrogen, with $m_H = 1.00782519(8)$ amu (atomic mass units). The total potential energy is given by (11.19), where

$$\lambda = \lambda_Q \tag{11.22a}$$

$$\alpha = -\alpha_Q \left(\frac{Q}{\mu}\right)_1 \left(\frac{Q}{\mu}\right)_2. \tag{11.22b}$$

There are a number of different approaches as far as the charge Q is concerned. As previously mentioned, Q could be given by the baryon number

$$Q = B. \tag{11.23}$$

For normal atoms or nuclei this corresponds to the sum of the nucleons.

Two bodies with the same mass ($m_1 = m_2 = m$) but different baryon numbers ($B_1 \neq B_2$) could then experience a different effect of the new force. The potential difference between two such objects in the field of a mass M is

$$\Delta V_B = \Delta \left(\frac{B}{\mu}\right) \alpha_B \frac{G_\infty m M}{r} \exp(-r/\lambda_B). \tag{11.24}$$

A more general approach [Fis86] assumes that the coupling of the fifth force is proportional to the hypercharge

$$Q = Y = B + S \tag{11.25}$$

i.e. it is the sum of the baryon number B and the strangeness S. Normal matter does not have a strangeness, so that $Y = B$ and equation (11.25) again becomes (11.23). However, small effects are expected in the framework of elementary particle physics, e.g. for K mesons. There are also proposals in which Q is written as a combination of the baryon number and the lepton number

$$Q = B\cos\theta + L\sin\theta \tag{11.26}$$

where the angle θ describes the mixing [Fay89].

A force which depends on the composition of the interacting bodies may depend not only on the baryon number but also on the difference between the neutron number and the proton number $N - Z$. For neutral atoms, by the way, the following relation holds

$$N - Z = B - 2L. \tag{11.27}$$

A dependence on $(N - Z)$ could imply a coupling to the third component of the isospin I_3. The following charge is frequently discussed [Fis88]

$$Q = B\sin\theta_5 + (N - Z)\cos\theta_5 \tag{11.28}$$

where the mixing angle θ_5 defines the relative strength of the couplings to B and $(N - Z)$. In the next section we shall discuss how new forces occur naturally in models of grand unification.

11.2.4 Quantum theories of gravity

The formulation of a renormalizable quantum theory of gravity (see supergravity, section 12.2) appears necessarily to imply additional new gravitational forces [Sch77, 79, Mac84, Gol86, Nie91]. In supersymmetric theories the standard graviton with spin 2 as a mediator of the known (tensorial) gravity should have two massive partners with spin 0 (graviscalar) and spin 1 (gravivector or graviphoton). These new particles could be the exchange quanta of new forces. If the masses are sufficiently small, the range in (11.18) could be large enough to give rise to macroscopic effects.

All field theories have in common the fact that the exchange of bosons with even spin ($S = 0, 2, \ldots$) generates attractive forces between unpolarized particles of the same type, while the exchange of bosons with odd spin ($S = 1, 3, \ldots$) results in a repulsive interaction. Thus, the graviscalar (S) interaction and the

gravivectorial (V) interaction would contribute an attractive and a repulsive component to the gravitational interaction, respectively. It also appears possible that the two components could depend on the composition of the interacting objects [Sch79, Gol86, Nie91]. Phenomenologically, we may write the potential energy between two point masses as

$$V(r) = -\frac{G_\infty m_1 m_2}{r}\left[1 - \alpha_V \exp(-r/\lambda_V) + \alpha_S \exp(-r/\lambda_S)\right]. \qquad (11.29)$$

In ordinary matter the two components could essentially cancel one another out, so that they would be impossible to detect. However, this situation changes dramatically when we consider the interaction between matter and antimatter (on the question of the interaction between matter and antimatter, see e.g. [Ade91a] and [Nie91]). The particle–antiparticle forces are attractive for both the graviscalar and the gravivector exchange (i.e. we have to invert the sign of α_V in (11.29)).

Thus, the gravivectorial interaction distinguishes between equal (matter–matter) and opposite (matter–antimatter) charges. We also find a corresponding phenomenon in the electromagnetic interaction, which is mediated by the exchange of a photon with spin 1. The exchange of a scalar, however, always results in an attraction. This is true in general for the case of bosons with even spin, whence also for the graviton. Thus, the normal, tensorial gravity should not distinguish between matter and antimatter.

Even if the last two terms in (11.29) essentially cancel one another out, their existence could be detected by comparing the interaction of matter with that of antimatter. An antiparticle in the Earth's gravitational field would fall with an acceleration a which is somewhat larger than g.

We shall now include a remark about the CPT theorem, in order to prevent possible misunderstandings (see [Nie91]). The weak equivalence principle requires the equivalence of inertial and gravitational mass

$$m_t = m_s. \qquad (11.30)$$

The CPT theorem [Lüd57] says that the inertial mass of a particle is equal to the inertial mass of its antiparticle

$$m_t = \overline{m}_t. \qquad (11.31)$$

However, this on no account implies that

$$\overline{m}_s = \overline{m}_t. \qquad (11.32)$$

We may only deduce that $m_s = m_t = \overline{m}_t$. Consequently, $m_s \neq \overline{m}_s$ does not necessarily imply a violation of the CPT theorem. The latter says only

that Newton's famous apple falls to the Earth in exactly the same way as an antiapple would fall to the Antiearth. It does not say anything about the fall of the antiapple to the Earth [Nie88].

However, the CPT theorem forms the basis for the derivation of the theorem that the exchange of bosons with even spin always leads to an attractive interaction. Thus, were the normal, tensorial gravity to admit a distinction between matter and antimatter, then this theorem would be violated [Nie88]. This remark about the relationship between the CPT theorem and gravity completes our discussion of the theoretical motivation for the search for a fifth force; we shall now turn to the experimental results.

11.3 THE EXPERIMENTAL SEARCH FOR A FIFTH FORCE

11.3.1 The geophysical window

The first question concerns the distances for which deviations from Newtonian gravity can be measured within the framework of the accuracy of current measurement techniques. A consideration of measurements of the $1/r^2$ law of gravity shows that there is a range of distances from a few centimetres to a few metres which is not covered by existing experiments ('geophysical window') [Mik77, Gib81].

There is no doubt about the $1/r^2$ dependence of gravity on astrophysical length scales. A new force with a very large or even infinite range may be essentially ruled out, since Newton's and Einstein's theories of gravity describe the dynamics of galaxies very well. Very accurate measurements of the orbits of planets and satellites in our planetary system give very sharp restrictions on an additional finite force for distances ranging from 10^5 to 10^9 km. Extremely precise tests of the weak equivalence principle were carried out by [Rol64] and [Bra72]. They observed the fall of test bodies towards the Sun and found no evidence for an interaction which would violate the equivalence principle. The sensitivity of these measurements was between 3×10^{-11} and 9×10^{-11}. Laser measurements of the distance between the Earth and the Moon have confirmed the equivalence principle with an accuracy of 5×10^{-12} [Nor82, All83]. However, all these experiments concern very large distances and thus do not provide information about possible violations at smaller distances.

On the other hand, subatomic and atomic systems also provide no evidence for a further force in addition to the four known interactions. This is true up to the length scales accessible in so-called Cavendish experiments. In a Cavendish experiment very sensitive torsion balances are used to determine the attractive force of sample bodies of various compositions at a distance of a few centimetres. The chemist Henry Cavendish used this method in 1798 to measure the constant of gravity for the first time [Cav98].

According to these considerations, the new fifth force would have a medium range and would play practically no role as far as problems in atomic or astrophysics are concerned. However, it would be important in relation to geophysical problems.

11.3.2 Verification of the $1/r^2$ law

Motivated by the lack of measurements in the middle-distance region, Long tested the $1/r^2$ dependence on the scale of 4–30 cm [Lon74, 76]. He determined the force of attraction of a ball on rings of various sizes and actually found deviations from Newton's law of gravity. The following parametrization of the constant of gravity was obtained

$$G(r) = G\left[1 + (2.0 \pm 0.4) \times 10^{-3} \ln\left(\frac{r}{1\ \text{cm}}\right)\right]. \tag{11.33}$$

However this result was not confirmed by later measurements [Spe80, Che84]. The deviations in Long's results, which gave rise to a large number of experiments, have now been traced back to systematic errors (see [Spe88]).

In essence, experiments which study the dependence of gravity on distance measure the change in the force of gravity with increasing depth in rock or in water or with increasing height from towers. It is also possible to study the variation of the force of gravity with time at a fixed point near moving masses of water (e.g. on reservoirs with a variable water level). Highly sensitive spring balances (gravimeters) are predominantly used to detect the force of gravity. Changes in the gravity lead to a displacement of the test mass from its equilibrium position. This displacement may be compensated using electrostatic feedback. Such relative measurements of the gravity achieve accuracies of below 10^{-9} g (here g denotes the acceleration due to gravity).

11.3.2.1 Airy experiments

An Australian working group attempted to detect a fifth force with a medium range from measurements of the force of gravity taken in a colliery shaft [Sta81, 87, Hol86]. They used a procedure first used by Airy in the last century to determine the average density of the Earth [Air56]. This method involves measuring the change in the acceleration of gravity with increasing depth in a mine. The difference in gravity between the Earth's surface and points at various depths in the shaft depends on the intervening layers of rock. Thus, the experiment involves measuring the vertical gradients of the acceleration due to gravity in a mine and searching for deviations from the expected dependence of the acceleration due to gravity on the depth z.

We shall use the example of a spherical Earth to describe the principle of an Airy measurement; for this, we shall for the moment assume a purely

Newtonian law of gravity. At an arbitrary distance from the centre of the Earth the acceleration due to the Earth's attraction amounts to

$$g(r) = \frac{Gm(r)}{r^2} \tag{11.34}$$

where $m(r)$ denotes the mass lying within a radius r

$$m(r) = 4\pi \int_0^r \rho(r')r'^2 \mathrm{d}r'. \tag{11.35}$$

Here, $\rho(r')$ denotes the density. The gradient of the acceleration due to gravity follows by differentiation of (11.34)

$$\begin{aligned} \frac{\mathrm{d}g(r)}{\mathrm{d}r} &= -\frac{2g(r)}{r} + \frac{4\pi G}{r^2}\frac{\partial}{\partial r}\int_0^r \rho(r')r'^2 \mathrm{d}r' \\ &= -\frac{2g(r)}{r} + 4\pi G\rho(r). \end{aligned} \tag{11.36}$$

Using (11.34), this may be rewritten as

$$\begin{aligned} \frac{\mathrm{d}g(r)}{\mathrm{d}r} &= -\frac{2Gm(r)}{r^3} + 4\pi G\rho(r) \\ &= 4\pi G\left(\rho(r) - \frac{2}{3}\overline{\rho}(r)\right) \end{aligned} \tag{11.37}$$

where $\overline{\rho}(r)$ denotes the average density within the radius r

$$\overline{\rho}(r) = \frac{m(r)}{(4\pi/3)r^3}. \tag{11.38}$$

The first term in (11.36) is called the 'free-air' gradient, the second term corresponds to double Bouguer correction [Sta87]. Airy experiments give the constant of gravity directly since the mass of the attracting body is explicitly taken into account in the second term. This is an important difference from the tower experiments which we shall discuss later.

We derived the expression (11.37) under the assumption of a spherical body with a $1/r^2$ force law. Consideration of the elliptical deformation of the Earth and the rotation leads to a similar expression with correction factors and an additional term [Sta81, 87, Dah82]. Consideration of contributions from a Yukawa term of the form (11.19) gives a deviation of the acceleration due to gravity from the expression according to Newton's law, which depends on the depth of the shaft z

$$\Delta g(z) = \frac{4\pi G_0\rho\alpha}{1+\alpha}\left(z - \frac{\lambda}{2}(1 - \mathrm{e}^{-z/\lambda})\right) \tag{11.39}$$

where G_0 is the constant of gravity determined in laboratory experiments. At a depth which is very much greater than the range of the force sought, an additional gradient occurs

$$\left(\frac{\mathrm{d}\Delta g}{\mathrm{d}z}\right)_5 = \frac{4\pi G_0 \rho \alpha}{1+\alpha} \quad (z \gg \lambda). \tag{11.40}$$

In the other limiting case ($z \ll \lambda$), since

$$\frac{\lambda}{2}\left(1 - \mathrm{e}^{-z/\lambda}\right) \approx \frac{z}{2} \tag{11.41}$$

the gradient is exactly half as large

$$\left(\frac{\mathrm{d}\Delta g}{\mathrm{d}z}\right)_5 = \frac{2\pi G_0 \rho \alpha}{1+\alpha} \quad (z \ll \lambda). \tag{11.42}$$

Between the two limiting values there is only a flat transitional region so that it is very difficult to determine a range λ even though an abnormal gradient may be comparatively easy to detect.

Stacey *et al* carried out measurements over a number of years at depths of up to 1 km in mines in Queensland, Australia [Sta71, 87, Hol86]. They were able to rule out the force discussed by Fujii [Fuj71] with $\lambda \sim 200$ m and $\alpha = 1/3$. However, they found an effect which was interpreted as possible evidence for a fifth force. It is very difficult to determine α and λ from mine experiments alone. The constant of gravity from measurements in the Hilton mine is

$$G_\infty = (6.720 \pm 0.002 \pm 0.024) \times 10^{-11}\ \mathrm{m^3\ kg^{-1}\ s^{-2}}. \tag{11.43}$$

This differs from the laboratory value [Coh87] obtained from measurement with torsion pendulums for distances in the centimetre range by less than one per cent (see also [Lut82])

$$G_0 = (6.67259 \pm 0.00085) \times 10^{-11}\ \mathrm{m^3\ kg^{-1}\ s^{-2}}. \tag{11.44}$$

If we take this small discrepancy between (11.43) and (11.44) seriously then the new force should be repulsive ($\alpha < 0$). Holding *et al* give $\alpha = -8 \times 10^{-3}$ and $\lambda \simeq 200$ m [Hol86]. However, it is highly likely that this result may be attributed to systematic errors [Bar89b, Ade91b, Nie91] (see also the discussion in section 11.3.2.3). Imprecise knowledge of the density in the region of the mine shafts is a problem for Airy experiments in mines. It is also extremely difficult to take account of density anomalies.

Measurements in boreholes in Nevada also gave a deviation from the $1/r$ potential, which, however, is much greater than that in the experiment described above [Tho90]. The authors conclude that this discrepancy between the two

measurements is not due to a fifth force but to the fact that the experiments involved large systematic uncertainties resulting from density anomalies.

Airy experiments in sea-water or ice have the advantage over mine measurements that the density distributions are very much easier to determine. The measurement point may be chosen so that the sea-bed is as devoid of structure as possible and covered by a thick, homogeneous layer of sediment. Thus, density anomalies may essentially be ruled out [Sta78, 87]. A first experiment of this type was carried out in summer 1987 in a 2 km deep borehole (Dye-3) in the Greenland ice [Cha87, And89]. There were differences between theory and experiment, but these were very probably a result of insufficient knowledge of the density distribution in the Earth's crust underneath the ice, since rocks with a much greater density may occur there. Thus, the interpretation is not clear. Other experiments of this type in the Antarctic and the Pacific are being planned, and in some cases data are already being collected (see [Mül91]).

11.3.2.2 Measurements on reservoirs

Experiments with gravimeters on fixed sites on reservoirs with a time-varying water level are less expensive than those described above. One major advantage is that the results cannot be falsified by unknown density anomalies in the Earth's crust. Since the measurement devices do not move, the results are independent of such effects. Pioneering work in this area was carried out by a research group on the Splityard Creek reservoir in Queensland, Australia, the water level of which was subject to daily variations of up to 10 m. Such reservoirs of peak-load power stations are particularly suitable. Since the water is circulated daily and thus is well mixed, the water density can be determined very accurately from a small number of sample recordings. Measurements of the acceleration due to gravity from a tower in the middle of the reservoir as the water surface sank or rose gave [Moo88]

$$G = (6.689 \pm 0.057) \times 10^{-11}\ \mathrm{m^3\ kg^{-1}\ s^{-2}} \qquad (11.45)$$

for an effective range of 22 m. This value is consistent with the expected value (11.44) within the bounds of error.

A further experiment of this type was carried out in August 1988 at the Hornberg reservoir in the southern Black Forest in Germany. No evidence for a possible fifth force was obtained in recordings of the force of gravity taken over a period of three weeks. The data could be explained by Newton's law of gravity taking into account the tidal forces. The deviation from the laboratory value was $(0.25 \pm 0.40)\%$ ($\lambda \sim 40$–70 m) [Mül89].

The most recent and most sensitive experiment of this type was carried out at the Gigerwald reservoir in Switzerland. It also gave no indication of a fifth force [Cor94].

11.3.2.3 Tower experiments

The law of gravity can also be tested by experiments on towers. These experiments measure the acceleration due to gravity along a vertical to the Earth's surface. These data are then compared with the expected values calculated assuming a purely Newtonian gravitational force. However, the determination of the theoretically expected decrease in the force of gravity with increasing height requires very precise measurements of gravity on the Earth's surface in the area around the tower up to distances of several hundred kilometres.

In comparison with measured values on the Earth's surface, new precise measurements of g in various heights at the WTVD television tower in Garner (North Carolina) showed significant deviations from the predictions of the Newtonian $1/r$ potential, amounting to

$$\Delta g = (-500 \pm 35) \times 10^{-8} \text{ m s}^{-2} \tag{11.46}$$

at the top of the tower, which appeared to point to an additional attractive force [Eck88]. The measurements included 257 measurements on the surface within a radius of 5 km and 1784 measurements on the surface within a radius of 220 km around the tower. The assumption of a simple Yukawa potential (11.19) led to an attractive force with the following parameters for the interaction strength α and the range λ

$$\alpha = +0.0204 \quad \lambda = 311 \text{ m}. \tag{11.47}$$

While the interpretation of mine experiments requires a precise knowledge of the surrounding mass density and assumptions about the unknown mass distribution deep under the Earth's surface, tower measurements provide a direct test of the $1/r$ form of the gravitational potential. The latter are essentially independent of the mass distribution, which determines g on the Earth's surface.

Unlike (11.47), the Airy experiment carried out by Stacey *et al* provided evidence for a repulsive perturbation, so that a simple Yukawa approach is not sufficient to explain both experiments. The 'scalar–vector model' discussed in section 11.2.4, with two components which essentially cancel each other out, permits a consistent description, provided both measurements reflect the physical reality. There are almost infinitely many two-parameter solutions. One typical fit would be [Eck88, Sta88]

$$\alpha_S = 1.000 \qquad \alpha_V = 1.007 \tag{11.48a}$$

$$\lambda_S = 103.0 \text{ m} \qquad \lambda_V = 97.0 \text{ m}. \tag{11.48b}$$

However, Bartlett and Tew [Bar89b, 90b] pointed out systematic errors in both experiments. A new analysis of the data from [Eck88] taking into account these effects showed that the data do not actually contain evidence for a deviation

from a $1/r$ potential [Jek90]. For a simple additional Yukawa potential, the bound on the coupling constant α is now

$$|\alpha| < 0.001 \text{ for } \lambda > 100 \text{ m}. \tag{11.49}$$

The same also appears to apply to the mine experiment [Ade91b]. Two other tower experiments on the 465 m BREN tower in Nevada [Tho89b] and on the 300 m NOAA meteorological tower in Erie (Colorado) [Spe90] were also consistent with the assumption of a pure $1/r$ potential. Thus, the decreases in gravity calculated using Newtonian gravity agree with the measured values of the decrease in gravity to within measurement errors for all three experiments on towers of height from 300 to 600 m (see table 11.1).

Table 11.1. Decrease in gravity in tower experiments. Δg_s denotes the difference between the measured and the expected acceleration due to gravity at the top of the tower.

Tower	Height [m]	Δg_s [10^{-8} m s^{-2}]	Ref.
WTVD (NC)	600	-500 ± 35	[Eck88]
WTVD (NC)[a]	600	-4 ● 43	[Jek90]
BREN (Nevada)	465	-60 ± 95	[Tho89b]
NOAA (Colorado)	300	10 ± 27	[Spe90]

[a] Re-analysis of the data of [Eck88].

In summary, we conclude that no indisputable deviation from a $1/r$ gravitational potential has yet been detected.

11.3.3 Substance dependence of gravity

Further evidence for a fifth force could come from a substance dependence of gravity, since Newtonian gravity is only coupled to the mass. The existence of such a substance-dependent force would also violate the equivalence principle which implies equal gravitational and inertial mass. As we have already seen, possible charges include the baryon number B, the hypercharge Y, the lepton number L, the isospin I_3 and a combination of these.

11.3.3.1 The Eötvös experiment and its reanalysis

In a famous experiment the Hungarian physicist Baron Loránt von Eötvös[1] studied the equivalence of inertial and gravitational mass [Eöt91]. This

[1] 'I can never forget the moment when my train rushed into the railroad station of Heidelberg along the banks of the Neckar ...', Baron L von Eötvös, 1887.

torsion-balance experiment was begun in 1890 and continued over many years. Originally, these measurements to determine irregularities of the gravitational gradient were thought of as a method of geophysical investigation. Eötvös *et al* compared the relative acceleration of pairs of material samples towards the Earth. After the formulation of Einstein's theory of general relativity, this work took on a new importance. The results were published after Eötvös's death [Eöt22] and were assessed as confirmation of the weak equivalence principle.

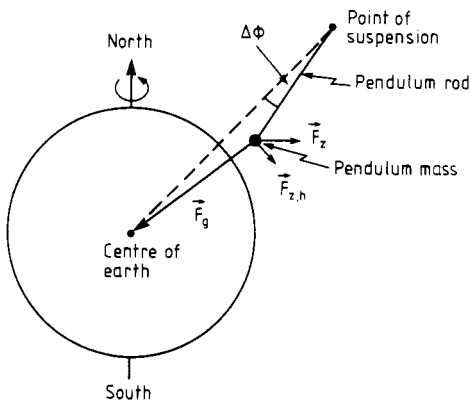


Figure 11.3. Deflection of a pendulum from the vertical by a small angle $\Delta\phi$ due to the centrifugal force, F_z, resulting from the rotation of the Earth. The horizontal component of the centrifugal force is $F_{z,h}$ and the gravitational force is F_g.

The Eötvös experiment is a classical method of determining the acceleration due to gravity and differences in the direction of the acceleration due to gravity for different sample bodies. The weak equivalence principle is tested by comparing the ratio of the gravitational to the inertial mass for bodies with different chemical compositions. To understand the principle of the torsion balance used by Eötvös, we consider a pendulum on the Earth's surface at a given (northern) geographical latitude θ (see figure 11.3). The pendulum is acted on by the force of gravity $F_g = m_s g$ acting towards the centre of the Earth and the centrifugal force $F_z = m_t \omega_E^2 R_E \sin\theta$ acting perpendicularly to the Earth's axis of rotation. Here, we are interested in the horizontal component of the centrifugal force $F_{z,h}$ which is given by

$$F_{z,h} = m_t \omega_E^2 R_E \sin\theta \cos\theta. \tag{11.50}$$

The centrifugal force resulting from the Earth's rotation gives rise to a deflection

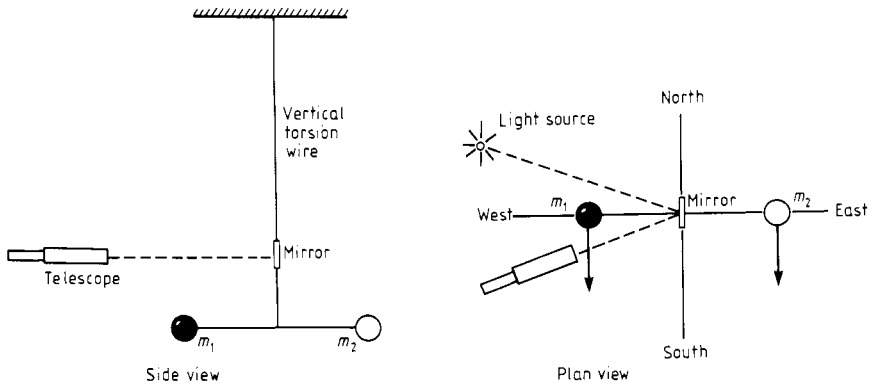


Figure 11.4. Torsion balance to determine the ratio of inertial to gravitational mass. m_1 and m_2 are different objects with the same gravitational mass. If the inertial masses m_1 and m_2 are equal then the horizontal components of the centrifugal force are also equal so that the whole stress acting on the string disappears (from [Kit79]).

of the pendulum from the vertical by a small angle $\Delta\phi$

$$\Delta\phi \simeq \frac{m_t \omega_E^2 R_E \sin\theta\cos\theta}{m_s g} = \frac{\omega_E^2 R_E \sin\theta\cos\theta}{g}\,\frac{m_t}{m_s}. \tag{11.51}$$

Figure 11.4 shows a torsion balance similar to that used by Eötvös to determine the ratio of inertial to gravitational mass. The balance with the two masses m_1 and m_2 hangs on a torsion wire. The two balls consist of different material with the same gravitational mass

$$m_s(1) = m_s(2). \tag{11.52}$$

If the inertial masses $m_t(1)$ and $m_t(2)$ are not equal the suspension is twisted as a result of the unequal centrifugal forces. The zero position of this balance is determined by repeating the measurement after rotating the apparatus by 180°. An effect is only observed for $m_t(1) \neq m_t(2)$. Eötvös *et al* found no difference between the forces for different pairs of material samples [Eöt22]. The relative error in the equation

$$\frac{m_s(1)}{m_t(1)} = \frac{m_s(2)}{m_t(2)} \tag{11.53}$$

was around 5×10^{-9}.

In an improvement on this experiment, Roll, Krotkov and Dicke used the smaller centrifugal effect of the Sun to reach an accuracy of 3×10^{-11} for samples of aluminium and gold [Rol64]. The experiment was later repeated by another group for aluminium and platinum [Bra72]. On a level of 0.9×10^{-11}

this measurement also showed no deviations from the equivalence principle. However, as previously mentioned, these studies provide information only about forces with very long ranges.

A new force with a medium range could, however, be apparent in experiments with torsion pendulums. A new analysis of the data from Eötvös's original experiment actually produced a surprising result. Study of Eötvös's samples from different materials showed they exhibited an unexpected systematic pattern which could imply a possible baryon-number-dependent or hypercharge-dependent fifth force [Fis86].

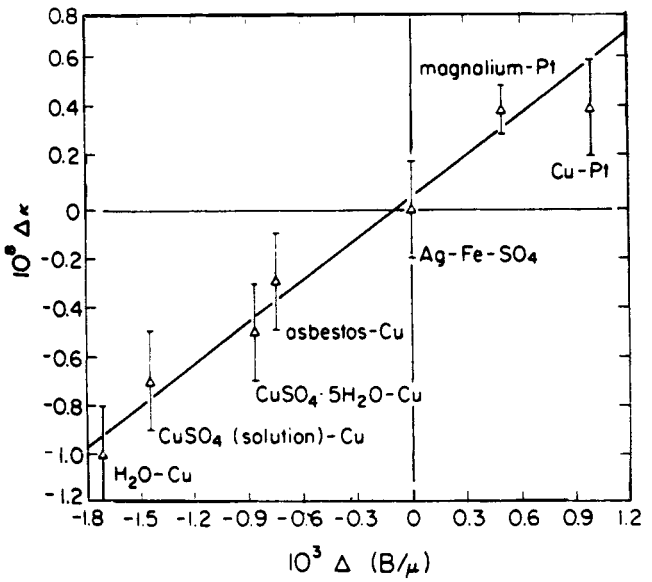


Figure 11.5. Analysis of the Eötvös data by Fischbach *et al* [Fis86]. The dependence of the interaction on the quotient of baryon number and mass is shown in a Δa–$\Delta(B/\mu)$ representation (from [Fis86]) ($\Delta K \equiv \Delta a$).

In (11.24), we saw that the potential difference between two bodies in the force field of a mass M (in this example, the Earth) is proportional to $\Delta(B/\mu)$ if the fifth force is coupled to the baryon number B. Since the binding energy per nucleon in the nucleus has a maximum for iron and decreases to both sides (i.e. for smaller and larger atomic numbers), the ratio B/μ changes over the periodic system of the elements. Figure 11.5 shows the deviation of the acceleration Δa from the value obtained from the interaction of identical substances, as a function of the quotient of baryon number and mass for the samples studied by Eötvös [Fis86]. It is clear that the result of the torsion balance experiment exhibits a small dependence on the nature of the substance used. Fischbach *et*

al deduced a power law of the form (see (11.21) and (11.22) with $Q = B$)

$$V = -\frac{G_\infty m_1 m_2}{r}\left[1 - \alpha_B \left(\frac{B}{\mu}\right)_1 \left(\frac{B}{\mu}\right)_2 \mathrm{e}^{-r/\lambda_B}\right]. \tag{11.54}$$

However, it was not possible to determine the parameters α_B and λ_B since the conditions under which the experiment was carried out could no longer be reconstructed [Fis86, Spe88]. More exact analyses showed that the results are strongly affected by the surroundings of the torsion balance equipment. This is ultimately a consequence of the relatively short range of the postulated fifth force in comparison with the normal gravity. Thus, the interpretation of the slope of the lines in figure 11.5 is uncertain.

The result found by Fischbach *et al* triggered a controversial discussion about the geometry and the sign and the strength of the interaction. We shall not discuss details of this here, but refer interested readers to [DeR86, Sta87, Nie89, 91]. The hypothesis that it is a vector force which couples to the hypercharge [Fis86] was ruled out, at least for the required strength of the interaction (see e.g. [Nie91]), by the non-observation of the decay

$$K^+ \rightarrow \pi^+ + \text{ unobserved neutral particles.} \tag{11.55}$$

Despite all the criticisms of various details the correlation found in the new analysis appeared to be physically real. This result has stimulated a number of modern experiments. Some working groups work, like Eötvös, with pairs of samples of different compositions suspended on torsion wires. While the classical experiment was static, modern versions are often dynamic experiments.

11.3.3.2 Modern experiments on a substance-dependent fifth force

One of the first modern experiments to search for a substance-dependent fifth force was carried out by Thieberger [Thi87]. The structure of the experiment was very different from that of Eötvös's original experiment. By appropriate dimensioning, a hollow copper ball ($m_{\mathrm{Cu}} = 4.925$ kg, diameter 21.11 cm) is floated in a container of water. It should not move under the effect of Newtonian gravity. In the equilibrium state, the weight of the ball is exactly compensated by the upthrust and the mass of the copper is exactly the same as that of the water displaced.

An additional force acting horizontally could perturb the equilibrium, as a result of the chemical differences between copper and water, and lead to a horizontal drift of the ball. The difference between the horizontal components of the acceleration should be proportional to the difference in charge

$$\Delta a \sim \alpha_Q \Delta \left(\frac{Q}{\mu}\right). \tag{11.56}$$

The horizontal acceleration may be determined from measurement of the drift velocity v_D with which the ball of radius r moves through the liquid with viscosity η. Using Stokes's law, the equation of motion is

$$m\Delta a - 6\pi\eta r v_D = 0. \tag{11.57}$$

With $m = (4\pi/3)r^3\rho$, it follows that

$$\Delta a = \frac{9}{2}\frac{\eta v_D}{\rho r^2}. \tag{11.58}$$

After correction for a non-negligible Reynolds number [Som50], we obtain

$$\Delta a = \frac{9}{2}\frac{v_D\eta}{\rho r^2}\left(1 + \frac{3}{8}\frac{v_D r\rho}{\eta}\right). \tag{11.59}$$

Thieberger placed this apparatus on the edge of a cliff, the New Jersey Pallisades Cliff. Thus, the average distribution of matter was sufficiently asymmetric that a substance-dependent force generated by the cliff could give the ball an acceleration relative to the water. The temperature of the water was kept at a constant value of 4.0 ± 0.2 °C so that convection currents were essentially ruled out. Distilled water was used to avoid perturbing chemical reactions and the content of dissolved oxygen was reduced by replacement with nitrogen.

In fact, Thieberger observed a systematic sidewards drift of the copper ball with a typical drift velocity of 4.7 mm h^{-1}. According to (11.59), this corresponds to an acceleration of [Thi87]

$$\Delta a = (8.5 \pm 1.3) \times 10^{-8} \text{ cm s}^{-2}. \tag{11.60}$$

This result is consistent with the assumption of a substance-dependent fifth force with a medium range. This new interaction would, accordingly, be more repulsive for copper than for water.

Assuming that the force couples to the baryon number, it is possible to obtain information about the parameters α and λ of (11.19). The following equation holds for the ratio of the baryon number to the mass B/μ

$$\left(\frac{B}{\mu}\right)_{\text{Cu}} = 1.00171\left(\frac{B}{\mu}\right)_{\text{H}_2\text{O}}. \tag{11.61}$$

It turns out that

$$\alpha\lambda = -(1.2 \pm 0.4) \text{ m} \qquad \text{for } 5 \text{ m} \ll \lambda < 100 \text{ m}. \tag{11.62}$$

However, no movement was observed in a similar experiment carried out at Vallombrosa near Florence [Biz89]. This experiment used a solid plastic cube

(weight fractions: 93.8% nylon ($[C_{12}H_{23}NO]_n$) and 6.2% $C_{10}H_{15}NO_2S$) of radius 6 cm, which was allowed to move freely in a salt solution with practically the same density. The charge difference $\Delta(B/\mu)$ in this experiment was

$$\Delta\left(\frac{B}{\mu}\right) = -3.8 \times 10^{-4}. \tag{11.63}$$

Assuming a vectorial force coupling to B, the authors obtained [Biz89]

$$|\alpha\lambda| < 0.25\ \text{m} \qquad \text{for } 60\,\text{m} < \lambda < 500\,\text{m}. \tag{11.64}$$

Boynton *et al* studied the torsion oscillations of a mass dipole in the gravitational field of a sideways-oriented mass formed by a 130 m high cliff near Index/Washington [Boy87]. The mass dipole was a half-aluminium, half-beryllium ring (see figure 11.6) which was suspended on a tungsten filament. Depending on the orientation of the dipole, a fifth force between the cliff and the ring would generate a stabilizing or destabilizing torque, which would be apparent via different periods of oscillation.

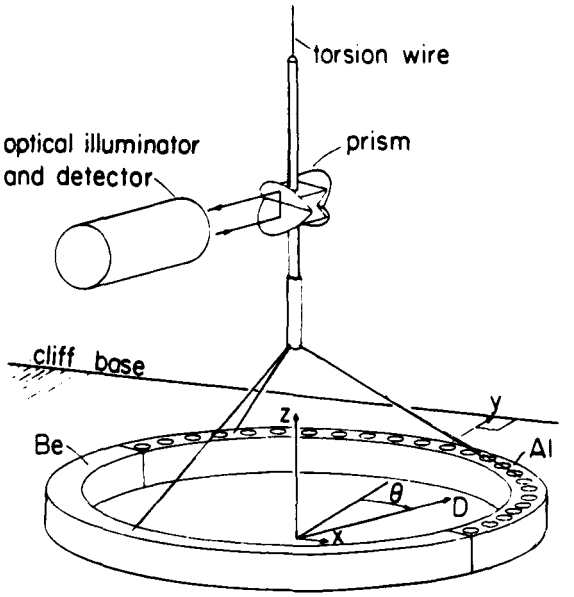


Figure 11.6. Experiment to measure the torsion oscillations of a mass dipole in the gravitational field of a laterally located mass (in this case, a cliff of height 130 m). D is the dipole axis (from [Boy87]).

Unlike Eötvös's original experiment the period of a torsion oscillation rather than the static deflection of the dipole axis was measured for various initial

conditions. The torsion pendulum was started once for an angle θ between the dipole axis and the surface of the cliff and once for an angle $\theta + \pi$. The difference between the two periods may be used to derive information about a fifth force. A significant effect was actually observed

$$\alpha\lambda = (-2.3 \pm 0.6) \times 10^{-2}\ \mathrm{m} \qquad \text{for } \lambda = 100\,\mathrm{m}. \tag{11.65}$$

This value is much smaller than that obtained by Thieberger. However, Boynton *et al* showed that the results could be harmonized if one postulated a force which was essentially only proportional to the z component of the nuclear isospin $I_3 = (N - Z) = (B - 2Z)$ (cf. (11.28) with $\theta_5 \approx 0$). The experimenters were not able to confirm this with improved apparatus [Boy90]; however, other materials were used.

A coupling to the isospin was strongly restricted by the non-observation of the decay of the K meson to the postulated exchange boson [Gol88b]. Later experiments essentially ruled out the isospin hypothesis [Cow88, 90, Spe88b, Stu89, Nel90].

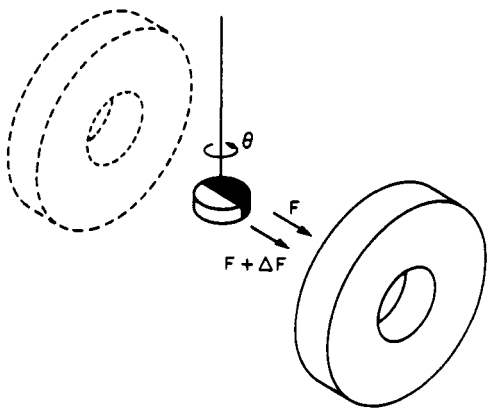


Figure 11.7. Torsion balance of Nelson, Graham and Newman (from [Nel90]) (see text).

Nelson, Graham and Newman [Nel90] used the measurement method illustrated schematically in figure 11.7. An attractive mass in the form of a 320 kg lead ring in an aluminium jacket is periodically moved from one side of the torsion balance to the other. A substance-dependent force on the sample mass, which essentially consisted of lead and copper would apply an angular momentum to the torsion balance, the sign of which would be inverted when the large attractive mass was moved.

Assuming the parametrization of the charge $Q = B \sin\theta_5 + (N - Z)\cos\theta_5$ given in (11.28), the measurement results give the following very sharp bounds

for the coupling parameter α_Q in (11.21) [Nel90]:

$$\alpha_B = (-1.2 \pm 1.3) \times 10^{-3} \text{ for } \theta_5 = 90° \quad (B \text{ coupling}) \tag{11.66a}$$

$$\alpha_{I_3} = (5.7 \pm 6.3) \times 10^{-5} \quad \text{for } \theta_5 = 0° \quad ((N-Z) \text{ coupling}). \tag{11.66b}$$

For $\lambda > 1$ m, the results are practically independent of the range of the interaction. Thus, coupling to the isospin is ruled out with great precision. Figure 11.8 shows the bounds at the 2σ level for the coupling strength α_Q as a function of the range for a Yukawa potential of the form (11.21) for $(N-Z)$ and

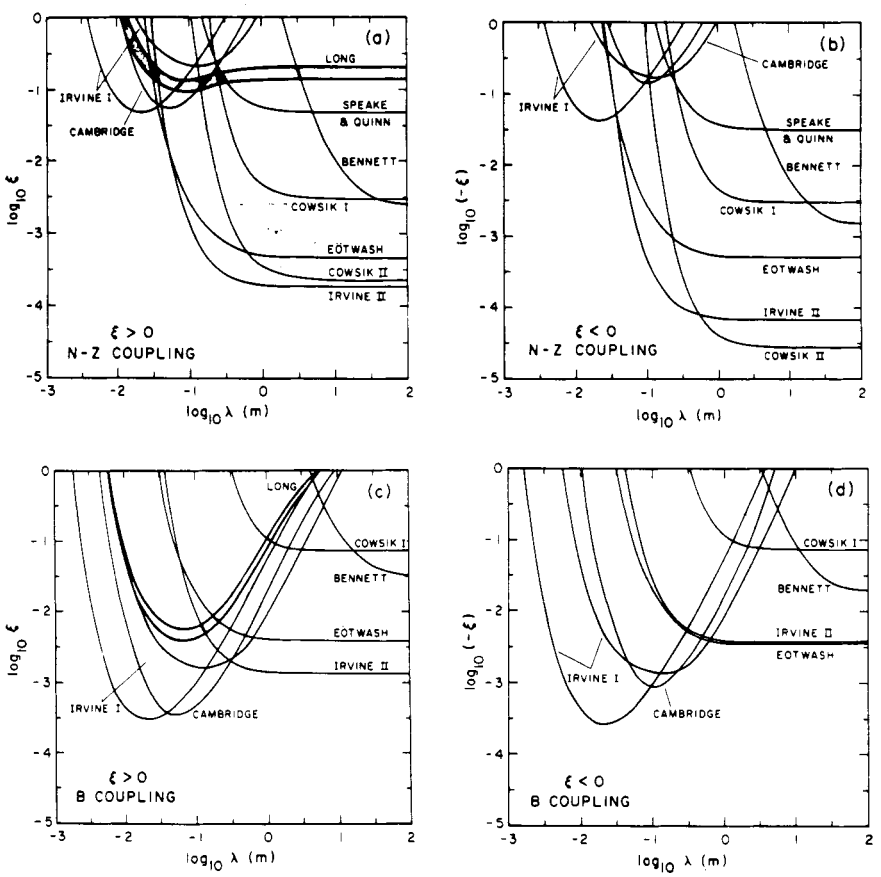


Figure 11.8. Measured bounds for the coupling strength $dQ = \xi$ assuming a Yukawa potential of the form (11.21) for $(N-Z)$ coupling and B coupling (from [Nel90]). The figure also shows the results of various other experiments: Long = [Lon76], Irvine I = [Hos85], Cambridge = [Che84], Irvine II = [Nel90], Bennett = [Ben89], Cowsik I = [Cow88], Cowsik II = [Cow90], Eöt–Wash = [Stu89] and Speake and Quinn [Spe88].

for B coupling. The figure also shows the results of various other measurements, not all of which can be discussed here.

Another 'zero-experiment' (an experiment which did not detect a positive effect) was carried out by Niebauer *et al.* They essentially repeated the free-fall experiment of Galileo Galilei said to have been carried out on the Leaning Tower of Pisa, but using modern apparatus [Nie87]. They determined the fall times of bodies with different baryonic compositions (copper and uranium) but the same mass. Both sample bodies were allowed to fall simultaneously in neighbouring vacuum chambers. The experimenters measured the differences in the acceleration using interferometers. The resolution of the apparatus was $\Delta g/g \simeq 5 \times 10^{-10}$. From the measured fall times it followed [Nie87]

$$|\alpha\lambda| = 1.6 \pm 6.0\,\mathrm{m}. \tag{11.67}$$

This freefall experiment is sensitive to the mass directly below the apparatus. Horizontal mass anomalies do not affect the result. The value given in (11.67) is valid for distances ranging from 100 m to 1000 km. The sensitivity for small distances is less than in experiments with torsion balances.

A group from the University of Washington (called the Eöt–Wash group, after the experiment carried out by Eötvös) carried out measurements with an apparatus of the Eötvös type. They used a torsion balance with four balls, two being of copper ($B/\mu = 1.00112$) and two of beryllium ($B/\mu = 0.99865$). The setup was located on the slope of a hill, to provide for an asymmetric mass distribution. The torsion balance was rotated relative to the hill in a vacuum container. A fifth force would be manifest via a signal of the appropriate phase in the deflection of the pendulum. No torque which would indicate a new force was found.

Assuming a force which couples to the baryon number, we have the following bounds for α_B from (11.21) and (11.22) [Stu87]:

$$|\alpha_B| < 2 \times 10^{-4} \quad \text{for } 250\ \mathrm{m} < \lambda < 1400\ \mathrm{m} \tag{11.68a}$$

$$|\alpha_B| < 1 \times 10^{-3} \quad \text{for } 30\ \mathrm{m} < \lambda < 250\ \mathrm{m}. \tag{11.68b}$$

Since the values $\Delta((B-L)/\mu) = -1.015 \times 10^{-2}$ and $\Delta(I_3/\mu) = 1.139 \times 10^{-2}$ for the pendulum are much larger than the $\Delta(B/\mu)$ value 2.468×10^{-3}, corresponding much sharper bounds for an interaction which couples to $(B-L)$ or I_3 are obtained (see also figure 11.8). A similar experiment with aluminium and beryllium test masses also failed to exhibit an anomalous torque [Ade87].

Using an improved apparatus (figure 11.9), the Eöt–Wash group measured the differential horizontal acceleration for Be/Al and Be/Cu pairs and obtained the following results ((11.69) [Hec89], (11.70) [Ade90]):

$$\Delta a(\mathrm{Al/Be}) = (1.5 \pm 2.3) \times 10^{-11}\ \mathrm{cm\ s^{-2}} \tag{11.69a}$$

$$\Delta a(\mathrm{Cu/Be}) = (0.9 \pm 1.7) \times 10^{-11}\ \mathrm{cm\ s^{-2}} \tag{11.69b}$$

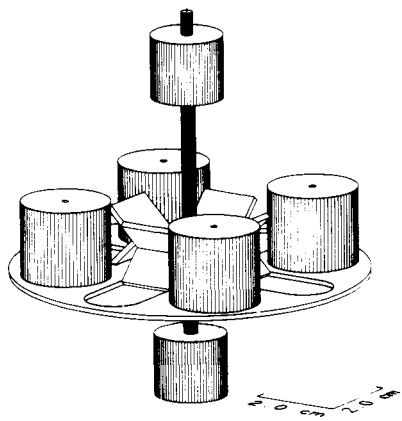


Figure 11.9. Torsion balance of the Eöt–Wash group to measure the differential horizontal acceleration of Be/Al and Be/Cu pairs (from [Ade90]).

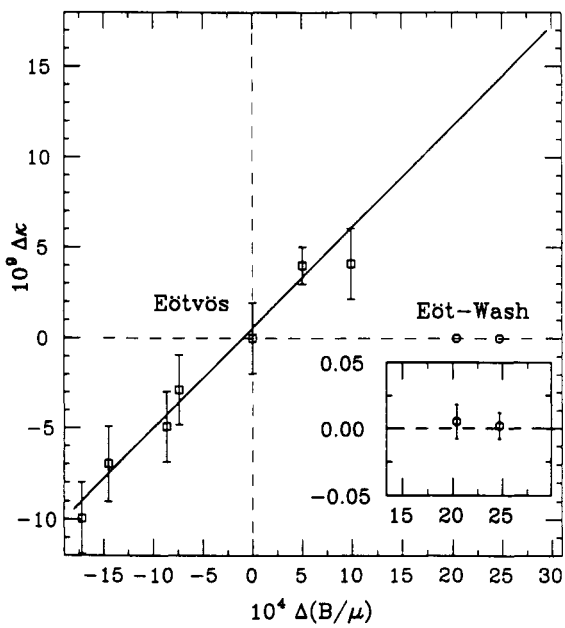


Figure 11.10. Results of the Eöt–Wash group compared with the analysis of the data from the Eötvös experiment in the Δa–$\Delta(B/\mu)$ plane (see figure 11.5). $\Delta K = \Delta a$ (from [Ade90]). In the inset, the vertical scale is magnified by a factor of 100 to show the error bands of the Eöt–Wash data. The new data contradict the existence of a B-dependent force.

and

$$\Delta a(\mathrm{Al/Be}) = (2.1 \pm 2.1) \times 10^{-11}\ \mathrm{cm\ s^{-2}} \qquad (11.70a)$$

$$\Delta a(\mathrm{Cu/Be}) = (0.8 \pm 1.7) \times 10^{-11}\ \mathrm{cm\ s^{-2}}. \qquad (11.70b)$$

No effects were observed which cannot be explained in the framework of known physics. This gives very sharp restrictions on the parameters of a new force and on the mass of the bosons – in the region between $3 \times 10^{-18}\mathrm{eV}/c^2$ and $1 \times 10^{-6}\ \mathrm{eV}/c^2$ – which mediate the Yukawa interaction. Figure 11.10, which gives the new Eöt–Wash data together with the results of the new analysis of the original Eötvös experiment, serves as an example of the enormous precision achieved in this experiment[2]. To within measurement accuracy, the new data contradict the existence of a B-dependent force.

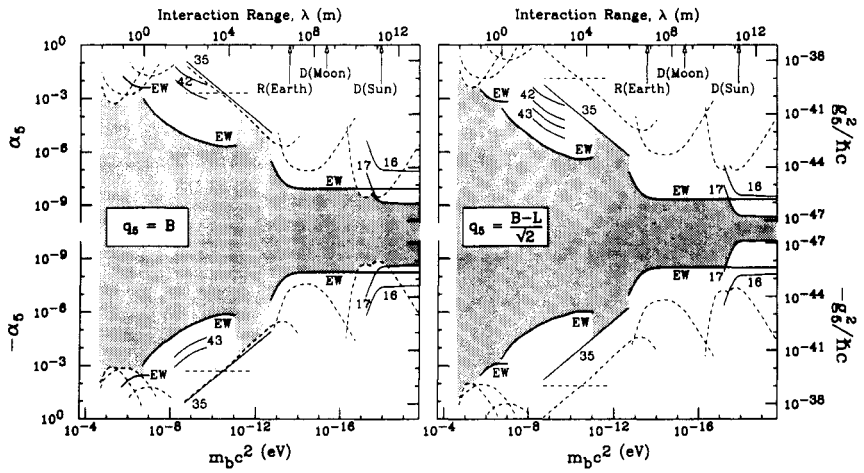


Figure 11.11. Experimental bounds for interactions via light vector bosons from the Eöt–Wash (EW) and other experiments. Bounds for the parameter α ($= \alpha_5$) as a function of the boson mass $m_b c^2$ (or the range λ), for different charges $Q = q_5$ (from [Ade90]; for details see the latter). For improved limits see [Su94].

Figure 11.11 shows the bounds for the parameters α and λ for various postulated charges (EW denotes the data of the experiment described above). To show the precision of the Eötvös experiment, we shall now restate a numerical example given by Adelberger [Ade91b]. It follows from the non-observation of an effect that the energy of the fundamental mode of oscillation (mode of oscillation with smallest energy) of the torsion pendulum is less than 5 μeV for the 70 g pendulum. This amounts to less than 10^{-29} eV per atom of the

[2] A considerably improved version of this experiment has been described recently [Su94].

pendulum. The energy 10^{-29} eV corresponds to the electrostatic energy between two electrons separated by a relative distance of 10 000 light years!

This completes our discussion of the experimental search for a new fundamental interaction. The list of measurements discussed is naturally incomplete. We refer interested readers to [Nie91] and the references given there. Overall, it appears that after initial positive evidence for a fifth force disillusionment has set in. Most of the new highly precise measurements are consistent with a substance-independent $1/r$ potential. The positive evidence found can probably be traced back to systematic errors, which only simulate a fifth force. Nevertheless, the existence of a new interaction cannot be completely ruled out.

11.3.3.3 The effect of matter–antimatter asymmetry

An experiment already proposed a long time ago at the antiproton ring LEAR at CERN [Gol82, 86] is of great interest as far as the verification of the scalar–vector model in the framework of quantum theories of gravity is concerned. This experiment compares the rise and fall of protons and antiprotons in the Earth's gravitational field (see the discussion of the effect of antimatter in section 11.2.4).

Finally, we point out an interesting relationship between a matter–antimatter dependence of the force of gravity and the search for neutron–antineutron oscillations [Lam91]. The observation of $n\bar{n}$ oscillations would give very restrictive bounds for a matter–antimatter asymmetry of gravity. The existence of such an interaction in the framework of current measurement accuracy would make the observation of $n\bar{n}$ oscillations appear highly unlikely.

In chapter 5, we saw that the probability $P_{n\bar{n}}(t)$ of finding an antineutron in a pure neutron beam after time t is given by (5.55). In the case of an additional gravitational interaction $V^G_{n\bar{n}}$ between neutrons and antineutrons, this probability is affected, since the new potential increases the splitting between the two states

$$2\Delta E \to 2\Delta E + V^G_{n\bar{n}}. \tag{11.71}$$

The quantity ΔMc^2 defined in (5.58) becomes

$$\Delta Mc^2 = \sqrt{\delta m^2 c^4 + (\Delta E + \tfrac{1}{2} V^G_{n\bar{n}})^2}. \tag{11.72}$$

If the quasi-free condition (5.70)

$$\Delta Mc^2 t \ll \hbar \tag{11.73}$$

is satisfied, we then have the simple relation

$$P_{n\bar{n}}(t) = \left(\frac{t}{\tau_{n\bar{n}}}\right)^2 \tag{11.74}$$

where

$$\tau_{n\bar{n}} = \frac{\hbar}{\delta m c^2}. \tag{11.75}$$

Since $\delta m \ll \Delta M$, it follows from (11.73) that

$$\left(\Delta E + \tfrac{1}{2} V_{n\bar{n}}^{G}\right) t \ll \hbar. \tag{11.76}$$

The following values for the energy shift ΔE due to the Earth's residual magnetic field and for the typical flight time t are valid for the experiment carried out at the ILL [Bal90, Bald94] (see chapter 5)

$$\Delta E \simeq 10^{-15}\ \text{eV} \qquad t \simeq 0.1\ \text{s}. \tag{11.77}$$

Using these values, we estimate the new potential having to satisfy the bound

$$V_{n\bar{n}}^{G} < 10^{-14}\,\text{eV}. \tag{11.78}$$

This corresponds to approximately 10^{-14} times the normal Newtonian gravitational potential of the neutron on the Earth's surface (0.64 eV). Thus, this circuitous route may also be used to achieve a very large sensitivity, provided the matrix element δm does not vanish.

Chapter 12

Time Dependence of Natural Constants

12.1 INTRODUCTION

The standard model of the electroweak theory contains 17 free parameters, coupling constants, mixing angles and masses (see chapter 1). Since the origin of these quantities is as yet unexplained, it is conceivable that some of these 'constants' may be subject to temporal variations. For simplicity, we shall continue to speak of natural *constants* in what follows even when discussing their temporal variation.

The question of the time dependence of natural constants was first raised by P A M Dirac in 1937 [Dir37]. There are many natural ratios of constants which, roughly speaking, have order of magnitude one, for example, $\alpha = 1/137$, $m_e/m_\mu \approx 1/200, \ldots$. On the other hand, there are also very large dimensionless constants such as the ratio of the electrostatic and the gravitational attraction between an electron and a proton, $e^2/Gm_pm_e \simeq 2 \times 10^{39}$. Dirac forwarded the hypothesis that the latter are not purely mathematical numbers but variable parameters which characterize the current state of the universe. In fact, he noted that the age of the universe, expressed in natural units e^2/m_ec^3 corresponds approximately to the quantity e^2/Gm_pm_e. This natural unit of time is defined to be the time taken by a beam of light in vacuum to travel a distance corresponding to the classical radius of the electron. This unit of time was previously called a 'tempon' [French], where 1 tempon $\approx 10^{-23}$ s. The age of the universe is approximately 10×10^9 years or 3×10^{40} tempons. This number is very close to the value of the ratio mentioned above.

Dirac also suggested that this equality was no accident, but that the two numbers should be practically the same at any time[1], i.e.

$$\frac{e^2}{Gm_em_p} \approx \frac{m_ec^3}{e^2}t. \tag{12.1}$$

[1] Dirac's large-number hypothesis.

This means that dimensionless constants of order of magnitude 10^{40} should vary linearly with time. Assuming that the atomic constants are not subject to temporal variations, it follows that the constant of gravity G decreases with the time t

$$G \propto t^{-1}. \tag{12.2}$$

The hypothesis may also be extended and pursued further, so that dimensionless numbers of order of magnitude $(10^{40})^n$ should vary with the age of the universe according to t^n. If one estimates the number of baryons in the universe by dividing the visible mass of the universe by the proton mass, one obtains an estimate of approximately 10^{78} baryons [Nor86]. Dirac predicted an increase in the number of baryons in the universe with t^2.

Some 10 years later, Teller showed that a decrease in G proportional to t^{-1} appeared to contradict findings about evolution [Tel48, Gam67]. He derived the following relationship between the luminosity L of a star of mass M and the gravitational constant G

$$L \propto G^7 M^5. \tag{12.3}$$

Thus, a G which was larger at an earlier time would have resulted in a greater luminosity of the Sun and a smaller radius of the Earth's orbit

$$R \propto G^{-1}. \tag{12.4}$$

According to the Dirac hypothesis, the variation of the constant of gravity with time is directly related to the expansion rate of the universe, the so-called Hubble constant, $H = 40$–100 km Mpc^{-1} s^{-1}:

$$|\dot{G}/G| = H = (4 \times 10^{-11}\text{–}1 \times 10^{-10})\ \text{year}^{-1}. \tag{12.5}$$

Since G decreases with time, we have $\dot{G}/G \simeq -5 \times 10^{-11}$ year^{-1}. The resulting surface temperature of the Earth around two billion years ago would have made the development of life on Earth impossible. According to this argument, evolution rules out equation (12.2). Astrophysical considerations also led to contradictions with $G \sim t^{-1}$ [Poc64].

Furthermore, in 1967, Gamow suggested that G in (12.1) could remain constant if the elementary charge were to increase with time [Gam67]

$$e^2 \propto t. \tag{12.6}$$

Working independently, Teller formulated [Tel48] the following hypothesis

$$\alpha^{-1} = \frac{\hbar c}{e^2} \propto \ln \frac{t m_e c^2}{\hbar}. \tag{12.7}$$

However, studies by Dyson of the ^{187}Re–^{187}Os system cast doubt on both approaches [Dys67].

Recently, it has been recognized that the above arguments which contradict the assumption $G \propto t^{-1}$ are no longer valid [Wes80, Can81, Nor86]. The astrophysical obstacle was based on the fact that the Sun should already be a red giant if the age of the universe was not at least 15 billion years. While this value was still viewed as unphysically high in 1964, it is now within the currently conjectured age range of from 10 to 20 billion years [Kla83]. Teller's argument about the luminosity of the Sun and the temperature of the Earth has also been invalidated. An accurate study of systems which are subject only to gravity showed that G and M always occur in the combination GM. In particular, the following secondary condition applies in considerations of, for example, the structure of the Sun [Can81]

$$GM = \text{constant} \tag{12.8}$$

whence the luminosity of a star is essentially time independent. This aspect was not taken into account by Teller and other authors; thus, equation (12.3) which predicts a strong variation of L with G and M is wrong.

Dirac's speculative ideas have led to a variety of experiments in the search for a possible time dependence of natural constants. The importance of these measurements has been increased by more recent theoretical models in which the values of the coupling constants are related to the radii of so-called compactified dimensions. The basic ideas underlying these Kaluza–Klein theories are discussed in the next section. Before that, however, we shall point out another interesting consequence of a possible variation of the constant of gravity G with time.

In the framework of Newtonian mechanics, a time-dependent G leads to a violation of the principle of the conservation of energy, as can be easily seen from the following consideration [Bis76]. A ring-shaped object and a small ball-shaped object may move towards each other from infinity as a result of the mutual mass attraction, fly through each other and separate again. If $G(t)$ decreases with time, the attractive force between the two objects at a given relative distance during the approach is greater than that during the separation; whence, the relative velocity and thus, also, the kinetic energy after the objects pass is greater than before. Since the potential energy vanishes for large relative distances the energy principle is violated in the case of interacting particles. Consequently, the requirement for energy conservation and Newton's law of gravity in the form

$$\boldsymbol{F}(\boldsymbol{r}) = -G(t)\frac{m_1 m_2 \boldsymbol{r}}{r^3} \tag{12.9}$$

are not compatible if $G(t) \neq$ constant. (Generally a time-dependent Hamiltonian violates energy conservation, see section 1.3.1.)

If we further assume that the conservation of energy is more fundamental than Newton's law in the form (12.9), a new force law may be derived. In what

follows, we shall adhere closely to the presentation in [Bis76]. We consider a system of N particles with masses m_i. The potential energy of this system is

$$V = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} m_i m_j \mathcal{V}(r_{ij}, t) \tag{12.10}$$

where $r_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j|$ denotes the relative distance between the particles i and j. Clearly, $\mathcal{V}(r_{ij}, t)$ has the form

$$\mathcal{V}(r_{ij}, t) = -\frac{G(t)}{r_{ij}}. \tag{12.11}$$

The energy principle gives

$$\frac{1}{2} \sum_{i=1}^{N} \sum_{\alpha=1}^{3} m_i \dot{x}_{i\alpha}^2 + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} m_i m_j \mathcal{V}(r_{ij}, t) = E \tag{12.12}$$

where E is a constant. Differentiation with respect to time gives

$$0 = \sum_{i=1}^{N} m_i \sum_{\alpha=1}^{3} \dot{x}_{i\alpha} \left[\ddot{x}_{i\alpha} + \sum_{\substack{j=1 \\ j \neq i}}^{N} m_j \left(\frac{\partial \mathcal{V}}{\partial r_{ij\alpha}} + \frac{\dot{r}_{ij\alpha}}{v_{ij}^2} \frac{\partial \mathcal{V}}{\partial t} \right) \right] \tag{12.13}$$

where $v_{ij} = |\dot{r}_{ij}|$. Equation (12.13) points to the following force law

$$\ddot{x}_{i\alpha} = -\sum_{\substack{j=1 \\ j \neq i}}^{N} m_j \left(\frac{\partial \mathcal{V}}{\partial r_{ij\alpha}} + \frac{\dot{r}_{ij\alpha}}{v_{ij}^2} \frac{\partial \mathcal{V}}{\partial t} \right). \tag{12.14}$$

How big is the correction term introduced in (12.14)? For a particle moving with velocity v on the Earth's surface (radius of the Earth r), by virtue of (12.11) we have

$$\left(\frac{\dot{r}}{c^2} \frac{\partial v}{\partial t} \right) \Big/ \left(\frac{\partial v}{\partial r} \right) = 1 - \frac{\dot{G}}{G} \frac{r}{v}. \tag{12.15}$$

Thus, the correction amounts to

$$|\xi| = \left| \frac{\dot{G}}{G} \right| \frac{r}{v} \approx \frac{10^{-9}\,\mathrm{cm\ s^{-1}}}{v}. \tag{12.16}$$

Analogously, for two atoms at a relative distance of $a = 10^{-8}$ cm in a solid body at temperature T, with $v \approx \sqrt{kT/m}$, we have

$$|\xi| = \sqrt{\frac{m}{kT}}\, a \left| \frac{\dot{G}}{G} \right| \approx \frac{10^{-30}\sqrt{K}}{\sqrt{T}}. \tag{12.17}$$

The given numerical values refer to the Hubble constant given in (12.5). The correction will generally be negligible.

Even if one does not use (12.11), the new force law implies that in the universe there are no two particles which are at rest relative to one another, if gravity is time dependent. This conclusion agrees with the observation that practically all physical systems are in a state of relative motion, from the zero point movement in microscopic systems to the expansion of the universe. Since the new force law does not necessarily lead to a radial force, angular momentum is not in general conserved.

12.2 THEORETICAL PREDICTIONS

There are as yet no basic theories which make quantitative statements about the size of a possible temporal variation of natural constants. However, a dependence on time is permitted in the framework of models with more than four dimensions, so-called Kaluza–Klein theories. Typically, these higher-dimensional theories have the structure $M^4 \otimes C^N$ where M^4 represents four-dimensional space–time and C^N is an N-dimensional compact space which at low energies allows a quantum field theory of Yang–Mills type.

As early as 1921, Kaluza attempted to unify gravity and electromagnetism in a five-dimensional model using Riemannian geometry [Kal21]. Corresponding to the description of gravity by the curvature in the Minkowski space, the curvature of the additional fifth dimension was to explain the electromagnetic interaction. These ideas were later developed further by Klein [Kle26]. However, the underlying space–time continuum could not be quantized so that this approach was initially relegated to the background in preference for the gauge theories introduced by Weyl. The concept of multidimensional unification theories has recently come to the fore again, since the Kaluza–Klein approach appears promising as far as the unification of gauge theories and the general theory of relativity is concerned [Wit81]. However, it is no longer sufficient to extend the Minkowski space by just one dimension, since the objective today is the unification of all fundamental forces.

By contrast to most classical works, the extra dimensions are viewed as equivalent, true, physical dimensions. The clear difference between the four observed dimensions and the extra dimensions could be the consequence of a spontaneous symmetry breaking ('spontaneous compactification') [Cre76]. The extra dimensions are assumed to be compact, i.e. the average radii are so small as to be inaccessible to direct observation.

Generalized Kaluza–Klein models are based on the extension of the four-dimensional space–time to $(4 + N)$ dimensions, so that the N new dimensions form a very small compact manifold with average radius R_{KK}. R_{KK} should be

of the order of magnitude of the Planck length

$$l_{\mathrm{Pl}} = \sqrt{\frac{G\hbar}{c^3}} \simeq 1.6 \times 10^{-33} \text{ cm}. \tag{12.18}$$

For energies $\ll \lambda\hbar/R_{\mathrm{KK}}$, the four-dimensional metric tensor $g_{\mu\nu}$ describes both the general theory of relativity and the gauge interactions. In particular, Witten showed that Kaluza-Klein models with $N \geq 7$ contain the full $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge group of the strong and electroweak interaction [Wit81].

The spontaneous compactification and the model of supergravitation could profitably be combined. According to our present knowledge, the known gauge interactions may be described by the gauge group $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$. This must be contained in the symmetry group $\mathcal{G}$ of the compact manifold

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1) \subset \mathcal{G}. \tag{12.19}$$

It was found, that the 11-dimensional manifold is the smallest with the desired symmetry. To construct a theory in which the $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge fields occur as components of the gravitational field in more than four dimensions, thus requires at least seven extra dimensions in addition to the four non-compact space–time dimensions. This result is remarkable since it is believed that 11 dimensions represent the maximum for supergravity theories[2] [Nah78, Gri77, Ber79, Ara80] or are at least mathematically very appealing[3].

It appears promising that the fewest possible dimensions for Kaluza–Klein theories are the same as the greatest possible dimensions for supergravity theories. However, this approach includes a number of unresolved problems related to the helicity of the quarks and the leptons.

Right- and left-handed fermions differ in their behaviour under gauge transformations; left-handed fermions form SU(2) doublets, while right-handed fermions form SU(2) singlets. This fact is on the one hand the reason why quarks and leptons do not have a manifest mass, but receive their mass via the Higgs mechanism as a result of the spontaneous symmetry breaking (see chapter 1). On the other hand, it also provides a theoretical explanation for the smallness of the lepton and quark masses in comparison with the mass scale of grand unification theories or with the Planck mass m_{Pl}.

However, in the Kaluza–Klein model discussed above, right- and left-handed fermions are transformed identically under $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ [Wit81]. Thus, these particles could be assigned a manifest mass, which could be

[2] SUGRA theories with $d > 11$ would contain massless particles with spin greater than two [Nah78]. However, there are theoretical reasons for supposing that there does not exist a consistent field theory with a gravitational interaction which couples to massless particles with spin greater than two [Gri77, Ber79, Ara80].

[3] A very good representation is given in [Fre85].

arbitrarily large. The renormalizability of the theory represents another problem; however, a discussion of this would be outside the scope of the present book and we shall not pursue it further.

At this point, we refer back to relationships mentioned in previous chapters. C, P and CP violations are possible in the 11-dimensional theory; however, the angle θ discussed in section 1.3 vanishes at the tree level so that the problem of the strong CP violation (θ problem) could be solved in these models. However, the contribution of quantum corrections to θ is difficult to estimate.

Moreover, the 11-dimensional SUGRA theory does not have a global symmetry which could be interpreted as the baryon number, so that the proton would be unstable. The mass scale for the nucleon decay is given by $1/R_{KK} \sim 1/l_{Pl}$ and thus corresponds to the Planck mass m_{Pl}. This implies a lifetime of around 10^{45} years for the nucleon, which is far above the experimentally achievable bounds.

The supersymmetry discussed initially is an attractive concept which could contribute to the solution of open questions relating to GUT models. However, the broad consensus is that the combination of supersymmetry and conventional quantum field theory alone does not lead to a consistent quantum theory of gravity. While the theory contains point-like fermions, it will diverge for energies greater than m_{Pl} (on the problem of the quantization of gravity, see e.g. [DeWit62]).

One attempt to solve this problem is based on fermions as one-dimensional extended objects or strings. It appears possible to construct superstring theories which are identical to a point-like quantum field theory with supersymmetry for energies considerably lower than the Planck mass, but which do not diverge for energies above the Planck mass as a result of the string nature of the fermions.

One promising superstring theory is based on the gauge group $E_8 \otimes E_8$ [Gre85]. However, this theory can only be formulated in a 10-dimensional geometrical space which must then compactify to the four observed space–time dimensions. For details, readers are referred to [Gre87, Moh86a].

Two questions now arise:

(i) Can the extra dimensions of the Kaluza–Klein and superstring theories be observed?

(ii) How are these extra dimensions related to a possible time dependence of the natural constants?

Since the average radii of the extra dimensions R_{KK} should be of the order of magnitude of the Planck length l_{Pl}, direct observation appears impossible. However, when there are $4 + N$ dimensions the coupling constants of the associated gauge theories in four dimensions are related to the size of the compact N-dimensional space. The true natural constants would be defined in the $(4 + N)$ dimensional space. The R_{KK} of the Kaluza–Klein theories are

related to the Newtonian constant of gravity G and the coupling constants

$$\alpha_i(R_{KK}) = \frac{g_i^2(R_{KK})}{4\pi} \qquad i = U(1), SU(2), SU(3) \tag{12.20}$$

of the three gauge groups U(1), SU(2) and SU(3) at very small distances by the following quantization conditions [Kal21, Kle26, Wei83, Mar84a]

$$\alpha_i(R_{KK}) = \frac{K_i G}{R_{KK}^2} = K_i G m_{KK}^2. \tag{12.21}$$

The numbers K_i come from the topology of the N-dimensional space and play no part in what follows.

The relationship between $\alpha_i(m_{KK})$ and the effective coupling constants for large distances $\alpha_i(\mu)$, with $\mu \ll m_{KK}$ is given by [Mar84a]

$$\alpha_i^{-1}(\mu) = \alpha_i^{-1}(m_{KK}) - \pi^{-1} \sum_j C_{ij} \left[\ln\left(\frac{m_{KK}}{m_j}\right) + \Theta(\mu - m_j) \ln\left(\frac{m_j}{\mu}\right) \right]. \tag{12.22}$$

For $i = 3$, (12.22) is only applicable for $\mu > 1$ GeV. The sum $\sum_j$ is taken over all leptons, quarks, gluons, $W^{\pm}$, ..., and the C_{ij} are known constants which depend on the spin and the group representation. The fine structure constant $\alpha(0) \simeq 1/137$, the Fermi coupling constant G_F and the mixing angle may be expressed in terms of the $\alpha_i(\mu)$

$$\alpha^{-1}(\mu) = \frac{5}{3}\alpha_1^{-1}(\mu) + \alpha_2^{-1}(\mu) \tag{12.23a}$$

$$G_F \simeq \frac{\pi \alpha(m_W)}{\sqrt{2} m_W^2} \tag{12.23b}$$

$$\tan^2 \theta_W(m_W) = \frac{3}{5} \frac{\alpha_1(m_W)}{\alpha_2(m_W)} \tag{12.23c}$$

where m_W denotes the mass of the $W^{\pm}$ boson.

While in Kaluza–Klein theories the gauge coupling constants evolve in our four-dimensional world like

$$\alpha_i \propto R_{KK}^{-2} \tag{12.24}$$

for the Newtonian constant of gravity G, we have [Bar87a]

$$G \propto R_{KK}^{-N}. \tag{12.25}$$

On the other hand, in the 10-dimensional superstring theories which are currently the subject of frequent discussion [Gre84, 85], the following dependence is expected [Kol86a]

$$\alpha_i \propto G \propto R_{KK}^{-6}. \tag{12.26}$$

Any cosmological evolution of the N extra dimensions would cause a variation with time and/or space of the usual constants observed in three dimensions, the size of which is determined by the average scale factor $R_{KK}(x, t)$.

If the radii R_{KK} were to vary with time, this would also induce a time dependence in the coupling constants. The radii could shrink, grow or even oscillate. There are several motivations for assuming that the radii are not constant in time. Firstly, it is not known how long the extra dimensions have been compactified so that the compactification could be ongoing. Secondly, in the framework of Kaluza–Klein models, as a natural but not compulsory consequence of the observed expansion of the universe, we have $\dot{R}_{KK} \neq 0$ [Cho80, Fre82].

Thus, multidimensional theories relate the average radii of the compact manifolds to the coupling constants of our four-dimensional world. The extra dimensions could reveal themselves via a time dependence of the coupling constants.

Under very special assumptions, superstring theories predict a temporal variation of the constant of gravity with $\dot{G}/G \simeq -1 \times 10^{-11\pm1}$ year^{-1} [Wu86].

12.3 EXPERIMENTS TO SEARCH FOR THE TIME DEPENDENCE OF NATURAL CONSTANTS

12.3.1 The design of experiments

Since the expected variation of natural constants is very small (cf. e.g. (12.5)) precise measurements are required. We note that such measurements often determine not only one coupling constant but combinations of several constants. Thus, the interpretation of the events is critically dependent on the constants to which the variation is assigned. Under certain circumstances, the dependences in such combinations may cancel out. In addition, one should also ensure that the measurement principle is not based on the constancy of the quantities, the time dependence of which is to be measured.

The experiments may be divided into two categories. One consists of measurements of the current variations of the natural constants and the other of geophysical and astronomical observations which provide for a comparison between the current value of a constant and its value at a given earlier point in time or an average value from earlier periods. For example, the results of a reaction a long time ago may be compared with those from the same reaction at a very much later time. The corresponding reaction cross sections provide information about the coupling constants. One problem in geophysical experiments is the dating of old samples, since the popular measurements of radioactivity also depend on the coupling constants.

There are many experiments of both categories and we are only able to

discuss a few of these. Table 12.1 gives an overview of the bounds obtained for various constants.

Table 12.1. Summary of the most important bounds for the temporal variation of natural constants ($H = 100$ km s^{-1} Mpc^{-1} with $0.4 < h < 1$).

Method	Quantity Q	$\mathrm{d}(\ln Q)/\mathrm{d}t$ [year^{-1}]	Reference
Planetary orbit	G	$(0.2 \pm 0.4) \times 10^{-11}$	[Hel83]
Clock comparison	$g_{Cs}\,(m_e/m_{Cs})\,\alpha^3$	$< 1.2 \times 10^{-11}$	[Tur76]
Reactor	α	$< 1 \times 10^{-17}$	[Shl76]
Fine structure	α	$< 1.3h \times 10^{-13}$	[Bah67]
HFS	α	$< 2h \times 10^{-14}$	[Tub80]
^{187}Re lifetime	α	$< 2 \times 10^{-15}$	[Lin86]
Nucleosynthesis	α	$< 1.5h \times 10^{-14}$	[Kol86a]
Nucleosynthesis	G	$< 9 \times 10^{-13}$	[Acc90]
Red shift	h	$(-3 \pm 4) \times 10^{-13}$	[Sol76]

12.3.2 Experiments to determine current variations

12.3.2.1 Comparison of gravitational and atomic clocks

The radii of the orbits of the planets and the Moon exhibit a sensitive dependence on the Newtonian constant of gravity. Modern measurement methods such as the radar echo procedure may be used to determine the distance to the Moon and other planets with very high accuracy. If such measurements are repeated over a long period of time one obtains information about the time dependence of G. Strictly speaking, one is comparing, for example, an atomic clock with a gravitational clock (figure 12.1).

Atomic clocks such as the caesium clock, the rubidium clock or the hydrogen maser are now used in applications in scientific and technical areas in which the time has to be measured with high accuracy. These clocks make use of the hyperfine structure splitting, which is based on the interaction of the spin of the atomic shell and the nuclear magnetic field. The magnetic energy depends on the orientation of the shell spin J, the energy difference between the states corresponds to the HFS splitting ν_{HFS}. ν_{HFS} is a graphic illustration of the Larmor frequency of the shell spin in the field of the atomic nucleus B_K

$$\nu_{\mathrm{HFS}} = g_J \mu_B \frac{B_K}{h} \tag{12.27}$$

where g_J denotes the so-called Landé g factor and μ_B is Bohr magneton.

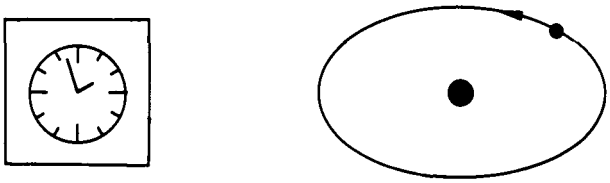


Figure 12.1. Atomic clock and gravitational clock.

The magnetic field of an atomic nucleus is extraordinarily stable. The ease with which it may be reproduced means that we can build clocks running practically synchronously at different locations and at different times. Therefore, the definition of the time unit, the second, is based on the hyperfine frequency of the ground state of the Cs isotope ^{133}Cs (caesium clock), which is given by [Ric87]

$$h\nu_{\mathrm{HFS}} \propto g_{\mathrm{Cs}} \frac{m_e}{m_{\mathrm{Cs}}} \alpha^4 m_e c^2 \tag{12.28}$$

where g_{Cs} and m_{Cs} are the g factor and the mass of the caesium nucleus, respectively.

The gravitational clock is defined by the time taken for a planet to revolve around the Sun. To explain the principle of such a measurement, we shall, for simplicity, assume a circular, rather than an elliptical orbit. According to Kepler's second law, the principle of the conservation of areas, a planet of mass m describes an area per unit time given by

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{1}{2}|\boldsymbol{R} \times \dot{\boldsymbol{R}}| = \frac{L}{2m} = \text{constant} \tag{12.29}$$

where $L = mvR$ is the orbital angular momentum. Thus, the time of revolution for the complete circular orbit is

$$T = \frac{2\pi R^2 m}{L}. \tag{12.30}$$

By virtue of the equilibrium condition for a circular orbit $GmM/R^2 = mv^2/R$ it follows that the orbital angular momentum is given by

$$L = m\sqrt{GMR}. \tag{12.31}$$

Thus, (12.30) may be written as follows

$$T = \frac{2\pi L^3}{m^3 M^2 G^2}. \tag{12.32}$$

The frequency of revolution $\nu_P = 1/T$ normalized by ν_{Cs} is given by

$$\frac{\nu_P}{\nu_{Cs}} = \frac{m^3 M^2 m_{Cs}}{m_e g_{Cs}} \frac{G^2 \hbar}{c^2 \alpha^4} \frac{1}{L^3}. \tag{12.33}$$

If the period of revolution of a planet T is repeatedly measured using an atomic clock, one obtains bounds for the temporal variation of the dimensionless combination of natural constants in (12.33), if one assumes that L = constant.

If one now assumes that the 'atomic time' is constant then the period of revolution provides information about $\dot{G}/G$. It follows from (12.32) that

$$\frac{\dot{T}}{T} = -2\frac{\dot{G}}{G} - 3\frac{\dot{m}}{m} - 2\frac{\dot{M}}{M} + 3\frac{\dot{L}}{L} \tag{12.34}$$

or, for constant masses and constant orbital angular momentum

$$\frac{\dot{G}}{G} = -\frac{1}{2}\frac{\dot{T}}{T}. \tag{12.35}$$

The precise relationship is model dependent; in general, one finds that

$$\frac{\dot{G}}{G} = f\frac{\dot{T}}{T} \tag{12.36}$$

where $f = -1/2$, if only G varies. However, in cosmological theories with a variable baryon number, it turns out that $f \neq -1/2$ [Fla81]. Either matter is continuously generated at the same rate throughout the universe (additive generation) or it is generated where the matter is already very dense (multiplicative generation) so that the mass of dense objects (planets and stars) increases with time. In Dirac theories with multiplicative mass generation, $f = 1$ and in those with additive generation, $f = -1$ [Fla81].

If one drops the restriction of circular orbits, then in a large class of theories of gravity including Newton's theory as a limiting case, one finds that the period of revolution of a two-body system is [Nor90]

$$T = \frac{2\pi l^3}{(Gm)^2} \frac{1}{(1-e^2)^{3/2}} \left[1 + O\left(\frac{G^2 m^2}{c^2 l^2}\right)\right]. \tag{12.37}$$

Here, l stands for $l = R^2\dot{\theta}$, m is a mass parameter, and e is the eccentricity ($e = \sqrt{a^2 - b^2}/a$). To a first approximation, we have

$$\frac{\dot{T}}{T} = -2\frac{\dot{G}}{G} - 2\frac{\dot{m}}{m} + 3\frac{\dot{l}}{l}. \tag{12.38}$$

We again obtain (12.35) if only the variation of G is taken into account.

Shapiro *et al* compared such a gravitational clock with a caesium atomic clock [Sha71] by determining the periods of Mercury and Venus using the radar echo procedure. They obtained

$$|\dot{G}/G| \leq 4 \times 10^{-10} \text{ year}^{-1}. \tag{12.39}$$

The orbit of the Moon can be studied in an exactly analogous way. Indeed, one experiment gave a positive result [Fla81]

$$\dot{G}/G = (-6.4 \pm 2.2) \times 10^{-11} \text{ year}^{-1}. \tag{12.40}$$

However, because of the large uncertainty associated with the tidal forces, this result is not very informative. Moreover, (12.40) contradicts more recent research (see e.g. (12.41)).

The best bounds obtained to date using this method came from measurements of the distance between the Earth and Mars in the Viking project [Hel83]. Between July 1976 and the end of July 1982, 1136 laser-beam measurements of the distance between the Earth and the Viking Lander on Mars were taken. The measurements themselves were accurate to within 2 m and the uncertainty in the calibration was around 9 m. Since the Earth's orbit must also be accurately known in the calculations several thousand additional astronomical data items relating to other objects in the Solar System which affect the movement of the Earth were also used. Assuming the constancy of all other constants, Hellings *et al* [Hel83] obtained

$$\dot{G}/G = (-0.2 \pm 0.4) \times 10^{-11} \text{ year}^{-1}. \tag{12.41}$$

The errors given in (12.41) are a consequence of the uncertainty in the asteroid masses.

Thus, Dirac's original assumption (12.5) appears to have been ruled out. The experimental uncertainties lie within the range of predictions of some superstring models [Wu86].

Although Dirac preferred a variation of G, a time dependence of other natural constants is also conceivable. In the 1950s and 1960s, the constancy of the elementary electrical charge, the fine structure constant and the Fermi coupling constant were investigated [Dys72, Nor86]. However, no evidence for a variation of these quantities was found at the level of 1×10^{-11} year^{-1}. More recent measurements in search of a time dependence of the fine structure constant, the Planck constant and the weak and strong coupling constants have also been unsuccessful (see [Ric87]).

12.3.2.2 Experiment on the fine structure constant

In the experiment described below, Turneaure and Stein [Tur76, 83], determined the relative difference in the running of two clocks, a standard caesium clock

and a superconducting cavity resonator. The frequency ν_{Cs} of the first clock is given by (12.28). The resonance frequency of a cavity resonator of length l has order of magnitude

$$\nu_R \sim \frac{c}{l}. \tag{12.42}$$

The length l is related to the interatomic distance of the resonator by

$$l \approx N a_0 \tag{12.43}$$

where $a_0 = \hbar/\alpha m_e c$ denotes the Bohr radius and N is the number of lattice planes along the resonator. It turns out that

$$\nu_R \sim \frac{\alpha m_e c^2}{N\hbar}. \tag{12.44}$$

The ratio of the two frequencies is proportional to α^3

$$R = \frac{\nu_{Cs}}{\nu_R} \propto g_{Cs}\frac{m_e}{m_{Cs}}\alpha^3. \tag{12.45}$$

The clocks remain synchronized if R is constant. Thus, comparison of the clocks may provide information about $\dot{R}/R$. Assuming that g_{Cs} and m_e/m_{Cs} are time independent, follows directly the time dependence of the fine structure constant α

$$\frac{1}{\alpha}\frac{d\alpha}{dt} = \frac{1}{3R}\frac{dR}{dt}. \tag{12.46}$$

Using $\nu_R \simeq \nu_{Cs}$, we obtain

$$\left|\frac{1}{\alpha}\frac{d\alpha}{dt}\right| \simeq \frac{1}{3\nu_R}\left|\frac{d}{dt}\langle\nu_{Cs} - \nu_R\rangle\right|. \tag{12.47}$$

No difference in the running of the clocks was found over a period of observation of 12 days; thus, the following bounds were derived for the current variation of R and α [Tur76, 83]

$$|\dot{R}/R| < 1.2 \times 10^{-11}\ \text{year}^{-1} \tag{12.48a}$$

$$|\dot{\alpha}/\alpha| < 4.1 \times 10^{-12}\ \text{year}^{-1}. \tag{12.48b}$$

12.3.2.3 On the constancy of the weak interaction

Interesting information about the constancy of G_F may be derived by comparing the lifetimes for the simple and the double-beta decay, since the half-lives depend on G_F in different ways (see [Irv86])

$$T_\beta \propto G_F^{-2} \qquad T_{\beta\beta} \propto G_F^{-4}. \tag{12.49}$$

One suitable system is the single-beta decay of ^{40}K in comparison with the double-beta decay of ^{82}Se. If it is possible to measure the lifetimes accurately to within 10%, it would be possible to achieve a sensitivity of

$$|\dot{G}_F/G_F| < 3 \times 10^{-10} \text{ year}^{-1}. \tag{12.50}$$

12.3.2.4 On the constancy of hc from a comparison of 'old' and 'young' photons

A number of measurements of the variation of $\hbar$, c and m (e.g. m_e, m_n, m_p ...) exist (see e.g. [Ric87]). Methods of modern astronomy are often used for this. One such technique is based on the following equation for photons

$$E\lambda = hc. \tag{12.51}$$

It is possible to search for variations of hc (or of h if c is constant) by measuring the energies E and wavelengths λ of 'young' and 'old' photons. Photons of

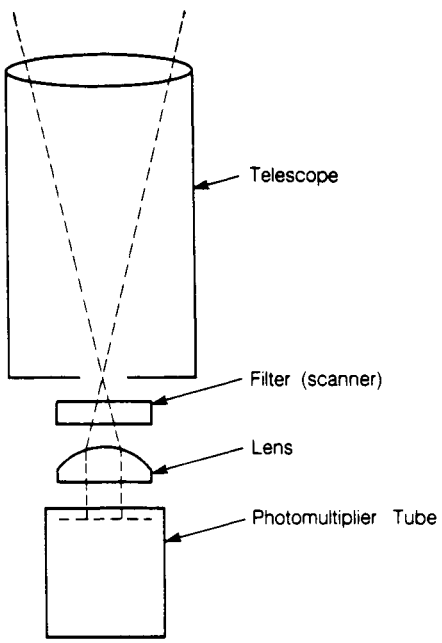


Figure 12.2. Measurement apparatus to determine the red shift of galaxies and quasars (schematic, from [Nor86]). The light from astronomical sources enters the telescope, the light wavelengths are selected by the filter and the photon energies are determined using the photomultiplier. The relationship between wavelength and energy for 'old' and 'young' photons is determined by comparing the results for photons from nearer and more distant sources.

various ages are known in astronomy. The distance covered by photons and thus the time of their emission may be determined from observations of the red shift z of galaxies and quasars using Hubble's law. Figure 12.2 shows the main components of a measurement device used for this purpose. Graphs of the Planck action quantum against the red shift z show no evidence of a dependence on z and thus on time. Accurate analysis gives

$$\frac{\dot{h}}{h} = (-3 \pm 4) \times 10^{-13} \text{ year}^{-1} \qquad \text{[Sol76]} \qquad (12.52a)$$

$$\frac{1}{hc}\frac{\mathrm{d}}{\mathrm{d}t}(hc) < 5 \times 10^{-13} \text{ year}^{-1} \qquad \text{[Bau76].} \qquad (12.52b)$$

To end this section, we note that there is also a search for creation of matter in the universe, since, according to Dirac's hypothesis, the baryon number should increase with t^2 (for an overview, see [Nor86]). This research has so far produced no evidence for a continuous synthesis of baryons.

12.3.3 Experiments relating to earlier variations

We now turn to the second class of experiments which compare the current value of a constant K with its value at a very early time in the development of the universe.

Experiments which look very far back in the past are particularly important if the quantity $\dot{K}/K$ does not follow a power law on the cosmological timescale. Thus, in this context, there arises the question as to when after the big bang the fundamental constants essentially took on their current values.

12.3.3.1 Primordial nucleosynthesis

The earliest reliable bounds are obtained from studies of the primordial nucleosynthesis, which set in around 100 s after the big bang. Primordial nucleosynthesis refers to the generation of light nuclei (H, D, He and Li) in the early phase of the evolution of the universe. The heavier elements were formed later during the development of the galaxies in the various star populations as a result of nuclear fusion up to iron and of neutron capture processes with subsequent β decays (s and r process) above this [Gro89, 90].

The mass fraction of the primordial elements, and of ^{4}He, in particular, provides information about the correctness of not only the big bang theory but also of elementary particle theories [Yan79a, Kol90]. One magnificent deduction from the primordial ^{4}He abundance was the prediction of a maximum of four light neutrino flavours long before the recent results from LEP [Yan79a, Blo84] (see also chapter 3).

The fraction Y_p of primordial ^{4}He in the total mass of the universe is heavily dependent on the ratio of protons to neutrons at the time of freezing of the reactions

$$n + e^+ \rightleftharpoons p + \overline{\nu}_e \tag{12.53a}$$

$$n + \nu_e \rightleftharpoons p + e^-. \tag{12.53b}$$

We have [Wei72]

$$Y_p \simeq \frac{2(N_n/N_p)_f}{1 + (N_n/N_p)_f} \tag{12.54}$$

where the subscript f stands for freezing out.

The effects of time-varying coupling constants may be studied using the dependence of the primordial ^{4}He generation on G, G_F and the mass difference between protons and neutrons $Q = (m_n - m_p)c^2$ [Kol86a]. According to (3.24), we have $(N_n/N_p)_f = \exp(-Q/kT_f)$. The temperature T_f is determined by equating the expansion rate of the universe $H \propto (G\rho)^{1/2}$ and the rate of the weak interaction $\Gamma_s \propto G_F^2$. Figure 12.3 illustrates the effect of G, G_F and Q on Y_p. An increasing G leads to an increase in the expansion rate and allows the weak interaction to freeze out earlier from the equilibrium. The increased T_f causes an increase in the primordial helium production. Because $\Gamma_s \propto G_F^2$, a decreasing G_F also results in a higher freeze-out temperature and thus an increase in Y_p. For a given T_f an increasing mass difference Q results in an exponential decrease in N_n/N_p and thus in the ^{4}He fraction.

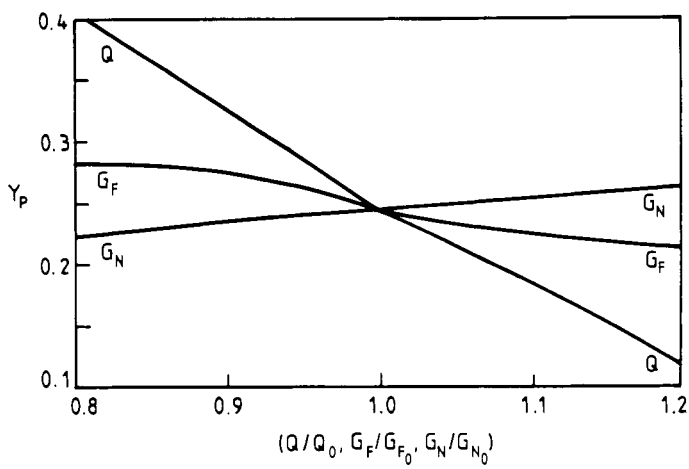


Figure 12.3. Dependence of the primordial ^{4}He fraction on the constant of gravity $G = G_N$, the Fermi constant G_F and the difference in the neutron and proton mass Q (from [Kol86a]). The index 0 denotes the present-day value.

Y_p is most sensitive to Q (see figure 12.3). For small changes in Q

$$Q = Q_0 + \Delta Q \tag{12.55}$$

where $Q_0 = 1.293$ MeV denotes the current value, ΔY is approximately linear, since for $\Delta Q \ll kT_f$

$$\begin{aligned}(N_n/N_p)_f &= \exp\left(\frac{-Q}{kT_f}\right) \\ &= \exp\left(\frac{-Q_0}{kT_f}\right)\exp\left(\frac{-\Delta Q}{kT_f}\right) \\ &= (N_n/N_p)_0\left(1 - \frac{\Delta Q}{kT_f}\right).\end{aligned} \tag{12.56}$$

A change in the proton–neutron mass difference should be associated with a change in the fine structure constant α [Kol86a]

$$\frac{Q}{Q_0} \simeq \frac{\alpha}{\alpha_0}. \tag{12.57}$$

Table 12.2 summarizes the relationships between the coupling constants and the radii of the extra dimensions R_{KK} in Kaluza–Klein and superstring theories. Thus, Y_p may be calculated as a function of R_{KK}/R_0. R_0 denotes the value of R_{KK} at the present time. Figure 12.4 contains the results for three models, one superstring model and two Kaluza–Klein models with $N = 2$ and $N = 7$, respectively. The observed mass fraction of primordial ^{4}He is $Y_p = 0.24 \pm 0.01$ [Yan84]. If we additionally require that the deuterium and helium-3 abundances should also agree with observations, a variation of α, G_F or G cannot be compensated by a change in the ratio of the baryon number to the photon number.

Table 12.2. Dependence of the fundamental coupling constants on the radius of the N extra dimensions R. The subscript $_0$ denotes the current value (from Kol86a]).

Theory	α/α_0	G/G_0	G_F/G_{F_0}
Kaluza–Klein	$(R/R_0)^{-2}$	$(R/R_0)^{-N}$	$(R/R_0)^{-2}$
Superstrings ($N = 10$)	$(R/R_0)^{-6}$	$(R/R_0)^{-6}$	$(R/R_0)^{-6}$

In the superstring model, we have $Y_p = 0.24 \pm 0.01$ for $1.005 \geq R_{KK}/R_0 \geq 0.995$ only; on the other hand this holds in Kaluza–Klein theories for $1.01 \geq R_{KK}/R_0 \geq 0.99$. This gives an upper bound for the time dependence

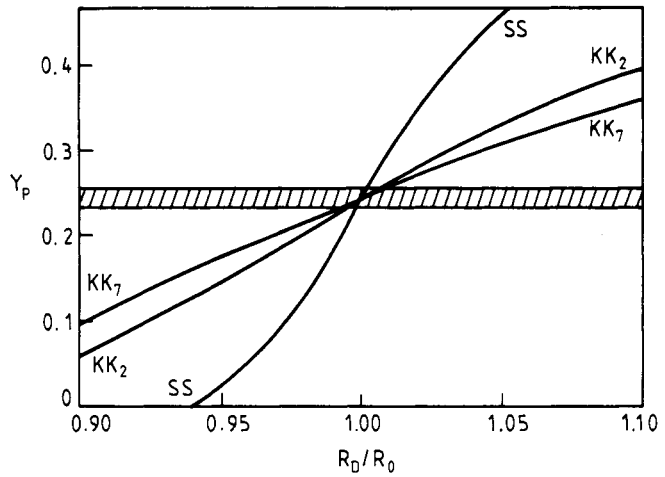


Figure 12.4. The primordial mass fraction of ^{4}He as a function of R_D/R_0 for a superstring model (SS) with $N = 6$ and two Kaluza–Klein models with $N = 2$ (KK_2) and $N = 7$ (KK_7), respectively. R_D is the radius of the extra dimensions at the time of the primordial nucleosynthesis, R_0 is the present-day value. The striped band shows the observed value of $Y_p = 0.24 \pm 0.01$ (from [Kol86a]).

of the fine structure constant of [Kol86a]

$$\left|\frac{1}{\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right| < 1.5 \times 10^{-14} h \text{ year}^{-1}, \quad h = 0.4\text{–}1. \tag{12.58}$$

While this bound is less restrictive than that which will be calculated below, it is nevertheless very important since it comes from a very early time.

Similar studies of the primordial nucleosynthesis, assuming a power law $G(t) \propto t^{-x}$, led Yang *et al* to the following bound for the Newtonian constant of gravity

$$-\left(\frac{1}{G}\frac{\mathrm{d}G}{\mathrm{d}t}\right) = \frac{x}{t_0} < 5 \times 10^{-13}\left(\frac{10^{10}}{t_0}\right) \text{ year}^{-1} \tag{12.59}$$

where t_0 is the current age of the universe ($t_0 \simeq (1\text{–}2) \times 10^{10}$ years); thus, the Dirac hypothesis appears to be ruled out.

A more recent analysis of the variation of G in the framework of the big bang nucleosynthesis using the most recent measurements of the lifetime of the neutron and the new results from LEP about the number of light neutrino flavours, for $Y_p = 0.22\text{–}0.25$ gives the following for the ratio of the gravitational constant during the nucleosynthesis G_{Nucl} to the current value G_0 [Acc90]

$$1.4 > G_{\text{Nucl}}/G_0 > 0.7. \tag{12.60}$$

For $G(t) \propto t^{-x}$, it follows that

$$\left|\frac{1}{G}\frac{dG}{dt}\right| < 9 \times 10^{-13}\ \text{year}^{-1}. \tag{12.61}$$

On the other hand, allowing for a time-dependent gravitational constant could lead to an early nucleosynthesis yielding a baryon density $\Omega = 0.1$–1 [Sta92c] (see chapter 3).

12.3.3.2 Lifetimes of radioactive nuclides (^{187}Re)

Another way of testing the time dependence of the natural constants involves comparing the lifetime of a radioactive nuclide measured in the laboratory with the lifetime calculated from isotopic abundances in geological samples [Dys67]. Since the age of such samples is determined by measuring the radioactivity (e.g. K–Ar method), strictly speaking, one measures the ratio between the lifetime of the nuclide in question and that of the nuclide used for dating, where, the half-life of the latter is averaged over the time since the sample was formed. These investigations require nuclides which satisfy the following conditions:

(i) long lifetimes, so that the oldest possible values of the constants can be determined;
(ii) the samples should be independently datable;
(iii) the decay products should not be volatile;
(iv) the lifetimes should exhibit a sensitive dependence on the coupling constants.

The rhenium isotope ^{187}Re best satisfies the conditions. ^{187}Re decays with a half-life of 4×10^{10} years via β emission to ^{187}Os. Because of the low decay energy of only 2.5 keV, the lifetime is a very sensitive function of α, since a small change in the electrostatic binding energy of Re or Os would result in a large change in the Q_β value. The electrostatic binding energy (the Coulomb term in the Weizsäcker mass formula) is of the form

$$E_C \sim \frac{Z^2}{A^{1/3}} \tag{12.62}$$

and thus depends on e^2 and α.

A sensitivity of the β decay rate λ_β to the fine structure constant may be defined as follows [Dys72]

$$s = \frac{\alpha}{\lambda_\beta}\frac{d\lambda_\beta}{d\alpha} \Rightarrow \lambda_\beta \propto \alpha^s. \tag{12.63}$$

For ^{187}Re, we have $s = 18\,000$; thus, the lifetime is proportional to $\alpha^{18\,000}$. In comparison with this strong dependence on α the lifetimes of other nuclides

may be considered to be practically independent of α, so that the latter may be used to determine the age of the sample. As a result of the much larger Q_{β}-value the sensitivity for ^{40}K ($T_{1/2} = 1.28 \times 10^9$ years, $Q_{\beta^-} = 1.31$ MeV) is only $s = 30$.

The β half-life of ^{187}Re was determined from the isotopic abundances of Re and Os in meteorites ($T_{1/2} = (4.56 \pm 0.12) \times 10^{10}$ years, [Luc83]) and in a laboratory experiment ($T_{1/2} = (4.35 \pm 0.13) \times 10^{10}$ years, [Lin86]). Comparison gives the following upper bound for the relative variation of α over the last 4.55×10^9 years [Lin86]

$$\left| \frac{1}{\alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}t} \right| < 2 \times 10^{-15} \text{ year}^{-1}. \tag{12.64}$$

12.3.3.3 Prehistoric reactor in Oklo (Gabon)

Even sharper bounds have been obtained in studies on the prehistoric reactor in Oklo (Gabon), which we shall discuss below. Shlyakhter found that the resonance energy for the capture of thermal (i.e. slow) neutrons is very sensitive to the value of the interaction coupling constant [Shl76]. Such a resonance is found, for example, in ^{149}Sm. As a result of this low-lying resonance the reaction cross section for the reaction

$$n + {}^{149}\mathrm{Sm} \rightarrow {}^{150}\mathrm{Sm}^* + \gamma \tag{12.65}$$

amounts to an unusually large 41 000 b (for comparison, the neutron capture cross section of ^{150}Sm amounts to 102 b). The energy of this resonance corresponds to a neutron kinetic energy of $T_n = 98$ meV with a width of 63 meV. The resonances of the other Sm isotopes lie in the eV region so that only neutrons from the high-energy tail of the Maxwell distribution are captured.

Because of the principle of the conservation of energy, T_n is given by the difference between the excitation energy E of ^{150}Sm* and the binding energy per nucleon $\Delta \simeq 8$ MeV

$$T_n = E - \Delta. \tag{12.66}$$

The occurrence of a resonance in the thermal region is thus associated with the fact that E and Δ are identical up to the eighth decimal place. Both energies depend on the strengths of the strong, electromagnetic and weak interactions. A change in one of the coupling constants could disturb the sensitive equilibrium between E and Δ and displace the resonance from the thermal region. The dependence on any one of the three coupling constants is given by [Ric87]

$$\frac{\mathrm{d}\Delta}{\Delta} = \beta_{\Delta} \frac{\mathrm{d}g}{g} \qquad \frac{\mathrm{d}E}{E} = \beta_E \frac{\mathrm{d}g}{g}. \tag{12.67}$$

The parameters β_Δ and β_E depend on the nuclear structure and may be calculated in the framework of the nuclear shell model.

It follows from (12.66) and (12.67) that

$$\begin{aligned}\frac{dT_n}{T_n} &= \frac{dg}{g}\frac{(\beta_E E - \beta_\Delta \Delta)}{E - \Delta} \\ &\approx \frac{dg}{g}(\beta_E - \beta_\Delta)\frac{E}{E - \Delta} \\ &\sim \frac{dg}{g}(\beta_E - \beta_\Delta) \times 10^8. \end{aligned} \tag{12.68}$$

It can be seen that a change in the coupling constants g results in a factor of 10^8 times greater variation in the resonant energy, provided we do not have $\beta_E \sim \beta_\Delta$.

Thus, we require ^{149}Sm which was exposed to a thermal neutron flux as long ago as possible. Present-day nuclear reactors have not yet been running long enough to provide sensitive measurements. In this context, the discovery of a prehistoric 'reactor' in a uranium deposit in Oklo (Gabon) in 1972 by French researchers [Okl75, Mau76, Kur82] was a piece of luck. Uranium ore from the deposit in Gabon was found to contain too little ^{235}U. Samples with an isotopic abundance of only 0.29% ^{235}U in comparison with the normal 0.7202±0.0010% for normal natural uranium were discovered. In addition, Nd and Sm isotopes produced as fission products were detected. From the analysis of the ^{235}U fraction and the isotopic abundance distribution of the fission products it was concluded that a natural chain reaction had occurred in this uranium mine some 2 billion years ago, which had possibly continued over a period of from 600 000 to 1500 000 years.

As a result of the different half-lives of the two uranium isotopes (^{238}U: 4.5×10^9 years, ^{235}U: 7×10^8 years) the ^{235}U content then amounted to 3–4%; thus a sufficient amount of material which could be subject to fission by thermal neutrons was present[4]. In the case of this natural reactor, the infiltration of water into the deposit may have given rise to conditions corresponding to those in a normal modern light water reactor. Water is used as a moderator to decelerate fast fission neutrons. Around 500 tonnes of uranium were involved in the chain reaction and the energy released amounted to approximately 10^{11} kWh. The overall neutron flux in several places exceeded 1.5×10^{21} neutrons/cm^2.

Table 12.3 gives the abundances of the Nd isotopes found in two different ore samples in the Oklo uranium mine (after [Neu72]). The Oklo M sample contained 0.4400 ± 0.0005% ^{235}U and the Oklo 310 sample 0.592 ± 0.001%, while the natural isotopic abundance contains 0.7202% ^{235}U. It is obvious that the Nd abundances differ from the ratios usually found, but are very similar

[4] Modern light water reactors operate with an enrichment of 3–3.5%.

Table 12.3. Abundances of the Nd isotopes in two different ore samples from the Oklo uranium mine (after the analysis of [Neu72]). The values marked by $^{(*)}$ are corrected to the fraction of the corresponding isotope which occurs naturally.

	Oklo M	Oklo 310	Natural	Oklo M$^{(*)}$	Oklo 310$^{(*)}$	Fission (^{235}U)
^{142}Nd	1.38	5.49	27.11	0	0	0
^{143}Nd	22.1	23.0	12.17	22.6	25.7	28.8
^{144}Nd	32.0	28.2	23.85	32.4	29.3	26.5
^{145}Nd	17.5	16.3	8.30	18.05	18.4	18.9
^{146}Nd	15.6	15.4	17.22	15.55	14.9	14.4
^{148}Nd	8.01	7.70	5.73	8.13	8.20	8.26
^{150}Nd	3.40	3.90	5.62	3.28	3.46	3.12

to the isotopic distribution generated by fission of ^{235}U (see the last column of table 12.3). Since the isotope ^{142}Nd is not a fission product, its abundance may be used as a measure of the natural Nd content in the Oklo M and Oklo 310 samples. The values corrected to this natural Nd content are shown by a star superscript. They agree well with the yields from neutron-induced fission of uranium, which indicates that a chain reaction did indeed take place.

A corresponding evaluation was carried out for samarium. Neuilly *et al* found that the ratio of ^{149}Sm to ^{147}Sm in the Oklo M sample was extremely small (~ 0.003). In the natural isotopic mixture this ratio is 0.924, a value of 0.475 is obtained in the fission of uranium. The cause of the ^{149}Sm deficit lies in the large cross section for the reaction

$$^{149}\mathrm{Sm}(n,\gamma)\,^{150}\mathrm{Sm} \tag{12.69}$$

which means that most ^{149}Sm nuclei are converted into ^{150}Sm nuclei. In comparison with that the cross section for the reaction

$$^{147}\mathrm{Sm}(n,\gamma)\,^{148}\mathrm{Sm} \tag{12.70}$$

is very small ($\simeq 60$ b). Some of the ^{147}Sm nuclei were converted into ^{148}Sm nuclei but the abundance in comparison with ^{149}Sm was only modified slightly by neutron capture. Strictly speaking, according to these considerations, one should compare the distributions of ^{147}Sm + ^{148}Sm and ^{149}Sm + ^{150}Sm. In their analysis Dozol and Neuilly [Doz75] found the results shown in table 12.4, corrected to the natural Sm content. The last column gives the values found in the fission of uranium. These data also point to the existence of a prehistoric reactor. The slight excess of ^{152}Sm is a result of neutron capture by ^{151}Eu which decays into ^{152}Sm.

Table 12.4. Abundances of the Sm isotopes in [%] averaged over various samples from Oklo [Doz75].

Sm isotope	Oklo	Natural	Fission (^{235}U)
147 + 148	61.22	26.21	61.56
149 + 150	27.33	21.27	29.3
152	10.51	26.72	7.23
154	1.77	22.71	1.94

The fact that only very little ^{149}Sm was found is explained by the energetically low-lying resonance for neutron capture which leads to the large reaction cross section for n capture by ^{149}Sm. This resonance must thus already have been located around 2 billion years ago in the thermal region. This implies a relative change in T_n of less than 10%.

Shlyakhter [Shl76] used this to derive the following bounds for the time dependence of the coupling constants of the strong, electromagnetic and weak interactions

$$\left|\frac{1}{g_s}\frac{\mathrm{d}g_s}{\mathrm{d}t}\right| < 5 \times 10^{-19}\ \text{year}^{-1} \tag{12.71a}$$

$$\left|\frac{1}{\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right| < 1 \times 10^{-17}\ \text{year}^{-1} \tag{12.71b}$$

$$\left|\frac{1}{g_w}\frac{\mathrm{d}g_w}{\mathrm{d}t}\right| < 2 \times 10^{-12}\ \text{year}^{-1}. \tag{12.71c}$$

12.3.3.4 Fine and hyperfine structure lines of remote galaxies

In further studies, the relative displacements of fine-structure and hyperfine-structure lines of remote galaxies were investigated. Since the light was emitted a long time ago this determines the value of α a long time ago [Bah67, Tub80].

12.3.4 Closing remarks and outlook

Table 12.1 summarizes the most important results as far as the constancy of natural constants is concerned. These data rule out the Dirac hypothesis. We note, however, that the bounds are in some cases only valid under the assumption that all other constants are time-invariant.

It has not yet been possible to detect a time dependence of natural constants. However, this does not contradict the multidimensional Kaluza–Klein theories in which time dependence is allowed but is not mandatory.

However, it may also be the case that the search has until now been considering the wrong timescale (cf. [Kol86a]). In general, it is assumed that the constants vary with a power of the cosmological time H^{-1}. However, it is also conceivable that the compactification of the extra dimensions ended very quickly and the radii today only oscillate around their equilibrium positions. The relevant timescale would in that case be defined by the Planck time $t_{\mathrm{Pl}} = 5.39 \times 10^{-44}$ s instead, so that the observed values would only represent average values over large numbers of oscillations.

A finite variation of the natural constants would have extensive consequences, as indicated in section 12.1. One of the most serious consequences would be that the laws of nature would no longer be invariant under time displacement. This inhomogeneity of time would result in a violation of the principle of the conservation of energy (see section 12.1).

Bibliography

[Aar87] Aardsma G *et al* 1987 *Phys. Lett.* **194B** 321

[Aba88] Abazov A I *et al* 1988 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 317

[Aba91] Abazov A I *et al* (SAGE Collaboration) 1991 *Phys. Rev. Lett.* **67** 3332

[Aba95] Abachi S *et al* 1995 *Phys. Rev. Lett.* **74** 2422

[Abb83] Abbott L and Sikivie P 1983 *Phys. Lett.* **120B** 133

[Abb88] Abbott L 1988 *Sci. Am.* **258**(5) 82

[Abd90] Abdullah K *et al* 1990 *Phys. Rev. Lett.* **65** 2347

[Abd95] Abdurashitov J N *et al* (SAGE Collaboration) 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 60

[Abe84] Abela R *et al* 1984 *Phys. Lett.* **146B** 431

[Abe90] Abe F *et al* 1990 *Phys. Rev. Lett.* **65** 2243

[Abe93] Abele H *et al* 1993 *Phys. Lett.* **316B** 26

[Abe94] Abe F *et al Preprint* FERMILAB-PUB94/116-E

[Abr86] Abramowicz H *et al* 1986 *Phys. Rev. Lett.* **57** 298

[Abr90] Abreu P *et al* 1990 *Phys. Lett.* **247B** 167; 1990 *Phys. Lett.* **252B** 149

[Acc90] Accetta F S, Krauss L M and Romanelli P 1990 *Phys. Lett.* **248B** 146

[Ack91] Acker A, Pakvasa S and Pantaleone J 1991 *Phys. Rev.* D **43** 1754

[Ada93] Adams F C *et al* 1993 *Phys. Rev. Lett.* **70** 2511

[Ade87] Adelberger E G *et al* 1987 *Phys. Rev. Lett.* **59** 849

[Ade90] Adelberger E G *et al* 1990 *Phys. Rev.* D **42** 3267

[Ade91a] Adelberger E G, Heckel B R, Stubbs C W and Su Y 1991 *Phys. Rev. Lett.* **66** 850

[Ade91b] Adelberger E 1991 *Nucl. Phys.* A **527** 223c

[Ade91c] Adeva B *et al* 1991 *Phys. Lett.* **257B** 450

[Afo83] Afonin A I *et al* 1983 *JETP Lett.* **38** 463

[Afo85] Afonin A I *et al* 1985 *JETP Lett.* **42** 285

[Agl87] Aglietta M *et al* 1987 *Europhys. Lett.* **3** 1315, 1321

[Agl89] Aglietta M A *et al* 1989 *Europhys. Lett.* **8** 611

[Agl91] Aglietta M A *et al* 1991 *Astrophys. J.* **382** 344

[Agl93] Aglietta M *et al* 1992 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 450

[Ahl83] Ahlen S P 1983 *Proc. Magnetic Monopoles 1982* (New York: Plenum) p 259
[Ahl84] Ahlen S P, Price P B, Guo S and Fleischer R L 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum) p 383
[Ahl90] Ahlen S P *et al* (MACRO Collaboration) 1990 *Phys. Lett.* **249B** 149
[Ahl92] Ahlen S P *et al* (MACRO Collaboration) 1992 *Phys. Rev. Lett.* **69** 1860
[Ahl93] Ahlen S P *et al* (MACRO Collaboration) 1993 *Nucl. Instrum. Methods* A **324** 337
[Ahl94] Ahlen S P *et al* (MACRO Collaboration) 1994 *Phys. Rev. Lett.* **72** 608
[Ahm94] Ahmed J *et al* (H1 Collaboration) 1994 *Preprint* DESY 94-154 and *Z. Phys.* C **64** 545
[Ahr85] Ahrens L *et al* 1985 *Phys. Rev.* D **31** 2732
[Ahr87] Ahrens L A *et al* 1987 (Singapore: World Scientific) p 99
[Air56] Airy G B 1856 *Phil. Trans. R. Soc., London*, **146** 297, 343
[Ait89] Aitchison I J R and Hey A J G 1989 *Gauge Theories in Particle Physics* (Bristol: Hilger)
[Akh88] Akhmedov E Kh and Khlopov M Yu 1988 *Mod. Phys. Lett.* A **3** 451
[Akr91] Akrawy M Z *et al* 1991 *Phys. Lett.* **253B** 511
[Aku72] Akulov D V and Volkov V P 1972 *JETP Lett.* **16** 438
[Alb82] Albrecht A and Steinhardt P J 1982 *Phys. Rev. Lett.* **48** 1220
[Alb85] Albrecht A and Turok N 1985 *Phys. Rev. Lett.* **54** 1868
[Alb88] Albrecht H *et al* 1988 *Phys. Lett.* **202B** 149
[Alb90] Albrecht H *et al* 1990 *Phys. Lett.* **234B** 409
[Alb91] Alberico W M, de Pace A and Pignone M 1991 *Nucl. Phys.* A **523** 488
[Alb92] Albrecht H *et al* 1992 *Phys. Lett.* **292B** 221
[Ale82] Alexeyev E N *et al* 1982 *Lett. Nuovo Cimento* **35** 413; 1982 *Proc. 17th Int. Cosmic Ray Conf.* vol 10 p 83
[Ale85] Alexeyev E N, Boliev M M, Chudakov A E and Mikheyev S P 1985 *Proc. 19th Int. Cosmic Ray Conf. (La Jolla)* vol 8 p 250
[Ale87] Alexeyev E N, Alexeyeva L N, Chudakov A E and Krivosheina I V 1987 *Proc. Int. Symp. Underground Physics (Baksan, USSR)* ed G V Domogatsky *et al* p 30
[Ale88] Alexeyev E N, Alexeyeva L N, Krivosheina I V and Volchenko V I 1988 *Phys. Lett.* **205B** 209
[Ale92] Alessandrello A *et al* 1992 *Phys. Lett.* **285B** 176
[Ale94] Alessandrello A *et al* 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 366 (Amsterdam: North-Holland); 1994 *Phys. Lett.* **335B** 519
[ALE89] ALEPH Collaboration 1989 *Phys. Lett.* **231B** 519
[ALE92] ALEPH Collaboration 1992 *Phys. Rep.* **216** 253
[Ali90] Alitti J *et al* 1990 *Phys. Lett.* **241B** 150
[All83] Alley C O 1983 *Quantum Optics, Experimental Gravity and Measurement Theory* ed P Meystre and M O Scully (New York: Plenum) p 429
[All86] Allaby J V *et al* 1986 *Phys. Lett.* **177B** 446

[All88] Allison W W M 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 50
[All90] Allen B and Shellard E P S 1990 *Phys. Rev. Lett.* **64** 119
[Als88] Alston-Garnjost M 1988 *Phys. Rev. Lett.* **60** 1928
[Als89] Alston-Garnjost M *et al* 1989 *Phys. Rev. Lett.* **63** 1671
[Als93] Alston-Garnjost M *et al* 1993 *Phys. Rev. Lett.* **71** 831
[Alt81] Altarev I S *et al* 1981 *Phys. Lett.* **102B** 13
[Alt85] Altzitzoglou T *et al* 1985 *Phys. Rev. Lett.* **55** 799
[Alt86] Altarev I S *et al* 1986 *JETP Lett.* **44** 460
[Alt90] Altarelli G 1990 *Proc. NEUTRINO'90 (Geneva) Nucl. Phys.* B (Proc. Suppl.) **19** 354 (Amsterdam: North-Holland)
[Alv63] Alvarez L W 1963 *LRL Physics Note* **470**
[Alv71] Alvarez L W *et al* 1971 *Rev. Sci. Instrum.* **42** 326
[Ama87] Amaldi U *et al* 1987 *Phys. Rev.* D **36** 1385
[Ama91] Amaldi U, de Boer W and Fürstenau H 1991 *Phys. Lett.* **260B** 447
[And32] Anderson C D 1932 *Science* **76** 238
[And89] Ander M E *et al* 1989 *Phys. Rev. Lett.* **62** 985
[Ang86] Angelini C *et al* 1986 *Phys. Lett.* **179B** 307
[Ano92] Anosov O L *et al* 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 111 (Amsterdam: North-Holland)
[Ans92] Anselmann P *et al* (GALLEX Collaboration) 1992 *Phys. Lett.* **285B** 376, 390
[Ans93] Anselmann P *et al* (GALLEX Collaboration) 1993 *Phys. Lett.* **314B** 445
[Ans94] Anselmann P *et al* (GALLEX Collaboration) 1994 *Phys. Lett.* **327B** 377
[Ant60] dell'Antonio G F and Fiorini E 1960 *Nuovo Cimento* Suppl. **17** 132
[Apa85] Apalikov A M *et al* 1985 *JETP Lett.* **42** 289
[Ara80] Aragone C and Deser S 1980 *Phys. Rev.* D **21** 352
[Arn83] Arnison G *et al* 1983 *Phys. Lett.* **122B** 103; 1983 **129B** 273; 1983 **126B** 398
[Arn87] Arnett W D and Rosner J 1987 *Phys. Rev. Lett.* **58** 1906
[Arn95] Arnold R *et al* (NEMO Collaboration) 1995 *Pis'ma Zh. Eksp. Teor. Fiz.* **61** 168
[Arp94] Arpesella C, Bellotti E and Bottino A (ed) 1994 *Proc. TAUP'93, Italy, Nucl. Phys.* B (Proc. Suppl.) **35** (Amsterdam: North-Holland)
[Art95] Artemiev V *et al* 1995 *Phys. Lett.* **345B** 564
[Asr81] Asratyan A E *et al* 1981 *Phys. Lett.* **105B** 301
[Ast89] Astier P *et al* 1989 *Phys. Lett.* **220B** 646
[Aul83] Aulakh C S and Mohapatra R N 1983 *Phys. Lett.* **119B** 136
[Avi87] Avignone F T *et al* 1987 *Phys. Rev.* D **35** 1713
[Avi88] Avignone F T and Brodzinski R L 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 147
[Avi89] Avignone F T *et al* 1989 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 66
[Avi90] Avignone F T *et al* 1990 *Phys. Rev. Lett.* **65** 3092
[Avi91] Avignone F T *et al* 1991 *Phys. Lett.* **256B** 559
[Bab91a] Babu K S, Mohapatra R N and Rothstein I Z 1991 *Phys. Rev.* D **44** 2265

[Bab91b] Babu K S, Mohapatra R N and Rothstein I Z 1991 *Phys. Rev. Lett.* **67** 545
[Bac91] Backe H *et al* 1991 *DPG-Tagung Physik der Hadronen und Kerne (Darmstadt)*
[Bac93] Backe H *et al* 1993 *Proc. NEUTRINO'92 (Granada), Nucl. Phys.* B (Proc. Suppl.) **31** 46 (Amsterdam: North-Holland)
[Bag83] Bagnaia P *et al* 1983 *Phys. Lett.* **129B** 130
[Bah67] Bahcall J N and Schmidt M 1967 *Phys. Rev. Lett.* **19** 1294
[Bah69] Bahcall J N and Frautschi S C 1969 *Phys. Lett.* **29B** 623
[Bah71] Bahcall J N and Ulrich R K 1971 *Astrophys. J.* **170** 593
[Bah78] Bahcall J N and Primakoff H 1978 *Phys. Rev.* D **18** 3463
[Bah82a] Bahcall J N *et al* 1982 *Rev. Mod. Phys.* **54** 767
[Bah82b] Bahcall J N and Davis R 1982 *Essays in Nuclear Astrophysics* (Cambridge: Cambridge University Press) p 242
[Bah87] Bahcall J N and Glashow S L 1987 *Nature* **326** 476
[Bah88] Bahcall J N and Ulrich R K 1988 *Rev. Mod. Phys.* **60** 297
[Bah89] Bahcall J N 1989 *Neutrino Astrophysics* (Cambridge: Cambridge University Press)
[Bah90] Bahcall J N and Bethe H A 1990 *Phys. Rev. Lett.* **65** 2233
[Bah92] Bahcall J N and Pinsonneault M H 1992 *Rev. Mod. Phys.* **64** 885
[Bah93] Bahcall J N and Kumar P 1993 *Astrophys. J.* **409** L73
[Bak84] Baker N J *et al* 1984 *Phys. Rev.* D **28** 2705
[Bal87] Baltz A J and Weneser J 1987 *Phys. Rev.* D **35** 528
[Bal88] Baltz A J and Weneser J 1988 *Phys. Rev.* D **37** 3364
[Bal90] Baldo-Ceolin M *et al* 1990 *Phys. Lett.* **236B** 95
[Bal91] Baltz A J and Weneser J 1991 *Phys. Rev. Lett.* **66** 520
[Bal92] Balysh A *et al* (Heidelberg–Moscow Collaboration) 1992 *Phys. Lett.* **283B** 32
[Bal93] Balysh A *et al* (Heidelberg–Moscow Collaboration) 1993 *Phys. Lett.* **298B** 278
[Bal94] Balysh A *et al* (Heidelberg–Moscow Collaboration) 1994 *Phys. Lett.* **322B** 176
[Bal94a] Balysh A *et al* (Heidelberg–Moscow Collaboration) 1994 *Proc. 27th Int. Conf. on High Energy Physics (Glasgow)* ed P J Bussey and I G Knowles (Bristol: Institute of Physics Publishing) p 939
[Bal95] Balysh A *et al* (Heidelberg–Moscow Collaboration) 1995 *Phys. Lett.* **356B** 450
[Bald92] Baldo-Ceolin M (ed) 1992 *Proc. 4th Int. Workshop on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin (Padova: Padova University)
[Bald94] Baldo-Ceolin M *et al* 1994 *Z. Phys.* C **63** 409
[Bam95] Bamert P, Burgess C P and Mohapatra R N 1995 *Nucl. Phys.* **449B** 25
[Bam95a] Bamert P, Burgess C P and Mohapatra R N 1995 *Nucl. Phys.* **438B** 3
[Ban83] Banner M *et al* 1983 *Phys. Lett.* **122B** 476
[Bar70] Bardin R K, Gollon P J, Ullman J D and Wu C S 1970 *Nucl. Phys.* A **158** 337
[Bar81] Barbieri R and Mohapatra R N 1981 *Phys. Lett.* **105B** 369
[Bar82] Barger V, Baer H, Keung W Y and Phillips R J N 1982 *Phys. Rev.* D **26** 218

[Bar83a] Barwick S W, Kinoshita K and Price P B 1983 *Phys. Rev.* D **28** 2338
[Bar83b] Barr S M, Reiss D B and Zee A 1983 *Phys. Rev. Lett.* **50** 317
[Bar85] Barabanov I R *et al* 1985 *Solar Neutrinos and Neutrino Astronomy (AIP Conf. Proc. 126)* (New York: AIP) p 175
[Bar86a] Bars I and Visser M 1986 *Phys. Rev. Lett.* **57** 25
[Bar86b] Barr S M and Mohapatra R N 1986 *Phys. Rev. Lett.* **57** 3129
[Bar86c] Barabanov I *et al* 1986 *Pis'ma Zh. Eksp. Teor. Fiz.* **43** 116
[Bar87a] Barrow J D 1987 *Phys. Rev.* D **35** 1805
[Bar87b] Barroso A, Chalabi R, Dombey N and Hamzaoui C 1987 *Phys. Lett.* **196B** 369
[Bar87c] Barabash A S 1987 *Preprint* ITEP 56-1987
[Bar87d] Barger V 1987 *Proc. VIIIth Summer School in Nuclear and Particle Physics (Tasmania)* ed R Delbourgo and J R Fox (Singapore: World Scientific) p 103
[Bar88a] Bari C *et al* 1988 *Nucl. Instrum. Methods* A **264** 5
[Bar88b] Barbieri R and Mohapatra R N 1988 *Phys. Rev. Lett.* **61** 27
[Bar88c] Barger V and Whisnant K 1988 *Phys. Lett.* **209B** 365
[Bar89] Barger V *et al* 1989 *Phys. Rev.* D **40** 2987
[Bar89a] Barr G, Gaisser T K and Stanev T 1989 *Phys. Rev.* D **39** 3532
[Bar89b] Bartlett D F and Tew W L 1989 *Phys. Rev.* D **40** 673
[Bar89c] Barabash A S *et al* 1989 *Phys. Lett.* **223B** 273
[Bar89d] Barabash A S 1989 *Phys. Lett.* **216B** 257
[Bar89e] Barabash A S 1989 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'89 (Montreal)* ed P Depommier (Gif-sur-Yvette: Frontières) p 249
[Bar90a] Barger V, Phillips R J N and Whisnant K 1990 *Phys. Rev. Lett.* **65** 3084
[Bar90b] Bartlett D F and Tew W L 1990 *J. Geophys. Res.* **95** 17363
[Bar90c] Barabash A S *et al* 1990 *Sov. J. Nucl. Phys.* **51** 1
[Bar90d] Barabash A S, Egorov O K, Klimenko A A, Kolganova E D, Pozharova E A, Skorodko T Yu, Smirnitsky V A and Smolnikov A A 1990 *Preprint* ITEP 90-131
[Bar92] Barry C 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 437 (Amsterdam: North-Holland)
[Bat82] Battistoni G *et al* 1982 *Phys. Lett.* **118B** 461
[Bat83] Battistoni G *et al* 1983 *Phys. Lett.* **133B** 454
[Bat88] Battistoni G *et al* 1988 *Nucl. Instrum. Methods* A **270** 185
[Bau76] Baum W A and Florentin-Nielsen R 1976 *Astrophys. J.* **209** 319
[Bau86] Baumann N *et al* 1986 *Proc. 12th Int. Conf. on Neutrino Physics and Astrophysics (Sendai)* ed T Kitagaki and H Yuta (Singapore: World Scientific)
[Bau97] Baudis *et al* (Heidelberg–Moscow Collaboration) 1997 *Phys. Lett.* **407B** 219, *Phys. Rev.* D **55** 54
[Baz84] Bazilevskaja G A *et al* 1984 *Sov. J. Nucl. Phys.* **39** 543
[Bea72] Beauchamp W T, Bowen T, Cox A J and Kalbach R M 1972 *Phys. Rev.* D **6** 1211

[Bec81] Becher P, Böhm M and Joos H 1981 *Eichtheorien der starken und elektroschwachen Wechselwirkung* (Stuttgart: Teubner)
[Bec91a] Becker H *et al* 1991 *Proc. XIth Moriond Workshop on Tests of Fundamental Laws in Physics (Les Arcs)*
[Bec91b] Beck M *et al* (Heidelberg–Moscow Collaboration) 1991 *Proc. XIth Moriond Workshop on Tests of Fundamental Laws in Physics (Les Arcs)*
[Bec92a] Beck M *et al* (Heidelberg–Moscow Collaboration) 1992 *Z. Phys.* A **343** 397
[Bec92b] Becker-Szendy R *et al* 1992 *Phys. Rev. Lett.* **69** 1010
[Bec93] Beck M *et al* (Heidelberg–Moscow Collaboration) 1993 *Phys. Rev. Lett.* **70** 2853
[Bec94] Beck M *et al* (Heidelberg–Moscow Collaboration) 1994 *Phys. Lett.* **336B** 141
[Bec95] Becker-Szendy R *et al* (IMB Collaboration) 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 331
[Bed94a] Bednyakov A, Klapdor-Kleingrothaus H V and Kovalenko S 1994 *Phys. Lett.* **329B** 5
[Bed94b] Bednyakov A, Klapdor-Kleingrothaus H V and Kovalenko S 1994 *Phys. Rev.* D **50** 7128
[Bei91] Beier E W, Davis jr R , Kim S B, Elliott S R and Jelley N 1991 *Nucl. Phys.* A **527** 653c
[Bel83] Bellotti E *et al* 1983 *Phys. Lett.* **121B** 72
[Bel88] Bellotti E 1988 *Nucl. Instrum. Methods* A **264** 1
[Bel89] Bellotti E *et al* 1989 *Phys. Lett.* **221B** 209
[Bel90] Bellotti E *et al* 1990 *Proc. Rencontres de la Vallée d'Aoste (La Thuile)*
[Bel91a] Bellotti E 1991 *Nucl. Phys. News* **1** 18
[Bel91b] Bellotti E *et al* 1991 *Phys. Lett.* **266B** 193
[Ben89] Bennett jr W R 1989 *Phys. Rev. Lett.* **62** 365
[Ben91] Bento L and Valle J W F 1991 *Phys. Lett.* **264B** 373
[Ben92a] Bennett C L *et al* 1992 *Astrophys. J.* **396** L7
[Ben92b] Benetti P *et al* 1992 *Proc. 4th Int. Workshop on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin p 279
[Ben93] Benetti P *et al* 1993 *Nucl. Instrum. Methods* A **327** 173
[Ben94] Benetti P *et al* 1994 *Nucl. Phys.* B (Proc. Suppl.) **35** 276
[Ber72] Bergkvist K E 1972 *Nucl. Phys.* B **39** 317
[Ber79] Berends F A, van Holten J W, de Wit B and Nieuwenhuizen P 1979 *Nucl. Phys.* B **154** 261; 1979 *Phys. Lett.* **83B** 188
[Ber85a] Berezinsky V S, Castagnoli C, Ryazhskaya O G and Saavedra O 1985 *Nucl. Phys.* B **262** 383
[Ber85b] Bermon S *et al* 1985 *Phys. Rev. Lett.* **55** 1850
[Ber85c] Bergkvist K E 1985 *Phys. Lett.* **159B** 408
[Ber85d] Bernstein M D and Brown L S 1985 *Comm. Nucl. Part. Phys.* **15** 35
[Ber86] Bernardi G *et al* 1986 *Phys. Lett.* **166B** 479; 1986 *Phys. Lett.* **181B** 173
[Ber89] Berger Ch *et al* 1989 *Nucl. Phys.* B **313** 509
[Ber90a] Berger Ch *et al* 1990 *Phys. Lett.* **240B** 237
[Ber90b] Berger Ch *et al* 1990 *Phys. Lett.* **245B** 305
[Ber90c] Bermon S *et al* 1990 *Phys. Rev. Lett.* **64** 839

[Ber90d] Bertani M *et al* 1990 *Europhys. Lett.* **12** 613
[Ber91a] Bertolini S and Santamaria A 1991 *Nucl. Phys.* B **357** 222
[Ber91b] Berger Ch *et al* 1991 *Z. Phys.* C **50** 385
[Ber92a] Bernatowicz T *et al* 1992 *Phys. Rev. Lett.* **69** 2341
[Ber92b] Berezhiani Z G, Smirnov A Y and Valle J W F 1992 *Phys. Lett.* **291B** 99
[Ber93] Berthomieu G, Provost J, Morch P and Lebreton Y 1993 *Astron. Astrophys.* **268** 775
[Ber93a] Bernatowicz T *et al* 1993 *Phys. Rev.* C **47** 806
[Bet86a] Bethge K and Schröder U E 1986 *Elementarteilchen und ihre Wechselwirkungen* (Darmstadt: Wissenschaftliche Buchgesellschaft)
[Bet86b] Bethe H A 1986 *Phys. Rev. Lett.* **56** 1305
[Bib87] Van Bibber K *et al* 1987 *Phys. Rev. Lett.* **59** 759
[Bil87] Bilenky S M and Petcov S T 1987 *Rev. Mod. Phys.* **59** 671
[Bil91] Bilenky S M, Masiero A and Petcov S T 1991 *Phys. Lett.* **263B** 448
[Bio87] Bionta R M *et al* 1987 *Phys. Rev. Lett.* **58** 1494
[Bio88] Bionta R M *et al* 1988 *Phys. Rev.* D **38** 768
[Bis76] Bishop N T and Landsberg P T 1976 *Nature* **264** 346
[Bit85] Bitter T and Dubbers D 1985 *Nucl. Instrum. Methods* A **239** 461
[Biz89] Bizzeti P G *et al* 1989 *Phys. Rev. Lett.* **62** 2901
[Bjo50] Bjorklund R *et al* 1950 *Phys. Rev.* **77** 213
[Bjo78] Bjorken J P and Drell S D 1978 *Relativistische Quantenfeldtheorie* (Mannheim: BI)
[Bla77] Bland R, Bocobo D, Eubank M and Royer J 1977 *Phys. Rev. Lett.* **39** 369
[Blo84] Blome H J and Priester W 1984 *Naturwissenschaften* **71** 456, 515, 528
[Blu92] Bludman S A, Kennedy D C and Langacker P G 1992 *Phys. Rev.* D **45** 1810
[Bob91] Bobbink G and Ratoff P 1991 *XXVIth Rencontres de Moriond*
[Boe85] Boesgaard A M and Steigman G 1985 *Ann. Rev. Astron. Astrophys.* **23** 319
[Boe87] Boehm F and Vogel P 1987 *Physics of Massive Neutrinos* (Cambridge: Cambridge University Press)
[Boe92] Boehm F and Vogel P 1992 *Physics of Massive Neutrinos* 2nd edn (Cambridge: Cambridge University Press)
[Boh75] Bohr A and Mottelson B R 1975 *Struktur der Atomkerne* vol 1 (Berlin: Akademie)
[Boo92] Booth N E and Salmon G L (ed) 1992 *Proc. Low Temperature Detectors for Neutrinos and Dark Matter IV, 1991* (Gif-sur-Yvette: Frontières)
[Bor85] Boris S *et al* 1985 *Phys. Lett.* **159B** 217
[Bor87] Boris S *et al* 1987 *Phys. Rev. Lett.* **58** 2019
[Bör88] Börner G 1988 *The Early Universe – Facts and Fiction, Texts and Monographs in Physics* (Berlin: Springer)
[Bor91] Bellini G, Campanella M, Giugni D and Raghavan R (ed) 1991 *Borexino experiment proposal*
[Bor92] Borodovsky L *et al* 1992 *Phys. Rev. Lett.* **68** 274
[Bot91] Bottino A 1991 *Proc. 3rd Int. Workshop on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin p 41
[Bot94] Bottino A *et al* 1994 *Astropart. Phys.* **2** 67, 77

[Bou86] Bouchez J *et al* 1986 *Z. Phys.* C **32** 499
[Bou88] Bourdarios C 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 42
[Bou89] Bouquet A, Kaplan J and Martin F 1989 *Astron. Astrophys.* **222** 103
[Boy85] Boyd R N 1985 *Solar Neutrinos and Neutrino Astronomy (AIP Conf. Proc. 126)* (New York: AIP) p 145
[Boy87] Boynton P E, Croby D, Ekstrom P and Szumilo A 1987 *Phys. Rev. Lett.* **59** 1385
[Boy90] Boynton P and Aronson S 1990 *Proc. XXVth Rencontres de Moriond, New and Exotic Phenomena* ed O Fackler and J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 207
[Bra59] Bradner H and Isbell W M 1959 *Phys. Rev.* **114** 603
[Bra72] Braginsky V B and Panov V I 1972 *Sov. Phys.–JETP* **34** 463
[Bra84] Bracci L *et al* 1984 *Phys. Lett.* **143B** 357; 1985 **155B** 468 Erratum
[Bra88] Braunschweig W *et al* 1988 *Z. Phys.* C **38** 543
[Bre89] Bressi G *et al* 1989 *Z. Phys.* C **43** 175
[Bru85] Bruenn S W 1985 *Astrophys. J.* Suppl. **58** 771
[Bru87] Bruenn S W 1987 *Phys. Rev. Lett.* **59** 938
[Buc90] Buckland K N *et al* 1990 *Phys. Rev.* D **41** 2726
[Buc92] Buchmüller W and Ingelman G 1991 *Proc. Workshop Physics at HERA (Hamburg)*
[Bug89] Bugaev E V and Naumov V A 1989 *Phys. Lett.* **232B** 391
[Bur57] Burbidge E M, Burbidge G R, Fowler W A and Hoyle F 1957 *Rev. Mod. Phys.* **29** 547
[Bur88] Burkhardt H *et al* 1988 *Phys. Lett.* **206B** 169
[Bur94] Burgess C P and Cline J M 1994 *Phys. Rev.* D **49** 5925
[Bus90] Busto J *et al* 1990 *Nucl. Phys.* A **513** 291
[But93] Butterworth J and Dreiner H 1993 *Nucl. Phys.* B **397** 3
[Cab82] Cabrera B 1982 *Phys. Rev. Lett.* **48** 1378
[Cab83] Cabrera B *et al* 1983 *Phys. Rev. Lett.* **51** 1933
[Cal76] Callan C, Dashen R and Gross D 1976 *Phys. Lett.* **63B** 334
[Cal82] Callan C G 1982 *Phys. Rev.* D **25** 2141; 1982 *Phys. Rev.* D **26** 2058
[Cal87] Caldwell D O *et al* 1987 *Phys. Rev. Lett.* **59** 419
[Cal88a] Caldwell D O *et al* 1988 *Phys. Rev. Lett.* **61** 510
[Cal88b] Calicchio M *et al* 1988 *Nucl. Instrum. Methods* A **264** 18
[Cal89] Caldwell D O 1989 *Int. J. Mod. Phys.* A **4** 1851
[Cal90a] Caldwell D O *et al* 1990 *Phys. Rev. Lett.* **65** 1305
[Cal90b] Caldwell D O *et al* 1990 *Nucl. Phys.* B (Proc. Suppl.) **13** 547
[Cal91a] Caldwell D O and Langacker P 1991 *Phys. Rev.* D **44** 823
[Cal91b] Caldwell D O 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S90, S137
[Cal91c] Caldwell D O 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S325
[Cal92] Caldwell D O 1992 *Proc. TAUP'91 Nucl. Phys.* B (Proc. Suppl.) A **28** 273 (Amsterdam: North-Holland)
[Cal93] Caldwell D O 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 371

[Cal94] Caldwell D O 1994 *Prog. Part. Nucl. Phys.* **32** 109
[Cal95] Caldwell D O and Mohapatra R N 1995 *Phys. Lett.* **354B** 371
[Cam91] Campbell B, Davidson S, Ellis J and Olive K 1991 *Phys. Lett.* **256B** 457
[Can81] Canuto V M and Hsieh S-H 1981 *Astrophys. J.* **248** 801
[Cap85] Caplin A D *et al* 1985 *Nature* **317** 234
[Cap86] Caplin A D, Hardiman M, Koratzinos M and Schouten J C 1986 *Nature* **321** 402
[Car74] Carrigan jr R A, Nezrick F A and Strauss B P 1974 *Phys. Rev.* D **10** 3867
[Car78] Cartwright B G, Shirk E K and Price P B 1978 *Nucl. Instrum. Methods* **153** 457
[Car82] Carrigan R A and Trower W P 1982 *Sci. Am.* **246** 91
[Car83] Carrigan R A and Trower W P (ed) 1983 *Magnetic Monopoles, Proc. NATO Advanced Study Institute (Wingspread, 1982)* (New York: Plenum)
[Car93] Carone C D 1993 *Phys. Lett.* **308B** 85
[Cav84] Cavaignac J F *et al* 1984 *Phys. Lett.* **148B** 387
[Cav98] Cavendish H 1798 *Phil. Trans. R. Soc., London*, pt II **88** 469
[Ces90] Cesarsky C *et al* 1990 *Preprint* DPhPE-90-22
[Cha39] Chandrasekhar S 1939 *An Introduction to the Study of Stellar Structure* (Chicago: University of Chicago Press)
[Cha82] Chang D and Pal P B 1982 *Phys. Rev.* D **26** 3113
[Cha85] Chang D *et al* 1985 *Phys. Rev. Lett.* **55** 2835
[Cha87] Chave A D *et al* 1987 *Nature* **326** 250
[CHA89] CHARM II Collaboration 1989 *Phys. Lett.* **232B** 539
[Che84] Chen Y T, Cook A H and Metherell A J F 1984 *Proc. R. Soc., London*, A **394** 47
[Che92] Chen M, Imel D A, Radcliffe T J and Boehm F 1992 *Phys. Rev. Lett.* **69** 3151
[Che94] Chen M *et al* 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 447 (Amsterdam: North-Holland)
[Chi80] Chikashige Y, Mohapatra R N and Peccei R D 1980 *Phys. Rev. Lett.* **45** 1926
[Chi81] Chikashige Y, Mohapatra R N and Peccei R D 1981 *Phys. Lett.* **98B** 265
[Cho80] Chodos A and Detweiler S 1980 *Phys. Rev.* D **21** 2167
[Chr64] Christenson J H, Cronin J W, Fitch V L and Turlay R 1964 *Phys. Rev. Lett.* **13** 138
[Chr94] Christensen-Dalsgaard J 1994 *Europhys. News* **25** 71
[Chu66] Chupka W A, Schiffer J P and Stevens C M 1966 *Phys. Rev. Lett.* **17** 60
[Chu88] Chu Y, Hoell J, Blome H J and Priester W 1988 *Astrophys. Space Sci.* **148** 119
[Cis71] Cisneros A 1971 *Astrophys. Space Sci.* **10** 87
[Civ87] Civitarese O, Faessler A and Tomoda T 1987 *Phys. Lett.* **194B** 11
[Cli87] Cline D B 1987 *Proc. Telemark IV Conf. on Neutrino Masses and Neutrino Astrophysics (Ashland, Wisconsin)* ed V Barger *et al* (Singapore: World Scientific)
[Coh87] Cohen E R and Taylor B N 1987 *Rev. Mod. Phys.* **59** 1121
[Coo69] Cook D D, DePasquall G, Frauenfelder H, Peacock R N, Steinrisser F and Wattenberg A 1969 *Phys. Rev.* **188** 2092

[Coo93] Cooper S *et al* 1993 *Proposal* MPI-PhE *Preprint* 93-29
[Cop95] Copi C J, Schramm D N and Turner M S 1995 *Science* **267** 192
[Cor94] Cornaz A, Hubler B and Kündig W 1994 *Phys. Rev. Lett.* **72** 1152
[Cos88] Costa G *et al* 1988 *Nucl. Phys.* B **297** 244
[Cow88] Cowsik R *et al* 1988 *Phys. Rev. Lett.* **61** 2179
[Cow90] Cowsik R *et al* 1990 *Phys. Rev. Lett.* **64** 336
[Cow91] Cowan J *et al* 1991 *Phys. Rep.* **208** 267
[Cra86] Craigie N S *et al* 1986 *Theory and Detection of Magnetic Monopoles in Gauge Theories* (Singapore: World Scientific)
[Cre76] Cremmer E and Scherk J 1976 *Nucl. Phys.* B **103** 393; 1976 *Nucl. Phys.* B **108** 409
[Cri86] Cribier M, Hampel W, Rich J and Vignaud D 1986 *Phys. Lett.* **182B** 89
[Cro86] Cromar M W *et al* 1986 *Phys. Rev. Lett.* **56** 2561
[Dah82] Dahlen F A 1982 *Phys. Rev.* D **25** 1735
[Dan74] Daniel H 1974 *Beschleuniger* (Stuttgart: Teubner)
[Dan89] Danevich F A *et al* 1989 *JETP Lett.* **49** 476
[Dan92] Danevich F A *et al* 1992 *Proc. 3rd Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'92 (Dubna, Russia)* (Singapore: World Scientific)
[Dan95] Danevich F A *et al* 1995 *Phys. Lett.* **344B** 72
[Das95] Dassie D *et al* 1995 *Phys. Rev.* D **51** 2090
[Dat85] Datar V M *et al* 1985 *Nature* **318** 547
[Dav68] Davis jr R, Harmer D S and Hoffman K C 1968 *Phys. Rev. Lett.* **20** 1205
[Dav84] Davis jr R, Cleveland B T and Rowley J K 1984 *Proc. Conf. on the Interaction between Particle and Nuclear Physics (Steamboat Springs)*
[Dav87a] Davis R 1987 *Proc. Int. Symp. on Underground Physics (Baksan, USSR)* ed G V Domogatsky *et al* p 6
[Dav87b] Davis jr R 1987 *Proc. 7th Workshop on Grand Unification, ICOBAN '86 (Toyama)* ed J Arafune (Singapore: World Scientific) p 237
[Dav88] Davis R, Lande K, Cleveland B T, Ullman J and Rowley J K 1988 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 518
[Dav92a] Davis M, Summers F J and Schlegel D 1992 *Nature* **359** 393
[Dav92b] Davies P (ed) 1992 *The New Physics* (Cambridge: Cambridge University Press)
[Dav94] Davis R 1994 *Prog. Part. Nucl. Phys.* **32** 13
[Dav94a] Davis R 1994 *Proc. 6th Int. Workshop on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin
[Dec95] Declais Y *et al* (Bugey III) 1995 *Nucl. Phys.* B **434** 503
[Dek93a] Dekel A 1993 *Ann. NY Acad. Sci.* **688** 558
[Dek93b] Dekel A *et al* 1993 *Astrophys. J.* **412** 1
[DEL89] DELPHI Collaboration 1989 *Phys. Lett.* **231B** 539
[DEL91] DELPHI Collaboration 1991 *Phys. Lett.* **255B** 466
[Den90] Denegri D, Sadoulet B and Spiro M 1990 *Rev. Mod. Phys.* **62** 1
[Der83] Derbin A V and Popeko L A 1983 *Sov. J. Nucl. Phys.* **38** 665

[Der86] Derendinger J P 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interaction in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer)
[DeR86] De Rújula A 1986 *Phys. Lett.* **180B** 213
[Des87] Deshpande N G 1987 *Proc. Telemark IV Conf. on Neutrino Masses and Neutrino Astrophysics (Ashland, Wisconsin)* ed V Barger *et al* (Singapore: World Scientific) p 78
[DeWit62] De Witt B S 1962 *An Introduction to Current Research* ed E Witten (New York: Wiley) p 49
[Die88] Dieterle B D 1988 *Proc. VIIIth Moriond Workshop on 5th Force – Neutrino Physics (Les Arcs)* ed O Fackler and J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 57
[Dim82] Dimopoulos S, Preskill J R and Wilczek F 1982 *Phys. Lett.* **119B** 320
[Din81] Dine M, Fischler W and Srednicki M 1981 *Phys. Lett.* **104B** 199
[Din83] Dine M and Fishler W 1983 *Phys. Lett.* **120B** 137
[Dir31] Dirac P A M 1931 *Proc. R. Soc.* A **133** 60
[Dir37] Dirac P A M 1937 *Nature* **139** 323
[Dir48] Dirac P A M 1948 *Phys. Rev.* **74** 817
[Doi83] Doi M *et al* 1983 *Prog. Theor. Phys.* **69** 602
[Doi85] Doi M, Kotani T and Takasugi E 1985 *Prog. Theor. Phys.* Suppl. **83** 1
[Doi88] Doi M, Kotani T and Takasugi E 1988 *Phys. Rev.* D **37** 2575
[Dok83] Doke T *et al* 1983 *Phys. Lett.* **129B** 370
[Dol81] Dolgov A D and Zeldovich Ya B 1981 *Rev. Mod. Phys.* **53** 1
[Dol92] Dolgov A D 1992 *Phys. Rep.* **222** 309
[Dol93] Dolgov A D and Rothstein I Z 1993 *Phys. Rev. Lett.* **71** 476
[Dom87] Dombeck T *et al* 1987 *Phys. Lett.* **194B** 491
[Dov79] Dover C B, Gaisser T K and Steigman G 1979 *Phys. Rev. Lett.* **42** 1117
[Dov83] Dover C B, Gal A and Richard M 1983 *Phys. Rev.* D **27** 1090
[Doz75] Dozol J-F and Neuilly M in [Okl75] p 357
[Dra87] Dragon N, Ellwanger U and Schmidt M G 1987 *Prog. Part. Nucl. Phys.* **18** 1
[Dre77] Dress W B *et al* 1977 *Phys. Rev.* D **15** 9
[Dre78] Dress W B *et al* 1978 *Phys. Rep.* **43** 410
[Dre83] Drell S D *et al* 1983 *Phys. Rev. Lett.* **50** 644
[Dre94] Dreiner H and Morawitz P 1994 *Nucl. Phys.* B **428** 31
[Dub88] Dubbers D 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 59
[Dyl73] Dylla H F and King J G 1973 *Phys. Rev.* A **7** 1224
[Dys67] Dyson F J 1967 *Phys. Rev. Lett.* **19** 1291
[Dys72] Dyson F J 1972 *Aspects of Quantum Theory* ed A Salam and E P Wigner (Cambridge: Cambridge University Press) p 213
[Ebe75] Eberhard P H *et al* in [Par90]
[Ebi84] Ebisu T and Watanabe T 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum) p 503
[Eck88] Eckhardt D H *et al* 1988 *Phys. Rev. Lett.* **60** 2567
[Edd26] Eddington A S 1926 *The Internal Constitution of the Stars* (Cambridge)

[Efs90] Efstathiou G *et al* 1990 *Nature* **348** 705
[Egg95] Eggert K *et al* 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 240
[Eis92] Eisele F and Wolf G 1992 *Phys. Bl.* **48** 787
[Eji87] Ejiri H *et al* 1987 *J. Phys. G: Nucl. Phys.* **13** 839
[Eji91] Ejiri H *et al* 1991 *Phys. Lett.* **258B** 17
[Eji91b] Ejiri H, Fushimi K, Kawasaki M, Kinoshita H, Ohsumi H, Okado K, Sano H, Shima T, Takasugi E, Tanaka J and Watanabe T 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S155
[Eji92] Ejiri H *et al* 1992 *Nucl. Phys.* B (Proc. Suppl.) A **28** 219
[Eji94] Ejiri H *et al* 1994 *Nucl. Phys.* B (Proc. Suppl.) **35** 372
[Ell81] Ellis J, Gaillard M K, Nanopoulos D V and Rudaz S 1981 *Phys. Lett.* **99B** 101
[Ell82] Ellis J, Nanopoulos D V and Olive K A 1982 *Phys. Lett.* **116B** 127
[Ell86] Elliott S R, Hahn A A and Moe M K 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer) p 692
[Ell87] Elliott S R, Hahn A A and Moe M K 1987 *Phys. Rev. Lett.* **59** 1649
[Ell88a] Ellis J and Flores R A 1988 *Nucl. Phys.* B **307** 883
[Ell88b] Ellis J 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 259
[Ell90a] Ellis J *et al* 1990 *Phys. Lett.* **245B** 251
[Ell90b] Ellis J, Kelley S and Nanopoulos D V 1990 *Phys. Lett.* **249B** 441
[Ell91a] Ellis J, Ridolfi G and Zwirner F 1991 *Phys. Lett.* **257B** 83
[Ell91b] Elliott S R *et al* 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S145
[Ell92] Elliott S R *et al* 1992 *Phys. Rev.* C **46** 1535
[Ell93] Ellis J and Flores R A 1993 *Phys. Lett.* **300B** 175
[Ell94] Ellis J 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 5 (Amsterdam: North-Holland)
[Els90] Elsworth Y *et al* 1990 *Nature* **347** 536
[Eng88] Engel J, Vogel P and Zirnbauer M R 1988 *Phys. Rev.* C **37** 731
[Eöt22] von Eötvös R, Pekár D and Fekete E 1922 *Ann. Phys., Lpz.* **68** 11
[Eöt91] Eötvös R B 1891 *Math. Naturwiss. Berichte Ungarn* **8** 65, 448
[Ern84] Ernwein J 1984 *Nucl. Instrum. Methods* **225** 583
[Err83] Errede S *et al* 1983 *Phys. Rev. Lett.* **51** 245
[Err84] Errede S M 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum)
[Ewa87] Ewan G T *et al* 1987 *Sudbury Neutrino Observatory Proposal*
[Fab79] Faber S M and Gallagher J S 1979 *Ann. Rev. Astron. Astrophys.* **17** 135
[Fae88] Faessler A 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer) p 164
[Fal94] Falk T, Olive K A and Srednicki M 1994 *Phys. Lett.* **339B** 248
[Fau85] Faulkner J and Gilliland R L 1985 *Astrophys. J.* **299** 994
[Fay86] Fayet P 1986 *Phys. Lett.* **171B** 261
[Fay89] Fayet P 1989 *Phys. Lett.* **227B** 127
[Fay91] Fayans S A 1991 private communication

[Fei82] von Feilitzsch F, Hahn A A and Schreckenbach K 1982 *Phys. Lett.* **118B** 162
[Fei88a] von Feilitzsch F 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 1
[Fei88b] von Feilitzsch F and Oberauer L 1988 *Phys. Lett.* **200B** 580
[Fer50] Fermi E, Orear J, Rosenfeld A H and Schluter R 1950 *Nuclear Physics* (Chicago: University of Chicago Press) p 201
[Fer94] Ferger P *et al* 1994 *Phys. Lett.* **323B** 95
[Fic91] Fich M and Tremaine S 1991 *Ann. Rev. Astron. Astrophys.* **29** 409
[Fid61] Fidecaro M, Finocchiaro G and Giacomelli G 1961 *Nuovo Cimento* **22** 657
[Fid85] Fidecaro G *et al* 1985 *Phys. Lett.* **156B** 122
[Fif93] 1993 *Proc. Fifth Int. Workshop on Low Temperature Detectors (Berkeley)*; *J. Low Temp. Phys.* **93** 3, 4
[Fio67] Fiorini E *et al* 1967 *Phys. Lett.* **45B** 602
[Fio84] Fiorini E and Niinikoski T O 1984 *Nucl. Instrum. Methods* **224** 83
[Fio91a] Fiorini E 1991 *Physica* B **167** 388
[Fio91b] Fiorini E 1991 *Nucl. Phys. News* **1**(5) 17
[Fir48] Fireman E L 1948 *Phys. Rev.* **74** 1238
[Fir49] Fireman E L 1949 *Phys. Rev.* **75** 323
[Fir52] Fireman E L and Schwarzer D 1952 *Phys. Rev.* **86** 451
[Fis86] Fischbach E *et al* 1986 *Phys. Rev. Lett.* **56** 3
[Fis88] Fischbach E *et al* 1988 *Ann. Phys., NY* **182** 1
[Fis89] Fisher P *et al* 1989 *Phys. Lett.* **218B** 257
[Fis93] Fisher K B *et al* 1993 *Astrophys. J.* **402** 42
[Fla81] Van Flandern T C 1981 *Astrophys. J.* **248** 813
[Fle58] Flerov G N *et al* 1958 *Sov. Phys. Dokl.* **3** 79
[Fle69] Fleischer R L *et al* 1969 *Phys. Rev.* **184** 1398
[Fle75] Fleischer R L, Price P B and Walker R M 1975 *Nuclear Tracks in Solids* (Berkeley: University of California Press)
[For94] Fornengo N 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 145 (Amsterdam: North-Holland)
[Fre82] Freund P G O 1982 *Nucl. Phys.* B **209** 146
[Fre83] Freese K, Turner M S and Schramm D N 1983 *Phys. Rev. Lett.* **51** 320
[Fre85] Freedman D Z and van Nieuwenhuizen P 1985 *Sci. Am.* **252**(3) 62
[Fre88] Freund P G O and Mahanthappa K T (ed) 1988 *Proc. NATO Advanced Research Workshop on Superstrings (Boulder, Colorado, 1987)* (New York: Plenum)
[Frem50] Fremlin J H and Walters M C 1950 *Proc. Phys. Soc.* A **63** 1178; 1952 *Proc. Phys. Soc.* A **65** 911, corrected for $Q_{2\beta}$ value
[Fri75] Fritzsch H and Minkowski P 1975 *Ann. Phys.* **93** 193
[Fri84] Frisch H J 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum) p 515
[Fri86] Fritschi M *et al* 1986 *Phys. Lett.* **173** 485
[Fri88] Frieman J A, Haber H E and Freese K 1988 *Phys. Lett.* **200B** 115
[Fri91] Fritschi M *et al* 1991 *Nucl. Phys.* B (Proc. Suppl.) **19** 205
[Fri92] Fritzsch H 1992 *Phys. Bl.* **48** 711
[Fuj71] Fujii Y 1971 *Nature* **234** 5; 1971 *Phys. Rev.* D **9** 874

[Fuj72] Fujii Y 1972 *Ann. Phys., NY* **69** 494
[Fuj80] Fujikawa K and Shrock R E 1980 *Phys. Rev. Lett.* **45** 963
[Fuk69] Fukushima Y *et al* 1969 *Phys. Rev.* **178** 2058
[Fuk87] Fukugita M and Yanagida T 1987 *Phys. Rev. Lett.* **58** 1807
[Fuk90] Fukugita M and Yanagida T 1990 *Phys. Rev.* D **42** 1285
[Fuk94] Fukuda Y *et al* 1994 *Phys. Lett.* **335B** 237
[Ful90] Fulton R *et al* 1990 *Phys. Rev. Lett.* **64** 16
[Fur39] Furry W H 1939 *Phys. Rev.* **56** 1184
[Gab84] Gabathuler K *et al* 1984 *Phys. Lett.* **138B** 449
[Gai94] Gaisser T K 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 209 (Amsterdam: North-Holland)
[Gal77] Gallinaro G, Marinelli M and Morpurgo G 1977 *Phys. Rev. Lett.* **38** 1255
[Gam67] Gamow G 1967 *Phys. Rev. Lett.* **19** 759
[Gan90] Gandhi R and Burrows A 1990 *Phys. Lett.* **246B** 149
[Gar91] Gardner R D, Cabrera B, Huber M E and Taber M A 1991 *Phys. Rev.* D **44** 622
[Gau86] Gauthier A 1986 *Proc. XXIIIrd Int. Conf. on High Energy Physics (Berkeley)* (Singapore: World Scientific)
[Gav90] Gavela M B 1990 *CP Violation in Particle Physics and Astrophysics* ed J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 249
[Gav91] Gavrin V N *et al* 1990 *Proc. 25th Int. Conf. on High Energy Physics (Singapore)* ed K K Phua and Y Yamaguchi (Singapore: World Scientific)
[Gav94] Gavrin V N *et al* 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 412 (Amsterdam: North-Holland)
[Gel64] Gell-Mann M 1964 *Phys. Lett.* **8** 214
[Gel79] Gell-Mann M, Ramond P and Slansky S 1979 *Supergravity* ed P van Nieuwenhuizen and D Z Freedman (Amsterdam: North Holland)
[Gel81] Gelmini G B and Roncadelli M 1981 *Phys. Lett.* **99B** 411
[Gel88] Gelmini G 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 309
[Gel89] Geller M J and Huchra J P 1989 *Science* **246** 892
[Gel91] Gelmini G, Nussinov S and Peccei R D 1991 *Preprint* UCLA/91/ TEP/15
[Gel92] Gelmini G and Yanagida T 1992 *Phys. Lett.* **294B** 53
[Geo74] Georgi H and Glashow S L 1974 *Phys. Rev. Lett.* **32** 438
[Geo75] Georgi H 1975 *Particles and Fields* ed C E Carlson (New York: AIP)
[Geo81] Georgi H M, Glashow S L and Nussunov S 1981 *Nucl. Phys.* B **193** 297
[Gia83] Giacomelli G 1983 *Proc. NATO Advanced Study Institute on Magnetic Monopoles (Wingspread, 1982)* ed R A Carrigan jr and W P Trower (New York: Plenum) p 41
[Gia88] Giaconelli G 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 83
[Gia94] Giammarchi M G 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 433 (Amsterdam: North-Holland)
[Gib81] Gibbons G W and Whiting B F 1981 *Nature* **291** 636
[Gil79] Gilman F J and Wise M B 1979 *Phys. Lett.* **83B** 83

[Gil86] Gilliland R L, Faulkner J, Press W H and Spergel D N 1986 *Astrophys. J.* **306** 703
[Giu91] Giuliani A 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S309
[Gla61] Glashow S L 1961 *Nucl. Phys.* **22** 579
[Gla70] Glashow S L, Iliopoulos J and Maiani L 1970 *Phys. Rev.* D **2** 1285
[Gla91] Glashow S L 1991 *Phys. Lett.* **256B** 255
[Goe35] Goeppert-Mayer M 1935 *Phys. Rev.* **48** 512
[Gol58] Goldhaber M, Grodzins L and Sunyar A W 1958 *Phys. Rev.* **109** 1015
[Gol61] Goldstone J 1961 *Nuovo Cimento* **19** 154
[Gol65] Goldhaber A S 1965 *Phys. Rev.* B **140** 1407
[Gol80] Goldhaber M, Langacker P and Slansky R 1980 *Science* **210** 851
[Gol82] Goldman T and Nieto M M 1982 *Phys. Lett.* **112** 437
[Gol86] Goldman T, Hughes R J and Nieto M M 1986 *Phys. Lett.* **171B** 217
[Gol88a] Goldman I, Aharonov Y, Alexander G and Nussinov S 1988 *Phys. Rev. Lett.* **60** 1789
[Gol88b] Goldman T, Hughes R J and Nieto M M 1988 *Mod. Phys. Lett.* A **3** 1243
[Gov71] Gove N B and Martin M J 1971 *Nucl. Data Tables* **10** 205
[Gre79] Greiner W 1979 *Quantenmechanik I + II* (Thun: Deutsch)
[Gre84] Green M B and Schwarz J H 1984 *Phys. Lett.* **149B** 117
[Gre85] Green M and Schwarz J 1985 *Phys. Lett.* **151B** 21
[Gre86] Greiner W and Müller B 1986 *Theoret. Physik 8, Eichtheorien der schwachen Wechselwirkung* (Thun: Deutsch)
[Gre87] Green M, Schwarz J and Witten E 1987 *Superstring Theory* vols I, II (Cambridge: Cambridge University Press)
[Gre89] Greiner W 1989 *Quantum Mechanics* (Heidelberg: Springer)
[Gri77] Grisaru M T, Pendleton H N and van Nieuwenhuysen P 1977 *Phys. Rev.* D **15** 996
[Gri87] Griest K and Seckel D 1987 *Nucl. Phys.* B **283** 681
[Gri90] Grifols J A and Massó A 1990 *Phys. Lett.* **242B** 77
[Gro46] Grover F W 1946 *Inductance Calculations* (New York: Van Nostrand)
[Gro85a] Gross D, Harvey J, Matinec E and Rohm R 1985 *Phys. Rev. Lett.* **54** 502
[Gro85b] Grotz K and Klapdor H V 1985 *Phys. Lett.* **153B** 1
[Gro85c] Grotz K and Klapdor H V 1985 *Phys. Lett.* **157B** 242
[Gro86a] Grotz K, Klapdor H V and Metzinger J 1986 *Phys. Rev.* C **33** 1263
[Gro86b] Groom D E 1986 *Phys. Rep.* **140** 323
[Gro86c] Grotz K and Klapdor H V 1986 *Nucl. Phys.* A **460** 395
[Gro89] Grotz K and Klapdor H V 1989 *Die Schwache Wechselwirkung in Kern-, Teilchen- und Astrophysik* (Stuttgart: Teubner)
[Gro90] Grotz K and Klapdor H V 1990 *The Weak Interaction in Nuclear, Particle and Astrophysics* (Bristol: Hilger)
[Gro92] Grotz K and Klapdor-Kleingrothaus H V 1992 *Slaboe wsaimodeistwie fisike jadra, tschastiz i astrofisike* (Moscow: Isdatelstvo Mir)
[Gun90] Gunion J F, Haber H E, Kane G and Dawson S 1990 *The Higgs Hunter's Guide* (Reading, MA: Addison-Wesley)
[Gut81] Guth A H 1981 *Phys. Rev.* D **23** 347

[Hab93] Haber H E 1993 *Proc. Workshop on Recent Advances in the Superworld (Houston, 1993) Preprint* hep-ph/9308209
[Hal84] Hall L J and Suzuki M 1984 *Nucl. Phys.* B **231** 419
[Hal86] Halprin A 1986 *Phys. Rev.* D **34** 3462
[Ham85] Hampel W 1985 *Solar Neutrinos and Neutrino Astronomy (AIP Conf. Proc. 126)* ed M L Cherry *et al* (New York: AIP) p 162
[Ham85a] Hamzaoui C and Barroso A 1985 *Phys. Lett.* **154B** 202
[Ham86a] Hampel W 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer) p 718
[Ham86b] Hampel W 1986 *Sterne und Weltraum*
[Ham88a] Hampel W 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer)
[Ham88b] Hampel W 1988 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 311
[Ham93] Hampel W 1993 *J. Phys. G: Nucl. Part. Phys.* **19** S209
[Ham94] Hampel W 1994 *Phil. Trans. R. Soc.* A **346** 3
[Har89] Harari H 1989 *Phys. Lett.* **216B** 413
[Hax82] Haxton W C, Stephenson G J and Strottman D 1982 *Phys. Rev.* D **25** 2360
[Hay50] Hayashi C 1950 *Prog. Theor. Phys.* **5** 224
[He89] He X G, McKellar B H and Pakvasa S 1989 *Int. J. Mod. Phys.* A **4** 5011
[Hec89] Heckel B R *et al* 1989 *Phys. Rev. Lett.* **63** 2705
[Hel83] Hellings R W, Adams P J, Anderson J D, Keesey M S, Lau E L, Standish E M, Canuto V M and Goldman I 1983 *Phys. Rev. Lett.* **51** 1609
[Hel95] Hellmig J 1995 *PhD Thesis* Universität Heidelberg in preparation
[Het87] Hetherington D W *et al* 1987 *Phys. Rev.* C **36** 1504
[Hid88] Hidaka K *et al* 1988 *Phys. Rev. Lett.* **61** 1537
[Hig64] Higgs P W 1964 *Phys. Lett.* **12** 132; 1964 *Phys. Rev. Lett.* **13** 508
[Him89] Hime A and Simpson J J 1989 *Phys. Rev.* D **39** 1837
[Him91a] Hime A and Jelley N A 1991 *Phys. Lett.* **257B** 441
[Him91b] Hime A *et al* 1991 *Phys. Lett.* **260B** 381
[Him92] Hime A 1992 *Preprint* LA-UR-92-3087
[Hir87] Hirata K S *et al* 1987 *Phys. Rev. Lett.* **58** 1490
[Hir88] Hirata K S *et al* 1988 *Phys. Lett.* **205B** 416
[Hir89] Hirata K S *et al* 1989 *Phys. Lett.* **220B** 308
[Hir90a] Hirata K S *et al* 1990 *Phys. Rev. Lett.* **65** 1297
[Hir90b] Hirata K S *et al* 1990 *Phys. Rev. Lett.* **65** 1301
[Hir91] Hirata K S *et al* 1991 *Phys. Rev. Lett.* **66** 9
[Hir92a] Hirata K S *et al* 1992 *Phys. Lett.* **280B** 146
[Hir92b] Hirsch M, Staudt A and Klapdor-Kleingrothaus H V 1992 *At. Data Nucl. Data Tables* **51** 243
[Hir93a] Hirsch M, Staudt A, Muto K and Klapdor-Kleingrothaus H V 1993 *At. Data Nucl. Data Tables* **53** 165
[Hir93b] Hirsch M, Wu X R, Klapdor-Kleingrothaus H V, Ching C R and Ho T H 1993 *Z. Phys.* A **345** 163

[Hir94] Hirsch M, Muto K, Oda T and Klapdor-Kleingrothaus H V 1994 *Z. Phys.* A **347** 151
[Hir94a] Hirsch M, Wu X R, Klapdor-Kleingrothaus H V, Ching C R and Ho T H 1994 *Phys. Rep.* **242** 403
[Hir95] Hirsch M, Klapdor-Kleingrothaus H V and Kovalenko S 1995 *Phys. Lett.* **352B** 1; *Phys. Rev. Lett.* **75** 17; *Phys. Bl.* **51** 418; 1996 *Phys. Rev.* D **53** 1329
[Hod81] Hodges D C *et al* 1981 *Phys. Rev. Lett.* **47** 1651
[Hoe90] Hoell J and Priester W 1990 *Sterne und Weltraum* **29** 638
[Hol86] Holding S C, Stacey F D and Tuck G J 1986 *Phys. Rev.* D **33** 3487
[Hol89] Holstein B R 1989 *Weak Interactions in Nuclei* (Princeton: Princeton University Press)
[Hom92] Homer G J *et al* 1992 *Z. Phys.* C **55** 549
[Hon94] Hong J T 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 261 (Amsterdam: North-Holland)
[Hos85] Hoskins J K *et al* 1985 *Phys. Rev.* D **32** 3084
[Hoy75] Hoyle F 1975 *Astronomy and Cosmology* (San Francisco: Freeman)
[Hub90] Huber M E, Cabrera B, Taber M A and Gardner R D 1990 *Phys. Rev. Lett.* **64** 835
[Hub91] Huber M E, Cabrera B, Taber M A and Gardner R D 1991 *Phys. Rev.* D **44** 636
[Huc90] Huchra J P *et al* 1990 *Astrophys. J.* Suppl. **72** 433
[Huc92] Huchra J P 1992 *Science* **256** 321
[Iac91] Iachello F, Krauss L M and Miano G 1991 *Phys. Lett.* **254B** 220
[Ilj83] Iljinov A S 1983 *Proc. Int. Conf. on Matter Non Conservation (Frascati)* ed E Bellotti and S Stipcich
[Inc84] Incandella J *et al* 1984 *Phys. Rev. Lett.* **53** 2067
[Ing49] Inghram M G and Reynolds J H 1949 *Phys. Rev.* **76** 1265
[Ing50] Inghram M G and Reynolds J H 1950 *Phys. Rev.* **78** 822
[Ing86] Ingelman G and Wetterich C 1986 *Phys. Lett.* **174B** 109
[Ioa94] Ioannissyan A and Valle J W F 1994 *Phys. Lett.* **332B** 93
[Irv86] Irvine J H and Humphreys 1986 *Prog. Part. Nucl. Phys.* **17** 59
[Jac75] Jackson J D 1975 *Classical Electrodynamics* (New York: Wiley)
[Jac76] Jachiw R and Rebbi C 1976 *Phys. Rev. Lett.* **37** 172
[Jaf80] Jaffe A and Taubes A 1980 *Vortices and Monopoles* (Boston: Birkhäuser)
[Jam93] James P A and Puxley P J 1993 *Nature* **363** 240
[Jar85] Jarlskog C 1985 *Phys. Rev. Lett.* **55** 1039
[Jar89] Jarlskog C (ed) 1989 *CP-Violation* (Singapore: World Scientific)
[Jar90a] Jarlskog C 1990 *Phys. Lett.* **241B** 579
[Jar90b] Jarlskog C and Rein D (ed) 1990 *Proc. Large Hadron Collider Workshop (Aachen)* CERN Report 90-10
[Jar90c] Jarlskog C 1990 *Proc. XXVth Rencontres de Moriond on Electroweak Interactions*; CERN-TH.5740/90
[Jec86] Jeckelmann B *et al* 1986 *Phys. Rev. Lett.* **56** 1444
[Jek90] Jekeli C, Eckhardt D H and Romaides A J 1990 *Phys. Rev. Lett.* **64** 1204

[Jon83] Jonson B *et al* 1983 *Nucl. Phys.* A **396** 479c
[Jon89] Jones W G *et al* 1989 *Z. Phys.* C **43** 349
[Jon90] Jones C and Melissinos A (ed) 1990 *Proc. Workshop on Cosmic Axions (Brookhaven, 1989)* (Singapore: World Scientific)
[Joy83] Joyce D C *et al* 1983 *Phys. Rev. Lett.* **51** 731
[Käp89] Käppeler F, Beer H and Wisshak K 1989 *Rep. Prog. Phys.* **52** 945
[Kaj84] Kajino F *et al* 1984 *Phys. Rev. Lett.* **52** 1373
[Kaj85] Kajino F *et al* 1985 *Nucl. Instrum. Methods* **228** 278
[Kaj89] KajitaT 1989 *Physics with the Superkamiokande Detector* ICR-Report-185-89-2
[Kal21] Kaluza Th 1921 *Sitzungsber. Preuss. Akad. Wiss. Berlin Math.-Phys. Kl* 966
[Kaw84] Kawagoe K *et al* 1984 *Lett. Nuovo Cimento* **41** 604
[Kaw87] Kawakami H *et al* 1987 *Phys. Lett.* **187B** 198
[Kaw88] Kawakami H *et al* 1988 *J. Phys. Soc. Japan* **57** 2873
[Kaw91] Kawakami H *et al* 1991 *Phys. Lett.* **256B** 105
[Kay81] Kayser B 1981 *Phys. Rev.* D **24** 110
[Kay84] Kayser B 1984 *Phys. Rev.* D **30** 1023
[Kay89] Kayser B, Gibrat-Debu F and Perrier F 1989 *The Physics of Massive Neutrinos (World Scientific Lecture Notes in Physics vol 25)* (Singapore: World Scientific)
[Kay92] Kayser R, Schramm T and Nieser L (ed) 1992 *Gravitational Lenses (Lecture Notes in Physics 406)* (Heidelberg: Springer)
[Key85] Keyser U 1985 *Z. Phys.* A **322** 529
[Key91] Ke You *et al* 1991 *Phys. Lett.* **265B** 53
[Kib67] Kibble T W B 1967 *Phys. Rev.* **155** 1554
[Kib76] Kibble T W B 1976 *J. Phys. A: Math. Gen.* **9** 1387
[Kim79] Kim J E 1979 *Phys. Rev. Lett.* **43** 103
[Kin81] Kinoshita K and Price P B 1981 *Phys. Rev.* D **24** 1707
[Kin89] Kinoshita K *et al* 1989 *Phys. Lett.* **228B** 543
[Kir67] Kirsten T, Gentner W and Schaeffer O A 1967 *Z. Phys.* **202** 273
[Kir68] Kirsten T *et al* 1968 *Phys. Rev. Lett.* **20** 1300
[Kir83a] Kirsten T, Richter H and Jessberger E 1983 *Phys. Rev. Lett.* **50** 474; 1983 *Z. Phys.* C **16** 189
[Kir83b] Kirsten T 1983 *Science Underground 1982 (AIP Conf. Proc. 96)* (New York: AIP) p 396
[Kir84] Kirsten T 1984 *Resonance Ionization Spectroscopy 1984 (Inst. Phys. Conf. Ser. 71)* ed G S Hurst and M G Payne (Bristol: Institute of Physics) p 251
[Kir86a] Kirsten T 1986 *Sterne und Weltraum* **25** 375
[Kir86b] Kirsten T 1986 *Proc. VIth Moriond Workshop on Massive Neutrinos in Astrophysics and Particle Physics* ed O Fackler and J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 119
[Kir86c] Kirsten T 1986 *Proc. Int. Symp. on Nuclear Beta Decays and Neutrino (Osaka)* ed T Kotani *et al* (Singapore: World Scientific) p 81
[Kir88a] Kirsten T 1988 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 742

[Kir88b] Kirsten T 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 221

[Kir93] Kirsten T 1993 *Sterne und Weltraum* **1** 16

[Kir95] Kirsten T 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 68

[Kit79] Kittel Ch, Knight W D, Ruderman M A, Helmholtz A C and Mayer B J 1979 *Berkeley Physikkurs* I (Braunschweig: Vieweg)

[Kla81] Klapdor H V 1981 *Phys. Rev.* C **23** 1269

[Kla82a] Klapdor H V and Metzinger J 1982 *Phys. Lett.* **112B** 22

[Kla82b] Klapdor H V and Metzinger J 1982 *Phys. Rev. Lett.* **48** 1127

[Kla83] Klapdor H V 1983 *Prog. Part. Nucl. Phys.* **10** 131

[Kla84] Klapdor H V and Grotz K 1984 *Phys. Lett.* **142B** 323

[Kla86a] Klapdor H V 1986 *Prog. Part. Nucl. Phys.* **17** 419

[Kla86b] Klapdor H V and Grotz K 1986 *Astrophys. J.* **304** L39

[Kla87] Klapdor H V 1987 *MPI-Bericht* MPI H-1987-V17 (Proposal)

[Kla89] Klapdor H V *et al* 1989 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'89 (Montreal)* (Gif-sur-Yvette: Frontières) p 701

[Kla91a] Klapdor-Kleingrothaus H V 1991 *Nuclei in the Cosmos* ed H Oberhummer (Heidelberg: Springer) p 199

[Kla91b] Klapdor-Kleingrothaus H V 1991 *Spektrum der Wissenschaft* (October) 20

[Kla91c] Klapdor-Kleingrothaus H V 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S129

[Kla91d] Klapdor-Kleingrothaus H V 1991 *Proc. Int. Symp. on γ-Ray Line Astrophysics (Paris, 1990) (AIP Conf. Proc. 232)* (New York: AIP) p 464

[Kla91e] Klapdor-Kleingrothaus H V 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S537

[Kla91f] Klapdor-Kleingrothaus H V 1991 *Phys. Bl.* **47** 206

[Kla92a] Klapdor-Kleingrothaus H V 1992 *Proc. 4th Int. Symp. on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin (Padova: Padova University) p 113

[Kla92b] Klapdor-Kleingrothaus H V and Zuber K 1992 *Phys. Bl.* **48** 1017

[Kla93a] Klapdor-Kleingrothaus H V 1993 *Proc. NEUTRINO'92 (Granada), Nucl. Phys.* B (Proc. Suppl.) **31** 72 (Amsterdam: North- Holland)

[Kla93b] Klapdor-Kleingrothaus H V 1993 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'92 (Dubna, Russia)* (Singapore: World Scientific) p 201

[Kla93c] Klapdor-Kleingrothaus H V 1994 *Proc. XIIIth Int. Conf. on Particles and Nuclei (PAN XIII) (Perugia, 1993)* (Singapore: World Scientific) p 283

[Kla93d] Klapdor-Kleingrothaus H V 1994 *Proc. NATO Advanced Study Institute, Frontier Topics in Nuclear Physics (Predeal, Romania, 1993)* ed W Scheid and A Sandulescu (New York: Plenum), Series B, Phys. vol 334, p 311

[Kla94] Klapdor-Kleingrothaus H V 1994 *Proc. Int. School on Neutrinos in Cosmology, Astro, Particle and Nuclear Physics (Erice, 1993)*; *Prog. Part. Nucl. Phys.* **32** 261

[Kla95] Klapdor-Kleingrothaus H V and Zuber K 1995 *Teilchenastrophysik* (Stuttgart: Teubner)

[Kla95a] Klapdor-Kleingrothaus H V and Stoica S (eds.) 1995 *Proc. Int. Workshop on Double Beta Decay and Related Topics (Trento, Italy, 24 April–5 May 1995)* (Singapore: World Scientific)

[Kla97] Klapdor-Kleingrothaus H V 1997 *Proc. Neutrino'96, Helsinki, June 1996* (Singapore: World Scientific) p 317

[Kla98] Klapdor-Kleingrothaus H V 1998 *Preprint* hep-ea/9802007; *Beyond the Desert 1997: Accelerator and Non-Accelerator Approaches (Proc. Workshop on Physics beyond the Standard Model, June 1997, Castle Ringberg, Tegensee, Germany)* ed H V Klapdor-Kleingrothaus and H Päs (Bristol: Institute of Physics Publishing)

[Kle26] Klein O 1926 *Z. Phys.* **37** 895

[Kle88] Kleiss R *et al* 1988 *Z. Phys.* C **39** 393

[Kli84] Klimenko A A, Pomansky A A and Smolnikov A A 1984 *Proc. XIth Int. Conf. on Neutrino Physics and Astrophysics (Dortmund)* (Singapore: World Scientific)

[Kli86] Klimenko A A *et al* 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer) p 701

[Kli86b] Klimenko A A, Pomansky A A and Smolnikov A A 1986 *Nucl. Instrum. Methods* B **17** 445

[Kli89] Klimenko A A *et al* 1989 *Proc. TAUP'89* ed A Bottino and P Monacelli (Gif-sur-Yvette: Frontières) p 241

[Kob73] Kobayashi M and Maskawa T 1973 *Prog. Theor. Phys.* **49** 652

[Kol71] Kolm H *et al* 1971 *Phys. Rev.* D **4** 3260

[Kol82] Kolb E W, Colgate S A and Harvey J A 1982 *Phys. Rev. Lett.* **49** 1373

[Kol84a] Kolb E W and Turner M S 1984 *Astrophys. J.* **286** 702

[Kol84b] Kolb E W 1984 *NEUTRINO'84, Proc. 11th Int. Conf. on Neutrino Physics and Astrophysics (Nordkirchen)* ed K Kleinknecht and E A Paschos (Singapore: World Scientific) p 243

[Kol85] Kolb E W, Seckel D and Turner M S 1985 *Nature* **314** 415

[Kol86a] Kolb E W, Perry M J and Walker T P 1986 *Phys. Rev.* D **33** 869

[Kol86b] Kolb E W and Olive K A 1986 *Phys. Rev.* D **33** 1202

[Kol86c] Kolb E W 1986 *Proc. Int. Symp. Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer) p 369

[Kol87] Kolb E W, Stebbins A J and Turner M S 1987 *Phys. Rev.* D **35** 3598

[Kol90] Kolb E W and Turner M S 1990 *The Early Universe* (Reading, MA: Addison-Wesley)

[Kol91] Kolb E W and Turner M S 1991 *Phys. Rev. Lett.* **67** 5

[Kon66] Konopinski E J 1966 *The Theory of Beta Radioactivity* (Oxford: Oxford University Press)

[Kos92] Koshiba M 1992 *Phys. Rep.* **220** 229

[Kra91] Krauss L M 1991 *Phys. Lett.* **263B** 441

[Kri81] Krishnaswamy M R *et al* 1981 *Phys. Lett.* **106B** 339

[Kri82] Krishnaswamy M R *et al* 1982 *Phys. Lett.* **115B** 349

[Kri94] Krivosheina I 1994 *Proc. Int. School on Neutrinos in Cosmology, Astro, Particle and Nuclear Physics (Erice, 1993)*; *Prog. Part. Nucl. Phys.* **32** 41
[Kro85] Krofcheck D *et al* 1985 *Phys. Rev. Lett.* **55** 1051
[Kün94] Kündig W and Holzschuh E 1994 *Prog. Part. Nucl. Phys.* **32** 131
[Kuo86] Kuo T K and Pantaleone J 1986 *Phys. Rev. Lett.* **57** 1805
[Kur82] Kuroda P K 1982 *The Origin of the Chemical Elements and the Oklo Phenomenon* (Berlin: Springer)
[Kuz66] Kuzmin V A 1966 *Sov. Phys.–JETP* **22** 1051
[Kuz70] Kuzmin V A 1970 *JETP Lett.* **12** 228
[Kuz85] Kuzmin V, Rubakov V and Shaposhnikov M 1985 *Phys. Lett.* **185B** 36
[Kwo81] Kwon H *et al* 1981 *Phys. Rev.* D **24** 1097
[L89] L3 Collaboration 1989 *Phys. Lett.* **231B** 509
[L90] L3 Collaboration *Phys. Lett.* **248B** 464
[Lac82] Lackner K S and Zweig G 1982 *Lett. Nuovo Cimento* **33** 65
[Lac83] Lackner K S and Zweig G 1983 *Phys. Rev.* D **28** 1671
[Lal94] Lalanne D (NEMO Collaboration) 1994 *Nucl. Phys.* B (Proc. Suppl.) **35** 369
[Lam87] Lamoreaux S K *et al* 1987 *Phys. Rev. Lett.* **59** 2275
[Lam91] Lamoreaux S K, Golub R and Pendlebury J M 1991 *Europhys. Lett.* **14** 503
[Lan52] Langer L M and Moffat R J D 1952 *Phys. Rev.* **88** 689
[Lan57] Landau L 1957 *Nucl. Phys.* **3** 127
[Lan79a] Landau L D and Lifschitz E M 1979 *Lehrbuch d. theoretischen Physik, Bd. I, Mechanik* (Berlin: Akademie)
[Lan79b] Landau L D and Lifschitz E M 1979 *Lehrbuch d. theoretischen Physik, Bd. III, Quantenmechanik* (Berlin: Akademie)
[Lan79c] Landau L D and Lifschitz E M 1979 *Lehrbuch d. theoretischen Physik, Bd. IV, Statistische Physik* (Berlin: Akademie)
[Lan81] Langacker P 1981 *Phys. Rep.* **72** 185
[Lan83] Langacker P *et al* 1983 *Phys. Rev.* D **27** 1228
[Lan85] Langacker P 1985 *Comm. Nucl. Part. Phys.* **15** 41
[Lan86] Langacker P 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Berlin: Springer) p 879
[Lan88] Langacker P 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 71
[Lan92a] Langacker P 1982 *Proc. 4th Int. Symp. on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin p 73
[Lan92b] Langacker P, Mingxing Luo and Mann A K 1992 *Rev. Mod. Phys.* **64** 87
[Lan93] Langacker P 1993 *Ann. NY Acad. Sci.* **688** 34
[LaR77] LaRue G S, Fairbank W M and Hebard A F 1977 *Phys. Rev. Lett.* **38** 1011
[LaR79] LaRue G S, Fairbank W M and Phillips J D 1979 *Phys. Rev. Lett.* **42** 142
[LaR81] LaRue G S, Phillips J D and Fairbank W M 1981 *Phys. Rev. Lett.* **46** 967
[Lat47] Lattes C M G *et al* 1947 *Nature* **159** 694; 1947 *Nature* **160** 453, 486
[Lat88] Lattimer J M and Cooperstein J 1988 *Phys. Rev. Lett.* **61** 23
[Lea79] Learned J, Reines F and Soni A 1979 *Phys. Rev. Lett.* **43** 907
[Lea88] Learned J G *et al* 1988 *Phys. Lett.* **207B** 79

[Lea93] Learned J G 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 456
[Lec83] Leccia F *et al* 1983 *Lett. Nuovo Cimento* A **78** 50
[Lee72] Lee B W and Zinn-Justin J 1972 *Phys. Rev.* D **5** 3121, 3137, 3155
[Lee77] Lee B and Schrock R S 1977 *Phys. Rev.* D **16** 1444
[Lee84a] Lee H L 1984 *An Introduction to Kaluza–Klein Theories* (Singapore: World Scientific)
[Lee84b] Lee I H 1984 *Phys. Lett.* **138B** 121
[Lee94] Lee D G and Mohapatra R N 1994 *Phys. Lett.* **329B** 463
[Lee95] Lee D G, Mohapatra R N, Parida M K and Rani M 1995 *Phys. Rev.* D **51** 229
[Lee95a Lee D G and Mohapatra R N 1995 *Phys. Rev.* D **51** 1353
[Lei85] Leibacher J W *et al* 1985 *Sci. Am.* **253**(9) 34
[Lev50] Levine C A *et al* 1950 *Phys. Rev.* **27** 296
[Lew80] Lewis R R 1980 *Phys. Rev.* D **21** 663
[Lew85] Lewin J D and Smith P F 1985 *Phys. Rev.* D **32** 1177
[Li82] Li L F and Wilczek F 1982 *Phys. Rev.* D **25** 143
[Lie83] Liebowitz D, Binder M and Ziock K O H 1983 *Phys. Rev. Lett.* **50** 1640
[Lim88] Lim C S and Marciano W J 1988 *Phys. Rev.* D **37** 1368
[Lin82] Linde A D 1982 *Phys. Lett.* **108B** 389
[Lin83] Lindgren M A *et al* 1983 *Phys. Rev. Lett.* **51** 1621
[Lin84] Linde A D 1984 *Rep. Prog. Phys.* **47** 925
[Lin86] Lindner M *et al* 1986 *Nature* **320** 246
[Lin88] Lin W J *et al* 1988 *Nucl. Phys.* A **481** 477, 484
[Lin90] Linde A 1990 *Particle Physics and Inflationary Cosmology* (Harward Publ.)
[Liu87] Liu J 1987 *Phys. Rev.* D **35** 3447
[Lob84] Lobashev V M and Serebrov A P 1984 *J. Physique* **45** 3
[Loh92] Lohrmann E 1992 *Phys. Bl.* **48** 33
[Lon74] Long D R 1974 *Phys. Rev.* D **9** 850
[Lon76] Long D R 1976 *Nature* **260** 417
[LoS85a] LoSecco J M 1985 *Comm. Nucl. Part. Phys.* **15** 23
[LoS85b] LoSecco J M, Reines F and Sinclair D 1985 *Spektrum der Wissenschaft* (August); 19– *Sci. Am.* **252**(6) 42
[Lou95] Louis W C (LSND collaboration) 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 229
[Lov72] Lovelocke D 1972 *J. Math. Phys.* **13** 874
[Lub80] Lubimov V A *et al* 1980 *Phys. Lett.* **94B** 266
[Luc83] Luck J-M and Allegre C J 1983 *Nature* **302** 130
[Luc86] Lucha W 1986 *Comm. Nucl. Part. Phys.* **16** 155
[Lüd54] Lüders G 1954 *Kgl. Danske Vidensk. Selsk. Mat.-Fys. Medd.* **28**(5)
[Lüd57] Lüders G 1957 *Ann. Phys., NY* **2** 1
[Lut82] Luther G G and Towler W R 1982 *Phys. Rev. Lett.* **48** 121
[Lyo85] Lyons L 1985 *Phys. Rep.* **129** 225
[Mac84] Macrae K I and Riegert R J 1984 *Nucl. Phys.* B **244** 513
[Mai94] Maier B 1994 *Nucl. Phys.* B (Proc. Suppl.) **35** 358
[Maj37] Majorana E 1937 *Nuovo Cimento* **14** 171

[Mal91] Malaney R A 1991 *Nuclei in the Cosmos* ed H Oberhummer (Heidelberg: Springer) p 127
[Mal93] Malaney R A and Mathews G J 1993 *Phys. Rep.* **229** 147
[Mam89] Mampe W *et al* 1989 *Phys. Rev. Lett.* **63** 593
[Man86] Manuel O K 1986 *Proc. Int. Symp. on Nuclear Beta Decays and Neutrino (Osaka)* ed T Kotani *et al* (Singapore: World Scientific) p 103
[Man88] Mann A K 1988 *Proc. 13th Int. Conf. on Neutrino Physics and Astrophysics (Boston)* ed J Schneps *et al* (Singapore: World Scientific) p 105
[Man91a] Manohar A V and Nelson A E 1991 *Phys. Rev. Lett.* **66** 2847
[Man91b] Manuel O K 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S221
[Mar70] Marmier P and Sheldon E 1970 *Physics of Nuclei and Particles* vol III (New York: Academic)
[Mar80a] Marshak R E and Mohapatra R 1980 *Phys. Rev. Lett.* **44** 1316
[Mar80b] Marinelli M and Morpurgo G 1980 *Phys. Lett.* **94B** 433
[Mar80c] Marciano W J and Sirlin A 1980 *Phys. Rev.* D **22** 2695
[Mar82] Marinelli M and Morpurgo G 1982 *Phys. Rep.* **85** 161
[Mar84a] Marciano W J 1984 *Phys. Rev. Lett.* **52** 489
[Mar84b] Marinelli M and Morpurgo G 1984 *Phys. Lett.* **137B** 439
[Mar85] Markey J and Boehm F 1985 *Phys. Rev.* C **32** 2215
[Mar86] Marciano W J and Parsa Z 1986 *Ann. Rev. Nucl. Part. Sci.* **36** 171
[MAR89] MARK II Collaboration 1989 *Phys. Rev. Lett.* **63** 2173
[Mas83] Mashomo T *et al* 1983 *Phys. Lett.* **128B** 327
[Mat66] Ter Mateosian E and Goldhaber M 1966 *Phys. Rev.* **146** 810
[Mat90a] Mather J C *et al* 1990 *Astrophys. J.* **354** L37
[Mat90b] Mathews G J and Cowan J J 1990 *Nature* **345** 491
[Mau76] Maurette M 1976 *Ann. Rev. Nucl. Sci.* **26** 319
[May84] Mayer-Kuckuk T 1984 *Kernphysik* (Stuttgart: Teubner)
[McC51] McCrea W H 1951 *Proc. R. Soc.* A **206** 562
[McC83] McCusker B 1983 *The Quest for Quarks* (Cambridge: Cambridge University Press)
[McD94] McDonald A B 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 340 (Amsterdam: North-Holland)
[McK80] McKellar B H J 1980 *Phys. Lett.* **97B** 93
[Mes76] Messiah A 1976 *Quantenmechanik I, II* (Berlin: de Gruyter)
[Mes86] Messiah A 1986 *Proc. VIth Moriond Workshop on Massive Neutrinos in Particle and Astrophysics* ed O Fackler and J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 373
[Mey86] Meyer H 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Berlin: Springer) p 846; 1986 *Proc. 12th Int. Conf. on Neutrino Physics and Astrophysics (Sendai)* ed T Kitagaki and H Yuta (Singapore: World Scientific) p 674
[Mik77] Mikkelsen D R and Newman M J 1977 *Phys. Rev.* D **16** 919
[Mik86a] Mikheyev S P and Smirnov A Yu 1986 *Sov. J. Nucl. Phys.* **42** 913
[Mik86b] Mikheyev S P and Smirnov A Yu 1986 *Sov. Phys.–JETP* **64** 4

[Mik86c] Mikheyev S P and Smirnov A Yu 1986 *Proc. 12th Int. Conf. on Neutrino Physics and Astrophysics (Sendai)* ed T Kitagaki and H Yuta (Singapore: World Scientific) p 177
[Mik88a] Mikheyev S P and Smirnov A Yu 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 239
[Mik88b] Mikheyev S P and Smirnov A Yu 1988 *Sov. Phys.–JETP* **65** 230
[Mil10] Millikan R 1910 *Phil. Mag.* **19** 209
[Mil85] Milner R G *et al* 1985 *Phys. Rev. Lett.* **54** 1472
[Mil87] Milner R G *et al* 1987 *Phys. Rev.* D **36** 37
[Mil90] Miley H S *et al* 1990 *Phys. Rev. Lett.* **65** 3092
[Mis73] Misner C W, Thorne T S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman)
[Mit88] Mitchell L W and Fisher P H 1988 *Phys. Rev.* C **38** 895
[Moe80] Moe M K and Lowenthal D D 1980 *Phys. Rev.* C **22** 2186
[Moe91a] Moe M K 1991 *Nucl. Phys.* B (Proc. Suppl.) **19** 158
[Moe91b] Moe M K 1991 *Phys. Rev.* C **44** 931
[Moe92] Moe M K *et al* 1992 *Preprint* UCI-Neutrino 92-1 presented at Franklin Symp. on Celebration of the Discovery of the Neutrino, Philadelphia
[Moe93a] Moe M K 1993 *Int. J. Mod. Phys.* E **2** 507
[Moe93b] Moe M K 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 68
[Moe94] Moe M K and Vogel P 1994 *Ann. Rev. Nucl. Part. Sci.* **44** 247
[Moe94a] Moe M K, Nelson M A and Vient M A 1994 *Prog. Part. Nucl. Phys.* **32** 247
[Moe95] Moe M K 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 36
[Moh74] Mohapatra R N and Pati J C 1974 *Phys. Rev.* D **11** 566
[Moh80] Mohapatra R N and Marshak R E 1980 *Phys. Lett.* **94B** 183
[Moh81] Mohapatra R N and Vergados J D 1981 *Phys. Rev. Lett.* **47** 1713
[Moh86a] Mohapatra R N 1986 *Unification and Supersymmetry* (New York: Springer)
[Moh86b] Mohapatra R N 1986 *Phys. Rev.* D **34** 3457
[Moh86c] Mohapatra R N 1986 *Phys. Rev.* D **34** 909
[Moh88a] Mohapatra R N and Takasugi E 1988 *Phys. Lett.* **211B** 192
[Moh88b] Mohapatra R N 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 117
[Moh88c] Mohapatra R N 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer)
[Moh89] Mohapatra R N 1989 *Nucl. Instrum. Methods* A **284** 1
[Moh91a] Mohapatra R N and Pal P B 1991 *Massive Neutrinos in Physics and Astrophysics* (Singapore: World Scientific)
[Moh91b] Mohapatra R N 1991 *Prog. Part. Nucl. Phys.* **26** 1
[Moh94] Mohapatra R N 1994 *Prog. Part. Nucl. Phys.* **32** 187
[Moo84] Moody J E and Wilczek F 1984 *Phys. Rev.* D **30** 130
[Moo88] Moore G I *et al* 1988 *Phys. Rev.* D **38** 1023
[Moo92] Moody K J, Longheed R W and Hulet E K 1992 *Preprint* UCRL–JC–110153
[Mor91] Mori M *et al* 1991 *Phys. Rev.* D **43** 2843
[Mös91] Mössbauer R L 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S1
[Mül89] Müller G, Zürn W, Lindner K and Rösch N 1989 *Phys. Rev. Lett.* **63** 2621

[Mül91] Müller G 1991 *Bild der Wissenschaft* **7** 102
[Mur89] Murthy S A, Krause jr D, Li Z L and Hunter L R 1989 *Phys. Rev. Lett.* **63** 965
[Mus83] Musset P, Price M and Lohrmann E 1983 *Phys. Lett.* **128B** 333
[Mut88a] Muto K and Klapdor H V 1988 *Neutrinos* ed H V Klapdor (Heidelberg: Springer) p 183
[Mut88b] Muto K and Klapdor H V 1988 *Phys. Lett.* **201B** 420
[Mut89a] Muto K, Bender E and Klapdor H V 1989 *Z. Phys.* A **334** 177
[Mut89b] Muto K, Bender E and Klapdor H V 1989 *Z. Phys.* A **334** 187
[Mut91] Muto K, Bender E and Klapdor H V 1991 *Z. Phys.* A **39** 435
[Nac86] Nachtmann O 1986 *Phänomene und Konzepte der Elementarteilchenphysik* (Braunschweig: Vieweg)
[Nah78] Nahm W 1978 *Nucl. Phys.* B **135** 149
[Nam60] Nambu Y 1960 *Phys. Rev. Lett.* **4** 380
[Nel90] Nelson P G, Graham D M and Newman R D 1990 *Phys. Rev.* D **42** 963
[Nem81] Nemethy P *et al* 1981 *Phys. Rev.* D **23** 262
[Neu72] Neuilly M *et al* 1972 *C. R. Acad. Sci. Paris* D **275** 1847
[Ng89] Ng J 1989 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'89 (Montreal)* ed P Depommier (Gif-sur-Yvette: Frontières) p 167
[Nie87] Niebauer T M, McHugh M P and Faller J E 1987 *Phys. Rev. Lett.* **59** 609
[Nie88] Nieto M M, Goldman T and Hughes R J 1988 *Aust. Phys.* **25** 259
[Nie89] Nieto M M, Hughes R J and Goldman T 1989 *Am. J. Phys.* **57** 397
[Nie91] Nieto M M and Goldman T 1991 *Phys. Rep.* **205** 221
[Nik93] Nikolaev M A and Klapdor-Kleingrothaus H V 1993 *Z. Phys.* A **345** 183, 373
[Noe18] Noether E 1818 *Kgl. Ges. Wiss. Nachrichten Math.-Phys. Klasse (Göttingen)* 235
[Nor82] Nordtvedt jr K 1985 *Rep. Prog. Phys.* **45** 631
[Nor84] Norman E B and DeFaccio M A 1984 *Phys. Lett.* **148B** 31
[Nor85] Norman E B 1985 *Phys. Rev.* C **31** 1937
[Nor86] Norman E B 1986 *Am. J. Phys.* **54** 317
[Nor90] Nordtvedt K 1990 *Phys. Rev. Lett.* **65** 953
[Nor91] Norman E B *et al* 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S291
[Nus76] Nussinov S 1976 *Phys. Lett.* **63B** 201
[Obe87] Oberauer L *et al* 1987 *Phys. Lett.* **198B** 113
[Obe88] Oberauer L and von Feilitzsch F 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer) p 142
[Ogo79] Ogorodnikov, Samoilov and Solntsev 1979 *JETP* **49** 953
[Ohi85] Ohi T *et al* 1985 *Phys. Lett.* **160B** 322
[Oka88] Okada K *et al* 1988 *Nucl. Phys.* A **478** 447c
[Okl75] 1975 Le Phenomène d'Oklo *C. R. d'un Colloque sur le Phenomène d'Oklo (Libreville, Gabon)* (Vienna: IAEA)
[Oku86] Okun L B 1986 *Sov. J. Nucl. Phys.* **44** 546
[Oku87] Okun L B 1987 *Elementarteilchen von α bis Z* (Berlin: Akademie)

[Oli81] Olive K A *et al* 1981 *Astrophys. J.* **246** 557
[Oli85] Olive K A and Schramm D N 1985 *Comm. Nucl. Part. Phys.* **15** 69
[Oli91] Olive K A 1991 *Science* **251** 1194
[Ono91] Ono Y and Suematsu D 1991 *Phys. Lett.* **271B** 165
[Oor83] Oort J H 1983 *Ann. Rev. Astron. Astrophys.* **21** 373
[OPA89] OPAL Collaboration 1989 *Phys. Lett.* **231B** 530
[Ori91] Orito S *et al* 1991 *Phys. Rev. Lett.* **66** 1951
[Pac86] Paczynski B 1986 *Astrophys. J.* **304** 1
[Pan94] Panella O and Srivastava Y 1994 *Preprint* LPC-94-39
[Pan95] Panella O 1996 *Proc. Int. Workshop on Double Beta Decay and Related Topics (Trento, Italy, 24 April–5 May 1995)* ed H V Klapdor-Kleingrothaus and S Stoica (Singapore: World Scientific) p 145
[Par70] Parker E N 1970 *Astrophys. J.* **160** 383
[Par84] Parker E N 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum) p 125
[Pas79] Pasierb E *et al* 1979 *Phys. Rev. Lett.* **43** 96
[Pas94] Paschos E A and Zioutas K 1994 *Phys. Lett.* **323B** 367
[Pat74] Pati J C and Salam A 1974 *Phys. Rev.* D **10** 275
[Pat90] Paterson J R *et al* 1990 *Phys. Rev. Lett.* **64** 1491
[Pau55] Pauli W 1955 *Niels Bohr and the Development of Physics* (London: Pergamon) p 30
[PDG90] Particle Data Group 1990 *Phys. Lett.* **239B** 1
[PDG92] Particle Data Group, Aguilar-Benitez M *et al* 1992 *Phys. Rev.* D **45** pt 2
[PDG94] Particle Data Group 1994 *Phys. Rev.* D **50** pt I
[Pec77] Peccei R D and Quinn H R 1977 *Phys. Rev. Lett.* **38** 1440; 1977 *Phys. Rev.* D **16** 1791
[Pec87] Peccei R D, Solà I and Wetterich C 1987 *Phys. Lett.* **195B** 183
[Pee84] Peebles P J E 1984 *Astrophys. J.* **284** 439
[Pen65] Penzias A A and Wilson R W 1965 *Astrophys. J.* **142** 419
[Pen84] Pendlebury J M *et al* 1984 *Phys. Lett.* **136B** 327
[Per82] Perkins D H 1982 *Introduction to High Energy Physics* (Reading, MA: Addison-Wesley)
[Per84] Perry M J 1984 *Proc. NATO Advanced Research Workshop Monopole 1983 (Ann Arbor)* ed J L Stone (New York: Plenum) p 29
[Pes88] Pessard H 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 130
[Pet87] Petcov S T and Toshev S 1987 *Phys. Lett.* **187B** 222
[Pet94] Petcov S T and Smirnov A Yu 1994 *Phys. Lett.* **322B** 109
[Phi88] Phillips J D, Fairbank W M and Navarro J 1988 *Nucl. Instrum. Methods* A **264** 125
[Phi89] Phillips T J *et al* 1989 *Phys. Lett.* **224B** 348
[Phy86] 1986 *Physics through the 1990's: Nuclear Physics* (Washington: National Academy Press)
[Pic92] Picard A *et al* 1992 *Nucl. Instrum. Methods* B **63** 345
[Poc64] Pochoda P and Schwarzschild M 1964 *Astrophys. J.* **139** 587

[Pol74] Polyakov A M 1974 *JETP Lett.* **20** 194
[Pom86] Pomanski A 1986 *Proc. NEUTRINO'86 (Sendai)* ed T Kitagaki and H Yuta (Singapore: World Scientific)
[Pon57] Pontecorvo B 1957 *Zh. Eksp. Teor. Fiz.* **33** 549 (1958 *Sov. Phys.–JETP* **6** 429)
[Pon58] Pontecorvo B 1958 *Zh. Eksp. Teor. Fiz.* **34** 247
[Pon68] Pontecorvo B 1968 *Phys. Lett.* **26B** 630
[Pou60] Pound R V and Rebka jr G A 1960 *Phys. Rev. Lett.* **4** 337
[Pre83] Preskill J, Wise M and Wilczek F 1983 *Phys. Lett.* **120B** 127
[Pre84] Preskill J 1984 *Ann. Rev. Nucl. Part. Sci.* **34** 461
[Pre87] Pretzl K *et al* (ed) 1987 *Proc. Low Temperature Detectors for Dark Matter Detection (Ringberg Castle)* (Heidelberg: Springer)
[Pre89] Press W H and Spergel D N 1989 *Phys. Today* **3** 29
[Pre90] Pretzl K P 1990 *Particle World* **1** 153
[Pre93] Pretzl K P 1993 *Europhys. News* **24** 167
[Pri83] Price P B 1983 *Proc. NATO Advanced Study Institute on Magnetic Monopoles (Wingspread, 1982)* ed R A Carrigan jr and W P Trower (New York: Plenum) p 307
[Pri84] Price P B 1984 *Phys. Lett.* **140B** 112
[Pri86] Price P B and Salamon M H 1986 *Phys. Rev. Lett.* **56** 1226
[Pri87a] Priester W, Blome H J and Hoell J 1987 *Sky and Telescope* **73** 237
[Pri87b] Price P B, Guoxiao R and Kinoshita K 1987 *Phys. Rev. Lett.* **59** 2523
[Pri88] Primack J R, Seckel D and Sadoulet B 1988 *Ann. Rev. Nucl. Part. Sci.* **38** 751
[Pug89] Puglieri G 1989 *Nucl. Instrum. Methods* A **284** 9
[Pur50] Purcell E M and Ramsey N F 1950 *Phys. Rev.* **78** 807
[Pur63] Purcell E M *et al* 1963 *Phys. Rev.* **129** 2326
[Put78] Putt G D and Yock P C M 1978 *Phys. Rev.* D **17** 1466
[Qui83] Quigg C 1983 *Gauge Theories of the Strong, Weak and Electromagnetic Interactions* (Benjamin/Cummings)
[Raf90] Raffelt G G 1990 *Phys. Rep.* **198** 1
[Raf95] Raffelt G G and Silk J 1995 *Preprint* HEP-PH-9502306; *Phys. Lett.* **366B** 429
[Rag94] Raghavan R S 1994 *Phys. Rev. Lett.* **72** 1411
[Ram80] Ramsey N F 1980 *Phys. Today* (July) 25
[Ram82] Ramsey N F 1982 *Ann. Rev. Nucl. Part. Sci.* **32** 211
[Ram90] Ramsey N F 1990 *Ann. Rev. Nucl. Part. Sci.* **40** 1
[Rau89] Raupach F 1989 *Proc. XXIV Int. Conf. on High Energy Physics (Munich)* ed R Kotthaus and J Kühn (Berlin: Springer) p 1290
[Rei54] Reines F, Cowan Jr C L and Goldhaber M 1954 *Phys. Rev.* **96** 1157
[Rei74] Reines F and Crouch M F 1974 *Phys. Rev. Lett.* **32** 493
[Rei80] Reines F, Sobel H W and Pasierb E 1980 *Phys. Rev. Lett.* **45** 1307
[Rei83] Reines F 1983 *Nucl. Phys.* A **396** 469c
[Reu91] Reusser D *et al* 1991 *Phys. Lett.* **255B** 143
[Ric87] Rich J, Lloyd Owen D and Spiro M 1987 *Phys. Rep.* **151** 239
[Rie87] Riepe P *et al* 1987 *Sterne und Weltraum* **3** 155

[Ril91] Riley S P and Irvine J M 1991 *J. Phys. G: Nucl. Part. Phys.* **17** 35
[Rob91] Robertson R G H *et al* 1991 *Phys. Rev. Lett.* **67** 957
[Röd72] Röde B and Daniel H 1972 *Lett. Nuovo Cimento* **5** 139
[Roh86] Rohlfs K 1986 *Sterne and Weltraum* **25** 467
[Roi92] Roi D P 1992 *Phys. Lett.* **283B** 270
[Rol64] Roll P G, Krotkov R and Dicke R H 1964 *Ann. Phys., NY* **26** 442
[Roo88] Roos M 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer) p 57
[Ros88] Rosen S P 1988 *Comm. Nucl. Part. Phys.* **18** 31
[Row85] Rowley J K, Cleveland B T and Davis jr R 1985 *Solar Neutrinos and Neutrino Astronomy (AIP Conf. Proc. 126)* ed M L Cherry *et al* (New York: AIP) p 1
[Rub81] Rubakov V 1981 *JETP Lett.* **33** 644; 1982 *Nucl. Phys.* B **203** 311
[Rub93] Rubbia C 1993 *Proc. XXVI Int. Conf. on High Energy Physics (AIP Conf. Proc. 272)* (New York: AIP) p 321
[Ruj81] De Rújula A 1981 *Nucl. Phys.* B **188** 414
[Sac67a] Sacharov A D 1967 *JETP Lett.* **5** 24
[Sac67b] Sacharov A D 1967 *JETP Lett.* **5** 32
[Sad94] Sadoulet B 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 117 (Amsterdam: North-Holland)
[Sal68] Salam A 1968 *Proc. 8th Nobel Symp. (Stockholm)* ed N Svartholm (Stockholm: Almqvist and Wiksell) p 367
[San83] Sandage A and Tammann G A 1983 *Proc. 1st ESO-CERN Symp.* ed G Setti and L Van Hove
[San90] Sanders R H 1990 *Astron. Astrophys. Rev.* **2** 1
[Sat87] Sato K and Suzuki H 1987 *Phys. Rev. Lett.* **58** 2722
[Sav86] Savage M L *et al* 1986 *Phys. Lett.* **167B** 481
[Sch51] Schwinger J 1951 *Phys. Rev.* **82** 914
[Sch77] Scherk J 1977 *La Recherche* **8** 878
[Sch78] Schiffer J P *et al* 1978 *Phys. Rev.* D **17** 2241
[Sch79] Scherk J 1979 *Phys. Lett.* **88B** 265
[Sch82] Schechter J and Valle J W F 1982 *Phys. Rev.* D **25** 2951
[Sch83] Schreckenbach K, Colvin G and von Feilitzsch F 1983 *Phys. Lett.* **129B** 265
[Sch84] Schreckenbach K 1984 *Technical Report 84SC26T* Institut Laue-Langevin, Grenoble
[Sch85a] Schreckenbach K, Colvin G, Gelletly W and von Feilitzsch F 1985 *Phys. Lett.* **160B** 325
[Sch85b] Schatzman E 1985 *Solar Neutrinos and Neutrino Astronomy (AIP Conf. Proc. 126)* ed M Cherry *et al* (New York: AIP) p 69
[Sch86a] Schreckenbach K *et al* 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Berlin: Springer) p 759
[Sch86b] Schramm D N 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Berlin: Springer) p 1033

[Sch89] Schulz H 1989 *Sterne and Weltraum* **28** 588, 656
[Sch90a] Schwarzschild B 1990 *Phys. Today* (October) 17
[Sch90b] Schramm D N and Truran J W 1990 *Phys. Rep.* **189** 89
[Sch90c] Schramm D N 1990 *Dark Matter in the Universe* ed P Galeotti and D N Schramm (Dordrecht: Kluwer) p 1
[Sch90d] Schramm D N 1990 *Proc. La Thuile Workshop (March 1990)*
[Sch91] Schwarzschild B 1991 *Phys. Today* (May) 17
[Sch93] Schwitters R F 1993 *Proc. XXVI Int. Conf. on High Energy Physics (AIP Conf. Proc. 272)* (New York: AIP) p 306
[Schm90a] Schmidt-Parzefall W 1990 *Phys. Bl.* **46** 442
[Schm90b] Schmüser P 1990 *Phys. Bl.* **46** 470
[Schn94] Schneider J R, Söding P, Voss G A, Wagner A and Wiik B H 1994 *Europhys. News* **25** 91
[Scho89] Schopper H 1989 *Materie und Antimaterie* (Piper)
[Scho91] Schopper H 1991 *Phys. Bl.* **47** 907
[Schr85a] Schramm D N and Vittorio N 1985 *Comm. Nucl. Part. Phys.* **15** 1
[Schr85b] Schrempp B and Schrempp F 1985 *Phys. Bl.* **41** 335
[Schw85] Schwarz J H 1985 *Comm. Nucl. Part. Phys.* **15** 9
[Sei88] Seidel S *et al* 1988 *Phys. Rev. Lett.* **61** 2522
[Sei93] Seidel W *et al* 1993 *J. Low Temp. Phys.* **93** 797
[Sel91] Selvin P 1991 *Science* **251** 1426
[Sex74] Sextro R G, Gough R A and Cerny J 1974 *Nucl. Phys.* A **234** 130
[Sha71] Shapiro I I *et al* 1971 *Phys. Rev. Lett.* **26** 27
[Sha86] Shafi Q 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer)
[Shl76] Shlyakhter A I 1976 *Nature* **264** 340
[Shr80] Shrock R E 1980 *Phys. Lett.* **96B** 159
[Shr81] Shrock R E 1981 *Phys. Rev.* D **24** 1232, 1275
[Sik83] Sikivie P 1983 *Phys. Rev. Lett.* **51** 1415; 1984 *Phys. Rev. Lett.* **52** 695
[Sik85] Sikivie P 1985 *Phys. Rev.* D **32** 2988
[Sik90] Sikivie P 1990 *Dark Matter in the Universe* ed H Sato and H Kodama (Berlin: Springer) p 94
[Sil92] Silk J 1992 *Nature* **356** 741
[Sil93] Silk J and Wyse R F G 1993 *Phys. Rep.* **231** 293
[Sim81] Simpson J J 1981 *Phys. Rev.* D **24** 2971
[Sim84] Simpson J J *et al* 1984 *Phys. Rev. Lett.* **53** 141
[Sim85] Simpson J J 1985 *Phys. Rev. Lett.* **54** 1891
[Sim86] Simpson J J 1986 *Phys. Lett.* **174B** 113
[Sim89] Simpson J J and Hime A 1989 *Phys. Rev.* D **39** 1825
[Sin87] Sinclair D 1988 *Neutrino Physics* ed H V Klapdor and B Povh (Heidelberg: Springer) p 239
[Sin91] Sinclair D 1991 *Proc. NEUTRINO'90 (Geneva) Nucl. Phys.* B (Proc. Suppl.) **19** 100 (Amsterdam: North-Holland)
[Sir80] Sirlin A 1980 *Phys. Rev.* D **22** 971
[Sir87] Sirlin A 1987 *Phys. Rev.* D **35** 3423

[Smi57] Smith J H, Purcell E M and Ramsey N F 1957 *Phys. Rev.* **108** 120
[Smi85] Smith P F *et al* 1985 *Phys. Lett.* **153B** 188
[Smi86] Smith P F *et al* 1986 *Phys. Lett.* **171B** 129; 1986 *Phys. Lett.* **181B** 407
[Smi87] Smith P F *et al* 1987 *Phys. Lett.* **197B** 447
[Smi89] Smith P F 1989 *Ann. Rev. Nucl. Part. Sci.* **39** 73
[Smi90a] Smith P F *et al* 1990 *Phys. Lett.* **234B** 191
[Smi90b] Smith P F and Lewin J D 1990 *Phys. Rep.* **187** 203
[Smi95] Smith D 1996 *Proc. Int. Workshop on Double Beta Decay and Related Topics (Trento, Italy, 24 April–5 May 1995)* ed H V Klapdor-Kleingrothaus and S Stoica (Singapore: World Scientific)
[Smo92] Smoot G F *et al* 1992 *Astrophys. J.* **396** L1
[Sob86] Sobel H 1986 *Proc. VIth Moriond Workshop on Massive Neutrinos in Particle and Astrophysics* ed O Fackler and J Tran Thanh Van (Gif-sur-Yvette: Frontières) p 339
[Sol76] Solheim J E, Barnes III T G and Smith H J 1976 *Astrophys. J.* **209** 330
[Som50] Sommerfeld A 1950 *Lectures on Theoretical Physics* vol 2 (New York: Academic)
[Spe80] Spero R *et al* 1980 *Phys. Rev. Lett.* **44** 1645
[Spe85] Spergel D N and Press W H 1985 *Astrophys. J.* **294** 663; 1985 *Astrophys. J.* **296** 679
[Spe88a] Speake C C 1988 *Proc. 9th Workshop on Grand Unification (Aix-Les-Bains)* ed R Barloutaud (Singapore: World Scientific) p 101
[Spe88b] Speake C C and Quinn T J 1988 *Phys. Rev. Lett.* **61** 1340
[Spe90] Speake C C *et al* 1990 *Phys. Rev. Lett.* **65** 1967
[Spr87] Springer P T, Bennett C L and Baisden P A 1987 *Phys. Rev.* A **35** 679
[Sta78] Stacey F D 1978 *Geophys. Res. Lett.* **5** 377
[Sta81] Stacey F D *et al* 1981 *Phys. Rev.* D **23** 1683;
Stacey F D and Tuck G J 1981 *Nature* **292** 230
[Sta87] Stacey F D *et al* 1987 *Rev. Mod. Phys.* **59** 157
[Sta88] Stacey F D, Tuck G J and Moore G I 1988 *J. Geophys. Res.* **93** 10575
[Sta89] Staudt A, Bender E, Muto K and Klapdor H V 1989 *Z. Phys.* A **334** 47
[Sta90a] Staudt A, Muto K and Klapdor-Kleingrothaus H V 1990 *Europhys. Lett.* **13** 31
[Sta90b] Staudt A, Kuo T T S and Klapdor-Kleingrothaus H V 1990 *Phys. Lett.* **242B** 17
[Sta90c] Staudt A, Hirsch M, Muto K and Klapdor-Kleingrothaus H V 1990 *Phys. Rev. Lett.* **65** 1543
[Sta91] Staudt A, Muto K and Klapdor-Kleingrothaus H V 1991 *Phys. Lett.* **268B** 312
[Sta92a] Staudt A and Klapdor-Kleingrothaus H V 1992 *Nucl. Phys.* A **549** 254
[Sta92b] Staudt A, Kuo T T S and Klapdor-Kleingrothaus H V 1992 *Phys. Rev.* C **46** 871
[Sta92c] Starkman G D 1992 *Phys. Rev.* D **45** 476
[Ste76] Stevens C M, Schiffer J P and Chupka W A 1976 *Phys. Rev.* D **14** 716

[Ste77] Steinberg R I and Evans jr J C 1977 *Proc. NEUTRINO'77* (Moscow: Academy of Sciences of the USSR) p 321
[Ste80] Stech B 1980 *Unification of Fundamental Particle Interactions* ed J Ellis *et al* (New York: Plenum) p 23
[Ste88] Stefanov M A 1988 *JETP Lett.* **47** 1
[Ste91] Steinberger J 1991 *Phys. Rep.* **203** 345
[Sto84] Stone J L (ed) 1984 *Proc. Monopole 1983 (Ann Arbor), NATO Advanced Research Workshop vol 111* (New York: Plenum)
[Stü39] Stückelberg E C G 1939 *Helv. Phys. Acta* **11** 225, 229
[Stu87] Stubbs C W *et al* 1987 *Phys. Rev. Lett.* **58** 1070
[Stu89] Stubbs C W *et al* 1989 *Phys. Rev. Lett.* **62** 609
[Su94] Su Y *et al* 1994 *Phys. Rev.* D **50** 3614
[Suh88] Suhonen J, Taigel T and Faessler A 1988 *Nucl. Phys.* A **486** 91
[Suh93] Suhonen J and Civitarese O 1993 *Phys. Lett.* **312B** 367
[Sur91] Sur B *et al* 1991 *Phys. Rev. Lett.* **66** 2444
[Suz92] Suzuki Y 1992 *Proc. 4th Int. Workshop on Neutrino Telescopes (Venice)* ed M Baldo-Ceolin (Padova: Padova University) p 237
[Suz94] Suzuki Y 1994 *Proc. TAUP'93, Italy* ed C Arpesella *et al Nucl. Phys.* B (Proc. Suppl.) **35** 273 (Amsterdam: North-Holland)
[Tak84] Takasugi E 1984 *Phys. Lett.* **149B** 372
[Tak86] Takita M *et al* 1986 *Phys. Rev.* D **34** 902
[Tak95] Takasugi E 1996 *Preprint* OU-HET 215, June 1995; *Proc. Int. Workshop on Double Beta Decay and Related Topics (Trento, Italy, 24 April–5 May 1995)* ed H V Klapdor-Kleingrothaus and S Stoica (Singapore: World Scientific) p 165
[Tas65] Tassie L J 1965 *Nuovo Cimento* **38** 1935
[Tay83] Taylor G N *et al* 1983 *Phys. Rev.* D **28** 2705
[Tay86] Taylor R J and Roy Q J 1986 *Astron. Soc.* **27** 383
[Tay92] Taylor A N and Rowan-Robinson M 1992 *Nature* **359** 396
[Tel48] Teller E 1948 *Phys. Rev.* **73** 801
[Ter80] Terazawa H 1980 *Phys. Rev.* D **22** 184
[Thi87] Thieberger P 1987 *Phys. Rev. Lett.* **58** 1066
[Thi91] Thielemann F 1991 *Nuclei in the Cosmos* ed H Oberhummer (Heidelberg: Springer) p 147
['tHo71] 't Hooft G 1971 *Phys. Lett.* **37B** 195
[t'Ho72] 't Hooft G and Veltman M 1972 *Nucl. Phys.* B **50** 318
['tHo74] 't Hooft G 1974 *Nucl. Phys.* B **79** 276
['tHo76] 't Hooft G 1976 *Phys. Rev. Lett.* **37** 8; 1976 *Phys. Rev.* D **14** 3432
[Tho89a] Thompson D 1989 *Nucl. Instrum. Methods* A **284** 40
[Tho89b] Thomas J *et al* 1989 *Phys. Rev. Lett.* **63** 1902
[Tho90] Thomas J and Vogel P 1990 *Phys. Rev. Lett.* **65** 1173
[Thr92] Thron J L *et al* 1992 *Phys. Rev.* D **46** 4846
[Thr93] Thron J L 1993 *Proc. XXVI Int. Conf. on High Energy Physics (AIP Conf. Proc. 272)* (New York: AIP) p 1232
[Tom86] Tomoda T *et al* 1986 *Nucl. Phys.* A **452** 591

[Tom87] Tomoda T and Faessler A 1987 *Phys. Lett.* **199B** 475
[Tom88] Tomoda T 1988 *Nucl. Phys.* A **484** 635
[Tom91] Tomoda T 1991 *Rep. Prog. Phys.* **54** 53
[Ton93] Tonry J L 1993 *Ann. NY Acad. Sci.* **688** 113
[Tot92] Totsuka Y 1993 *Proc. NEUTRINO'92 (Granada) Nucl. Phys.* B (Proc. Suppl.) **31** 428 (Amsterdam: North-Holland)
[Tra90] Tran Thanh Van J (ed) 1990 *CP Violation in Particle Physics and Astrophysics* (Gif-sur-Yvette: Frontières)
[Tre91] Treichel M *et al* 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S193
[Tre95] Tretyak V I and Zdesenko Yu G 1995 *At. Data Nucl. Data Tables* **61** 43
[Tri87] Trimble V 1987 *Ann. Rev. Astron. Astrophys.* **25** 425
[Tri88] Trimble V 1988 *Rev. Mod. Phys.* **60** 859
[Tro83] Trowers W P 1983 *Acta Phys. Austr. Suppl.* **25** 101
[Tub80] Tubbs A D and Wolfe A M 1980 *Astrophys. J.* **236** L105
[Tup87] Tupper G, Danos M, Müller B and Rafelski J 1987 *Phys. Rev.* D **35** 394
[Tur76] Turneaure J P and Stein S R 1976 *Atomic Masses and Fundamental Constants* vol 5 ed J Sanders and A Wapstra (New York: Plenum) p 636
[Tur81] Turner M S 1981 *Proc. Int. Conf. on Neutrino Physics and Astrophysics, NEUTRINO'81 (Hawaii)* vol I ed R J Cence *et al* p 95
[Tur83] Turneaure J P *et al* 1983 *Phys. Rev.* D **27** 1705
[Tur84] Turner M S, Steigman G and Krauss L M 1984 *Phys. Rev. Lett.* **52** 2090
[Tur88] Turck-Chièze S, Cahen S, Cassé M and Doom C 1988 *Astrophys. J.* **268** 415
[Tur89a] Turok N 1989 *Phys. Rev. Lett.* **63** 2625
[Tur89b] Turkevich A L *et al* 1989 *Proc. Workshop on Fundamental Symmetry and Nuclear Structure (Santa Fe, 1988)* ed J N Ginocchio and S P Rosen (Singapore: World Scientific) p 86
[Tur90] Turner M S 1990 *Phys. Rep.* **197** 67
[Tur91] Turkevich A L, Economou T E and Cowan G A 1991 *Phys. Rev. Lett.* **67** 3211
[Tur92] Turner M S 1992 *Phys. Rev.* D **45** 1066
[Tur93a] Turck-Chièze S *et al* 1993 *Phys. Rep.* **230** 57
[Tur93b] Turck-Chiéze S and Lopes I 1993 *Astrophys. J.* **408** 347
[Uns81] Unsöld A and Baschek B 1981 *Der neue Kosmos* (Heidelberg: Springer)
[Ush81] Ushida N *et al* 1981 *Phys. Rev. Lett.* **47** 1694
[Val86] Valle J W F 1986 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer)
[Val87] Valle J W F 1987 *Phys. Lett.* **186B** 73; **196B** 157
[Val90] Valentijn E A 1990 *Nature* **346** 153
[Val93] Valle J W F 1993 *Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei, WEIN'92 (Dubna, Russia)* (Singapore: World Scientific) p 131
[Val94] Valle J W F 1994 *Nucl. Phys.* B (Proc. Suppl.) **35** 309
[Van68] Vant-Hull L L 1968 *Phys. Rev.* **173** 1412
[Van82] van den Bergh S 1982 *Nature* **299** 297
[Vau86] de Vaucouleur G and Corwin H G 1986 *Astrophys. J.* **308** 487
[Vas90] Vasenko A A *et al* 1990 *Mod. Phys. Lett.* A **5** 1299

[Ver87] Vergados J D 1987 *Phys. Lett.* **184B** 55
[Vid94] Vidyakin G S *et al* 1994 *JETP Lett.* **59** 364
[Vil87] Vilenkin A 1987 *Sci. Am.* **12** 52
[Vil88] Vilenkin A 1988 *Spektrum der Wissenschaft* **2** 94
[Vog81] Vogel P, Schenter G K, Mann F M and Schenter R E 1981 *Phys. Rev.* C **24** 1543
[Vog84] Vogel P 1984 *Phys. Rev.* D **30** 1505
[Vog86] Vogel P and Zirnbauer M R 1986 *Phys. Rev. Lett.* **57** 3148
[Vog95] Vogel P 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 204
[Vol84] Vold T G, Raab F J, Heckel B R and Fortson E N 1984 *Phys. Rev. Lett.* **52** 2229
[Vol86] Voloshin M B, Vysotskii M I and Okun L B 1986 *Sov. J. Nucl. Phys.* **44** 440; 1986 *Sov. Phys.–JETP* **64** 446
[Vui82] Vuilleumier J-L *et al* 1982 *Phys. Lett.* **114B** 298
[Vui92a] Vuilleumier J-C *et al* 1992 *Proc. XIIth Moriond Workshop on Massive Neutrinos (Les Arcs)*
[Vui92b] Vuilleumier J-L *et al* 1992 private communication
[Vui93] Vuilleumier J-C *et al* 1993 *Phys. Rev.* D **48** 1009
[Vyl84] Vylov Ts *et al* 1984 *Isv. Akad. Nauk Ser. Fiz.* **48** 1809
[Wad84] Wada, Yamashita and Yamamoto 1984 *Lett. Nuovo Cimento* **40** 329
[Wad88] Wada, Yamashita and Yamamoto 1988 *Nuovo Cimento* C **11** 229
[Wal84] Walsh T F, Weisz P and Wu T T 1984 *Nucl. Phys.* B **232** 349
[Wal91] Walker T P *et al* 1991 *Astrophys. J.* **376** 51
[Wan84] Wang K C *et al* 1984 *NEUTRINO'84, Proc. 11th Conf. on Neutrino Physics and Astrophysics (Nordkirchen)* ed E K Kleinknecht and E A Paschos (Singapore: World Scientific) p 177
[Wap85] Wapstra A H and Audi G 1985 *Nucl. Phys.* A **432** 1
[Wei35] von Weizsäcker C F 1935 *Z. Phys.* **96** 431
[Wei67] Weinberg S 1967 *Phys. Rev. Lett.* **19** 1264
[Wei72] Weinberg S 1972 *Gravitation and Cosmology* (New York: Wiley)
[Wei77] Weinberg S 1977 *The First Three Minutes* (Thun: Deutsch)
[Wei78] Weinberg S 1978 *Phys. Rev. Lett.* **40** 223
[Wei79a] Weinberg S 1979 *Phys. Rev. Lett.* **42** 850
[Wei79b] Weinberg S 1979 *Phys. Rev. Lett.* **43** 1566
[Wei83] Weinberg S 1983 *Phys. Lett.* **125B** 265
[Wei89] Weinberg S 1989 *Rev. Mod. Phys.* **61** 1
[Wei93] Weinheimer Ch *et al* 1993 *Phys. Lett.* **300B** 210
[Wes74] Wess J and Zumino B 1974 *Phys. Lett.* **49B** 52; 1974 *Nucl. Phys.* B **70** 39
[Wes80] Wesson P 1980 *Phys. Today* **33**(7) 32
[Wet86] Wetterich C 1986 *Phys. Lett.* **167B** 325
[Wey29] Weyl H 1929 *Z. Phys.* **56** 330
[Wig49] Wigner E P 1949 *Proc. Am. Phil. Soc.* **93** 521
[Wil78] Wilczek F 1978 *Phys. Rev. Lett.* **40** 279
[Wil79] Wilczek F and Zee A 1979 *Phys. Rev. Lett.* **43** 1571
[Wil80] Willis S E *et al* 1980 *Phys. Rev. Lett.* **44** 522

[Wil87] Wilkerson J F *et al* 1987 *Phys. Rev. Lett.* **58** 2023
[Wil92] Wille K 1992 *Physik der Teilchenbeschleuniger und Synchrotronstrahlquellen* (Stuttgart: Teubner)
[Wil93] Wilczek F 1993 *Ann. NY Acad. Sci.* **688** 94
[Win52] Winter R G 1952 *Phys. Rev.* **85** 687
[Win88] Winter K 1988 *Phys. Bl.* **44** 73
[Win95] Winter K 1995 *Nucl. Phys.* B (Proc. Suppl.) **38** 211
[Wit81] Witten E 1981 *Nucl. Phys.* B **186** 412
[Wit84] Witten E 1984 *Phys. Rev.* D **30** 272
[Wla59] Wlassow N A 1959 *Neutronen* (Berlin: VEB Deutscher Verlag der Wissenschaften)
[Wol64] Wolfenstein L 1964 *Phys. Rev. Lett.* **13** 562
[Wol78] Wolfenstein L 1978 *Phys. Rev.* D **17** 2369
[Wol79a] Wolfenstein L 1979 *Phys. Rev.* D **20** 2634
[Wol79b] Wolfram S 1979 *Phys. Lett.* **82B** 65
[Wol81] Wolfenstein C 1981 *Phys. Lett.* **107B** 77
[Won91] Wong H T *et al* 1991 *Phys. Rev. Lett.* **67** 1218
[Won92] Wong H T *et al* 1992 *Vortrag auf der Tagung der Deutschen Physikalischen Gesellschaft (DPG)*
[Wri92] Wright E L *et al* 1992 *Astrophys. J.* **396** L13
[Wu57] Wu S *et al* 1957 *Phys. Rev.* **105** 1413
[Wu66] Wu C S and Moszkowski S A 1966 *Beta Decay* (New York: Interscience)
[Wu86] Wu Y-S and Wang Z 1986 *Phys. Rev. Lett.* **57** 1978
[Wu91] Wu X R, Staudt A, Klapdor-Kleingrothaus H V, Ching C R and Ho T H 1991 *Phys. Lett.* **272B** 169
[Wu92] Wu X R, Staudt A, Kuo T T S, Klapdor-Kleingrothaus H V, Ching C R and Ho T H 1992 *Phys. Lett.* **276B** 272
[Wue89] Wuensch W U *et al* 1989 *Phys. Rev.* D **40** 3153
[Yan79a] Yang J, Schramm D N, Steigman G and Rood R T 1979 *Astrophys. J.* **227** 697
[Yan79b] Yanagida T 1979 *Proc. Workshop on Unified Theory and the Baryon Number of the Universe KEK* ed O Sawada and A Sugimoto
[Yan84] Yang J *et al* 1984 *Astrophys. J.* **281** 493
[Yas83] Yasumi S *et al* 1983 *Phys. Lett.* **122B** 461
[Yas86] Yasumi S *et al* 1986 *Phys. Lett.* **181B** 169
[You58] Young D A 1958 *Nature* **182** 375
[You91] Ke You *et al* 1991 *Phys. Lett.* **265B** 53
[Yuk35] Yukawa H 1935 *Proc. Phys. Math. Soc. Japan* **17** 48
[Zac85] Zacek V *et al* 1985 *Phys. Lett.* **164B** 193
[Zac86a] Zacek G *et al* 1986 *Phys. Rev.* D **34** 2621
[Zac86b] Zacek V 1987 *Physik in unserer Zeit* **4** 114
[Zac86c] Zacek V 1986 *Proc. Int. Symp. on Weak and Electroweak Interactions in Nuclei, WEIN'86 (Heidelberg)* ed H V Klapdor (Heidelberg: Springer) p 750
[Zde81] Zdesenko Yu *et al* 1981 *Izv. Akad. Nauk SSSR Ser. Fiz.* **45** 1856

[Zde91] Zdesenko Yu 1991 *J. Phys. G: Nucl. Part. Phys.* **17** S243
[Zde93] Zdesenko Yu 1993 private communication
[Zel68] Zel'dovich Ya B 1968 *Sov. Phys.–Usp.* **11** 381
[Zel83] Zel'dovich Ya B and Novikov I D 1983 *The Structure and Evolution of the Universe* (Chicago: University of Chicago Press)
[Zer92] Zerwas P M (ed) 1992 *Proc. e^+e^- Collisions at 500 GeV: the Physics Potential* DESY-Report 92-123 A+B
[Zer93] Zerwas P M 1993 *Phys. Bl.* **49** 187
[Zhi80] Zhitnitsky A R 1980 *Sov. J. Nucl. Phys.* **31** 260
[Zli91] Zlimen I *et al* 1991 *Phys. Rev. Lett.* **67** 560
[Zub93] Zuber K and Klapdor-Kleingrothaus H V 1993 *Phys. Bl.* **49** 125
[Zwe64] Zweig G 1964 CERN 8182 - TH401; CERN 8419 - TH412