Particle Physics

THIRD EDITION

B.R. Martin and G. Shaw

WILEY

SUGGESTED SHORT COURSE

A brief overview is given in Chapters 1–3. For a short course, follow these Chapters plus 4 and 5, Sections 6.1, 6.2, 6.3, 7.1, 7.2, 10.1 and 10.2 and Chapters 8 and 9, omitting all 'starred' material throughout.

The starred items contain more specialized or advanced material that is not required elsewhere in the book. They may be omitted at a first reading and used for reference when you want to know about a particular topic.

This book has a website: www.hep.manchester.ac.uk/u/graham/ppbook.html. Any misprints or other necessary corrections brought to our attention will be listed on this page. We would also be grateful for any other comments about the book.

The pictures on the front cover show computer simulations of the detection of a Higgs boson in the ATLAS experiment (large circle) and CMS experiment (small circle), both at the CERN Laboratory in Geneva. (Reproduced by permission of CERN.)

PARTICLE PHYSICS

Third Edition

B. R. Martin
Department of Physics and Astronomy
University College London

G. Shaw
School of Physics & Astronomy
University of Manchester

A John Wiley and Sons, Ltd, Publication

Particle Physics

Third Edition

The Manchester Physics Series

General Editors
F. K. LOEBINGER: F. MANDL: D. J. SANDIFORD

Department of Physics and Astronomy, Faculty of Science,
University of Manchester

Properties of Matter:	B. H. Flowers and E. Mendoza
Optics: *Second Edition*	F. G Smith and J. H. Thomson
Statistical Physics: *Second Edition*	F. Mandl
Electromagnetism: *Second Edition*	I. S. Grant and W. R. Philips
Statistics:	R. J. Barlow
Solid State Physics: *Second Edition*	J. R. Hook and H. E. Hall
Quantum Mechanics:	F. Mandl
Particle Physics: *Third Edition*	B. R. Martin and G. Shaw
The Physics of Stars: *Second Edition*	A. C. Phillips
Computing for Scientists:	R. J. Barlow and A. R. Barnett

PARTICLE PHYSICS

Third Edition

B. R. Martin
Department of Physics and Astronomy
University College London

G. Shaw
School of Physics & Astronomy
University of Manchester

A John Wiley and Sons, Ltd, Publication

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex,
PO19 8SQ, United Kingdom

Library of Congress Cataloging in Publication Data

Martin, B. R. (Brian Robert)
Particle physics / B. R. Martin, G. Shaw. — 3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-03293-0 (cloth) — ISBN 978-0-470-03294-7 (pbk.)
1. Particles (Nuclear physics) I. Shaw, G. (Graham), 1942– II. Title.
QC793.2.M38 2008
539.7′2—dc22
2008024775

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

ISBN 978-0-470-03293-0 (HB)
ISBN 978-0-470-03294-7 (PB)

Set in 10/12pt Times by Integra Software Services Pvt. Ltd, Pondicherry, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

‘But why are such terrific efforts made just to find new particles?’ asked Mr Tompkins.

‘Well, this is science,’ replied the professor, ‘the attempt of the human mind to understand everything around us, be it giant stellar galaxies, microscopic bacteria, or these elementary particles. It is interesting and exciting, and that is why we are doing it.’

From *Mr Tompkins Tastes a Japanese Meal,* by George Gamow (*Mr Tompkins in Paperback*, Cambridge University Press (1965), p.186).

Contents

★ Starred sections may be omitted on a first reading as they are not required later in the book.

Editors' Preface to the Manchester Physics Series

The Manchester Physics Series is a series of textbooks at first degree level. It grew out of our experience at the University of Manchester, widely shared elsewhere, that many textbooks contain much more material than can be accommodated in a typical undergraduate course; and that this material is only rarely so arranged as to allow the definition of a short self-contained course. In planning these books we have had two objectives. One was to produce short books so that lecturers would find them attractive for undergraduate courses, and so that students would not be frightened off by their encyclopaedic size or price. To achieve this, we have been very selective in the choice of topics, with the emphasis on the basic physics together with some instructive, stimulating and useful applications. Our second objective was to produce books which allow courses of different lengths and difficulty to be selected with emphasis on different applications. To achieve such flexibility we have encouraged authors to use flow diagrams showing the logical connections between different chapters and to put some topics in starred sections. These cover more advanced and alternative material which is not required for the understanding of latter parts of each volume.

Although these books were conceived as a series, each of them is self-contained and can be used independently of the others. Several of them are suitable for wider use in other sciences. Each Author's Preface gives details about the level, prerequisites, etc., of that volume.

The Manchester Physics Series has been very successful since its inception 40 years ago, with total sales of more than a quarter of a million copies. We are extremely grateful to the many students and colleagues, at Manchester and elsewhere, for helpful criticisms and stimulating comments. Our particular thanks go to the authors for all the work they have done, for the many new ideas they have contributed, and for discussing patiently, and often accepting, the suggestions of the editors.

Finally we would like to thank our publishers, John Wiley & Sons, Ltd., for their enthusiastic and continued commitment to the Manchester Physics Series.

F. K. Loebinger
F. Mandl
D. J. Sandiford
August 2008

Authors' Preface

Particle Physics is the study of the fundamental constituents of matter and the forces between them. For the past 25 years, these have been described by the so-called *standard model* of particle physics, which provides, at least in principle, a basis for understanding most particle interactions, except gravity. The purpose of this book is to provide a short introduction to particle physics, which emphasizes the foundations of the standard model in experimental data, rather than its more formal and theoretical aspects. The book is intended for undergraduate students who have previously taken introductory courses in nonrelativistic quantum mechanics and special relativity. No prior knowledge of particle physics is assumed.

The introductory nature of the book, and the need to keep it reasonably short, have influenced both the level of the treatment and the choice of material. We have tried to take a direct approach, and while we have made many historical comments and asides, we have not felt bound by the historical development of ideas. Also we have not given any detailed theoretical calculations involving the mathematical formalism of quantum field theory, since these are well beyond the scope of a typical undergraduate course. Rather, we have focused on the interpretation of experimental data in terms of the basic properties of quarks and leptons, and extensive use has been made of symmetry principles and Feynman diagrams, which are introduced early in the book.

The structure of the book is simple. The first chapter gives a very brief overview of the subject and an introduction to some of the basic ideas that are used extensively throughout the rest of the book. This is followed by two chapters that introduce the basic entities of particle physics – quarks, leptons and hadrons – and their interactions. The remaining chapters discuss a wide selection of important topics in more detail. These include experimental methods, space–time symmetries, the quark model of hadrons, quantum chromodynamics and jet physics, the weak interaction, including its unification with the electromagnetic interaction and *CP* violation and other related symmetries, and a brief account of some of the important open questions 'beyond the standard model' that are currently being investigated in laboratories around the world.

Since publication of the Second Edition of this book, there has been substantial progress in particle physics. This includes, among other things, the discovery of neutrino mixing and nonzero neutrino masses; new results on heavy quark states, which have greatly increased our understanding of quark mixing and *CP* violation; rapid growth in the field of particle astrophysics and cosmology; and new developments in experimental methods as the LHC begins to explore a new energy frontier. Our main

aim in producing this new edition is again to bring the book thoroughly up to date in the light of these and other new developments, while leaving its basic philosophy unchanged.

Finally, while preparing this book we have benefited greatly from discussions with colleagues too numerous to mention individually. We are grateful to them all.

B. R. Martin
G. Shaw
April 2008

Notes

References

References are referred to in the text by giving the name of the author(s) and date of publication. A list of references with full publication details is given at the end of the book.

Data

A short list of particle data is given in Appendix E. A comprehensive compilation may be obtained from the biannual publications of the Particle Data Group (PDG). The 2006 edition of their definitive 'Review of particle properties' is in W.-M. Yao *et al*., *Journal of Physics*, **G33** (2006), 1. This also contains useful topical reviews of specific topics in particle physics. The PDG Review is available at http://pdg.lbl.gov and this site also contains links to other sites where compilations of particle data may be found. It is important that students have some familiarity with these data compilations.

Problems

Problems are provided for all chapters and appendices (except Appendices E and F). They are an integral part of the text. The problems are often numerical and require values of physical constants that are given on the inside back cover. A few also require input data that may be found in Appendix E and the references given above. Solutions to the problems are given in Appendix F.

Illustrations

Some illustrations in the text have been adapted from diagrams that have been published elsewhere. In a few cases they have been reproduced exactly as previously published. We acknowledge, with thanks, permission from the relevant copyright holders to use such illustrations and this is confirmed in the captions.

1

Some Basic Concepts

1.1 INTRODUCTION

Particle physics is the study of the fundamental constituents of matter and their interactions. However, which particles are regarded as fundamental has changed with time as physicists' knowledge has improved. Modern theory – called the *standard model* – attempts to explain all the phenomena of particle physics in terms of the properties and interactions of a small number of particles of three distinct types: two spin–$\frac{1}{2}$ families of fermions called *leptons* and *quarks*, and one family of spin-1 bosons – called *gauge bosons* – which act as 'force carriers' in the theory. In addition, at least one spin-0 particle, called the *Higgs boson*, is postulated to explain the origin of mass within the theory, since without it all the particles in the model are predicted to have zero mass. All the particles of the standard model are assumed to be *elementary*; i.e. they are treated as point particles, without internal structure or excited states.

The most familiar example of a lepton is the *electron* e^- (the superscript denotes the electric charge), which is bound in atoms by the *electromagnetic interaction,* one of the four fundamental forces of nature. A second well-known lepton is the *electron neutrino* ν_e, which is a light, neutral particle observed in the decay products of some unstable nuclei (the so-called β-decays). The force responsible for the β-decay of nuclei is called the *weak interaction.*

Another class of particles called *hadrons* is also observed in nature. Familiar examples are the neutron *n* and proton *p* (collectively called *nucleons*) and the *pions* (π^+, π^-, π^0), where the superscripts again denote the electric charges. These are not elementary particles, but are made of quarks bound together by a third force of

Particle Physics Third Edition B. R. Martin and G. Shaw

nature, the *strong interaction.* The theory is unusual in that the quarks themselves are not directly observable, only their bound states. Nevertheless, we shall see in later chapters that there is overwhelming evidence for the existence of quarks and we shall discuss the reason why they are unobservable as free particles. The strong interaction between quarks gives rise to the observed strong interaction between hadrons, such as the nuclear force that binds nucleons into nuclei. There is an analogy here with the fundamental electromagnetic interaction between electrons and nuclei that also gives rise to the more complicated forces between their bound states, i.e. between atoms.

In addition to the strong, weak and electromagnetic interactions between quarks and leptons, there is a fourth force of nature – gravity. However, the gravitational interaction between elementary particles is so small that it can be neglected at presently accessible energies. Because of this, we will often refer in practice to the *three* forces of nature. The standard model also specifies the origin of these forces. Consider, firstly, the electromagnetic interaction. In classical physics this is propagated by electromagnetic waves, which are continuously emitted and absorbed. While this is an adequate description at long distances, at short distances the quantum nature of the interaction must be taken into account. In quantum theory, the interaction is transmitted discontinuously by the exchange of spin-1 photons, which are the 'force carriers', or gauge bosons, of the electromagnetic interaction and, as we shall see presently, the long-range nature of the force is related to the fact that photons have zero mass. The use of the word 'gauge' refers to the fact that the electromagnetic interaction possesses a fundamental symmetry called *gauge invariance*. This property is common to all three interactions of nature and has profound consequences, as we shall see.

The weak and strong interactions are also associated with the exchange of spin-1 particles. For the weak interaction, they are called W and Z *bosons,* with masses about 80–90 times the mass of the proton. The resulting force is very short range, and in many applications may be approximated by an interaction at a point. The equivalent particles for the strong interaction are called *gluons g.* There are eight gluons, all of which have zero mass and are electrically neutral, like the photon. Thus, by analogy with electromagnetism, the basic strong interaction between quarks is long range. The 'residual' strong interaction between the quark bound states (hadrons) is not the same as the fundamental strong interaction between quarks (but is a consequence of it) and is short range, again as we shall see later.

In the standard model, which will play a central role in this book, the main actors are the leptons and quarks, which are the basic constituents of matter; and the 'force carriers' (the photon, the W and Z bosons, and the gluons) that mediate the interactions between them. In addition, because not all these particles are directly observable, the quark bound states (i.e. hadrons) will also play a very important role.

In particle physics, high energies are needed both to create new particles and to explore the structure of hadrons. The latter requires projectiles whose wavelengths λ are at least as small as hadron radii, which are of order 10^{-15} m. It follows that their momenta, $p = h/\lambda$, and hence their energies, must be several hundred MeV/c ($1\,\text{MeV} = 10^6\,\text{eV}$). Because of this, any theory of elementary particles must combine the requirements of both special relativity and quantum theory. This has startling consequences, as we shall now show.

1.2 ANTIPARTICLES

For every charged particle of nature, whether it is one of the elementary particles of the standard model, or a hadron, there is an associated particle of the same mass, but opposite charge, called its *antiparticle*. This result is a necessary consequence of combining special relativity with quantum mechanics. This important theoretical prediction was made by Dirac and follows from the solutions of the equation he first wrote down to describe relativistic electrons. Thus we start by considering how to construct a relativistic wave equation.

1.2.1 Relativistic wave equations

We start from the assumption that a particle moving with momentum $\mathbf{p}$ in free space is described by a de Broglie wavefunction[1]

$$\Psi(\mathbf{r},t) = N\, e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}, \tag{1.1}$$

with frequency $\nu = E/h$ and wavelength $\lambda = h/p$. Here $p \equiv |\mathbf{p}|$ and N is a normalization constant that is irrelevant in what follows. The corresponding wave equation depends on the assumed relation between the energy E and momentum $\mathbf{p}$. Nonrelativistically,

$$E = p^2/2m \tag{1.2}$$

and the wavefunction (1.1) obeys the nonrelativistic Schrödinger equation

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t). \tag{1.3}$$

Relativistically, however,

$$E^2 = p^2c^2 + m^2c^4, \tag{1.4}$$

where m is the rest mass,[2] and the corresponding wave equation is

$$-\hbar^2\frac{\partial^2\Psi(\mathbf{r},t)}{\partial t^2} = -\hbar^2c^2\nabla^2\Psi(\mathbf{r},t) + m^2c^4\Psi(\mathbf{r},t), \tag{1.5}$$

as is easily checked by substituting (1.1) into (1.5) and using (1.4). This equation was first proposed by de Broglie in 1924, but is now more usually called the *Klein–Gordon equation*.[3] Its most striking feature is the existence of solutions with negative energy. For every plane wave solution of the form

$$\Psi(\mathbf{r},t) = N\exp\left[i(\mathbf{p}\cdot\mathbf{r} - E_pt)/\hbar\right], \tag{1.6a}$$

[1] We use the notation $\mathbf{r} = (x_1, x_2, x_3) = (x, y, z)$.
[2] From now on, the word *mass* will be used to mean the rest mass.
[3] These authors incorporated electromagnetic interactions into the equation, in a form now known to be appropriate for charged spin-0 bosons.

with momentum **p** and positive energy

$$E = E_p \equiv +(p^2c^2 + m^2c^4)^{1/2} \geq mc^2$$

there is also a solution

$$\tilde{\Psi}(\mathbf{r},t) \equiv \Psi^*(\mathbf{r},t) = N^* \exp\left[i(-\mathbf{p}\cdot\mathbf{r} + E_p t)/\hbar\right], \tag{1.6b}$$

corresponding to momentum –**p** and negative energy

$$E = -E_p = -(p^2c^2 + m^2c^4)^{1/2} \leq -mc^2.$$

Other problems also occur, indicating that the Klein–Gordon equation is not, in itself, a sufficient foundation for relativistic quantum mechanics. In particular, it does not guarantee the existence of a positive-definite probability density for position.[4]

The existence of negative energy solutions is a direct consequence of the quadratic nature of the mass–energy relation (1.4) and cannot be avoided in a relativistic theory. However, *for spin-$\frac{1}{2}$ particles* the other problems were resolved by Dirac in 1928, who looked for an equation of the familiar form

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = H(\mathbf{r},\hat{\mathbf{p}})\Psi(\mathbf{r},t), \tag{1.7}$$

where H is the Hamiltonian and $\hat{\mathbf{p}} = -i\hbar\nabla$ is the momentum operator. Since (1.7) is first order in $\partial/\partial t$, Lorentz invariance requires that it also be first order in spatial derivatives. Dirac therefore proposed a Hamiltonian of the general form

$$H = -i\hbar c\sum_{i=1}^{3}\alpha_i\frac{\partial}{\partial x_i} + \beta mc^2 = c\,\boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + \beta mc^2, \tag{1.8}$$

in which the coefficients β and $\alpha_i (i = 1, 2, 3)$ are determined by requiring that solutions of the Dirac equation (1.8) are also solutions of the Klein–Gordon equation (1.5). Acting on (1.7) with $i\hbar\partial/\partial t$ and comparing with (1.5) leads to the conclusion that this is true if, and only if,

$$\alpha_i^2 = 1, \quad \beta^2 = 1, \tag{1.9a}$$

$$\alpha_i\beta + \beta\alpha_i = 0 \tag{1.9b}$$

and

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 0 \quad (i \neq j). \tag{1.9c}$$

These relations cannot be satisfied by ordinary numbers and the simplest assumption is that β and $\alpha_i (i = 1, 2, 3)$ are matrices, which must be Hermitian so that the

[4] For a discussion of this point, see pp. 467–468 of Schiff (1968).

Hamiltonian is Hermitian. The smallest matrices satisfying these requirements have dimensions 4×4 and are given in many books,[5] but are not required below. We thus arrive at an interpretation of the Dirac equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi = -i\hbar c \sum_i \alpha_i \frac{\partial \Psi}{\partial x_i} + \beta mc^2 \Psi, \tag{1.10}$$

as a four-dimensional matrix equation in which the Ψ are four-component wavefunctions

$$\Psi(\mathbf{r}, t) = \begin{pmatrix} \Psi_1(\mathbf{r}, t) \\ \Psi_2(\mathbf{r}, t) \\ \Psi_3(\mathbf{r}, t) \\ \Psi_4(\mathbf{r}, t) \end{pmatrix}, \tag{1.11}$$

called *spinors.* Plane wave solutions take the form

$$\Psi(\mathbf{r}, t) = u(\mathbf{p}) \exp\left[i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar\right], \tag{1.12}$$

where $u(\mathbf{p})$ is also a four-component spinor satisfying the eigenvalue equation

$$H_{\mathbf{p}} u(\mathbf{p}) \equiv (c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2) u(\mathbf{p}) = Eu(\mathbf{p}), \tag{1.13}$$

obtained by substituting (1.11) into (1.10). This equation has four solutions:[6] two with positive energy $E = +E_p$ corresponding to the two possible spin states of a spin-$\frac{1}{2}$ particle (called 'spin up' and 'spin down', respectively) and two corresponding negative energy solutions with $E = -E_p$.

The problem of the negative-energy solutions will be resolved in the next section. Here we note that the positive-energy solutions of the Dirac equation lead to many predictions that have been verified experimentally to a very high precision. Notable among these are relativistic corrections in atomic spectroscopy, including spin–orbit effects, and the prediction that point-like spin-$\frac{1}{2}$ particles of mass m and charge q have a Dirac magnetic moment

$$\boldsymbol{\mu}_D = q\mathbf{S}/m, \tag{1.14}$$

where $\mathbf{S}$ is the spin vector. This is a key result. It not only yields the correct value for the electron, but provides a simple test for the point-like nature of any other spin-$\frac{1}{2}$ fermion. For the proton and neutron, the experimental values are

$$\boldsymbol{\mu}_p = 2.79e\mathbf{S}/m_p \quad \text{and} \quad \boldsymbol{\mu}_n = -1.91\, e\,\mathbf{S}/m_n, \tag{1.15}$$

[5] See, for example, pp. 473–475 of Schiff (1968).
[6] A proof of these results is given in, for example, pp. 475–477 of Schiff (1968).

in disagreement with Equation (1.14). Historically, the measurement of the proton magnetic moment by Frisch and Stern in 1933 was the first indication that the proton was not a point-like elementary particle.

1.2.2 Hole theory and the positron

The problem of the negative energy states remains. They cannot be ignored, since their existence leads to unacceptable consequences. For example, if such states are unoccupied, then transitions from positive to negative energy states could occur, leading to the prediction that atoms such as hydrogen would be unstable. This problem was resolved by Dirac, who postulated that the negative energy states are almost always filled. For definiteness consider the case of electrons. Since they are fermions, they obey the Pauli exclusion principle, and the Dirac picture of the vacuum is a so-called 'sea' of negative energy states, each with two electrons (one with spin 'up' and the other with spin 'down'), while the positive energy states are all unoccupied (see Figure 1.1). This state is indistinguishable from the usual vacuum with $E_V = 0$, $\mathbf{p}_V = \mathbf{0}$, etc. This is because for each state of momentum $\mathbf{p}$ there is a corresponding state with momentum $-\mathbf{p}$, so that the momentum of the vacuum $\mathbf{p}_V = \Sigma\, \mathbf{p} = \mathbf{0}$. The same argument applies to spin, while, since energies are measured relative to the vacuum, $E_V \equiv 0$ by definition. Similarly, we may define the charge $Q_V \equiv 0$, because the constant electrostatic potential produced by the negative energy sea is unobservable. Thus this state has all the *measurable* characteristics of the naive vacuum and the 'sea' is unobservable.

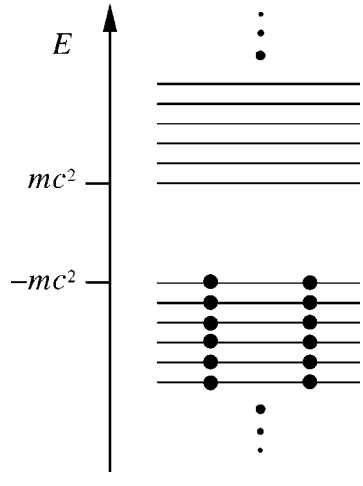


Figure 1.1 Dirac picture of the vacuum. The sea of negative energy states is totally occupied with two electrons in each level, one with spin 'up' and one with spin 'down'. The positive energy states are all unoccupied.

Dirac's postulate solves the problem of unacceptable transitions from positive energy states, but has other consequences. Consider what happens when an electron is added to, or removed from, the vacuum. In the former case, the electron is confined to

the positive energy region since all the negative energy states are occupied. In the latter case, *removing* a negative energy electron with $E = -E_p < 0$, momentum –**p**, spin –**S** and charge $-e$ from the vacuum (which has $E_V = 0, \mathbf{p}_V = 0, \mathbf{S}_V = 0, Q_V = 0$) leaves a state (the sea with a 'hole' in it) with positive energy $E = E_p > 0$, momentum **p**, spin **S** and charge $+e$. This state cannot be distinguished by any measurement from a state formed by *adding* to the vacuum a particle with momentum **p**, energy $E = E_p > 0$, spin **S** and charge $+e$. The two cases are equivalent descriptions of the same phenomena. Using the latter, Dirac predicted the existence of a spin-$\frac{1}{2}$ particle e^+ with the same mass as the electron, but opposite charge. This particle is called the *positron* and is referred to as the *antiparticle* of the electron.[7]

The positron was subsequently discovered by Anderson and by Blackett and Ochialini in 1933. The discovery was made using a device of great historical importance, called a *cloud chamber*. When a charged particle passes through matter, it interacts with it, losing energy. This energy can take the form of radiation or of excitation and ionization of the atoms along the path.[8] It is the aim of *track chambers* – of which the cloud chamber is the earliest example – to produce a visible record of this trail and hence of the particle that produced it.

The cloud chamber was devised by C. T. R. Wilson, who noticed that the condensation of water vapour into droplets goes much faster in the presence of ions. It consisted of a vessel filled with air almost saturated with water vapour and fitted with an expansion piston. When the vessel was suddenly expanded, the air cooled and became supersaturated, and droplets were formed preferentially along the trails of ions left by charged particles passing through the chamber. The chamber was illuminated by a flash of light immediately after expansion, and the tracks of droplets so revealed were photographed before they had time to disperse.

Figure 1.2 shows one of the first identified positrons tracks observed by Anderson in 1933. The band across the centre of the picture is a 6 mm lead plate inserted to slow particles down. The track is curved due to the presence of a 1.5 T applied magnetic field **B**, and since the curvature of such tracks increases with decreasing momentum, we can conclude that the particle enters at the bottom of the picture and travels upwards. The sign of the particle's charge q then follows from the direction of the Lorentz force $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, where **v** is the particle's velocity, and hence of the curvature; it is positive.

That the particle is a positron and not a proton follows essentially from the range of the upper track. The rate of energy loss of a charge particle in matter depends on its charge and velocity. (This will be discussed in Chapter 4.) From the curvature of the tracks, one can deduce that the momentum of the upper track is 23 MeV/c, corresponding to either a slow moving proton $\upsilon \ll c$ or a relativistic ($\upsilon \approx c$) positron. The former would lose energy rapidly, coming to rest in a distance of about 5 mm, comparable with the thickness of the lead plate. The observed track length is more than 5 cm, enabling a limit $m_+ \leq 20m_e \ll m_p$ to be set on the mass m_+ of the particle, which

[7] This prediction, which is now regarded as one of the greatest successes of theoretical physics, was not always so enthusiastically received at the time. For example, in 1933 Pauli wrote: 'Dirac has tried to identify holes with antielectrons. We do not believe that this explanation can be seriously considered'. (Pauli, 1933)

[8] This will be discussed in detail in Chapter 4.

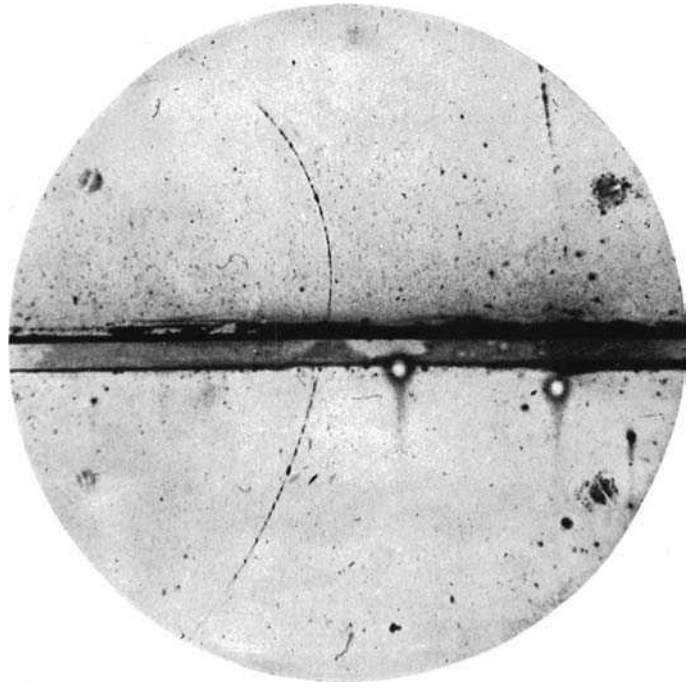


Figure 1.2 One of the first positron tracks observed by Anderson in a Wilson cloud chamber. The band across the centre of the picture is a lead plate, inserted to slow down particles. The positive sign of the electric charge and the particle's momentum are deduced from the curvature of the tracks in the applied magnetic field. That it is a positron follows from the long range of the upper track. (Reprinted Figure 1 with permission from C. D. Anderson, *Phys. Rev.*, **43**, 491. Copyright 1933 by American Physical Society.)

Anderson suggested was a positron. Many other examples were found, especially by Blackett and Ochialini, and by 1934 Blackett, Ochialini and Chadwick had established that $m_+ = m_e$ within experimental errors of order 10 %. The interpretation of the light positive particles as positrons was thus established beyond all reasonable doubt.

The Dirac equation applies to any spin-$\frac{1}{2}$ particle, and hole theory predicts that *all* charged spin-$\frac{1}{2}$ particles, whether they are elementary or hadrons, have distinct antiparticles with opposite charge, but the same mass. The argument does not extend to bosons, because they do not obey the exclusion principle on which hole theory depends, and to show that charged bosons also have antiparticles of opposite charge requires the formal apparatus of relativistic quantum field theory.[9] We shall not pursue this here, but note that the basic constituents of matter – the leptons and quarks – are not bosons, but are spin-$\frac{1}{2}$ fermions. The corresponding results on the antiparticles of hadrons, irrespective of their spin, can then be found by considering their quark constituents, as we shall see in Chapter 3. For neutral particles, there is no general rule governing the existence of antiparticles, and while some neutral

[9] For an introduction, see, for example, Mandl and Shaw (1993).

particles have distinct antiparticles associated with them, others do not. The photon, for example, does not have an antiparticle (or, rather, the photon and its antiparticle are identical) whereas the neutron does. Although the neutron has zero charge, it has a non-zero magnetic moment, and distinct antineutrons exist in which the sign of this magnetic moment is reversed relative to the spin direction. The neutron is also characterized by other quantum numbers (which we will meet later) that change sign between particle and antiparticle.

In what follows, if we denote a particle by P, then the antiparticle is in general written with a bar over it, i.e. $\overline{P}$. For example, the antiparticle of the proton is the antiproton $\overline{p}$, with negative electric charge, and associated with every quark, q, is an antiquark, $\overline{q}$. However, for some particles the bar is usually omitted. Thus, for example, in the case of the positron e^+, the superscript denoting the charge makes explicit the fact that the antiparticle has the opposite electric charge to that of its associated particle.

1.3 INTERACTIONS AND FEYNMAN DIAGRAMS

By analogy with chemical reactions, interactions involving elementary particles and/or hadrons are conveniently summarized by 'equations', in which the different particles are represented by symbols. Thus, in the reaction $\nu_e + n \rightarrow e^- + p$, an electron neutrino ν_e collides with a neutron n to produce an electron e^- and a proton p while the equation $e^- + p \rightarrow e^- + p$ represents an electron and proton interacting to give the same particles in the final state, but in general travelling in different directions. The forces producing the above interactions are due to the exchange of particles and a convenient way of illustrating this is to use the pictorial technique of *Feynman diagrams*. These were introduced by Feynman in the 1940s and are now one of the cornerstones of the analysis of elementary particle physics. Associated with them are mathematical rules and techniques that enable the calculation of the quantum mechanical probabilities for given reactions to occur. Here we will avoid the mathematical detail, but use the diagrams to understand the main features of particle reactions. We will introduce Feynman diagrams by firstly discussing electromagnetic interactions.

1.3.1 Basic electromagnetic processes

The electromagnetic interactions of electrons and positrons can all be understood in terms of eight basic processes. In hole theory, they arise from transitions in which an electron jumps from one state to another, with the emission or absorption of a single photon. The interpretation then depends on whether the states are both of positive energy, both of negative energy or one of each.

The basic processes whereby an electron either emits or absorbs a photon are

$$\text{(a) } e^- \rightarrow e^- + \gamma \text{ and (b) } \gamma + e^- \rightarrow e^-.$$

They correspond in hole theory to transitions between positive energy states of the electron and are represented pictorially in Figure 1.3. They may also be represented diagrammatically by Figures 1.4(a) and (b), where by convention time runs from left to right. These are examples of Feynman diagrams.

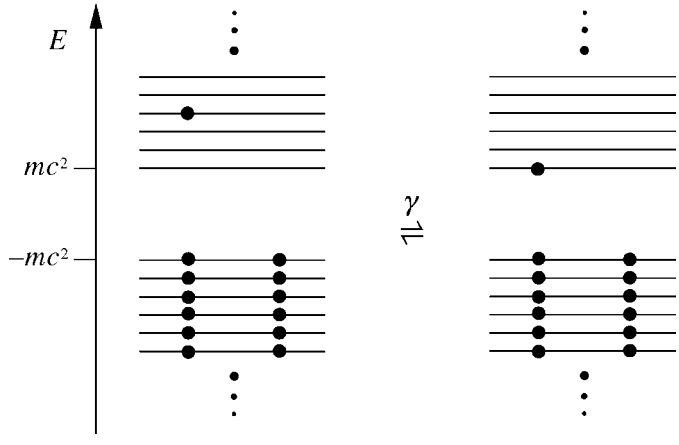


Figure 1.3 Hole theory representation of the processes $e^- \rightleftarrows e^- + \gamma$.

Similar diagrams may be drawn for the corresponding positron processes

$$\text{(c) } e^+ \to e^+ + \gamma \text{ and (d) } \gamma + e^+ \to e^+,$$

and are shown in Figures 1.4(c) and (d). Time again flows to the right, and we have used the convention that an arrow directed towards the right indicates a particle (in this case an electron) while one directed to the left indicates an antiparticle (in this case a positron). The corresponding hole theory diagram, analogous to Figure 1.3, is left as an exercise for the reader. Finally, there are processes in which an electron is excited from a negative energy state to a positive energy state, leaving a 'hole' behind, or in which a positive energy electron falls into a vacant level (hole) in the negative energy sea. These are illustrated in Figure 1.5, and correspond to the production or annihilation of e^+e^- pairs. In both cases, a photon may either be absorbed from the initial state or emitted to the final state, giving the four processes

$$\begin{array}{ll} \text{(e) } e^+ + e^- \to \gamma, & \text{(f) } \gamma \to e^+ + e^-, \\ \text{(g) vacuum } \to \gamma + e^+ + e^-, & \text{(h) } \gamma + e^+ + e^- \to \text{vacuum}, \end{array}$$

represented by the Feynman diagrams of Figures 1.4(e) to (h).

This exhausts the possibilities in hole theory, so that there are just eight basic processes represented by the Feynman diagrams of Figures 1.4(a) to (h). Each of these processes has an associated probability proportional to the strength of the electromagnetic fine structure constant

$$\alpha \equiv \frac{1}{4\pi\varepsilon_0}\frac{e^2}{\hbar c} \approx \frac{1}{137}. \tag{1.16}$$

1.3.2 Real processes

In each of the diagrams of Figures 1.4(a) to (h), each vertex has a line corresponding to a single photon being emitted or absorbed, while one fermion line has the arrow

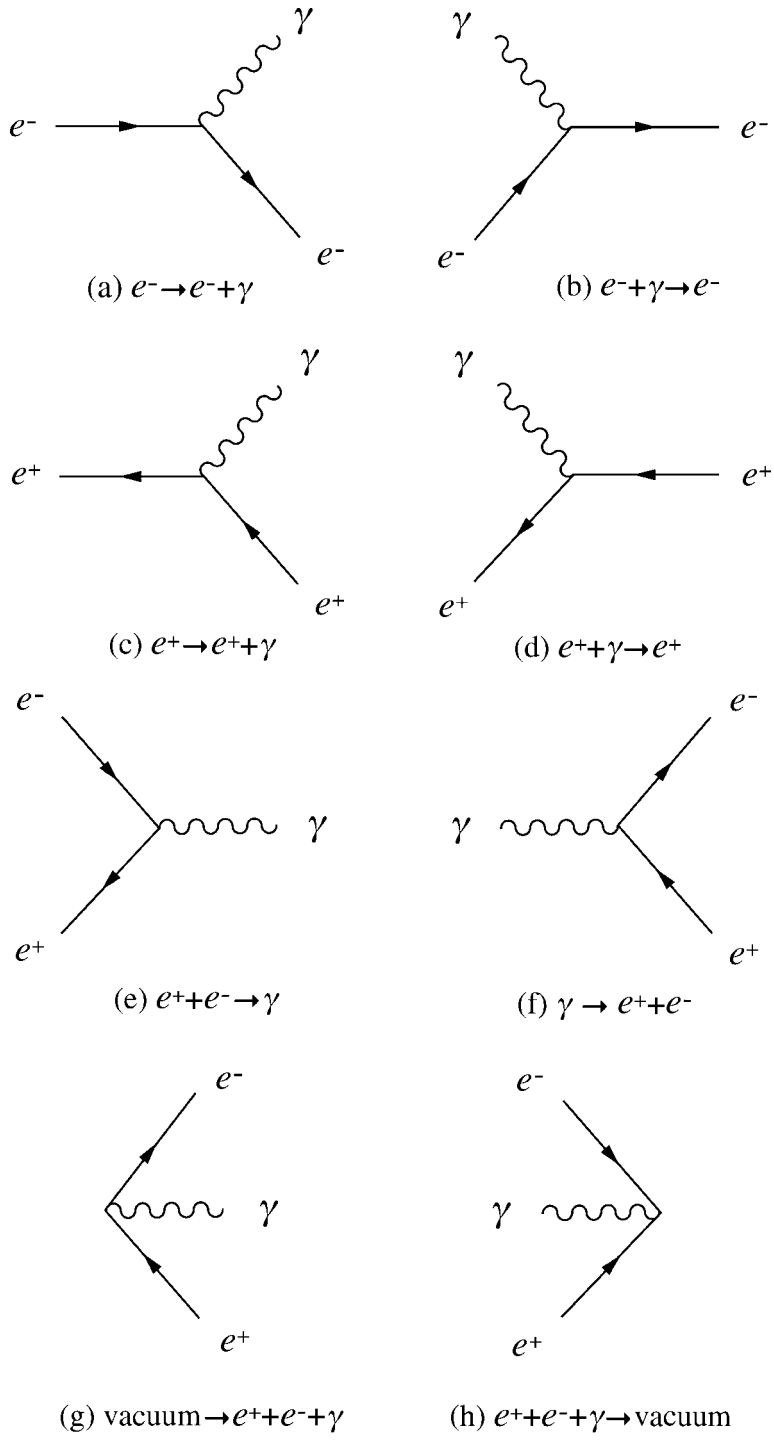


Figure 1.4 Feynman diagrams for the eight basic processes whereby electrons and positrons interact with photons. In all such diagrams, time runs from left to right, while a solid line with its arrow pointing to the right (left) indicates an electron (positron).

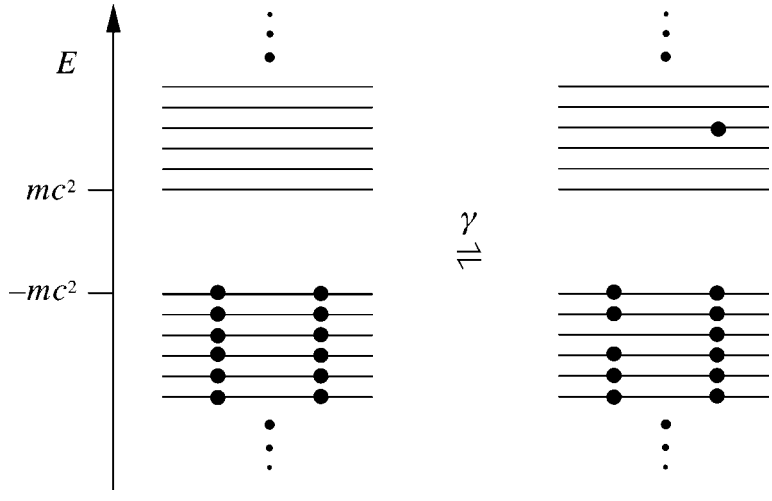


Figure 1.5 Hole theory representation of the production or annihilation of e^+e^- pairs.

pointing towards the vertex and the other away from the vertex, implying charge conservation at the vertex.[10] For example, a vertex like Figure 1.6 would correspond to a process in which an electron emitted a photon and turned into a positron. This would violate charge conservation and is therefore forbidden.

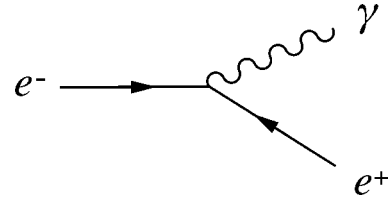


Figure 1.6 The forbidden vertex $e^- \to e^+ + \gamma$.

Momentum and angular momentum are also assumed to be conserved at the vertices. However, in free space, energy conservation is violated. For example, if we use the notation $(E, \mathbf{k})$ to denote the total energy and three-momentum of a particle, then in the rest frame of the electron, reaction (a) is

$$e^-(E_0, \mathbf{0}) \to e^-(E_k, -\mathbf{k}) + \gamma(ck, \mathbf{k}), \tag{1.17}$$

where $k \equiv |\mathbf{k}|$ and momentum conservation has been imposed. In free space, $E_0 = mc^2$, $E_k = (k^2c^2 + m^2c^4)^{1/2}$ and $\Delta E \equiv E_k + kc - E_0$ satisfies

[10] Compare Kirchhoff's laws in electromagnetism.

$$kc < \Delta E < 2kc \tag{1.18}$$

for all finite k.

Similiar arguments show that energy conservation is violated for all the basic processes. They are called *virtual* processes to emphasize that they cannot occur in isolation in free space. To make a real process, two or more virtual processes must be combined in such a way that energy conservation is only violated for a short period of time compatible with the energy–time uncertainty principle

$$\tau \Delta E \sim \hbar. \tag{1.19}$$

In particular, the initial and final states – which in princple can be studied in the distant past ($t \to -\infty$) and future ($t \to +\infty$), respectively – must have the same energy. This is illustrated by Figure 1.7(a), which represents a process whereby an electron emits a photon that is subsequently absorbed by a second electron. Although energy conservation is violated at the first vertex, this can be compensated by a similiar violation at the second vertex to give exact energy conservation overall. Figure 1.7(a) represents a contribution to the physical elastic scattering process

$$e^- + e^- \to e^- + e^-$$

from single-photon exchange. There is also a second contribution, represented by Figure 1.7(b) in which the other electron emits the exchanged photon. Both processes contribute to the observed scattering.

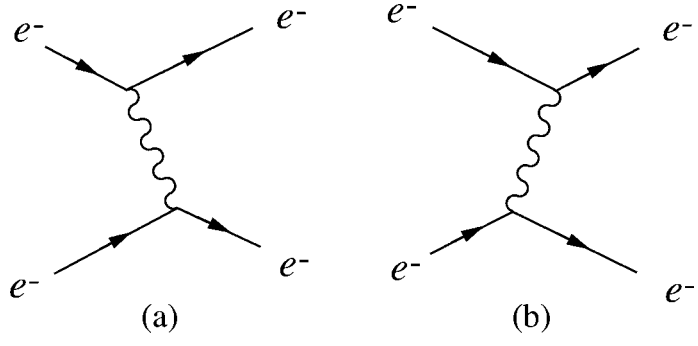


Figure 1.7 Single-photon exchange contributions to electron–electron scattering. Time as usual runs from left to right.

Scattering can also occur via multiphoton exchange and, for example, one of the diagrams corresponding to two-photon exchange is shown in Figure 1.8. The contributions of such diagrams are, however, far smaller than the one-photon exchange contributions. To see this, we consider the number of vertices in each diagram, called its *order*. Since each vertex represents a basic process whose probability is of order $\alpha \approx \frac{1}{137} \ll 1$, any diagram of order n gives a contribution of order α^n. By comparing Figures 1.7 and 1.8, we see that single-photon exchange is of order α^2,

two-photon exchange is of order α^4 and, more generally, n-photon exchange is of order α^{2n} To a good approximation multiphoton exchanges can be neglected, and we would expect the familiar electromagnetic interactions used in atomic spectroscopy, for example, to be accurately reproduced by considering only one-photon exchange. Detailed calculation confirms that this is indeed the case.

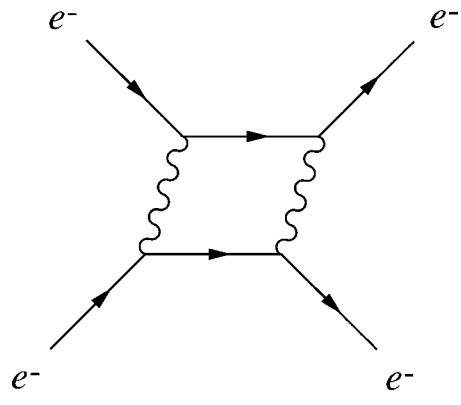


Figure 1.8 A contribution to electron–electron scattering from two-photon exchange.

1.3.3 Electron–positron pair production and annihilation

Feynman diagrams like Figures 1.7 and 1.8 play a central role in the analysis of elementary particle interactions. The contribution of each diagram to the probability amplitude for a given physical process can be calculated precisely by using the set of mathematical rules mentioned previously (the *Feynman rules*), which are derived from the quantum theory of the corresponding interaction. For electromagnetic interactions, this theory is called *quantum electrodynamics* (or *QED* for short) and the resulting theoretical predictions have been verified experimentally with quite extraordinary precision. The Feynman rules are beyond the scope of this book, and our use of

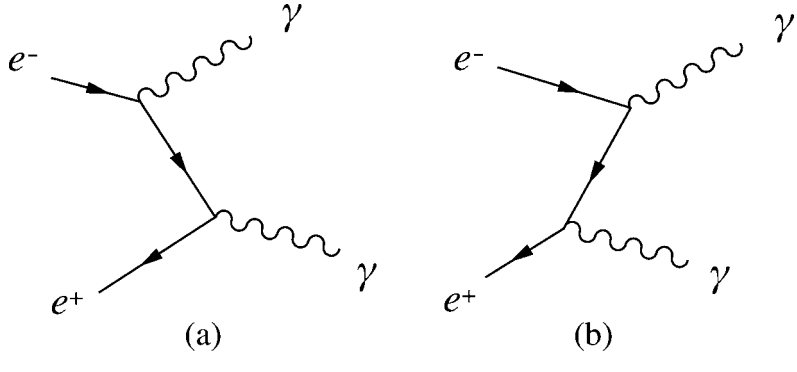


Figure 1.9 Lowest-order contributions to $e^+ + e^- \rightarrow \gamma + \gamma$. The two diagrams are related by 'time ordering', as explained in the text.

Feynman diagrams will be much more qualitative. For this, we need consider only the lowest-order diagrams contributing to a given physical process, neglecting higher-order diagrams like Figure 1.8 for electron–electron scattering.

To illustrate this we consider the reactions $e^+ + e^- \to p\gamma$, where to conserve energy at least two photons must be produced, $p \geq 2$. In lowest order, to produce p photons we must combine p vertices from the left-hand side of Figure 1.4. For $p = 2$,

$$e^+ + e^- \to \gamma + \gamma$$

and there are just two such diagrams, Figure 1.9(a), which is obtained by combining Figures 1.4(a) and (e), and Figure 1.9(b), which combines Figures 1.4(c) and (e). These two diagrams are closely related. If the lines of Figure 1.9(a) were made of rubber, we could imagine deforming them so that the top vertex occurred after, instead of before, the bottom vertex, and it became Figure 1.9(b). Figures 1.7(a) and (b) are related in the same way. Diagrams related in this way are called different 'time orderings' and in practice it is usual to draw only one time ordering (e.g. Figure 1.9(a)), leaving the other(s) implied. Thus for p = 3,

$$e^+ + e^- \to \gamma + \gamma + \gamma$$

and a possible diagram is that of Figure 1.10, obtained by combining Figures 1.4(a),(c) and (e). Since there are three vertices, there are $3! = 6$ different ways of ordering them in time. We leave it as an exercise for the reader to draw the other five time-ordered diagrams whose existence is implied by Figure 1.10.

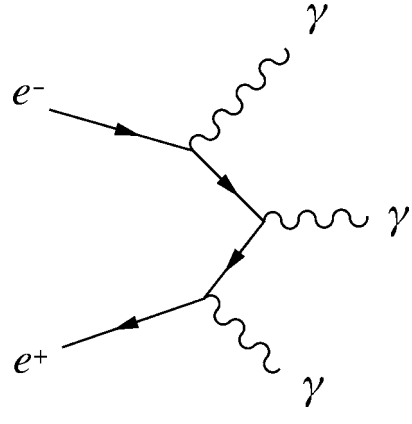


Figure 1.10 The process $e^+ + e^- \to \gamma + \gamma + \gamma$ in lowest order. Only one of the six possible time orderings is shown, leaving the other five implied.

In general, the process $e^+ + e^- \to p\gamma$ is of order p, with an associated probability of order α^p. From just the order of the diagrams (i.e. the number of vertices), we

therefore expect that many-photon annihilation is very rare compared to few-photon annihilation, and that[11]

$$R \equiv \frac{\text{Rate } (e^+e^- \to 3\gamma)}{\text{Rate } (e^+e^- \to 2\gamma)} = O(\alpha). \tag{1.20}$$

For very low energy e^+e^- pairs, this prediction can be tested by measuring the 2γ and 3γ decay rates of positronium, which is a bound state of e^+ and e^- analogous to the hydrogen atom.[12] In this case, the experimental vaue of R is 0.9×10^{-3}. This is somewhat smaller than $\alpha = 0.7 \times 10^{-2}$, and indeed some authors argue that $\alpha/2\pi = 1.2 \times 10^{-3}$ is a more appropriate measure of the strength of the electromagnetic interaction. We stress that (1.20) is only an order-of-magnitude prediction, and we will in future use $10^{-2} - 10^{-3}$ as a rough rule-of-thumb estimate of what is meant by a factor of order α.

As a second example, consider the pair production reaction $\gamma \to e^+ + e^-$. This basic process cannot conserve both energy and momentum simultaneously, but is allowed in the presence of a nucleus, i.e.

$$\gamma + (Z,A) \to e^+ + e^- + (Z,A),$$

where Z and A are the charge and mass number (the number of nucleons) of the nucleus, respectively. Diagrammatically, this is shown in Figure 1.11, where the two diagrams are not different time orderings. (There is no way that (a) can be continuously deformed into (b).) Since one of the vertices involves a charge Ze, the corresponding factor $\alpha \to Z^2\alpha$ and the expected rate is of order $Z^2\alpha^3$. This Z dependence is confirmed experimentally, and is exploited in devices for detecting

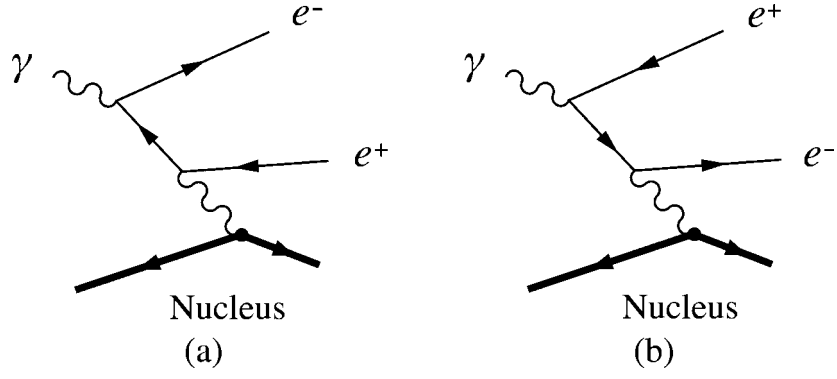


Figure 1.11 The pair production process $\gamma + (Z,A) \to e^+ + e^- + (Z,A)$ in lowest order. The two diagrams represent 'distinct' contributions and are not related by time ordering.

[11] We will often use the symbol O to mean 'order' in the sense of 'order-of-magnitude'.
[12] This is discussed in Section 5.5.

high-energy photons by the pair production that occurs so readily in the field of a heavy nucleus.[13]

1.3.4 Other processes

Although we have introduced Feynman diagrams in the context of electromagnetic interactions, their use is not restricted to this. They can also be used to describe the fundamental weak and strong interactions. This is illustrated by Figure 1.12(a), which shows a contribution to the elastic weak scattering reaction $e^- + \nu_e \rightarrow e^- + \nu_e$ due to the exchange of a Z^0, and by Figure 1.12(b), which shows the exchange of a gluon g (represented by a coiled line) between two quarks, which is a strong interaction.

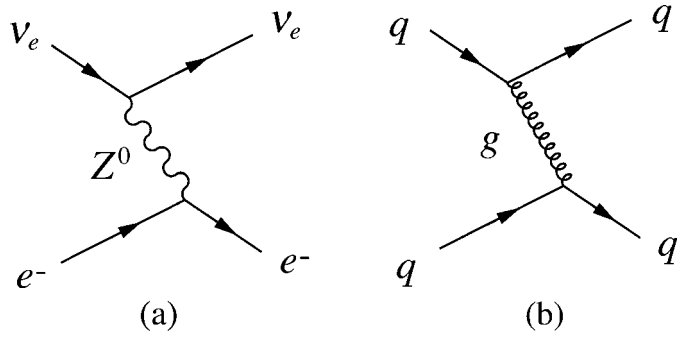


Figure 1.12 (a) Contributions of (a) Z^0 exchange to the elastic weak scattering reaction $e^- + \nu_e \rightarrow e^- + \nu_e$ and (b) the gluon exchange contribution to the strong interaction $q + q \rightarrow q + q$.

Feynman diagrams that involve hadrons can also be drawn. As an illustration, Figure 1.13 shows the decay of a neutron via an intermediate charged W boson. In later chapters we will make extensive use of Feynman diagrams in discussing particle interactions.

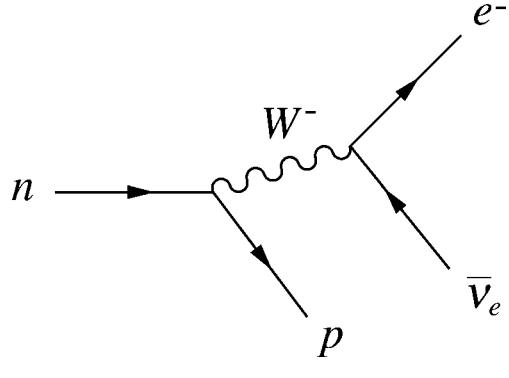


Figure 1.13 The decay $n \rightarrow p + e^- + \overline{\nu}_e$ via an intermediate W meson.

[13] These detectors will be discussed in Chapter 4.

1.4 PARTICLE EXCHANGE

In Section 1.1 we said that the forces of elementary particle physics were associated with the exchange of particles. In this section we will explore the fundamental relation between the range of a force and the mass of the exchanged particle, and show how the weak interaction can frequently be approximated by a force of zero range.

1.4.1 Range of forces

The Feynman diagram of Figure 1.14 represents the elastic scattering of two particles A and B of masses M_A and M_B, i.e. $A + B \to A + B$, via the exchange of a third particle X of mass M_X, with equal coupling strengths g to particles A and B. In the rest frame of the incident particle A, the lower vertex represents the virtual process

$$A(M_A c^2, \mathbf{0}) \to A(E_A, \mathbf{p}) + X(E_X, -\mathbf{p}), \tag{1.21}$$

where

$$E_A = (p^2c^2 + M_A^2 c^4)^{1/2} \quad \text{and} \quad E_X = (p^2c^2 + M_X^2 c^4)^{1/2}, \tag{1.22}$$

and $p \equiv |\mathbf{p}|$. The energy difference between the final and initial states is given by

$$\begin{aligned} \Delta E = E_X + E_A - M_A c^2 &\to 2pc, \qquad p \to \infty \\ &\to M_X c^2, \quad p \to 0 \end{aligned} \tag{1.23}$$

and thus $\Delta E \geq M_X c^2$ for all p. By the uncertainty principle, such an energy violation is allowed, but only for a time $\tau \approx \hbar/\Delta E$, so we immediately obtain

$$r \approx R \equiv \hbar/M_X c \tag{1.24}$$

as the maximum distance over which X can propagate before being absorbed by particle B. This constant R is called the *range* of the interaction and gives the sense of the word used in Section 1.1.

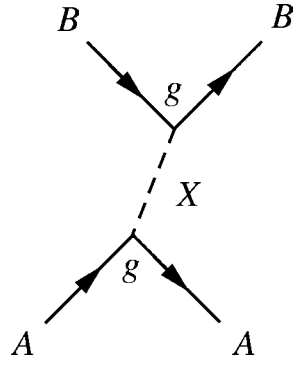


Figure 1.14 Contribution to the reaction $A + B \to A + B$ from the exchange of a particle X.

The electromagnetic interaction has an infinite range, because the exchanged particle is a massless photon. The strong force between quarks also has infinite range because the exchanged particles are massless gluons.[14] In contrast, the weak interaction is associated with the exchange of very heavy particles, the W and Z bosons, with masses

$$M_W = 80.4\,\mathrm{GeV/c^2} \quad \text{and} \quad M_Z = 91.2\,\mathrm{GeV/c^2} \quad (1\,\mathrm{GeV} = 10^9\,\mathrm{eV}) \tag{1.25}$$

corresponding to ranges that from (1.24) are of order

$$R_{W,Z} \equiv \frac{\hbar}{M_W c} \approx 2 \times 10^{-3}\,\mathrm{fm} \quad (1\,\mathrm{fm} = 10^{-15}\,\mathrm{m}). \tag{1.26}$$

In many applications, this range is very small compared with the de Broglie wavelengths of all the particles involved. The weak interaction can then be approximated by a zero-range or point interaction, corresponding to the limit $M_X \to \infty$, as shown in Figure 1.15.

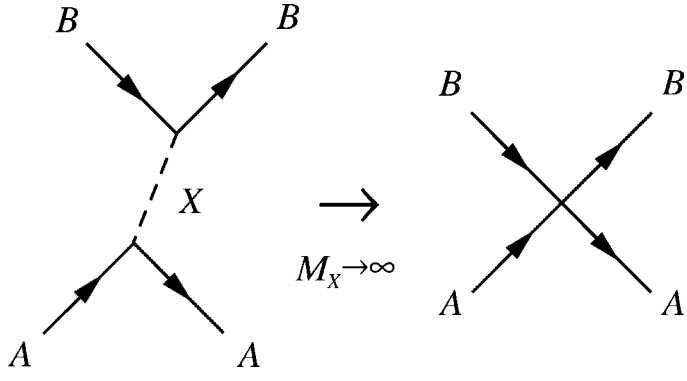


Figure 1.15 The zero-range or point interaction resulting from the exchange of a particle X in the limit $M_X \to \infty$.

1.4.2 The Yukawa potential

In the limit where M_A becomes large, we can regard B as being scattered by a static potential of which A is the source. This potential will in general be spin-dependent, but its main features can be obtained by neglecting spin and considering X to be a spin-0 boson, in which case it will obey the Klein–Gordon equation,

$$-\hbar^2 \frac{\partial^2 \phi(\mathbf{r},t)}{\partial t^2} = -\hbar^2 c^2 \nabla^2 \phi(\mathbf{r},t) + M_X^2 c^4 \phi(\mathbf{r},t). \tag{1.27}$$

[14] The force between hadrons is much more complicated, because it is not due to single-gluon exchange. It has a range of approximately $(1-2) \times 10^{-15}$ m.

The static solution of this equation satisfies

$$\nabla^2\phi(\mathbf{r}) = \frac{M_X^2 c^4}{\hbar^2}\phi(\mathbf{r}), \tag{1.28}$$

where $\phi(\mathbf{r})$ is interpreted as a static potential. For $M_X = 0$ this equation is the same as that obeyed by the electrostatic potential, and for a point charge $-e$ interacting with a point charge $+e$ at the origin, the appropriate solution is the Coulomb potential

$$V(r) = -e\phi(r) = -\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r}, \tag{1.29}$$

where $r = |\mathbf{r}|$ and ε_0 is the dielectric constant. The corresponding solution in the case where $M_X^2 \neq 0$ is easily verified by substitution to be

$$V(r) = -\frac{g^2}{4\pi}\frac{\mathrm{e}^{-r/R}}{r}, \tag{1.30}$$

where R is the range defined in (1.24) and we assume equal coupling constants g for particle X to particles A and B. It is conventional to introduce a dimensionless strength parameter

$$\alpha_X = \frac{g^2}{4\pi\hbar c} \tag{1.31}$$

that characterizes the strength of the interaction at short distances, in analogy to the fine structure constant of QED.

The form of $V(r)$ in (1.30) is called a *Yukawa potential*, after the physicist who first introduced the idea of forces due to massive particle exchange in 1935. As $M_X \to 0$, $R \to \infty$ and the Coulomb potential is recovered from the Yukawa potential, while for very large masses the interaction is approximately point-like (zero range). The effective coupling strength in this latter approximation and its range of validity are best understood by considering the corresponding scattering amplitude.

1.4.3 The zero-range approximation

In lowest-order perturbation theory, the probability amplitude for a particle with initial momentum $\mathbf{q}_i$ to be scattered to a final state with momentum $\mathbf{q}_f$ by a potential $V(\mathbf{r})$ is proportional to[15]

$$\mathcal{M}(\mathbf{q}) = \int \mathrm{d}^3\mathbf{r}\, V(\mathbf{r}) \exp(i\mathbf{q}\cdot\mathbf{r}/\hbar), \tag{1.32}$$

[15] This is called the Born approximation. For a discussion, see, for example, Section 10.2.2 of Mandl (1992) or pp. 397–399 of Gasiorowicz (1974).

where $\mathbf{q} \equiv \mathbf{q}_i - \mathbf{q}_f$ is the momentum transfer. The integration may be done using polar co-ordinates. Taking $\mathbf{q}$ in the x direction gives

$$\mathbf{q} \cdot \mathbf{r} = |\mathbf{q}|\ r \cos\theta \tag{1.33}$$

and

$$\mathrm{d}^3\mathbf{r} = r^2 \sin\theta\ \mathrm{d}\theta\ \mathrm{d}r\ \mathrm{d}\phi, \tag{1.34}$$

where $r \equiv |\mathbf{r}|$. For the Yukawa potential, the integral (1.32) gives

$$\mathcal{M}(\mathbf{q}) = \frac{-g^2\hbar^2}{|\mathbf{q}|^2 + M_X^2c^2}. \tag{1.35}$$

In deriving (1.35) for the scattering amplitude we have used potential theory, treating particle A as a static source. Particle B then scatters through some angle without loss of energy, so that $|\mathbf{q}_i| = |\mathbf{q}_f|$ and the initial and final energies of particle B are equal, $E_i = E_f$. While this is a good approximation at low energies, at higher energies the recoil energy of the target particle cannot be neglected, so that the initial and final energies of B are no longer equal. A full relativistic calculation taking account of this is beyond the scope of this book, but the result is surprisingly simple. Specifically, in lowest-order perturbation theory, one obtains

$$\mathcal{M}(q^2) = \frac{g^2\hbar^2}{q^2 - M_X^2c^2}, \tag{1.36}$$

where

$$q^2 \equiv (E_f - E_i)^2 - (\mathbf{q}_f - \mathbf{q}_i)^2c^2 \tag{1.37}$$

is the squared energy–momentum transfer, or squared four-momentum transfer.[16] In the low-energy limit, $E_i = E_f$ and (1.36) reduces to (1.35). However, in contrast to (1.35), which was derived in the rest frame of particle A, the form (1.36) is explicitly Lorentz-invariant and holds in all inertial frames of reference.

This amplitude (1.35) corresponds to the exchange of a single particle, as shown, for example, in Figure 1.14. (Multiparticle exchange corresponds to higher orders in perturbation theory and higher powers of g^2.) In the zero-range approximation, (1.36) reduces to a constant. To see this, we note that this approximation is valid when the range $R = \hbar/M_Xc$ is very small compared with the de Broglie wavelengths of all the particles involved. In particular, this implies $q^2 \ll M_X^2c^2$ and neglecting q^2 in (1.36) gives

$$\mathcal{M}(q^2) = -G, \tag{1.38}$$

[16] A resumé of relativistic kinematics is given in Appendix A.

where the constant G is given by

$$\frac{G}{(\hbar c)^3} = \frac{1}{\hbar c}\left(\frac{g}{M_X c^2}\right)^2 = \frac{4\pi\alpha_X}{(M_X c^2)^2} \tag{1.39}$$

and has the dimensions of inverse energy squared. Thus we see that in the zero-range approximation, the resulting point interaction between A and B (see Figure 1.15) is characterized by a single dimensioned coupling constant G and not g and M_X separately. As we shall see in the next chapter, this approximation is extremely useful in weak interactions, where the corresponding *Fermi coupling constant*, measured, for example, in nuclear β-decay, is given by

$$\frac{G_F}{(\hbar c)^3} = 1.166 \times 10^{-5}\,\text{GeV}^{-2}. \tag{1.40}$$

1.5 UNITS AND DIMENSIONS

In previous sections, we have quoted numerical values for some constants, for example G_F above, and given formulas for others. Before continuing our discussion, we consider more carefully the question of units.

Units are a perennial problem in physics, since most branches of the subject tend to adopt a system that is convenient for their own purpose. Elementary particle physics is no exception, and adopts so-called *natural units,* chosen so that the fundamental constants

$$\hbar = 1 \quad \text{and} \quad c = 1. \tag{1.41}$$

In other words, c and $\hbar$ are used as fundamental units of velocity and action (or angular momentum) respectively. To complete the definition, a third unit must be fixed, which is chosen to be the unit of energy. This is taken to be the electronvolt (eV), defined as the energy required to raise the electric potential of an electron or proton by one volt. The abbreviations keV(10^3 eV), MeV(10^6 eV), GeV (10^9 eV) and TeV(10^{12} eV) are also in general use.

Quantum mechanics and special relativity play crucial roles in elementary particle physics, so that $\hbar$ and c occur frequently in formulas. By choosing natural units, all factors of $\hbar$ and c may be omitted from equations using (1.41), which leads to considerable simplifications. For example, the relativistic energy relation

$$E^2 = p^2c^2 + m^2c^4$$

becomes

$$E^2 = p^2 + m^2,$$

while the Fermi coupling constant (1.40) becomes

$$G_F = 1.166 \times 10^{-5}\,\text{GeV}^{-2}. \tag{1.42}$$

In natural units (nu) *all* quantities have the dimensions of a power of energy, since they can all be expressed in terms of $\hbar, c$ and an energy. In particular, masses, lengths and times can be expressed in the forms

$$M = E/c^2, \quad L = \hbar c/E, \quad T = \hbar/E,$$

so that a quantity with metre–kilogram–second (mks) dimensions $M^p L^q T^r$ has the nu dimensions E^{p-q-r}. Since $\hbar$ and c are supressed in nu, this is the only dimension that is relevant, and dimensional checks and estimates are very simple. The nu and mks dimensions are listed for some important quantities in Table 1.1. We note that in natural units many different quantites (e.g. mass, energy and momentum) all have the same dimension.

TABLE 1.1 The mks dimension $M^p L^q T^r$ and the nu dimensions $E^n = E^{p-q-r}$ of some quantities.

	mks			nu
Quantity	p	q	r	n
Action $\hbar$	1	2	−1	0
Velocity c	0	1	−1	0
Mass	1	0	0	1
Length	0	1	0	−1
Time	0	0	1	−1
Momentum	1	1	−1	1
Energy	1	2	−2	1
Fine structure constant α	0	0	0	0
Fermi coupling constant G_F	1	5	−2	−2

Natural units are very convenient for theoretical arguments, as we will see. However, we must still know how to convert from nu to the 'practical' units in which experimental results are invariably stated. This is done in two steps. We first restore the $\hbar$ and c factors by dimensional arguments and then use the conversion factors

$$\hbar = 6.582 \times 10^{-22}\,\text{MeV s} \tag{1.43a}$$

and

$$\hbar c = 1.973 \times 10^{-13}\,\text{MeV m} \tag{1.43b}$$

to evaluate the result.

We illustrate this by using the nu expression for the cross-section[17] for Thomson scattering, i.e. for Compton scattering from free electrons when the photon energy is much less than the electron rest energy. This is given by[18]

[17] Cross-sections and related quantities are formally defined in Appendix B.
[18] See, for example, p. 22 of Mandl and Shaw (1993).

$$\sigma = \frac{8\pi\alpha^2}{3m_e^2}.$$

To convert it to practical units, we write

$$\sigma = \frac{8\pi\alpha^2}{3m_e^2}\hbar^a c^b$$

and demand that σ has the dimensions of length squared. This gives $a = 2, b = -2$, so that

$$\sigma = \frac{8\pi\alpha^2(\hbar c)^2}{3(m_e c^2)^2} = 6.65 \times 10^{-29}\,\mathrm{m}^2, \tag{1.44}$$

using (1.43b) and $m_e = 0.51\,\mathrm{MeV/c^2}$.

In practice, cross-sections are usually quoted in *barns* b($10^{-28}\,\mathrm{m}^2$), *millibarns* mb ($10^{-31}\,\mathrm{m}^2$) or *microbarns* mb ($10^{-34}\,\mathrm{m}^2$), rather than square metres. Thus the Thomson cross-section (1.44) is 0.665 b, while the total cross-section for *np* scattering is typically of order 30–40 mb, depending on the energy. Similiarly, lengths are often quoted in *fermis* fm ($10^{-15}\,\mathrm{m}$) rather than metres. In these units, the radius of the proton is about 0.8 fm, while the range of the weak force is of order 10^{-3} fm. Energies are measured in MeV, GeV, etc., while momenta are measured in MeV/c, etc., and masses in $\mathrm{MeV/c^2}$, etc., as in Equation (1.25) for the weak boson mass. This should be compared to natural units, where energy, momentum and mass all have the dimension of energy (see Table 1.1), and are all measured in, for example, MeV.

A list of some useful physical constants and conversion factors is given inside the back cover of this book. *From now on natural units will be used throughout the book unless explicitly stated otherwise.*

PROBLEMS 1

1.1 Write down 'equations' in symbol form that represent the following interactions:

(a) elastic scattering of an electron antineutrino and a positron;
(b) annihilation of an antiproton with a neutron to produce three pions.

1.2 Draw the topologically distinct Feynman diagrams that contribute to the following processes in lowest order:

(a) $\gamma + e^- \to \gamma + e^-$
(b) $e^+ + e^- \to e^+ + e^-$
(c) $\nu_e \overline{\nu}_e$ elastic scattering

(*Hint.* There are two such diagrams for each reaction.)

1.3 Draw a fourth-order Feynman diagram for the reaction $\gamma + \gamma \to e^+ + e^-$.

1.4 Show that the Yukawa potential of Equation (1.30) is the only spherically symmetric solution of the static Klein–Gordon equation (1.28) that vanishes as r goes to infinity.

1.5 In lowest order, the process $e^+ + e^- \to \gamma + \gamma$ is given by the Feynman diagrams of Figure 1.9. Show that for electrons and positrons almost at rest, the distance between the two vertices is typically of order m^{-1} in natural units, where m is the electron mass. Convert this result to practical units and evaluate it in fermis.

1.6 In lowest order, the process $e^+ + e^- \to \mu^+ + \mu^-$ is given by the Feynman diagram of Figure 1.16. Estimate the typical distance between the vertices at energies much larger

than the masses of any of the particles in(a) the rest frame of the electron and (b) the centre-of-mass frame. Check the consistency of these estimates by considering the Lorentz contraction in going from the rest frame, in which the initial electron is stationary, to the centre-of-mass frame.

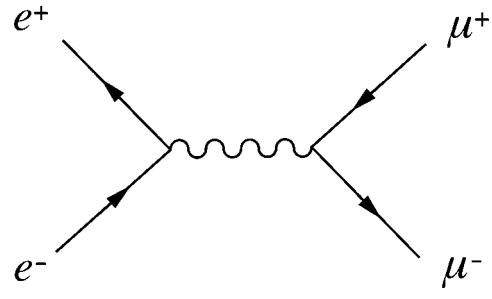


Figure 1.16 Lowest-order Feynman diagram for the process $e^+ + e^- \to \mu^+ + \mu^-$.

1.7 Parapositronium is an unstable bound state of an electron and a positron. Its lifetime is given in natural units by $\tau = 2/m\alpha^5$, where m is the mass of the electron and α is the fine structure constant. Restore the factors of h and c by dimensional arguments and evaluate τ in seconds.

2

Leptons and the Weak Interaction

In Chapter 1 we introduced some of the ideas underlying the discussion of particle interactions. In doing this, we concentrated mainly on electrons and positrons and their electromagnetic interactions. In this chapter we extend that discussion to include other leptons, starting with a review of their basic properties, followed by a discussion of the weak interactions between them. The final section gives an account of an interesting phenomenon that neutrinos can exhibit if they have nonzero masses, which results in their changing their properties as they travel over long distances. We conclude by considering the experimental evidence for nonzero neutrino masses and its implications.

2.1 LEPTON MULTIPLETS AND LEPTON NUMBERS

Leptons are one of the three classes of fundamental particles in the standard model. They are spin-$\frac{1}{2}$ fermions without strong interactions. There are six known leptons, which occur in pairs called *generations*, written as doublets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}. \tag{2.1}$$

The three charged leptons (e^-, μ^-, τ^-) are the familiar electron and two new particles, the *mu-lepton* or *muon* and the *tau-lepton* or *tauon*. All have electric charge $Q = -e$. Associated with them are three neutral leptons, or *neutrinos*, called the *electron neutrino, mu-neutrino* and *tau-neutrino*, respectively, all of which have

Particle Physics Third Edition B. R. Martin and G. Shaw

very small masses. The six distinct types of leptons are also referred to as having different 'flavours'. In addition to the leptons, there are six corresponding antiparticles (antileptons):

$$\begin{pmatrix} e^+ \\ \bar{\nu}_e \end{pmatrix}, \begin{pmatrix} \mu^+ \\ \bar{\nu}_\mu \end{pmatrix}, \begin{pmatrix} \tau^- \\ \bar{\nu}_\tau \end{pmatrix}. \tag{2.2}$$

The charged leptons interact via both electromagnetic and weak forces, whereas for neutral leptons only weak interactions have been observed. Because of this, neutrinos can be detected only with considerable difficulty. We will discuss each of these particles in turn presently.

Firstly, however, we note that each generation of leptons has associated with it a quantum number. The first of these *lepton numbers* is the *electron number*, defined for any state by

$$L_e \equiv N(e^-) - N(e^+) + N(\nu_e) - N(\bar{\nu}_e)\,, \tag{2.3a}$$

where $N(e^-)$ is the number of electrons present, and so on. For single-particle states, $L_e = 1$ for e^- and ν_e, $L_e = -1$ for e^+ and $\bar{\nu}_e$ and $L_e = 0$ for all other particles. In electromagnetic interactions, electron number conservation reduces to the conservation of $N(e^-) - N(e^+)$, since neutrinos are not involved. This implies that electrons and positrons can only be created or annihilated in pairs. In weak interactions more general possibilities are allowed. For example, an electron could be created together with an antineutrino $\bar{\nu}_e$, rather than a positron. Similar remarks apply to *muon number* and *tauon number*, defined by

$$L_\mu \equiv N(\mu^-) - N(\mu^+) + N(\nu_\mu) - N(\bar{\nu}_\mu) \tag{2.3b}$$

and

$$L_\tau \equiv N(\tau^-) - N(\tau^+) + N(\nu_\tau) - N(\bar{\nu}_\tau), \tag{2.3c}$$

respectively. In the standard model, lepton numbers are individually conserved in all known interactions.

2.1.1 Electron neutrinos

Neutrinos, they are very small.
  They have no charge and have no mass
And do not interact at all.
The earth is just a silly ball
  To them, through which they simply pass,
Like dustmaids down a drafty hall
  Or photons through a sheet of glass.
J. Updike[1]

[1] From *Telephone Poles and Other Poems,* André Deutch, London (1964).

The existence of the *electron neutrino* was first postulated by Pauli in 1930 in order to understand the observed β-decays

$$(Z, A) \rightarrow (Z+1\,, A) + e^- + \bar{\nu}_e \tag{2.4a}$$

and

$$(Z', A') \rightarrow (Z'-1\,, A') + e^+ + \nu_e\,, \tag{2.4b}$$

where (Z,A) denotes a nucleus with atomic number Z and mass number A. These reactions are actually decays of bound neutrons and protons via the basic processes

$$n \rightarrow p + e^- + \bar{\nu}_e \tag{2.5a}$$

and

$$p \rightarrow n + e^+ + \nu_e, \tag{2.5b}$$

where only the neutron decay can occur in free space since $m_n > (m_p + m_e)$. For bound protons the decay (2.5b) *can* occur because the apparent energy violation can be compensated by changes in the nuclear binding energy. The neutrinos in these decays have been written as ν_e or $\bar{\nu}_e$ by using electron number conservation. They were not initially observed experimentally, but inferred from energy and angular momentum conservation. In the case of energy, if the antineutrino were not present in (2.4a), the reaction would be a two-body decay, and the energy E_e of the emitted electron would have the unique value

$$E_e = \Delta M = M(Z\,, A) - M(Z+1\,, A)\,,$$

where we have neglected the nuclear recoil energy. However, if the antineutrino is present, the electron energy will not be unique, but will lie in the range

$$m_e \leq E_e \leq (\Delta M - m_{\bar{\nu}_e})\,. \tag{2.6}$$

Experimentally, the observed spectrum spans the whole range (2.6), with $m_{\bar{\nu}_e} \approx 0$. Careful study of the spectrum near the end point $E_e = \Delta M - m_{\bar{\nu}_e}$ allows an upper limit to be set on the neutrino mass. The best results come from the β-decay of tritium

$$^3\mathrm{H} \rightarrow {}^3\mathrm{He} + e^- + \bar{\nu}_e,$$

which gives

$$m_{\bar{\nu}_e} < 2\,\mathrm{eV/c^2} \approx 4 \times 10^{-6} m_e\,. \tag{2.7}$$

Such an extremely small mass can be ignored in most calculations,[2] although we shall see later that there is strong evidence that neutrino masses are nonzero.

Neutrinos and antineutrinos can in principle be detected by observing the *inverse β-decay* processes

$$\nu_e + n \rightarrow e^- + p \tag{2.8a}$$

and

$$\bar{\nu}_e + p \rightarrow e^+ + n\,. \tag{2.8b}$$

However, the cross-sections for these processes are extremely small. For the neutrinos and antineutrinos emitted in β-decays, with energies of order 1 MeV, the cross-sections are of order $10^{-47}\,\mathrm{m}^2$, corresponding to mean free paths in matter of many millions of kilometres. Nevertheless, if the neutrino flux is intense enough, the reactions can be observed.

This was first done in a classic experiment by Reines and Cowan in 1959 using antineutrinos emitted from a nuclear reactor at Savannah River, North Carolina. Uranium fission fragments are neutron rich, and decay by (2.4a) to give an antineutrino flux of order $10^{17}\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ in the vicinity of the reactor. These occasionally interacted with protons in a large detector to give examples of reaction (2.8b) at rates of approximately two per hour. To distinguish these from 'background' (i.e. spurious) events from other sources, a very distinctive signal was required. This was achieved in an ingenious manner. The detector, which is shown schematically in Figure 2.1, comprised two target tanks, each containing an aqueous solution of cadmium chloride, sandwiched between three tanks of liquid scintillator.[3] When the reaction (2.8b) occurred in one of the target tanks, the emitted positron rapidly annihilated with an atomic electron to produce photons, giving a 'prompt coincidence' signal in the two adjacent tanks of scintillator. In contrast, the neutron would suffer multiple scattering from protons in the water until, somewhat later, it reached thermal energies. It could then be radiatively captured by a cadmium nucleus, releasing more photons and giving rise to a second 'delayed coincidence' in the same two adjacent tanks of scintillator.

This course of events is illustrated schematically in Figure 2.2. The distinctive signal is thus two coincidences in the same pair of scintillator tanks, separated by a time, which was calculated to be a few microseconds. In their first series of measurements lasting just under 200 hours, 567 such events were found, compared to an estimated 209 from background accidentals.[4] Detailed analysis[5] confirmed that the remaining signals were due to inverse β-decay events. Antineutrinos had for the first time been detected.

---

[2] The exceptions occur mostly in astrophysics and cosmology. We will discuss this briefly in Section 2.3.3 and Chapter 11.

[3] A charged particle passing through matter will excite atomic electrons, and in some substances, called 'scintillators', a small fraction of this excitation energy re-emerges as visible light. Scintillators and scintillation counters are discussed in Chapter 4.

[4] There are single coincidences due to unaccompanied positron or neutron counts from other sources. From their measured rates, one can estimate how often two of these will occur separated by a few microseconds, mimicking a genuine neutrino event. These are the so-called 'accidentals'.

[5] A simple account is given in Trigg (1975).

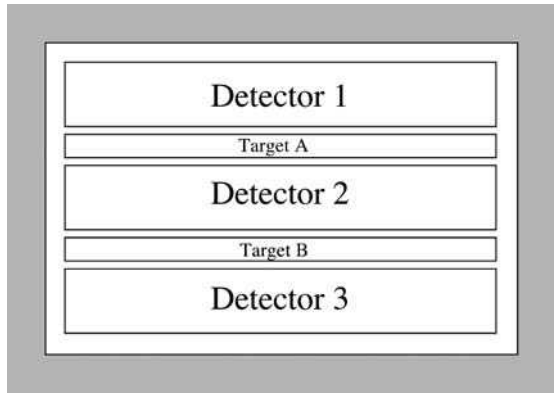


Figure 2.1 Schematic diagram of the apparatus used by Reines and Cowen to detect antineutrinos. The two target tanks, containing an aqueous solution of cadmium chloride, are sandwiched between three detector tanks of liquid scintillator. The whole apparatus is surrounded by heavy shielding to eliminate all incident particles except neutrinos and antineutrinos.

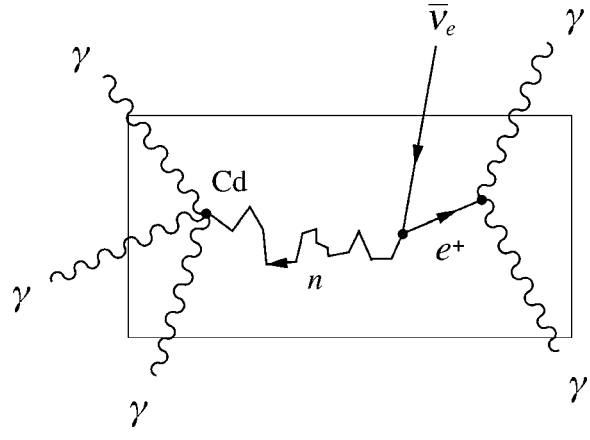


Figure 2.2 Schematic picture of an event corresponding to reaction (2.8b) in the Reines–Cowan experiment. The emitted positron rapidly annihilates with an atomic electron, and the emitted photons are detected in the adjacent tanks of scintillator as a 'prompt coincidence'. The neutron is reduced to thermal energies by repeated scattering from protons in the water before being radiatively captured by the cadmium nucleus. The emitted photons are then detected as a 'delayed coincidence' in the same pair of scintillator tanks.

2.1.2 Further generations

In addition to the (e^-, ν_e) pair, we have seen that there are two more lepton generations known: the muon μ^- and its associated mu-neutrino ν_μ, and the tauon τ^- and its associated tau-neutrino ν_τ.

The muon is a very penetrating particle of mass 105.7 MeV/c^2 that was first identified in cosmic ray experiments by Anderson and Neddermeyer in 1936. Cosmic ray *primaries* are high-energy particles, mostly protons, incident on the Earth's atmosphere from all directions in space. Other particles, called *secondaries,* are produced

when the primaries collide with nuclei in the Earth's atmosphere, and some penetrate to sea level. It was among these that muons were discovered. Subsequently, they have been copiously produced at accelerator laboratories, enabling their properties to be studied in great detail. Muons have the Dirac magnetic moment

$$\boldsymbol{\mu} = \frac{e}{m_\mu}\mathbf{S} \tag{2.9}$$

appropriate to point-like spin-$\frac{1}{2}$ particles, and in general their electromagnetic properties are identical with those of electrons, *provided the mass difference is taken into account.* In particular, it is their much greater mass that is the reason for their much greater penetrating power in matter compared to electrons.[6]

The tauon is even heavier ($m_\tau = 1777\,\mathrm{MeV/c^2}$) and in 1975 was discovered in electron–positron annihilation experiments at high energies. Its properties have been studied in less detail than those of the muon, but are compatible with a point-like spin-$\frac{1}{2}$ particle whose electromagnetic interactions are identical with those of the electron and muon.

Because the electron is the lightest charged particle, conservation of electric charge means it is necessarily stable. Both the muon and the tauon, however, are unstable, with lifetimes 2.2×10^{-6} s and 2.9×10^{-13} s, respectively. Both decay by weak interactions, and the great difference in their lifetimes is a consequence of the mass difference, as we shall see below. In the case of the muon, the decay is purely leptonic

$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu; \quad \mu^- \to e^- + \bar{\nu}_e + \nu_\mu \tag{2.10}$$

and conserves both charge and lepton number. For the tauon, many decay modes are observed, most of them involving hadrons in the final state, although the purely leptonic modes

$$\tau^- \to e^- + \bar{\nu}_e + \nu_\tau \tag{2.11a}$$

and

$$\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau \tag{2.11b}$$

are also observed. The relative importance of any decay mode is measured by its *branching ratio B*, defined as the fraction of all decays leading to that particular final state. For the leptonic decays (2.11a) and (2.11b), the measured branching ratios are 0.178 and 0.174, respectively.

The neutrinos emitted in the decays (2.10) and (2.11) were not directly observed, but, like those in β-decay, are inferred from energy and angular momentum conservation. Both ν_μ and ν_τ have, however, subsequently been detected in other reactions.

---

[6] High-energy electrons lose energy in matter dominantly by radiative collisions. The rate of energy loss in such collisions is approximately proportional to m^{-2}. Consequently, it is heavily suppressed for muons, which travel much further. This is discussed in Chapter 4.

Well-defined muon–neutrino beams can be created in the laboratory and used to study reactions like the elastic scattering from electrons

$$\nu_\mu + e^- \rightarrow \nu_\mu + e^-, \tag{2.12}$$

inverse muon decay

$$\nu_\mu + e^- \rightarrow \mu^- + \nu_e \tag{2.13}$$

and reactions on nucleons, such as

$$\nu_\mu + p \rightarrow \nu_\mu + p \tag{2.14a}$$

and

$$\bar{\nu}_\mu + p \rightarrow \mu^+ + n. \tag{2.14b}$$

Well-defined tau-neutrino beams are not available in the laboratory and it was not until 2000 that tau-neutrinos were directly detected. Only a few direct interaction events have ever been observed and they supplement evidence based on energy and angular momentum conservation. The masses of both ν_μ and ν_τ can be inferred from the e^- and μ^- energy spectra in the decays (2.10) and (2.11) using energy conservation. The results from this and other more accurate methods are consistent with zero masses, although the errors are much larger than those for electron neutrinos. The present limits are (at 90 % confidence limit)

$$m_\mu < 0.19\,\mathrm{MeV/c^2}, \quad m_\tau < 18.2\,\mathrm{MeV/c^2}. \tag{2.15}$$

Underlying the whole of this discussion is the assumption that lepton numbers are individually conserved. It is these that fix the identity (as ν_μ, ν_τ, etc.) of the undetected neutrinos in, for example, muon decay (2.10). One way of testing their validity is to search for reactions that violate lepton number conservation, but satisfy all other conservation laws, such as $\mu^\pm \rightarrow e^\pm + \gamma$ and $\nu_e + e^- \rightarrow \mu^- + \nu_e$. Examples of lepton number violating decays of this type and the experimental upper limits on the branching ratios, above which they would have been detected, are shown in Table 2.1. These tests are impressive evidence for lepton number conservation, and no violations have ever been detected in any particle decays.

TABLE 2.1 Examples of leptonic decays that violate conservation of lepton numbers and the experimental upper limits on their branching ratios B.

Decay	Violates	B
$\mu^- \rightarrow e^- + e^+ + e^-$	L_μ, L_e	$<1.0 \times 10^{-12}$
$\mu^- \rightarrow e^- + \gamma$	L_μ, L_e	$<1.2 \times 10^{-11}$
$\tau^- \rightarrow e^- + \gamma$	L_τ, L_e	$<1.1 \times 10^{-7}$
$\tau^- \rightarrow \mu^- + \gamma$	L_τ, L_μ	$<6.8 \times 10^{-8}$
$\tau^- \rightarrow e^- + \mu^+ + \mu^-$	L_τ, L_e	$<2 \times 10^{-7}$

2.2 LEPTONIC WEAK INTERACTIONS

We begin the study of the dynamics of weak interactions by focusing for the moment on simple leptonic reactions like (2.10) to (2.13). Our discussion will be based on the general ideas of particle exchange introduced in Sections 1.3 and 1.4, and in particular on the zero-range approximation of Figure 1.15.

2.2.1 $W^{\pm}$ and Z^0 exchange

In Section 1.3, we saw that elastic electron–electron scattering is described by the exchange diagrams of Figures 1.7(a) and (b), in which a photon is emitted by one electron and absorbed by another. In the same way, weak reactions involving only leptons are described by exchange processes in which a $W^{\pm}$ or Z^0 is emitted by one lepton and absorbed by another. In drawing such diagrams, we must remember that these absorption and emission processes conserve the lepton numbers (2.3), as well as electric charge, where the $W^{\pm}$ and Z^0 have zero lepton numbers.

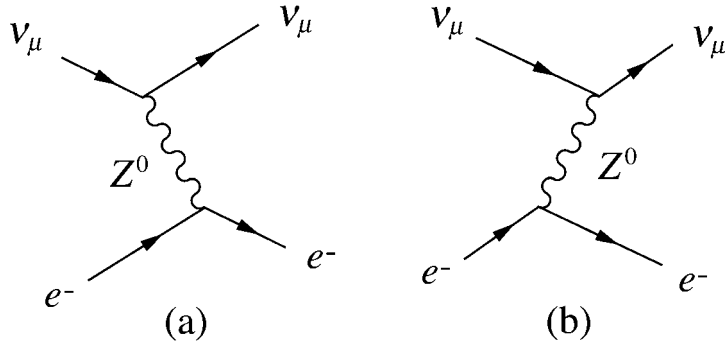


Figure 2.3 Two time-ordered diagrams for elastic $\nu_\mu e^-$ scattering (2.12), where as usual time flows from left to right. By convention, only one of these diagrams is usually drawn, leaving the existence of the other implied.

A simple example is elastic $\nu_\mu e^-$ scattering (2.12), for which the two possible time-ordered Feynman diagrams are shown in Figure 2.3. These are clearly closely analogous to the two time-ordered diagrams of Figures 1.7(a) and (b) for e^-e^- scattering by photon exchange and, as in that case, it is conventional to draw one such diagram, leaving the other implied.[7] The corresponding diagrams for inverse muon decay (2.13) are shown in Figure 2.4. In Figure 2.4(a), the initial process corresponds to the lower vertex, in which an electron converts into an electron–neutrino by emitting a W^- boson, conserving both electric charge and lepton numbers. The W^- is then absorbed at the upper vertex, again conserving both charge and lepton numbers. Figure 2.4(a) therefore corresponds to W^- exchange. On the other hand, application of the conservation laws to the second time-ordered diagram of Figure 2.4(b) shows that it corresponds to W^+ exchange. Nevertheless, it is still conventional to draw just

[7] This convention was introduced for electromagnetic interactions in Section 1.3.3.

one such diagram, leaving the other implied. Thus inverse muon decay is described by just one of the Feynman diagrams of Figure 2.4, while muon decay itself is described by the analogous diagram of Figure 2.5. In the latter, the first vertex corresponds to a muon emitting a W^- boson, followed by the conversion of the W^- boson into a lepton pair in the lepton-number conserving process $W^- \to e^- \bar{\nu}_e$, giving the muon decay reaction (2.10) overall.

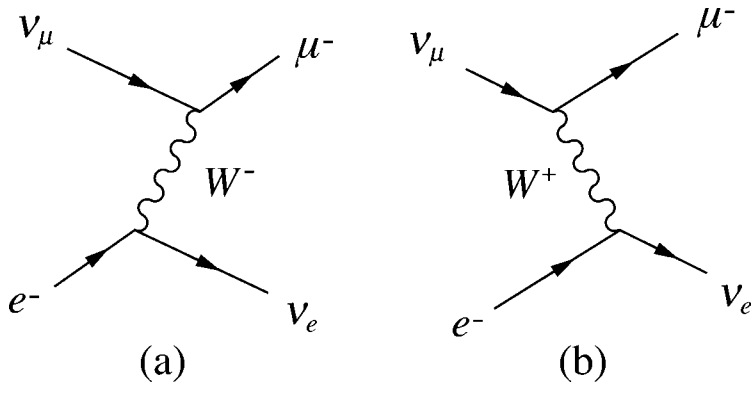


Figure 2.4 Dominant Feynman diagrams for inverse muon decay.

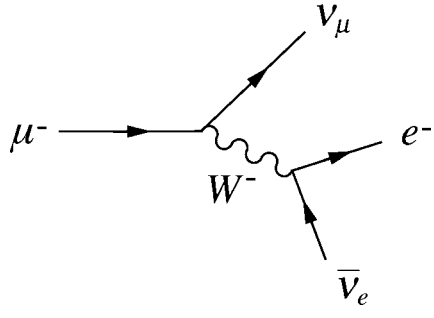


Figure 2.5 Dominant Feynman diagram for muon decay.

These processes can also occur via higher-order diagrams (i.e. ones containing more vertices) like that shown in Figure 2.6 for inverse muon decay, involving the exchange of two vector bosons. In such diagrams, each vertex corresponds to a basic process whose intrinsic probability is much less than unity, as we shall see shortly. Thus the contribution of Figure 2.6, which has four vertices, is expected to be much smaller that that of Figure 2.5, which has only two.[8] Detailed calculations confirm the dominance of Figure 2.5 and from now on we shall only consider the dominant lowest-order diagrams contributing to a given process.

At low energies, the de Broglie wavelengths of all the particles involved are large compared with the range $R_W \approx 2 \times 10^{-3}$ fm of the W-exchange interaction. This can

[8] This should be compared with the discussion of multiphoton exchange at the end of Section 1.3.3.

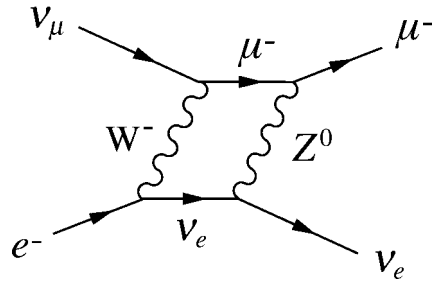


Figure 2.6 Example of a higher-order contribution to inverse muon decay.

then be approximated by a zero-range point interaction whose strength is characterized by the Fermi coupling constant (1.40), i.e.

$$G_F = 1.166 \times 10^{-5}\ \text{GeV}^{-2}. \tag{2.16}$$

The origin of this approximation is illustrated for inverse muon decay in Figure 2.7, and the corresponding diagram for muon decay itself is shown in Figure 2.8. Figure 2.7 is similar to Figure 1.15 that was discussed for spin-0 particles in Section 1.4, and it can be shown that when spin complications are taken into account, the coupling constant relation (1.39) derived for spinless particles becomes

$$\frac{G_F}{\sqrt{2}} = \frac{g_W^2}{M_W^2} = \frac{4\pi\alpha_W}{M_W^2} \tag{2.17}$$

in natural units. Here, g_W is the coupling constant associated with the W boson–lepton vertices and

$$\alpha_W \equiv g_W^2/4\pi \tag{2.18}$$

is the dimensionless strength parameter introduced by analogy with the fine structure constant $\alpha = e^2/4\pi\varepsilon_0$ of electromagnetism. On substituting the measured values for M_W and G_F into (2.17), we find that

$$\alpha_W = 4.2 \times 10^{-3} = 0.58\alpha. \tag{2.19}$$

Hence the weak and electromagnetic interactions are of comparable intrinsic strength, and weak interaction rates are only small at low energies because of the large mass of the W boson, which enters the low-energy effective coupling constant (2.17) as the inverse square.

2.2.2 Lepton decays and universality

All known experimental data are consistent with the assumption that the interactions of the electron and its neutrino are identical with those of the muon and its associated neutrino and the tauon and its neutrino, provided the mass differences are taken into account. This fundamental assumption is called the *universality* of lepton

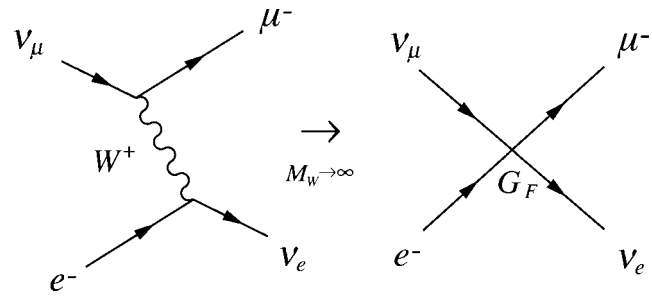


Figure 2.7 Origin of the low-energy zero-range interaction in inverse muon decay (2.13).

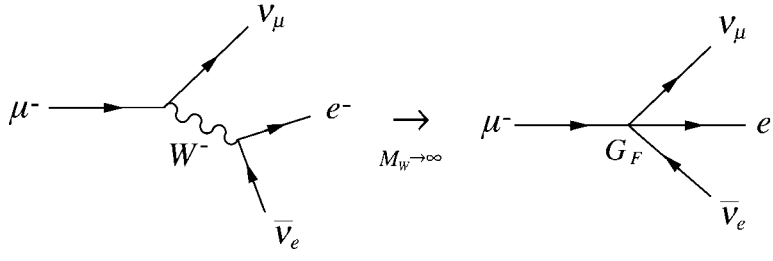


Figure 2.8 Origin of the low-energy zero-range interaction in muon decay (2.10).

interactions and was discussed briefly, for the case of electromagnetic interactions, in Section 2.1.2. It was also implicit in the previous section, where we assumed that the coupling strengths g_W and G_F were independent of the nature of the leptons to which the W boson coupled.

In the rest of this subsection, we will illustrate universality by considering the leptonic decays (2.10) and (2.11) of the muon and tauon at rest. As usual, we will work to lowest order only and we will use the zero-range approximation, since the masses of the leptons, and hence the energies of all the decay products, are very small compared with the rest energy of the W-bosons. We shall also assume that we can, to a good approximation, neglect the masses of the leptons in the final state relative to the mass of the decaying μ or τ. Since $m_e \ll m_\mu \ll m_\tau$, we would naively expect this approximation to give rise to errors of order $(m_e/m_\mu) \approx 5 \times 10^{-3}$ in the decay (2.10), of order $(m_e/m_\tau) \approx 2 \times 10^{-5}$ in the electronic decay of the tauon (2.11a) and of order $(m_\mu/m_\tau) \approx 5 \times 10^{-2}$ in the muonic decay of the tau (2.11b). In fact the errors are smaller than this, but this cannot be seen without a full calculation.

We start by considering muon decay (2.10). In the zero-range approximation, the amplitude is of order G_F, which has a natural dimension $[E]^{-2}$. The only other dimensional constant involved is the muon mass, because we are setting the electron and neutrino masses to zero. It follows that the decay rate must be given by an expression of the form

$$\Gamma(\mu^- \to e^- + \bar{\nu}_e + \nu_\mu) = K G_F^2 m_\mu^5, \tag{2.20a}$$

because the rate has a natural dimension $[E]$. Here, K is a dimensionless constant whose value will depend on the precise form of the interaction. If we assume

this is the same for muons and tauons ($\mu - \tau$ universality) the same argument gives

$$\Gamma(\tau^- \to e^- + \bar{\nu}_e + \nu_\tau) = KG_F^2 m_\tau^5, \tag{2.20b}$$

where K is the same constant, while $e - \mu$ universality gives

$$\Gamma(\tau^- \to e^- + \bar{\nu}_e + \nu_\tau) = \Gamma(\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau).$$

This explains why the experimental branching ratios for the two leptonic decay modes of the tauon (2.11a) and (2.11b) are, to a good approximation, equal. It also gives a relation between the μ and τ lifetimes. The lifetimes τ_ℓ are related to decay rates by[9]

$$\tau_l = \frac{1}{\Gamma_{tot}} = \frac{B(l^- \to e^- \bar{\nu}_e \nu_l)}{\Gamma(l^- \to e^- \bar{\nu}_e \nu_l)}, \tag{2.21}$$

where Γ_{tot} is the total decay rate and

$$B(l^- \to e^- \bar{\nu}_e \nu_l) \equiv \frac{\Gamma(l^- \to e^- \bar{\nu}_e \nu_l)}{\Gamma_{tot}}$$

is the branching ratio previously defined. Experimentally, $B = 1$ and 0.178 ± 0.004 for $l = \mu$ and τ, respectively. Thus from (2.20a) and (2.20b) we have

$$\frac{\tau_\tau}{\tau_\mu} = \frac{B(\tau^- \to e^- \bar{\nu}_e \nu_\tau)}{B(\mu^- \to e^- \bar{\nu}_e \nu_\mu)} \left(\frac{m_\mu}{m_\tau}\right)^5 = (1.328 \pm 0.004) \times 10^{-7},$$

which is consistent with the ratio of the experimental lifetimes $(1.3227 \pm 0.0005) \times 10^{-7}$.

This agreement, involving lifetimes that differ by seven orders of magnitude, is impressive evidence of the universality of lepton interactions. More generally, all known experimental data are consistent with the assumption that the interactions of the electron and its associated neutrino are identical with those of the muon and its associated neutrino and of the tauon and its neutrino. The three generations of leptons tell not three stories but, in all essential points, one story three times.

2.3 NEUTRINO MASSES AND NEUTRINO MIXING

At present we have no understanding of why the observed leptons have the masses they do. The neutrinos could have very small but finite masses compatible with the experimental bounds given in Equations (2.7) and (2.15), or zero masses as was originally assumed in the standard model. This question is important, because nonzero neutrino masses can lead to nontrivial effects that are not possible for zero masses,

[9] See Equation (B.31) of Appendix B.

and because these effects, and the nonzero masses themselves, can have important consequences in astrophysics and cosmology. (The latter will be discussed briefly in Chapter 11).

Firstly, however, it will be convenient to extend our quantum mechanical notation slightly. So far we have described a single-particle state by naming the particle, e.g. π^+, and specifying its wavefunction Ψ, which determines its other properties (spin, momentum, etc.). We now combine these in a single notation, i.e. $|\pi^+, \Psi\rangle$. Alternatively, instead of a wavefunction, we may just designate those observables that are sufficient to uniquely determine the wavefunction. For a spinless particle, momentum is sufficient, so that $|\pi^+, \mathbf{p}\rangle$ specifies both the particle and its wavefunction up to an arbitrary overall phase.[10] Multiparticle states are specified by, for example,

$$\left|\pi^+, \Psi_1; \pi^-, \Psi_2\right\rangle \equiv \left|\pi^+, \Psi_1\right\rangle \left|\pi^-, \Psi_2\right\rangle,$$

by analogy with product wavefunctions

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2).$$

2.3.1 Neutrino mixing

One of the new phenomena that can occur if neutrinos have nonzero masses is *neutrino mixing*. This is the assumption that the neutrino states ν_e, ν_μ and ν_τ that couple to electrons, muons and tauons, respectively, do not have definite masses; instead they are linear combinations of three other states ν_1, ν_2 and ν_3 that do have definite masses m_1, m_2 and m_3.

Instead of considering mixing between all three 'flavour' states ν_e, ν_μ and ν_τ, it is simpler, and often a good approximation, to consider the mixing between just two of them, which we will denote ν_α and ν_β. Then, in order to preserve the orthonormality of the states, we can write

$$\nu_\alpha = \nu_i \cos\theta_{ij} + \nu_j \sin\theta_{ij} \tag{2.22a}$$

and

$$\nu_\beta = -\nu_i \sin\theta_{ij} + \nu_j \cos\theta_{ij}, \tag{2.22b}$$

where ν_i and ν_j are the two mass eigenstates involved. Here ν_α is shorthand for $|\nu_\alpha, \psi\rangle$, etc., and θ_{ij} is a *mixing angle* that must be determined from experiment.

If $\theta_{ij} = 0$, then $\nu_\alpha = \nu_i$, $\nu_\beta = \nu_j$ and there is no mixing. However, if $\theta_{ij} \neq 0$ then some interesting predictions follow. When, for example, a ν_α neutrino is produced with momentum $\mathbf{p}$ at time $t = 0$, the ν_i and ν_j components will have slightly different energies E_i and E_j due to their slightly different masses. In quantum mechanics, their associated waves will therefore have slightly different frequencies, giving rise to a

[10] This is part of the 'Dirac notation' in quantum mechanics. However, no knowledge of this notation is assumed or required beyond that stated above.

phenomenon somewhat akin to the 'beats' heard when two sound waves of slightly different frequency are superimposed. As a result of this, one finds that the original beam of ν_α particles develops a ν_β component whose intensity oscillates as it travels through space, while the intensity of the ν_α neutrino beam itself is correspondingly reduced. These are called *neutrino oscillations* and their existence follows from simple quantum mechanics.

To illustrate this we will consider a ν_α produced with momentum $\mathbf{p}$ at time $t = 0$, so that the initial state (2.22a) is

$$|\nu_\alpha, \mathbf{p}\rangle = |\nu_i, \mathbf{p}\rangle \cos\theta_{ij} + \left|\nu_j, \mathbf{p}\right\rangle \sin\theta_{ij}. \tag{2.23a}$$

After time t this will become

$$a_i(t)\ |\nu_i, \mathbf{p}\rangle \cos\theta_{ij} + a_j(t)\left|\nu_j, \mathbf{p}\right\rangle \sin\theta_{ij}, \tag{2.24}$$

where

$$a_i(t) = e^{-iE_i t} \quad \text{and} \quad a_j(t) = e^{-iE_j t} \tag{2.25}$$

are the usual oscillating time factors associated with any quantum mechanical stationary state.[11] For $t \neq 0$, the linear combination (2.24) does not correspond to a pure ν_α state, but can be written as a linear combination

$$A(t)\,|\nu_\alpha, \mathbf{p}\rangle + B(t)\left|\nu_\beta, \mathbf{p}\right\rangle, \tag{2.26}$$

of ν_α and ν_β states, where the latter is given by (2.22b), i.e.

$$\left|\nu_\beta, \mathbf{p}\right\rangle = -|\nu_i, \mathbf{p}\rangle \sin\theta_{ij} + \left|\nu_j, \mathbf{p}\right\rangle \cos\theta_{ij}\,. \tag{2.23b}$$

The functions $A(t)$ and $B(t)$ are found by solving (2.23a) and (2.23b) for $|\nu_i, \mathbf{p}\rangle$ and $\left|\nu_j\,, \mathbf{p}\right\rangle$, substituting the results into (2.24) and comparing with (2.26). This gives

$$A(t) = a_i(t)\cos^2\theta_{ij} + a_j(t)\sin^2\theta_{ij} \tag{2.27a}$$

and

$$B(t) = \sin\theta_{ij}\cos\theta_{ij}\left[a_j(t) - a_i(t)\right]. \tag{2.27b}$$

The probability of finding a ν_β state is therefore, using (2.25),

$$P(\nu_\alpha \to \nu_\beta) = |B(t)|^2 = \sin^2(2\theta_{ij})\sin^2\left[\tfrac{1}{2}(E_j - E_i)\,t\right] \tag{2.28}$$

[11] See, for example, Chapter 1 of Mandl (1992).

and thus oscillates with time, while the probability of finding a ν_α particle is reduced by a corresponding oscillating factor. Irrespective of which neutrino states ν_α and ν_β are involved, the oscillations vanish if the mixing angle is zero, or if the corresponding mass eigenstates $\nu_{i,j}$ have equal masses, and hence equal energies, as can be seen explicitly from (2.28). In particular, such oscillations are not possible if ν_i and ν_j both have zero masses.

These formulas assume that the neutrinos are propagating in a vacuum. This is usually a very good approximation, because of the enormous mean free paths for neutrinos to interact with matter. However, it can be shown that such oscillations can be enhanced when neutrinos traverse very long distances in matter, such as from the interior of the Sun to its surface. This is a result of the fact that electron neutrinos interact with electrons in a different way from muon and tauon neutrinos (see Problem 2.4), and is an important factor when solar neutrino data are analysed, as we shall see.

2.3.2 Neutrino oscillations

Attempts to detect neutrino oscillations rest upon the use of lepton number conservation to identify the neutrinos emitted or absorbed in any given reaction as electron, muon or tau-neutrinos.[12] For example, electron neutrinos can produce electrons via reactions like $\nu_e + n \to e^- + p$ but cannot produce muons; whereas muon neutrinos can produce muons via reactions like $\nu_\mu + n \to \mu^- + p$, but not electrons. In addition, the time t is determined by the distance L of the neutrino detector from the source of the neutrinos, since their momenta are always much greater than their possible masses and they travel, to a very good approximation, at the speed of light. In this approximation, $t = L$,

$$E_j - E_i = \left(m_j^2 + p^2\right)^{1/2} - \left(m_i^2 + p^2\right)^{1/2} \approx \frac{m_j^2 - m_i^2}{2p}, \tag{2.29}$$

where $p = |\mathbf{p}|$, so that (2.28) may be written

$$P(\nu_\alpha \to \nu_\beta) \approx \sin^2(2\theta_{ij}) \sin^2 [L/L_0], \tag{2.30a}$$

where the oscillation length

$$L_0 = \frac{4E}{(m_j^2 - m_i^2)} \tag{2.31}$$

and $E = p$, with

$$P(\nu_\alpha \to \nu_\alpha) = 1 - P(\nu_\alpha \to \nu_\beta). \tag{2.30b}$$

[12] We emphasize again that it is ν_e, ν_μ and ν_τ that have definite values of the lepton numbers and not the mass eigenstates ν_1, ν_2 and ν_3.

As we shall see, these lengths are typically of order 100 km or more, so that oscillations can be safely neglected under normal laboratory conditions. Nevertheless, neutrino oscillations have been detected in several experiments. These experiments are conveniently divided into those that start from muon neutrinos or antineutrinos (*muon neutrino oscillations*) and those that start with electron neutrinos or antineutinos (*electron neutrino oscillations*).

2.3.2(a) Muon neutrino oscillations

The first experiment to produce definitive evidence for neutrino oscillations was that of the *SuperKamiokande* group, who in 1998 used a giant detector to study *atmospheric neutrinos*. When cosmic ray protons collide with nuclei in the upper atmosphere they create many pions and, as we shall see presently, the decay sequence of pions in turn creates neutrinos, with two muon neutrinos being produced for every electron neutrino.

The *SuperKamiokande* detector is shown in Figure 2.9. It consists of a stainless steel cylindrical tank of roughly 40 m in diameter and 40 m in height, containing about 50 000 metric tons of very pure water. The detector is situated deep underground in the Japanese Alps, at a depth equivalent to 2700 m of water. This is done to use the rocks above to shield the detector from cosmic ray muons. The volume is separated into two regions. The walls of the large inner region are lined with 11 200 light-sensitive devices called photomultipliers. When neutrinos with energies above 1 GeV interact with nuclei in the water, the velocities of the electrons and muons produced are greater than the speed of light in water. Because of this, a shock wave of light, called Čerenkov radiation, is emitted. This is analogous to the shock wave emitted when an aircraft exceeds the speed of sound in air. This Čerenkov radiation is detected by the photomultipliers and used to infer properties of the particles that produced it.[13] The outer region of water acts as a shield against low-energy particles entering the detector from outside. An additional 1200 photomultipliers are located there to detect muons that enter or exit the detector.

As indicated above, in the absence of neutrino oscillations, one would expect twice as many muon neutrinos as electron neutrinos to be detected. However, the measured ratio was about 1.3 on average, suggesting the presence of oscillations. This was confirmed by exploiting the fact that the detector could measure the direction of the detected neutrinos to study the azimuthal dependence of the effect. Since the flux of cosmic rays that lead to neutrinos with energies above about 1 GeV is isotropic, the production rate for neutrinos should be the same all around the Earth. In particular, one can compare the measured flux from neutrinos produced in the atmosphere directly above the detector, which have a relatively short flight path before detection, with those incident from directly below, which have travelled a long way through the Earth before detection, and so have had plenty of time to oscillate (perhaps several cycles). Experimentally, it was found that the yield of electron neutrinos from above and below were the same within errors and consistent with expectation for no oscillations. However, while the yield of muon neutrinos from above accorded with the expectation

[13] Čerenkov radiation and other aspects of particle detection are discussed in more detail in Chapter 4.

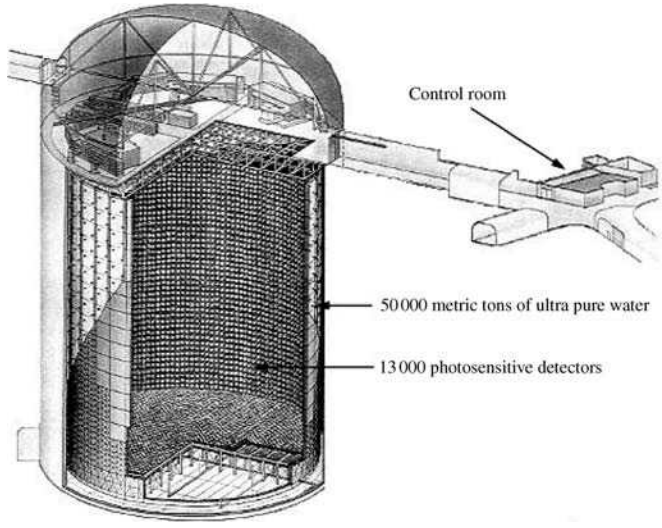


Figure 2.9 A schematic diagram of the SuperKamiokande detector. (Adapted from an original illustration used with permission from University of Hawaii, Manoa.)

for no significant oscillations, the flux of muon neutrinos from below was a factor of about two lower. This is clear evidence for muon neutrino oscillations.

In a later development of the experiment, the flux of muon neutrinos was measured as a function of L/E by estimating L from the reconstructed neutrino direction. Values of L range from 15 km to 13 000 km. The results are shown in Figure 2.10 in the form of the ratio of $P(\nu_\mu \rightarrow \nu_\mu)$ of the observed number of events to the theoretical expectation if there were no oscillations. The data show clear evidence for a deviation of this ratio from unity, particularly at large values of L/E.

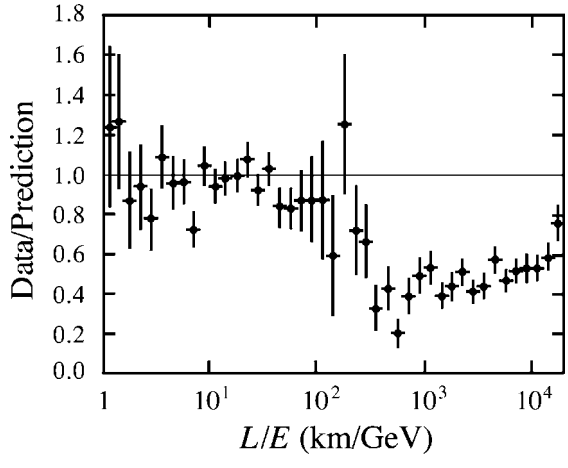


Figure 2.10 Data from the SuperKamiokande detector showing evidence for neutrino oscillations in atmospheric neutrinos. See text for details. (Reprinted Figure 4 with permission from Y. Ashie *et* al., *Phys. Rev. Lett*., **93**, 101801. Copyright 2004 American Physical Society.)

Other experiments set limits on $P(\nu_\mu \to \nu_e)$ and taking these into account the most plausible hypothesis is that muon neutrinos are changing into tau-neutrinos, which for the neutrino energies concerned could not be detected by SuperKamiokande. The data are consistent with this hypothesis. If we define

$$\Delta(m_{ij}^2) \equiv m_i^2 - m_j^2,$$

then the experiment yields the values

$$1.9 \times 10^{-3} \lesssim \left|\Delta(m_{32}^2)\right| \lesssim 3.0 \times 10^{-3}\ (\text{eV})^2, \quad \sin^2(2\theta_{23}) \gtrsim 0.9, \tag{2.32}$$

i.e. $\theta_{23} \gtrsim 36^\circ$, at 90 % confidence level.[14] This conclusion is supported by results obtained in 2006 from the MINOS laboratory-based experiment that starts with a beam of ν_μ and measures the flux at a large distance (250 km) from the place of origin of the beam. Analysis of the data yields parameters consistent with those above.

2.3.2(b) Electron neutrino oscillations

Electron neutrino oscillations were first established by a series of observations of *solar neutrinos*. The energy of the Sun is due to various nuclear reactions and these produce a huge flux of low-energy ($E \leq 15$ MeV) electron neutrinos that can be detected at the surface of the Earth. This was first done by Davis and co-workers in an experiment that began in 1968 and detected neutrinos via the reaction

$$\nu_e + {}^{37}\text{Cl} \to e^- + {}^{37}\text{Ar}. \tag{2.33}$$

The principal component of this experiment is a huge tank containing tetrachloroethylene (C_2Cl_4) in which a single argon-37 atom is produced on average every few days by the reaction (2.33). These atoms are unstable, with a half-life of 35 days. They are extracted by flushing the tank with helium gas every few weeks and counted by observing their decays. The experiment was located deep underground, in a gold mine in South Dakota, to reduce the number of background events in which argon-37 atoms are produced by reactions involving cosmic ray muons rather than neutrinos. The inferred neutrino flux is expressed in terms of the 'solar neutrino unit' SNU (pronounced 'snew'), defined as one capture event per second for every 10^{36} target atoms. After more than twenty years of running, the neutrino flux was measured to be $2.55 \pm 0.17 \pm 0.18$ SNU, where the first error is statistical and the second systematic. This measured rate is much smaller than the expected rate, 7.3 ± 2.3 SNU, which is predicted by the accepted model of the Sun, called the 'standard solar model'. The discrepancy between these two values constituted the 'solar neutrino problem' and was confirmed by a second experiment, called Kamiokande II, which was a smaller precursor of the SuperKamiokande detector described above. Unlike the ^{37}Cl experiment, this experiment is sensitive to the direction of the incoming neutrinos, and verifies that they come from the direction of the Sun.

[14] We label the two neutrinos involved in this experiment as ν_2 and ν_3, rather than ν_1 and ν_2, to conform with the convention used by the Particle Data Group (Yao *et al.*, 2006).

The solution of the solar neutrino problem could be neutrino oscillations or errors in the solar model. In regard to the latter, it is worth noting that both the argon-37 experiment and Kamiokande II could only detect neutrinos if they had an energy in excess of approximately 1 MeV. Such neutrinos come predominantly from the weak β-decay

$$^8\mathrm{B} \to {}^8\mathrm{Be} + e^+ + \nu_e, \tag{2.34}$$

with an average energy of about 7 MeV. However, this decay contributes only about 10^{-4} of the total solar neutrino flux. It is therefore important to detect neutrinos from other reactions and in particular from the reaction

$$p + p \to d + e^+ + \nu_e, \tag{2.35}$$

which is the primary reaction that produces the energy of the Sun and contributes approximately 90 % of the flux of solar neutrinos. The neutrinos produced in this reaction have average energies of about 0.26 MeV and so cannot be detected in the above experiments. Instead, the reaction

$$\nu_e + {}^{71}\mathrm{Ga} \to {}^{71}\mathrm{Ge} + e^- \tag{2.36}$$

has been used, which has a threshold of 0.23 MeV. Just as for the original experiments of Davis *et al*., there were formidable problems in identifying the radioactive products from this reaction, which produces only about 1 atom of ^{71}Ge per day in a target of 30 tons of gallium. Nevertheless, results from these experiments, called SAGE and GALLEX, confirmed the deficit of electron neutrinos and found between 60 and 70 % of the flux expected in the absence of oscillations. Since the predicted neutrino flux arising from (2.35) is insensitive to the details of the standard solar model, it is unlikely that shortcomings in the latter could be the source of this discrepancy.[15]

That neutrino oscillations really are the solution to the solar neutrino problem was definitively established by an experiment at the Sudbury Neutrino Observatory (SNO) in Canada in 2002. The experiment consisted of a stainless steel sphere of radius 9 m located 6800 ft underground, and is shown during its construction in coloured Plate 8. It used a water Čerenkov detector, like Kamiokande and SuperKamiokande, but instead of normal water it used heavy water D_2O and was therefore able to study the reactions

$$\text{(a) } \nu_e + d \to e^- + p + p, \quad \text{(b) } \nu_x + d \to \nu_x + p + n, \quad \text{(c) } \nu_x + e^- \to \nu_x + e^-, \tag{2.37}$$

where x denotes any lepton flavour (e, μ, τ) and d is the deuteron. The cross-section for (b) is independent of the lepton flavour (this is a consequence of 'lepton universality' discussed in Section 2.2.2) and hence independent of any possible oscillations. Since the observed flux was consistent with expectations, this confirms the correctness of the solar model. On the other hand, the observed flux from (a) was only about

[15] For a simple discussion of energy and neutrino generation in stars, see, for example, Section 4.2 of Phillips (1994).

one-third of expectations, implying that about two-thirds of the electron neutrinos originally produced had transformed to μ and /or τ neutrinos before being detected at the surface of the Earth. The flux for (c) would then be due to a mixture of approximately one-third of electron neutrinos and two-thirds of μ/τ neutrinos. Because the cross-section for $\nu_x e^- \to \nu_x e^-$ is different for ν_e and $\nu_{\mu,\tau}$ (see Problem 2.4), the rate for (2.37c) is below what would be expected if there were no oscillations. The data were consistent with this assumption.

When these solar neutrino results were analysed, taking into account their interactions with matter, they implied that a substantial fraction of a beam of $\bar{\nu}_e$ from a nuclear reactor would change to antineutrinos of other flavours after travelling a distance of order 100 km from its source. This prediction has been tested by the KamLAND group in Japan. They have studied the $\bar{\nu}_e$ flux from more than 60 reactors in Japan and South Korea after the neutrinos have travelled distances of between 150 and 200 km. They found that the $\bar{\nu}_e$ flux was only about 60 % of that expected from the known characteristics of the reactors. A simultaneous analysis of the data from this experiment and the solar neutrino data using two-component mixing yields the results

$$7.6 \times 10^{-5} \lesssim \left|\Delta(m^2_{21})\right| \lesssim 8.6 \times 10^{-5}\,(\text{eV})^2, \quad 0.32 \gtrsim \tan^2(\theta_{12}) \lesssim 0.48, \tag{2.38}$$

i.e. $29^\circ \lesssim \theta_{12} \lesssim 35^\circ$.

The existence of neutrino oscillations (flavour changing), and by implication nonzero neutrino masses, is now generally accepted on the basis of the above set of experiments.

2.3.3 Neutrino masses

In order to extract reliable information from all the oscillation data, including that described above, it is necessary to extend the discussion of Section 2.3.1 to include mixing between all three neutrino mass states ν_1, ν_2 and ν_3, rather than just between two. We will not discuss this in detail, except to say that the extended scheme requires two squared mass differences Δm^2_{21} and Δm^2_{32} and three mixing angles θ_{12}, θ_{13} and θ_{23} to describe the mixing.[16] A global fit to all the data yields values of $\Delta m^2_{32}, \theta_{23}, \Delta m^2_{21}$ and θ_{12} that are consistent with those given in Equations (2.32) and (2.38), together with

$$\sin^2(2\theta_{13}) \lesssim 0.19, \tag{2.39}$$

the latter constraint coming mainly from measurements on $\bar{\nu}_e$ using the Chooz nuclear reactor in France. We will discuss the implications of these results in what follows.

We first note that, for the solar neutrino data, the interactions with matter play an important role. As a bonus, this enables the sign of Δm^2_{21} to be measured, whereas oscillations in free space only determine the magnitude of Δm^2_{ij}, as can be seen from (2.30) and (2.31). Because of this, the sign of Δm^2_{32} is not determined and two solutions for the mass hierarchy are possible: the so-called 'normal' mass hierarchy, $m_3 > m_2 > m_1$, and the 'inverted' mass hierarchy, $m_2 > m_1 > m_3$. The former case is

[16] There is also a phase angle δ that we will discuss later in Chapter 10 in the context of possible violation of *CP* invariance.

shown in Figure 2.11, where we also show the approximate flavour decomposition of the mass eigenstates resulting from three-component mixing with parameters compatible with Equations (2.32), (2.38) and (2.39). As can be seen, the lighter of the two states that dominate solar neutrino oscillations is predominately an electron neutrino, while ν_3 has large ν_μ and ν_τ components, but little or no electron neutrino component.

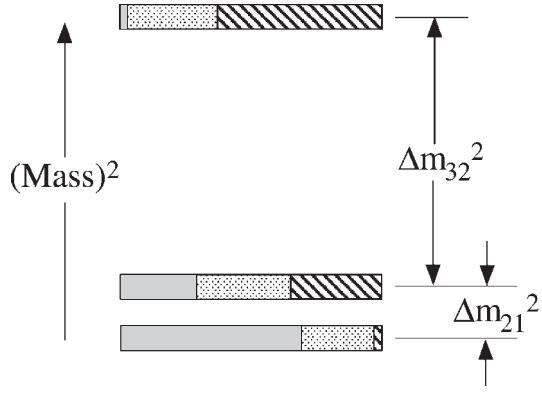


Figure 2.11 A three-neutrino squared-mass spectrum, assuming the normal mass hierarchy, that is consistent with the values for masses and mixing angles given in Equations (2.32), (2.38) and (2.39). Also shown is the approximate fractional flavour compositions of ν_1, ν_2 and ν_3 (ν_e (solid), ν_μ (dotted), ν_τ (hatched)).

We can now return to the interpretation of the bound (2.7) and consider its consequences for the 'mass' of the electron neutrino. The point here is that neutrinos with definite flavours, like the electron neutrino, are superpositions of the mass eigenstates ν_1, ν_2 and ν_3 and do not themselves have definite masses. Rather, in accord with the standard theory of measurement in quantum mechanics, a measurement of the mass of the electron neutrino can yield any one of the three values m, m_2 or m_3. However, if $m_i^2 > m_j^2$, one can easily show that $(m_i - m_j)^2 < m_i^2 - m_j^2$. Hence Equations (2.32) and (2.38) set upper limits on the neutrino mass differences that are approximately

$$m_2 - m_1 \lesssim 10^{-2}\,\mathrm{eV/c^2}, \quad |m_3 - m_2| \lesssim 5 \times 10^{-2}\,\mathrm{eV/c^2}. \tag{2.40}$$

In other words, the neutrino masses are almost equal compared to a mass scale of $1\,\mathrm{eV/c^2}$. Hence it is safe to interpret (2.7) as implying

$$m_i \leq 2\,\mathrm{eV/c^2} \tag{2.41}$$

for all three neutrinos ν_1, ν_2 and ν_3. This limit is very small compared to the mass of the lightest charged particle, the electron, and is a great improvement on the limits (2.15) obtained from muon and tauon decay experiments.

Finally, it is worth remarking that it is possible to obtain bounds on neutrino masses from cosmology. The most exacting of these comes from using the current standard cosmological model to analyse the large-scale structure of the universe. The bound is

$$\sum m_\nu = (0.5 - 1.0)\,\mathrm{eV/c^2}, \tag{2.42}$$

where the sum is over all neutrino flavours. This is compatible with (2.7), but unlike the latter is not a direct measurement of mass. We will return to these cosmological arguments briefly in Chapter 11.

2.3.4 Lepton numbers revisited

In all the above discussions, we have assumed that lepton number conservation holds and can be used to identify the neutrino flavour emitted or absorbed in any weak reaction. However, in principle, lepton number violation can be induced in such reactions by the existence of neutrino oscillations. In practice, such effects are totally negligible in the standard model due to the short range of the weak interaction and can indeed be safely ignored.

To illustrate this, we will make a very rough estimate of the branching ratio for the lepton number violating decay[17]

$$\tau^- \to \mu^- + \gamma, \tag{2.43}$$

which can proceed via the mechanism of Figure 2.12, in which the tau-neutrino emitted at the first vertex oscillates into a muon neutrino before being reabsorbed at the second vertex. This diagram involves a weak interaction vertex, the emission of a photon and an oscillation, while the leptonic decay modes (2.11) are purely weak interactions. Therefore, the branching fraction of the decay (2.43) is expected to be of the order of magnitude

$$B(\tau^- \to \mu^- + \gamma) = O[\alpha P(\nu_\tau \to \nu_\mu)\mathrm{B}(\tau^- \to e^- \nu_\tau \bar{\nu}_e)], \tag{2.44}$$

where $P(\nu_\tau \to \nu_\mu)$ is the probability of oscillation and α is the fine structure constant. Thus, using $B(\tau^- \to e^- \nu_\tau \bar{\nu}_e) \sim O(10^{-1})$, we have

$$B(\tau^- \to \mu^- + \gamma) = O[10^{-3} P(\nu_\tau \to \nu_\mu)]. \tag{2.45}$$

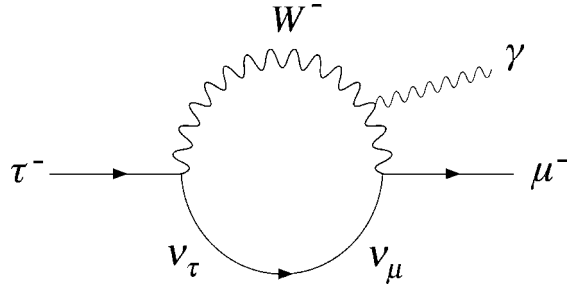


Figure 2.12 A Feynman diagram contributing to the decay $\tau^- \to \mu^- + \gamma$. There are two other diagrams, where the photon is emitted by either the μ^- or the τ^-.

[17] The details of the argument that follows are not required for later work.

The term $P(\nu_\tau \to \nu_\mu)$ can be estimated from the two-component mixing model of $\nu_\mu - \nu_\tau$ oscillations used to describe the atmospheric neutrino data in Section 2.3.2. From Equations (2.30), we then have

$$P(\nu_\tau \to \nu_\mu) \approx \sin^2(2\theta_{23}) \sin^2(L/L_0) \approx (L/L_0)^2, \tag{2.46a}$$

where we have used $\sin^2(2\theta_{23}) \gtrsim 0.9$ and assumed $L \ll L_0$. Here L is the typical distance travelled by the neutrino and is of the same order as the range M_W^{-1} of the weak interaction. The oscillation length is given by

$$L_0 = 4E/\Delta m_{32}^2, \tag{2.46b}$$

where E is the typical energy of the neutrinos in Figure 2.12. Substituting (2.46) into (2.45) gives

$$B(\tau \to \mu + \gamma) = O\left[\frac{10^{-3}(m_3^2 - m_2^2)^2}{16E^2 M_W^2}\right]. \tag{2.47}$$

For E we make the crude dimensional estimate $m_\mu \leq E \leq m_\tau$ for a tauon decaying at rest. Therefore, using the experimental value (2.32) for Δm_{32}^2, we have the rough estimate

$$B(\tau \to \mu + \gamma) \approx 10^{-48} - 10^{-50}. \tag{2.48}$$

This is utterly negligible for all practical purposes and should be compared to the experimental upper limit given in Table 2.1.

The above conclusion is confirmed by a more detailed treatment of both this and other reactions. Lepton number violation in weak interactions is completely negligible within the standard model, but could occur if one abandons the standard model. Some particularly interesting possibilities and the experiments required to investigate them will be discussed in Chapter 11.

PROBLEMS 2

2.1 Which of the following reactions are allowed, and which are forbidden, by the conservation laws appropriate to weak interactions?

(a) $\nu_\mu + p \to \mu^+ + n$
(b) $\nu_e + p \to e^- + \pi^+ + p$
(c) $\nu_\tau + e^- \to \tau^- + \nu_e$
(d) $\tau^+ \to \mu^+ + \bar{\nu}_\mu + \nu_\tau$

2.2 Draw a lowest-order Feynman diagram for $\nu_e \nu_\mu$ elastic scattering.

2.3 Draw the two fourth-order Feynman diagrams for the weak reaction $e^- + \mu^+ \to \nu_e + \bar{\nu}_\mu$.

2.4 In Section 2.3.1 it is stated that electron neutrinos interact with electrons in a different way from muon and tauon neutrinos. Justify this remark by considering the lowest order Feynman diagrams for $\nu_e + e^- \to \nu_e + e^-$ and $\nu_\mu + e^- \to \nu_\mu + e^-$.

2.5 Show that the oscillation length in Equation (2.31) may be written $L_0 = E/(1.27\Delta m_{ij}^2)$, where L_0 is expressed in km, E in GeV and Δm_{ij}^2 in $(\mathrm{eV/c^2})^2$.

2.6 A KamLAND-type experiment detects $\bar{\nu}_e$ neutrinos at a distance of 300 m from a nuclear reactor and finds that the flux is 90 ± 10 % of that expected if there were no oscillations. Assuming a two-component model with maximal mixing ($\theta = 45°$) and a mean neutrino energy of 2.5 MeV, estimate the squared mass difference of the electron antineutrino and its oscillating partner.

2.7 If the Sun is assumed to be a uniform spherical plasma consisting of nucleons, with radius 7×10^5 km and total mass 2×10^{30} kg, calculate the mean free path $\lambda = 1/n\sigma$ of solar neutrinos from the dominant reaction (2.35). Here n is the number of nucleons per unit volume and σ, the neutrino–nucleon cross-section, may be written $\sigma = 0.7E_L \times 10^{-42}\,\mathrm{m}^2$, where E_L is the neutrino laboratory energy in GeV.

3

Quarks and Hadrons

We turn now to the strongly interacting particles – the quarks and their bound states, the hadrons. These also interact by the weak and electromagnetic interactions, although such effects can often be neglected compared to the strong interactions. To this extent we are entering the realm of 'strong interaction physics'.

Strong interactions are most familiar in nuclear physics, where the interactions of neutrons and protons are studied at relatively low energies of a few tens of MeV. However, in 1947 new types of hadrons, not present in ordinary matter, were discovered in cosmic rays by groups from the universities of Bristol and Manchester. To create these new particles required high energies, in accordance with Einstein's mass–energy relation $E = mc^2$, and as intense beams of particles of increasingly higher energies became available at accelerator laboratories, more and more hadrons were discovered. By the late 1960s several dozen were known, and some unifying theoretical framework was urgently needed to interpret this multitude of states if any progress was to be made. The result was the *quark model*. In 1964, Gell-Mann, and independently Zweig, noted that all the observed hadrons could be simply interpreted as bound states of just three fundamental spin-$\frac{1}{2}$ particles, together with their antiparticles. These particles were required to have fractional electric charges of $\frac{2}{3}$, $-\frac{1}{3}$ and $-\frac{1}{3}$, in units of e, and were called *quarks* by Gell-Mann, who later cited the now famous quotation 'Three quarks for Muster Mark' in James Joyce's book *Finnegans Wake*.[1]

In the following years, the success of the quark model grew ever more impressive as more and more states were discovered. Nonetheless, the existence of quarks as real particles, rather than convenient mathematical entities, was seriously doubted because all attempts to detect free quarks, or any other fractionally charged particles, met with failure. These doubts were subsequently removed in two ways. Firstly, a series of experimental results, starting in 1968 with the scattering of high-energy

[1] In fact more than three quarks are known, but this was not discovered until ten years later (see below).

Particle Physics Third Edition B. R. Martin and G. Shaw

electrons from protons, showed the dynamical effects of individual quarks within the proton. These experiments will be described in Chapter 7. Secondly, a detailed theory of strong interactions was constructed, which both successfully described the experimental results and offered an explanation of why isolated free quarks could not be observed. This theory is called *quantum chromodynamics* (*QCD*), and will also be discussed in Chapter 7. As a result of these developments, the quark hypothesis is now universally accepted, and is central to the interpretation of a wide range of phenomena in particle physics.

In what follows, we take the existence and properties of quarks as our basic assumptions, interpreting hadrons in terms of them. We begin by briefly introducing the quarks, together with a few of their best-known hadronic bound states, leaving more detailed considerations to later chapters.

3.1 QUARKS

Although only three quarks were initially proposed, six are now known to exist. Like the leptons, these six distinct types, or *flavours,* occur in pairs, or generations, denoted

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix}. \tag{3.1a}$$

Each generation consists of a quark with charge $+\frac{2}{3}$ (u, c or t) together with a quark of charge $-\frac{1}{3}$ (d, s, b), in units of e. They are called the *down* (d), *up* (u), *strange* (s), *charmed* (c), *bottom* (b) and *top* (t) quarks. The corresponding antiquarks are denoted

$$\begin{pmatrix} \bar{d} \\ \bar{u} \end{pmatrix}, \quad \begin{pmatrix} \bar{s} \\ \bar{c} \end{pmatrix}, \quad \begin{pmatrix} \bar{b} \\ \bar{t} \end{pmatrix}, \tag{3.1b}$$

with charges $+\frac{1}{3}(\bar{d}, \bar{s}, \bar{b})$ and $-\frac{2}{3}(\bar{u}, \bar{c}, \bar{t})$. There is good experimental evidence for the existence of all six flavours (*d, u, s, c, b* and *t*). Approximate quark masses are given in Table 3.1. The masses of the (*d, u, s, c* and *b*) quarks are inferred indirectly from the observed masses of their hadron bound states, together with models of quark binding. The top quark is too short-lived to form observable hadron states and its mass is inferred from its decay products, as we shall see.[2] None of the masses can be obtained from measurements on free quarks because free quarks have never been seen, despite many experiments to find them.

Free quarks would be most readily identified via their fractional electric charges. One consequence of this is that the lightest quark would be stable, as electric charge has to be conserved in any decay. In matter, such stable quarks would give rise to 'exotic atoms' with fractional charges that could be identified by techniques like mass spectroscopy. Many searches for pre-existing quarks in matter have been made and many strange materials have been investigated (including moon rock, crushed oyster shells and deep sea sludge!), all with null results. The best upper limit on a possible quark density is 10^{-24} quarks/nucleon in sea water.

[2] This will be discussed in Section 8.2.3.

TABLE 3.1 The approximate masses of the quarks in GeV/c^2 and their electric charges Q in units of e. Also shown are the values of the baryon number B, strangeness S, charm C, bottom $\tilde{B}$ and top T, as defined in Section 3.2. The values for the corresponding antiquarks are equal in magnitude, but opposite in sign.

Name	Symbol	Mass	Q	B	S	C	$\tilde{B}$	T
Down	d	$m_d \approx 0.3$	$-1/3$	1/3	0	0	0	0
Up	u	$m_u \approx m_d$	2/3	1/3	0	0	0	0
Strange	s	$m_s \approx 0.5$	$-1/3$	1/3	-1	0	0	0
Charmed	c	$m_c \approx 1.5$	2/3	1/3	0	1	0	0
Bottom	b	$m_b \approx 4.5$	$-1/3$	1/3	0	0	-1	0
Top	t	$m_t \approx 174$	2/3	1/3	0	0	0	1

Another technique for detecting possible pre-existing free quarks is to do experiments similar to the classic Millikan oil drop experiment that first measured precisely the charge on the electron. The first experiment of this type was by Gallinaro and Morpurgo in 1966, and later work by Larue, Fairbank and Hubbard in 1977, using niobium balls at low temperatures to replace Millikan's oil drops, claimed evidence for several examples of fractional charges consistent with $\pm\frac{1}{3}e$. However, further experiments using a variety of materials, including niobium, failed to confirm the original results, and one can only conclude that the possible concentration of free quarks in such metals is less than 10^{-20} quarks per nucleon.

There have also been attempts to find quarks that might have been created in high-energy collisions in cosmic rays, or at accelerator laboratories. The aim here is to use the fact that the ionization produced by quarks would be either $\frac{1}{9}$ or $\frac{4}{9}$ of that produced by a particle of charge e. In cosmic ray searches it is concluded that the concentration of quarks must be less than one per 10^{10} primary cosmic rays. More than this would have produced a detectable signal in the secondary cosmic rays observed at sea level. At accelerators, a wide range of reactions has been studied, all with null results.

In summary, there is no convincing evidence for the existence of isolated free quarks, or any other fractionally charged particles, despite great efforts to find them.

3.2 GENERAL PROPERTIES OF HADRONS

Although no isolated quarks have been found, more than two hundred of their bound states have been discovered, all with integer electric charges. The reason for this is closely associated with a new degree of freedom that exists for quarks, but not for leptons, called *colour*. We shall postpone discussion of this for the moment, and follow the simple quark model in assuming that only three types of quark bound states are allowed. These are the *baryons,* which have half-integer spin and are assumed to be bound states of three quarks ($3q$); the *antibaryons,* which are their antiparticles and assumed to be bound states of three antiquarks ($3\bar{q}$); and the *mesons,* which have integer spin and are assumed to be bound states of a quark and an antiquark ($q\bar{q}$). Some examples of baryons and mesons that we shall meet later in this chapter, together with their quark compositions, are shown in Tables 3.2 and 3.3.

TABLE 3.2 Some examples of baryons, with their quark compositions and the corresponding values of their electric charge Q, strangeness S, charm C and bottom $\tilde{B}$.

Particle		Mass (MeV/c^2)	Q	S	C	$\tilde{B}$
p	uud	938	1	0	0	0
n	udd	940	0	0	0	0
Λ	uds	1116	0	-1	0	0
Λ_c	udc	2285	1	0	1	0
Λ_b	udb	5624	0	0	0	-1

Tables 3.2 and 3.3 also show the values of several quantum numbers which are associated with any state and which refer to its quark content. For example, the *strangeness* S is defined by

$$S \equiv -N_s \equiv -[N(s) - N(\bar{s})], \tag{3.2a}$$

where $N(s)$ and $N(\bar{s})$ are the number of s quarks and $\bar{s}$ antiquarks present in the state. Clearly $S = -1$ for an s quark, $S = 1$ for an $\bar{s}$ antiquark and $S = 0$ for all other quarks and antiquarks. The *charm* (C), *bottom* ($\tilde{B}$) and *top* (T) quantum numbers are similarly defined by[3]

$$C \equiv N_c \equiv N(c) - N(\bar{c}), \tag{3.2b}$$

$$\tilde{B} \equiv -N_b \equiv -[N(b) - N(\bar{b})] \tag{3.2c}$$

and

$$T \equiv N_t \equiv N(t) - N(\bar{t}), \tag{3.2d}$$

respectively, giving values for the individual quarks which are summarized for convenience in Table 3.1.

TABLE 3.3 Some examples of mesons, with their quark compositions and the corresponding values of their electric charge Q, strangeness S, charm C and bottom $\tilde{B}$.

Particle		Mass (MeV/c^2)	Q	S	C	$\tilde{B}$
π^+	$u\bar{d}$	140	1	0	0	0
K^-	$s\bar{u}$	494	-1	-1	0	0
D^-	$d\bar{c}$	1869	-1	0	-1	0
D_s^+	$c\bar{s}$	1969	1	1	1	0
B^-	$b\bar{u}$	5279	-1	0	0	-1
Υ	$b\bar{b}$	9460	0	0	0	0

[3] *Bottom* was originally called *beauty* and *top* was called *truth*, but these names are seldom used now.

The resulting values of S, C and $\tilde{B}$ for our specimen hadrons are given in Tables 3.2 and 3.3 together with the hadron's electric charge Q, which is just the sum of the appropriate quark charges. The top quantum number $T = 0$ for all known hadrons. The remaining quark numbers,

$$N_u = N(u) - N(\bar{u}) \tag{3.3a}$$

and

$$N_d = N(d) - N(\bar{d}), \tag{3.3b}$$

do not have special names since, given $S, C, \tilde{B}, T$ and Q, their values can be inferred from the *baryon number* B of the state, defined by

$$B \equiv \tfrac{1}{3}[N(q) - N(\bar{q})], \tag{3.4a}$$

where $N(q)$ and $N(\bar{q})$ are the total numbers of quarks and antiquarks present, irrespective of their flavours. The baryon number $B = 1$ for baryons, $B = -1$ for antibaryons and $B = 0$ for mesons. In terms of the quark numbers (3.2) and (3.3)

$$\begin{aligned} B &= \tfrac{1}{3}(N_u + N_d + N_s + N_c + N_b + N_t) \\ &= \tfrac{1}{3}(N_u + N_d - S + C - \tilde{B} + T)\,, \end{aligned} \tag{3.4b}$$

while the electric charge is given by

$$\begin{aligned} Q &= \tfrac{2}{3}(N_u + N_c + N_t) - \tfrac{1}{3}(N_d + N_s + N_b) \\ &= \tfrac{2}{3}(N_u + C + T) - \tfrac{1}{3}(N_d - S - \tilde{B}). \end{aligned} \tag{3.5}$$

All these quantum numbers are called *internal quantum numbers*, because they are not associated with motion or the spatial properties of wave functions. They are important, because in strong and electromagnetic interactions quarks and antiquarks are only created or destroyed in particle–antiparticle pairs. For example, the quark description of the strong interaction process

$$p + p \to p + n + \pi^+ \tag{3.6a}$$

is

$$(uud) + (uud) \to (uud) + (udd) + (u\bar{d}). \tag{3.6b}$$

On counting the quarks of each flavour, we see that the final state contains the same number of quarks of each flavour as the initial state, plus an additional $(d\bar{d})$, so

that the quark numbers N_u and N_d are separately conserved. This is characteristic of strong and electromagnetic processes, in which all the quark numbers (3.2) to (3.5) are separately conserved. However, in neutron β-decay,

$$n \to p + e^- + \bar{\nu}_e, \tag{3.7a}$$

whose quark interpretation is

$$(udd) \to (uud) + e^- + \bar{\nu}_e, \tag{3.7b}$$

a d quark is replaced by a u quark. This decay is conveniently represented by the *quark diagram* of Figure 3.1 This corresponds to the Feynman diagram of Figure 1.13, except that each hadron is represented by lines corresponding to its constituent quarks and antiquarks, with quarks and antiquarks labelled by arrows pointing to the right and left, respectively. The fact that N_u and N_d are not conserved is characteristic of the weak interactions, in which the quark flavours can change and only the baryon number (3.4a) and the total electric charge are in general conserved.

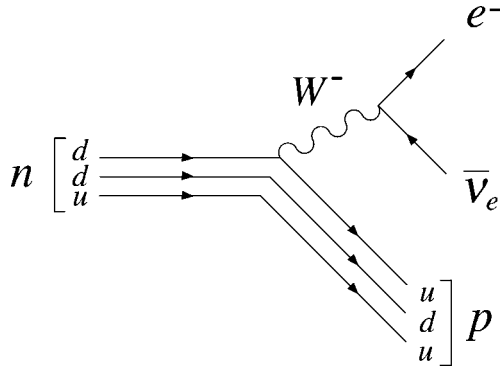


Figure 3.1 Quark diagram for the decay $n \to p + e^- + \bar{\nu}_e$.

The quark numbers (3.2) to (3.5) play an important role in understanding the long lifetimes of some hadrons. Hadrons have typical effective radii r of order 1 fm, with an asociated time scale r/c of order 10^{-23} s. The vast majority are highly unstable and decay to lighter hadrons by the strong interaction with lifetimes of this order. However, each hadron is characterized by a set of values for B, Q, S, C, $\tilde{B}$ and T, and in some cases there are no lighter hadron states with the same values of these quantum numbers to which they can decay. These hadrons, which cannot decay by strong interactions, are long-lived on a timescale of order 10^{-23} s and are often called *stable particles.* Here we shall call them *long-lived particles*, because except for the proton they are not absolutely stable, but decay by either the electromagnetic or weak interaction.[4] Electromagnetic decay rates are suppressed by powers of the fine structure constant

[4] Some modern theories predict that even the proton is unstable, albeit with a lifetime of order $10^{29} - 10^{33}$ y. This is discussed briefly in Chapter 11.

α relative to strong decays, leading to observed lifetimes in the range$10^{-16} - 10^{-21}$ s. Weak decays give longer lifetimes that depend sensitively on the characteristic energy of the decay, as we have seen in the case of the μ and τ leptons (cf. Section 2.2.2). A useful measure of this characteristic energy is the *Q-value*, which is the kinetic energy released in the decay of the particle at rest. For neutron decay (3.7a),

$$Q = m_n - m_p - m_e - m_{\bar{\nu}_e} = 0.79\,\text{MeV} \tag{3.8}$$

and the lifetime is about 10^3 s. However, Q-values of order $10^2 - 10^3$ MeV are more typical, and the observed lifetimes for other weak hadron decays lie in the range $10^{-7} - 10^{-13}$ s. Thus hadron lifetimes span some 27 orders of magnitude, from about 10^{-24} s to about 10^3 s. The typical ranges corresponding to each interaction are summarized in Table 3.4.

TABLE 3.4 Typical lifetimes of hadrons decaying by the three interactions.

Interaction	Lifetime (s)
Strong	$10^{-22} - 10^{-24}$
Electromagnetic	$10^{-16} - 10^{-21}$
Weak *	$10^{-7} - 10^{-13}$

* The neutron lifetime is an exception, for reasons explained in Section 3.2.

3.3 PIONS AND NUCLEONS

The lightest known mesons are the pions or pi-mesons $\pi^{\pm}(140), \pi^0(135)$, where we indicate the masses in MeV/c^2 in parentheses. These particles are produced copiously in many hadronic reactions that conserve both charge and baryon number, e.g. in proton–proton collisions

$$p + p \to p + n + \pi^+, \tag{3.9a}$$

$$\to p + p + \pi^0 \tag{3.9b}$$

and

$$\to p + p + \pi^- + \pi^+. \tag{3.9c}$$

The charged pions decay predominantly by the reactions

$$\pi^+ \to \mu^+ + \nu_\mu, \quad \pi^- \to \mu^- + \bar{\nu}_\mu, \tag{3.10}$$

with lifetimes of 2.6×10^{-8} s, typical of weak interactions (cf. Table 3.4). They were first discovered in cosmic rays by a Bristol group (Lattes, Muirhead, Occhialini and Powell) in 1947 using photographic emulsions containing a silver halide. The ionization energy deposited by a particle passing through the emulsion causes the formation

of a latent image, and the silver grains resulting from subsequent development form a visual record of the path of the particle. As in a cloud chamber, the range of a stopping particle can be used to estimate its energy. Some historic examples of stopping π^+ decays observed in this way are shown in Figure 3.2. Because these are two-body decays, the muons in the final state all have the same energy. Consequently, they all travel approximately the same distance before coming to rest and decaying to give a positron by reaction (3.10). Neutral pions were discovered somewhat later and decay by the electromagnetic interaction

$$\pi^0 \to \gamma + \gamma, \tag{3.11}$$

with a lifetime of 0.8×10^{-16} s. Because they are neutral they do not leave tracks, and must be detected via their decay photons.

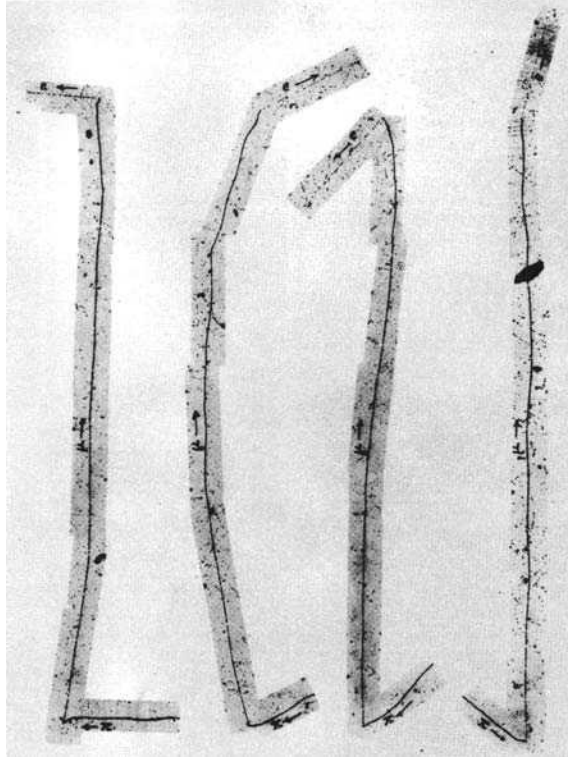


Figure 3.2 Some early examples of pion decays observed in a photographic emulsion by the Bristol group (Lattes *et al*., 1947). The positive pions (short lines at the bottom of the picture) come to rest before decaying by the reaction $\pi^+ \to \mu^+ + \nu_\mu$ to give monoenergtic muons. These travel about 600 mm in the emulsion before coming to rest and decaying by the reaction $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$ to give a positron.

Pions and nucleons are the lightest known mesons and baryons, respectively, and are bound states of the lightest quarks (u, d) and their antiquarks $(\bar{u}, \bar{d})$. The only $3q$ and $q\bar{q}$ combinations that give the observed electric charges are then

$$p = uud, \quad n = udd \tag{3.12}$$

for the nucleons, and

$$\pi^+ = u\bar{d}, \quad \pi^0 = u\bar{u}, d\bar{d}, \quad \pi^- = d\bar{u} \tag{3.13}$$

for the pions, where the ambiguity in the case of the π^0 decays will be resolved in Chapter 6. The detailed evidence that this quark picture is correct will also be described in later chapters. For the moment, we shall assume it is, and interpret various reactions in terms of it. The quark interpretation of (3.9a) has already been discussed, following (3.6), and we leave it to the reader to show that the quark numbers N_u and N_d are also conserved in (3.9b) and (3.9c). However, in π^+ decay (3.10), whose quark interpretation is

$$(u\bar{d}) \to \mu^+ + \nu_\mu,$$

and is shown in Figure 3.3, a u quark annihilates with a $\bar{d}$ antiquark, demonstrating again that individual quark numbers like N_u and N_d are not conserved in weak interactions.

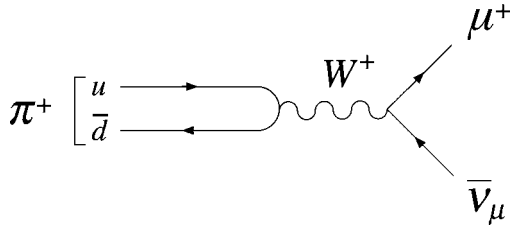


Figure 3.3 Quark diagram for the decay $\pi^+ \to \mu^+ + \nu_\mu$.

Pions play an important role in nuclear forces. In 1935 Yukawa proposed that these were due to the exchange of spin-0 mesons, and from the range of the forces (which was not precisely known at that time) predicted that these mesons should have a mass of approximately 200 MeV/c^2 (cf. the discussion of Section 1.4.2). The discovery of pions was a great triumph for the Yukawa theory. In it, the nuclear forces are given by Figure 3.4, where the nucleons and pions are treated as point particles. Neutral pion exchange gives rise to normal direct forces, while $\pi^\pm$ exchange gives rise to exchange forces whereby the neutron and proton are interchanged. The resulting potential is of the general Yukawa type (1.30) and reproduces the longest range part of the nuclear force $r \geq 2$ fm very accurately, including spin effects when these are taken into account. At short ranges, $r \leq 1$ fm, comparable with the sizes of the nucleons and pions themselves, this agreement breaks down completely. At these distances, the internal quark structure of the hadrons can no longer be neglected and the situation is much more complicated.

Finally, we note that the atmospheric neutrinos discussed in Section 2.3.2 are produced almost entirely from pion decays. When cosmic ray protons collide with atoms in the upper atmosphere, they create many charged pions in reactions like (3.9).

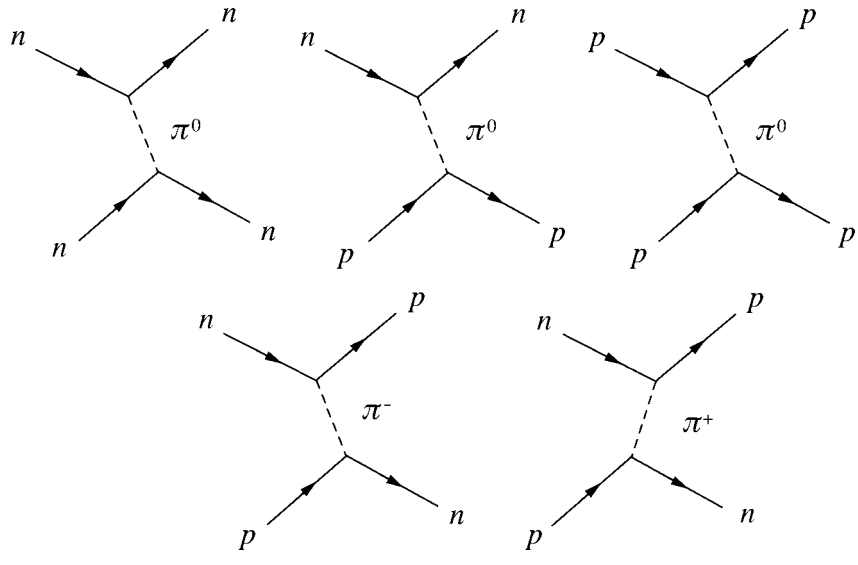


Figure 3.4 The Yukawa model for nuclear forces: direct forces (upper three diagrams) and exchange forces (lower two diagrams).

These decay almost entirely by (3.10), followed by muon decay (2.10) to give two muon neutrinos or antineutrinos for each electron neutrino or antineutrino, as noted in Section 2.3.2.

3.4 STRANGE PARTICLES, CHARM AND BOTTOM

Soon after the discovery of the pion, members of the Manchester cosmic ray group discovered other hadrons that were produced in strong interactions, but decayed by weak interactions. This was unexpected, as there was apparently no reason why they should not decay by the strong interaction with lifetimes of order 10^{-23} s. For this reason they were christened 'strange particles'. Figure 3.5 shows one of the first observed events, in which a positively charged particle of mass around 500 MeV/c^2 decays to a muon and an unseen neutrino. The decaying particle is actually a K^+ meson (or kaon), with mass 494 MeV/c^2 and a lifetime of 1.2×10^{-8} s. Charged kaons have many decay modes, but the principle ones and their branching ratios, B, are

$$K^+ \to \mu^+ + \nu_\mu, \quad B = 0.63, \tag{3.14a}$$

as observed in Figure 3.5, and

$$K^+ \to \pi^+ + \pi^0, \quad B = 0.21. \tag{3.14b}$$

Another example of a strange particle is the Λ (lambda) baryon, which has a mass of 1116 MeV/c^2 and decays mainly into pions and nucleons

$$\Lambda \to \pi^- + p, \quad B = 0.64 \tag{3.15a}$$

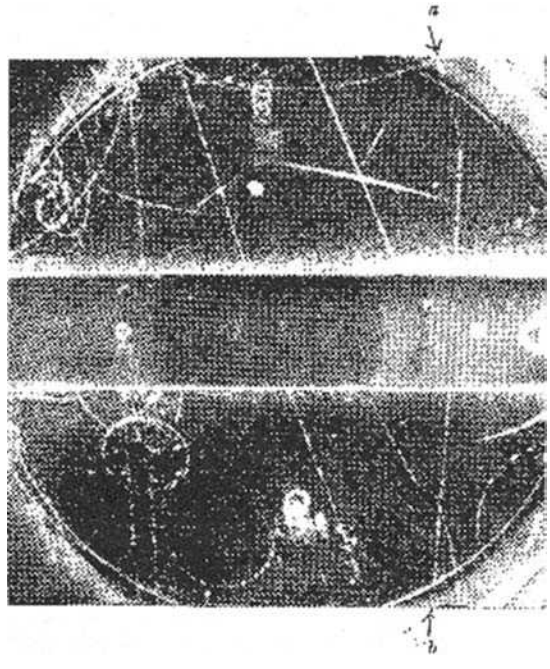


Figure 3.5 One of the first strange particle events observed in a cloud chamber by Rochester and Butler. A K^+ meson (track *a*) enters the top of the chamber before decaying by the reaction $K^+ \to \mu^+ + \nu_\mu$ to give a muon (track *b*) and an unseen neutrino. The dark band across the chamber is a 3 cm thick lead plate, and track *b* is identified by its ability to penetrate this plate with little loss of energy. (Reprinted from Rochester and Butler, *Nature* **160**, 855, 1947, by permission of Macmillan Publishers Ltd.)

and

$$\Lambda \to \pi^0 + n, \quad B = 0.36, \tag{3.15b}$$

with a lifetime of 2.6×10^{-10} s.

It is clear from the long lifetimes of the K^+ and Λ that they both decay via the weak interaction. This strongly suggests that these particles are not made of *u* and *d* quarks alone, because if this were the case then, for example, the neutral Λ would be a (*udd*) state just like the neutron. At the quark level, the decay (3.15a) would then be

$$(udd) \to (d\bar{u}) + (uud),$$

which conserves the quark numbers N_u and N_d. We would therefore expect (3.15a) to be a strong decay, with a lifetime of order 10^{-23} s, in contradiction to experiment. The solution is to assign the quark structure *uds* to the Λ, so that the decay is

$$\begin{array}{cccc} (uds) & \to & (d\bar{u}) & + (uud) \\ S = -1 & & 0 & 0 \end{array}$$

and neither the quark number N_d nor the strangeness S is conserved. As both the strong and electromagnetic interactions conserve all quark numbers, the decay can only proceed by the weak interaction, in which such quark numbers are not conserved.

Strange particles are now defined as any particle with a nonzero value of the strangeness quantum number. Most of them, like most hadrons with $S = 0$, decay by the strong interactions. However, conservation of quark numbers in strong and electromagnetic interactions means that if a particle is the lightest state with a given non-zero set of B, Q and S values,[5] it can only decay by weak interactions, and so will be relatively long lived. From its quark structure, Λ has $B = 1$, $Q = 0$ and $S = -1$. It is the lightest strange baryon. The lightest strange mesons are the kaons. In addition to the K^+, which we have already met, there are negative and neutral kaons, which also decay via the weak interaction. Their quark structures are

$$K^+(494) = u\bar{s}, \quad K^0(498) = d\bar{s} \tag{3.16a}$$

and

$$K^-(494) = s\bar{u}, \quad \bar{K}^0(498) = s\bar{d}, \tag{3.16b}$$

where K^+ and K^0 have $S = +1$ and K^- and $\bar{K}^0$ have $S = -1$.

The production of strange particles in strong interactions is an example of *associated production.* In such processes, more than one strange particle is produced, giving strangeness conservation overall. A beautiful example of such an event is given in Figure 3.6, which shows the reaction

$$\begin{array}{cccccc} \pi^- + p & \rightarrow & K^0 + & \Lambda \\ S = 0 & 0 & 1 & -1, \end{array} \tag{3.17}$$

together with the subsequent decays of the kaon and lambda. This picture was taken in a bubble chamber, which is a detection device similar in principle to a cloud chamber, but where the working medium is a liquid rather than a gas. Bubble chambers were invented by Glaser in 1952 and largely superceded cloud chambers in the late 1950s. The liquid is initially held in a vessel at a pressure above that of the equilibrium vapour pressure at that temperature. When the pressure is suddenly reduced by an expansion, the liquid is left in a 'superheated' state, where its temperature is above the boiling point at the reduced pressure. The passage of a charged particle through the liquid produces ion pairs, and bubbles form preferentially along the ionization trails. The liquid in the chamber (hydrogen in the case of Figure 3.6) is both target and detector, and by placing the whole chamber in a magnetic field the particle momenta can be obtained. The curvatures of the tracks are clearly visible in Figure 3.6.

In the thirty years following the discovery of the pions and kaons, a great many hadrons were discovered. Until 1974 all could be accounted for as bound states of just the three quarks, u, d and s, originally proposed by Gell-Mann and Zweig.

[5] The lightest hadron with the zero set $B = Q = S = 0$ is the π^0 meson, which decays to photons by the electromagnetic interaction, as we have seen.

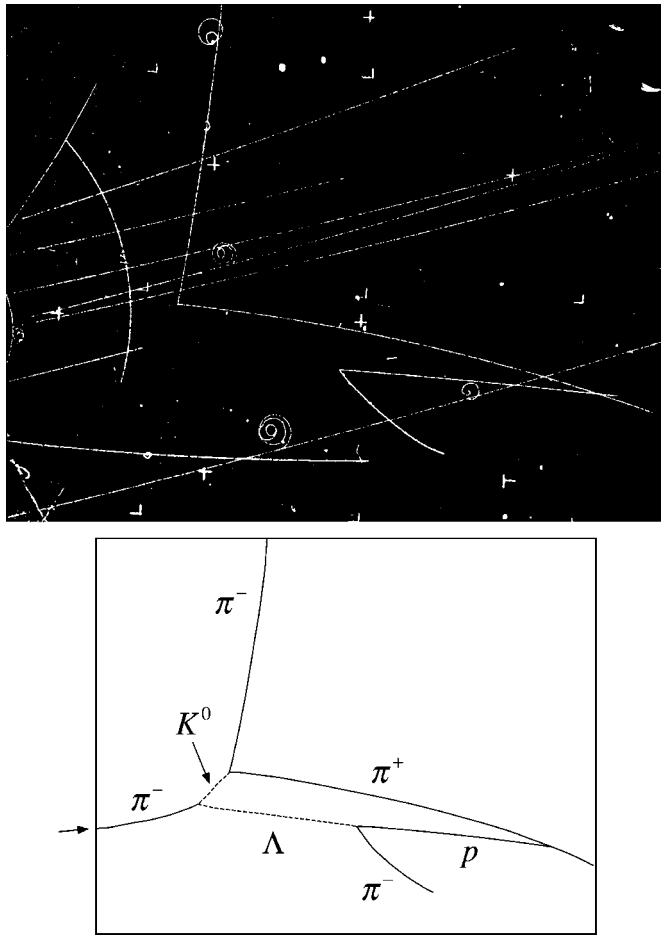


Figure 3.6 A bubble chamber picture of the associated production reaction $\pi^- + p \to K^0 + \Lambda$. The incoming pion is indicated by the arrow, and the unseen neutrals are detected by their decays $K^0 \to \pi^+ + \pi^-$ and $\Lambda \to p + \pi^-$. This picture was taken in the 10 in (25 cm) bubble chamber at the Lawrence Berkeley National Laboratory. (Lawrence Berkeley National Laboratory photo, with permission.)

However, in that year a relatively heavy particle was discovered in two independent experiments, one at the Brookhaven National Laboratory (BNL) and the other at the Stanford Linear Accelerator Laboratory (SLAC). The BNL group, led by Ting, named this new particle J, while the SLAC group, led by Richter, chose Ψ. It is now known as J/Ψ, and its properties show that it is one of the lightest of a family of particles that are bound states of a charmed quark and its antiparticle, i.e.

$$J/\Psi(3097) = c\bar{c} \qquad (C = 0).$$

Such $c\bar{c}$ states are collectively known as *charmonium* by analogy with positronium, the bound state of an electron and positron that we will discuss in Chapter 6. Since $C = 0$, these states are often said to contain 'hidden charm'.

Particles with 'naked charm' $C \neq 0$ were also discovered at SLAC shortly after the discovery of the J/Ψ. Because charm is a quark number, like strangeness, it should be conserved in strong and electromagnetic interactions, and the lightest charmed particles should decay by weak interactions. This is indeed the case. For example, the lightest charmed mesons are the D mesons with quark structures

$$D^+(1869) = c\bar{d}, \quad D^0(1865) = c\bar{u} \quad (C = +1) \tag{3.18a}$$

and

$$D^-(1869) = d\bar{c}, \quad \bar{D}^0(1865) = u\bar{c} \quad (C = -1) \tag{3.18b}$$

and D_s mesons with quark structures

$$D_s^+(1969) = c\bar{s} \quad (C = +1, S = +1)$$

and

$$D_s^-(1969) = s\bar{c} \quad (C = -1, S = -1),$$

while the lightest charmed baryon is

$$\Lambda_c^+(2285) = udc \quad (C = +1).$$

These particles all have lifetimes of order 10^{-13} s, which is in the expected range for weak decays. Charmed particles can be produced in strong and electromagnetic interactions by associated production reactions, just like strange particles, and a bubble chamber picture of such an event is shown in Figure 3.7. However, because the charmed particles have much shorter lifetimes than the strange particles (3.15) and (3.16), they travel much shorter distances before decaying, and very good spatial resolution is needed to observe their tracks. This was achieved in the present case by only allowing the bubbles to develop to a size of about 0.05 mm before photographing them, compared to sizes of about 0.5 mm used in a typical bubble chamber.

Historically, the discovery of strange particles caused great excitement because they clearly represented a new form of matter that was completely unexpected at the time. The discovery of charmed particles caused equally great excitement because their existence *was* expected, having been predicted from the then newly formulated theory of electroweak interactions. Their discovery was a decisive event in confirming the essential correctness of this theory, which is a unified theory of both weak and electromagnetic interactions. In its present form it requires that the number of leptons and quarks should be the same, implying that there should be six quarks to match the six known leptons. Evidence for the fifth quark – the bottom quark b with its associated quantum number $\tilde{B}$ – came from the discovery in 1977 of one of the lightest *bottomium* states

$$\Upsilon(9460) = b\bar{b} \quad (\tilde{B} = 0),$$

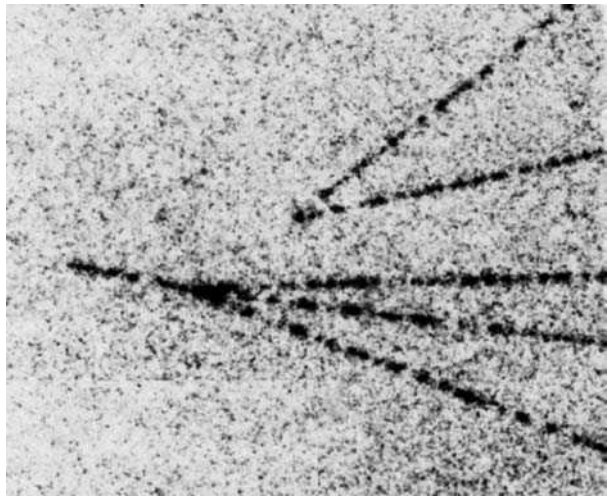


Figure 3.7 An example of the associated production of two charmed particles observed in a high-resolution bubble chamber at SLAC (Abe *et al.*, 1982). An incident unseen photon enters from the left and interacts with a proton in the liquid hydrogen of the chamber to produce a charged, charmed particle, which travels about 2 mm before decaying to give three charged particles, and an unseen, neutral charmed particle, which travels about 3 mm before decaying to give two charged particles. (Reprinted Figure 1 with permission from K. Abe *et al.*, *Phys. Rev. Lett.*, 48, 1526. Copyright 1982 American Physical Society.)

which is a 'hidden bottom' state called the upsilon. Subsequently mesons with 'naked bottom' $\tilde{B} \neq 0$ were discovered, the lightest examples of which are the B mesons

$$B^+(5278) = u\bar{b}, \quad B^0(5279) = d\bar{b} \quad (\tilde{B} = +1) \tag{3.19a}$$

and

$$B^-(5278) = b\bar{u}, \quad \bar{B}^0(5279) = b\bar{d} \quad (\tilde{B} = -1) \tag{3.19b}$$

with lifetimes of order 10^{-12} s, consistent with their decay via the weak interaction. Examples of baryons with $\tilde{B} \neq 0$, such as

$$\Lambda_b^0(5624) = udb \tag{3.20}$$

and the doublet

$$\Xi_b^0(5792) = usb, \quad \Xi_b^-(5792) = dsb, \tag{3.21}$$

which also decay via the weak interaction, with lifetimes of order 10^{-12} s, have also been discovered.

Research in the field of spectroscopy is ongoing. For example, recent experiments have found evidence for the first examples of a 'charmed strange' and a 'bottom–strange' baryon, i.e. ones containing a strange quark in combination with a charmed or bottom quark, respectively.

3.5 SHORT-LIVED HADRONS

The early history of hadron physics was dominated by cosmic ray studies, in which new particles, such as pions and kaons, were produced in collisions between cosmic ray protons and atoms in the upper atmosphere. Such processes are random and hard to observe, making detailed studies of hadronic interactions extremely difficult. However, by the 1960s the situation had been transformed by the creation of high-energy beams of pions and kaons in the laboratory, leading to the discovery of large numbers of hadrons that decayed by the strong interaction, with lifetimes of order 10^{-23} s. This profusion of states was the 'last nail in the coffin' for the idea that hadrons were in any sense fundamental and led rapidly to the development of the quark model to explain their properties. In this section we will firstly explain how such states are detected and then discuss their interpretation within the quark model.

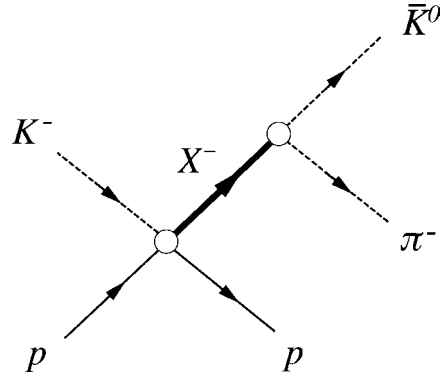


Figure 3.8 Production of a negatively charged resonance X^- in the reaction $K^- + p \to X^- + p$, and its subsequent decay, $X^- \to \bar{K}^0 + \pi^-$.

Hadrons that decay by strong interactions are called *resonances*. They are far too short-lived to be directly observed, even by electronic methods, and their existence must be inferred from observations on the more stable hadrons to which they decay. A typical example is the production of a boson resonance X^-in the reaction

$$K^- + p \to X^- + p \tag{3.22a}$$

and its subsequent decay by the reaction

$$X^- \to \bar{K}^0 + \pi^-, \tag{3.22b}$$

where $\bar{K}^0(498)$ is the neutral boson with $S = -1$ listed in Equation (3.16b). This is illustrated in Figure 3.8, where the distance between the points at which the resonance is produced and decays is too small to be measured (see Problem 3.5). The observed reaction is therefore

$$K^- + p \to \bar{K}^0 + \pi^- + p \tag{3.23}$$

and the fleeting existence of the X^- must be inferred from measurements of the energies E_K and E_π and the momenta $\mathbf{p}_K$ and $\mathbf{p}_\pi$ of the $\bar{K}^0$ and π^- mesons to which it decays. If the decaying particle has mass M, energy E and momentum $\mathbf{p}$, then by energy–momentum conservation the invariant mass[6] W of the $\bar{K}^0\pi^-$ pair is given by

$$W^2 \equiv (E_K + E_\pi)^2 - (\mathbf{p}_K + \mathbf{p}_\pi)^2 = E^2 - \mathbf{p}^2 = M^2. \tag{3.24}$$

Thus, if (3.22) proceeds by the mechanism of Figure 3.8, a plot of the invariant mass distribution will show a sharp peak at the resonance mass M. On the other hand, if uncorrelated $\bar{K}^0$ and π^- particles are produced by some other mechanism, a smooth distribution would be expected.

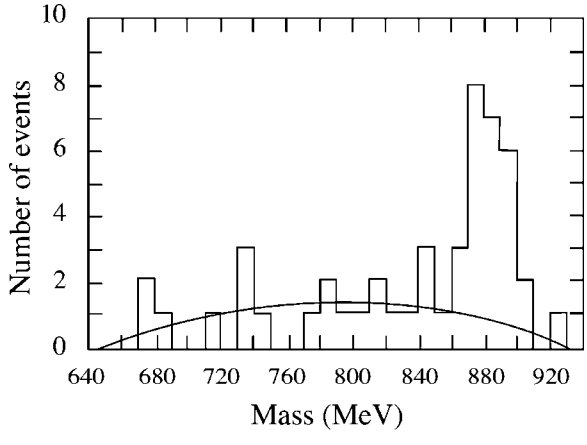


Figure 3.9 Invariant mass distribution of the $\bar{K}^0\pi^-$ pairs in the reaction $K^- + p \to X^- + p$ for an initial beam momentum of 1.15 GeV/c. (Reprinted Figure 2 with permission from M. H. Alston *et al.*, *Phys. Rev. Lett.* **6**, 520. Copyright 1961 American Physical Society.)

Figure 3.9 shows a histogram of the invariant mass distribution obtained in an early experiment to detect a meson resonance. In this experiment, a 1.15 GeV/c K^- beam was directed into a hydrogen bubble chamber and 47 events corresponding to the reaction (3.23) were identified. Despite the limited statistics, a clear peak is seen with a mass corresponding to a resonance called the $K^{*-}(890)$. The peak sits on a background arising from uncorrelated pairs produced by some other, non-resonant, mechanism, and has a width of about 50 MeV. Widths of this order are characteristic of particles that decay by the strong interaction. For a particle at rest, $W = E$ by (3.24) and the energy–time uncertainty principle leads to a natural line width

$$\Delta W = \Delta E \approx \Gamma \equiv 1/\tau, \tag{3.25}$$

where τ is the mean lifetime at rest. The constant Γ is called the *decay width* (or just the *width*) of the state. If τ is of order 10^{-23} s, as expected for a strong decay, then $\Gamma = O(100\,\text{MeV})$, which is similar to the width of the resonance peak in Figure 3.9.

[6] See Appendix A, Section A.2.

The shape of an isolated resonance peak is conveniently approximated by the *Breit–Wigner formula*

$$N(W) = \frac{K}{(W - W_r)^2 + \Gamma^2/4}, \tag{3.26}$$

derived in Section B.5.[7] Here K is a constant that depends on the total number of decays observed and W_r is the position of the maximum, as shown in Figure 3.10.

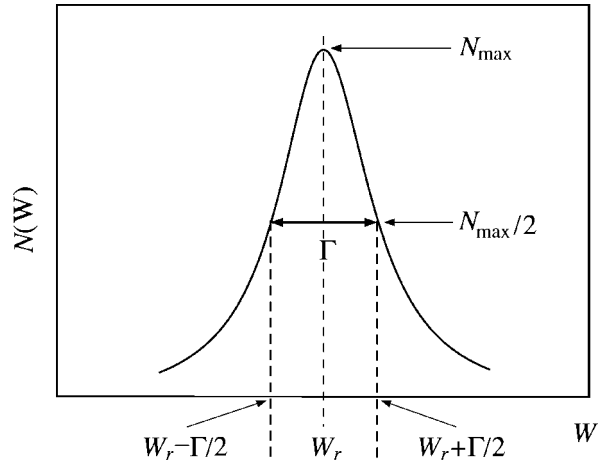


Figure 3.10 Plot of the Breit–Wigner formula (3.26).

Since the invariant mass W is identical to the rest mass for a single particle,[8] this can be regarded as a mass distribution corresponding to a mean value W_r and an uncertainty $\Gamma/2$. In practice, it is more usual to redefine the mass M as being equal to W_r, so that the Breit–Wigner formula becomes (cf. Equation (B.41))

$$N(W) = \frac{K}{(W - M)^2 + \Gamma^2/4}. \tag{3.27}$$

---

[7] This formula is closely analogous to that used to describe the natural line width of an excited state of an atom, which is an unstable particle made of a nucleus and electrons, rather than quarks and antiquarks. In this context, we note that (3.26) becomes

$$N(E) = \frac{K}{(E - E_r)^2 + \Gamma^2/4}$$

in the rest frame of the decaying particle, and putting $E = \omega$ (remember $\hbar = 1$ here) and $\Gamma = 1/\tau$ we obtain

$$N(\omega) = \frac{K}{(\omega - \omega_r)^2 + 1/4\tau^2}.$$

This is just the Lorentzian line shape used in atomic spectroscopy. (A simple semiclassical discussion of its origin in this context is given in Sections 3.14 to 3.26 of Wichmann (1967).)

[8] See Appendix A, Equation (A.10b).

The masses and widths of resonances are determined by fitting this formula to the experimentally observed peaks. In doing this, careful account must be taken of nonresonant backgrounds. When this is done for K^{*-}, using later data with much higher statistics than those shown in Figure 3.10, a mass and width of 892 MeV and 51 Mev, respectively, are obtained. A similar resonance is also observed in the reaction sequence

$$K^- + p \to \bar{K}^{*0} + n \tag{3.28a}$$

and

$$\bar{K}^{*0} \to K^- + \pi^+ \tag{3.28b}$$

and found to have mass and width values of 896 and 51 MeV, respectively, which are almost identical to those of K^{*-}.

Resonances are characterized not only by their masses and widths but also by their various quantum numbers. These are determined from observations on their decay products by exploiting the appropriate conservation laws. In particular, their spins can be determined from observations on the angular distributions of their decay products, using angular momentum conservation. We shall not discuss this here, but merely quote the result that both $K^{*-}(896)$ and $K^{*0}(892)$ have spin-1, like the photon. Their internal quantum numbers are obtained much more easily from (3.6) and (3.22), using the known values for the other particles given in Tables 2.3 and 2.4. In this way, one obtains

$$B = 0\,, S = -1, C = \tilde{B} = T = 0 \tag{3.29}$$

for both resonances. Within the quark model, the only allowed combinations compatible with (3.29) are $s\bar{u}$ and $s\bar{d}$, with charges $Q = -1, 0$, respectively. We thus arrive at the unique quark assignments

$$K^{*-}(896) = s\bar{u} \text{ and } K^{*0}(892) = s\bar{d}, \quad S = -1 \tag{3.30}$$

so that the quark description of the strong decay (3.28b) is

$$(s\bar{d}) \to (s\bar{u}) + (u\bar{d}),$$

involving the creation of a $u\bar{u}$ pair. This is summarized by the quark diagram of Figure 3.11. The continuity of quark lines is a general feature of quark diagrams for strong decays and reflects the fact that quarks and antiquarks can only be created or destroyed as particle–antiparticle pairs in strong interactions. A similar quark diagram, involving the creation of a $d\bar{d}$ pair, describes the decay (3.22b) (where $X^- \equiv K^{*-}$) and is left as an exercise for the reader.

At this point we note that the quark assignments (3.30) are identical to those of the kaons given in (3.16). Subsequently, many other states with these quantum numbers have been discovered and Figure 3.12 shows the spectrum of states corresponding to the $s\bar{u}$ system observed to date. The lightest state (the ground state) is the long-lived

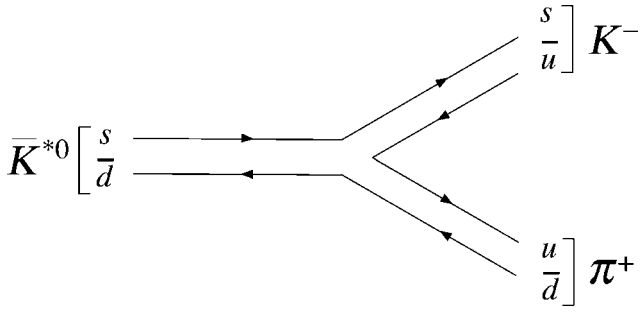


Figure 3.11 Quark diagram for the decay $\bar{K}^{*0} \to K^- + \pi^+$.

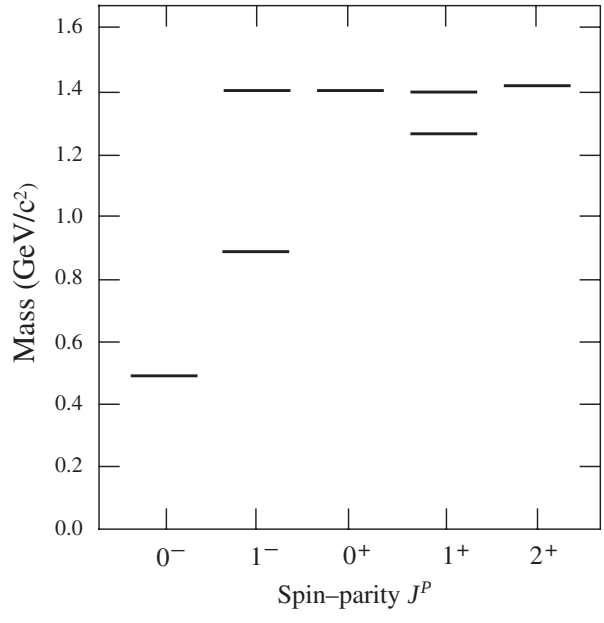


Figure 3.12 Observed bound states of the $s\bar{u}$ system with masses below 1.5 GeV/c^2, together with values of their spin-parities[9] J^P. The ground state is the $K^-(494)$ and the others can be interpreted as its excited states.

K^- meson discussed in Section 3.4, which decays by the weak interaction processes $K^- \to \mu^- + \bar{\nu}_\mu, K^- \to \pi^- + \pi^0$. The heavier states (excited states) are resonances that decay by the strong interaction with widths typically of order 50–250 MeV.

This picture is not restricted to strange mesons, but applies qualitatively to all quark systems, $u\bar{d}, uud, uds$, etc. Each system has a ground state, which is usually a long-lived particle decaying by weak or electromagnetic interactions, and a number of excited (resonance) states. The resulting spectra are qualitatively similar to that

[9] Parity is not discussed until Chapter 5, but is included here for later convenience.

shown in Figure 3.12 and the analogy with energy-level diagrams of other composite systems, like atoms and nuclei, is obvious.

3.6 ALLOWED QUANTUM NUMBERS AND EXOTICS

So far we have emphasized the quark interpretation of the internal quantum numbers

$$Q, B, C, S \text{ and } \tilde{B}, \tag{3.31}$$

but when a hadron is discovered the identities of its quark constituents are not known. Here we will show how the quantum numbers (3.31) can be determined for hadrons without assuming anything about their quark contents, and how the resulting values can be used to test the simple quark model assumption that mesons, baryons and antibaryons have the quark compositions

$$q\bar{q}, qqq \text{ and } \bar{q}\bar{q}\bar{q}, \tag{3.32}$$

respectively, rather than more complicated possibilities like

$$qq\bar{q}\bar{q} \text{ and } qqqq\bar{q}. \tag{3.33}$$

We first consider the allowed values of charge and strangeness corresponding to the usual quark compositions (3.32), and since the vast majority of known hadrons have

$$C = \tilde{B} = 0, \tag{3.34}$$

we confine ourselves initially to this case only.[10] For baryons, the possible qqq combinations are then

$$sss, \quad ssq_i, \quad sq_iq_j, \quad q_iq_jq_k, \tag{3.35}$$

where s is the strange quark and $q_i = u, d$, because of (3.34). The allowed strangeness values are therefore $S = -3, -2, -1$ and 0, and the allowed charges are obtained by adding the quark charges. Thus a baryon sss with $S = -3$ must have charge $Q = -1$, while a baryon

$$ssq_j \equiv ssu, ssd$$

with $S = -2$, can have $Q = 0$ or $Q = -1$. The remaining possibilities are

$$sl_il_j \equiv suu\,, sud\,, sdd,$$

[10] The generalization to hadrons with nonzero charm, which are not so well explored experimentally, is discussed in Problem 3.3.

with $S=-1$ and $Q=1$, 0 or -1, and

$$l_i l_j l_k \equiv uuu\,, uud\,, udd\,, ddd,$$

with $S=0$ and $Q=2$, 1, 0 or -1. These values are collected together in Table 3.5(a), and the corresponding values for meson combinations $q\bar{q}$, left as an exercise for the reader, are given in Table 3.5(b). Hadrons with combinations of S and Q that do not occur in these tables are called *exotic*, since they can arise only from quark combinations like (3.33) that are not included in the simple quark model. Examples of exotic hadrons would be a meson $dd\bar{u}\bar{s}$ with $S=1$ and $Q=-1$, or a baryon $uuud\bar{s}$ with $S=1$ and $Q=2$.

TABLE 3.5 The combinations of charge and strangeness allowed for hadrons in the simple quark model.

(a) Baryons		(b) Mesons	
S	Q	S	Q
0	2, 1, 0, −1	1	1, 0
−1	1, 0, −1	0	1, 0, −1
−2	0, −1	−1	0, −1
−3	−1		

To determine whether a given hadron is exotic or not, we must measure its internal quantum numbers without assuming anything about its quark content. This can be done by starting from the values of the quantum numbers B, S and $\tilde{B}$ for the particles[11]

$$p(938), n(940), K^-(494), D^-(1869), B^-(5278) \tag{3.36}$$

shown in Tables 3.2 and 3.3. Given these values, those for all other hadrons can be determined by exploiting the associated conservation laws. For example, by applying strangeness conservation to the observed strong interaction processes

$$p + p \to p + n + \pi^+, \tag{3.37a}$$

$$p + p \to p + p + \pi^0, \tag{3.37b}$$

$$\pi^- + p \to \pi^0 + n, \tag{3.37c}$$

$$K^- + p \to \pi^0 + \Lambda \tag{3.37d}$$

and

$$\pi^- + p \to K^+ + \pi^- + \Lambda \tag{3.37e}$$

[11] This choice of reference hadrons is not unique, but is convenient.

and assuming the strangeness values of Table 2.6, we find that the pions have strangeness $S = 0$, K^+ has $S = 1$ and Λ has $S = -1$ without using their quark composition. Similar arguments apply to the other quantum numbers, and enable us to construct a catalogue of their values for all known hadrons by considering a sufficient number of observed reactions like (3.37).[12]

More than two hundred hadrons have been identified and their internal quantum numbers deduced by the above methods. Without exception, they conform to the values allowed by the simple quark model assumption that only meson states $q\bar{q}$ and baryon states $3q$ occur. Despite extensive experimental searches for exotic hadrons, which could only arise from more complicated combinations like those of (3.33), not a single well-established example has been found.[13]

PROBLEMS 3

3.1 Classify the following experimentally observed processes into strong, electromagnetic and weak reactions by considering the particles involved and the appropriate selection rules.

(a) $\pi^- \to \pi^0 + e^- + \bar{\nu}_e$
(b) $\gamma + p \to \pi^+ + n$
(c) $p + \bar{p} \to \pi^+ + \pi^- + \pi^0$
(d) $D^- \to K^+ + \pi^- + \pi^-$
(e) $\Lambda + p \to K^- + p + p$
(f) $\pi^- + p \to n + e^+ + e^-$

3.2 The particles $X^0(1193)$ and $Y^-(1321)$ can be produced in the strong interaction processes

$$K^- + p \to \pi^0 + X^0$$

and

$$K^- + p \to K^+ + Y^-,$$

respectively. Deduce the baryon number, strangeness, charm and bottom in each case and, using these, their quark contents. The $X^0(1193)$ and $Y^-(1321)$ decay by the reactions

$$X^0 \to \Lambda + \gamma \text{ and } Y^- \to \Lambda + \pi^-,$$

respectively. Use the 'typical values' of Table 3.5 to give rough estimates of their lifetimes.

3.3 Six observed hadrons have the quantum numbers $(Q, B, S, C, \tilde{B}) = (2,1,0,1,0)$, $(0,1,-2,1,0)$, $(0,0,1,0,-1)$, $(0,-1,1,0,0)$, $(0,1,-1,1,0)$ and $(-1,1,-3,0,0)$, respectively. Identify their quark constituents.

3.4 Draw a lowest-order Feynman diagram at the quark level for the following decays:

(a) $D^- \to K^0 + \pi^-$
(b) $\Lambda \to p + e^- + \bar{\nu}_e$

3.5 Estimate the mean distance travelled by a resonance produced in a reaction like (3.22a) with $\gamma = E/m \approx 10$, before it decays by a process like (3.22b). Compare this with the

[12] The known particles all have top $T = 0$.
[13] From time to time experiments have claimed evidence for exotic states, but as more experiments have been performed and far greater quantities of data produced, the evidence for exotic hadrons has receded. The present consensus is that there is no convincing evidence for the existence of such states.

shortest distance that can be resolved experimentally (of order 1 μm) and with the range of the strong interactions. Why is this latter comparison significant?

3.6 Derive the allowed combinations of charm C and electric charge Q for mesons and baryons in the quark model.

3.7 Which of the following possible hadronic states with quantum numbers $(Q, B, S, C, \tilde{B}) = (1,0,0,1,1)$, $(-1,1,-2,0,-1)$, $(0,0,1,0,1)$ and $(-1,1,0,1,-1)$ are compatible with the quark model and which are not?

4

Experimental Methods

It may be so, there is no arguing against facts and experiment

Isaac Newton[1]

In this chapter we will take a brief look at experimental methods. This is a very extensive subject and the aim will not be to give a comprehensive review, but rather to emphasize the physical principles behind the methods. More details may be found in specialized texts.[2]

4.1 OVERVIEW

In particle physics, high energies are needed both to create new and unstable particles and to explore the structure of hadrons. Until the early 1950s the only source of such high-energy particles was cosmic rays, and studies of them led to many notable discoveries, including positrons and pions. However, cosmic rays are now used only in special circumstances, and the majority of experiments are conducted using beams of particles produced by machines called *accelerators.* This has the great advantage that the projectiles are of a single type and have energies that may be controlled by the experimenter. For example, beams that are essentially monoenergetic may be prepared, and can be used to study the energy dependence of interactions.

The beam, once established, is directed on to a target so that interactions may be produced. In a *fixed-target* experiment the target is stationary in the laboratory. Although

[1] In reply to a friend on being told of observations that were thought – wrongly – to contradict the Newtonian system. See p. 59 of Andrade (1961).

[2] See, for example, Fernow (1986), Kleinknecht (1986) and Leo (1994). There is also a review in Yao *et al.* (2006).

Particle Physics Third Edition B. R. Martin and G. Shaw

historically of great importance, the main disadvantage of this type of experiment is apparent with the need to work at higher centre-of-mass energies. The centre-of-mass energy is important because it is a measure of the energy available to create new particles. In the laboratory frame at least some of the final-state particles must be in motion to conserve momentum. Consequently, at least some of the initial beam energy must reappear as kinetic energy of the final-state particles, and is therefore unavailable for particle production. The centre-of-mass energy E_{CM} is given in terms of the projectile's laboratory energy E_L by[3]

$$E_{CM} = \left(m_b^2 c^4 + m_t^2 c^4 + 2m_t c^2 E_L\right)^{1/2}. \tag{4.1}$$

At high energies this increases only like $E_L^{1/2}$ and so an increasingly smaller fraction of the beam energy is available for particle production, the rest going to impart kinetic energy to the target.

In contrast, if we could work in the centre-of-mass system, all the available energy would in principle be available for particle production. This is achieved in a *colliding-beam* experiment, where two beams of particles travelling in almost opposite directions are made to collide at a small, or zero, crossing angle. If for simplicity we assume the particles in the two beams have the same mass and laboratory energy and collide at zero crossing angle, then the total centre-of-mass energy is

$$E_{CM} = 2E_L. \tag{4.2}$$

This increases linearly with the energy of the accelerated particles and is a significant improvement on the fixed-target result. Colliding-beam experiments are not, however, without their own disadvantages. The colliding particles have to be stable, which limits the interactions that can be studied, and the collision rate in the intersection region is smaller than that achieved in fixed-target experiments, because the particle densities in the beams are low compared to a solid or liquid target.

In addition to the beam energy, an accelerator is characterized by its *luminosity* L, which gives the rate for a given reaction in terms of the cross-section for the process.[4]

Finally, details of the particles produced in the collision (e.g. their energies) are deduced by observing their interactions with the material of *detectors*, which are placed in the vicinity of the interaction region. A wide range of detectors is available. Some have a very specific characteristic; others serve more than one purpose. Modern experiments in particle physics typically use several types in a single experiment.

In this chapter we start by describing some of the different types of accelerator that have been built, the beams that they can produce and also how beams of neutral and unstable particles can be prepared. Then we discuss the ways in which particles interact with matter and review how these mechanisms are exploited in the construction

[3] This formula is derived as Equation (A.12) in Appendix A, where a brief summary of relativistic kinematics may be found. In this chapter, since we are concerned with practical matters, we use practical units, retaining factors of $\hbar$ and c.

[4] Luminosity is formally defined in Appendix B, Equation (B.4).

of a range of particle detectors. Finally, by considering some examples, we illustrate how these individual detectors are combined into modern complex detector systems.

4.2 ACCELERATORS AND BEAMS

All accelerators use electromagnetic forces to boost the energy of stable charged particles. These are injected into the machine from a device that provides a high-intensity source of low-energy particles, for example an electron gun (a hot filament), or a proton ion source. The accelerators used for particle physics use radio frequency (r.f.) electric fields to accelerate the particles and may conveniently be divided into *linear* and *cyclic* varieties.

4.2.1 Linear accelerators

In a linear accelerator (or *linac*) for accelerating protons, the particles pass through a series of metal pipes called *drift tubes* that are located in a vacuum vessel and connected successively to alternate terminals of an r.f. oscillator, as shown in Figure 4.1. The positively charged particles accelerated by the field move towards the first drift tube. If the alternator can change its direction as the particles pass through that tube, they will then be accelerated again on their way between the exit of the first and entry to the second tube, and so on. Thus the particles will form bunches. Because the particles are accelerating, their speed is increasing and hence the lengths of the drift tubes have to increase to ensure continuous acceleration. To produce a useful beam, the particles must keep in phase with the r.f. field and remain focused. Proton linacs of this type are often used as *injectors*; i.e. they produce protons of moderate energy that are then injected into a more powerful machine, usually a synchrotron, where they are accelerated to much higher energies.

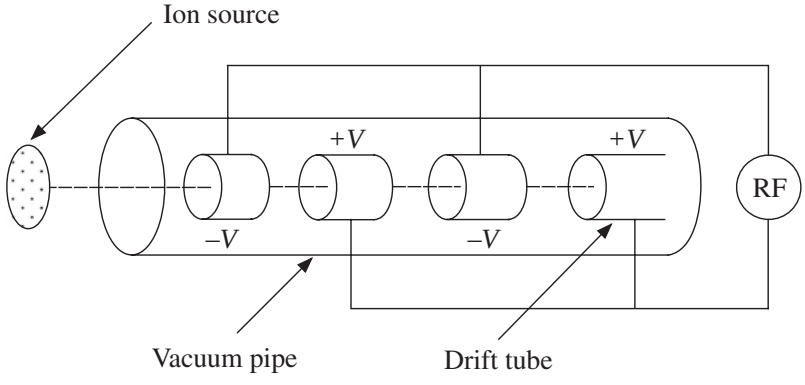


Figure 4.1 Acceleration in a linear proton accelerator.

For electrons, a variation of this method is used. In this case the accelerator consists of a straight tube in the form of a series of cylindrical metal cavities. Power is fed to the accelerator from a series of devices called klystrons, which produce electromagnetic radiation in the form of microwave pulses that are transported via waveguides to

the accelerator. There they generate an oscillating electric field pointing along the direction of the metal tube and a magnetic field in a circle around the interior of the accelerating tube. The magnetic field helps to keep the beam focused, and the frequency of the microwaves is adjusted so that the electrons arrive at each cavity of the accelerator at the optimal time to receive the maximum energy boost from the electric field. As long as this phase relationship can be maintained, the particles will be continuously accelerated. The largest electron linac is the SLC at the SLAC Laboratory in Stanford, California, USA, and has a maximum energy of 50 GeV. It consists of 80 000 copper cavities separated by copper discs with a small hole at the centre to transmit the beam, and is over 3 km long.

An ingenious way of reducing the enormous lengths of high-energy linacs has been developed at the Continuous Electron Beam Accelerator Facility (CEBAF) at the Jefferson Laboratory in the USA. This utilizes the fact that above about 50 MeV, electron velocities are very close to the speed of light and thus electrons of very different energies can be accelerated in the same device. Instead of a single long linac, the CEBAF machine consists of two much shorter linacs and the beam from one is bent and passed through the other. This can be repeated for up to four cycles. Even with the radiation losses inherent in bending the beams, very intense beams can be produced with energies between 0.5 and 6.0 GeV. CEBAF is proving to be an important machine for studying the details of nucleon structure.

4.2.2 Cyclic accelerators

Cyclic accelerators used in particle physics are of a type called *synchrotrons*. The principle of a synchrotron is analogous to that of a linear accelerator, but the acceleration takes place in a near-circular orbit rather than in a straight line. The beam of particles travels in an evacuated tube called the *beam pipe* and is constrained in a circular or near-circular path by an array of dipole magnets called bending magnets (see Figure 4.2(a)).

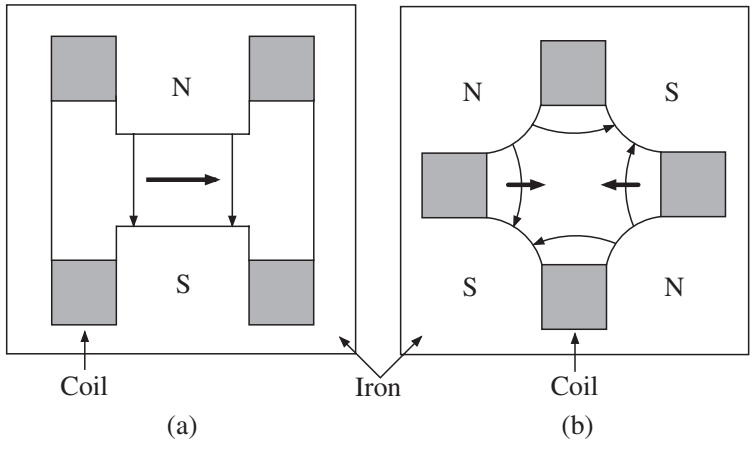


Figure 4.2 Cross-section of a typical bending (dipole) magnet (left diagram) and a focusing (quadrupole) magnet (right diagram). The thin arrows indicate field directions and the thick arrows indicate the force on a negative particle travelling into the paper.

Acceleration is achieved as the beam repeatedly traverses a number of cavities placed in the ring where energy is given to the particles. Since the particles travel in a circular orbit they continuously emit radiation, called in this context *synchrotron radiation*. The amount of energy radiated per turn by a relativistic particle of mass m and charge q is given by[5]

$$\Delta E = \frac{4\pi q^2 \beta^3 \gamma^4}{3\varepsilon_0 \rho}, \tag{4.3}$$

where $\beta \equiv \upsilon/c$, υ is the particle's velocity, $\gamma \equiv (1-\beta^2)^{-1/2}$ and ρ is the radius of curvature of the orbit. For relativistic particles ($\beta \sim 1$) of a given energy γmc^2, the energy loss is inversely proportional to the fourth power of the particle's mass. The losses for electrons are thus very severe, and the need to compensate for these by the input of large amounts of r.f. power limits the energies of electron synchrotrons.

The momentum in GeV/c of an orbiting particle assumed to have unit charge is given by $p = 0.3B\rho$, where B is the magnetic field in tesla and ρ, the radius of curvature, is measured in metres. Because p is increased during acceleration, B must also be steadily increased if ρ is to remain constant, and the final momentum is limited both by the maximum field available and by the size of the ring. With conventional electromagnets, the largest field attainable over an adequate region is about 1.5 T, and even with superconducting coils it is only of order 10 T. Hence the radius of the ring must be very large to achieve very high energies. For example, the large hadron collider (LHC) accelerator, located at the CERN Laboratory, Geneva, which accelerates protons to energies of 7 TeV, has a radius of about 4 km.[6] A large radius is also important to limit synchrotron radiation losses in electron machines.

In the course of its acceleration, a beam may make typically 10^5 traversals of its orbit before reaching its maximum energy. Consequently, stability of the orbit is vital, both to ensure that the particles continue to be accelerated and that they do not strike the sides of the vacuum tube. In practice, the particles are accelerated in bunches, each being synchronized with the r.f. field. In equilibrium, a particle increases its momentum just enough to keep the radius of curvature constant as the field **B** is increased during one rotation, and the circulation frequency of the particle is in step with the r.f. of the field. This is illustrated in Figure 4.3(a). Particle B is assumed to be in an equilibrium orbit, synchronous with the r.f. field. Particle A, behind the r.f. phase, receives a lower momentum increase from the field than particle B. This will reduce the radius of its orbit and (since $\upsilon \approx c$) increase its rotational frequency relative to particle B. Conversely, a particle C, ahead of the r.f. phase, receives a greater momentum increase and a decrease in its rotational frequency. With obvious changes, a similar principle is used in linear accelerators.

In practice, the particles remain in the bunch, but their trajectories oscillate about the stable orbits. These oscillations are controlled by a series of focusing magnets, usually of the quadrupole type, which are placed at intervals around the beam and act

[5] See, for example, p. 661 of Jackson (1975). His results must be divided by $4\pi\varepsilon_0$ to convert from Gaussian to SI units.

[6] The beam line under construction in the LHC tunnel is shown in coloured Plate 1.

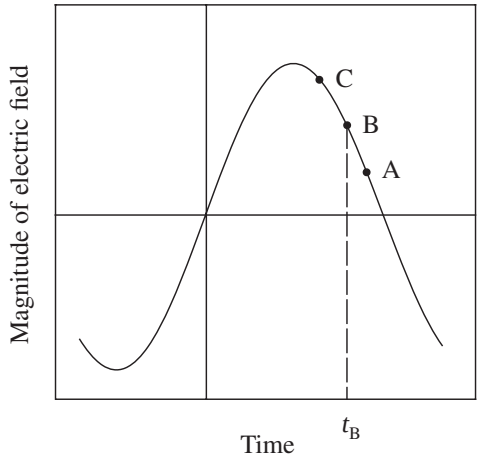


Figure 4.3 Magnitude of the electric field as a function of time at a fixed point in the r.f. cavity. Particle B is synchronous with the field and arrives at time t_B. Particle A (C) is behind (ahead of) B and receives an increase (decrease) in its rotational frequency. Thus particles oscillate about the equilibrium orbit.

like optical lenses. A schematic diagram of one of these is shown in Figure 4.2(b). Each focuses the beam in one direction and de-focuses it in the orthogonal direction, so alternate magnets have their field directions reversed to keep the particles in a stable orbit.

In addition to the energy of the beam, one is also concerned to produce a beam of high intensity, so that interactions will be plentiful. The intensity is ultimately limited by de-focusing effects, e.g. the mutual repulsion of the particles in the beam, and a number of technical problems have to be overcome that are outside the scope of this brief account.

4.2.3 Fixed-target machines and colliders

Both linear and cyclic accelerators can be divided into *fixed-target* and *colliding-beam* machines. The latter are also known as *colliders,* or sometimes in the case of cyclic machines, *storage rings*.[7] In fixed-target machines, particles are accelerated to the highest operating energy and then the beam is extracted from the machine and directed on to a stationary target, which is usually a solid or liquid. Much higher energies have been achieved for protons than electrons, because of the large radiation losses inherent in electron machines, mentioned earlier. The intensity of the beam is such that large numbers of interactions can be produced, which can either be studied in their own right or used to produce secondary beams.

The main disadvantage of fixed-target machines for particle physics has been mentioned earlier: the need to achieve large centre-of-mass energies to produce new particles. Almost all new machines for particle physics are therefore colliders,

[7] The use of the terms *storage rings* and *colliders* as synonymous is not strictly correct, because the former can also describe a machine that stores a single beam for use on both internal and external fixed targets.

although some fixed-target machines for specialized purposes are still constructed.[8] The largest collider currently operational is the LHC at CERN, mentioned above. This is a massive *pp* accelerator of circumference 27 km, with a design luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and with each beam having an energy of 7 TeV. A schematic diagram of the CERN site showing the LHC and some of its other accelerators is shown in Figure 4.4. The acceleration process starts with a linac, whose beam is boosted in energy in the PSB (proton synchrotron booster) and passed to the PS (proton synchrotron), a machine that is still the source of beams for lower-energy experiments. The beam energy is increased still further in the SPS (super proton synchrotron), which also provides beams for a range of experiments as well as the injection beams for the LHC itself. Four beam intersection points are shown in the LHC and experiments

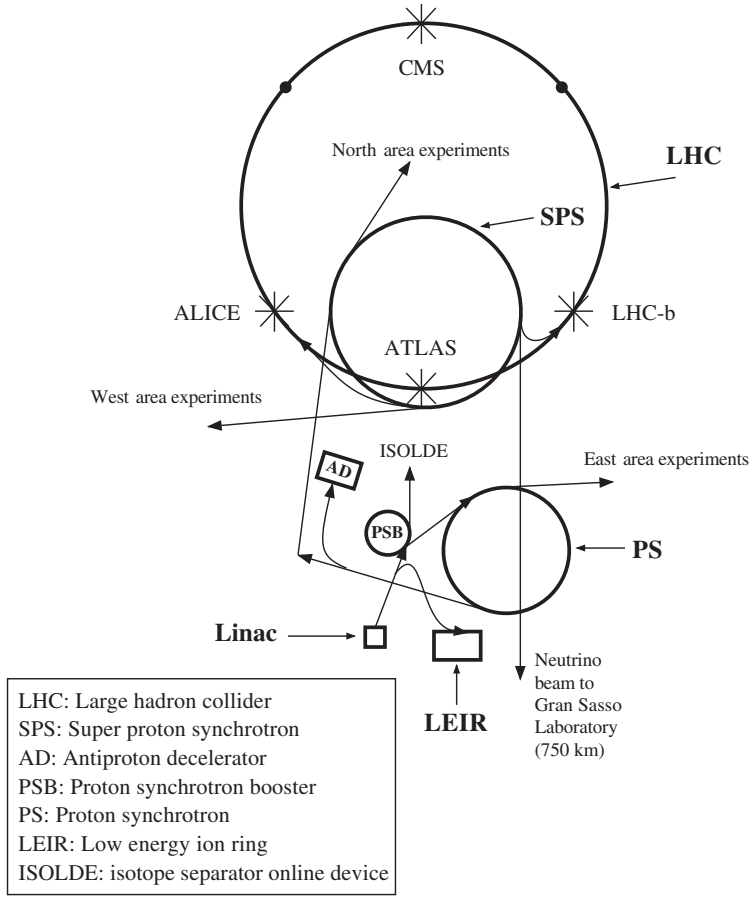


Figure 4.4 A schematic diagram of the CERN site showing the LHC and some of its other accelerators.

[8] The operational details of colliders currently operational or under construction may be found in Yao *et al*. (2006).

(ALICE, CMS, LHC-b and ATLAS) are located at each of these.[9] The extracted neutrino beam shown at the bottom of the diagram is sent to the Gran Sasso Laboratory 750 km away where neutrino experiments are performed.

Other large colliders currently operational include the Tevatron, mentioned above, and the relativistic heavy ion collider (RHIC), located at Brookhaven National Laboratory, USA. The latter machine, which began operation in 2000 following 10 years of development and construction, is the first collider capable of accelerating heavy ions to high energies. Like the LHC, there are several stages before the ions are injected into the main machine. There they form two counter-circulating beams controlled by two 4 km rings of superconducting magnets and are accelerated to an energy of 100 GeV/nucleon. Thus the total centre-of-mass energy is 200 GeV/nucleon. Collisions occur at six intersection points, where major experiments are sited. RHIC primarily accelerates ions of gold and is used to study matter at extreme energy-densities, where a new state of matter called a 'quark–gluon plasma' is predicted to occur. We will return to this briefly in Chapter 7.

4.2.4 Neutral and unstable particle beams

The particles used in accelerators must be stable and charged, but one is also interested in the interaction of neutral particles, e.g. photons and neutrons, as well as those of unstable particles (such as charged pions). Beams appropriate for performing such experiments are produced in a number of ways.

Beams of unstable particles can be formed, provided the particles live long enough to travel appreciable distances in the laboratory. One way of doing this is to direct an extracted primary beam on to a heavy target. In the resulting interactions with the target nuclei, many new particles are produced, which may then be analysed into secondary beams of well-defined momentum. Such beams will ideally consist predominantly of particles of one type, but if this cannot be achieved, then the wanted species may have to be identified by other means. In addition, if these secondary beams are composed of unstable particles, they can themselves be used to produce further beams formed from their decay products. Two examples will illustrate how in principle such secondary particle beams can be formed.

Consider, firstly, the construction of a π^+ beam from a primary beam of protons. By allowing the protons to interact with a heavy target, secondary particles, most of which will be pions, will be produced. A collimator can be used to select particles in a particular direction, and the π^+ component can subsequently be removed and focused into a mono-energetic beam by selective use of electrostatic fields and bending and focusing magnets.

The pion beam may also be used to produce a beam of neutrinos. For example, the π^- is unstable and, as we have seen, its dominant decay mode is $\pi^- \to \mu^- + \bar{\nu}_\mu$. So if pions are passed down a long vacuum pipe, many will decay in flight to give muons and antineutrinos, which will mostly travel in essentially the same direction as the initial beam. The muons and any remaining pions can then be removed by passing the beam through a very long absorber, leaving the neutrinos. In this case, the

[9] Pictures of the ATLAS and CMS detectors during their construction phase are shown in coloured Plates 2 and 3, respectively.

final neutrino beam will not be monoenergetic, but instead will have a momentum spectrum reflecting the initial momentum spectrum of the pions, and since neutrinos are neutral, no further momentum selection using magnets is possible.

4.3 PARTICLE INTERACTIONS WITH MATTER

In order to be detected, a particle must undergo an interaction with the material of a detector. In this section we discuss these interactions, but only in sufficient detail to be able to understand the detectors themselves.

The first possibility is that the particle interacts with an atomic nucleus. For example, this could be via the strong nuclear interaction if it is a hadron or by the weak interaction if it is a neutrino. We know from the work of Chapter 1 that both are *short-range interactions*. If the energy is sufficiently high, new particles may be produced, and such reactions are often the first step in the detection process. In addition to these short-range interactions, a charged particle will also excite and ionize atoms along its path, giving rise to *ionization energy losses*, and emit radiation, leading to *radiation energy losses*. Both of these processes are due to the long-range electromagnetic interaction. They are important because they form the basis of most detectors for charged particles. Photons are also directly detected by electromagnetic interactions, and at high energies their interactions with matter lead predominantly to the production of e^+e^- pairs via the *pair production* process $\gamma \rightarrow e^+ + e^-$, which has to occur in the vicinity of a nucleus to conserve energy and momentum. (Recall the discussion in Section 1.4.1 on the range of forces.) All these types of interactions are described in the following sections.

4.3.1 Short-range interactions with nuclei

For hadrons, the most important short-range interactions with nuclei are due to the strong nuclear force, which unlike the electromagnetic interaction is as important for neutral particles as for charged ones. For the simplest nucleus, the proton, the resulting reactions are of two types: *elastic scattering* such as

$$\pi^- + p \rightarrow \pi^- + p, \tag{4.4a}$$

where for illustration we have taken the incident particle to be a π^-; and *inelastic reactions* such as

$$\pi^- + p \rightarrow \pi^+ + \pi^- + \pi^0 + n \tag{4.4b}$$

and

$$\pi^- + p \rightarrow K^0 + \Lambda, \tag{4.4c}$$

in which the final state particles differ from those in the initial state. At high energies, many inelastic reactions are possible, most of them involving the production of several particles in the final state. The total cross-section

$$\sigma_{tot} = \sigma_{el} + \sigma_{inel} \tag{4.5}$$

is the sum of the cross-section for elastic scattering σ_{el} and that for inelastic reactions σ_{inel}, where the latter is itself a sum over all possible inelastic processes allowed by conservation laws.[10] The behaviours of the total and elastic cross-sections for $\pi^- p$ scattering are shown in Figure 4.5.

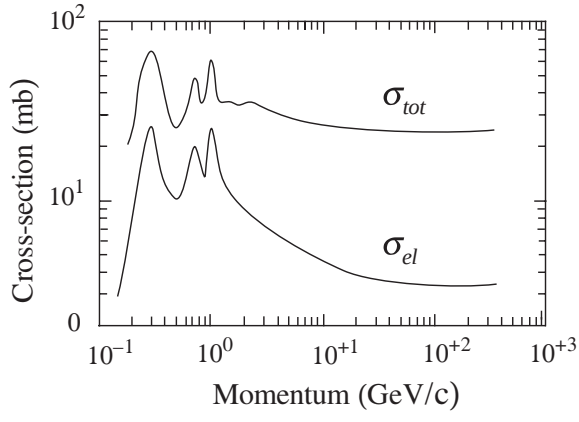


Figure 4.5 Total and elastic cross-sections for $\pi^- p$ scattering as functions of the pion laboratory momentum.

There is considerable structure at low energies due to the production of resonances, but for momenta above about 3 GeV/c the total cross-section is slowly varying and much larger than the elastic cross-section. The same general features are found for other incident hadrons at high energies, with total cross-sections typically in the range 10–50 mb. This is of the same order of magnitude as the 'geometrical' cross-section $\pi r^2 \approx 30$ mb, where $r \approx 1$ fm is the approximate range of the strong interaction between hadrons. Total cross-sections on neutrons are similar to those on protons, but total cross-sections on nuclei are much larger, increasing roughly like the square of the nuclear radius.

The probability of a hadron–nucleus interaction occurring as the hadron traverses a small thickness dx of material is given by $n\sigma_{tot}\,\mathrm{d}x$, where n is the number of nuclei per unit volume in the material. Consequently, the mean distance travelled before an interaction occurs is given by

$$l_c = 1/n\sigma_{tot}. \tag{4.6}$$

This is called the *collision length*.[11] An analogous quantity is the *absorption length*, defined by

$$l_a = 1/n\sigma_{inel}, \tag{4.7}$$

[10] The various types of cross-section and related quantities are formally defined in Appendix B.
[11] There is a further discussion of this in Section B.2, Appendix B.

that governs the probability of an inelastic collision. In practice, $l_c \approx l_a$ at high energies. This is illustrated in Table 4.1, where l_a and l_c are given for incident neutrons on several common materials.

TABLE 4.1 Nuclear cross-sections and the associated collision lengths l_c and absorption lengths l_a for incident neutrons with energies in the range 80–300 GeV. The values for protons are approximately the same.

Element	Z	Nuclear cross-section (b)		Interaction length (cm)	
		σ_{tot}	σ_{inel}	l_c	l_a
H*	1	0.039	0.033	687	806
C	6	0.331	0.231	26.6	38.1
Al	13	0.634	0.421	26.1	39.4
Fe	26	1.120	0.703	10.5	16.8
Pb	82	2.960	1.770	10.2	17.1

* Liquid hydrogen at 26 K. The other materials are solids.

Neutrinos and antineutrinos can also be absorbed by nuclei, leading to reactions of the type

$$\bar{\nu}_\ell + p \to \ell^+ + X, \tag{4.8}$$

where ℓ is a lepton and X denotes any hadron or set of hadrons allowed by the conservation laws. Such processes are weak interactions (because they involve neutrinos) and the associated cross-sections are extremely small compared to the cross-sections for strong interaction processes. The corresponding interaction lengths are therefore enormous. Nonetheless, in the absence of other possibilities such reactions are the basis for detecting neutrinos. Finally, photons can be absorbed by nuclei, giving *photoproduction* reactions such as $\gamma + p \to X$. However, these electromagnetic interactions are only used to detect photons at low energies, because at higher energies there is a far larger probability for e^+e^- pair production in the Coulomb field of the nucleus. We will return to this in Section 4.3.3 below.

4.3.2 Ionization energy losses

Ionization energy losses are important for all charged particles, and for particles other than electrons and positrons they dominate over radiation energy losses at all but the highest attainable energies. The theory of such losses, which are due dominantly to Coulomb scattering from the atomic electrons,[12] was worked out by Bethe, Bloch and others in the 1930s.[13] The result is called the Bethe–Bloch formula, and for spin-0

[12] The analogous scattering from the nucleus gives a negligible contribution to the energy loss since the recoil energy of the nucleus is much smaller because of its much greater mass.
[13] For a semiclassical discussion, see, for example, Chapter 13 of Jackson (1975).

bosons with charge $\pm q$ (in units of e), mass M and velocity υ, it takes the approximate form

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{D\,q^2 n_e}{\beta^2}\left[\ln\left(\frac{2m_e c^2\beta^2\gamma^2}{I}\right) - \beta^2 - \frac{\delta(\gamma)}{2}\right], \tag{4.9a}$$

where x is the distance travelled through the medium,

$$D = \frac{4\pi\,\alpha^2 h^2}{m_e} = 5.1 \times 10^{-25}\,\mathrm{MeV\,cm^2}, \tag{4.9b}$$

m_e is the electron mass, $\beta = \upsilon/c$ and $\gamma = (1-\beta^2)^{-1/2}$. The other constants refer to the properties of the medium: n_e is the electron density, I is the mean ionization potential of the atoms averaged over all electrons, which is given approximately by $I = 10Z$ eV for Z greater than 20, and δ is a dielectric screening correction that is important only for highly relativistic particles. The corresponding formula for spin-$\frac{1}{2}$ particles

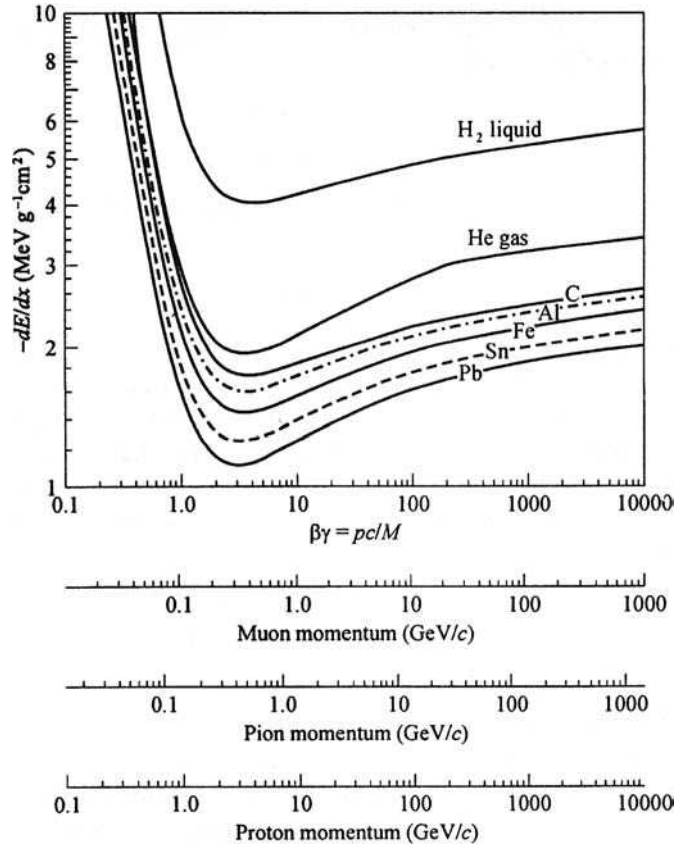


Figure 4.6 Ionization energy loss for muons, pions and protons on a variety of materials. The units of $\mathrm{d}E/\mathrm{d}x$ are explained in the text. (Reprinted by permission of Institute of Physics (IOP), Fig. 27.7, W.-M. Yao *et al.*, *Journal of Physics*, **G33**, 1, 2006.)

differs from this, but in practice the differences are small and may be neglected when discussing the main features of ionization energy losses.

It is common practice to divide (4.9) by the mass density ρ and to redefine x as the density multiplied by the distance travelled, so that

$$-\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x} \rightarrow -\frac{\mathrm{d}E}{\mathrm{d}x}, \tag{4.10}$$

where x is now measured in $\mathrm{g\,cm^{-2}}$. Examples of the behaviour of (4.10) for muons, pions and protons traversing a range of materials are shown in Figure 4.6, using this convention. As can be seen, $-\mathrm{d}E/\mathrm{d}x$. falls rapidly as the velocity increases from zero because of the $1/\beta^2$ factor in the Bethe–Bloch equation. All particles have a region of 'minimum ionization' for $\beta\gamma$ in the range 3 to 4. Beyond this, β tends to unity, and the logarithmic factor in the Bethe–Bloch formula gives a 'relativistic rise' in $-\mathrm{d}E/\mathrm{d}x$.

The magnitude of the energy loss depends on the medium. The electron density is given by $n_e = \rho N_A Z/A$, where N_A is Avogadro's number, and ρ and A are the mass density and atomic weight of the medium respectively, so the mean energy loss is proportional to the density of the medium. The remaining dependence on the medium is relatively weak because $Z/A \approx 0.5$ for all atoms except the very light and the very heavy elements, and because the ionization energy I only enters the Bethe–Bloch formula logarithmically. In the 'minimum ionization' region where $\beta\gamma \approx 3-4$, the minimum value of $-\mathrm{d}E/\mathrm{d}x$ can be calculated from (4.9) and for a particle with unit charge is given approximately (in the units of Figure 4.6) by

$$\frac{1}{\rho}\left(-\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{min} \approx 3.5\frac{Z}{A}\,\mathrm{MeV\,g^{-1}cm^2}, \tag{4.11}$$

where we revert to using x to denote the distance travelled, as usual.[14] Typical values for a minimum ionizing particle with unit charge are given in Table 4.2 and are in good agreement with (4.11) for a wide range of elements except the very lightest. Also shown are values of $(-\mathrm{d}E/\mathrm{d}x)_{min}$ in units of $\mathrm{MeV\,cm^{-1}}$ and the corresponding value of the density.

Ionization losses are proportional to the squared charge of the particle, so that a fractionally charged particle with $\beta\gamma \geq 3$ would have a much lower rate of energy loss than the minimum energy loss of any integrally charged particle. This has been used as a means of identifying possible free quarks, but without success.

From the rate of energy loss as a function of β, we can calculate the rate of energy loss as a function of the distance x travelled in the medium. This is called the *Bragg curve*. Because of the factor β^{-2} in (4.9), most of the ionization loss occurs near the end of the path where the speed is smallest, and the curve has a pronounced peak

[14] In this book we only use x in $\mathrm{g\,cm^{-2}}$ in the right-hand side of (4.10) and in Figure 4.6.

TABLE 4.2 The minimum ionization energy losses $(-\mathrm{d}E/\mathrm{d}x)_{min}$ for various materials and their dependence on the density ρ in $\mathrm{g\,cm^{-3}}$.

Element	Z	ρ	$\left(-\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{min}$ $(\mathrm{MeV cm^{-1}})$	$-\frac{1}{\rho}\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{min}$ $(\mathrm{MeV g^{-1} cm^2})$
H*	1	0.063	0.26	4.12
C	6	2.26	4.02	1.78
Al	13	2.70	4.37	1.62
Fe	26	7.87	11.6	1.48
Pb	82	11.35	12.8	1.13

* Liquid hydrogen at 26 K. The other materials are solids.

close to the end point before falling rapidly to zero at the end of the particle's path length. For particles whose energy loss is dominated by ionization, the *range* R, i.e. the mean distance a particle travels before it comes to rest, is given by

$$R = \int_0^R \mathrm{d}x = \int_0^{\beta_\mathrm{i}} \left(-\frac{\mathrm{d}E}{\mathrm{d}x}\right)^{-1} \frac{\mathrm{d}E}{\mathrm{d}\beta}\mathrm{d}\beta = \frac{M}{q^2 n_e} F(\beta_\mathrm{i}), \tag{4.12}$$

where F is a function of the initial velocity β_i and we have used the relation $E = \gamma Mc^2$ to show the dependence on the projectile mass M. For nonrelativistic particles $\beta_i = 1$, the rate of energy loss is very high and the particle comes to rest very quickly. This was the basis for excluding the possibility that the upper track in Figure 1.2 could be due to a proton of momentum 23 MeV/c. Such a particle would be highly nonrelativistic ($\beta \approx 0.02$) and would come to rest in a short distance, which detailed calculation using (4.9) shows is of order 5 mm, which is much less than the length of the observed track.

Finally, we note that the range as given by (4.12) is actually an average value because scattering is a statistical process and there will therefore be a spread of values for individual particles. This is illustrated by the muon tracks in Figure 3.2. Here the four muons all have the same energy, since they all arise from the decay of pions at rest and all travel approximately the same distances in the emulsion before themselves coming to rest and subsequently decaying.

4.3.3 Radiation energy losses

When a charged particle traverses matter it can also lose energy by radiative collisions, especially with nuclei. The electric field of a nucleus will accelerate and decelerate the particles as they pass, causing them to radiate photons, and hence lose energy. This process is called *bremsstrahlung* (literally 'braking radiation' in German) and is a particularly important contribution to the energy loss for electrons and positrons.

The dominant Feynman diagrams for electron bremsstrahlung in the field of a nucleus, i.e.

$$e^- + (Z, A) \rightarrow e^- + \gamma + (Z, A), \tag{4.13}$$

are shown in Figure 4.7 and are of order $Z^2\alpha^3$, like the pair creation diagrams of Figure 1.11.

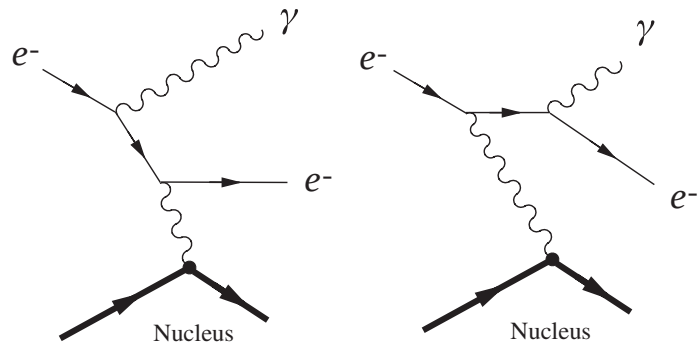


Figure 4.7 Dominant Feynman diagrams for the bremsstrahlung process $e^- + (Z, A) \rightarrow e^- + \gamma + (Z, A)$.

There are also contributions from bremsstrahlung in the fields of the atomic electrons, each of order α^3. Since there are Z atomic electrons for each nucleus, these give a total contribution of order $Z\alpha^3$, which is small compared to the contribution from the nucleus for all but the lightest elements. A detailed calculation shows that for relativistic electrons with $E \gg mc^2/\alpha Z^{1/3}$, the average rate of energy loss is given by

$$-\mathrm{d}E/\mathrm{d}x = E/L_R. \tag{4.14}$$

The constant L_R is called the *radiation length* and is given by

$$\frac{1}{L_R} = 4\left(\frac{\hbar}{mc}\right)^2 Z(Z+1)\alpha^3 n_a \ln\left(\frac{183}{Z^{1/3}}\right), \tag{4.15}$$

where n_a is the number density of atoms/cm^3 in the medium. Integrating (4.14) gives

$$E = E_0 \exp\left(-x/L_R\right), \tag{4.16}$$

where E_0 is the initial energy. It follows that the radiation length is the average thickness of material that reduces the mean energy by a factor e. Values of L_R for various materials are shown in Table 4.3.

From these results, we see that at high energies the radiation losses are proportional to E/m_P^2 for an arbitrary charged particle of mass m_p. On the other hand, the ionization energy losses are only weakly dependent on the projectile mass and energy at very high energies. Consequently, radiation losses completely dominate the energy losses for electrons and positrons at high enough energies, but are much smaller than ionization losses for all particles other than electrons and positrons at all but the highest energies. In Table 4.3 we show the energy E_C at which radiation losses become as large as ionization losses for electrons in various materials. For the next lightest particles, muons, E_C will be increased by a factor of approximately $(m_\mu/m)^2 \approx 2 \times 10^4$, which

TABLE 4.3 The radiation length L_R and the energy E_C, at which electron radiation energy losses equal their ionization energy losses, for various elements. Except for the very lightest elements, $E_C \approx 600/Z$ Mev.

Element	Z	L_R (cm)	E_C (MeV)
H_2*	1	≈ 1000	340
C	6	18.8	103
Al	13	8.9	47
Fe	26	1.8	24
Pb	82	0.56	7

* Liquid hydrogen at 26 K. The other materials are solids.

justifies our earlier statement that radiation losses are much smaller than ionization losses for all particles other than electrons and positrons at all but the highest energies.

4.3.4 Interactions of photons in matter

In contrast to heavy charged particles, photons have a high probability of being absorbed or scattered through large angles by the atoms in matter. Consequently, a collimated monoenergetic beam of I photons per second traversing a thickness dx of matter will lose

$$\mathrm{d}I = -I\frac{\mathrm{d}x}{\lambda} \tag{4.17}$$

photons per second, where

$$\lambda = (n_a \sigma_\gamma)^{-1} \tag{4.18}$$

is the mean free path before absorption or scattering out of the beam and σ_γ is the total photon interaction cross-section with an atom. The mean free path λ is analogous to the collision length for hadronic reactions. Integrating (4.17) gives

$$I(x) = I_0\, \mathrm{e}^{-x/\lambda} \tag{4.19}$$

for the intensity of the beam as a function of distance, where I_0 is the initial intensity.

The main processes contributing to σ_γ at the energies relevant for particle physics are: the *photoelectric effect*, in which the photon is absorbed by the atom as a whole with the emission of an electron; *Compton scattering*, where the photon scatters from an atomic electron; and *electron–positron pair production* in the field of a nucleus or of an atomic electron. The corresponding cross-sections on lead are shown in Figure 4.8, where it can be seen that above a few MeV the cross-section is dominated by pair production from the nucleus.[15]

[15] The photoelectric, Compton and pair production cross-sections are roughly proportional to Z^5/E_γ, Z/E_γ and Z^2, respectively, at high energies (see Section 2.5 of Fernow, 1986). Hence, for smaller atomic numbers, the cross-sections are smaller, and the region of pair production dominance is shifted to somewhat higher energies.

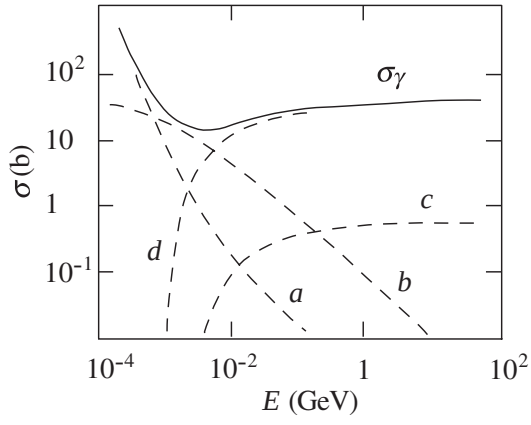


Figure 4.8 Total experimental photon cross-section σ_γ on a lead atom, together with the contributions from (a) the photoelectric effect, (b) Compton scattering, (c) pair production in the field of the atomic electrons and (d) pair production in the field of the nucleus. (Reprinted by permission of Institute of Physics (IOP), Fig. 27.14, W.-M. Yao *et al.*, *Journal of Physics*, **G33**, 1, 2006.)

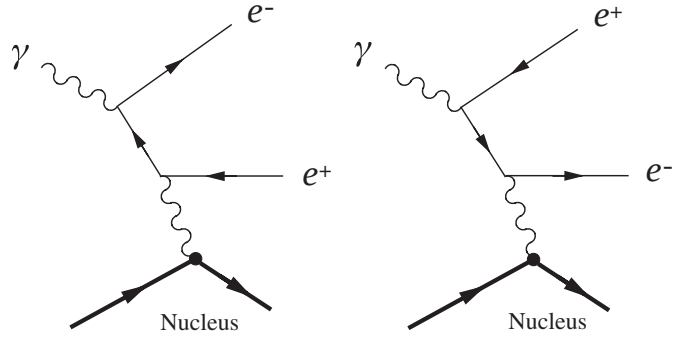


Figure 4.9 The pair production process $\gamma + (Z, A) \rightarrow e^- + e^+ + (Z, A)$.

The pair production process is closely related to electron bremsstrahlung, as can be seen by comparing the Feynman diagrams shown in Figures 4.7 and 4.9. The cross-section for pair production rises rapidly from threshold, and is given to a good approximation by

$$\sigma_{pair} = \frac{7}{9}\frac{1}{n_a L_R} \tag{4.20}$$

for $E_\gamma \gg mc^2/\alpha Z^{1/3}$, where L_R is the radiation length. Substituting these results into (4.19) gives

$$I(x) = I_0 \exp(-7x/9L_R), \tag{4.21}$$

so that at high energies, photon absorption, like electron radiation loss, is characterized by the radiation length L_R.

4.3.5 Ranges and interaction lengths

For nonrelativistic charged particles ($\beta = 1$), ionization energy losses are very high and the particles come to rest relatively quickly, as we saw in Section 4.3.4. However, as β becomes larger, the range (4.12) becomes comparable with the interaction lengths (4.6) and (4.7) and the radiation length (4.15). The situation is complicated, but for highly relativistic particles a relatively simple picture emerges for any given material. We will briefly summarize the situation here by discussing particles in order of their increasing penetration power in condensed matter, since this is an important factor in the design of experiments at high energies, as we shall see in Section 4.5.

4.3.5(a) Electrons and photons

For energies above 1 GeV, these are the least penetrating particles, whose interactions are controlled by the radiation lengths listed in Table 4.3.

4.3.5(b) Hadrons

Radiation energy losses are unimportant for charged hadrons because they are suppressed, relative to those of electrons, by the square of their mass. Their strong interactions with atomic nuclei are controlled by the interaction lengths of Table 4.1, which are typically much larger than the radiation lengths for electrons in the same material.

4.3.5(c) Muons

Again, radiation energy losses are suppressed by the square of their mass and muons have no strong interactions with nuclei. They are easily the most penetrating of all charged particles.

4.3.5(d) Neutrinos

As these only interact by the weak interaction, they can pass through enormous distances of matter without interacting, as discussed earlier.

4.4 PARTICLE DETECTORS

The detection of a particle means more than simply its localization. To be useful this must be done with a resolution sufficient to enable particles to be separated in both space and time in order to determine which are associated with a particular event. We also need to be able to identify each particle and measure its energy and momentum. No single detector is optimal with respect to all these requirements and modern particle physics experiments commonly use very large multicomponent detector systems that integrate many different sub-detectors in a single device. Such systems rely heavily on fast electronics and computers to monitor and control the subdetectors, and to coordinate, classify and record the vast amount of information flowing in from different parts of the apparatus. Such multicomponent detectors

will be considered in Section 4.5. Here we will briefly introduce some of the most important individual detectors of which they are composed. Detector development is a rapidly moving major area of research and new devices are frequently developed, so the list below is by no means exhaustive.[16]

4.4.1 Introduction

We start by discussing the large family of *gas detectors*, which convert the ionization produced by the passage of a charged particle through a gas into an electronic signal. Such detectors are primarily used to provide accurate measurements of a particle's position or, by a sequence of such measurements, a record of the particle's trajectory. In this context they are also called *tracking detectors*.

Tracking detectors are very often placed in a magnetic field, in which case they can provide a measurement of the particle's momentum from the resulting curvature of its track. An apparatus that is dedicated to measuring momentum is called a *spectrometer.* It consists of a magnet and a series of detectors to track the passage of the particles. The precise design depends on the nature of the experiment being undertaken. For example, in a fixed-target experiment at high energies, the reaction products are usually concentrated in a narrow cone about the initial beam direction, whereas in colliding beam experiments spectrometers must completely surround the interaction region to obtain full angular coverage. However, in this case the beam will also be deflected, and so at colliders so-called 'compensating magnets' are added to correct for this.

Next we discuss three more types of charged particle detectors: *solid-state detectors scintillation counters* and *Čerenkov counters*. Solid-state detectors exploit the properties of semiconductors. They are in some respects the solid-state analogue of gas detectors and have, to some extent, replaced the latter in current experiments. Scintillation counters have excellent time resolution and are sometimes used for 'triggering' other devices in multicomponent detector systems, i.e. to decide whether or not to activate other detectors or whether to record the information from a particular event. Čerenkov counters measure the velocity of a charged particle and can be used to distinguish between different particles having a given very high momentum by using their velocities to determine their masses.

All the above detectors can only detect charged particles, and all leave the nature of the particle unchanged as it passes through the detector. In contrast, the final detectors we discuss, *calorimeters,* can detect both neutral and charged particles. They totally absorb the detected particle to yield a measurement of its energy. Calorimeters can also have good spatial and time resolutions, which is particularly important for neutral particles that often cannot be detected in any other way.

4.4.2 Gas detectors

Most gas detectors detect the ionization produced by the passage of a charged particle through a gas, typically an inert one such as argon, either by collecting the ionization products or induced charges on to electrodes or (historically) by making

[16] For more detailed discussions of particle detectors see, for example, Grupen (1996) and the references in Footnote 2.

the ionization track visible in some form. The average energy needed to produce an electron–ion pair is $30 \pm 10\,\text{eV}$, with a weak dependence on the gas used and the energy of the incident particle. In practice, the output is a pulse at the anode (which is amplified by electronic means), with the bulk of the signal being due to the positive ions because of their longer drift distance. For a certain range of applied voltages – the so-called 'proportional region' (see below) – these devices are primarily used to provide accurate measurements of a particle's position. As position detectors, gas detectors largely replaced earlier detectors that used visual techniques, such as cloud chambers, bubble chambers and stacks of photographic emulsions. These detectors were discussed in Chapter 2. Although historically important, none of these visual devices are now in general use and they have all been superceded by electronic detectors.

To understand the principles of gas detectors we refer to Figure 4.10, which shows the number of ion pairs produced per incident charged particle (the *gas amplification*

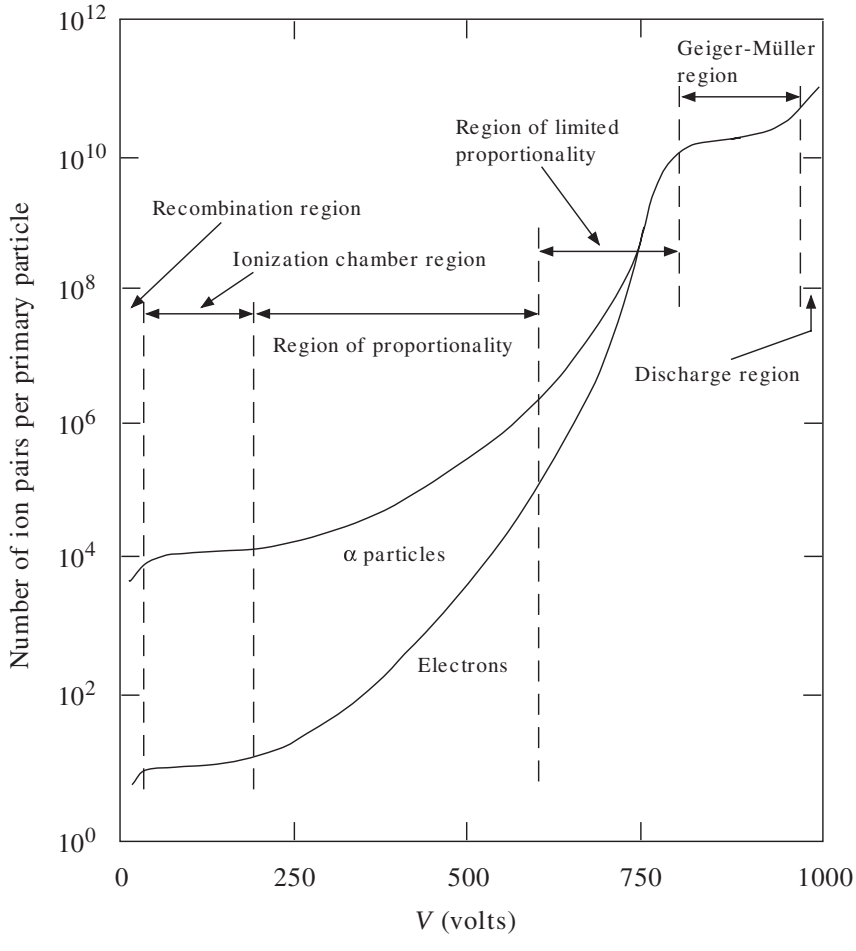


Figure 4.10 Gas amplification factor as a function of voltage V applied in a single-wire gas detector, with a wire radius typically $20\,\mu\text{m}$, for a strongly ionizing particle (α) and a weakly ionizing particle (electron).

factor) as a function of the applied voltage V for two cases: a heavily ionizing particle (e.g. an alpha particle – upper curve) and a lightly ionizing particle (e.g. an electron – lower curve).

4.4.2(a) Ionization chamber

At low applied voltages, the output signal is very small because electron–ion pairs recombine before reaching the electrodes, but as the voltage increases the number of pairs increases to a saturation level representing complete collection. This is the region of the *ionization chamber*. The simplest type of chamber is a parallel plate condenser filled with an inert gas and having an electric field $E = V/d$, where d is the distance between the plates. In practice the gas mixture must contain at least one 'quenching' component that absorbs ultraviolet light and stops a plasma forming and spreading throughout the gas.

Another arrangement is cylindrical with an inner anode of radius r_a and an outer cathode of radius r_c, giving an electric field

$$E(r) = \frac{V}{r \ln(r_c/r_a)} \tag{4.22}$$

at a radial distance r from the centre of the anode wire. The output signal is proportional to the number of ions formed and hence the energy deposited by the radiation, but is independent of the applied voltage. However, the signal is very small compared to the noise of all but the slowest electronic circuits and requires considerable amplification to be useful. Overall, the energy resolution and the time resolution of the chamber are relatively poor, and so ionization chambers are of very limited use in recording individual pulses. They are used, for example, as beam monitors, where the particle flux is very large.

4.4.2(b) Wire chambers

If the voltage is increased beyond the region of operation of the ionization chamber, we move into the *proportional region*. In this region, a cylindrical arrangement as used in the ionization chamber will produce electric field strengths or order $(10^4 - 10^5)\,\mathrm{V/cm}$ near the wire, which is strong enough for electron–ion pairs released in the primary ionization to gain sufficient energy to cause secondary ionization. The rapid increase in amplification due to secondary ionization is called a *Townsend avalanche*. The output signal at the anode is still proportional to the energy lost by the original particle. There are a number of different types of device working in the proportional region and they are sometimes generically referred to as *track chambers* or simply *wire chambers*.

The earliest detector using this idea was the *proportional counter*, which consists of a cylindrical tube filled with gas (again a quenching component in the gas is

required) and maintained at a negative potential, and a fine central anode wire at a positive potential. Subsequently, the resolution of proportional counters was greatly improved as a result of the discovery that if many anode wires were arranged in a plane between a common pair of cathode plates, each wire acts as an independent detector. This device is called a *multiwire proportional chamber* (MWPC) and was introduced in 1968. An MWPC can achieve spatial resolutions of $50 - 200\,\mu$m and has a typical time resolution of about 2 ns.

A schematic diagram of an MWPC is shown in Figure 4.11. The planes (a) have anode wires into the page and those in plane (b) are at right angles. The wire spacing is typically 2 mm. The cathodes are the faces of the chambers. A positive voltage applied to the anode wires generates a field as shown in the upper corner. A particle crossing the chamber ionizes the gas and the electrons drift along the field lines to the anode wires. In this particular example, there would be signals from one wire in the upper (a) chamber and two in the lower (a) chamber.

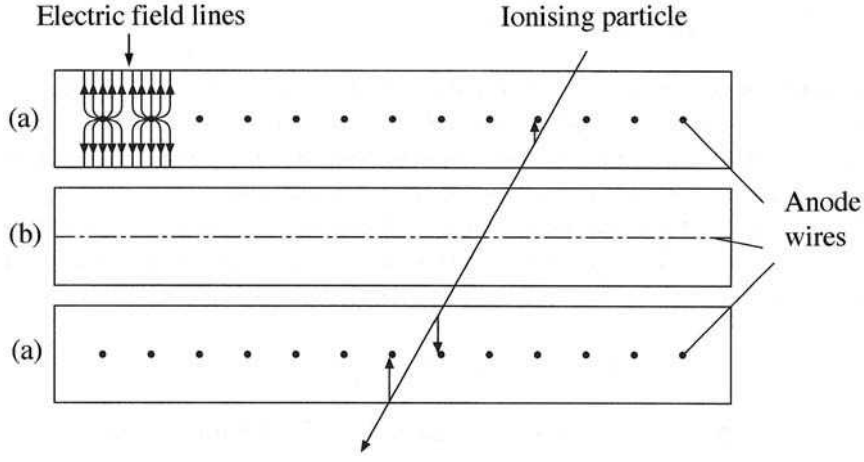


Figure 4.11 A group of three planes of a MWPC (see text for details). (From *Particles and Nuclei*, Povh, Rith, Scholz and Zetsche, Fig. A7, 1999. With kind permission of Springer Science and Business Media.)

A related device is the *drift chamber*, which has largely replaced the MWPC as a general detector. This uses the fact that the liberated electrons take time to drift from their point of production to the anode. Thus the time delay between the passage of a charged particle through the chamber and the creation of a pulse at the anode is related to the distance between the particle trajectory and the anode wire. In practice, additional wires are incorporated to provide a relatively constant electric field in each cell in a direction transverse to normal incidence. A reference time has to be defined, which, for example, could be done by allowing the particle to pass through a scintillator positioned elsewhere in the experiment. (Scintillation counters are discussed in Section 4.4.4 below.) The electrons drift for a time and are then collected at the anode, thus providing a signal that the particle has passed. If the drift time can be measured accurately (to within a few ns) and if the drift velocity is known, then spatial and temporal resolutions similar to those of the MWPC can easily be achieved.

Drift chambers are constructed in a variety of geometries to suit the nature of the experiment, and arrangements where the wires are in planar, radial or cylindrical configurations have all been used. The latter type is also called a 'jet chamber'. Such a chamber was used by the JADE collaboration at an e^+e^- collider at DESY, Hamburg. It was a cylindrical array of drift chambers with the beam direction as the axis and the collision region, at which the e^+e^- interactions occurred, at the centre. The anode wires ran parallel to the axis, and the whole detector was divided into 24 segments, with 64 anode wires in each. Figure 4.12 shows an end view along the beam direction of a computer reconstruction of an event observed in the chamber in which the e^+ and e^- annihilate to produce two 'jets' of hadrons. The solid lines indicate the reconstructed charged particle trajectories taking into account the known magnetic field, which is also parallel to the beam direction; the dotted lines indicate the reconstructed trajectories of neutral particles, which were detected outside the chamber by other means.

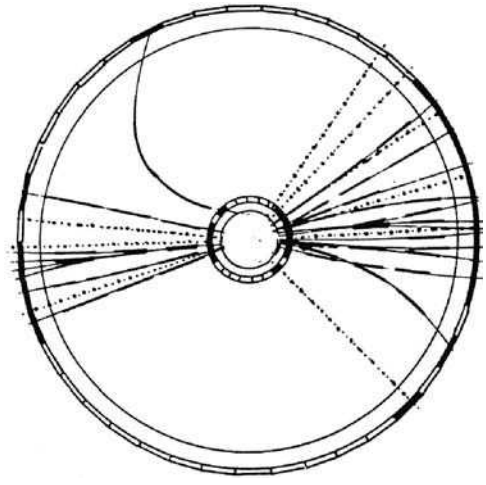


Figure 4.12 Computer reconstruction of a typical 'two-jet' event observed in the JADE chamber. The solid lines indicate the trajectories of charged particles and the dotted lines those of neutral particles, which were detected outside the chamber by other means. (After Orito, 1979, Fermilab. Courtesy of Brookhaven National Laboratory.)

One of the most advanced applications of proportional and drift chamber principles is embodied in the *time projection chamber* (TPC) illustrated schematically in Figure 4.13. This device consists of a cylindrical barrel, typically 2 m long and 1 m in diameter, surrounding the beam pipe of a collider. At each end of the chamber is a segmented layer of proportional counters. The electric drift field **E**, due to a negative high-voltage electrode plane at the centre of the chamber, and a strong magnetic field **B** are aligned parallel and antiparallel to each other in the two sections of the chamber with reference to the axis of the cylinder. Electrons formed along the track of an ionizing particle emerging from the interaction point at the centre of the barrel drift under the action of the electric field towards one of the endcaps along helical trajectories whose direction is parallel to the axis of the barrel. Their locations are

measured by a set of anode wires located between rectangular cathodes in the endcaps. The remaining third coordinate necessary to reconstruct the position of a point on the track is found from the time it takes for the electrons to drift from the point of production to the endcaps where they are detected. The TPC has excellent spatial resolution and has been used in e^+e^- annihilation experiments. A TPC is also at the heart of the STAR detector at the RHIC as will be illustrated in Section 4.5.

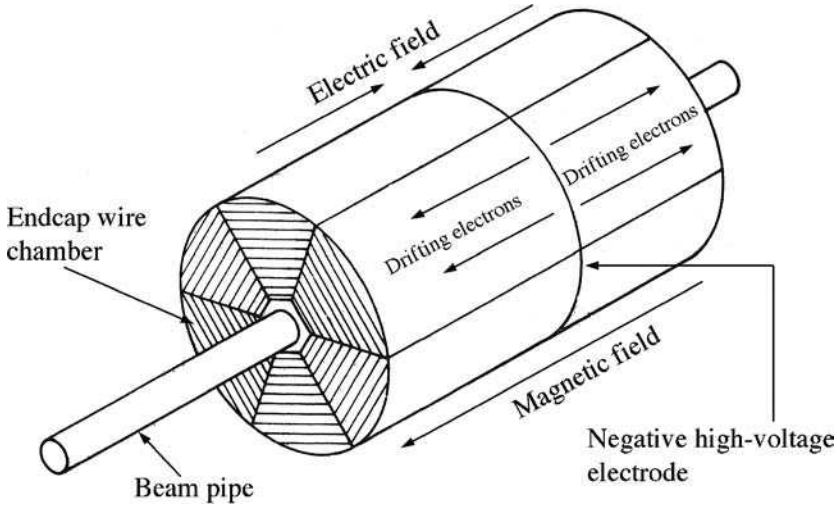


Figure 4.13 Schematic diagram of a time projection chamber. (From Konrad Kleinknecht, *Detectors for Particle Radiation,* 1986. With permission from Cambridge University Press.)

A more robust form of chamber has also evolved, in which the wires are replaced by conductive metal strips on a printed circuit board. This is called a *microstrip gas chamber* (MSGC) and is incorporated in experiments at modern large accelerators.

4.4.2(c) Beyond the region of proportionality

Referring again to Figure 4.10, by increasing the external voltage still further one moves into a region where the output signal ceases to be proportional to the number of ion pairs produced and hence the incident energy. This is the region of *limited proportionality*. In this region a type of gas detector called a streamer tube operates, but this will not be discussed here. Eventually the process runs out of control and we enter the *Geiger–Müller region* where the output signal is independent of the energy lost by the incident particle. In this region a quenching agent is not used. Detectors working in this region are called Geiger–Müller counters. Physically they are similar to the simple cylindrical proportional counter and are widely used as portable radiation monitors in the context of health physics.

For completeness, we can mention that if the gas amplification factor is taken beyond the Geiger–Müller region, the avalanche develops moving plasmas or streamers. Recombination of ions then leads to visible light that can be made to generate an

electrical output. Eventually complete breakdown occurs and a spark is emitted as the incident particle traverses the gas. Detectors in this region, called streamer and spark chambers (these were of parallel plate construction, rather than cylindrical), were widely used in the 1970s and 1980s and played an important role in hadron physics, but are no longer in general use.

4.4.3 Semiconductor detectors

Solid-state detectors operate through the promotion of electrons from the valence band of a solid to the conduction band as a result of the entry of the incident particle into the solid. The resulting absence of an electron in the valence band (a 'hole') behaves like a positron. Semiconductor detectors are essentially solid-state ionization chambers with the electron–hole pairs playing the role of electron–ion pairs in gas detectors, with the number of pairs being proportional to the energy loss. In the presence of an externally applied electric field, the electrons and holes separate and collect at the electrodes, giving a signal proportional to the energy loss of the incident charged particle.

Most semiconductors are manufactured from silicon (although the use of germanium is being explored) and use the principle of the junction diode. The band gap in some solids is as small as 1 eV and the energy loss required to produce a pair is only 3–4 eV on average, compared to the 30 eV ionization energy required in a gas detector, or the approximately 300 eV to extract an electron from a photocathode coupled to a plastic scintillator (to be discussed in Section 4.4.4). Thus a very large number of electron–hole pairs with only a small statistical fluctuation will be produced by a low-energy particle. Solid-state detectors are therefore very useful in detecting low-energy particles. Semiconductors are used as a compromise between materials that have residual conductivity sufficient to enable conduction pulses due to single particles to be distinguished above background and those in which the charge carriers are not rapidly trapped in impurities in the material. Such detectors have excellent energy resolution and linearity coupled with a fast response time and have long been used in nuclear physics, but thin planar detectors only became important in particle physics at a later date, because of the expense of covering large areas. Nevertheless, more than 200 square metres of semiconductor detector are being used in experiments at the LHC.

One example of a solid-state detector is a *silicon microstrip detector*, where narrow strips of active detector are etched on to a thin slice of silicon, with gaps of order l0 μm, to give a tiny analogue of an MWPC. Arrays of such strips can then be used to form detectors with resolutions of order 10 μm. These are often placed close to the interaction vertex in a colliding beam experiment, with a view to studying events involving the decay of very short-lived particles. Another example is the *pixel detector*. A single-plane strip detector only gives position information in one dimension (orthogonal to the strip). A pixel detector improves on this by giving information in two dimensions from a single plane. Solid-state 'vertex detectors' have become increasingly important in particle physics and have been incorporated in most of the multicomponent detectors used at the largest colliders. Their main advantage is their superb spatial resolution; a disadvantage is their limited ability to withstand radiation damage.

4.4.4 Scintillation counters

For charged particles we have seen that energy losses occur due to excitation and ionization of atomic electrons in the medium of the detector. In suitable materials, called *scintillators*, a small fraction of the excitation energy re-emerges as visible light (or sometimes in the ultraviolet (UV) region) during de-excitation. In a scintillation counter this light passes down the scintillator and on to the face of a *photodetector* – a device that converts a weak photon signal to a detectable electric impulse. Short-wavelength light may initially be collected on to the material of a *wavelength shifter*, which is a device that shifts the wavelength to higher values that are better matched to the frequency sensitivity of the photodetector.

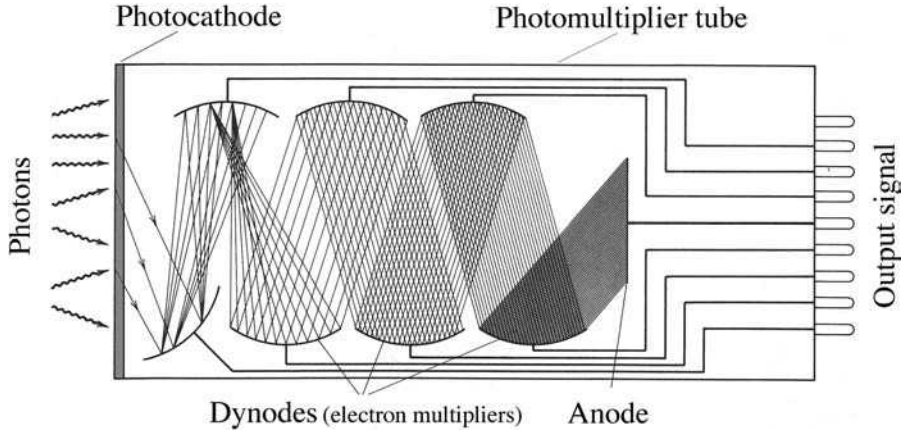


Figure 4.14 Schematic diagram of the main elements of a photomultiplier tube. (From Krane, *Introductory Nuclear Physics,* copyright 1988. Reprinted with permission of John Wiley & Sons, Inc.)

An important example of a photodetector is the *photomultiplier tube*, a schematic diagram of which is shown in Figure 4.14. Electrons are emitted from the cathode of the photomultiplier by the photoelectric effect and strike a series of focusing dynodes. These amplify the electrons by secondary emission at each dynode and accelerate the particles to the next stage. The final signal is extracted from the anode at the end of the tube. The electronic pulse can be as short as 100 ps if the scintillator has a short decay time. The scintillation counter is thus an ideal timing device and is widely used for 'triggering' other detectors; i.e. its signal is used to decide whether or not to record information from the event. They can also be used to distinguish between different types of particle with a common momentum by measuring the 'time-of-flight', and hence the velocity, between two scintillation counters a known distance apart. However, the method is limited to relatively low-energy particles because all high-energy particles have velocities close to the speed of light.

Commonly used scintillators are inorganic single crystals (e.g. cesium iodide) or organic liquids and plastics, with the latter two generally used in particle physics.

Some complex detector systems have used several tons of detector in combination with thousands of photomultiplier tubes. The robust and simple nature of the scintillation counter, together with its low cost and the ease with which plastic scintillator can be fabricated, has made it a mainstay of experimental particle physics since the earliest days of the subject.

4.4.5 Čerenkov counters

An important identification method for high-energy particles is based on the Čerenkov effect. When a charged particle with velocity υ traverses a dispersive medium of refractive index n, excited atoms in the vicinity of the particle become polarized, and if υ is greater than the speed of light in the medium c/n, a part of the excitation energy reappears as coherent radiation emitted at a characteristic angle θ to the direction of motion. The necessary condition $\upsilon > c/n$ implies $\beta n > 1$. The angle is found by using Huygen's construction, as used in optics, as illustrated in Figure 4.15. In Figure 4.15(a), the points 1, 2 and 3 on the axis represent three successive positions of the particle at equally spaced times t_1, t_2 and t_3; the curves labelled 2 and 3 represent the positions of the wavelets emitted from those points at times t_2 and t_3. The resulting wavefront is indicated by the dashed line and moves in the direction of the arrows. Figure 4.15(b) shows the triangle 13p, where $t = t_3 - t_1$ and the Čerenkov radiation is emitted at an angle θ. It follows that $\cos\theta = 1/\beta n$ for the angle θ, where $\beta = \upsilon/c$ as usual. A determination of θ is thus a direct measurement of the velocity.

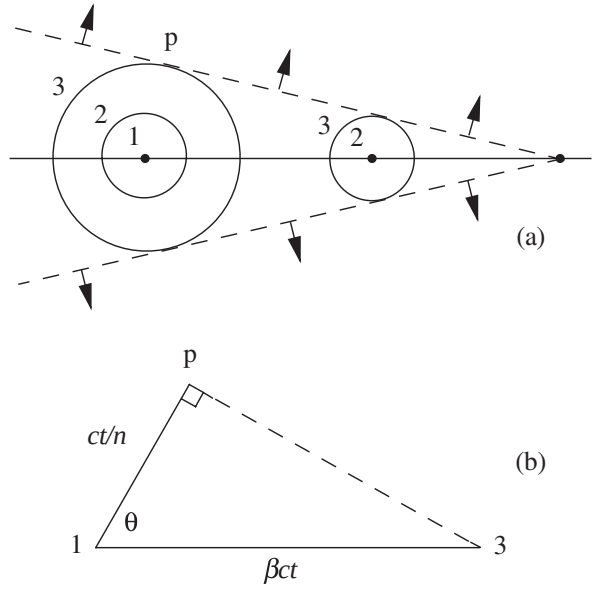


Figure 4.15 Hyugen's construction for the Čerenkov radiation emitted by a particle travelling with speed υ greater than c/n, the speed of light in the medium.

Čerenkov radiation appears as a continuous spectrum and may be collected on to a photosensitive detector. (A wavelength shifter may also be used for the same reason as in scintillation counters.) Its main limitation from the point of view of particle detection is that very few photons are produced. The number of photons $N(\lambda)\,\mathrm{d}\lambda$ radiated per unit path length in a wavelength interval $\mathrm{d}\lambda$ can be shown to be

$$N(\lambda)\,\mathrm{d}\lambda = 2\pi\alpha\left(1-\frac{1}{\beta^2 n^2}\right)\frac{\mathrm{d}\lambda}{\lambda^2} < 2\pi\alpha\left(1-\frac{1}{n^2}\right)\frac{\mathrm{d}\lambda}{\lambda^2} \tag{4.23}$$

and so vanishes rapidly as the refractive index approaches unity. The maximum value occurs for $\beta = 1$, and is given for visible Čerenkov radiation in various media in Table 4.4. These numbers should be compared to the 10^4 photons/cm emitted by a typical scintillator. Because the yield is so small, appreciable lengths are needed to give enough photons, and gas Čerenkov counters in fixed-target experiments can be several metres long.

TABLE 4.4 The refractive index n and the threshold value of γ_{th} for some commonly used Čerenkov radiators, together with the number of photons/cm emitted in the visible region 300–700 nm by a particle with unit charge and $\beta \approx 1$.

Medium	$n-1$	γ_{th}	Photons/cm
Helium (STP)	3.5×10^{-5}	120	0.03
CO_2(STP)	4.1×10^{-4}	35	0.40
Silica aerogel	0.025 – 0.075	4.6 – 2.7	24 – 66
Water	0.33	1.52	213
Glass	0.46 – 0.75	1.37 – 1.22	261 – 331

Čerenkov counters are used in two different modes. The first is as a *threshold counter* to detect the presence of particles whose velocities exceed some minimum value. Suppose that two particles with β values β_1 and β_2 at some given momentum p are to be distinguished. If a medium can be found such that $\beta_1 n > 1 \geq \beta_2 n$, then particle 1 will produce Čerenkov radiation but particle 2 will not. Clearly, to distinguish between highly relativistic particles with $\gamma >> 1$ also requires $n \approx 1$, so that from (4.23) very few photons are produced. Nevertheless, common charged particles can be distinguished in this way up to at least 30 GeV/c.

Another device is the so-called *ring-image* Čerenkov detector and is a very important device at both fixed-target machines and colliders. If we assume that the particles are not all travelling parallel to a fixed axis, then the radiating medium can be contained within two concentric spherical surfaces of radii R and $2R$ centred on the target or interaction region where the particles are produced, as illustrated in Figure 4.16. The outer surface is lined with a mirror, which focuses the Čerenkov radiation into a ring at the inner detector surface. The radius of this ring depends on the angle θ at which the Čerenkov radiation is emitted, and hence on the particle velocity υ. It is determined by constructing an image of the ring electronically. This was the technique used in the SuperKamiokande detector discussed in Chapter 2 to detect relativistic electrons and muons produced by neutrino interactions. In that experiment the radiating medium was pure water.

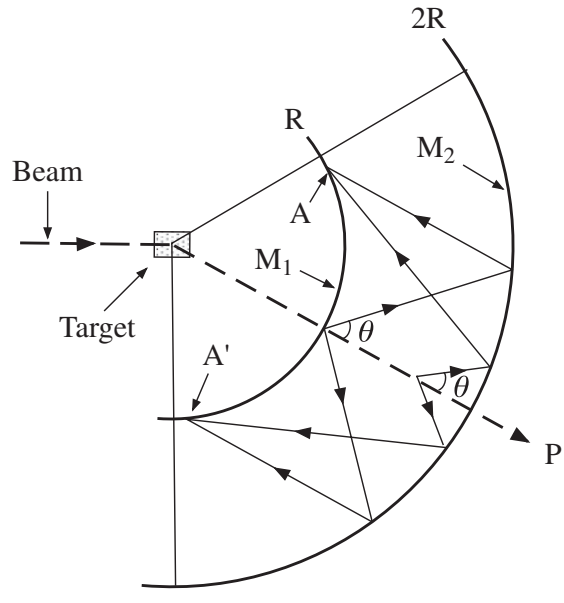


Figure 4.16 A particle P, produced from the target, emits Čerenkov radiation at an angle θ on traversing a medium contained between two spheres of radius R and $2R$. The mirror M_2 on the outer sphere focuses the radiation into a ring image at A and A′ on the inner detector sphere M_1. The radius of the ring image depend on the angle of emission of the Čerenkov radiation and hence on the velocities of the particles.

4.4.6 Calorimeters

Calorimeters are an important class of detector used for measuring the energy and position of a particle by its total absorption. They differ from most other detectors in that the nature of the particle is changed by the detector and by the fact that they can detect neutral as well as charged particles. A calorimeter may be a homogeneous absorber/detector (e.g. a scintillator such as CsI). Alternatively, it can be a sandwich construction with separate layers of absorber (e.g. a metal such as lead) and detector (scintillator, MWPC, etc.). The latter are also known as 'sampling calorimeters'. During the absorption process, the particle will interact with the material of the absorber, generating secondary particles, which will themselves generate further particles and so on, so that a cascade or shower develops. For this reason calorimeters are also called 'shower counters'. The shower is predominantly in the longitudinal direction due to momentum conservation, but will be subject to some transverse spreading due both to multiple Coulomb scattering and the transverse momentum of the produced particles. Eventually all, or almost all, of the primary energy is deposited in the calorimeter, and gives a signal in the detector part of the device.

There are several reasons why calorimeters are important, especially at high energies:

1. They can detect neutral particles, by detecting the charged secondaries.
2. The absorption process is statistical and governed by the Poisson distribution, so that the relative precision of energy measurements $\Delta E/E$ varies like $E^{-1/2}$ for

large E. This is a great improvement on high-energy spectrometers where $\Delta E/E$ varies like E^2.
3. The signal produced can be very fast, of order 10–100 ns, and is ideal for making triggering decisions.

Since the characteristics of electromagnetic and hadronic showers are somewhat different it is convenient to describe each separately. In practice, it is common to have both types in one experiment with the hadron calorimeter stacked behind the electromagnetic one, although it is possible to design calorimeters that detect both classes of particle.

4.4.6(a) Electromagnetic showers

When a high-energy electron or positron interacts with matter we have seen that the dominant energy loss is due to bremsstrahlung, and for the photons produced the dominant absorption process is pair production. Thus the initial electron will, via these two processes, lead to a cascade of $e^{\pm}$ pairs and photons, and this will continue until the energies of the secondary electrons fall below the critical energy E_C where ionization losses equal those from bremsstrahlung. This energy is roughly given by $E_C \approx 600\,\text{MeV/Z}$.

Most of the correct qualitative features of shower development may be obtained from the following very simple model. We assume:

1. Each electron with $E > E_C$ travels one radiation length and then gives up half of its energy to a bremsstrahlung photon.
2. Each photon with $E > E_C$ travels one radiation length and then creates an electron–positron pair with each particle having half the energy of the photon.
3. Electrons with $E < E_C$ cease to radiate and lose the rest of their energy by collisions.
4. Ionization losses are negligible for $E > E_C$.

A schematic diagram of the approximate development of a shower in an electromagnetic calorimeter assuming this simple model is shown in Figure 4.17.

If the initial electron has energy $E_0 >> E_C$, then after t radiation lengths the shower will contain 2^t particles, which consist of approximately equal numbers of electrons, positrons and photons, each with an average energy

$$E(t) = E_0/2^t. \tag{4.24}$$

The multiplication process will cease abruptly when $E(t) = E_C$, i.e. at $t = t_{max}$ where

$$t_{max} = t\,(E_C) \equiv \frac{\ln\,(E_0/E_C)}{\ln 2}, \tag{4.25}$$

and the number of particles at this point will be

$$N_{max} = \exp\,(t_{max}\,\ln\,2) = E_0\,\big/E_C\,. \tag{4.26}$$

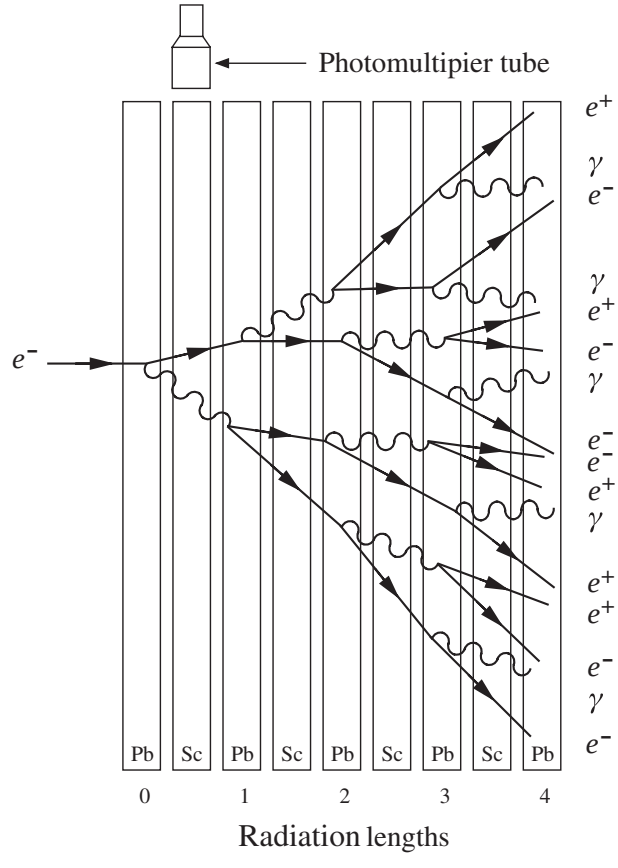


Figure 4.17 Approximate development of an electromagnetic shower in a sampling calorimeter assuming the simple model of the text. The calorimeter consists of alternate layers of lead (Pb) and a scintillator (Sc), the latter attached to photomultipliers (one only shown).

The main features of this simple model are observed experimentally, and in particular the maximum shower depth increases only logarithmically with primary energy. Because of this, the physical sizes of calorimeters need increase only slowly with the maximum energies of the particles to be detected. The energy resolution of a calorimeter, however, depends on statistical fluctuations, which are neglected in this simple model. For an electromagnetic calorimeter it is typically $\Delta E/E \approx 0.05/E^{1/2}$, where E is measured in GeV.

4.4.6(b) Hadronic showers

Although hadronic showers are qualitatively similar to electromagnetic ones, shower development is far more complex because many different processes contribute

to the inelastic production of secondary hadrons. The scale of the shower is determined by the nuclear absorption length defined earlier. Since this absorption length is larger than the radiation length, which controls the scale of electromagnetic showers, hadron calorimeters are thicker devices than electromagnetic ones. Another difference is that some of the contributions to the total absorption may not give rise to an observable signal in the detector. Examples are nuclear excitation and leakage of secondary muons and neutrinos from the calorimeter. The loss of 'visible' or measured energy for hadrons is typically 20–30 % greater than for electrons.

The energy resolution of calorimeters is in general much worse for hadrons than for electrons and photons because of the greater fluctuations in the development of the hadron shower. Depending on the proportion of π^0's produced in the early stages of the cascade, the shower may develop predominantly as an electromagnetic one because of the decay $\pi^0 \to \gamma\,\gamma$. These various features lead to an energy resolution typically a factor of 5–10 poorer than in electromagnetic calorimeters.

4.5 DETECTOR SYSTEMS AND EXPERIMENTS

In this final section, we briefly illustrate how the basic detectors we have discussed are combined into multicomponent systems and used in collider experiments. We take as our examples the historic discovery of the $W^{\pm}$ and Z^0 bosons at CERN in 1983, and the CDF, ATLAS and STAR detectors, which are among those currently in operation. A further historic collider experiment, the discovery of the top quark at FNAL in 1995, is discussed in some detail in Section 8.2.3.

Before proceeding, however, we should emphasize that while most modern experiments exploit the high centre-of-mass energies available at colliders, these are not the only important experiments. Fixed-target and nonaccelerator experiments have also played, and continue to play, an extremely important role in the development of particle physics. Several such experiments, both historical and modern, have been discussed in Chapters 2 and 3, and several more will be covered later. The latter include the inelastic scattering of neutrinos by nucleons (to be discussed in Section 7.5) and the detection of *CP* violation in the decays of neutral mesons (to be discussed in Chapter 10).

4.5.1 Discovery of the $W^{\pm}$ and Z^0 bosons

The $W^{\pm}$ and Z^0 bosons are highly unstable particles, which were first produced in the reactions

$$p + \bar{p} \to W^+ + X^-, \tag{4.27a}$$

$$p + \bar{p} \to W^- + X^+ \tag{4.27b}$$

and

$$p + \bar{p} \to Z^0 + X^0, \tag{4.27c}$$

where $X^{\pm}$ and X^0 are arbitrary hadronic states allowed by the conservation laws. The heavy bosons were detected via their subsequent decays

$$W^+ \to \ell^+ + \nu_\ell, \tag{4.28a}$$

$$W^- \to \ell^- + \bar{\nu}_\ell \tag{4.28b}$$

and

$$Z^0 \to \ell^+ + \ell^-, \tag{4.28c}$$

where the charged leptons $\ell^{\pm}$ were either muons or electrons. (These are not the only possible decay modes, as we shall see later.) The production mechanism that gives rise to the reactions (4.27) is shown in Figure 4.18 and involves quark–antiquark annihilation processes such as[17]

$$u + \bar{d} \to W^+, \quad d + \bar{u} \to W^- \tag{4.29a}$$

and

$$u + \bar{u} \to Z^0, \quad d + \bar{d} \to Z^0 \tag{4.29b}$$

for charged and neutral vector bosons, respectively. In the centre-of-mass of the $q\bar{q}$ system, the total energy of the $q\bar{q}$ pair must be at least 80 or 91 GeV, corresponding to the production of either a $W^{\pm}$ or a Z^0 boson at rest. However, the energy of the $p\bar{p}$ system has to be considerably higher for a reasonable reaction rate to occur, because each quark (or antiquark) has only a fraction of the parent proton (or antiproton) energy.

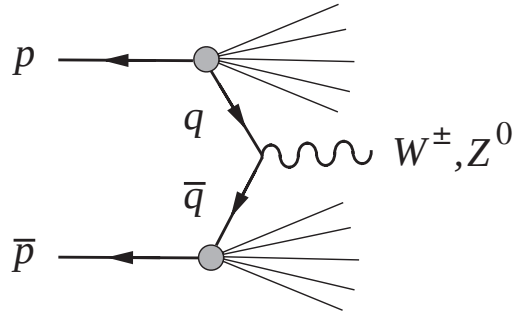


Figure 4.18 Mechanism for producing $W^{\pm}$ and Z^0 bosons in proton–antiproton annihilations.

The $W^{\pm}$ and Z^0 bosons were first discovered using a proton–antiproton collider at CERN, which was specifically built for this purpose. At the time it had proton and

[17] The interactions of the $W^{\pm}$ and Z^0 bosons with leptons and quarks will be discussed more systematically in Chapters 8 and 9.

antiproton beams with maximum energies of 270 GeV each, giving a total centre-of-mass energy of 540 GeV. Two independent experiments were mounted, called UA1 and UA2.[18] In each case the proton and antiproton beams were brought together in an intersection region that lay at the centre of a very large and elaborate detector system. We will not describe this in detail, but simply list the main components, which are shown schematically in Figure 4.19 for the case of the UA1 experiment. The order of the detectors reflects the penetrating power of the various types of particles discussed in Section 4.3.5. As one moves out from the centre there are:

1. a central tracking detector used to observe charged particles and to measure their momentum from the curvature of the tracks in an applied magnetic field;
2. a set of electromagnetic shower counters which absorb and detect both electrons, which are also observed in the central detector, and photons, which are not;
3. a set of hadron calorimeters, which absorb and detect both neutral and charged hadrons;
4. a set of counters to identify muons, which are the only charged particles to penetrate the hadron calorimeters.

Thus only neutrinos escape detection. An example of a reconstructed event from the UA1 detector is shown in Figure 4.20. We shall see below that it is in fact a Z^0 production event.

One of the main problems facing the experimenters was that for each event in which a $W^{\pm}$ or Z^0 is produced and decays to leptons, there are more than 10^7 events in which hadrons alone are produced. The extraction of a 'signal' in the presence of such a large

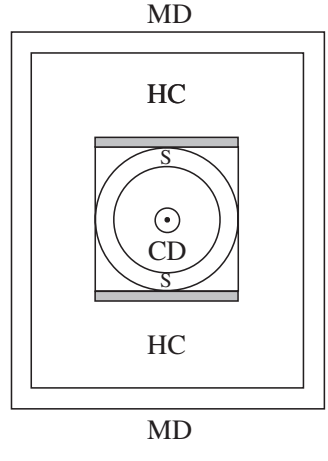


Figure 4.19 Schematic diagram showing a cross-section of the UA1 detector seen along the beam direction. The beams collide at the central point that is surrounded by a central tracking detector CD, shower counters S, hadron calorimeters HC and muon detectors MD. The shaded areas indicate the coils used to produce the magnetic field in the central detector.

[18] The accelerator lies below ground level; UA1 and UA2 stand for Underground Areas 1 and 2, where the experiments were performed.

'background' is only possible because the leptons arising from the decays (4.28) have large momenta (because the $W^{\pm}$ or Z^0 bosons are very heavy) and are often emitted at wide angles to the initial beam directions. In other words, the leptons often have large 'transverse momenta'. In contrast, the hadrons produced in proton–antiproton collisions, and the leptons that arise from their decays, very rarely have such large transverse momenta. Hence the background of purely hadronic reactions could almost be entirely eliminated by retaining only those events containing at least one electron or muon with a large transverse momentum.[19] This subset of events was then examined to see if any corresponded to $W^{\pm}$ or Z^0 production followed by the decays (4.28).

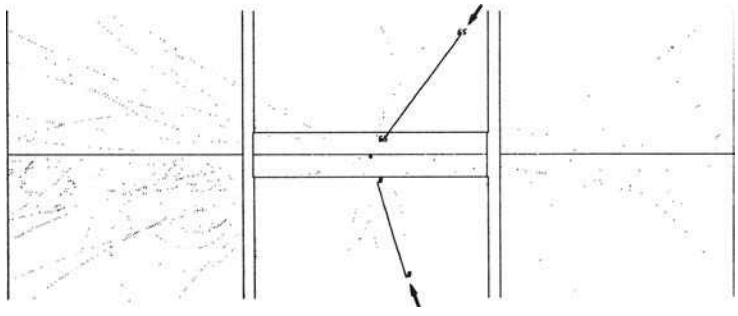


Figure 4.20 Computer reconstruction of a Z^0 production event observed in the UA1 detector. The dotted lines correspond to hadrons produced together with the Z^0 and the solid lines to the electron and positron produced in its decay. (B. Sadoulet, *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies,* Cornell University.)

4.5.1(a) Z^0 events

A typical event in which a Z^0 was produced and decayed to give an e^+e^- pair is shown in Figure 4.20. The electron and positron were both detected in the central detector, where they gave rise to tracks that were almost straight, despite the applied magnetic field, because they had very large momenta. They were identified as electrons, and not muons or hadrons, from their signal in the electromagnetic shower counters that surrounded the central detector. The e^+e^- pair stands out dramatically in the output from them, as shown in Figure 4.21.

Finally, if the observed e^+e^- pair indeed arose from the decay of a Z^0, then its invariant mass should correspond to the mass of the decaying particle, which was predicted in the electroweak theory[20] to be about $90\,\mathrm{MeV/c^2}$. For the event shown, this is indeed the case within experimental errors. It was also confirmed for the other twelve events observed by the UA1 and UA2 collaborations in their first experiments. Subsequently, the experiment has been repeated with much higher statistics, and in Figure 4.22 we show the e^+e^- and $\mu^+\mu^-$ mass distributions obtained in an experiment at Fermilab. A clear peak is seen in both cases, corresponding to a mass of

[19] For example, in the UA1 experiment this lepton was required to be emitted at more than 5° to the beam direction and to have a transverse momentum in excess of 10 GeV/c.
[20] This prediction is discussed in Chapter 9.

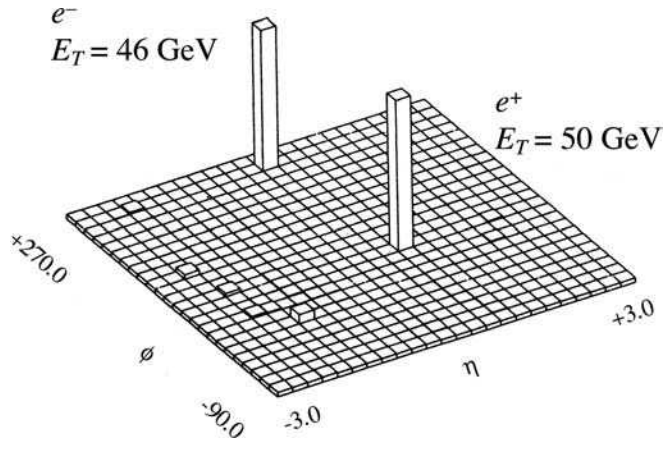


Figure 4.21 Output from the electromagnetic shower counters corresponding to the Z^0 event shown in Figure 4.20. Each square element corresponds to a particular direction in space and the height of the 'tower' to the energy of the particle detected. (B. Sadoulet, *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies*, Cornell University.)

$90.9 \pm 0.4\,\mathrm{GeV/c^2}$ and a width of order 3 GeV. These parameters have since been measured much more accurately by other methods,[21] leading to a mass

$$M_Z = 91.1876 \pm 0.0021\,\mathrm{GeV/c^2} \tag{4.30a}$$

and a width

$$\Gamma_Z = 2.4952 \pm 0.0023\,\mathrm{GeV}, \tag{4.30b}$$

corresponding to a lifetime of 2.6×10^{-25} s. The branching ratios to e^+e^- and $\mu^+\mu^-$ are both approximately 3 %.

4.5.1(b) $W^\pm$ events

In this case, the $W^\pm$bosons decay by (4.28a) and (4.28b) to give a charged lepton plus an unseen neutrino. Thus a typical $W^\pm$ event looks just like a typical Z^0 event (e.g. that shown in Figure 4.20) except that one of the electron or muon tracks is replaced by an invisible neutrino. Because the neutrino is expected to have a large transverse momentum, its presence is revealed by summing the transverse momenta of all the observed particles. If the sum is not zero, as required by momentum conservation, the 'missing transverse momentum' can only be due to a neutrino, as this is the only particle that is not detected. In summary, an event in which a $W^\pm$ is created and then decays via (4.28a) and (4.28b) is characterized by a charged lepton

[21] See Chapters 8 and 9.

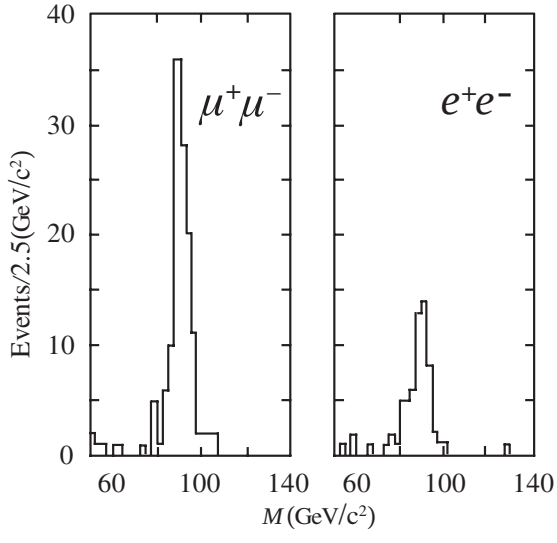


Figure 4.22 Z^0 peaks observed in the e^+e^- and $\mu^+\mu^-$ mass distributions from a Fermilab experiment by the CDF collaboration. (M. Campbell, *Proceedings of the 1989 International Symposium on Lepton and Photon Interactions at High Energies*, Stanford University.)

with a large transverse momentum,[22] and overall there is a large missing transverse momentum.

In all, a total of 43 events were observed in the original UA1 experiment, where the charged lepton was either an electron or a positron. If the $W^{\pm}$ were produced with zero transverse momentum, then momentum conservation would imply that the transverse momentum of the $e^{\pm}$ produced in the decay would be precisely equal and opposite to the missing transverse momentum ascribed to the neutrino. Since the lepton masses are negligible, this in turn implies that the corresponding transverse energies $E_T = (p_T{}^2c^2 + m^2c^4)^{1/2} \approx p_T c$ are approximately equal, and this is approximately true for the observed events, as shown in Figure 4.23. This result had been predicted theoretically, and the observed events conformed in every way to what was expected for the production and decay of a $W^{\pm}$ boson with a mass of approximately $80\,\mathrm{GeV/c^2}$. We shall not discuss this further because the analysis is rather complicated, but merely note that the results have been confirmed by later experiments with higher statistics. These yield a mass value of

$$M_W = 80.403 \pm 0.029\,\mathrm{GeV/c^2} \tag{4.31a}$$

and a decay width

$$\Gamma_W = 2.141 \pm 0.041\,\mathrm{GeV}, \tag{4.31b}$$

[22] Occasionally a high-momentum $e^{\pm}$ or $e^{\pm}$ from the decay of a $W^{\pm}$ will be emitted close to the beam direction. Such events are lost in the background and escape detection.

which corresponds to a lifetime of 3.2×10^{-25} s. The branching ratio for each of the decay modes ev_e and $\mu\nu_\mu$ is approximately 10 %.

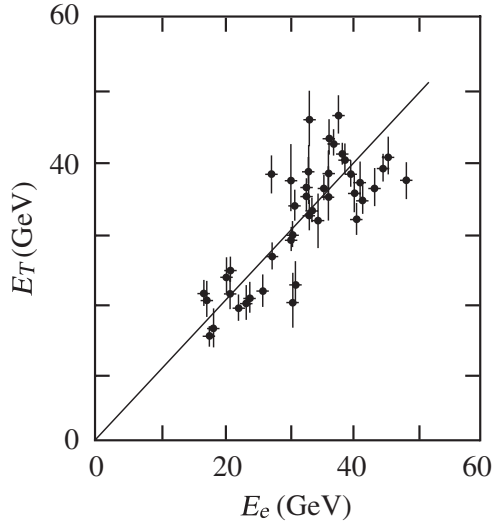


Figure 4.23 The 'missing transverse energy' E_T plotted as a function of the measured transverse energy E_e of the electron for the first $W^{\pm}$ production events observed in the UA1 experiment. (M. Spiro, *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies*, Cornell University.)

4.5.2 Some modern detector systems

The order of the detectors shown schematically in Figure 4.19 is dictated by the penetrating power of the various particles discussed in Section 4.3.5, and is typical of detector systems at high-energy colliders. However, actual detectors are much more complicated when considered in detail, and often include other features, like scintillation counters used for triggering, as discussed in Section 4.4.4. We will illustrate this by some examples.

The first is the CDF detector, in use at the $p\bar{p}$ Tevatron collider at Fermilab, which is shown schematically in Figure 4.24. The detection of the top quark and the measurement of its mass were first made using this device in 1995. The dashed lines indicate some particles produced in the collision. CDF is a large device, being approximately 12 m in all three directions and weighing about 5000 tonnes. The beams of protons and antiprotons enter from each end through focusing quadrupole magnets and interact in the central intersection region where there is a silicon vertex detector (1) to detect the decay of very short-lived particles.[23] The intersection point is surrounded by a 2000 tonne detector system that, in addition to the vertex detector, consists of a tracking detector composed of inner drift chambers (2), electromagnetic calorimeters

[23] The installation of the silicon vertex detector in CDF is shown in coloured Plate 5.

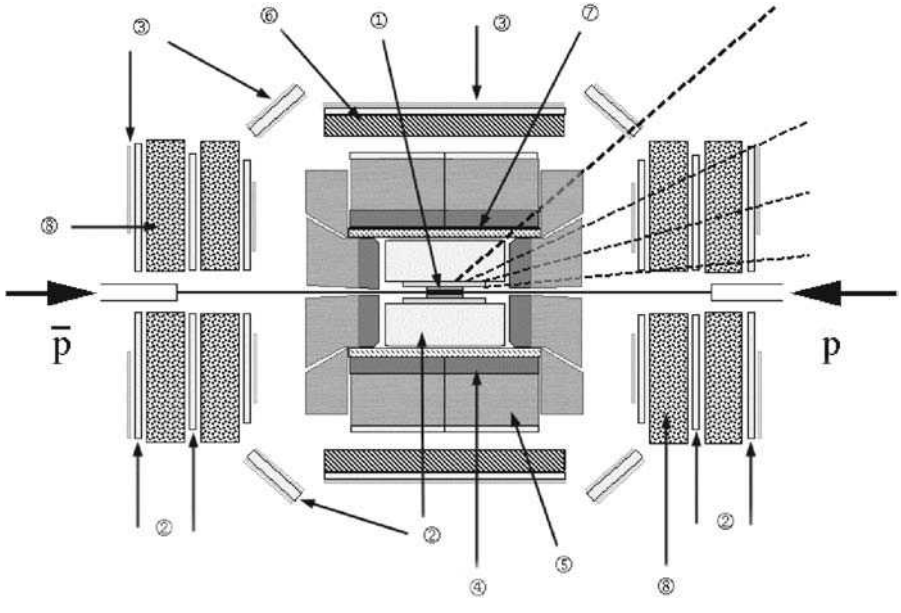


Figure 4.24 The CDF detector used at the Tevatron collider at FNAL. The various subdetectors shown are: (1) a silicon vertex detector; (2) drift chambers; (3) scintillation counters; (4) electromagnetic calorimeters; (5) hadron calorimeters. Steel shielding (6) is also shown, while magnetic fields are produced by a solenoid coil (7) and magnetic toroids (8). (Fermilab Graphic.)

(4), hadron calorimeters (5), time-of-flight detectors (not indicated) and further drift chambers (2) on the outside to detect muons. The whole system is in a magnetic field with the solenoid coil shown at (7) and steel shielding at (6). The rest of the detector consists of two symmetrical sets of drift chambers (2) sandwiched between scintillation counters (3) and magnetic toroids (8) to provide momentum measurements, primarily for muons.

The second example is the STAR detector, shown in Figure 4.25, which has been in use at the RHIC accelerator at Brookhaven National Laboratory since 2000. In this case, the central tracker is a time projection chamber, rather than the drift chambers used in CDF. It is used, among other things, to detect events resulting from the collisions of heavy ions, typically those of fully stripped gold nuclei, where the final state may contain many thousands of particles.[24]

The final example is the ATLAS detector, which is one of the detectors built for the large hadron collider (LHC) and has been operational from 2008. It is hoped that this and other detectors at the LHC will be able to detect the Higgs boson, if it exists, and so help solve one the outstanding current problems in particle physics – the origin of mass. A perspective view of the ATLAS detector is shown in Figure 4.26 (see also Plate 2). As can be seen, it has the usual basic arrangement of a tracking detector, calorimeters and muon counters, but the tracking detector is a solid-state detector

[24] A view of the STAR detector during its construction is shown in coloured Plate 7.

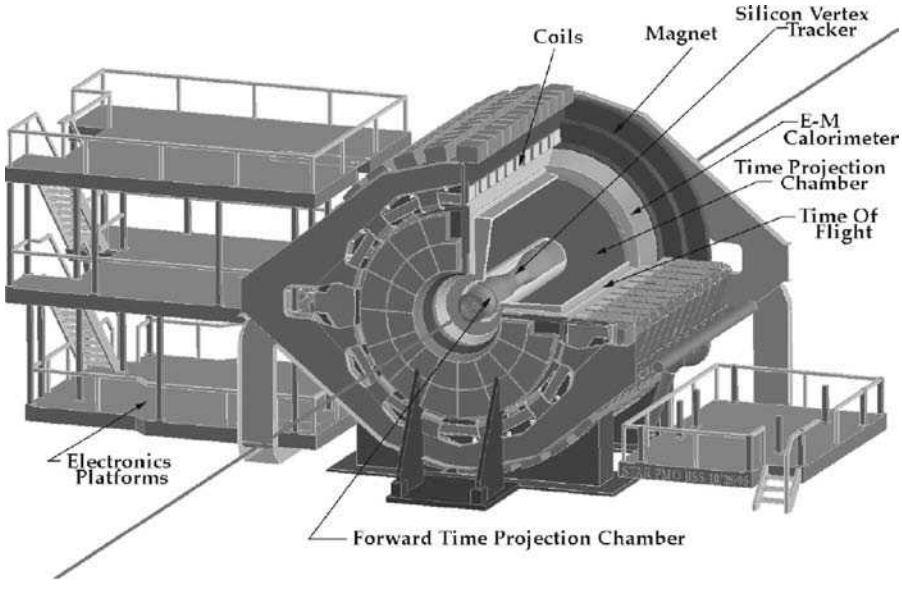


Figure 4.25 The STAR detector at the RHIC accelerator at Brookhaven National Laboratory, USA. (Courtesy of Brookhaven National Laboratory.)

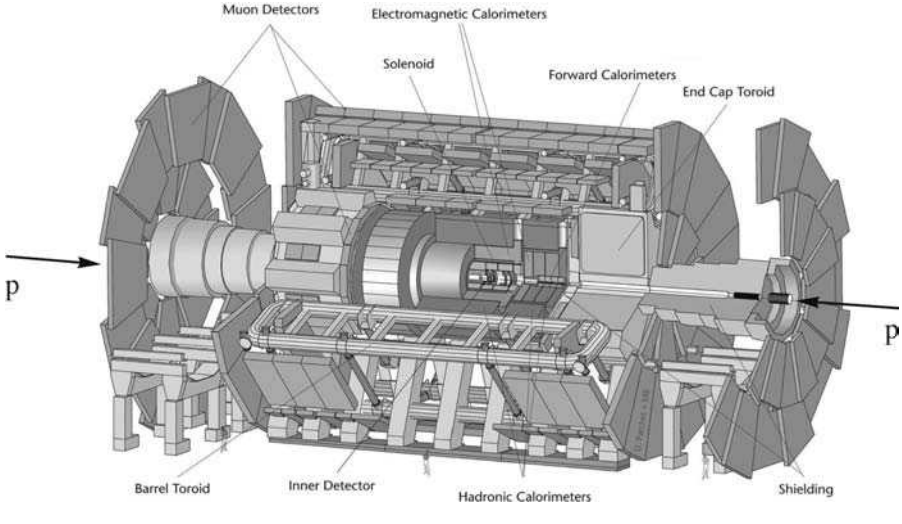


Figure 4.26 The ATLAS detector used at the *pp* LHC at CERN, Geneva. (Reproduced by permission of CERN.)

rather than a drift or time projection chamber. The Atlas detector is even larger than CDF and measures about 25 m in diameter and 46 m long, with an overall weight of approximately 7000 tonnes.

PROBLEMS 4

4.1 What is the length L of the longest drift tube in a linac, which, operating at a frequency $f = 50\,\text{MHz}$, is capable of accelerating protons to a maximum energy of $E = 300\,\text{MeV}$?

4.2 Show that the momentum p in GeV/c for a relativistic particle of charge e moving in a circular orbit of radius ρ metres in a magnetic field of magnitude $B = |\mathbf{B}|$ teslas is given by $p = 0.3\,B\rho$.

4.3 At a collider, two high-energy particles A and B with energies E_A and E_B, which are much greater than either of the masses, collide at a crossing angle θ. Show that the total centre-of-mass energy E_{CM} is given by

$$E_{CM}^2 = 2E_A E_B(1 + \cos\theta)$$

if the rest masses are neglected. At the HERA collider (which was operational until 2007) in the DESY Laboratory in Hamburg, a 30 GeV electron beam collided with an 820 GeV proton beam at zero crossing angle. Evaluate the total centre-of-mass energy and show that a fixed-target electron accelerator would require a beam energy of approximately 5×10^4 GeV to achieve the same total centre-of-mass energy.

4.4 Charged pions with momentum 50 GeV/c are deflected through a collimator slit 3 mm wide by a bending magnet 2 m long which produces a uniform field of 1.5 T. Calculate how far from the magnet the slit should be placed so that it accepts particles with momentum within about 1 % of the central value.

4.5 Estimate the thickness of iron through which a beam of neutrinos with energy 200 GeV must travel if 1 in 10^9 of them is to interact. Assume that at high energies the neutrino–nucleon total cross-section is given approximately by $\sigma_\nu \approx 10^{-38} E_\nu\,\text{cm}^2$, where E_ν is given in GeV. The density of iron is $\rho = 7.9\,\text{g cm}^{-3}$.

4.6 Derive the result of Equation (4.12) for the range R.

4.7 Two particles of masses m_1 and m_2, and common momentum p travel between two scintillation counters that are a distance L apart. Show that the difference in their flight times decreases like p^{-2} for large momenta. Calculate the minimum flight path necessary to distinguish pions from kaons if they have momentum 3 GeV/c and the time-of-flight can be measured with an accuracy of 200 ps.

4.8 Estimate the minimum length of a gas Čerenkov counter that could be used in the threshold mode to distinguish between pions and kaons with momentum 20 GeV/c. Assume that 200 photons need to be radiated to ensure a high probability of detection. Assume also that the radiation covers the whole visible spectrum and neglect the variation with wavelength of the refractive index of the gas.

4.9 Use the model for electromagnetic showers of Section 4.4.6 to show that (a) the energy spectrum of all the secondaries contained in the shower falls approximately like E^{-2} for $E_0 \gg E > E_C$ and (b) the sum of the track lengths of all charged secondaries is approximately proportional to the primary energy. (*Hint.* Start by writing an expression for the total number of secondaries whose energy is greater than E and approximate the sum involved by an integral.)

4.10 The reaction $e^+ + e^- \to \mu^+ + \mu^-$ is studied using a collider with equal beam energies of 3 GeV. The differential cross-section *in natural units* is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{4E_{CM}^2}(1 + \cos^2\theta),$$

where E_{CM} is the total centre-of-mass energy and θ is the angle between the initial e^- and the final μ^-. If the detector can only record an event if the $\mu^+\mu^-$ pair makes an angle of at least 25° relative to the beam line, what fraction of events will be recorded? What is the total cross-section for this reaction in nanobarns? If the reaction is recorded for 10^7 s at a luminosity of $L = 10^{31}\,\text{cm}^{-2}\text{s}^{-1}$, how many events are expected?

5

Space–time Symmetries

Symmetries and conservation laws are a constant theme in physics, and nowhere more so than in particle physics. In this chapter we concentrate on those laws that are associated with space–time symmetries and their applications in strong and electromagnetic interactions. Such conservation laws are particularly important in spectroscopy. In atomic physics, spectroscopy is a vital ingredient in understanding the structure of atoms in terms of nuclei and electrons, and it plays an analogous role in elucidating the internal structure of hadrons in terms of their constituent quarks. In both cases, each state in the spectrum has not only a specific energy but also well-defined values of 'good' quantum numbers associated with conserved observables, like angular momentum, whose quantum mechanical operators commute with the Hamiltonian of the system. In atomic physics these quantum numbers are crucial to understanding the degeneracies of the energy levels, their behaviours in the presence of electric and magnetic fields, and the selection rules that govern transitions between them. Conservation laws have an equally important place in hadron spectroscopy, and the first task in studying hadrons is to determine which are the appropriate conserved quantities and to measure their values for observed states.

Some of the conservation laws discussed in this chapter – those for linear and angular momentum – are universal laws of nature, valid for all interactions. Others we shall meet, like parity, are only conserved in the approximation that weak interactions are neglected. Their violation will be discussed in Chapter 10. Here we shall neglect weak interactions and concentrate on the strong and electromagnetic interactions with which we will be primarily concerned in the next two chapters.

The conservation laws we discuss have their origin in the symmetries and invariance properties of the underlying interactions. To understand the connection between these we consider firstly a simple example, that of translational invariance, and show that it leads directly to the conservation of linear momentum.

Particle Physics Third Edition B. R. Martin and G. Shaw

5.1 TRANSLATIONAL INVARIANCE

Translational invariance expresses the fact that all positions in space are physically indistinguishable. By this we mean that when a closed system of particles (i.e. one on which no external forces are acting) is moved as a whole from one position in space to another, its physical properties are unaltered. In quantum mechanics this is expressed as an invariance of the Hamiltonian. If the system is displaced a distance $\mathbf{a}$, then the position vector vector $\mathbf{r}_i$ of a particle i becomes $\mathbf{r}_i' = \mathbf{r}_i + \mathbf{a}$, i.e.

$$\mathbf{r}_i \to \mathbf{r}_i' = \mathbf{r}_i + \mathbf{a}, \tag{5.1}$$

where it is sufficient for what follows to consider only infinitesimal displacements, $\mathbf{a} = \delta\mathbf{r}$. The Hamiltonian will similarly be replaced by a new Hamiltonian

$$H(\mathbf{r}_1', \mathbf{r}_2', \ldots) = H(\mathbf{r}_1 + \delta\mathbf{r}, \mathbf{r}_2 + \delta\mathbf{r}, \ldots).$$

For the case of a single free particle of mass m,

$$H = -\frac{1}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

and one easily sees that the two Hamiltonians are equal, i.e.

$$H(\mathbf{r}') = H(\mathbf{r} + \delta\mathbf{r}) = H(\mathbf{r}). \tag{5.2}$$

In general,

$$H(\mathbf{r}_1', \mathbf{r}_2', \ldots) = H(\mathbf{r}_1, \mathbf{r}_2, \ldots) \tag{5.3}$$

for a closed system, and the Hamiltonian is said to be invariant under the translation (5.1).

We now specialize to a system containing just a single particle and introduce a translation operator $\hat{D}$ defined by

$$\hat{D}\psi(\mathbf{r}) \equiv \psi(\mathbf{r} + \delta\mathbf{r}), \tag{5.4}$$

where $\psi(\mathbf{r})$ is an arbitrary wavefunction. Expanding the right-hand side of (5.4) to first order in $\delta\mathbf{r}$ gives

$$\psi(\mathbf{r} + \delta\mathbf{r}) = \psi(\mathbf{r}) + \delta\mathbf{r} \cdot \nabla\psi(\mathbf{r}).$$

Thus

$$\hat{D} = 1 + i\delta\mathbf{r} \cdot \hat{\mathbf{p}}, \tag{5.5}$$

where $\hat{\mathbf{p}} = -i\nabla$ is the operator for linear momentum. The conservation law for linear momentum follows by first applying $\hat{D}$ to the wavefunction

$$\psi'(\mathbf{r}) = H(\mathbf{r})\psi(\mathbf{r}),$$

giving

$$\hat{D}\psi'(\mathbf{r}) = \hat{D}H(\mathbf{r})\psi(\mathbf{r}), \tag{5.6}$$

and then comparing with the results obtained from the definition (5.4), i.e.

$$\begin{aligned}\hat{D}\psi'(\mathbf{r}) &= \psi'(\mathbf{r}+\delta\mathbf{r}) = H(\mathbf{r}+\delta\mathbf{r})\psi(\mathbf{r}+\delta\mathbf{r}) \\ &= H(\mathbf{r})\psi(\mathbf{r}+\delta\mathbf{r}) = H(\mathbf{r})\hat{D}\,\psi(\mathbf{r}),\end{aligned} \tag{5.7}$$

where we have used the invariance property of the Hamiltonian (5.3). Comparing (5.6) and (5.7), we see that

$$\left(\hat{D}H(\mathbf{r}) - H(\mathbf{r})\hat{D}\right)\psi(\mathbf{r}) = 0,$$

and since $\psi(\mathbf{r})$ is an arbitrary wavefunction it follows that $\hat{D}$ commutes with the Hamiltonian, i.e.

$$\left[\hat{D}, H\right] = 0.$$

Substituting (5.5) into this commutation relation leads to the conservation law for linear momentum

$$\left[\hat{\mathbf{p}}, H\right] = 0 \tag{5.8}$$

for a single-particle state whose Hamiltonian is invariant under the translation (5.1). The generalization to an N-particle state, described by the wavefunction $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$, is straightforward (see Problem 5.1) and leads to conservation of the total linear momentum

$$\mathbf{p} = \sum_{i=1}^{N} \mathbf{p}_i,$$

where $\mathbf{p}_i$ is the momentum of the ith particle.

The chain of arguments leading to the conservation of linear momentum is characteristic of the way conservation laws in general arise in quantum mechanics. Associated with a symmetry of the system is a transformation of the particle coordinates,

$$\mathbf{r}_i \to \mathbf{r}_i', \tag{5.9}$$

which leaves the Hamiltonian invariant:

$$H(\mathbf{r}_1, \mathbf{r}_2, \ldots) = H(\mathbf{r}'_1, \mathbf{r}'_2, \ldots). \tag{5.10}$$

The associated observable is then identified by examining the corresponding transformation of an arbitrary wavefunction $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots)$, and the conservation law established by applying the same transformation to the wavefunction $\psi' = H\psi$, as illustrated in Equations (5.6) to (5.8).

5.2 ROTATIONAL INVARIANCE

Angular momentum conservation is a consequence of rotational invariance in the same way that linear momentum conservation follows from translational invariance. Rotational invariance expresses the fact that all directions in space are physically indistinguishable. In particular, when a closed system of particles is rotated as a whole about its centre-of-mass to a new orientation in space, its physical properties are unchanged. Under such a rotation, the position vector $\mathbf{r}_i$ of any particle i will change to a new value $\mathbf{r}'_i$, where, for example, if the rotation is through an angle θ about the z axis,

$$\begin{aligned} x_i &\to x'_i = x_i \cos\theta - y_i \sin\theta, \\ y_i &\to y'_i = x_i \sin\theta + y_i \cos\theta, \\ z_i &\to z'_i = z_i. \end{aligned} \tag{5.11}$$

Correspondingly, the Hamiltonian $H(\mathbf{r}_1, \mathbf{r}_2, \ldots)$ will be replaced by a new Hamiltonian $H(\mathbf{r}'_1, \mathbf{r}'_2, \ldots)$, and if the system is rotationally invariant, these two Hamiltonians are the same, i.e.

$$H(\mathbf{r}_1, \mathbf{r}_2, \ldots) = H(\mathbf{r}'_1, \mathbf{r}'_2, \ldots). \tag{5.12}$$

This invariance property holds for any closed system, and also for the case of a particle moving in a central potential $V(r)$ with the Hamiltonian

$$H = -\frac{1}{2m}\nabla^2 + V(r),$$

where

$$r = (x^2 + y^2 + z^2)^{1/2}$$

as usual. The derivation of angular momentum conservation from this invariance is similar to the derivation of linear momentum conservation above, but is more complicated. We will therefore give the derivation only for the case of spinless particles, and state without proof the corresponding result for particles with spin.

5.2.1 Angular momentum conservation

We start by deriving the law of angular momentum conservation for a single spin-0 particle moving in a central potential. To do this, we need only consider infinitesimal rotations, and for an infinitesimal rotation through an angle $\delta\theta$ about the z axis, (5.11) reduces to

$$x' = x - y\,\delta\theta,\ y' = x\,\delta\theta + y,\ z' = z.$$

We next introduce a rotation operator $\hat{R}_z$ defined by analogy with the displacement operator (5.4) by

$$\hat{R}_z(\delta\theta)\psi(\mathbf{r}) \equiv \psi(\mathbf{r}') = \psi(x - y\delta\theta,\, x\delta\theta + y,\, z). \tag{5.13}$$

Expanding the right-hand side to first order in $\delta\theta$ gives

$$\psi(\mathbf{r}') = \psi(\mathbf{r}) - \delta\theta\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right) = (1 + i\,\delta\theta\,\hat{L}_z)\psi(\mathbf{r}), \tag{5.14}$$

where

$$\hat{L}_z = -i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

is the z component of the orbital angular momentum operator $\hat{\mathbf{L}}$. Comparing (5.13) and (5.14) then gives the explicit form

$$\hat{R}_z(\delta\theta) = (1 + i\delta\theta\,\hat{L}_z)$$

for the rotation operator to first order in $\delta\theta$. There is of course nothing special about the z direction, and for a rotation about an arbitrary axis specified by the unit vector $\mathbf{n}$, we have similarly

$$\hat{R}_{\mathbf{n}}(\delta\theta)\psi(\mathbf{r}) \equiv \psi(\mathbf{r}'), \tag{5.15}$$

where $\hat{L}_z$ in (5.14) is replaced by the component of $\hat{\mathbf{L}}$ in the new direction $\mathbf{n}$, i.e.

$$\hat{R}_{\mathbf{n}}(\delta\theta) = (1 + i\delta\theta\,\hat{\mathbf{L}}\cdot\mathbf{n}) \tag{5.16}$$

to first order in the rotation angle $\delta\theta$.

The conservation law for orbital angular momentum can now be obtained in the same way that linear momentum conservation was derived in Equations (5.6) to (5.8). Applying $\hat{R}_n(\delta\theta)$ to the wavefunction

$$\psi'(\mathbf{r}) = H(\mathbf{r})\psi(\mathbf{r})$$

gives

$$\hat{R}_{\mathbf{n}}\psi'(\mathbf{r}) = \hat{R}_{\mathbf{n}}H(\mathbf{r})\psi(\mathbf{r}), \tag{5.17}$$

whereas the definition (5.15) gives

$$\begin{aligned}\hat{R}_{\mathbf{n}}\psi'(\mathbf{r}) &= \psi'(\mathbf{r}') = H(\mathbf{r}')\psi(\mathbf{r}') \\ &= H(\mathbf{r})\psi(\mathbf{r}') = H(\mathbf{r})\hat{R}_{\mathbf{n}}\psi(\mathbf{r}),\end{aligned} \tag{5.18}$$

where we have used the invariance of the Hamiltonian

$$H(\mathbf{r}') = H(\mathbf{r}).$$

Comparing (5.17) and (5.18) gives, for an arbitrary wavefunction $\psi(\mathbf{x})$,

$$\left(\hat{R}_{\mathbf{n}}H(\mathbf{r}) - H(\mathbf{r})\hat{R}_{\mathbf{n}}\right)\psi(\mathbf{r}) = 0,$$

so that $\hat{R}_{\mathbf{n}}$ commutes with the Hamiltonian,

$$\left[\hat{R}_{\mathbf{n}}\,,\, H\right] = 0. \tag{5.19}$$

Finally, substituting (5.16) into this commutation relation leads to the conservation law for orbital angular momentum,

$$\left[\hat{\mathbf{L}}\,, H\right] = 0 \tag{5.20}$$

for a spinless particle whose Hamiltonian is invariant under rotation.

The above argument applies to a spin-0 particle in a central potential. If the particle has spin, then the total angular momentum $\mathbf{J}$ is the sum of the orbital and spin angular momenta,

$$\mathbf{J} = \mathbf{L} + \mathbf{S},$$

and the wavefunction Ψ may be written as the product of a space wavefunction $\psi(\mathbf{r})$ and a spin wavefunction χ, i.e.

$$\Psi = \psi(\mathbf{r})\chi.$$

For the case of spin-$\frac{1}{2}$ particles we shall use the notation $\chi = \alpha$ to denote a spin 'up' state ($S_z = \frac{1}{2}$) and $\chi = \beta$ to denote a spin 'down' state ($S_z = -\frac{1}{2}$). It is convenient to write the spin-$\frac{1}{2}$ operator $\hat{\mathbf{S}}$ in terms of the *Pauli spin matrices* $\boldsymbol{\sigma}$ by

$$\hat{\mathbf{S}} = \tfrac{1}{2}\boldsymbol{\sigma}, \tag{5.21a}$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.21b}$$

Then α and β may be represented by

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{5.22}$$

which satisfy the eigenvalue equations

$$\hat{S}_z\alpha = \tfrac{1}{2}\alpha\,, \qquad \hat{S}_z\beta = -\tfrac{1}{2}\beta,$$

as required. Under a rotation, both spin and space wavefunctions will alter. We will not discuss the rotation properties in detail, but simply state without proof the result that under a rotation through a small angle $\delta\theta$ about a direction $\mathbf{n}$, the rotation operator of Equation (5.16) generalizes to

$$\hat{R}_{\mathbf{n}}(\delta\theta) = (1 + i\delta\theta\hat{\mathbf{J}}\cdot\mathbf{n}), \tag{5.23}$$

so that

$$\Psi \to \Psi' = (1 + i\delta\theta\,\hat{\mathbf{J}}\cdot\mathbf{n})\Psi. \tag{5.24}$$

On the right-hand side, $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ act only on the space and spin wavefunctions, respectively, i.e.

$$\hat{\mathbf{J}}\Psi = (\hat{\mathbf{L}} + \hat{\mathbf{S}})\Psi = \left(\hat{\mathbf{L}}\psi(\mathbf{r})\right)\chi + \psi(\mathbf{r})\left(\hat{\mathbf{S}}\chi\right).$$

The arguments that led to the conservation of orbital angular momentum in the absence of spin may be applied in the present case using Equation (5.23) for the rotation operator, and lead directly to conservation of the total angular momentum,

$$\left[\hat{\mathbf{J}}, H\right] = 0. \tag{5.25}$$

Rotational invariance does *not* lead to the conservation of $\mathbf{L}$ and $\mathbf{S}$ separately, and in general the orbital and spin angular momenta are *not* conserved, i.e.

$$\left[\hat{\mathbf{L}}, H\right] = -\left[\hat{\mathbf{S}}, H\right] \neq 0,$$

because of the existence of spin-dependent forces.

5.2.2 Classification of particles

In this book, we use the word *particle* to describe both elementary particles, which have no internal structure, and composite particles, which are bound states of smaller constituents. In either case, one of the quantum numbers that characterize the particle is its spin, which is defined as the particle's angular momentum in its own rest frame. For a composite particle, this frame corresponds to the centre-of-mass frame of its constituents. For example, mesons are $q\bar{q}$ bound states, with the rest frame of the meson corresponding to the centre-of-mass frame of the $q\bar{q}$ pair. Consequently, the spin $\mathbf{S}_P$ of a particle P is the total angular momentum $\mathbf{J}$ of its constituents in this frame. The spin quantum number $\mathbf{S}_P = \mathbf{J}$ can be either integer or half-integer, while its z component can take on any of the $2J + 1 = 2S_P + 1$ possible values

$$(S_P)_z = J_z = J, J - 1, \ldots, -J,$$

depending on the orientation of the particle's spin relative to the z axis.

In addition to its spin, a composite particle may have other 'good' quantum numbers, related to its internal structure, which take on definite values and therefore characterize the particle. 'Good' quantum numbers are defined as those that are constants of the motion, so that the associated operators commute with the Hamiltonian. In other words, they are associated with conserved observables. In the rest frame of the particle, the total angular momentum $\mathbf{J}$ of the constituents that make up the particle is conserved, but the total orbital angular momentum $\mathbf{L}$ and the total spin angular momentum $\mathbf{S}$ are not separately conserved. However, despite this, it is often a very good approximation to assume that $\mathbf{L}^2$ and $\mathbf{S}^2$, defined by

$$\mathbf{L}^2 \equiv L_x^2 + L_y^2 + L_z^2$$

and

$$\mathbf{S}^2 \equiv S_x^2 + S_y^2 + S_z^2,$$

are conserved, i.e. to assume that

$$\left[H, \hat{\mathbf{L}}^2\right] = \left[H, \hat{\mathbf{S}}^2\right] = 0. \tag{5.26}$$

This corresponds to assuming that the forces can, for example, flip the direction of the spin, thereby changing its components but not its magnitude. In this approximation, the orbital angular momentum and spin quantum numbers L and S are also good quantum numbers, so that the particle is characterized by $S_P = J$, L and S, while a fourth J_z depends on the orientation of its spin. This is the basis of the so-called spectroscopic notation, in which states are denoted by

$$^{2S+1}L_J, \tag{5.27}$$

and where, instead of the numerical value of L, it is conventional to write $S, P, D, F, \ldots$ for $L = 0, 1, 2, 3, \ldots$. Thus a composite particle denoted by $^3S_1$ has $L = 0, S = 1$ and spin $S_P = J = 1$, while one denoted $^2D_{3/2}$ has $L = 2$, $S = \frac{1}{2}$ and spin $S_P = J = \frac{3}{2}$. In all of this it is of course crucial to distinguish between the spin $S_P = J$ of the particle and the spin S of its constituents.

Before applying these ideas to hadrons, it is instructive to consider a simple example from nuclear physics. This is the deuteron, which is a spin-1 bound state of a neutron and a proton. As usual in nuclear physics, we shall neglect the quark substructure of the nucleons and treat them as elementary spin-$\frac{1}{2}$ particles. Since the deuteron has spin 1, the total angular momentum of the pn system in its centre-of-mass frame must be $J = 1$. It is also reasonable to assume that $L = 0$, as the lowest lying bound states of any system are usually S waves. Thus we deduce that $S = 1$ for the sum of the proton and neutron spins, so that in the spectroscopic notation of (5.27), the deuteron is a $^3S_1$ bound state of the pn system.

This assignment may be tested by considering the magnetic moment of the deuteron. If $L = 0$, as we have assumed, there will be no contribution from the orbital motion of the proton's charge. Hence, the deuteron's magnetic moment must arise solely from the spin magnetic moments of the nucleons and must be given by

$$\mu_d = \mu_p + \mu_n = 2.793 - 1.913 = 0.880$$

nuclear magnetons, since the proton and neutron spins are parallel ($S = 1$). The experimental value is $\mu_d = 0.857$ nuclear magnetons, which is close to, but not exactly equal to, the above values. The difference is due to the fact that the deuteron is not a pure $^3S_1$ state, but has a few percent admixture of the $^3D_1$ state with $L = 2$. Thus, while the assumption that L is a good quantum number in the presence of nuclear or strong forces is a good approximation and very useful, it is not exact. With this caveat, we now use the spectroscopic notation to classify hadrons in the framework of the quark model.

5.2.3 Angular momentum in the quark model

In the simple quark model we assume that it is a good approximation to treat L and S as good quantum numbers and that the lightest states for any combinations of $q\bar{q}$ and $3q$ have zero orbital angular momenta. The spectral classification of mesons and baryons is then as follows.

Mesons are $q\bar{q}$ bound states, with the rest frame of the meson corresponding to the centre-of-mass frame of the $q\bar{q}$ system. In this frame there is a single orbital angular momentum $\mathbf{L}$, but two constituent spins, so that in an obvious notation

$$\mathbf{S} = \mathbf{S}_q + \mathbf{S}_{\bar{q}}.$$

Since both q and $\bar{q}$ have spin-$\frac{1}{2}$, the only possibilities are $S=0$ and $S=1$. For $L=0$, $J=S$ and in the spectroscopic notation (5.27), the possible states are

$$^{2S+1}L_J = {}^1S_0\,, {}^3S_1 \qquad (L=0). \tag{5.28a}$$

For $L \geq 1$, the possible J values are $J=L$ for singlet states ($S=0$) and $J=L\pm 1\,, L$ for triplet states ($S=1$), giving the four possibilities

$$^{2S+1}L_J = {}^1L_L\,, {}^3L_{L+1}\,, {}^3L_L\,, {}^3L_{L-1} \quad (L \geq 1). \tag{5.28b}$$

Here J refers to the total angular momentum in the $q\bar{q}$ centre-of-mass frame, which from the discussion in Section 5.2.2 we identify with the meson spin

$$\mathbf{S}_M = \mathbf{J}.$$

If the lightest states have $L=0$, then by (5.28a) they can have spin 0 or 1, so that for any flavour combination $u\bar{d}, u\bar{s}, c\bar{c}$, etc., we would expect two low-mass states with spin-0 and spin-1. This is exactly what is found experimentally, as we shall discuss in Chapter 6, and the lower lying of the two states always has spin-0. Examples are the pions, kaons and D mesons discussed in Chapter 3, which are all spin-0 mesons corresponding to ${}^1S_0$ states of the appropriate $q\bar{q}$ systems given in Equations (3.13), (3.16) and (3.18).

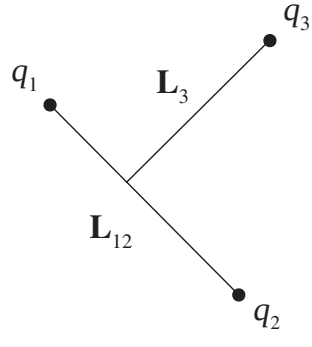


Figure 5.1 Internal orbital angular momenta of a three-quark state.

Baryons are $3q$ bound states, so in the rest frame of the baryon, corresponding to the centre-of-mass frame of the $3q$ system, there are two orbital angular momenta associated with the relative motion of the three quarks. The first is conveniently taken to be the orbital angular momentum $\mathbf{L}_{12}$ of a chosen pair of quarks in their mutual centre-of-mass frame. The second is then the orbital angular momentum $\mathbf{L}_3$ of the third quark about the centre-of-mass of the pair in the overall centre-of-mass frame. This is illustrated in Figure 5.1. The total orbital angular momentum is given by

$$\mathbf{L} = \mathbf{L}_{12} + \mathbf{L}_3,$$

while the spin is the sum of the spins of the three quarks

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3,$$

so that $S = \frac{1}{2}$ or $S = \frac{3}{2}$. It is now straightforward to show that the possible baryon states in the spectroscopic notation are

$$^2S_{1/2}, {}^4S_{3/2} \quad (L = 0), \tag{5.29a}$$

$$^2P_{1/2}, {}^2P_{3/2}, {}^4P_{1/2}, {}^4P_{3/2}, {}^4P_{5/2} \quad (L = 1), \tag{5.29b}$$

and

$$^2L_{L+1/2}, {}^2L_{L-1/2}, {}^4L_{L-3/2}, {}^4L_{L-1/2}, {}^4L_{L+1/2}, {}^4L_{L+3/2} \quad (L \geq 2). \tag{5.29c}$$

The baryon spectrum is clearly very complicated! However, if we restrict ourselves to the lowest lying states with zero orbital angular momentum $\mathbf{L}_{12} = \mathbf{L}_3 = 0$, then only those of (5.29a) can occur, and the baryon spin $S_B = J$ can be $\frac{1}{2}$ or $\frac{3}{2}$. We therefore expect the lightest baryons to have spin-$\frac{1}{2}$ or spin-$\frac{3}{2}$. Again this is in accord with experiment, as we shall see in Chapter 6. For the moment, we merely note that we have already met some of the low-lying spin-$\frac{1}{2}$ baryons in Chapter 3, namely p, n, Λ, Λ_c, which we now identify with the $^2S_{1/2}$ states of the appropriate quark constituents given in Table 3.2.

5.3 PARITY

The parity transformation is

$$\mathbf{r}_i \to \mathbf{r}_i' = -\mathbf{r}_i, \tag{5.30}$$

so the position vector $\mathbf{r}_i$ of every particle i is reflected in the origin. A system is said to be invariant under parity if the Hamiltonian remains unchanged by this transformation, i.e. if

$$H(\mathbf{r}_1', \mathbf{r}_2', \ldots) = H(-\mathbf{r}_1, -\mathbf{r}_2, \ldots) = H(\mathbf{r}_1, \mathbf{r}_2, \ldots). \tag{5.31}$$

Unlike translational and rotational invariance, parity is not an exact symmetry of nature that holds for all closed systems, but is violated by the weak interaction. This will be discussed in Chapter 10. For the present, we will ignore the weak interaction and only consider systems that interact by the strong and electromagnetic interactions. We will assume that these interactions are invariant under the parity transformation and explore its consequences.

We start by specializing to a single particle and introduce a parity operator $\hat{P}$ defined by

$$\hat{P}\psi(\mathbf{r}, t) \equiv P_a \psi(-\mathbf{r}, t), \tag{5.32a}$$

where the label a identifies the particle type (e.g. $a = e^-$ for an electron, $a = \mu^-$ for a muon, etc.) and P_a is a constant phase factor. Since two successive parity transformations leave the system unchanged, we require that

$$\hat{P}^2\psi(\mathbf{r},t) = \psi(\mathbf{r},t),$$

implying

$$P_a = \pm 1 \tag{5.32b}$$

for the possible values of P_a. If we consider an eigenfunction of momentum

$$\psi_{\mathbf{p}}(\mathbf{r},t) = \exp[i(\mathbf{p}\cdot\mathbf{r} - Et)],$$

then

$$\hat{P}\psi_{\mathbf{p}}(\mathbf{r},t) = P_a\psi_{\mathbf{p}}(-\mathbf{r},t) = P_a\psi_{-\mathbf{p}}(\mathbf{r},t),$$

so that a particle at rest, with $\mathbf{p} = \mathbf{0}$, is an eigenstate of the parity operator with eigenvalue P_a. For this reason P_a is called the *intrinsic parity* of particle a, or more usually just the *parity* of particle a, with the words *at rest* left implicit. The values of P_a for various particle types will be determined below; here we merely note that if we tried to set $P_a = 1$ for all particles, we would be led to contradictions with experiments in which particles are either created or destroyed. For many-particle systems the appropriate generalization of (5.32a) is

$$\hat{P}\,\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,t) \equiv P_1P_2\ldots\psi(-\mathbf{x}_1,-\mathbf{r}_2,\ldots,t),$$

with one intrinsic parity factor occurring for each particle present.

In addition to a particle at rest, a particle with a definite orbital angular momentum is also an eigenstate of parity. The wavefunction for such a particle has the form

$$\psi_{nlm}(\mathbf{r}) = R_{nl}(r)\,Y_l^m(\theta,\phi), \tag{5.33a}$$

where (r,θ,ϕ) are spherical polar coordinates, $R_{nl}(r)$ is a function of the radial variable r only, and the $Y_l^m(\theta,\phi)$ are spherical harmonics. For example, the first few spherical harmonics are

$$Y_0^0 = 1/\sqrt{4\pi}, \tag{5.33b}$$

and

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta\,; \quad Y_1^{\pm1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\,\mathrm{e}^{\pm i\phi}. \tag{5.33c}$$

Using the relations

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\cos\phi, \quad z = r\cos\theta$$

between Cartesian and polar coordinates, the parity transformation becomes

$$r \to r' = r\,, \quad \theta \to \theta' = \pi - \theta\,, \quad \phi \to \phi = \pi + \phi,$$

from which it can be shown that

$$Y_l^m(\theta,\phi) \to Y_l^m(\pi - \theta, \pi + \phi) = (-1)^l Y_l^m(\theta,\phi).$$

Hence

$$\hat{P}\psi_{nlm}(\mathbf{r}) = P_a\psi_{nlm}(-\mathbf{r}) = P_a(-1)^l\psi_{nlm}(\mathbf{r}), \tag{5.34}$$

and $\psi_{nlm}(\mathbf{r})$ is an eigenstate of parity with an eigenvalue $P_a(-1)^l$.

If the Hamiltonian is invariant under the parity transformation, then an argument analogous to those used to demonstrate linear and orbital angular momentum conservation in Sections 5.1 and 5.2 leads to the conclusion that parity is conserved, i.e.

$$\left[\hat{P}, H\right] = 0.$$

This has two obvious consequences: in a reaction the initial and final parities are equal, $P_i = P_f$, and for bound states parity is a good quantum number. In particular, atomic and nuclear states have been shown to be eigenstates of parity to very high accuracy by searching for parity-violating transitions, verifying that parity invariance indeed holds for strong (nuclear) and electromagnetic interactions.

5.3.1 Leptons and antileptons

Our discussion of parity invariance has been based so far on nonrelativistic quantum mechanics. In this approximation, electrons and positrons are described by distinct wavefunctions that satisfy the Schrödinger equation, and there is no obvious connection between their intrinsic parities defined by

$$\hat{P}\psi(\mathbf{r},t) = P_{e^\pm}\psi(-\mathbf{r},t)$$

for electron and positron wavefunctions, respectively. However, relativistically a single four-component wavefunction (1.11) describes both electrons and positrons. One would therefore expect their intrinsic parities to be related, and indeed an analysis[1] shows that the Dirac equation is only compatible with parity invariance if

$$P_{e^+}P_{e^-} = -1, \tag{5.35}$$

[1] See, for example, p. 372 of Perkins (1987).

i.e. if electrons and positrons have opposite parities. The same argument applies to other spin-$\frac{1}{2}$ particles, giving

$$P_f P_{\bar{f}} = -1, \tag{5.36}$$

where f labels a particular fermion and $\bar{f}$ the corresponding antifermion.

The prediction (5.35) can be confirmed experimentally by studying the reaction

$$e^+ + e^- \to \gamma + \gamma,$$

where the e^+e^- are bound together in an S-wave state called parapositronium.[2] The initial state has zero orbital momentum, so that its parity is just the product of the intrinsic parities $P_i = P_{e^+} P_{e^-}$. If the final state has relative orbital angular momentum l_γ, then its parity is $P_f = P_\gamma^2 (-1)^{l_\gamma} = (-1)^{l_\gamma}$, where P_γ is the intrinsic parity of the photon and $P_\gamma^2 = 1$ by (5.32b). Parity conservation $P_i = P_f$ then gives

$$P_{e^+} P_{e^-} = (-1)^{l_\gamma},$$

so that a measurement of the orbital angular momentum in the final state l_γ can determine the product $P_{e^+} P_{e^-}$. Such a measurement can be made indirectly by studying the polarizations of the emitted photons,[3] which unambiguously confirms the prediction (5.35).

The above argument determines the product $P_{e^+} P_{e^-}$, but says nothing about the parity of a single electron or positron. Other reactions, such as

$$\gamma + e^- \to \gamma + e^-$$

and

$$e^- + e^- \to e^- + e^-,$$

contain the same factors of P_{e^-} in both the initial and final states, so that they cancel on equating initial and final state parities. In general, since single electrons can be neither created nor destroyed in electromagnetic interactions, similar cancellations occur in all electromagnetic processes and therefore P_{e^-} cannot be determined. Similarly, the intrinsic parities of muons and taus cannot be measured, since, like electrons, they can only be created or destroyed in electromagnetic interactions as particle–antiparticle pairs. Because of this it is conventional to *define*

$$P_{e^-} \equiv P_{\mu^-} \equiv P_{\tau^-} \equiv 1. \tag{5.37a}$$

---

[2] This is the ground state of the e^+e^- system. Both it and the excited states are discussed further in Section 5.5.

[3] See, for example, pp. 89–90 of Perkins (1987).

These choices are arbitrary, but having made them it follows from (5.36) that

$$P_{e^+} = P_{\mu^+} = P_{\tau^+} = -1. \tag{5.37b}$$

5.3.2 Quarks and hadrons

In Section 3.2 we saw that quarks could only be created or destroyed in the strong and electromagnetic interactions as particle–antiparticle pairs. Consequently, the quark intrinsic parities, like those of electrons and positrons, are indeterminate and must be fixed by convention. The usual convention is

$$P_u \equiv P_d \equiv P_s \equiv P_c \equiv P_b \equiv P_t \equiv 1, \tag{5.38a}$$

implying

$$P_{\bar{u}} = P_{\bar{d}} = P_{\bar{s}} = P_{\bar{c}} = P_{\bar{b}} = P_{\bar{t}} = -1 \tag{5.38b}$$

for the antiquarks. With these conventions, we can predict the parities of hadrons from their assumed quark structures, remembering that the hadron rest frame is the same as the centre-of-mass frame of the bound state quark system. Hence the (intrinsic) parity P_M of a meson $M = a\bar{b}$ is

$$P_M = P_a P_{\bar{b}} (-1)^L = (-1)^{L+1}. \tag{5.39a}$$

where L is the orbital angular momentum of the $q\bar{q}$ pair and the quark labels a, b can each be u, d, s, c, b or t. Low-lying mesons, corresponding to $L = 0$, are thus predicted to have negative parity, which is consistent with the observed parities of the spin-0 pions (3.13), kaons (3.16) and D mesons (3.18). The parity of a baryon $B = abc$ is similarly given by

$$P_B = P_a P_b P_c (-1)^{L_{12}} (-1)^{L_3} = (-1)^{L_{12}+L_3}, \tag{5.39b}$$

where L_{12} and L_3 are the internal angular momenta illustrated in Figure 5.1. For the corresponding antibaryon $\bar{B} = \bar{a}\bar{b}\bar{c}$ we have

$$P_{\bar{B}} = P_{\bar{a}} P_{\bar{b}} P_{\bar{c}} (-1)^{L_{12}} (-1)^{L_3} = -(-1)^{L_{12}+L_3} = -P_B, \tag{5.39c}$$

so that baryons and antibaryons have opposite parities, consistent with (5.36). For the low-lying baryons with $L_{12} = L_3 = 0$, we predict positive parity, which is consistent with the observed parities of the spin-$\frac{1}{2}$ baryons p, n, Λ, Λ_c of Table 3.2.

In the above discussion, we have fixed the parities of the known hadrons by making the choice (5.38a) for the first five quarks u, d, s, c, b, since hadrons containing the t quark have not yet been discovered. We could equally well have defined our

conventions by fixing the parities of five suitably chosen hadrons. A convenient choice is

$$P_p \equiv P_n \equiv 1 \tag{5.40a}$$

for the proton and neutron, which are *uud* and *udd* states with $L_{12} = L_3 = 0$, and

$$P_{K^-} \equiv P_{D^-} \equiv P_{B^-} \equiv -1 \tag{5.40b}$$

for $K^-(494)$, $D^-(1869)$ and $B^-(5278)$, which are $L = 0$ mesons with the quark compositions $s\bar{u}, d\bar{c}$ and $b\bar{u}$ given in Table 3.3. The conventions (5.40) are equivalent to the conventions (5.38a) within the context of the quark model. However, they have the considerable advantage that they refer to directly observable particles and can be used to determine experimentally the parity of any hadron without reference to its assumed quark substructure. This is done by using suitably chosen strong and electromagnetic reactions to relate the unknown parity to the values (5.40) by parity conservation. That the resulting values agree with the quark model predictions (5.39) is a significant test of the quark model and of the assignment of the state within it. This will be illustrated immediately for the π^- meson, which should have negative parity if it is really an $L = 0\, u\bar{d}$ state as we have assumed.

5.3.3 Parity of the charged pion

The parity of the π^- meson was first demonstrated to be negative by the observation of the absorption reaction

$$\pi^- + d \to n + n \tag{5.41}$$

for stopped pions in deuterium. We have already met the deuteron in Section 5.2.2 as an S-wave bound state of a proton and a neutron with spin-1 resulting from the parallel spins of the two nucleons. The deuteron parity is therefore $P_d = P_p P_n = 1$. The pions are captured into atomic orbitals to form 'pionic deuterium', and then cascade down through the levels with emission of 'pionic X-rays'. The absorption takes place from S states, because only they have a nonvanishing wavefunction at the origin and so a reasonable chance of finding the π^- and d close together. For S-wave capture the total angular momentum is $J = 1$, since the pions have spin-0 and the deutrons have spin-1, and the parity of the initial state is $P_i = P_{\pi^-} P_d = P_{\pi^-}$. By parity conservation, we then have

$$P_{\pi^-} = P_f = P_n P_n (-1)^L = (-1)^L, \tag{5.42}$$

where L is the orbital angular momentum of the final two-neutron state, which must have $J = 1$ by angular momentum conservation.

In most reactions, the L-values, and hence parities, are found by measuring the angular distributions of the particles involved. In the present case, however, the result can be established by exploiting the Pauli principle, which requires that the final-state wavefunction (i.e. the product of space and spin parts) must be antisymmetric under

the interchange of the two neutrons. The possibilities for the spin state are $S=0$ and $S=1$. We examine each in turn.

The spin triplet state $S=1$ has the symmetric spin wavefunctions[4]

$$\begin{array}{ll} \alpha_1\alpha_2 & (S=1,S_z=1), \\ \dfrac{1}{\sqrt{2}}(\alpha_1\beta_2+\beta_1\alpha_2) & (S=1,S_z=0), \\ \beta_1\beta_2 & (S=1,S_z=-1), \end{array} \tag{5.43}$$

where we have used our previous notation; i.e. $\alpha_i(\beta_i)$ corresponds to a particle i having spin up (down) in the z direction. By the Pauli principle, the space wavefunction must in this case be antisymmetric, which requires L to be odd. Since both $J=1$ and $S=1$, the only possibility is $L=1$, which would then imply $P_{\pi^-}=-1$ by (5.42). The other possibility is the spin singlet state

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\beta_1\alpha_2) \quad (S=S_z=0), \tag{5.44}$$

which is antisymmetric. Consequently, the space wavefunction must be symmetric with L even, implying $J=0,2,\ldots$ since $S=0$. Thus it is not possible to satisfy the Pauli principle and angular momentum conservation with $S=0$, since the initial state has $J=1$. The only possibility is $S=1$, and consequently

$$P_{\pi^-}=-1,$$

in agreement with the assignment of the pion to an $L=0$ state in the quark model.

5.3.4 Parity of the photon

In contrast to the intrinsic parity of an electron or a quark, the intrinsic parity of the photon can be deduced theoretically by invoking the correspondence principle. Classically, the electric field obeys Poisson's equation

$$\nabla\cdot\mathbf{E}(\mathbf{r},t)=\frac{1}{\varepsilon_0}\rho(\mathbf{r},t),$$

where $\rho(\mathbf{r},t)$ is the charge density. Under a parity transformation, the value of the charge density at $\mathbf{r}$ is replaced by its value at $-\mathbf{r}$, i.e.

$$\rho(\mathbf{r},t)\to\rho(-\mathbf{r},t),$$

and since ∇ changes sign under (5.30), the electric field itself must transform according to

$$\mathbf{E}(\mathbf{r},t)\to-\mathbf{E}(-\mathbf{r},t), \tag{5.45}$$

[4] See, for example, Equation (5.80) in Section 5.5.1 of Mandl (1992).

if Poisson's equation is to remain invariant. Alternatively, the electromagnetic field can be described by the vector and scalar potentials $\mathbf{A}(\mathbf{r},t)$ and $\phi(\mathbf{r},t)$ respectively. These are related to the electric field by[5]

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \tag{5.46}$$

and in the absence of charges the scalar potential may be set to zero. In quantum mechanics, it is the vector potential that corresponds to the photon wavefunction,[6] and a photon propagating freely through space with momentum $\mathbf{k}$ is described by a wavefunction

$$\mathbf{A}(\mathbf{r},t) = N\,\boldsymbol{\varepsilon}(\mathbf{k})\,\exp[i(\mathbf{k}\cdot\mathbf{r} - Et)], \tag{5.47}$$

where $\boldsymbol{\varepsilon}(\mathbf{k})$ is a unit polarization vector and N is a normalization constant, which is irrelevant for what follows. Under the parity transformation

$$\mathbf{A}(\mathbf{r},t) \to P_\gamma\,\mathbf{A}(-\mathbf{r},t), \tag{5.48}$$

where the intrinsic parity of the photon P_γ is defined by analogy with Equation (5.32), implying

$$\mathbf{E}(\mathbf{r},t) \to P_\gamma \mathbf{E}(-\mathbf{r},t)$$

on substituting into Equation (5.46). Consistency with (5.45) in the classical limit then requires that the intrinsic parity of the photon $P_\gamma = -1$.

5.4 CHARGE CONJUGATION

Charge conjugation ($\hat{C}$) is the operation that replaces all particles by their antiparticles in the same state, so that momenta, positions, etc., are unchanged. Since the charge and magnetic moment of every particle are reversed, the electromagnetic interaction remains invariant under this operation. Charge conjugation is also a symmetry of the strong interaction, but is violated by the weak interaction, as we shall see in Chapter 10. Again, we shall neglect weak interactions for the present and assume invariance of the Hamiltonian and the associated conservation law

$$\left[\hat{C}, H\right] = 0 \tag{5.49}$$

for strong and electromagnetic interactions only.

We now distinguish between those particles $a = \pi^+, K^+, p, N, \ldots$ that have distinct antiparticles and those particles $\alpha = \pi^0, \gamma, \ldots$ that do not. Charged particles of all types are obvious examples of the former, but some neutral particles, like the neutron

[5] See, for example, p. 229 of Grant and Phillips (1990).
[6] A full discussion of the quantization of electromagnetic radiation may be found in, for example, Chapter 1 of Mandl and Shaw (1993).

which has a nonzero magnetic moment, also belong to this class. Under $\hat{C}$ we then have

$$\hat{C}\,|\alpha\Psi\rangle = C_\alpha|\alpha\Psi\rangle \tag{5.50a}$$

and

$$\hat{C}|a\Psi\rangle = |\bar{a}\Psi\rangle, \tag{5.50b}$$

where C_α is a phase factor analogous to that introduced in the discussion of parity (cf. Equation (5.32)) and we use the Dirac notation summarized at the beginning of Section 2.3. Since a second transformation turns antiparticles back into particles, we require $\hat{C}^2 = 1$ and hence

$$C_\alpha = \pm 1. \tag{5.51}$$

A similar phase factor C_α could have been introduced into (5.50b), but it would have no physical consequences as the relative phase of the two states $|a\Psi\rangle$ and $|\bar{a}\Psi\rangle$ cannot be measured.[7] We therefore omit such a factor. For multiparticle states we then have

$$\hat{C}|\alpha_1,\alpha_2,\ldots\,;a_1,a_2,\ldots\,;\Psi\rangle = C_{\alpha_1}C_{\alpha_2}\ldots|\alpha_1,\alpha_2,\ldots\,;\bar{a}_1,\bar{a}_2,\ldots\,;\Psi\rangle, \tag{5.50c}$$

where Ψ denotes the same multiparticle wavefunction in both sides. In future, when the wavefunction is not our primary concern, we shall omit it.

From (5.50a) we see that the particles $\alpha = \pi^0, \gamma, \ldots$ are eigenstates of $\hat{C}$ with eigenvalues $C_\alpha = \pm 1$, called their *C-parities*. These values are readily measured using C conservation, as we shall see below. Otherwise, eigenstates can only be constructed from particle–antiparticle pairs, when $\hat{C}$ in effect interchanges particle with antiparticle. If the state is symmetric or antisymmetric under $a \leftrightarrow \bar{a}$, then

$$\hat{C}\,|a\Psi_1\,,\bar{a}\Psi_2\rangle = |\,\bar{a}\Psi_1\,,a\Psi_2\rangle = \pm\,|a\Psi_1,\bar{a}\Psi_2\rangle$$

and hence $|a\Psi_1\,,\bar{a}\Psi_2\rangle$ is an eigenstate of $\hat{C}$. A simple example is given by a $\pi^+\pi^-$ pair in a state of definite orbital angular momentum L, when

$$\hat{C}|\pi^+\pi^-\,;L\rangle = (-1)^L|\pi^+\pi^-\,;L\rangle, \tag{5.52}$$

since interchanging the π^+ and π^- reverses their relative position vector in the spatial wavefunction. The same factor occurs for spin-$\frac{1}{2}$ fermions f, like quarks and leptons, together with two further factors: a factor $(-1)^{S+1}$ that arises from interchanging the particles in the spin wavefunctions, which may be verified from the explicit

[7] See Problem 5.6.

wavefunctions (5.43) and (5.44), and a factor (-1) that arises whenever fermions and antifermions are interchanged.[8] Overall, we therefore obtain

$$\hat{C}|f\bar{f};J,L,S\rangle = (-1)^{L+S}|f\bar{f};J,L,S\rangle, \tag{5.53}$$

for a fermion–antifermion pair with total, orbital and spin angular momentum quantum numbers J, L and S, respectively. Thus if the π^0, for example, is really a $^1S_0$ state of $u\bar{u}$ or $d\bar{d}$, as assumed in the quark model, then it must have C-parity $C_{\pi^0} = 1$.

5.4.1 π^0 and η decays

The quark model prediction $C_{\pi^0} = 1$ is confirmed experimentally by the observation of the dominant decay mode

$$\pi^0 \to \gamma + \gamma.$$

If we represent the states in the Dirac notation by just the particle symbol, we have

$$\hat{C}\,|\pi^0\rangle = C_{\pi^0}\,|\pi^0\rangle$$

and

$$\hat{C}\,|\gamma\gamma\rangle = C_\gamma\; C_\gamma\,|\gamma\gamma\rangle = |\gamma\gamma\rangle,$$

since $C_\gamma^2 = 1$ by (5.51). Since the final state has $C = 1$, so does the initial state, implying

$$C_{\pi^0} = 1,$$

in agreement with the prediction of the quark model.

The C-parity of the photon can, like its parity, be inferred from the behaviour of the classical electromagnetic field. Under charge conjugation

$$\mathbf{A}(\mathbf{r},t) \to C_\gamma \mathbf{A}(\mathbf{r},t), \tag{5.54a}$$

by analogy with the corresponding result for parity (5.48). However, since all electric charges change sign under charge conjugation, it is clear that the electric field $\mathbf{E}$ and the scalar potential ϕ must also change sign, i.e.

$$\mathbf{E}(\mathbf{r},t) \to -\mathbf{E}(\mathbf{r},t), \quad \phi(\mathbf{r},t) \to -\phi(\mathbf{r},t). \tag{5.54b}$$

[8] This factor can only be properly understood in the context of quantum field theory. See, for example, pp. 259–260 of Gottfried and Weisskopf (1986).

Substituting (5.54a) and (5.54b) into (5.46) gives the desired result

$$C_\gamma = -1 \tag{5.55}$$

for the C-parity of the photon. This value is confirmed, and C conservation tested, by considering the possible decay mode

$$\pi^0 \to \gamma + \gamma + \gamma. \tag{5.56}$$

Since

$$\hat{C}|\gamma\gamma\gamma\rangle = (C_\gamma)^3|\gamma\gamma\gamma\rangle = C_\gamma|\gamma\gamma\gamma\rangle,$$

the reaction (5.56) is forbidden if $C_\gamma = -1$ *and* C-parity is conserved. If, on the contrary, $C_\gamma = +1$ *or* if C-parity were not conserved in electromagnetic interactions, we would expect

$$R \equiv \frac{\Gamma(\pi^0 \to 3\gamma)}{\Gamma(\pi^0 \to 2\gamma)} = O(\alpha),$$

since no other selection rule forbids this decay. Experimentally, the 3γ mode has never been seen, and the measured upper limit on this ratio is

$$R < 3 \times 10^{-8} \ll \alpha.$$

We therefore conclude that (5.56) is indeed forbidden as an electromagnetic interaction, thus confirming both the assignment $C_\gamma = -1$ and C conservation.

Further tests of$C$ invariance can be made in the decays of the eta (η) particle. This is a neutral spin-0 meson of mass 558 MeV/c^2 that, like the neutral pion, is detected via its decay products. The most important decay modes, with their branching ratios B, are

$$\eta \to \gamma + \gamma, \quad B = 0.39, \tag{5.57a}$$

$$\eta \to \pi^0 + \pi^0 + \pi^0, \quad B = 0.33, \tag{5.57b}$$

and

$$\eta \to \pi^+ + \pi^- + \pi^0, \quad B = 0.23. \tag{5.57c}$$

Reaction (5.57a) is obviously electromagnetic, and since the other two reactions have comparable rates, it is reasonable to assume that the are also electromagnetic. The measured lifetime of approximately 6×10^{-19} s is consistent with this assumption. The C-parity of η follows directly from the existence of the 2γ mode, and is confirmed by the existence of the $3\pi^0$ mode; it is $C_\eta = 1$. In addition, tests of C invariance may be made by considering, for example, the reaction

$$\eta \to \pi^+(\mathbf{p}_1) + \pi^-(\mathbf{p}_2) + \pi^0(\mathbf{p}_3), \tag{5.58}$$

where we have labelled the final state particles by their momenta. If the interaction is C-invariant, then the charge-conjugate reaction

$$\eta \to \pi^-(\mathbf{p}_1) + \pi^+(\mathbf{p}_2) + \pi^0(\mathbf{p}_3)$$

should be indistinguishable from (5.58), implying identical momentum distributions for the π^+ and π^- mesons. Experimentally, the $\pi^\pm$ spectra are indeed the same within experimental errors, and C-violating asymmetries are ruled out to better than one percent.

★5.5 POSITRONIUM

Positronium is a hydrogen-like bound state of an electron and positron. It has much in common with meson systems, which are bound states of quarks and antiquarks, and provides striking examples of the application of conservation laws. In particular, observations on positronium can be used to verify experimentally the fundamental result (5.53) on the C-parity of fermion–antifermion systems, which has been assumed so far on the basis of theory alone.

A good approximation to the spectrum of positronium can be obtained by firstly considering only the Coulomb interaction between the electron and positron

$$V(r) = -\frac{\alpha}{r}.$$

The energy levels then have the same form as those for the hydrogen atom

$$E_n = -\frac{m_r \alpha^2}{2n^2} = -\frac{R_\infty}{2n^2}, \tag{5.59}$$

where $R_\infty = 13.6\,\text{eV}$ is the Rydberg constant, n is the principal quantum number and $m_r = m_e/2$ is the reduced mass. The level spacings are thus half those in hydrogen, which is confirmed by detecting photons from the $2p \to 2s$ transition with the predicted energy $3R_\infty / 8 = 5.1\,\text{eV}$. In each level, the orbital angular momentum L is restricted to $L \leq n-1$ as in hydrogen, while the total spin is the sum of the electron and positron spins

$$\mathbf{S} = \mathbf{S}_+ + \mathbf{S}_-, \tag{5.60}$$

so that $S = 0$ or $S = 1$. The resulting states corresponding to the $n = 1$ and $n = 2$ bands are easily deduced, and are summarized in Table 5.1 together with the corresponding values of the parity

$$P = P_{e^+} P_{e^-} (-1)^L = -(-1)^L$$

and C-parity

$$C = (-1)^{L+S},$$

TABLE 5.1 The states of positronium with the principal quantum number $n = 1$ and $n = 2$.

		J	P	C
$n = 1$	$^1S_0$	0	-1	1
	$^3S_1$	1	-1	-1
$n = 2$	$^1S_0$	0	-1	1
	$^3S_1$	1	-1	-1
	$^1P_1$	1	1	-1
	$^3P_2$	2	1	1
	$^3P_1$	1	1	1
	$^3P_0$	0	1	1

which follow from (5.35) and (5.53), respectively. The $^3S_1$ state with $n = 1$ is called *orthopositronium* and the $^1S_0$ state with $n = 1$ is called *parapositronium*.

★5.5.1 Fine structure

In the positronium spectrum, different states corresponding to the same values of n are not exactly degenerate, but are split due to small spin-dependent interactions, as shown in Figure 5.2. In particular, orthopositronium and parapositronium are split by an amount that is measured to be

$$\Delta E \equiv E(n = 1, {}^3S_1) - E(n = 1, {}^1S_0) = 8.45 \times 10^{-4}\,\text{eV}, \tag{5.61}$$

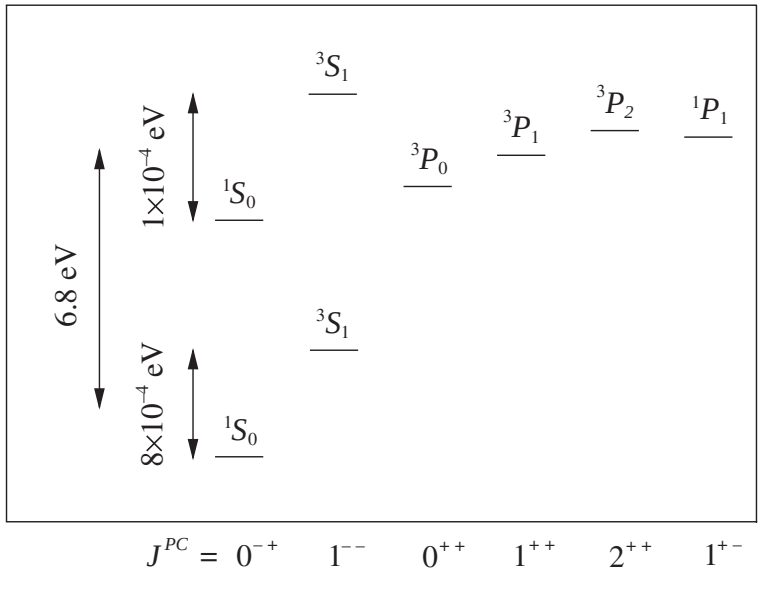


Figure 5.2 The energy levels of positronium corresponding to principal quantum numbers $n = 1$ and $n = 2$.

so that parapositronium is the ground state. This splitting is a sum of two contributions. The first of these arises because the magnetic moment of the positron $\boldsymbol{\mu}_+ = e\mathbf{S}_+/m$ gives rise to a magnetic field that interacts with the magnetic moment of the electron $\boldsymbol{\mu}_- = -e\mathbf{S}_-/m$. A detailed calculation shows that the resulting shifts in the $n = 1$ energy levels are given by

$$\delta E = -\frac{8}{3}\left(\frac{\mu_0}{4\pi}\right)\frac{\langle \boldsymbol{\mu}_+ \cdot \boldsymbol{\mu}_- \rangle}{a_0^3}, \tag{5.62}$$

where μ_0 is the magnetic permeability of free space ($\mu_0 = 1/\varepsilon_0$ in natural units), $a_0 = 2/\alpha m$ is the analogue of the Bohr radius and the expectation value is given by

$$\begin{aligned}\langle \boldsymbol{\mu}_+ \cdot \boldsymbol{\mu}_- \rangle &= -\frac{e^2}{m^2}\langle \mathbf{S}_+ \cdot \mathbf{S}_- \rangle = -\frac{e^2}{2m^2}\langle \mathbf{S}^2 - \mathbf{S}_+^2 - \mathbf{S}_-^2 \rangle \\ &= -\frac{e^2}{2m^2}\left[S(S+1) - \frac{3}{2}\right].\end{aligned}$$

Straightforward substitution in (5.62) then gives

$$\Delta E_1 = 2\alpha^2 R_\infty/3 \tag{5.63}$$

as the contribution to the splitting (5.61), which is just over one-half of the measured value. The corresponding effect also exists in hydrogen, but is very much less important because the magnetic moment of the proton is very much smaller than that of the positron.

The second contribution to the energy shift has no analogue in hydrogen. In the latter, the $e^- p$ interaction is due to one-photon exchange, which gives rise to the spin–spin and spin–orbit interactions as well as the Coulomb force. While the same diagram occurs for e^+e^- (Figure 5.3(a)), there is an additional 'annihilation diagram' (Figure 5.3(b)). Because annihilation occurs at a point, the contribution from the latter will be proportional to $|\psi_{nlm}(0)|^2$ and will vanish except for S waves; since the photon has spin-1 and angular momentum is conserved, only e^+e^- states with $J = 1$ will be affected. It thus shifts the energies of the ${}^3S_1$ states only. A perturbation theory calculation, beyond the scope of this book, gives a contribution of $\Delta E_2 = \alpha^2 R_\infty/2$ for orthopositronium, and adding this to (5.63) gives

$$\Delta E = \Delta E_1 + \Delta E_2 = \frac{2}{3}\alpha^2 R_\infty + \frac{1}{2}\alpha^2 R_\infty = \frac{7}{6}\alpha^2 R_\infty$$

for the orthopositronium and parapositronium splitting (5.61). This prediction is in excellent agreement with the experimental value, thus confirming both the existence of the annihilation term and the Dirac values of the magnetic moments used. The latter are also confirmed by studying P-wave splittings, which depend on both spin–spin and spin–orbit interactions and have been measured for the three ${}^3P$ states with $n = 2$.

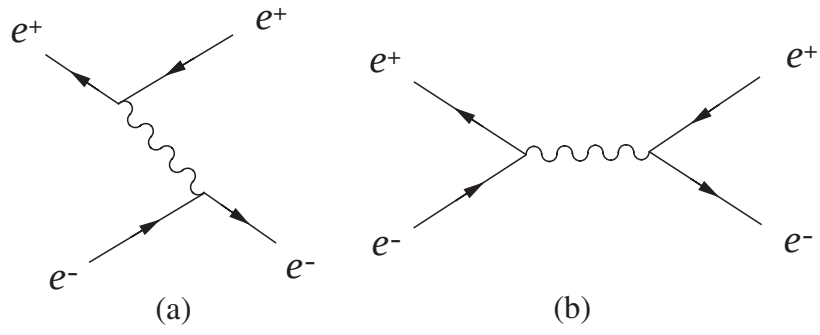


Figure 5.3 Contributions to the e^+e^- interaction from (a) one-photon exchange and (b) one-photon annihilation.

★5.5.2 C-parity and annihilations

The excited states of positronium can decay to lower-lying states by photon emission, as in hydrogen. In contrast to hydrogen, however, the ground state is itself unstable, and can decay via electron–positron annihilation. Because these annihilations only occur when the electron and positron are very close together (see Problem 1.5), the rate is to a good approximation proportional to the squared wavefunction at the origin $|\psi_{nlm}(0)|^2$, which vanishes for all except S waves. The main features of the decay can be understood by considering C conservation. The C-parities of orthopositronium ($^3S_1$) and parapositronium ($^3S_0$) are -1 and $+1$, respectively, while $C=(-1)^N$ for a final state of N photons. Thus, while the two-photon decay

$$^1S_0 \to \gamma + \gamma \tag{5.64}$$

of parapositronium is allowed, the decay $^1S_0 \to 3\gamma$ is forbidden. Searches for the latter have been unsuccessful, giving an experimental upper limit on the ratio of decay rates

$$\frac{\Gamma(^1S_0 \to 3\gamma)}{\Gamma(^1S_0 \to 2\gamma)} < 1 \times 10^{-5}.$$

This should be compared with a value of order α that would be expected if both decays were allowed, since two-photon annihilations can take place via the second-order diagrams of Figure 1.9, while three-photon annihilations can only proceed via third-order diagrams like Figure 1.10. Correspondingly, for orthopositronium the decay $^3S_1 \to \gamma\gamma$ is forbidden by C conservation, while the third-order process

$$^3S_1 \to \gamma + \gamma + \gamma \tag{5.65}$$

is allowed. Thus we would expect the ratio of the lifetimes of parapositronium and orthopositronium to be

$$\frac{\tau(^1S_0; n=1)}{\tau(^3S_1; n=1)} = O(\alpha), \tag{5.66}$$

where we allow the right-hand side to be interpreted also as $O(\alpha / 2\pi)$, as discussed in Section 1.3.3.[9] Experimentally,

$$\frac{\tau(^1S_0; n=1)}{\tau(^3S_1; n=1)} = \frac{1.25 \times 10^{-10}\,\text{s}}{1.42 \times 10^{-7}\,\text{s}} = 0.88 \times 10^{-3},$$

in reasonable agreement with (5.66) and in precise agreement with more detailed theoretical calculations.

★5.6 TIME REVERSAL

Time-reversal invariance, or T invariance, is defined as invariance under the transformation

$$t \to t' = -t, \tag{5.67}$$

leaving all position vectors unchanged. Like parity and charge conjugation invariance, it is a symmetry of the strong and electromagnetic interactions, but is violated by the weak interactions. However, unlike parity and charge conjugation, there is no associated quantum number that is conserved when weak interactions are neglected. To understand this we consider the transformation of a single-particle wavefunction, which must satisfy

$$|\psi(\mathbf{r}, t)|^2 \xrightarrow{T} \left|\psi'(\mathbf{r}, t)\right|^2 = |\psi(\mathbf{r}, -t)|^2 \tag{5.68}$$

if the system is T-invariant, so that the probability of finding the particle at position $\mathbf{r}$ at time $-t$ becomes the probability of finding it at position $\mathbf{r}$ at time t in the transformed system. In addition, since linear and angular momenta change sign under (5.67) in classical mechanics, we would expect the same result,

$$\mathbf{p} \xrightarrow{T} \mathbf{p}' = -\mathbf{p}; \quad \mathbf{J} \xrightarrow{T} \mathbf{J}' = -\mathbf{J}, \tag{5.69}$$

to hold in quantum mechanics by the correspondence principle. Hence a free-particle wavefunction

$$\psi_{\mathbf{p}}(\mathbf{r}, t) = \exp[i(\mathbf{p} \cdot \mathbf{r} - Et)],$$

corresponding to momentum $\mathbf{p}$ and energy $E = p^2/2m$, must transform into a wavefunction corresponding to momentum $-\mathbf{p}$ and energy E, i.e.

$$\psi_{\mathbf{p}}(\mathbf{r}, t) \xrightarrow{T} \psi'_{\mathbf{p}}(\mathbf{r}, t) = \psi_{-\mathbf{p}}(\mathbf{r}, t) = \exp[-i(\mathbf{p} \cdot \mathbf{r} + Et)]. \tag{5.70}$$

[9] See the discussion following Equation (1.20).

A suitable transformation that satisfies both (5.68) and (5.70) is

$$\psi(\mathbf{r},t) \xrightarrow{T} \psi'(\mathbf{r},t) = \psi^*(\mathbf{r},-t) \equiv \hat{T}\,\psi(\mathbf{r},t), \tag{5.71}$$

where we have introduced the time reversal operator $\hat{T}$ by analogy with the parity operator $\hat{P}$ introduced in Equation (5.32a). However, quantum mechanical operators $\hat{O}$ that correspond to physical observables must be both linear

$$\hat{O}(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1(\hat{O}\,\psi_1) + \alpha_2(\hat{O}\,\psi_2) \tag{5.72a}$$

and Hermitian

$$\int \mathrm{d}x(\hat{O}\psi_1)^*\psi_2 = \int \mathrm{d}x\psi_1^*(\hat{O}\psi_2), \tag{5.72b}$$

where $\psi_{1,2}$ are arbitrary wavefunctions and $\alpha_{1,2}$ are arbitrary complex numbers.[10] In contrast, the definition (5.71) implies

$$\hat{T}(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1^*(\hat{T}\psi_1) + \alpha_2^*(\hat{T}\psi_2) \neq \alpha_1(\hat{T}\psi_1) + \alpha_2(\hat{T}\psi_2)$$

for complex α_1 and α_2, and one easily verifies that (5.72b) is also not satisfied by $\hat{T}$. Thus the time-reversal operator does not correspond to a physical observable and there is no observable analogous to parity that is conserved as a consequence of T invariance.

Although T invariance does not give rise to a conservation law, it does lead to a relation between any reaction and the 'time-reversed' process related to it by (5.67). This is familiar in classical mechanics where, because Newton's equations are second order in time, they are invariant under time reversal $t \to -t$. Consequently, for every process corresponding to a given set of particle trajectories, there is a corresponding process in which all the trajectories are reversed. A similar result holds in quantum mechanics, and relates reactions like

$$a(\mathbf{p}_a, m_a) + b(\mathbf{p}_b, m_b) \to c(\mathbf{p}_c\,, m_c) + d(\mathbf{p}_d\,, m_d) \tag{5.73a}$$

to their time-reversed counterparts

$$c(-\mathbf{p}_c\,, -m_c) + d(-\mathbf{p}_d, -m_d) \to a(-\mathbf{p}_a\,, -m_a) + b(-\mathbf{p}_b\,, -m_b), \tag{5.73b}$$

in which the initial and final states are interchanged and the particle momenta ($\mathbf{p}_a$, etc.) and z components of their spins (m_a, etc.) are reversed in accordance with (5.69). In particular, if weak interactions are neglected, the rates for reactions (5.73a) and (5.73b) must be equal.

[10] Linearity ensures that the superposition principle holds, while hermiticity ensures that the the eigenvalues of $\hat{O}$, i.e. the observed values, are real.

★5.6.1 Principle of detailed balance

A more useful relation between reaction rates can be obtained if we combine time reversal with parity invariance. Under the parity transformation (5.30), momenta $\mathbf{p}$ change sign while orbital angular momenta $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ do not. If we assume the same behaviour holds for spin angular momenta, then

$$\mathbf{p} \xrightarrow{P} \mathbf{p}' = -\mathbf{p}\,; \quad \mathbf{J} \xrightarrow{P} \mathbf{J}' = \mathbf{J} \tag{5.74}$$

under parity. The parity-transformed reaction corresponding to (5.73b) is

$$c(\mathbf{p}_c\,, -m_c) + d(\mathbf{p}_d\,, -m_d) \to a(\mathbf{p}_a\,, -m_a) + b(\mathbf{p}_b\,, -m_b) \tag{5.73c}$$

so that if both P and T invariance holds, all three reactions (5.73a), (5.73b) and (5.73c) must have the same rate. If we average over all spin projections

$$m_i = -S_i, -S_i + 1, \ldots S_i \quad (i = a, b, c, d),$$

where S_i is the spin of particle i, then reactions (5.73a) and (5.73c) differ only by the interchange of initial and final states. Consequently, the rates for the reactions

$$i \equiv a(\mathbf{p}_a) + b(\mathbf{p}_b) \leftrightarrow c(\mathbf{p}_c) + d(\mathbf{p}_d) \equiv f \tag{5.75}$$

should be equal, provided that we average over all possible spin states. This relation is called the *principle of detailed balance*, and has been confirmed experimentally in a variety of strong and electromagnetic reactions. Here we will assume its validity and illustrate its use in a classic application.

★5.6.2 Spin of the charged pion

Although we have assumed throughout that pions have spin zero, we have not so far discussed any experimental evidence for this assignment. The first evidence was provided by using detailed balance to relate the cross-section for the reaction

$$p + p \to \pi^+ + d \tag{5.76a}$$

to the time-reversed process

$$\pi^+ + d \to p + p \tag{5.76b}$$

at the same centre-of-mass energy. The experiments were carried out using unpolarized beams and targets, so that an *average* over the possible initial spin states is required. However, all final state particles were detected, irrespective of their spin projections, so the final spin states are *summed*. Because of this, the spin-averaged reaction rates must be multiplied by a factor $(2S_f + 1)$ for each final state particle of spin S_f. In

addition, flux and phase space factors must be taken into account. When this is done, the differential cross-sections are given by[11]

$$\frac{\mathrm{d}\sigma(p+p\to\pi^+ + d)}{\mathrm{d}\cos\theta} = \frac{(2S_\pi+1)(2S_d+1)}{2\pi}\frac{p_\pi^2}{\upsilon_i\upsilon_f}\overline{\left|\mathcal{M}_{if}\right|^2} \tag{5.77a}$$

and

$$\frac{\mathrm{d}\sigma(\pi^+ + d\to p+p)}{\mathrm{d}\cos\theta} = \frac{(2S_p+1)^2}{2\pi}\frac{p_p^2}{\upsilon_i\upsilon_f}\overline{\left|\mathcal{M}_{fi}\right|^2}, \tag{5.77b}$$

where θ is the scattering angle, p_π and p_p are the magnitudes of the pion and proton momenta and υ_i, υ_f are the relative velocities of the pp and $\pi^+ d$ pairs, respectively, in the centre-of-mass frame. Here $\overline{\left|\mathcal{M}_{fi}\right|^2}$ is the spin average of the squared scattering amplitude $\left|\mathcal{M}_{fi}\right|^2$, which is in turn proportional to the transition rate between specific quantum states i and f. In terms of it, the principle of detailed balance takes the form

$$\overline{\left|\mathcal{M}_{if}\right|^2} = \overline{\left|\mathcal{M}_{fi}\right|^2}, \tag{5.78}$$

and taking the ratio of cross-sections (5.77b) gives

$$\frac{\mathrm{d}\sigma(p+p\to\pi^+ + d)/\,\mathrm{d}\cos\theta}{\mathrm{d}\sigma(\pi^+ + d\to p+p)/\,\mathrm{d}\cos\theta} = 3\left(\frac{p_\pi}{p_p}\right)^2(2S_\pi+1),$$

where we have set S_p and S_d equal to their known values of $\frac{1}{2}$ and 1, respectively. This ratio is sensitive to the pion spin, and measurements of both cross-sections in the early 1950s clearly established the familiar result $S_\pi = 0$.

PROBLEMS 5

5.1 Generalize the discussion leading to the conservation of linear momentum (5.8) to the case of an N-particle system described by the wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_N)$.

5.2 Use the standard commutation relations for angular momentum operators to show that L and S remain good quantum numbers if the spin-dependent forces arise from a simple spin–orbit interaction, i.e. if

$$H = H_0 + \alpha\hat{\mathbf{L}}\cdot\hat{\mathbf{S}},$$

where

$$\left[H_0, \hat{\mathbf{L}}\right] = \left[H_0, \hat{\mathbf{S}}\right] = 0$$

and α is a constant.

5.3 The deuteron is a bound state of two nucleons with spin-1 and positive parity. Show that it may only exist in the $^3S_1$ and $^3D_1$ states of the np system.

[11] These formulas are derived in Appendix B.

5.4 If $p\bar{p}$ annihilation at rest proceeds via S states, explain why the reaction $p + \bar{p} \to \pi^0 + \pi^0$ is forbidden as a strong or electromagnetic interaction.

5.5 The ground state baryons have $L_{12} = L_3 = 0$ in the simple quark model, where $\mathbf{L}_{12}$ and $\mathbf{L}_3$ are the orbital angular momenta defined in Figure 5.1. If the next lowest-lying band of baryons has $L_{12} = 1, L_3 = 0$ *or* $L_{12} = 0, L_3 = 1$, what values of the spin J and parity P occur?

5.6 Suppose that an intrinsic C-parity factor C_a is introduced into (5.50b), which then becomes

$$\hat{C}\,|a, \Psi\rangle = C_a\,|\bar{a}, \Psi\rangle\,.$$

Show that the eigenvalue corresponding to any eigenstate of $\hat{C}$ is independent of C_a, so that C_a cannot be measured.

5.7 The $\eta(549)$ meson has spin-0 and is observed to decay to three-pion final states by the electromagnetic processes (5.57b) and (5.57c). Use this information to deduce the parity of η, and hence explain why the decays $\eta \to \pi^+ + \pi^-$ and $\eta \to \pi^0 + \pi^0$ have never been observed.

6

The Quark Model

> Young man, if I could remember the names of all these particles, I would have been a botanist.
>
> *Enrico Fermi*

The existence and properties of quarks were first inferred from the study of hadron spectroscopy, and the close correspondence between the experimentally observed hadrons and those predicted by the quark model remains one of the strongest reasons for our belief in the existence of quarks. In this chapter, we will explore this correspondence in more detail, and use it to infer some fundamental properties of the quarks and the strong interactions between them. In particular, we will present evidence that the strong force is 'flavour independent', by which we mean that its properties do not depend on the flavour $u, d, s, \ldots$ of the quark on which it is acting, so that, for example, the strong forces between $qq = ud, us, sd, \ldots$ pairs are identical.[1] We will also make estimates of quark masses and introduce a new property of quarks, called *colour*, which lies at the heart of our understanding of strong interactions.

We shall not attempt a comprehensive discussion of the quark model,[2] but will concentrate on just four particularly important topics. To begin, we introduce an approximate symmetry of hadrons, called *isospin symmetry*, which has its origins in flavour independence and the almost equal masses of the u and d quarks. We next consider hadrons made of the light quarks u, d, s, and their antiquarks, restricting ourselves for simplicity to the lightest particles, in which the quarks have zero orbital

[1] The same principle applies to quark–antiquark forces, but these are not identical to quark–quark forces because in the former case annihilations can occur.
[2] A fuller description may be found in Yao *et al*. (2006).

Particle Physics Third Edition B. R. Martin and G. Shaw

angular momentum. In this case, the quark model is strikingly successful in accounting for the observed states, but apparently violates the Pauli principle. This paradox is removed in the next section, where we introduce colour, with emphasis on its role in quark confinement. Finally, we consider the spectra of the quark–antiquark systems $c\bar{c}$ and $b\bar{b}$, called *charmonium* and *bottomium*, respectively. These are interesting because they are essentially nonrelativistic, and a simple discussion based on the Schrödinger equation leads to an understanding of the level structure and a determination of the quark–antiquark potential.

6.1 ISOSPIN SYMMETRY

Perhaps the most striking fact about hadrons is that they occur in families of particles with approximately equal masses. Within a given family, all particles have the same spin, parity, baryon number, strangeness, charm and bottom, but differ in their electric charges. Many examples of this were seen in Chapter 3, including the nucleons

$$p(938) = uud, \quad n(940) = udd, \tag{6.1}$$

and the K mesons

$$K^+(494) = u\bar{s}, \quad K^0(498) = d\bar{s}. \tag{6.2}$$

This behaviour reflects a symmetry between u and d quarks: if

$$m_u = m_d \tag{6.3}$$

and if the forces acting on u and d quarks were exactly equal, then replacing the u quark in the K^+ by a d quark would produce another particle – presumably the K^0 – with exactly the same mass. The observed masses are slightly different, indicating that this symmetry in practice is not exact. The strong forces on u and d quarks are the same, as far as we know, but the electromagnetic forces are different because the quarks have different electric charges. In addition, the d quark is a few MeV/c^2 heavier than the u quark, as we shall show below. However, this quark mass difference is small compared with typical hadron masses, and electromagnetic forces are weak compared with strong forces. Consequently, *isospin symmetry*, as it is called, is a good approximation and greatly simplifies the interpretation of hadron physics. Families of particles like (6.1) and (6.2) are called *isospin multiplets*.

6.1.1 Isospin quantum numbers

In order to formulate isospin symmetry more precisely, we introduce three more quantum numbers that are conserved in strong interactions. Two of these are not really new, but are combinations of ones previously introduced. The first is the *hypercharge* Y defined by

$$Y \equiv B + S + C + \tilde{B} + T, \tag{6.4}$$

where, as usual, B, S, C, $\tilde{B}$ and T are the baryon number, strangeness, charm, bottom and top, respectively. As these quantum numbers have the same values for all members of a given isospin multiplet, so does the hypercharge. The second combination is defined by

$$I_3 \equiv Q - Y/2, \tag{6.5}$$

where Q is the electric charge. Alternatively, I_3 can be expressed in terms of the quark numbers

$$N_u = N(u) - N(\bar{u}) \quad \text{and} \quad N_d = N(d) - N(\bar{d}),$$

and we leave it as an exercise for the reader to show that (6.5) is equivalent to

$$I_3 = \frac{1}{2}(N_u - N_d) \tag{6.6}$$

using Equations (3.2) to (3.5).

Whereas the hypercharge Y takes the same value for each member of a multiplet, I_3 takes different values, and if we call the maximum value within a multiplet I, i.e.

$$I \equiv (I_3)_{max}, \tag{6.7}$$

then we find that all observed multiplets contain precisely $(2I + 1)$ members, with I_3 values

$$I_3 = I, I - 1, \ldots, -I. \tag{6.8}$$

This is very reminiscent of the formalism for the spin quantum numbers S and S_3, where

$$S_3 = S, S - 1, \ldots, -S, \tag{6.9}$$

and, by analogy, I is called the *isospin* quantum number and I_3 the third component of isospin. This is illustrated in Table 6.1, where the values of these various quantum numbers are listed for four multiplets that we met in Chapter 3, including the 'isosinglet' Λ with $I = I_3 = 0$. The theory of isospin symmetry is formally very similar to the theory of angular momentum. We shall not pursue this in detail here, but argue by analogy with the latter. (For the interested reader a more formal treatment of isospin is given in Appendix C.) In particular, the rules for addition of isospins are identical to those for the addition of ordinary spins or orbital angular momenta. Thus the sum of two isospins I^a and I^b can give the values

$$I^a + I^b,\ I^a + I^b - 1,\ \ldots, |I^a - I^b|, \tag{6.10a}$$

while the sum of their third components is

$$I_3 = I_3^a + I_3^b. \tag{6.10b}$$

The latter result follows directly from the definition (6.5), since both Q and Y are additive quantum numbers.

TABLE 6.1 Values of the baryon number B, hypercharge Y, charge Q and isospin quantum numbers I and I_3 for some typical hadron isospin multiplets.

Particle	B	Y	Q	I_3	I
$\Lambda(1116)$	1	0	0	0	0
$p(938)$	1	1	1	1/2	1/2
$n(940)$	1	1	0	$-1/2$	1/2
$K^+(494)$	0	1	1	1/2	1/2
$K^0(498)$	0	1	0	$-1/2$	1/2
$\pi^+(140)$	0	0	1	1	1
$\pi^0(135)$	0	0	0	0	1
$\pi^-(140)$	0	0	-1	-1	1

6.1.2 Allowed quantum numbers

The isospin quantum numbers I and I_3 could in principle take on any integer or half-integer values subject to (6.8). In practice, the only values observed are those that can arise from the simple quark combinations $q\bar{q}$ and qqq, to which we now turn.

TABLE 6.2 Values of the baryon number B, hypercharge Y, charge Q and isospin quantum numbers I and I_3 for quarks. For the corresponding antiquarks, B, Y, Q and I_3 are reversed in sign, while I remains unchanged.

Quark	B	Y	Q	I_3	I
d	1/3	1/3	$-1/3$	$-1/2$	1/2
u	1/3	1/3	2/3	1/2	1/2
s	1/3	$-2/3$	$-1/3$	0	0
c	1/3	4/3	2/3	0	0
b	1/3	$-2/3$	$-1/3$	0	0
t	1/3	4/3	2/3	0	0

The values of Y and I_3 for the quarks themselves can be deduced from Equations (6.4) and (6.6) using the values given in Table 3.1, and are given in Table 6.2. In this table, the u and d quarks have been assigned to an isodoublet (i.e. an isospin doublet with $I = \frac{1}{2}$), since they have $I_3 = \frac{1}{2}$ and $-\frac{1}{2}$, respectively, and approximately equal masses. For antiquarks, the sign of I_3 is reversed, so that the $\bar{d}$ and $\bar{u}$ antiquarks have

$I_3 = \frac{1}{2}$ and $-\frac{1}{2}$, respectively, and are also assigned to an isodoublet. The remaining quarks and antiquarks all have $I_3 = 0$, with no partners of similar mass, and are assigned to isosinglets (i.e. $I = 0$). With these assignments, the isospins of hadrons may be deduced using the rules for the addition of isospin given in (6.10).

The allowed isospins in the quark model are correlated to the values of the strangeness, charm, bottom and top. For simplicity, we will confine ourselves to zero values of charm, bottom and top, since these are zero for most known hadrons. For baryons the possibilities are

$$sss,\ ssl_i,\ sl_il_j,\ l_il_jl_k, \tag{6.11}$$

where s is the strange quark with $I = 0$, and $l_i = u$ or d is a quark with $I = \frac{1}{2}$. The s quark contributes nothing to the total isospin, so that the states sss with $S = -3$ and ssl_i with $S = -2$ must have $I = 0$ and $I = \frac{1}{2}$, respectively. The isospin of the states with $S = -1$ is just the sum of the isospins of the quarks l_i and l_j, both with $I = \frac{1}{2}$. By Equation (6.10a) this is $I = 0$ or 1. Finally, for states $l_il_jl_k$ with $S = 0$, we have to combine a third $I = \frac{1}{2}$ isospin with the latter results, giving $I = \frac{1}{2}$ or $I = \frac{3}{2}$, respectively. These values are given in Table 6.3(a). The corresponding values for mesons with $C = \tilde{B} = T = 0$, summarized in Table 6.3(b), are left as an exercise for the reader. These values are the only ones allowed if mesons are $q\bar{q}$ and baryons qqq states, provided $C = \tilde{B} = T = 0$, and so far they are the only values observed in nature.

TABLE 6.3 Values of the isospin I and strangeness S allowed in the quark model for hadrons with $C = \tilde{B} = T = 0$.

(a) Baryons		(b) Mesons	
S	I	S	I
0	3/2, 1/2	1	1/2
−1	1, 0	0	1, 0
−2	1/2	−1	1/2
−3	0		

6.1.3 An example: the sigma (Σ) baryons

We will illustrate the use of internal quantum numbers, particularly isospin, by considering the Σ baryons.[3] The $\Sigma^+(1189)$ is a particle that can be produced in the strong interaction

$$K^- + p \to \pi^- + \Sigma^+ \tag{6.12}$$

and which decays to pions and nucleons,

$$\Sigma^+ \to \pi^+ + n, \quad \pi^0 + p, \tag{6.13}$$

[3] Baryons with nonzero strangeness are also called *hyperons*.

with a lifetime of 0.8×10^{-10} s, typical of weak interactions. By applying the conservation laws of baryon number, strangeness, charm, bottom and top to (6.12), and using the known values for $K^-(B=0, S=-1, C=\tilde{B}=T=0)$, $p(B=1, S=C=\tilde{B}=T=0)$ and $\pi^-(B=0, S=C=\tilde{B}=T=0)$, it follows that for the Σ^+,

$$B=1; S=-1; C=\tilde{B}=T=0 \tag{6.14}$$

and hence, from (6.4) and (6.5), $Y=0$ and $I_3=1$.

If a particle is found with a nonzero value of I_3, it follows as a simple consequence of isospin symmetry that other particles belonging to the same isospin multiplet must exist. Thus the Σ^+ is a member of a multiplet in which all the particles have the quantum numbers (6.14), and hence, by (6.7), $Q=I_3$. Within this multiplet it follows that there must be states with charges

$$Q=I_3=I,\ I-1,\ \ldots, -I,$$

where I must be an integer greater than or equal to one, since we already know that the Σ^+ has $I_3=1$. The smallest multiplet possible is therefore an isotriplet $\Sigma^+\ \Sigma^0 \Sigma^-$. Other partners would exist in addition to these if $I \geq 2$, but this would contradict the quark model (see Table 6.3(a)).

The above argument shows that isospin symmetry unambiguously predicts the existence of Σ^0 and Σ^- particles with approximately the same mass as the Σ^+ and the quantum numbers (6.14). These particles do indeed exist and can be produced in the reactions

$$K^- + p \rightarrow \pi^0 + \Sigma^0 \tag{6.15a}$$

and

$$K^- + p \rightarrow \pi^+ + \Sigma^-, \tag{6.15b}$$

which are similar to (6.12). The $\Sigma^-(1197)$ decays by the strangeness-violating weak reaction

$$\Sigma^- \rightarrow \pi^- + n, \tag{6.16}$$

analogous to (6.13), while the $\Sigma^0(1193)$ decays by the strangeness-conserving electromagnetic reaction

$$\Sigma^0 \rightarrow \Lambda + \gamma, \tag{6.17}$$

because the $\Lambda(1116)$ has the same internal quantum numbers (6.14) as the Σ^0 and is lighter. No doubly charged partners have been found, confirming that the Σ^+, Σ^- and Σ^0 are an isotriplet, as expected from the quark model.

6.1.4 The *u*, *d* quark mass splitting

We next deduce the quark composition of the Σ baryons in order to discuss the origins of the small mass differences between them and obtain an estimate of the mass difference between the *u* and *d* quarks. From (6.14) we see that each Σ particle contains one *s* quark but no *c*, *b* or *t* quarks. The identity of the remaining quarks as *u* or *d* can be inferred from the observed electric charges, or from (6.6). Either way, we find the compositions

$$\Sigma^+(1189) = uus, \quad \Sigma^0(1193) = uds, \quad \Sigma^-(1197) = dds. \tag{6.18}$$

We now assume that the mass differences arise from quark mass differences and from electromagnetic interactions between pairs of quarks, and to obtain a rough estimate of the former we assume that the latter are proportional to the products of the quark charges. This requires that pairs of quarks are on average separated by the same distance and that their magnetic moments are proportional to their charges. This is a reasonable first approximation.[4] We then have

$$M(\Sigma^-) = M_0 + m_s + 2m_d + \delta(e_d^2 + e_d e_s + e_d e_s) = M_0 + m_s + 2m_d + \delta/3, \tag{6.19a}$$

$$\begin{aligned} M(\Sigma^0) &= M_0 + m_s + m_d + m_u + \delta(e_u e_d + e_u e_s + e_d e_s) \\ &= M_0 + m_s + m_d + m_u - \delta/3 \end{aligned} \tag{6.19b}$$

and

$$M(\Sigma^+) = M_0 + m_s + 2m_u + \delta(e_u^2 + e_u e_s + e_u e_s) = M_0 + m_s + 2m_u, \tag{6.19c}$$

where e_q are the quark charges, δ is a constant and M_0 is the contribution to the Σ masses arising from the strong interactions between the quarks. This latter contribution will be the same in all three cases if the strong interactions of *u* and *d* quarks are identical, as we have assumed. From these equations, we directly obtain

$$m_d - m_u = \frac{1}{3}\left[M(\Sigma^-) + M(\Sigma^0) - 2M(\Sigma^+)\right] = 4.0\,\mathrm{MeV/c^2}$$

as a rough estimate of the quark mass difference from this simple model. This estimate agrees with the range of values

$$2 \lesssim m_d - m_u \lesssim 4\,\mathrm{MeV/c^2} \tag{6.20}$$

found in other more sophisticated determinations, and justifies our assertion that this quark mass difference is small compared with typical hadron masses.

[4] Magnetic moments of quarks are discussed in Section 6.2.4.

6.2 THE LIGHTEST HADRONS

The lowest-lying states of any quark system have zero orbital angular momentum, and the lightest quarks are the u, d and s quarks. We therefore confine our attention to states composed of only these three quarks and/or their antiquarks, and having zero orbital angular momentum. As we shall see, there are 36 such states–18 mesons and 18 baryons–whose quantum numbers match exactly those of the lightest experimentally observed hadrons.

Since all the particles we discuss have zero charm, bottom and top, the hypercharge (6.4), which will be used repeatedly in this section, reduces to

$$Y = B + S. \tag{6.21}$$

6.2.1 The light mesons

The lightest mesons observed experimentally are a family of nine particles with spin-parity $J^p = 0^-$, called the *pseudoscalar meson nonet*; and a family of nine particles with spin-parity 1^-, called the *vector meson nonet*. The members of each nonet are conveniently summarized on the so-called *weight diagrams* of Figures 6.1(a) and (b). In these diagrams, each particle is associated with a dot at a position corresponding to its quantum numbers (Y, I_3) and if two or more dots coincide they are displaced slightly for clarity. The pseudoscalar nonet of Figure 6.1(a) includes the familiar pions and kaons, together with the $\eta(549)$ meson, which was discussed in Section 5.4.1, and a heavier meson called the $\eta'(958)$. The vector mesons (Figure 6.1(b)) are all resonances and include the K^* mesons discussed in Section 3.5.

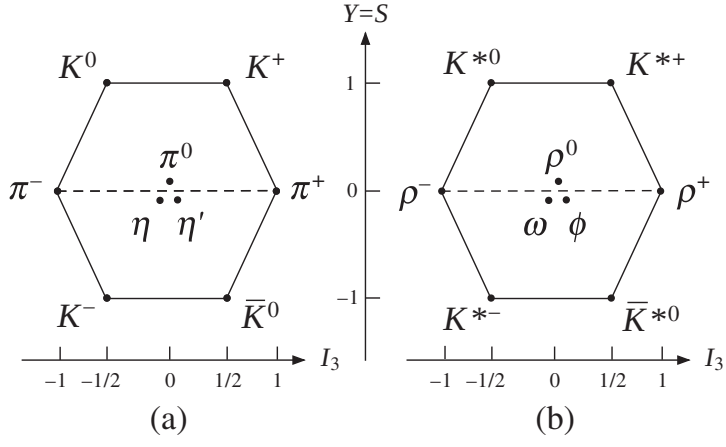


Figure 6.1 Weight diagrams for (a) the 0^- meson nonet and (b) the 1^- meson nonet.

Different members of the same isospin multiplet, like the pions or the kaons in Figure 6.1(a), lie along horizontal lines in these weight diagrams. On comparing Figures 6.1(a) and (b) we see that precisely the same set of internal quantum number values occurs for both nonets. These striking results are exactly what is expected if we are dealing with mesons $M = a\bar{b}$ made of the lightest quarks $a = u, d, s$ and antiquarks

$\bar{b} = \bar{u}, \bar{d}, \bar{s}$ with zero orbital angular momentum $L = 0$. For $L = 0$, the parity of the meson is

$$P = P_a P_{\bar{b}} (-1)^L = (-1)^{L+1} = -1$$

by (5.39a), while the spin of the mesons is just the sum of the quark and antiquark spins, and can be 0 or 1. Since there are nine possible combinations $a\bar{b}$, we arrive at two nonets of mesons with spin-parities 0^- and 1^-, as is experimentally observed. The Y and I_3 values for each quark combination $a\bar{b}$ are obtained by adding the values for a and $\bar{b}$ listed in Table 6.2. The resulting values correspond to those for the observed mesons with the quark assignments listed in Table 6.4. In most cases these are completely unambiguous: for example, there is only one quark–antiquark pair $u\bar{s}$ with the values $I_3 = \frac{1}{2}$, $Y = S = 1$ corresponding to the K^+ meson in the 0^- nonet and the K^{*+} meson in the 1^- nonet. However, there are three particles in each nonet with

$$I_3 = Y = S = 0,$$

which could correspond to any of the particle–antiparticle pairs

$$u\bar{u}, \quad d\bar{d}, \quad \text{or} \quad s\bar{s}, \tag{6.22}$$

or to any linear combination of them. It can be shown that the linear combinations of (6.22) corresponding to definite isospins $I = 0$ or $I = 1$ are[5]

$$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}) \quad (I = 1, I_3 = 0), \tag{6.23}$$

TABLE 6.4 The states of the light $L = 0$ meson nonets.

Quark content	0^- state	1^- state	I_3	I	$Y = S$
$u\bar{s}$	$K^+(494)$	$K^{*+}(892)$	1/2	1/2	1
$d\bar{s}$	$K^0(498)$	$K^{*0}(896)$	−1/2	1/2	1
$u\bar{d}$	$\pi^+(140)$	$\rho^+(768)^{\#}$	1	1	0
$\frac{(u\bar{u} - d\bar{d})}{\sqrt{2}}$	$\pi^0(135)$	$\rho^0(768)^{\#}$	0	1	0
$d\bar{u}$	$\pi^-(140)$	$\rho^-(768)^{\#}$	−1	1	0
$s\bar{d}$	$\bar{K}^0(498)$	$\bar{K}^{*0}(896)$	1/2	1/2	−1
$s\bar{u}$	$K^-(494)$	$K^{*-}(892)$	−1/2	1/2	−1
See text	$\eta(549)$	$\omega(782)$	0	0	0
See text	$\eta'(958)$	$\phi(1019)$	0	0	0

[#] The measured mass difference between the neutral and charged ρ mesons is $m(\rho^0) - m(\rho^+) = 0.3 \pm 2.2\,\text{MeV}/c^2$.

[5] These results are derived as Equations (C.24b) and (C.26) in Appendix C.

$$\frac{1}{\sqrt{2}}(u\bar{u}+d\bar{d}) \quad (I=0, I_3=0) \tag{6.24a}$$

and

$$s\bar{s} \quad (I=0, I_3=0), \tag{6.24b}$$

so that the quark assignments of the $I=1$, $I_3=S=0$ states corresponding to the π^0 and ρ^0 mesons in Table 6.4 are unique. In contrast, the assignment of the $I=I_3=S=0$ particles to the quark states (6.24a) and (6.24b), or to linear combinations of them, is more complicated, and we will not pursue this further.[6]

6.2.2 The light baryons

Families of hadrons with the same baryon number, spin and parity, like the 0^- and 1^- meson nonets, are called *supermultiplets*, and all hadrons belong to such families. For mesons, these supermultiplets are all nonets, whereas baryon supermultiplets can have 1, 8 or 10 members and are called *singlets, octets* and *decuplets*, respectively. Here we shall consider only the lightest baryon supermultiplets that are experimentally observed, which are an octet of $J^P=\frac{1}{2}^+$ particles and a decuplet of $J^P=\frac{3}{2}^+$ particles. These are summarized in the weight diagrams of Figure 6.2, where for baryons the hypercharge is given by $Y=S+1$ by Equation (6.21). The $\frac{1}{2}^+$ octet includes the nucleons, the $\Lambda(1116)$ and the Σ particles discussed in Section 6.1.3, together with the xi or 'cascade' particles Ξ^0 and Ξ^-, which decay by weak interactions with lifetimes of order 10^{-10} s. The $\frac{3}{2}^+$ states are all resonances, except for the $\Omega^-(1672)$, which decays by the weak interaction[7] with a lifetime of order 10^{-10} s.

We shall now show that the remarkable patterns of experimental states shown in Figures 6.2(a) and (b) are predicted by the quark model if we assume that the combined space and spin wavefunctions are *symmetric* under the interchange of like quarks. At first sight this seems impossible, since quarks are spin-$\frac{1}{2}$ particles and so are *fermions*. The explanation will emerge in Section 6.3, where we introduce the fundamental concept of colour. Here we shall simply assume that only overall symmetric space–spin wavefunctions are allowed, and show that this assumption leads to the octet and decuplet states which are observed.

The baryons of Figure 6.2 correspond to three-quark states $B=abc$ containing only the light quarks u, d, s with zero orbital angular momentum. Hence their parity

$$P=P_aP_bP_c=1, \tag{6.25}$$

by (5.39b), while the baryon spin is equal to the sum of the quark spins. In addition, for simple systems the lowest-lying states almost invariably have spatial wavefunctions that are symmetric under the interchange of like particles, and we will assume this to be the case here. Since we are assuming symmetric space–spin wavefunctions, this

[6] See, for example, pp. 155–157 of Perkins (1987).

[7] This particle has strangeness $S=-3$. It cannot decay by strong interactions, unlike the other members of the decuplet, because there is no lighter state with $S=-3$ to which it can decay.

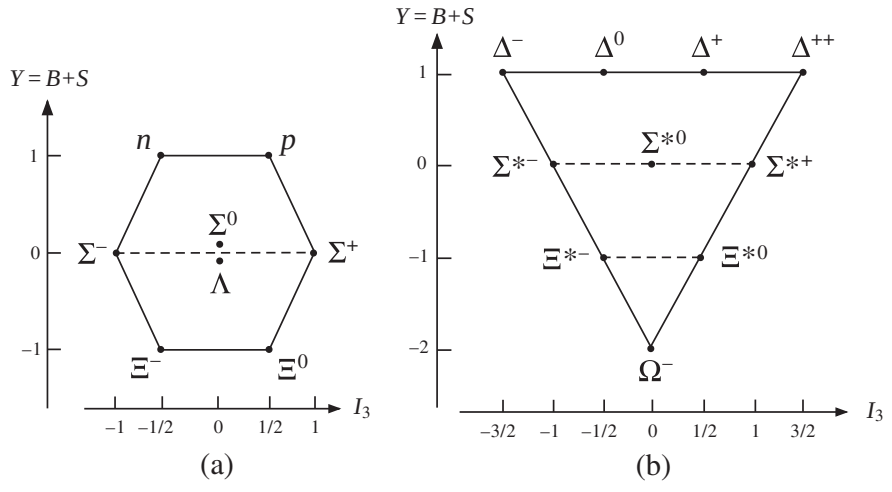


Figure 6.2 Weight diagrams for (a) the $J^P = \frac{1}{2}^+$ octet of light baryons and (b) the $J^P = \frac{3}{2}^+$ baryon decuplet.

implies that the spin wavefunction must also be symmetric under the exchange of like quarks; i.e. any like pair of quarks *aa* must have spin-1, corresponding to parallel quark spins and the symmetric spin wavefunction (5.43), and may not have spin-0, corresponding to antiparallel spins and the antisymmetric spin wavefunction (5.44). Hence, in the six combinations

$$uud,\ uus,\ ddu,\ dds,\ ssu,\ ssd \tag{6.26a}$$

of the form $B = aab$, where $a \neq b$, the *aa* pair must be in a spin-1 state. Adding the spin of the third quark gives

$$J = S = \frac{1}{2},\ \frac{3}{2} \tag{6.26b}$$

as the possible values for the baryon spin, since the orbital angular momentum is zero. For the three combinations

$$uuu\,,\quad ddd\,,\quad sss \tag{6.27a}$$

the only way to have parallel spins for all pairs is for all three quarks to have their spins parallel, giving

$$J = S = \frac{3}{2} \tag{6.27b}$$

as the only possibility for the baryon spin. Finally, for the remaining case

$$uds, \tag{6.28a}$$

in which all the quarks are different, there are two spin-$\frac{1}{2}$ baryons and one spin-$\frac{3}{2}$ baryon; i.e.

$$J = S = \frac{1}{2}, \ \frac{1}{2}, \ \frac{3}{2}. \tag{6.28b}$$

These arise because the *ud* pair can have spin-0, leading to $S = \frac{1}{2}$ on adding the spin of the *s* quark, or spin-1, leading to either $S = \frac{1}{2}$ or $S = \frac{3}{2}$ on adding the *s* quark spin.

The rest is book-keeping. All the above states have positive parity by (6.25), and on counting the various spin states in (6.26), (6.27) and (6.28) we find that there are eight with $J^P = \frac{1}{2}^+$ and ten with $J^P = \frac{3}{2}^+$, as observed experimentally.[8] Furthermore, the Y and I_3 values of the various states, obtained by adding the values for the individual quarks given in Table 6.2, correspond exactly to those of the observed states with the assignments given in Tables 6.5 and 6.6.

TABLE 6.5 The states of the $L = 0$, $\frac{1}{2}^+$ octet of light baryons.

Quark composition	Observed state	I_3	I	S
uud	$p(938)$	1/2	1/2	0
udd	$n(940)$	$-1/2$	1/2	0
uds	$\Lambda(1116)^{\#}$	0	0	-1
uus	$\Sigma^+(1189)$	1	1	-1
uds	$\Sigma^0(1193)^{\#}$	0	1	-1
dds	$\Sigma^-(1197)$	-1	1	-1
uss	$\Xi^0(1315)$	1/2	1/2	-2
dss	$\Xi^-(1321)$	$-1/2$	1/2	-2

$^{\#}$ The $\Lambda(1116)$ and the $\Sigma^0(1193)$ differ in that in the former the *ud* pair is in a spin-0 state, while in the latter it is in a spin-1 state, as discussed at the end of Section 6.2.2.

Although we have considered just the three lightest quarks, it is straightforward to extend these ideas to include more quark flavours, although the weight diagrams become increasingly complex. Nevertheless, the striking fact remains that no states have been found that lie outside the quark model scheme.

We finally comment on one detailed point that will be required in our discussion of baryon magnetic moments in Section 6.2.4 below.[9] In the argument following (6.28a), we saw that there were two *uds* states with spin-$\frac{1}{2}$, corresponding to $S = 0$ and $S = 1$ for the spin of the *ud* pair alone. These states have $Y = I_3 = 0$, corresponding to the

[8] If we had assumed that the space–spin wavefunction is antisymmetric, then we would instead have obtained a quite different spectrum (see Problem 6.6).

[9] Readers who plan to omit this starred section may go directly to Section 6.2.3.

TABLE 6.6 The states of the $L = 0$, $\frac{3}{2}^+$ decuplet of light baryons.

Quark composition	Observed state	I_3	I	S
uuu	$\Delta^{++}(1232)^{\#}$	3/2	3/2	0
uud	$\Delta^{+}(1232)^{\#}$	1/2	3/2	0
udd	$\Delta^{0}(1232)^{\#}$	$-1/2$	3/2	0
ddd	$\Delta^{-}(1232)^{\#}$	$-3/2$	3/2	0
uus	$\Sigma^{+}(1383)$	1	1	-1
uds	$\Sigma^{0}(1384)$	0	1	-1
dds	$\Sigma^{-}(1387)$	-1	1	-1
uss	$\Xi^{0}(1532)$	1/2	1/2	-2
dss	$\Xi^{-}(1535)$	$-1/2$	1/2	-2
sss	$\Omega^{-}(1672)$	0	0	-3

We do not show the small mass differences between the various charged and neutral states of the Δ as these are not all precisely determined.

observed particles $\Lambda(1116)$ and $\Sigma^0(1193)$ in Table 6.5, and the question is: 'Which is which?' To resolve this ambiguity, we recall that the Σ^0 is a member of the isotriplet

$$\Sigma^+(1189) = uus\,, \quad \Sigma^0(1193) = uds\,, \quad \Sigma^-(1197) = dds,$$

discussed in Section 6.1.3, and that the members of isospin multiplets differ in the replacement of u quarks by d quarks (or vice versa), with the wavefunctions unchanged. Since in the Σ^+ the two like quarks uu must be in the symmetric spin-1 state, it follows that in the Σ^0 the ud pair must also have spin-1. The remaining spin-$\frac{1}{2}$ uds state, in which the ud pair is in a spin singlet state, is then assigned to the $\Lambda(1116)$. These assignments are crucial in understanding the baryon magnetic moments, as we shall see in Section 6.2.4.

6.2.3 Baryon mass splittings

The mass differences between different members of a given supermultiplet are conveniently separated into the small mass differences between members of the same isospin multiplet and the large mass differences between members of different isospin multiplets. Here we will concentrate on the latter, ignoring the small mass differences within isospin multiplets;[10] i.e. we will work in the approximation that all members of an isospin multiplet have the same mass, so that, for example,

$$m_p = m_n \equiv m_N$$

for the nucleons.

From Tables 6.5 and 6.6 we see that within each supermultiplet the heaviest isospin multiplets are those that contain the most s quarks, suggesting that the main contribution to the mass differences may be the mass difference between the s quark and

[10] These have been discussed in Section 6.1.4.

the lighter u and d quarks. We therefore explore the simple model in which this is assumed to be the only contribution and the mass differences are calculated directly from the sums of the masses of the constituent quarks. For the $\frac{3}{2}^+$ decuplet this gives

$$M_\Omega - M_\Xi = M_\Xi - M_\Sigma = M_\Sigma - M_\Delta = m_s - m_{u,d}$$

and for the $\frac{1}{2}^+$ octet

$$M_\Xi - M_\Sigma = M_\Xi - M_\Lambda = M_\Lambda - M_N = m_s - m_{u,d}.$$

Thus we have six independent estimates of the mass difference between the strange quark and the u, d quarks. Numerically these are (142, 145, 153) MeV/c^2 from the $\frac{3}{2}^+$ decuplet and (123, 202, 177) MeV/c^2 from the $\frac{1}{2}^+$ octet, where the mass of each isospin multiplet has been taken to be the mean mass of its members. There is considerable spread among these values, but nevertheless they do suggest that the main contribution to the mass splittings within a supermultiplet is indeed the quark mass difference. Averaging the above numbers gives the rough estimate

$$m_s - m_{u,d} \approx 160\,\mathrm{MeV/c^2}, \tag{6.29}$$

which is consistent with that obtained in more detailed analyses.

★6.2.4 Baryon magnetic moments

We now consider the magnetic moments of the $\frac{1}{2}^+$ octet of baryons, which have all been measured except for that of the very short-lived $\Sigma^0(1193)$. If the quarks have zero orbital angular momenta, then these magnetic moments are just the sums of contributions from the quark magnetic moments, which we assume to be of the Dirac form (1.14); i.e. we assume

$$\mu_q \equiv \left\langle q, S_z = \tfrac{1}{2} \right| \hat{\mu}_z \left| q, S_z = \tfrac{1}{2} \right\rangle = e_q e/2m_q = (e_q M_p/m_q)\,\mu_N, \tag{6.30a}$$

where e_q is the quark charge in units of e and

$$\mu_N \equiv e/2M_p \tag{6.30b}$$

is the nuclear magneton. In the case of the $\Lambda(1116) = uds$, the ud pair is in a spin-0 state, and hence makes no contribution to the Λ spin or magnetic moment, so that

$$\mu_\Lambda = \mu_s = -\frac{1}{3}\frac{M_p}{m_s}\mu_N, \tag{6.31a}$$

where we have used (6.30). For $\frac{1}{2}^+$ baryons B with the quark configuration aab, the aa pair is in the symmetric spin-1 state[11] with parallel spins and magnetic moment $2\mu_a$. The spin-up baryon state is given by[12]

$$\left|B; S=\tfrac{1}{2},\ S_z=\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\left|b;\ S=\tfrac{1}{2},\ S_z=-\tfrac{1}{2}\right\rangle\left|\ aa;\ S=1,\ S_z=1\right\rangle - \sqrt{\tfrac{1}{3}}\left|b;\ S=\tfrac{1}{2},\ S_z=\tfrac{1}{2}\right\rangle\left|\ aa;\ S=1,\ S_z=0\right\rangle.$$

The first term corresponds to a state with magnetic moment $2\mu_a - \mu_b$, since the b quark has $S_z = -\frac{1}{2}$; the second term corresponds to a state with magnetic moment μ_b, since the aa pair has $S_z = 0$ and does not contribute. Hence the baryon magnetic moment is given by

$$\mu_B = \frac{2}{3}(2\mu_a - \mu_b) + \frac{1}{3}\mu_b = \frac{4}{3}\mu_a - \frac{1}{3}\mu_b. \tag{6.31b}$$

In particular, the proton magnetic moment is given by

$$\mu_p = \frac{4}{3}\mu_u - \frac{1}{3}\mu_d = \frac{M_p}{m_{u,d}}\mu_N, \tag{6.31c}$$

where we have neglected the small mass difference between the u and d quarks. The predictions (6.31a) and (6.31c) agree with the measured values of the p and Λ magnetic moments for quark masses

$$m_s = 510\,\mathrm{MeV/c^2}, \quad m_{u,d} = 336\,\mathrm{MeV/c^2}, \tag{6.32}$$

which agree with our previous rough estimate of the mass difference (6.29). Substituting these values into (6.30) and (6.31b) and using the quark assignments of Table 6.5 leads to predictions for the other magnetic moments, which are compared with experiment in Table 6.7. The agreement is reasonable, but not exact. This suggests that while the assumption that the observed baryons are three-quark states with zero orbital angular momentum is a good approximation, there could be a small admixture of other states with the same quantum numbers. One possibility is suggested by our discussion of the deuteron magnetic moment in Section 5.2.2. This showed that the orbital angular momentum quantum number is only approximately conserved in strong interactions and that the ground states of strongly interacting systems are not in general pure S waves, but contain small admixtures of states with nonzero orbital angular momenta. Such admixtures are neglected in the simple quark model, but may well occur. A second possibility is that in addition to the dominant qqq state, there may be small admixtures of the more complicated $qqqq\bar{q}$ states. This possibility will be returned to briefly in the following section.

[11] This result was obtained at the end of Section 6.2.2.
[12] This result is derived in Appendix C, Equation (C.36c).

TABLE 6.7 A comparison of the observed magnetic moments of the $\frac{1}{2}^+$ baryon octet, and the predictions of the simple quark model, Equations (6.31a) and (6.31b), for $m_u = m_d = 336\,\mathrm{MeV/c^2}$, $m_s = 510\,\mathrm{MeV/c^2}$.

Particle	Prediction (μ_N)	Experiment (μ_N)
$p(938)$	2.79	2.793[#]
$n(940)$	−1.86	−1.913[#]
$\Lambda(1116)$	−0.61	-0.613 ± 0.004
$\Sigma^+(1189)$	2.69	2.42 ± 0.05
$\Sigma^-(1197)$	−1.04	-1.157 ± 0.025
$\Xi^0(1315)$	−1.44	-1.250 ± 0.014
$\Xi^-(1321)$	−0.51	-0.679 ± 0.031

[#] The errors on the proton and neutron magnetic moments are of the order 6×10^{-8} and 5×10^{-7}, respectively.

6.3 COLOUR

The quark theory of hadrons is conspicuously successful, but appears to contradict the Pauli principle. This states that the wavefunction of any state must be antisymmetric under the interchange of any two identical spin-$\frac{1}{2}$ fermions, and is one of the fundamental assumptions of quantum mechanics. Despite this, in discussing the lightest baryons we assumed that their combined space and spin wavefunctions were symmetric under the interchange of any two quarks of the same flavour contained within them. This assumption is crucial to the success of the model, and the assumption of antisymmetry leads to a spectrum that is incompatible with experiment (cf. Problem 6.6). The same remarks apply to the excited states of baryons, which are also well accounted for by the quark model provided the combined space and spin wavefunctions are assumed to be symmetric.

The apparent contradiction between the quark model and the Pauli principle was resolved in 1964 by Greenberg, who argued that in addition to space and spin degrees of freedom, quarks must possess another attribute, which he called *colour*. The total wavefunction is then written as the product of a spatial part $\psi(\mathbf{r})$, a spin part χ and a colour wavefunction χ_C, i.e.

$$\Psi = \psi(\mathbf{r})\chi\,\chi_C, \tag{6.33}$$

where the colour wavefunction will be described shortly. The Pauli principle is now interpreted as applying to the total wavefunction, including the colour part χ_C. The combined space and spin wavefunctions can then be symmetric under the interchange of quarks of the same flavour (to accord with experiment) provided the colour wavefunction is antisymmetric. Associated with the colour wavefunction are several conserved quantum numbers, called *colour charges*, which play a similar role in strong interactions to that played by the electric charge in electromagnetic interactions. This will be discussed in the next chapter. Here we will introduce the colour charges and wavefunctions, and investigate the hypothesis of *colour confinement*. This is the hypothesis that hadrons can only exist in states, called *colour singlets*, that have zero values for all colour charges, while quarks, which

have nonzero colour charges, can only exist confined within them. As we shall see shortly, this explains why hadrons have integer electric charges, while fractionally charged combinations like qq and $qq\bar{q}$ are forbidden, in accordance with experimental observation.

6.3.1 Colour charges and confinement

The basic assumption of the colour theory is that any quark $q = u, d, s, \ldots$ can exist in three different colour states $\chi_C = r, g, b$, standing for 'red', 'green' and 'blue', respectively.[13] Just as the spin states $\chi = \alpha, \beta$ correspond to different values of the spin component S_3, the colour states χ_C correspond to different values of two of the colour charges called the *colour hypercharge* and the *colour isospin charge*.[14] They are denoted Y^C and I_3^C, and their values for the single quark states $\chi_C = r, g, b$ are listed in Table 6.8(a). Their values for other states composed of quarks and antiquarks then follow by using the fact that they are additive quantum numbers, like electric charge, whose values for particles and antiparticles are equal in magnitude but opposite in sign. Thus under charge conjugation a quark q in the r-state is transformed into a quark $\bar{q}$ in the colour state $\chi_C = \bar{r}$ with reversed values of Y^C and I_3^C, as listed in Table 6.8(b).

TABLE 6.8 Values of the colour charges I_3^C and Y^C for the colour states of quarks and antiquarks.

(a) Quarks			(b) Antiquarks		
	I_3^C	Y^C		I_3^C	Y^C
r	1/2	1/3	$\bar{r}$	−1/2	−1/3
g	−1/2	1/3	$\bar{g}$	1/2	−1/3
b	0	−2/3	$\bar{b}$	0	2/3

Direct evidence for the three colour states of quarks will be presented in Chapter 7, where we will also discuss the physical interpretation of the colour charges as sources of the strong interaction. Here we will discuss the consequences of assuming that any hadron that can be observed as an isolated particle in free space must have colour charges

$$I_3^C = Y^C = 0, \tag{6.34}$$

in accordance with the hypothesis of colour confinement.

[13] Explicit forms for these colour wavefunctions, analogous to the Pauli representation (5.22) for the spin wavefunctions $\chi = \alpha, \beta$, will be given in starred Section 6.3.2.

[14] The origin of the colour charges is discussed in more detail in starred Section 6.3.2, but is not required here, and the values of Table 6.8 may be regarded as definitions. The names colour hypercharge and colour isospin arise because of formal similarities with the 'ordinary' hypercharge and isospin discussed previously. There is, however, no physical connection between the latter quantum numbers and the two colour charges introduced here.

We first consider baryons, assumed to be composed of three quarks. From Table 6.8(a) it is clear that the colour charges of the three quarks can only add up to give $I_3^C = Y^C = 0$ if baryons are composed of one r, one g and one b quark. It is this observation that gave rise to the term 'colour', since in the theory of colour vision, white (i.e. colourless) light is formed by combining three primary colours, which can be chosen to be red, green and blue.[15] The most general colour wavefunction for a baryon is thus a linear superposition of six possible combinations; i.e.

$$\chi_B^C = \alpha_1 r_1 g_2 b_3 + \alpha_2 g_1 r_2 b_3 + \alpha_3 b_1 r_2 g_3 + \alpha_4 b_1 g_2 r_3 + \alpha_5 g_1 b_2 r_3 + \alpha_6 r_1 b_2 g_3, \qquad (6.35)$$

where, for example, r_3 means that the third quark is in an r state and the $\alpha_i (i = 1, 2, \ldots, 6)$ are constants. In particular, if we choose the totally antisymmetric combination

$$\chi_B^C = \frac{1}{\sqrt{6}}(r_1 g_2 b_3 - g_1 r_2 b_3 + b_1 r_2 g_3 - b_1 g_2 r_3 + g_1 b_2 r_3 - r_1 b_2 g_3), \qquad (6.36)$$

then the Pauli principle will indeed require the combined space and spin wavefunction of the baryon to be symmetric under the interchange of quarks of the same flavour, as assumed in our earlier discussion of baryons. In fact, it can be shown that the choice (6.36) is not only allowed by colour confinement but is required by it. However, this needs a fuller treatment than the one given here.[16]

We next consider the combinations $q^m \bar{q}^n$ of m quarks and n antiquarks that are allowed by the confinement condition (6.34). In doing this, it will be sufficient to consider only combinations with baryon number $B \geq 0$, and hence $m \geq n$, since the particles with $B < 0$ are just antiparticles of those with $B > 0$. The corresponding colour wavefunctions can be written

$$r^\alpha g^\beta b^\gamma \bar{r}^{\bar{\alpha}} \bar{g}^{\bar{\beta}} \bar{b}^{\bar{\gamma}}, \qquad (6.37)$$

where r^α means that there are α quarks in the r state, etc.,[17] and

$$m = \alpha + \beta + \gamma \geq n = \bar{\alpha} + \bar{\beta} + \bar{\gamma}.$$

We obtain the colour hypercharge Y^C and colour isospin I_3^C for the colour wavefunction (6.37) by adding the colour charges listed in Table 6.8. Imposing the confinement condition (6.34) then gives

$$I_3^C = (\alpha - \bar{\alpha})/2 - (\beta - \bar{\beta})/2 = 0$$

[15] This is of course only an analogy, and the quark colours have nothing whatsoever to do with ordinary colours!

[16] It is given, for the interested reader, in starred Section 6.3.2, but is not required elsewhere in the book.

[17] There is no need to consider the order of the factors here, since, for example, *rrb* has the same values of I_3^C and Y^C as *rbr* or *brr*.

and

$$Y^C = (\alpha - \bar{\alpha})/3 + (\beta - \bar{\beta})/3 - 2(\gamma - \bar{\gamma})/3 = 0,$$

implying

$$\alpha - \bar{\alpha} = \beta - \bar{\beta} = \gamma - \bar{\gamma} \equiv p$$

and hence

$$m - n = 3p,$$

where p is a nonnegative integer. Thus the only combinations $q^m \bar{q}^n$ allowed by colour confinement are of the form

$$(3q)^p (q\bar{q})^n \quad (p, n \geq 0), \tag{6.38}$$

where we recall that we have restricted ourselves to states with baryon number $B \geq 0$.

It follows from this that hadrons with fractional electric charges such as

$$qq,\ qq\bar{q},\ qqqq, \ldots$$

are forbidden by colour confinement, while the combinations $q\bar{q}$ and $3q$ used in the simple quark model are allowed. Nuclei composed of A nucleons and hence $3A$ quarks are also examples of (6.38) with $p = A$ and $n = 0$, although for most purposes the quark substructure may be neglected. More unusual are combinations like

$$qq\bar{q}\,\bar{q} \quad \text{and} \quad qqqq\bar{q},$$

which could give rise to exotic mesons and baryons, respectively. As we have seen in Section 3.6, there is no convincing experimental evidence for such states, and the simple quark model gives a good account of hadrons without considering them. Nonetheless, they are not forbidden by colour confinement and may well play a minor role in hadron physics. For example, baryons could in principle contain small components of $qqqq\bar{q}$ states with the same conserved quantum numbers as the dominant qqq states, and such admixtures might be partly responsible for the small discrepancies between the measured baryon magnetic moments and those predicted by the simple quark model with zero orbital angular momentum.

★6.3.2 Colour wavefunctions and the Pauli principle

In this section we give explicit expressions for the colour wavefunctions and operators, and show that for baryons the totally antisymmetric colour wavefunction (6.36) is not only allowed by colour confinement but is required by it.

The three independent colour wavefunctions $\chi^C = r, g, b$ of a quark are conveniently represented by the 'colour spinors'

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{6.39}$$

in the same way that the two spin wavefunctions $\chi = \alpha, \beta$ were represented by the spinors

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

in Section 5.2.1. Just as the spin wavefunctions are acted on by spin operators, the colour wavefunctions are acted on by 'colour operators'. The latter are represented by three-dimensional matrices in the same way that the spin operators are represented by the two-dimensional matrices (5.21b). There are eight such independent colour operators

$$\hat{F}_i = \frac{1}{2}\lambda_i \quad (i = 1, 2, \ldots, 8), \tag{6.40a}$$

where the matrices

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{6.40b}$$

can be regarded as the three-dimensional analogues of the two-dimensional Pauli matrices (5.21b).[18] The colour wavefunctions (6.39) are eigenfunctions of the operators $\hat{F}_3$ and $\hat{F}_8$ with, for example,

$$\hat{F}_3 r = \frac{1}{2} r, \quad \hat{F}_8 r = \frac{1}{2\sqrt{3}} r \tag{6.41}$$

[18] In both cases, they are traceless Hermitian matrices, and the most general Hermitian matrix in two or three dimensions can be written as a linear combination of the unit matrix and the σ_i or λ_i respectively.

for the r state. The colour charges I_3^C and Y^C discussed in the previous section are the eigenvalues of the operators

$$\hat{I}_3^C \equiv \hat{F}_3 \quad \text{and} \quad \hat{Y}^C \equiv \frac{2}{\sqrt{3}}\hat{F}_8, \tag{6.42}$$

where the factor $2/\sqrt{3}$. was introduced for historical reasons. From (6.41) and (6.42) it follows that $I_3^C = \frac{1}{2}$ and $Y^C = \frac{1}{3}$ for the r state, and the corresponding values for the g and b states listed in Table 6.8(a) are obtained in a similar fashion. The remaining operators $\hat{F}_1, \hat{F}_2, \hat{F}_4, \hat{F}_5, \hat{F}_6$ and $\hat{F}_7$ mix the colour states (6.39), and it can easily be shown using (6.40) that, for example,

$$\hat{F}_1\, r = \frac{1}{2} g, \quad \hat{F}_1\, g = \frac{1}{2} r, \quad \hat{F}_1\, b = 0. \tag{6.43}$$

The observables associated with the operators $\hat{F}_i$ are all believed to be exactly conserved in nature, i.e.

$$\left[\hat{F}_i\,, H\right] = 0 \quad (i = 1, 2, \ldots, 8),$$

and for this reason they are all called colour charges. Colour confinement is the requirement that all eight colour charges vanish for any observed hadron h, implying

$$\hat{F}_i\, \chi_h^C = 0 \quad (i = 1, 2, \ldots, 8) \tag{6.44}$$

for the corresponding colour wavefunction χ_h^C. This is reminiscent of a spin singlet state, in which all three spin components vanish, so that the spin wavefunctions satisfies

$$\hat{S}_x\, \chi = 0, \quad \hat{S}_y\, \chi = 0, \quad \hat{S}_z\, \chi = 0. \tag{6.45}$$

It is for this reason that colour states that satisfy (6.44) are called colour singlets. Equation (6.44) implies, but is more restrictive than, the conditions (6.34) that were exploited extensively in the previous section.

Finally, we shall use colour confinement to fix completely the colour wavefunction for any baryon. We have already seen that the most general colour wavefunction compatible with (6.34) is

$$\chi_B^C = \alpha_1\, r_1\, g_2\, b_3 + \alpha_2\, g_1\, r_2\, b_3 + \alpha_3\, b_1\, r_2\, g_3 + \alpha_4\, b_1\, g_2\, r_3 + \alpha_5\, g_1\, b_2\, r_3 + \alpha_6\, r_1\, b_2\, g_3, \tag{6.46}$$

where the $\alpha_i (i = 1, 2, \ldots, 6)$ are constants. The values of the latter follow from applying (6.44), which we will illustrate for the case of $\hat{F}_1$. It follows from (6.43) that acting with $\hat{F}_1$ on the first term of (6.46) gives

$$\alpha_1 \hat{F}_1 (r_1 g_2 b_3) = \alpha_1 (\hat{F}_1 r_1) g_2 b_3 + \alpha_1 r_1 (\hat{F}_1 g_2) b_3 + \alpha_1 r_1 g_2 (\hat{F}_1 b_3) = \frac{\alpha_1}{2} \left(g_1 g_2 b_3 + r_1 r_2 b_3\right).$$

Similar contributions are obtained by acting with $\hat{F}_1$ on the other terms in (6.46), and collecting these together we obtain

$$\hat{F}_1\,\chi_B^C = \frac{(\alpha_1+\alpha_2)}{2}\,(g_1g_2b_3 + r_1r_2b_3) + \frac{(\alpha_3+\alpha_4)}{2}\,(b_1g_2g_3 + b_1r_2r_3) + \frac{(\alpha_5+\alpha_6)}{2}\,(g_1b_2g_3 + r_1b_2r_3)\,.$$

This is only compatible with (6.44) for $i=1$ if the conditions

$$\alpha_1 = -\alpha_2, \qquad \alpha_3 = -\alpha_4, \qquad \alpha_5 = -\alpha_6 \tag{6.47}$$

are satisfied. Analogous relations follow from considering the other operators $\hat{F}_2, \hat{F}_4$, etc., in (6.44), and together they uniquely determine the wavefunction (see Problem 6.10) to be the totally antisymmetric combination (6.36)

$$\chi_B^C = \frac{1}{\sqrt{6}}(r_1g_2b_3 - g_1r_2b_3 + b_1r_2g_3 - b_1g_2r_3 + g_1b_2r_3 - r_1b_2g_3),$$

up to an arbitrary normalization constant. Since the Pauli principle requires the total wavefunction to be antisymmetric under the interchange of identical quarks, the combined space and spin part must therefore be symmetric, as assumed in our discussion of baryons in Section 6.2.2.

★6.4 CHARMONIUM AND BOTTOMIUM

Finally, we consider the spectra of charmonium ($c\bar{c}$) and bottomium ($b\bar{b}$). The c and b quarks are much heavier than the u, d and s quarks (see Table 3.1) and the lower-lying $c\bar{c}$ and $b\bar{b}$ states are essentially nonrelativistic systems, in which the quark kinetic energies are much smaller than their rest masses. There is thus a strong analogy with the e^+e^- bound system positronium, except that the potential energy is different. As in positronium, states can be labelled by their angular momentum quantum numbers, J, L and S, where the spin S can be 0 or 1, and by their principal quantum number $n = \nu + L + 1$, where ν is the number of radial nodes, not counting nodes at the origin. Alternatively, instead of n, the radial quantum number $n_r = \nu + 1 = n - L = 1, 2, 3, \ldots$ is sometimes used to label the states.[19] The resulting states for positronium with principal quantum number $n=1$ or $n=2$ have been listed in Table 5.1, together with the values of their parity

$$P = P_{e^+}P_{e^-}(-1)^L = (-1)^{L+1} \tag{6.48}$$

and C-parity

$$C = (-1)^{L+S}, \tag{6.49}$$

[19] This is done, for example, in the tables of the Particle Data Group (Yao *et al.*, 2006).

which follow from (5.35) and (5.53) respectively. The same results apply to charmonium and bottomium, and are listed in Table 6.9, together with the observed mesons that correspond to them, where these have been identified experimentally. As we shall see, the measured energies of these states can be used to partially determine the potential energies in the two systems, leading to the important conclusion that the forces acting in the $c\bar{c}$ and $b\bar{b}$ systems are the same within errors.

TABLE 6.9 Predicted $c\bar{c}$ and $b\bar{b}$ states with principal quantum numbers $n = 1$ and 2 and radial quantum number $n_r = n - L$, compared with experimentally observed states. Masses are given in MeV/c^2.

$^{2S+1}L_J$	n	n_r	J^{PC}	$c\bar{c}$ state	$b\bar{b}$ state
$^1S_0$	1	1	0^{-+}	$\eta_c(2980)$	$\eta_b(9300)$#
$^3S_1$	1	1	1^{--}	$J/\psi(3097)$	$\Upsilon(9460)$
$^3P_0$	2	1	0^{++}	$\chi_{c0}(3415)$	$\chi_{b0}(9859)$
$^3P_1$	2	1	1^{++}	$\chi_{c1}(3511)$	$\chi_{b1}(9893)$
$^3P_2$	2	1	2^{++}	$\chi_{c2}(3556)$	$\chi_{b2}(9913)$
$^1P_1$	2	1	1^{+-}	$h_c(3526)$#	
$^1S_0$	2	2	0^{-+}	$\eta_c(3638)$	
$^3S_1$	2	2	1^{--}	$\psi(3686)$	$\Upsilon(10023)$

State is not well established and its quantum number assignments are unknown.

★6.4.1 Charmonium

The charmonium system is most conveniently studied by first forming those states $V^0 = J/\psi(3097)$, $\psi(3686), \ldots$ that have the same J^{PC} values 1^{--} as the photon (cf. Table 6.9) and then observing their radiative decays. The basic formation process is shown in Figure 6.3: the electron–positron pair annihilates to a virtual photon, which converts to the V^0, which then decays, mainly into hadrons. This is only possible when the total centre-of-mass energy is equal to the V^0 mass, leading to a peak in the total cross-section for electron–positron annihilation into hadrons. This cross-section is usually plotted in terms of the ratio

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}, \tag{6.50a}$$

where the measured cross-section for muon pair production,[20]

$$\sigma(e^+e^- \to \mu^+\mu^-) = \frac{4\pi\alpha^2}{3E_{\text{CM}}^2}, \tag{6.50b}$$

is a smoothly varying function of the centre-of-mass energy E_{CM}.

[20] This cross-section can be calculated theoretically from the mechanism of Figure 1.3 (see, for example, pp. 146–150 of Mandl and Shaw, 1993). However, such calculations are beyond the scope of this book, and it is sufficient for our purposes to take this result from experiment.

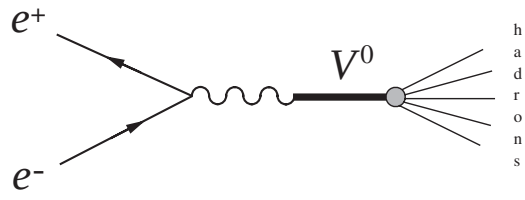


Figure 6.3 Mechanism for the formation of mesons V^0 with quantum numbers $J^{PC} = 1^{--}$ in electron–positron annihilation, and their subsequent decay to hadrons.

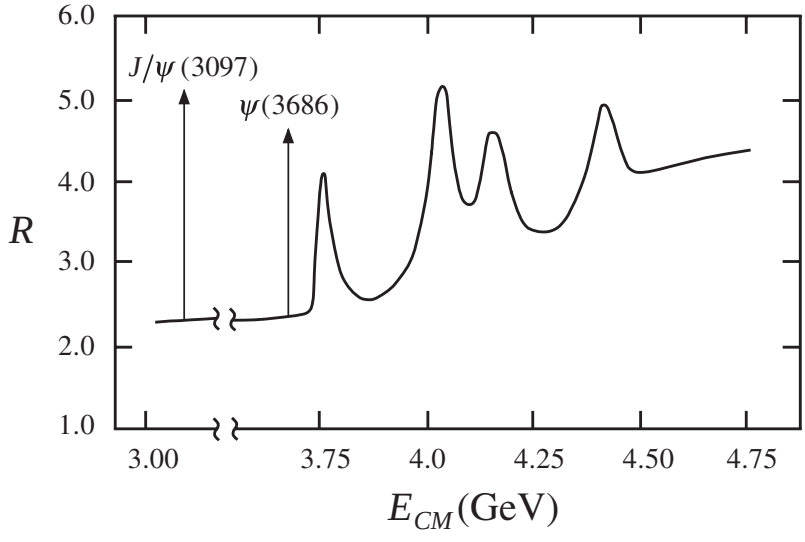
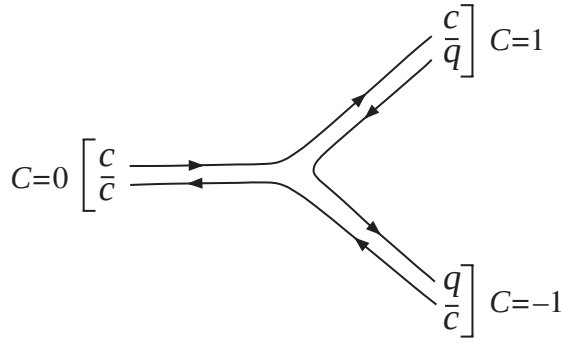


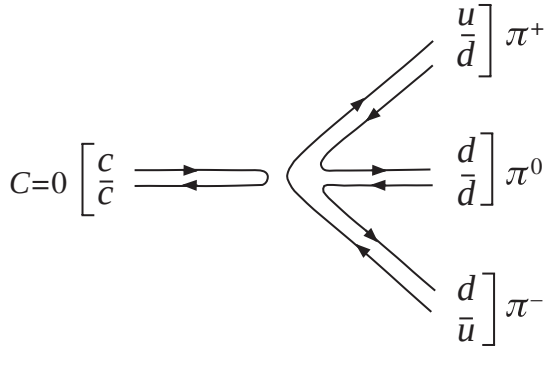
Figure 6.4 Measured cross-section ratio (6.50a) in the vicinity of the threshold for producing pairs of charmed particles. The extremely narrow peaks associated with the $J/\psi\,(3097)$ and $\psi\,(3686)$ are indicated by arrows because they extend far above the scale shown. In the partially suppressed region between these peaks, the ratio is essentially constant. (Data from Brandelik *et al.*, 1978, and Schindler *et al.*, 1980.)

As can be seen, there are two sharp peaks in the ratio shown in Figure 6.4, corresponding to the masses of the $J/\psi\,(3097)$ and $\psi\,(3686)$, followed by several relatively low, broad peaks at energies above 3730 MeV. This latter energy is called the *charm threshold* and is twice the mass of the lightest charmed particles, which are the D mesons (3.18). Charmonium states that lie above this threshold are resonances that decay readily into pairs of charmed particles by the mechanism of Figure 6.5(a), involving the creation of a light quark–antiquark pair, and the broad peaks in Figure 6.4 are evidence for the existence of several such states. In contrast, charmonium states that lie below this threshold are all very narrow and cannot decay by the mechanism of Figure 6.5(a) because of energy conservation; instead they decay much more slowly to noncharmed hadrons by mechanisms like that of Figure 6.5(b), involving

the annihilation of a charmed quark–antiquark pair.[21] As narrow states are much easier to identify experimentally than broad ones, the spectrum is much better known below the charmed threshold than above it, and we will therefore mainly confine ourselves to this part of the spectrum in what follows.



(a)



(b)

Figure 6.5 Quark diagrams for (a) the decay of a charmonium state to a pair of charmed mesons and (b) an example of a decay to noncharmed mesons.

The $J/\psi(3097)$ and $\psi(3686)$ are the only well-established charmonium states listed in Table 6.9 that can be produced by the mechanism of Figure 6.3, because they are the only ones with the same quantum numbers $J^{PC} = 1^{--}$ as the photon. In addition, there are suggestions of other $J^{PC} = 1^{--}$ states at higher masses. Some of the other states shown in Table 6.9, however, can be found by studying the radiative

[21] The suppression of processes involving the creation or annihilation of $c\bar{c}$ and $b\bar{b}$ pairs, as opposed to those involving the creation or annihilation of light quark–antiquark pairs ($u\bar{u}, d\bar{d}, s\bar{s}$), is a general characteristic of strong interactions. It is often referred to as the OZI rule, after Okubo, Zweig and Iizuka. It does not apply to weak or electromagnetic interactions.

decays of the $J/\psi(3097)$ and $\psi(3686)$. In particular, the $\psi(3686)$ decays mostly into hadrons, but about 25 % of the time it decays by the radiative transitions

$$\psi(3686) \to \chi_{ci} + \gamma \quad (i = 0, 1, 2) \tag{6.51}$$

where $\chi_{c0}(3415)$, $\chi_{c1}(3511)$ and $\chi_{c2}(3556)$ are three new charmonium states, called chi-states. These particles themselves decay, mostly to give hadrons in the final state, and an analysis of these decays shows that they have J^{PC} values 0^{++}, 1^{++} and 2^{++}, respectively. We therefore identify them with the $^3P_0$, $^3P_1$ and $^3P_2$ states of charmonium, which have these predicted J^{PC} values, as can be seen from Table 6.9. In addition, the $n = 1$, $^1S_0$ ground state $\eta_c(2980)$ has been observed in the radiative transitions

$$\psi(3686) \to \eta_c(2980) + \gamma \tag{6.52a}$$

and

$$J/\psi(3097) \to \eta_c(2980) + \gamma, \tag{6.52b}$$

despite the small branching ratios of about 3×10^{-3} and 10^{-2}, respectively. The remaining charmonium states in Table 6.9 with $n \leq 2$ are the $^1S_0$ state $\eta_c(3638)$ and the $^1P_1$ state $h_c(3526)$. The former has been seen in $\gamma\gamma$ interactions and the latter in $p\bar{p}$ annihilations, although it is not as well established. There is also evidence for a small number of higher mass states with $n = 3$.

To summarize, all eight predicted states of charmonium with principal quantum number $n = 1$ or $n = 2$ have been identified and there is some evidence for states with $n > 2$. The resulting experimental spectrum is shown in Figure 6.6. Despite the difference in energy scales, it bears a striking resemblance to the spectrum of positronium shown in Figure 5.2, except that the $2S$ levels are somewhat higher in energy than the $2P$ levels, instead of being degenerate with them.

★6.4.2 Bottomium

The bottomium spectrum is observed in much the same way as the charmonium spectrum, and we confine ourselves here to stating the results, which are summarized in Figure 6.6. In this case there are three bands of narrow states, corresponding to principal quantum number $n = 1$, 2 and 3, lying below the bottom threshold of $10.56\,\mathrm{GeV/c^2}$. This is just twice the mass of the B mesons (3.19), which are the lightest particles with nonzero bottom, and bottomium states above this threshold are broad resonances that decay to pairs of particles, each with $\tilde{B} \neq 0$, by a mechanism analogous to that of Figure 6.5(a). As can be seen in Figure 6.6, the ordering of the levels and the magnitudes of their splittings are similar in charmonium and bottomium, indicating an underlying similarity between the forces acting in the two systems.

★6.4.3 The quark–antiquark potential

The experimental spectra of charmonion and bottomium shown in Figure 6.6 bear a striking resemblance to the spectrum of positronium shown in Figure 5.2, and like it

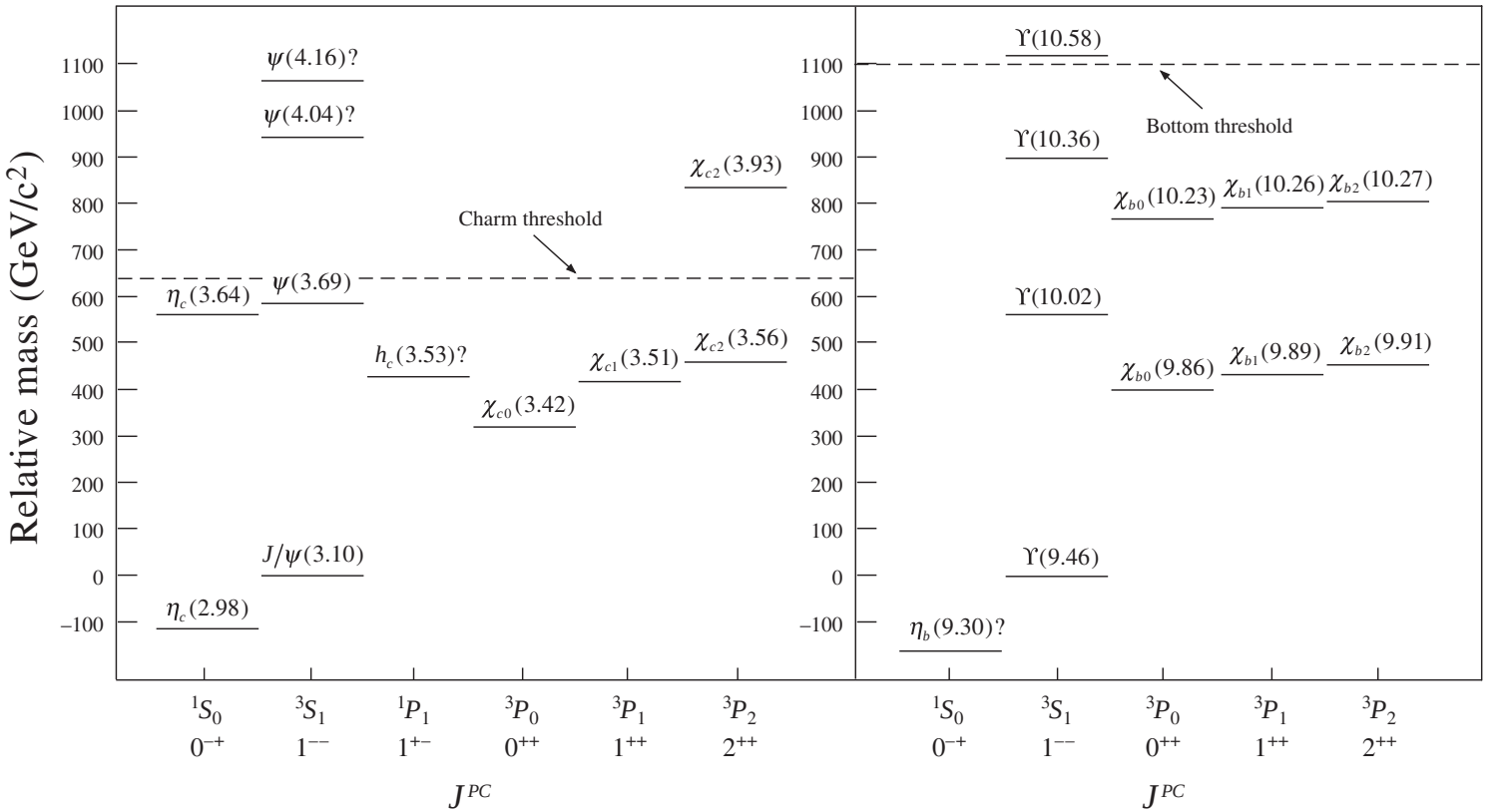


Figure 6.6 The observed states of the charmonium ($c\bar{c}$) and bottomium ($b\bar{b}$) for $L \leq 1$. The masses are given in units of GeV/c^2 and are plotted relative to that of the $^3S_1$ ground state.

can be understood in terms of a simple nonrelativistic treatment. In the centre-of-mass frame of the $q\bar{q} = c\bar{c}$ or $b\bar{b}$ pair, the Schrödinger equation is

$$-\frac{1}{2\mu}\nabla^2\psi(\mathbf{r}) + V(r)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{6.53}$$

where $r = |\mathbf{r}|$ is the distance between the quarks, $\mu = m_q/2$ is their reduced mass and we have neglected spin-dependent effects. In this approximation the energies E and the particle masses

$$M(q\bar{q}) = 2m_q + E \tag{6.54}$$

depend only on the principal quantum number n and the orbital angular momentum quantum number L, for a given potential $V(r)$. In the particular case of a Coulomb-like potential

$$V(r) \propto r^{-1}, \tag{6.55}$$

the energies depend on n only, but for other r-dependencies this degeneracy is broken.

This is illustrated in Figure 6.7 where the spectrum of S- and P-wave states resulting from (6.55) is compared with that resulting from a simple harmonic oscillator potential

$$V(r) \propto r^2. \tag{6.56}$$

The heavy quarkonia spectra of Figure 6.6 are intermediate between these possibilities, and can be fitted using a potential of the form

$$V(r) = -\frac{a}{r} + b\,r, \tag{6.57a}$$

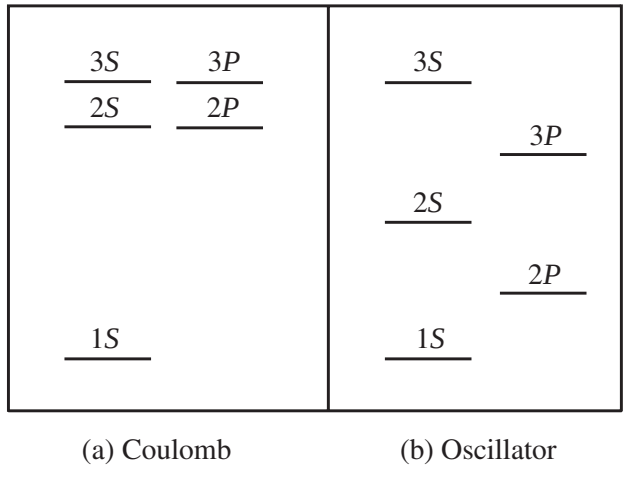


Figure 6.7 Patterns of S- and P-wave energy levels arising from (a) Coulomb-like (r^{-1}) and (b) simple harmonic oscillator (r^2) potentials for principal quantum numbers $n = 1, 2, 3$. The strengths of the potentials are assumed to be such that the $3s$ states have the same energy, measured relative to the $1s$ state, in each case.

which behaves like a Coulomb potential at small r, but rises linearly at large r values. The constants a and b are determined by solving the Schrödinger equation numerically and adjusting their values until good fits to the experimental spectra are obtained.[22] In this way one finds that the values

$$a = 0.30 \quad \text{and} \quad b = 0.23\,\text{GeV}^2 \tag{6.57b}$$

can be used to fit both the charmonium and bottomium spectra, and the resulting potential is shown in Figure 6.8. The potential is not completely determined by this analysis, because of uncertainties in the quark masses and because equally satisfactory results can be obtained using other forms, e.g.

$$V(r) = a\,\ln(br), \tag{6.58a}$$

where

$$a = 0.75\,\text{GeV} \quad \text{and} \quad b = 0.80\,\text{GeV}. \tag{6.58b}$$

However, all the potentials that successfully explain the spectrum for given quark masses are in good agreement with each other in the range

$$0.2\,\text{fm} \lesssim r \lesssim 1.0\,\text{fm}, \tag{6.59}$$

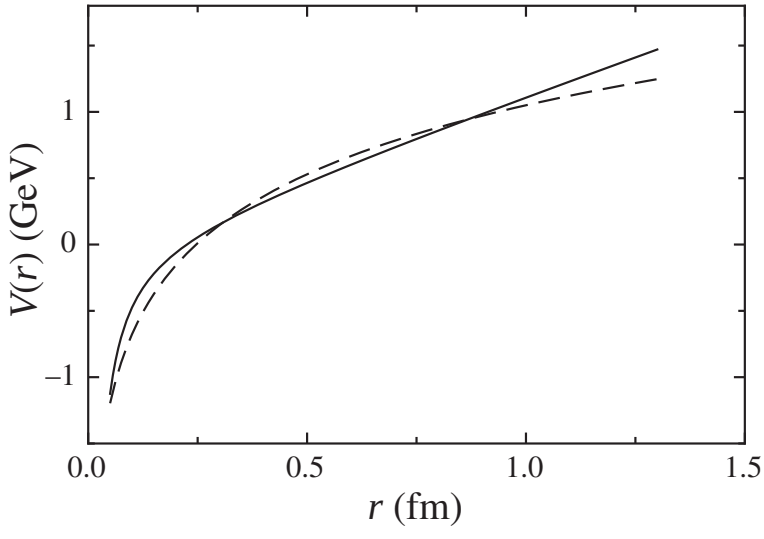


Figure 6.8 Heavy quark–antiquark potentials obtained from fitting the energy levels of charmonium and bottomium. The solid and dashed lines show the results obtained from the forms (6.57) and (6.58), respectively.

[22] The agreement is not perfect, because relativistic effects are not completely negligible. These include a spin-dependent interaction to account for the fine structure seen in Figure 6.4. This interaction is similar to that used to describe the positronium fine structure (cf. Section 5.5.1).

as illustrated in Figure 6.8 for the cases (6.57) and (6.58). We therefore conclude that the potential is well-determined in the range (6.59) and that, within this range at least, the potentials for charmonium ($c\bar{c}$) and bottomium ($b\bar{b}$) are, within small experimental uncertainties, the same, as expected from flavour independence.[23]

PROBLEMS 6

6.1 The hadron $\Sigma_c^+(2455)$ is observed to decay by $\Sigma_c^+ \to \Lambda_c^+ + \pi^0$ with a rate typical of strong interactions, where the $\Lambda_c^+(2286) = udc$ is the isosinglet hadron listed in Appendix E, Table E.4. Deduce the values of the quantum numbers (3.31) and hence the quark content of the Σ_c^+. Does it possess any isospin partners and, if so, what is their quark content?

6.2 In the simple model of the Σ baryon mass splittings discussed in Section 6.1.4, the electromagnetic interaction energy between two quarks a and b was assumed to be of order $\delta e_a e_b$, where e_a and e_b are the quark charges in units of e. Deduce the value of δ obtained from the observed Σ baryon masses in this approximation and use it to make a rough estimate of the typical distance r between the quarks by assuming that the Coulomb interaction dominates.

6.3 Show that a meson which decays to $\pi^+\pi^-$ pairs by the strong interaction must have $C = P = (-1)^J$, where J is the spin of the meson.
The $\rho^0(775)$ and $f_2^0(1275)$ mesons decay by strong interactions to give $\pi^+\pi^-$ pairs and have spin-1 and spin-2, respectively. Which of the decays $\rho^0 \to \pi^0\gamma$ and $f_2^0 \to \pi^0\gamma$ is forbidden in electromagnetic interactions? Which of the decays $\rho^0 \to \pi^0\pi^0$ and $f_2^0 \to \pi^0\pi^0$ is forbidden in any interaction?

6.4 The K^+ and its excited states with masses below $1.5\,\mathrm{GeV/c^2}$ are shown in Figure 3.12. Identify these mesons with states ${}^{2S+1}L_J$ of the appropriate quark–antiquark system, specifying the value of the principal quantum number n in each case.

6.5 In general, a meson with spin J can have $C = (-1)^J$ or $C = (-1)^{J+1}$ and $P = (-1)^J$ or $P = (-1)^{J+1}$, giving four possible combinations of C and P in all. Which of these combinations can occur in the simple quark model? List the forbidden J^{PC} values explicitly for $J = 0, 1, 2$ and 3.

6.6 In Section 6.2.2 we assumed that the combined space–spin wavefunction of a baryon was symmetric under the interchange of any pair of like quarks. What spectrum of low-lying baryon states composed of light quarks with zero orbital angular momenta $L_{12} = L_3 = 0$ would be predicted if, instead, we had assumed that the wavefunctions were antisymmetric under the exchanges of like quarks?

6.7 The lightest charmed baryons have the quark compositions cab with zero orbital angular momenta $L_{12} = L_3 = 0$, where c is the charmed quark and a and b can be any of the light quarks u, d, s. Show that the resulting states can be classified into three families:

(a) $\frac{1}{2}^+$ baryons in which the light quark pair ab has spin-0;
(b) $\frac{1}{2}^+$ baryons in which the light quark pair has spin-1;
(c) $\frac{3}{2}^+$ baryons in which the light quark pair has spin-1.

List the quark combinations cab that occur in each of these families and classify them into isospin multiplets by deducing I_3 from (6.4) and (6.5).

---

[23] These analyses also determine the mean square radii

$$r_0^2 \equiv \overline{r^2} = \int \mathrm{d}^3\mathbf{r}\, r^2\, |\psi(r)|^2 \tag{6.60}$$

of the various states, where the wavefunctions are determined by the solutions of the Schrödinger equation for successful potentials like (6.57) and (6.58). The resulting values of r_0 for the various states shown in Figure 6.6 span just the range (6.59) in which the potential is well-determined.

6.8 The following low-lying charmed baryons are known to exist: the isosinglet $\Lambda_c^+(2286)$ with $C = 1, S = 0$; the isodoublet $\Xi_c^+, \Xi_c^0(2468)$ with $C = 1, S = -1$; the isotriplet $\Sigma_c^{++}, \Sigma_c^+, \Sigma_c^0(2455)$ with $C = 1, S = 0$; and the isosinglet $\Omega_c^0(2700)$ with $C = 1, S = -2$. Assume that these particles are all $\frac{1}{2}^+$ states and identify them with states predicted to exist in the previous question. Make an estimate of the mass of the remaining spin-$\frac{1}{2}$ state needed to complete the supermultiplet of $\frac{1}{2}^+$ states.

6.9 Calculate the quark model prediction for the magnetic moment of the $\Sigma^0(1193)$.

6.10 In Section 6.3.2 we showed that the confinement condition $\hat{F}_1 \chi_B^C = 0$ led to the conditions

$$\alpha_1 = -\alpha_2, \quad \alpha_3 = -\alpha_4, \quad \alpha_5 = -\alpha_6$$

for the coefficients α_i in the general colour wavefunction for baryons (6.35). Find the conditions resulting from the other confinement conditions (6.44). Show that they are consistent with each other, and uniquely determine the colour wavefunction for baryons to be the totally antisymmetric form (6.36), up to an arbitrary overall normalization constant.

7

QCD, Jets and Gluons

In Chapter 6, we have seen how the static properties of hadrons can be understood in terms of their quark substructure. In this chapter we will consider the dynamical effects of this substructure as seen in scattering experiments. Such effects are observed in many reactions, but here we will concentrate on two topics where the quark interpretation of the experiments is relatively direct. These are: high-energy electron–positron annihilation, in which quarks and gluons are closely associated with jets of hadrons observed in the final state; and high-energy scattering experiments, in which beams of leptons, which are themselves structureless, are used to probe the internal structures of the proton and neutron. Firstly, however, we introduce the theory of strong interactions, whose qualitative features will underline almost our whole account.

7.1 QUANTUM CHROMODYNAMICS

The theory that describes strong interactions in the standard model is called *quantum chromodynamics*, or QCD for short. Although this is not tested to the same extent or precision as quantum electrodynamics (QED), it is nevertheless in impressive agreement with a large body of experimental data and is not contradicted by any known experiment. QCD is similar to QED in that both describe interactions that are mediated by massless spin-1 bosons coupling to conserved charges. Theories of this type are called *gauge theories* because they have a characteristic symmetry called gauge invariance. Gauge invariance plays a fundamental role in theoretical treatments of QED and QCD, where it can be used to infer the detailed forms of the interactions.[1] Here we shall adopt a more phenomenological approach. The spin-1 bosons

[1] This approach is called the *principle of minimal gauge interactions*, or the *gauge principle* for short. It is discussed explicitly in Appendix D, but is not required in what follows.

Particle Physics Third Edition B. R. Martin and G. Shaw

are called gauge bosons. In QED they are photons, and in QCD they are called gluons. Gluons have zero electric charge, like photons, but couple to the colour charges that were discussed in Section 6.3.1, rather than to the electric charge. This leads immediately to the so-called *flavour independence* of strong interactions; i.e. the different quark flavours $a = u, d, s, c, b$ and t must have identical strong interactions, because they exist in the same three colour states r, g, b with the same possible values of the colour charges. This has its most striking consequences for u and d quarks, which have almost equal masses, where it leads to the phenomenon of isospin symmetry described in Section 6.1. Flavour independence also implies the equality of the potentials in the charmonium and bottomium systems, which emerged from experiment in Section 6.4.3.

A second property of strong interactions that follows from the above picture without detailed argument is that the forces between the quarks must be long range, because the gluons have zero mass. This does not imply that the forces between hadrons are also long range, because hadrons have zero colour charges overall, as we saw in the last chapter. The forces between the 'colourless' hadrons are the residues of the forces between their quark constituents, and cancel when the hadrons are far apart.

We have noted that QED and QCD both describe interactions, albeit of very different strengths, which are mediated by massless spin-1 bosons that couple to conserved charges. However, there is a crucial difference between them which profoundly affects the character of the resulting forces. It is that while the photons that couple to the electric charge are themselves neutral, gluons have nonzero values of the colour charges to which they couple. This is illustrated in Figure 7.1, which shows a particular example of a quark–quark interaction by gluon exchange, where the gluon is represented by a 'corkscrew' line to distinguish it from a photon.

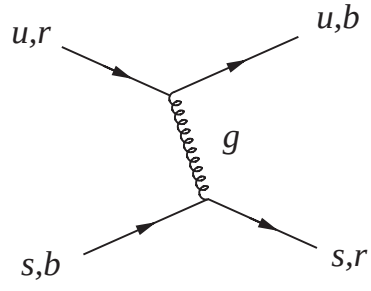


Figure 7.1 Example of quark–quark scattering by gluon exchange, where the gluon is represented by a 'corkscrew' line to distinguish it from a photon. In this diagram the quark flavour u or s is unchanged on gluon emission, but the colour state can change, as shown.

In this diagram, the colour states of the two quarks are interchanged, and the gluon has the colour quantum numbers

$$I_3^C = I_3^C(r) - I_3^C(b) = \tfrac{1}{2} \tag{7.1a}$$

and

$$Y^C = Y^C(r) - Y^C(b) = 1, \tag{7.1b}$$

which follow from applying colour conservation at, for example, the lower vertex and using the values for the quark colour states r, b listed in Table 6.8. Just as quarks exist in three different colour states, gluons can exist in eight different colour states,[2] and the I_3^C, Y^C values (7.1) correspond to just one of them. We do not need to consider these states in detail for what follows, since we shall not attempt any detailed calculations in QCD. Rather, we note that if gluons couple to particles with nonzero colour charges, and if gluons themselves also have nonzero charges, then by implication gluons couple to other gluons. The two types of gluon self-coupling that occur in QCD are illustrated in Figure 7.2, which shows the two lowest-order contributions to gluon–gluon scattering. The first is a gluon exchange process analogous to Figure 7.1 for quark–quark scattering, while the second involves a zero-range 'contact' interaction.

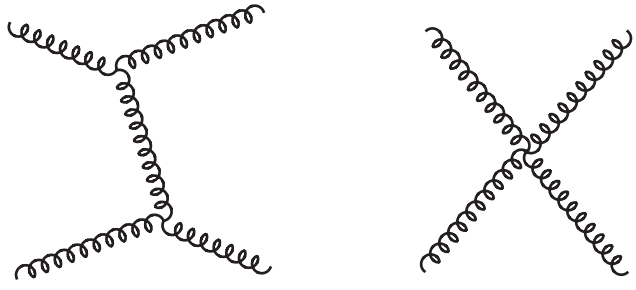


Figure 7.2 The two lowest-order contributions to gluon–gluon scattering in QCD.

The gluon–gluon interactions of Figure 7.2 have no analogue in QED, and it can be shown that they lead to properties of the strong interaction that differ markedly from those of the electromagnetic interaction. These properties are *colour confinement* and *asymptotic freedom*. Colour confinement is the requirement that observed states have zero colour charges. This was discussed in Section 6.3.1, where, for example, it was shown to imply that all quark bound states (hadrons) must have integral electric charges. It also implies that gluons, like quarks, cannot be observed as isolated free particles, because they have nonzero values of the colour charges. Bound states of two or more gluons with zero colour charges overall can be formed in principle, due to the strong interaction between gluons themselves. Such states are called *glueballs* and will be briefly discussed at the end of this subsection.

Asymptotic freedom means that the interaction gets weaker at short distances, and at distances less than about 0.1 fm the lowest-order diagrams dominate. At these distances quark–quark scattering, for example, is given approximately by one-gluon exchange diagrams like Figure 7.1. However, as the distance between the quarks increases, the interaction gets stronger and many higher-order diagrams become important. In this strong interaction regime the situation is very complicated, and

[2] This can be shown to follow as a consequence of the assumed symmetry properties of QCD. This is discussed in Section D.6.

it has not yet been possible to evaluate the theory precisely. We therefore have to rely on approximate results obtained by numerical simulations of the theory on very large computers, and the demonstration of confinement in QCD rests largely on such simulations. They are done using an approach called *lattice gauge theory* in which space (and sometimes time) is approximated by a finite lattice of discrete points. The exact theory can then in principle be recovered by letting the lattice spacing go to zero and the number of lattice points become infinite.[3]

The above features are conveniently illustrated by considering the static potential between a heavy quark and its antiquark in a colour singlet state. This is the appropriate potential for a nonrelativistic discussion of charmonium and bottomium, and was determined empirically for the limited range $0.2 \leq r \leq 0.8$ fm in Section 6.4.3. Here we are concerned with the behaviour predicted on the basis of QCD. At short interquark distances $r \leq 0.1$ fm, the interaction is dominated by one-gluon exchange and we might expect a Coulomb-like potential analogous to that arising from one-photon exchange in QED. In fact it can be shown that the potential is given by

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r} \quad (r \leq 0.1\,\text{fm}), \tag{7.2a}$$

where the strong coupling constant α_s is a measure of the strength of the interaction analogous to the fine structure constant α in QED.[4] Because of asymptotic freedom, the strength of the interaction, and hence α_s, decreases with decreasing r, but for $r \leq 0.1$ fm this variation is slight and can in many applications be neglected, as we shall see below. At distances beyond 0.1 fm, however, the strength of the interaction increases more rapidly and one-gluon exchange no longer dominates. In this region, we are forced to rely on lattice gauge theory calculations of limited precision. These are consistent with the empirical potential of Figure 6.8 in the range where this is determined, and show that at large distances the potential rises approximately linearly

$$V(r) \approx \lambda r \quad (r \geq 1\,\text{fm}), \tag{7.2b}$$

where the constant λ cannot be calculated precisely, but is of order $1\,\text{GeV fm}^{-1}$. This is an example of a confining potential in that it does not die away with increasing separation and the force between the quark and antiquark cannot be neglected, even when they are very far apart.

Lattice gauge theory also throws light on the possible existence of the *glueballs* mentioned earlier. Since these are composed of gluons, they will be strongly interacting neutral bosons with $S = C = \tilde{B} = 0$, if they exist, and one would, for example, expect their electromagnetic interactions to be much weaker than those of 'ordinary' mesons composed of charged quarks and antiquarks. Unfortunately, precise theoretical calculations of glueball properties are not yet possible because of our limited understanding of confinement, but lattice gauge theory calculations that ignore the

[3] A simple account of such calculations is given in Weingarten (1996).

[4] The numerical factor multiplying α_s (i.e. $-4/3$ in this case) depends on the colour state chosen, and we will not discuss it further.

existence of quarks suggest that glueballs do exist and that the lightest glueball is a scalar meson with $J^{PC} = 0^{++}$ and a mass of around $1.5 - 1.7\,\mathrm{GeV/c^2}$. However, when quarks are included in the theory, such states are expected to mix with $q\bar{q}$ mesons of the same quantum numbers and similar masses. In other words, the observed states are unlikely to be pure glueballs but states with both glueball and $q\bar{q}$ components, which are difficult to distinguish from ordinary mesons. There is at present no conclusive evidence for pure glueballs, despite many experimental searches, but there is some evidence for mixed states containing both gluon and $q\bar{q}$ components. This is still very much an active area of research.[5]

7.1.1 The strong coupling constant

The strong interaction derives its name from the strong forces acting at distances of order 1 fm which, among other things, bind quarks into hadrons. However, the remarkable phenomena discussed in this chapter depend on the fact that the interaction gets weaker at short distances, i.e. on asymptotic freedom. Such short-distance interactions are associated with large momentum transfers $|\mathbf{q}|$ between the particles, with

$$|\mathbf{q}| = O(r^{-1}), \tag{7.3}$$

where $r = |\mathbf{r}|$ is the distance at which the interaction occurs. For example, the amplitude (1.32) for scattering from a spherically symmetric potential $V(r)$ becomes

$$\mathcal{M}(q) = 4\pi \int_0^\infty V(r) \left[\frac{\sin(qr)}{qr} \right] r^2 \mathrm{d}r \tag{7.4}$$

on integrating over all angular directions. The dominant contributions arise from r values of order q^{-1} as asserted, since for smaller r the integrand is suppressed by the factor r^2, while for large r it is suppressed by the average over the rapidly oscillating sine factor. Hence in discussing scattering from the static potential (7.2a), in which the strong coupling constant α_s decreases with decreasing r, we can equally well regard it as decreasing with increasing momentum transfer, since the one implies the other by (7.3).

In this discussion, we have considered scattering from a static potential leaving the energy of the particle unchanged. Particle energies are also unchanged in elastic scattering in the centre-of-mass frame, but in other reference frames, and in other processes, energy as well as momentum can be exchanged between the particles. In such cases, the strength of the interaction can be shown to depend on

$$\mu^2 \equiv \left| \mathbf{q}^2 - E_q^2 \right|, \tag{7.5}$$

[5] A discussion of such states is given in Close (1997 and 2007).

which is Lorentz-invariant and reduces to $\mathbf{q}^2$ when the energy exchanged E_q is zero. More specifically, it can be shown that the QCD coupling constant α_s is given to a good approximation by

$$\alpha_s(\mu) = \alpha_s(\mu_0)\left[1 + \frac{(33 - 2N_f)}{6\pi}\alpha_s(\mu_0)\ln(\mu/\mu_0)\right]^{-1} \tag{7.6}$$

for $\mu^2 \gg 1\,\text{GeV}^2$. Here N_f is the number of quark flavours $u, d, s, \ldots$,[6] with $4m_q^2 < \mu^2$, and $\alpha_s(\mu_0)$ is the value of the coupling at a chosen reference value μ_0, which is usually taken to be $\mu_0 = M_Z$, the mass of the Z boson. Measured values of $\alpha_s(\mu)$, obtained from a variety of different processes, are shown in Figure 7.3, where the curves show the predicted behaviours corresponding to the 'best-fit' value

$$\alpha_s(M_Z) = 0.118 \pm 0.002 \tag{7.7}$$

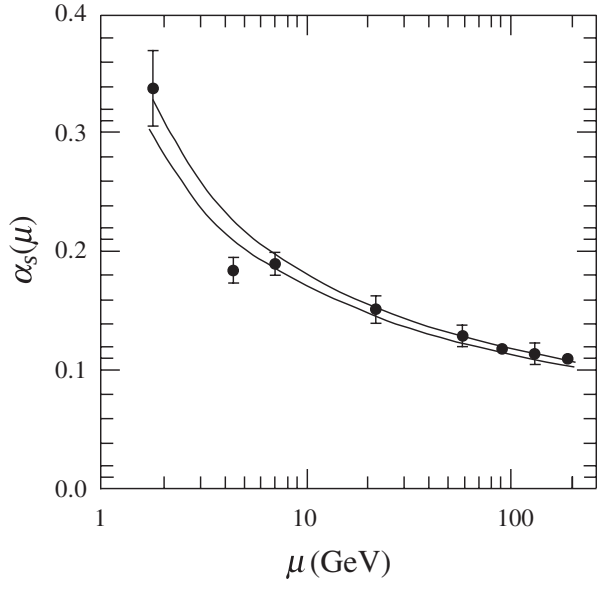


Figure 7.3 Values of the running coupling constant α_s obtained from the following experiments at increasing values of μ: τ decay; Υ decay; deep inelastic lepton scattering; e^+e^- annihilation at 22 and 50 GeV; Z^0 decay; and e^+e^- annihilation at 135 and 189 GeV. The solid curves show the evolution of α_s with μ, as predicted by QCD, assuming the value (7.7) for $\alpha_s(M_Z)$. (Reprinted by permission of Institute of Physics (IOP), Fig. 9.2, W.-M. Yao *et al.*, *Journal of Physics*, **G33**, 1, 2006.)

[6] The reason for the dependence on N_f is discussed qualitatively in Section 7.1.2 below. The change in $\alpha_s(\mu)$ at $\mu^2 = 4m_c^2$, $4m_b^2$ is, of course, not really discontinuous as implied by the approximation (7.6), but is 'smoothed out' over a threshold region.

at the reference value $\mu_0 = M_Z$. The decrease in $\alpha_s(\mu)$ as μ increases, corresponding to shorter distances, is clearly seen, and because of this variation, $\alpha_s(\mu)$ is often referred to as the *running coupling constant*.

★7.1.2 Screening, antiscreening and asymptotic freedom

A derivation of asymptotic freedom is far beyond the scope of this book, and we will content ourselves here with a qualitative discussion, making plausible its connection to gluon self-interactions. Firstly, however, it will be helpful to take a closer look at QED, where the analogous photon self-interactions are absent.

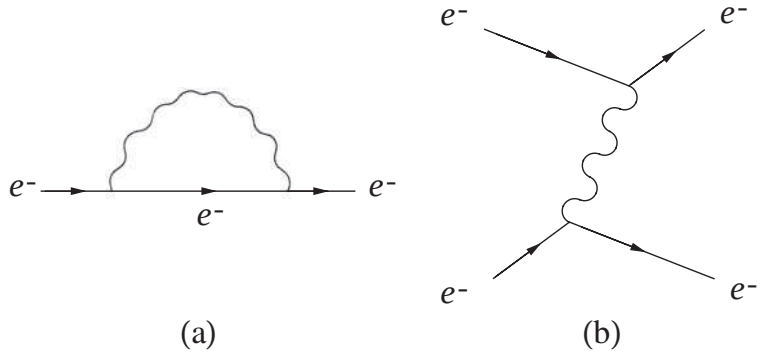


Figure 7.4 The simplest quantum fluctuation of an electron and the associated exchange process.

Students studying QED for the first time, when meeting the one-photon exchange contribution to electron scattering of Figure 1.7, often ask the question: 'When the first electron emits the photon, how can it know that there is a second electron close enough to absorb it in the short time allowed by the energy–time uncertainty principle?' Of course the electron knows no such thing, and the question suggests itself because our earlier account is incomplete. According to QED, what happens is that the electron emits and reabsorbs virtual photons all the time, as shown in Figure 7.4(a), whether another electron is nearby or not. If, however, a second electron is nearby, it will sometimes absorb the photon before it is reabsorbed by the first electron, giving rise to the familiar exchange diagram of Figure 7.4(b).[7]

Processes in which one particle is converted for a short time to two or more particles are called *quantum fluctuations*. Figure 7.4(a) is only the simplest quantum fluctuation of an electron; others are possible, such as that shown in Figure 7.5(a). This is similar to Figure 7.4(a), except that in this case the emitted photon itself fluctuates to an electron–positron pair before it is reabsorbed by the initial electron. Thus an electron

[7] In interpreting these diagrams, it is important to remember that the particles do not travel on well-defined classical trajectories. For example, an argument analogous to that given in Section 1.4.1 shows that the de Broglie wavelength of the exchanged photon in Figure 7.4(b) is greater than, or about equal to, the distance between the two vertices; and the uncertainty in the position of the photon must be of at least the same order. It is possible, therefore, for it to be absorbed by either electron, despite the distance between them.

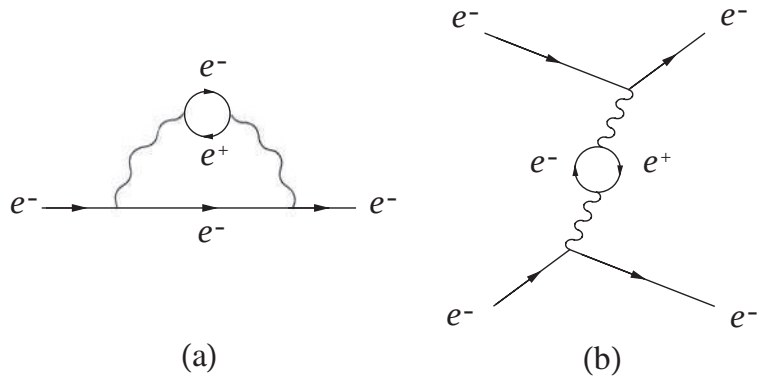


Figure 7.5 A more complicated quantum fluctuation of the electron, together with the associated exchange process.

continuously emits and reabsorbs not only photons but also, indirectly, electron–positron pairs. These pairs have small but measurable consequences called *vacuum polarization effects*[8] by analogy with those observed when a charge is immersed in a dielectric medium. If the charge is positive, then the molecules of the dielectric will tend to align themselves as shown in Figure 7.6, and if a small test charge is introduced at a distance that is large compared with the size of the molecules, it will experience a force that is the same as if the charge Q had been replaced by a smaller 'effective charge' Q_{eff}. However, at distances between the charges that are of the same order as the molecular sizes, this screening becomes less effective, so that Q_{eff} increases as the charges get closer together. There is a somewhat similar effect in QED, even in a vacuum, associated with the electron–positron pairs produced by quantum fluctuations. For example, if a second electron is close by, it will sometimes absorb the second photon in Figure 7.5(a) before it is reabsorbed by the initial electron, giving rise to Figure 7.5(b). This diagram contributes to electron–electron scattering, and a detailed calculation shows that at short distances,

$$r \leq r_c \equiv \hbar/mc = 3.9 \times 10^{-13}\,\mathrm{m}, \tag{7.8}$$

it leads to corrections that are of order α smaller than the familiar Coulomb potential arising from the one-photon exchange diagram of Figure 7.4(b). Specifically, if we write the effective potential in the form

$$\phi_{eff} = \alpha_{eff}/r, \tag{7.9}$$

then

$$\alpha_{eff} = \alpha \approx 1/137 \quad (r \gg r_c), \tag{7.10}$$

[8] The calculation of these and other effects due to quantum fluctuations presents considerable difficulties, which took many years to overcome. Here we will simply state the results of such calculations without further comment. They are described in, for example, Chapter 9 of Mandl and Shaw (1993).

but for $r < r_c$ it is somewhat larger, and increases as r becomes smaller. In other words, the interaction strength increases at very short distances, and because of this one can show, for example, that the $2s$ state in hydrogen is more tightly bound (by 2.2×10^{-17} eV) than it would be in a pure Coulomb potential. This effect is tiny, but nonetheless is present and must be taken into account if the very precise experimental data on the hydrogen spectrum are to be fully understood. [9]

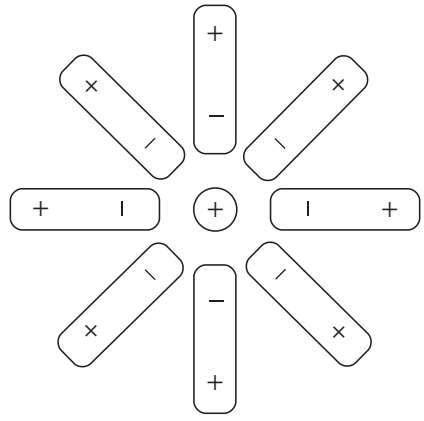


Figure 7.6 Schematic diagram representing the polarization of the molecules of a dielectric by a positive charge placed within it.

Quantum fluctuations also exist in QCD, and as in QED they also lead to a variation in the interaction strength with distance. Specifically, if we consider quark–quark scattering there are two lowest-order vacuum polarization corrections to the one-gluon exchange diagram of Figure 7.1. These are shown in Figure 7.7. The first of these is a direct analogue of Figure 7.5(b) in electron–electron scattering, and like it leads to a screening correction. If this were the only contribution, the interaction would grow stronger at short distances as in QED. However, there is also the second diagram of Figure 7.7(b), involving a gluon–gluon pair produced in a gluon self-interaction of the type shown in Figure 7.2(a). This diagram has no counterpart in QED, and the nature of its contribution is far from obvious. In fact, detailed calculation shows that it leads to an *antiscreening* effect, meaning that it causes the interaction to grow weaker at short distances. This effect is bigger than the screening correction from Figure 7.7(a), and the net result is that the interaction grows weaker at short distances; i.e. there is asymptotic freedom.

★7.1.3 The quark–gluon plasma

In ordinary matter, quarks are confined within hadrons and an atomic nucleus, for example, can be thought of as a bound state of protons and neutrons. However, as the energy density is increased, a phase transition can occur to a state in which

[9] This is discussed in pp. 207–208 of Mandl and Shaw (1993).

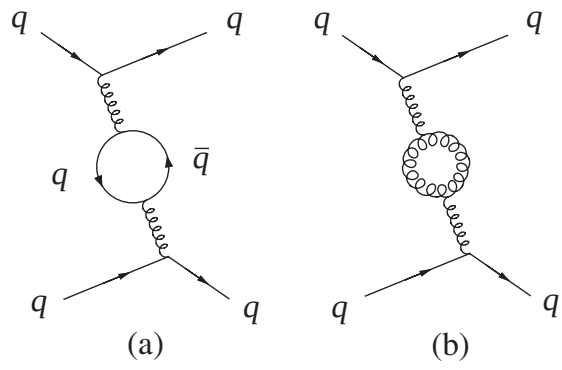


Figure 7.7 The two lowest-order vacuum polarization corrections to one-gluon exchange in quark–quark scattering.

individual hadrons lose their identities, and quarks and gluons become free to move across a volume that is large compared to a hadron. Approximate lattice gauge theory calculations suggest this should occur at an energy density of order 1 GeV fm^{-3}, i.e. about 6 times the energy density at the centre of a heavy nucleus, and the resulting new state of matter is called a *quark–gluon plasma.* A quark–gluon plasma is believed to have existed in the first few microseconds after the 'big bang' and it may exist today at the centre of neutron stars. More prosaically, a quark–gluon plasma may be created briefly in collisions between heavy ions, if the collision energy is large enough. The steps in the formation of such a plasma, and its subsequent expansion and cooling to yield many hadrons, are illustrated schematically in Figure 7.8.

High-energy collisions between heavy ions have been studied at CERN and with the relativistic heavy ion collider (RHIC) at the Brookhaven National Laboratory. The RHIC typically collides two counter-circulating beams of fully stripped gold ions at a maximum energy of 200 GeV per nucleon. If the ions collide centrally (i.e. head-on) several thousand final state particles are produced. An example of an event seen in the STAR detector (see Figure 4.20 and Plate 7) is shown in Figure 7.9. A key question is whether the energy-density in the collisions is sufficient to have created a quark–gluon plasma and its subsequent cooling phases. There are many signatures for this, including the relative abundances of different final state particle types. For example, the large numbers of gluons in the plasma would lead to copious production of $s\bar{s}$ pairs via gluon fusion $gg \to s\bar{s}$, and hence production of strange particles in excess of that expected from nucleon–nucleon collisions at very high energies. On the other hand, the production of J/Ψ would be suppressed because the c and $\bar{c}$ quarks produced (also from gluon fusion) would be separated by many quarks of other flavours, leading instead to the production of charmed mesons, e.g. the D mesons. In practice, these arguments depend on how long the quarks remain in the central region of the plasma, and this will lead to angular dependences that provide the basis for more detailed tests. Present measurements are all consistent with the predicted energy density at which hadrons would be formed, while that of the initial fireball is considerably higher.

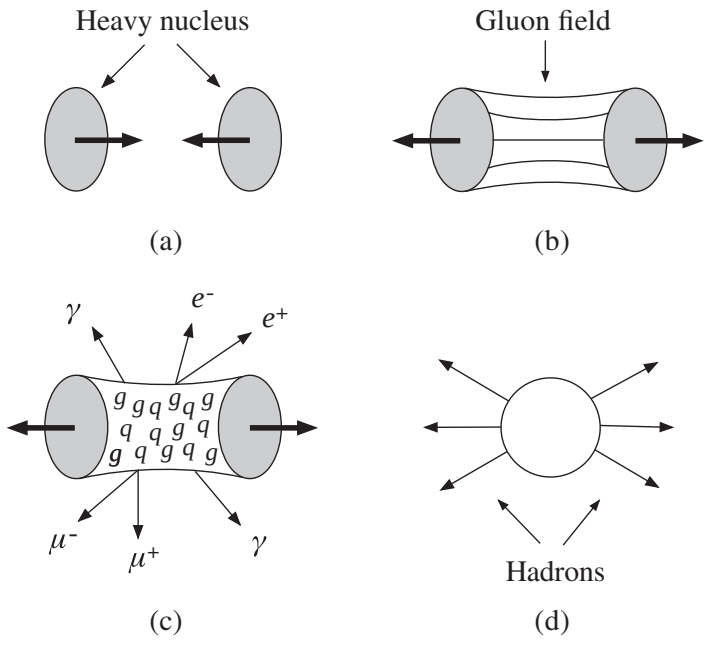


Figure 7.8 Stages in the formation of a quark-gluon plasma and subsequent hadron emission: (a) two heavy nuclei collide at high energies and (b) interact via the gluon field; (c) the very high energy-density produced causes the quarks and gluons to deconfine and form a plasma that can radiate photons and lepton pairs; (d) finally, as the plasma cools, hadrons condense and are emitted. (Reprinted with permission from the National Academies Press, Copyright 1999, National Academy of Sciences.)

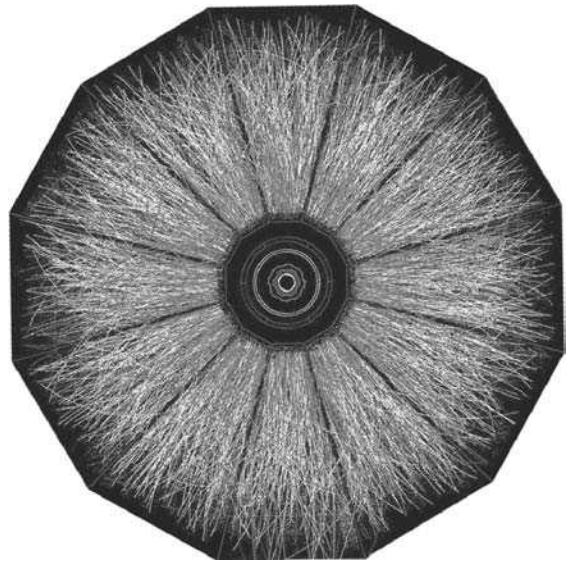


Figure 7.9 View of a 200 GeV gold–gold interaction in the STAR detector at the RHIC accelerator. (Courtesy of Brookhaven National Laboratory.)

Future experiments at RHIC and at the LHC will play a crucial role in understanding the basic nature of deconfinement. Questions to be addressed include: 'What is the nature of matter at the highest densities (experiments at RHIC suggest that the plasma behaves more like a liquid than a gas)?', 'Under what conditions can a quark–gluon plasma be made?' and 'What are the rules governing the evolution and the transition to and from this kind of matter?'

7.2 ELECTRON–POSITRON ANNIHILATION

In this section we shall consider reactions of the type

$$e^+ + e^- \to \text{hadrons} \tag{7.11}$$

in the centre-of-mass energy range 15–40 GeV. Such reactions have been extensively studied in electron–positron colliding-beam experiments, and give clear evidence for the existence of colour and gluons. Their study will also introduce us to jets of hadrons, like those shown in Figure 4.12, which are produced in a wide variety of high-energy processes. These jets are closely related to the underlying quark and gluon interactions, and are the closest thing to quark and gluon 'tracks' that we are ever likely to see.

7.2.1 Two-jet events

In the centre-of-mass energy range 15–40 GeV, electron–positron annihilation into hadrons (7.11) is dominated by processes like that shown in Figure 7.10.[10] These can be regarded as occurring in two stages: a primary electromagnetic process

$$e^+ + e^- \to q + \bar{q}, \tag{7.12}$$

leading to the production of a quark–antiquark pair, followed by a strong interaction process, called *fragmentation*, which converts the high-energy $q\bar{q}$ pair into two jets of hadrons. These jets are emitted in opposite directions in the centre-of-mass frame in order to conserve momentum, and a typical example of such an event observed in an electron–positron colliding-beam experiment is shown in Figure 4.12.

The fragmentation process that converts the quarks into hadrons is very complicated, and the composition of the jets (i.e. the numbers and types of particles in the jet and their momenta) varies from event to event. However, the direction of a jet, defined by the total momentum

$$\mathbf{P} = \sum_i \mathbf{p}_i, \tag{7.13}$$

where the sum extends over all the particles within the jet, reflects closely the parent quark or antiquark direction. This is because the QCD interaction is relatively weak at very short distances, and the quark and antiquark do not interact strongly until they are

[10] At lower energies, resonance formation processes like Figure 6.3 contribute, as we saw in Section 6.4.1, while at higher energies, effects due to weak interactions become important, as we shall see in Chapter 9.

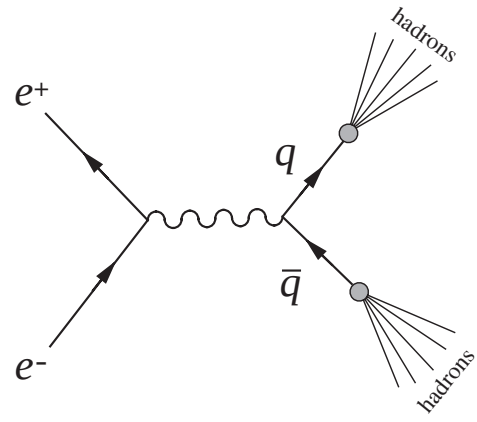


Figure 7.10 Basic mechanism of two-jet production in electron–positron annihilation.

separated by a distance r of order 1 fm, which according to (7.3) gives rise to a typical momentum transfer of order 200 MeV/c between them. This is small compared with the momentum of the quark and antiquark produced in the initial reaction (7.12), and the total momenta (7.13) of the jets that subsequently develop point almost exactly in the initial quark and antiquark directions; i.e. the jet angular distributions reflect the angular distributions of the quark and antiquark in the basic reaction (7.12).

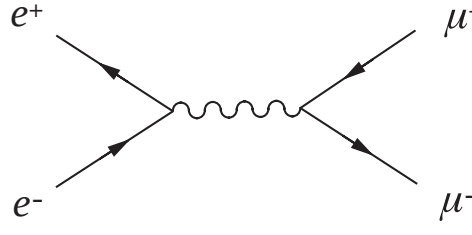


Figure 7.11 Dominant mechanism for electron–positron annihilation to muon pairs.

The above interpretation of the two-jet events is confirmed by comparing their angular distributions with those of muons produced in the reaction

$$e^+ + e^- \rightarrow \mu^+ + \mu^- \tag{7.14}$$

by the mechanism of Figure 7.11. This mechanism is identical to that assumed for quark pair production in Figure 7.10, so that the two reactions are very closely related. At high beam energies $E \gg m_\mu$ the differential cross-section for (7.14) is[11]

[11] This cross-section can be calculated theoretically from the mechanism of Figure 7.11 (see, for example, pp. 146–150 of Mandl and Shaw (1993). However, it is sufficient for our purposes to take (7.15) from experiment.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta}\left(e^+ + e^- \to \mu^+ + \mu^-\right) = \frac{\pi\alpha^2}{8E^2}\left(1+\cos^2\theta\right), \tag{7.15}$$

where θ is the production angle of either muon with respect to the initial electron direction in the centre-of-mass frame.[12] The corresponding cross-section for producing quark–antiquark pairs $a\bar{a}$ of a given flavour $a = u, d, s, \ldots$ is then obtained by replacing the muon charge e with the quark charge ee_a, and since this charge is the same for all three colour states r, g and b, we must multiply by three to account for the three colour states $r\bar{r}, g\bar{g}$ and $b\bar{b}$, giving

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta} = \frac{3\pi\, e_a^2\alpha^2}{8E^2}\left(1+\cos^2\theta\right) \tag{7.16}$$

overall. In particular, the angular distributions of the jets should be proportional to

$$(1+\cos^2\theta), \tag{7.17}$$

irrespective of which quark is produced, if the jet directions reflect the initial quark and antiquark directions, as we have assumed. This is indeed the case, as is shown in Figure 7.12.

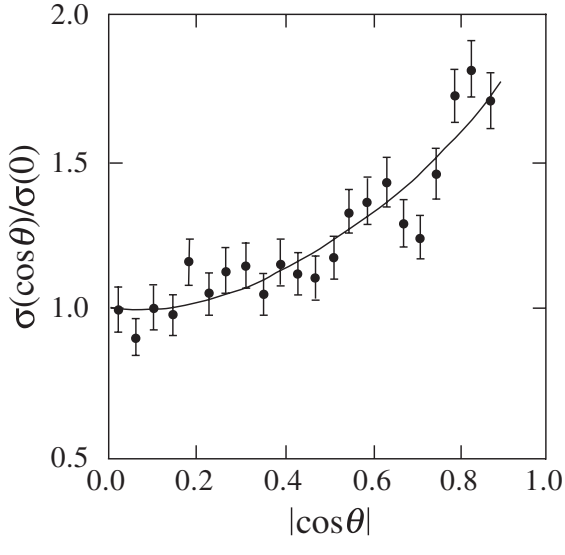


Figure 7.12 Comparison between the angular distribution of two-jet events observed in an electron–positron colliding-beam experiment by the CELLO collaboration at DESY, and the theoretically predicted $(1+\cos^2\theta)$ behaviour (7.17). (Reprinted from Behrend, H.-J., *et al.*, *Physics Letters B*, **183**, 400. Copyright 1987, with permission from Elsevier.)

[12] The μ^+ and μ^- emerge in opposite directions in the centre-of-mass frame, so that $\theta_+ = \pi - \theta_-$, in an obvious notation, implying $\cos^2\theta_+ = \cos^2\theta_- = \cos^2\theta$.

7.2.2 Three-jet events

The dominant process in e^+e^- annihilation to hadrons at high energies is the formation of 'back-to-back' jets in the centre-of-mass frame, as discussed above and illustrated in Figure 7.13(a). However, sometimes we might expect a high-momentum gluon to be emitted at a wide angle by the quark or antiquark before fragmentation occurs, leading to the formation of a three-jet event as illustrated in Figure 7.13(b).

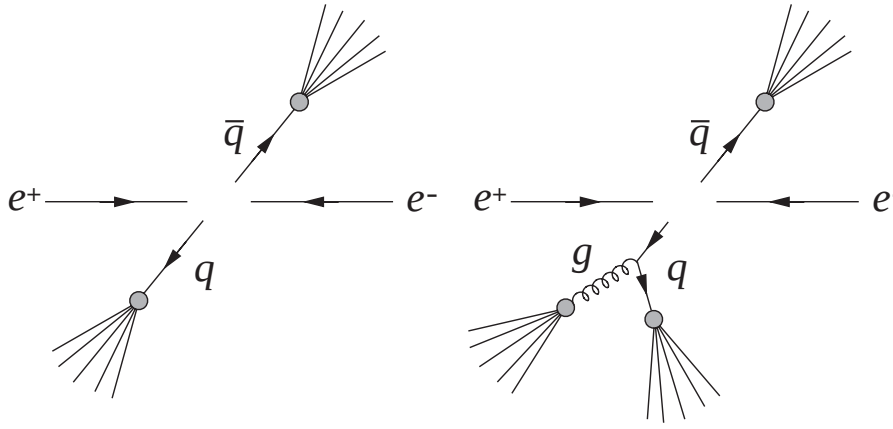


Figure 7.13 Schematic diagrams representing (a) two-jet and (b) three-jet formation in electron–positron annihilation in the centre-of-mass frame.

Such events are indeed observed, as shown in Figure 7.14, and provide strong evidence for the spin-1 gluons whose existence is assumed in QCD. The analysis is not straightforward because it is not obvious which of the jets in Figure 7.14 is due to the gluon. Instead one focuses on the total energy of each jet. The jets are classified in the order $E_1 > E_2 > E_3$ in the overall centre-of-mass frame, before transforming to the mutual centre-of-mass frame of jets 2 and 3. The angular distribution of jet 1 is then plotted with respect to the angle ϕ between its direction and the mutual line-of-flight of jets 2 and 3. The latter is almost certainly coincident with the line-of-flight of the gluon, since from Figure 7.13(b) we would expect the highest-energy jet to be associated with the quark or antiquark which does not emit the gluon, and detailed calculations show that the distribution described is sensitive to the gluon spin. The measured distribution is compared with the theoretical expectations for spin-0 and spin-1 gluons in Figure 7.15, and clearly favours the latter.

The probability that a quark will emit a gluon is determined by the strong coupling constant α_s in the same way that the probability that an electron will emit a photon is determined by the fine structure constant α. Hence the observed rate for three-jet events can be used to determine a value of the strong coupling constant α_s. The analysis is, however, very complicated, and we will content ourselves with noting that one of the experimental values shown in Figure 7.3 was obtained in this way.

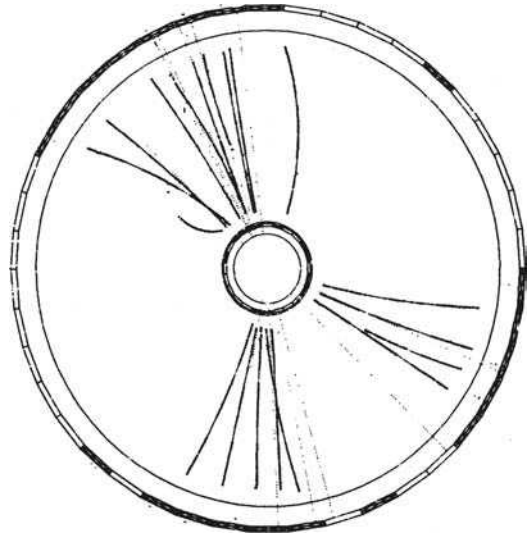


Figure 7.14 Computer reconstruction of a three-jet event observed by the JADE collaboration at DESY. The curved tracks correspond to charged particles and the dotted lines to neutral particles whose trajectories are unaffected by the magnetic field. (Reprinted from Wu, S.-L., *Physics Reports*, **107**, 59. Copyright 1984, with permission from Elsevier.)

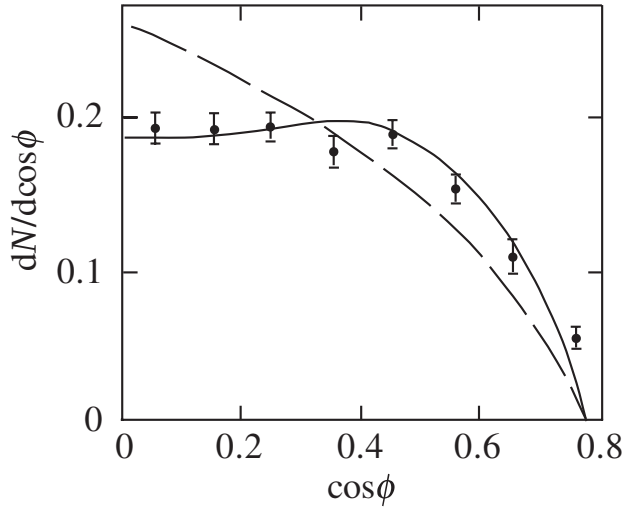


Figure 7.15 Angular distribution of three-jet events measured in the TASSO collaboration, where ϕ is the angle defined in the text. The dashed and solid lines show the theoretical expectations for spin-0 and spin-1 gluons respectively. (After Brandelik *et al.*, 1980, and Wu, 1984. Copyright 1984, with permission from Elsevier.)

7.2.3 The total cross-section

The measured values of the total cross-section for electron–positron annihilation to hadrons are shown in Figure 7.16 in terms of the ratio

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}, \tag{7.18}$$

where the total cross-section for muon production

$$\sigma(e^+e^- \to \mu^+\mu^-) = \frac{\pi\alpha^2}{3E^2} \tag{7.19}$$

follows from integrating the differential cross-section (7.15) over all angles θ. The near constancy of this ratio follows from the dominance of the two-step mechanism of Figure 7.10, with the total annihilation rate being determined by that of the initial reaction (7.12). while the value of the ratio directly confirms the existence of three colour states, each with the same electric charge, for a given quark flavour.

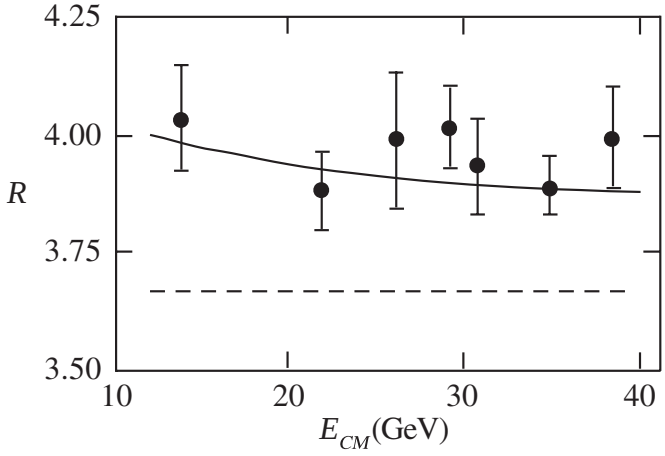


Figure 7.16 Comparison between the measured values of the cross-section ratio R of Equation (7.18) and the theoretical prediction (7.22) for three colours, $N_C = 3$. The dashed line shows the corresponding prediction (7.21) omitting small contributions of order α_s. (Data from the compilations of Wu, 1984, and Behrend *et al.*, 1987.)

To understand this, let us suppose that each quark flavour $a = u, d, s, \ldots$ exists in N_C colour states so that $N_C = 3$ according to QCD, while $N_C = 1$ if the colour degree of freedom does not exist. Since the different colour states all have the same electric charge, they will all be produced equally readily by the mechanism of Figure 7.10, and the cross-section for producing quark pairs of any given flavour $a = u, d, s, \ldots$ will be proportional to the number of colours N_C. The differential cross-section for three colours is given by (7.16), and on multiplying by $N_C/3$ and integrating over all angles, we obtain

$$\sigma(e^+e^- \to q\bar{q}) = N_C\, e_a^2 \sigma(e^+e^- \to \mu^+\mu^-) \tag{7.20}$$

for the total cross-section, where the muon pair production cross-section is given by (7.19). Hence if hadron production were completely dominated by the two-step process of Figure 7.10, we would have

$$R = R_0 \equiv N_C(e_u^2 + e_d^2 + e_s^2 + e_c^2 + e_b^2) = 11N_C/9, \tag{7.21}$$

because quark–antiquark pairs of five flavours can be produced at the energies we are considering. The small contribution from the three-jet events of Figure 7.13(b) is much harder to calculate, but it can be shown that when it and other corrections of order α_s are taken into account, (7.21) is modified to

$$R = R_0(1 + \alpha_s / \pi), \tag{7.22}$$

where α_s is given by (7.6) evaluated at μ^2 equal to the centre-of-mass energy squared. Although these corrections of order α_s are very small compared with the dominant contribution (7.21), they must be included if the most precise experimental data on R, which have errors of order 2–3 %, are to be accounted for. As can be seen from Figure 7.16, these data are in excellent agreement with the theoretical prediction (7.22) for the value $N_C = 3$, in accord with the basic assumption of QCD.

7.3 ELASTIC ELECTRON SCATTERING: THE SIZE OF THE PROTON

Electron and muon scattering have played a crucial role in exploring the structure of the hadrons. Elastic scattering led to the first measurements of the size of the proton, while inelastic scattering experiments gave the first clear evidence for scattering from individual quarks within the nucleon. Here we will discuss the elastic scattering reactions

$$\ell^- + p \to \ell^- + p \qquad (\ell = e, \mu), \tag{7.23}$$

where the same arguments apply to electrons and muons, since they have identical electromagnetic interactions. In particular, both reactions (7.23) are dominated by the single-photon exchange mechanisms of Figure 7.17, where the 'blob' indicates an interaction whose nature we seek to explore. In this section we shall consider only these single-photon exchange contributions, neglecting completely the small corrections of higher order in α arising from multiphoton exchange.

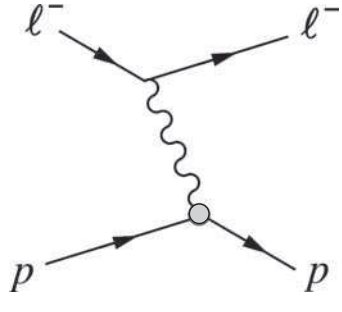


Figure 7.17 Dominant one-photon exchange mechanism for elastic lepton–proton scattering, where $\ell = e$ or μ.

7.3.1 Static charge distributions

At high energies, electron–proton scattering is quite complicated, as we shall see shortly. It is therefore useful first to consider nonrelativistic electron scattering from static charge distributions. The starting point for this is the *Rutherford cross-section*

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_R = \frac{m^2\alpha^2}{4p^4\sin^4(\theta/2)} \tag{7.24}$$

for scattering of an electron of momentum p through an angle θ by a static point charge e. This formula is derived in starred Section 7.3.3, where we also show how it is modified if the same total charge is spread out in a spherically symmetric density distribution $e\rho(r)$, where

$$\int \rho(r)\,\mathrm{d}^3\mathbf{r} = 1 \tag{7.25}$$

and $r = |\mathbf{r}|$ as usual. In this case the cross-section (7.24) is replaced by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_R G_E^2(q^2), \tag{7.26}$$

where the *form factor*

$$G_E(q^2) = \int \mathrm{d}^3\mathbf{r}\rho(r)e^{i\mathbf{q}\cdot\mathbf{r}} \tag{7.27}$$

is the Fourier transform of the charge distribution with respect to the momentum transfer

$$\mathbf{q} = \mathbf{p} - \mathbf{p}', \tag{7.28}$$

where $\mathbf{p}$ and $\mathbf{p}'$ are the initial and final electron momenta.[13] For $\mathbf{q} = \mathbf{0}$, (7.27) reduces to the integral for the density $\rho(r)$ so that

$$G_E(0) = 1 \tag{7.29}$$

with our normalization (7.25). On the other hand, for very large momentum transfers

$$G_E(q^2) \to 0 \quad \text{as} \quad q^2 \to \infty, \tag{7.30}$$

because of the increasingly rapid oscillation of the exponential factor in (7.27).

Clearly, measurements of the cross-section (7.26) determine the form factor, and hence the Fourier transform of the charge distribution that has caused the scattering.

[13] Despite appearances, the right-hand side of (7.27) is a function of $q^2 = |\mathbf{q}^2|$ only, as can be seen by explicitly integrating over the angles (cf. Problem 7.2).

In particular, it follows from (7.27) that the root-mean-square (rms) charge radius r_E is given by (see Problem 7.2)

$$r_E^2 \equiv \overline{r^2} = \int \mathrm{d}^3\mathbf{r}\; r^2 \rho(r) = -6 \frac{\mathrm{d}G_E(q^2)}{\mathrm{d}q^2}\bigg|_{q^2=0}. \tag{7.31}$$

7.3.2 Proton form factors

In discussing electron scattering from the proton, several other factors must be taken into account. In addition to the *electric form factor* G_E associated with the charge distribution, there is also a *magnetic form factor* G_M associated with the magnetic moment distribution within the proton. For momentum transfers that are much smaller than the proton mass M, the recoil energy of the proton is negligible, and the electric form factor is again given by (7.27), with an analogous expression for the magnetic form factor G_M. However, at high momentum transfers the recoil energy of the proton cannot be neglected, so that the initial and final electron energies are no longer equal. In this regime, it can be shown that the form factors are functions of the Lorentz-invariant generalization of the momentum transfer

$$Q^2 \equiv (\mathbf{p} - \mathbf{p}')^2 - (E - E')^2, \tag{7.32}$$

where the initial and final electrons have momenta and energies $(\mathbf{p}, E)$ and $(\mathbf{p}', E')$, respectively, and the interpretation of the form factor in terms of static charge and magnetic moment distributions breaks down. It is, however, regained in the limit of low momentum transfers, when Q^2 reduces to our previous variable q^2; in particular, Equation (7.31) for the rms radius of the proton remains valid. Finally, in addition to recoil effects and the magnetic form factor of the proton, we must also take account of the magnetic moment of the electron, which is assumed to be a point particle without internal structure, throughout the analysis. When all these factors are taken into account, the cross-section formula becomes quite complicated, and we shall simply quote it in the limit that the electron mass can be neglected compared with its energy. The differential cross-section is then given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \left(\frac{E'}{E}\right) \left[G_1(Q^2) \cos^2(\theta/2) + 2\tau\, G_2(Q^2) \sin^2(\theta/2)\right]. \tag{7.33}$$

Here

$$G_1(Q^2) = \frac{G_E^2 + \tau G_M^2}{1+\tau}, \quad G_2(Q^2) = G_M^2 \tag{7.34a}$$

with

$$\tau = Q^2/4M^2, \tag{7.34b}$$

and the form factors are normalized so that

$$G_E(0) = 1 \quad \text{and} \quad G_M(0) = \mu_p, \tag{7.35}$$

where $\mu_p = 2.79$ is the proton magnetic moment in nuclear magnetons. The same formulas hold for muon elastic scattering in the limit that the muon mass can be neglected compared with its energy.

The values of the form factors G_E and G_M have been measured in many muon and electron scattering experiments, starting with the pioneering work of Hofstadter and co-workers in 1956, which gave the first measurement of the finite size of the proton. The results are conveniently divided into three Q^2 regions.

At low Q^2, the contributions of the electric form factor G_E dominate the cross-section (7.33), since the contributions from the magnetic form factor are suppressed by a factor $\tau \ll 1$. Hence G_E can be precisely measured, enabling the size of the proton – or more correctly its rms charge radius – to be determined by (7.31). The average value from different analyses is

$$r_E = 0.85 \pm 0.02\,\text{fm}. \tag{7.36}$$

In the intermediate range $0.02 \geq Q^2 \leq 3\,\text{GeV}^2$, both form factors make measurable contributions, which can be separated by varying the scattering angle θ, as can be seen from (7.33). The results in this region can be summarized to within a few percent by the approximate 'dipole fit'

$$G_E(Q^2) \approx \frac{G_M(Q^2)}{\mu_p} \approx G_D \equiv \left(\frac{\beta^2}{\beta^2 + Q^2}\right)^2, \tag{7.37}$$

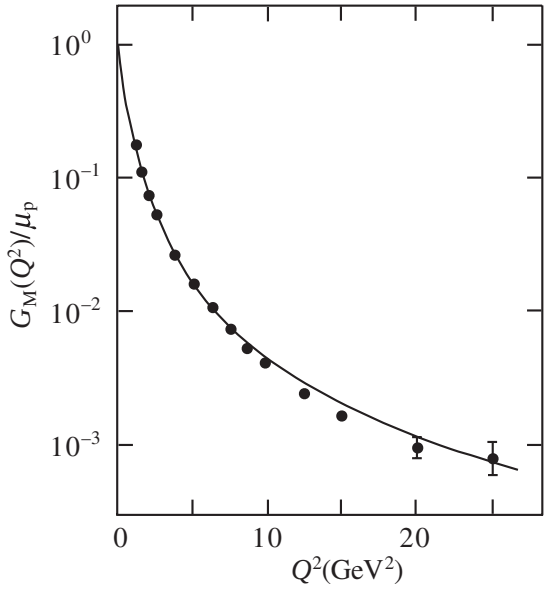


Figure 7.18 Comparison between the measured values of the proton magnetic form factor and the 'dipole fit' of Equation (7.37). (After Panofsky, 1968.)

where $\beta = 0.84\,\text{GeV}$ and Q^2 is measured in GeV^2. Finally, for $Q^2 > 3\,\text{GeV}^2$, the contributions of the electric form factor become negligible, and only the magnetic form factor G_M can be measured with reasonable accuracy.

The resulting values for $G_M(Q^2)$ are compared with the predictions of the dipole fit (7.37) in Figure 7.18. As can be seen, the form factor continues to fall rapidly with Q^2, decreasing by three orders of magnitude over the Q^2 range for which it has been measured. This is in marked contrast to the behaviour of a hypothetical point proton, with a Dirac magnetic moment, for which the form factors reduce to[14]

$$G_E(Q^2) = 1 \quad \text{and} \quad G_M(Q^2) = 1, \tag{7.38}$$

independent of Q^2.

★7.3.3 The basic cross-section formulas

It remains to give the promised derivations for the nonrelativistic cross-sections that were the starting point for our whole discussion. These are the Rutherford cross-section (7.24) for scattering from a Coulomb potential

$$V_C(\mathbf{r}) = -\frac{e^2}{4\pi\varepsilon_0}.\frac{1}{|\mathbf{r}|} = -\frac{\alpha}{r}, \tag{7.39}$$

and the modified cross-section (7.26) for scattering from the electrostatic potential

$$V_\rho(\mathbf{r}) = -\frac{e^2}{4\pi\varepsilon_0}\int \frac{\mathrm{d}^3\mathbf{r}'\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{7.40}$$

due to a spherically symmetric charge distribution

$$\rho(\mathbf{r}) \equiv \rho(r). \tag{7.41}$$

These cross-sections follow directly from the Born approximation[15]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{m^2}{4\pi^2}\,|\mathcal{M}(\mathbf{q})|^2 \tag{7.42a}$$

for scattering from an arbitrary potential $V(\mathbf{r})$ in lowest-order perturbation theory,[16] where

$$\mathcal{M}(\mathbf{q}) = \int \mathrm{d}^3\mathbf{r}V(\mathbf{r})\exp(i\mathbf{q}\cdot\mathbf{r}) \tag{7.42b}$$

---

[14] These results follow directly from (7.27) and the analogous relation for G_M in the nonrelativistic limit. They continue to hold relativistically.

[15] For a discussion see, for example, pp. 397–399 of Mandl (1992).

[16] This gives cross-sections of order α^2, corresponding to the lowest-order diagram of Figure 7.17. Higher orders in perturbation theory give additional powers of α corresponding to higher-order diagrams, which we are neglecting (cf. the discussion for scattering from a Yukawa potential in Section 1.4.2).

and **q** is the momentum transfer. Unfortunately the integral (7.42b) is ill-defined for a pure Coulomb potential, because it diverges at large r. However, charges in nature are in practice always screened at large distances by intervening matter, and it is reasonable to regard the pure Coulomb potential as a screened potential in the limit where the screening is removed; i.e. we interpret

$$\mathcal{M}_{\mathrm{C}}(\mathbf{q}) = \int \mathrm{d}^3\mathbf{r} \left(-\frac{\alpha}{r}\right) \exp(i\mathbf{q}\cdot\mathbf{r})$$

as

$$\mathcal{M}_{\mathrm{C}}(\mathbf{q}) = \underset{\mu\longrightarrow 0}{\mathrm{Lt}} \int \mathrm{d}^3\mathbf{r} \left(-\frac{\alpha e^{-\mu r}}{r}\right) \exp(i\mathbf{q}\cdot\mathbf{r}),$$

which gives

$$\mathcal{M}_C(\mathbf{q}) = -\frac{4\pi\alpha}{q^2} \tag{7.43}$$

on evaluating the integral (cf. Problem 7.3). Finally, we note that the magnitude of the electron momentum is unchanged by scattering from a static potential, so that $|\mathbf{p}| = |\mathbf{p}'| = p$, $\mathbf{p}\cdot\mathbf{p}' = p^2\cos\theta$ and

$$q^2 = (\mathbf{p}-\mathbf{p}')^2 = 4p^2\sin^2(\theta/2). \tag{7.44}$$

The Rutherford cross-section formula (7.24) then follows on substituting Equations (7.43) and (7.44) into (7.42a).

The corresponding cross-section for scattering from a static charge distribution is obtained by substituting the potential (7.40) into (7.42b). This gives

$$\mathcal{M}_\rho(\mathbf{q}) = \int \mathrm{d}^3\mathbf{r} V_\rho(\mathbf{r}) \exp(i\mathbf{q}\cdot\mathbf{r}) = -\frac{e^2}{4\pi\varepsilon_0} \int \mathrm{d}^3\mathbf{r}'\rho(\mathbf{r}') \left\{ \int \mathrm{d}^3\mathbf{r} \frac{\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}}{|\mathbf{x}-\mathbf{x}'|} \right\}, \tag{7.45}$$

where we have changed the order of integration in the double integral. Changing variables from **x** to $\mathbf{x}-\mathbf{x}'$ and using (7.43) then gives

$$\mathcal{M}_\rho(\mathbf{q}) = \mathcal{M}_C(\mathbf{q}) G_E(\mathbf{q}), \tag{7.46}$$

where

$$G_E(\mathbf{q}) \equiv \int \mathrm{d}^3\mathbf{r}\rho(\mathbf{r}) \exp(i\mathbf{q}\cdot\mathbf{r}). \tag{7.47}$$

For a spherically symmetric charge distribution, $G_E(\mathbf{q})$ reduces to a function of q^2 (see Problem 7.2) and the desired cross-section formula (7.26) follows on substituting (7.46) into (7.42a) and using (7.43) and (7.44).

7.4 INELASTIC ELECTRON AND MUON SCATTERING

The study of inelastic electron scattering at high energies is of great historical importance because it led to the first clear evidence for scattering from individual quark constituents confined within the proton. Here we shall consider both electron and muon scattering reactions

$$\ell^- + p \to \ell^- + X \qquad (\ell = e, \mu), \tag{7.48}$$

where X as usual denotes any set of hadrons allowed by the conservation laws.[17] These reactions are dominated by the one-photon exchange mechanism of Figure 7.19, where the energies and momenta of the leptons and proton are also defined. This is a direct generalization of Figure 7.17 for elastic scattering $X \equiv p$; as in our discussion of elastic scattering, we shall consider only these contributions, neglecting completely the small corrections arising from multiphoton exchanges.

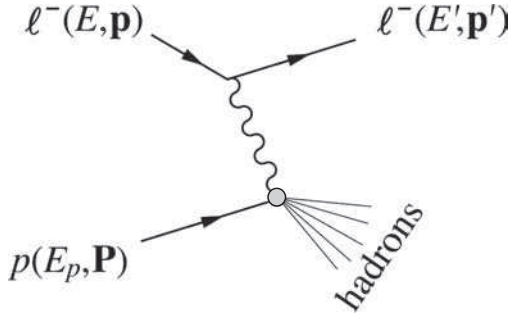


Figure 7.19 Dominant one-photon exchange mechanism for inelastic lepton–proton scattering, where $\ell = e$ or μ.

In Figure 7.19 the exchanged photon acts as a probe of the proton structure, and to resolve individual quarks within it requires large momentum transfers $\mathbf{q} = \mathbf{p} - \mathbf{p}'$, corresponding to photon wavelengths that are very small compared with the size of the proton. In this region, inelastic scattering is dominated by processes like that shown in Figure 7.20, i.e. by two-step processes in which the first step is elastic lepton scattering

$$\ell^- + q \to \ell^- + q \qquad (\ell = e, \mu) \tag{7.49}$$

from one of the quarks contained in the proton, while the second is a complicated 'fragmentation' process, which converts the recoil quark and the remaining constituents of the proton into hadrons. In the experiments we shall discuss, the final-state hadrons are not observed, and we shall not need to consider this fragmentation process further. Instead we shall concentrate on the energy and angular distributions of the recoil leptons, which directly reflect the properties of the quarks from which they scatter.

[17] The corresponding neutrino scattering reactions are also of considerable importance and will be discussed in Section 7.5.

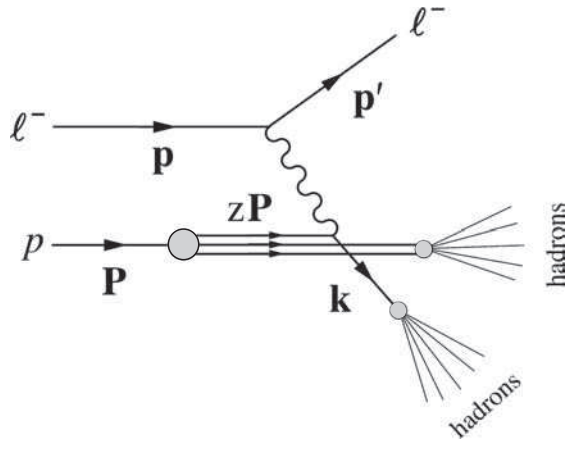


Figure 7.20 Dominant contribution to deep inelastic lepton–proton scattering in the quark model, where $\ell = e$ or μ.

In what follows, we firstly describe the striking features exhibited by the data and then show how these are explained by the mechanism of Figure 7.19. Before doing that, however, it is convenient to introduce two variables that will prove useful in the discussion. These are the variable ν defined by

$$2M\nu \equiv W^2 + Q^2 - M^2 \tag{7.50}$$

and the dimensionless variable

$$x \equiv Q^2 / 2M\nu. \tag{7.51}$$

Here, M is the proton mass, W is the invariant mass of the hadron state X and Q^2 is again the squared energy–momentum transfer (7.32). In the rest frame of the proton, the variable ν reduces to[18]

$$\nu = E - E', \tag{7.52}$$

so that it can be regarded as a Lorentz-invariant generalization for the energy transferred from the lepton to the proton. The physical interpretation of the variable x, which is called the *scaling variable*, will become clear later.

7.4.1 Bjorken scaling

In a typical high-energy lepton scattering experiment, the hadron state is left unidentified and the scattered lepton momentum, and hence its energy, is measured in a spectrometer. The measured quantities are therefore the initial and final lepton energies E and E' and the lepton scattering angle θ. For elastic scattering the invariant

[18] See Problem 7.4.

mass W of the hadrons is just the proton mass, and for a given beam energy E, the final energy E' is determined by the scattering angle θ.[19] However, in inelastic scattering W can vary continuously, depending on the energies and momenta of the produced hadrons, so that both E' and θ are independent variables for a given initial energy E. We therefore have to consider the differential cross-section for scattering to an energy E' and angle θ, and it can be shown that the appropriate generalization of the elastic scattering formula (7.33) is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E'\mathrm{d}\Omega'} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\frac{1}{\nu}\left[\cos^2(\theta/2)\,F_2\left(x,Q^2\right) + \sin^2(\theta/2)\,\frac{Q^2}{xM^2}F_1\left(x,Q^2\right)\right]. \tag{7.53}$$

Here the dimensionless *structure functions* $F_1(x,Q^2)$ and $F_2(x,Q^2)$ parameterize the interaction at the lower vertex in Figure 7.19 in the same way that the dimensionless functions $G_1(Q^2)$ and $G_2(Q^2)$ in (7.33) parameterize the interaction at the lower vertex in elastic scattering of Figure 7.17. We have written them as functions of the scaling variable x and Q^2, rather than the energy transfer ν and Q^2, because for fixed x the measured values of F_1 and F_2 become approximately independent of Q^2 at large Q^2; i.e.

$$F_{1,2}(x,Q^2) \approx F_{1,2}(x) \qquad (Q^2 \gg M^2). \tag{7.54}$$

This is called *Bjorken scaling* or *scale invariance*, because it implies that the structure functions are left unchanged by a scale transformation, i.e. by a transformation in which all particle masses, energies and momenta are multiplied by a scale factor κ, so that $Q^2 \to \kappa^2 Q^2$ but x remains unchanged. Scale invariance is not exact, as we have already said, and small deviations from it are observed, which are called *scaling violations*. They can be clearly seen in Figure 7.21, where we show measured values of F_2 in the range $5 < Q^2 < 200\,\mathrm{GeV}^2$. This behaviour should be compared with that of the analogous function $G_2 \equiv G_M^2$ for elastic scattering, which falls by more than two orders of magnitude over the much smaller range $5 < Q^2 < 30\,\mathrm{GeV}^2$, as can be seen in Figure 7.18.

The approximate scaling behaviour (7.54) was first observed in inelastic scattering experiments at the Stanford Linear Accelerator Centre (SLAC) in 1969, and caused considerable excitement because of the contrast with the strong Q^2 dependence of the elastic form factors. However, if the proton were a point particle, the elastic form factors would also be independent of Q^2, as can be seen from Equations (7.38), and it was soon realized that the observed scaling behaviour of inelastic scattering could be successfully accounted for by scattering from point-like constituents within the proton, rather than from the proton as a whole, as we shall discuss below. This was historically the first dynamical evidence for quarks, whose existence had previously been inferred solely on the basis of static quantities, like the masses and quantum numbers of the hadrons discussed in the previous chapter. Somewhat later, the development of QCD led to the prediction of small deviations from scaling. These scale violations had not been observed at the time because of the limited accuracy

[19] See Problem 7.5.

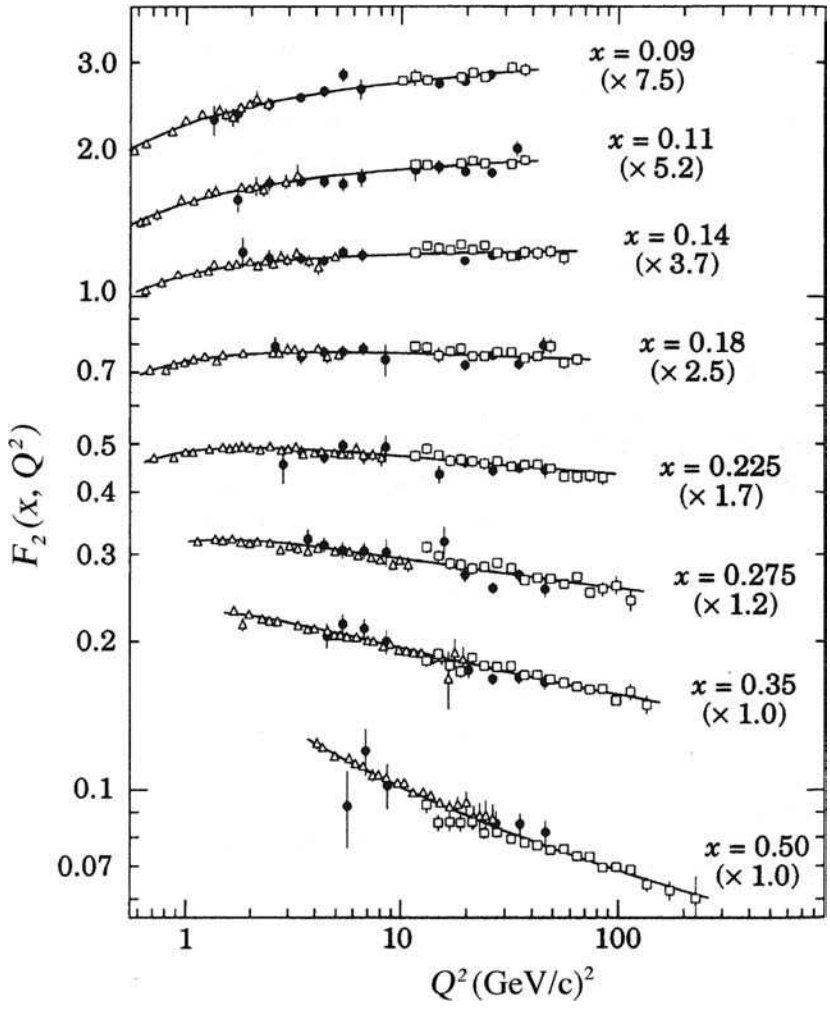


Figure 7.21 Measured values of the structure function $F_2(x, Q^2)$ from a deep inelastic scattering experiment using muons. The data points at the lower x values have been multiplied by the factors in brackets so that they can be displayed on a single diagram. (Reprinted Figure 32 with permission from L. Montanet *et al.*, *Phys. Rev. D*, **50**, 1173. Copyright 1994 American Physical Society.)

and range of the data that were then available, and their subsequent detection with the predicted form was in turn one of the earliest triumphs of QCD.

In what follows we shall concentrate on understanding the phenomenon of approximate scale invariance, deferring the discussion of the small deviations from it, which are predicted by QCD, to starred Section 7.4.3.

7.4.2 The parton model

The approximate scaling behaviour (7.54) was quickly understood, as we have said, in terms of scattering from individual charged point-like constituents within the

proton. These were called *partons* because the existence of quarks, with which they are now identified, was still controversial at the time. Correspondingly, the simplest model that explains the main features of the data is called the *parton model*. In discussing it we shall not initially assume that the partons are quarks, in order to see how far this identification can be justified by experimental data.

The parton model is defined by assuming the mechanism of Figure 7.20 and completely neglecting the corrections to it, which were subsequently predicted by QCD. It takes its simplest form in a reference frame in which the target proton has a very large momentum. In such a frame, the momenta of the constituents will be almost collinear with the proton momentum, so that to a good approximation the target can be viewed as a stream of partons, each carrying a fraction of the proton momentum **P**. Furthermore, for $Q^2 \gg M^2$ it can be shown that the fraction z of the proton momentum carried by the struck parton in Figure 7.20 is given by[20]

$$z = x = Q^2 / 2Mv. \tag{7.55}$$

This is a key result in interpreting the data. It implies that the measured cross-section at a given value of $x = Q^2 / 2Mv$ is proportional to the probability of finding a parton with a fraction $z = x$ of the proton momentum. The full calculation is, however, quite complicated, and we will simply quote below the results obtained for the structure functions in the parton model.[21] We will do this for both spin-0 and spin-$\frac{1}{2}$ partons with arbitrary electric charges in order to see which spin and charge values, if any, are selected by the data.

We start with the structure function F_2, which is given for both spin-0 and spin-$\frac{1}{2}$ partons by

$$F_2(x, Q^2) = \sum_a e_a^2 x f_a(x), \tag{7.56}$$

where $f_a(x)\,\mathrm{d}x$ is the probability of finding a parton a of charge e_a with fractional momentum between x and $x + \mathrm{d}x$, and the sum extends over all parton types a. The parton distributions $f_a(x)$ are not determined by the model, so that the x-dependence of the structure functions must be determined by experiment. However, they are predicted to be the same at all Q^2; i.e. one has exact scaling for point-like partons, irrespective of their spin.

In contrast, the predictions obtained for the structure function F_1 do depend on the spin of the partons. They are

$$F_1(x, Q^2) = 0 \quad (\text{spin} - 0) \tag{7.57a}$$

and

$$2xF_1(x, Q^2) = F_2(x, Q^2) \quad (\text{spin} - \tfrac{1}{2}), \tag{7.57b}$$

[20] This is derived in Problem 7.6.
[21] See, for example, pp. 191–192 of Halzen and Martin (1984).

where the latter result is called the *Callan–Gross relation*. The data rule decisively in favour of the latter relation, as one would expect if the partons were spin-$\frac{1}{2}$ quarks.

Finally, if the partons are quarks, their squared charges e_a^2 must correspond to the squared quark charges $\frac{1}{9}$ and $\frac{4}{9}$. Unfortunately, this cannot be confirmed from measurements of the proton structure functions alone, because of the occurrence of the unknown probability distributions $f_a(x)$ in (7.56). However, by combining measurements of the proton structure functions with those for neutron structure functions, and with similar results obtained in neutrino scattering,[22] it is possible to check that the squared parton charges are consistent with those expected for quarks.

In summary, the scaling behaviour of F_2 is evidence for point-like constituents with the proton, the Callan–Gross relation shows that they are spin-$\frac{1}{2}$ particles and the above checks on their squared charges confirm their identification with quarks beyond all reasonable doubt.

★7.4.3 Parton distributions and scaling violations

The parton model described in Section 7.4.2 was developed before the advent of QCD and completely ignores the quark–gluon interaction. This interaction is, of course, what holds the proton together, and it manifests itself in deep inelastic scattering in two important ways. The first is that the leptons not only scatter from the three so-called *valence* quarks *uud* that give rise to the internal quantum numbers

$$B = Q = 1, \quad S = C = \tilde{B} = T = 0$$

of the proton, but also from a so-called *sea* of quark–antiquark pairs $a\bar{a}$, which are created by the interaction. The second is that the resulting structure functions become weakly dependent on Q^2, as seen experimentally in Figure 7.21. Here we shall discuss how each of these features arises in turn.

In the absence of quark–gluon interactions, the interaction of the exchanged photon with the struck quark in the nucleon may be depicted by Figure 7.22(a). When the quark–gluon interactions are taken into account, other processes besides those of Figure 7.22(a) can contribute, like those shown in Figure 7.22(b) and (c). In Figure 7.22(b) the quark emits a gluon before interacting with the exchanged photon. In Figure 7.22(c) the emitted gluon is converted to a quark–antiquark pair, which forms part of the sea of quark–antiquark pairs mentioned above, and it is one of these particles that absorbs the exchanged photon. Detailed calculations show that the parton model remains a good first approximation when such processes are taken into account, provided that the parton distribution functions $f_a(x)$ are extended to include contributions from antiquarks as well as quarks; i.e. Equation (7.56) becomes

$$F_2\left(x, Q^2\right) \approx \sum_a \left[e_a^2 x f_a(x) + e_a^2 x f_{\bar{a}}(x)\right], \tag{7.58}$$

where the sum extends over all quark flavours $a = u, d, s, \ldots$. The Callan–Gross relation (7.57b) remains unchanged, since both quarks and antiquarks have spin-$\frac{1}{2}$.

[22] This is discussed in starred Section 7.5 below.

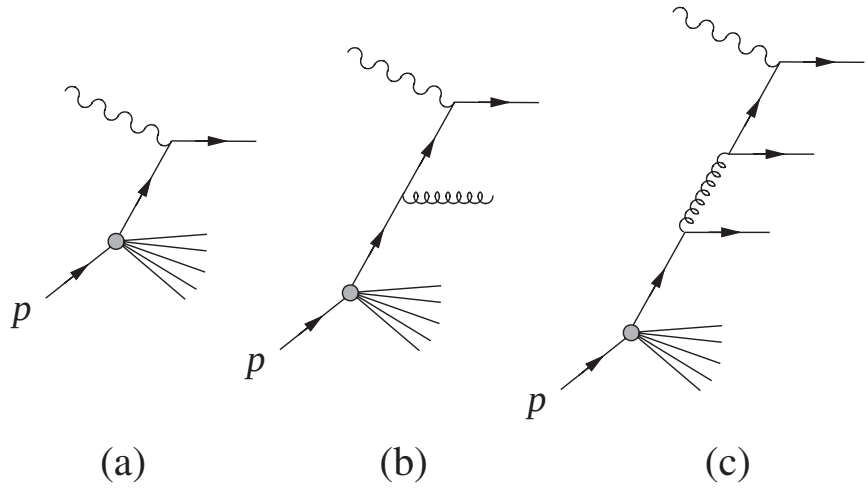


Figure 7.22 (a) The interaction of the exchanged photon with the struck quark in the parton model, together with (b, c) two of the additional processes that occur when quark–gluon interactions are taken into account.

Electron and muon scattering cannot separate the quark and antiquark contributions because the structure functions depend on the squared charges of the constituents, which are the same for quarks and antiquarks. However, neutrino scattering can distinguish between quarks and antiquarks, and is sensitive to the same distribution functions $f_a(x)$ and $f_{\bar{a}}(x)$ as electron and muon scattering.[23] It is important to determine these distribution functions, both in order to understand the nature of the scaling violations discussed below and because the values measured in deep inelastic lepton scattering can be used to predict the rates of many other processes.[24] We will not go into details, but simply show in Figure 7.23 the measured quark and antiquark distributions

$$x\,q(x) \equiv \sum_a x f_a(x), \quad x\,\bar{q}(x) \equiv \sum_a x f_{\bar{a}}(x), \tag{7.59a}$$

where the sum extends over all quark flavours $a = u,\ d,\ s, \ldots$. Also shown is the difference

$$xV(x) \equiv xq(x) - x\bar{q}(x), \tag{7.59b}$$

which corresponds to the valence quark distribution, if we assume that the quarks belonging to the sea of $q\bar{q}$ pairs produced by the gluons have the same momentum distribution as the antiquarks.[25] The important features for what follows are that the

[23] This is discussed briefly in the next section.
[24] Two such processes are discussed in detail in Sections 7.7 and 7.9 of Perkins (1987). Others are $W^{\pm}$ and Z^0 production, which we discussed in Section 4.5.1.
[25] This follows from C invariance if we assume $q\bar{q}$ pairs are produced in the basic process $g \to q\bar{q}$.

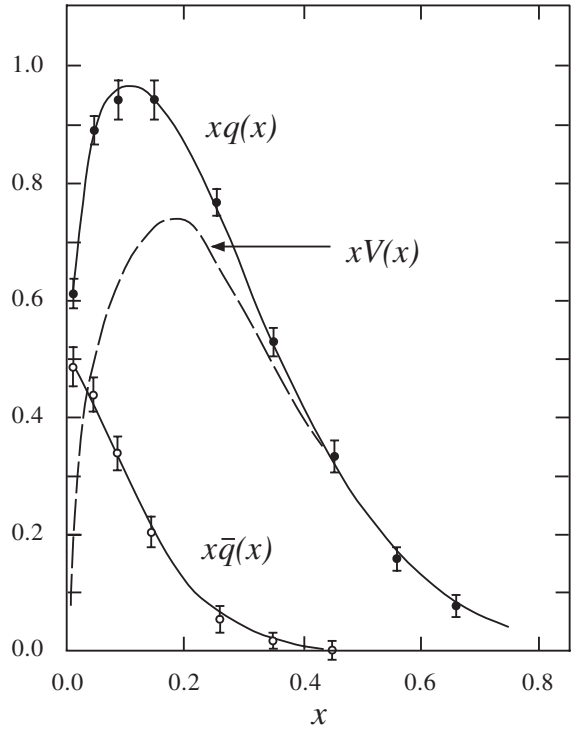


Figure 7.23 Quark and antiquark distributions (7.59a), together with the valence quark distribution (7.59b), measured at a Q^2 value of about $10\,\mathrm{GeV}^2$, from neutrino experiments at CERN and Fermilab.

valence quarks dominate the distributions for $x \geq 0.3$, while the sea of $q\bar{q}$ pairs is dominant at very small x values.

The second important result of the quark–gluon interaction is that the effective parton distributions, which are shown in Figure 7.23 for Q^2 values of about $10\,\mathrm{GeV}^2$, depend, albeit weakly, on Q^2. At larger x values, where the valence quarks dominate, the most important correction to the parton model diagram of Figure 7.22(a) is the process shown in Figure 7.22(b), in which the struck quark emits a gluon before interacting with the exchanged photon. Detailed calculations show that, as Q^2 increases, this process increases in importance[26] and that the emitted gluon carries off some of the initial quark momentum.[27] Thus the effective quark distribution shrinks towards smaller values of the fractional momentum x as Q^2 increases. This is the reason for the decrease in the structure functions with increasing Q^2 seen for fixed $x > 0.2$ in Figure 7.21. In addition, because gluons are emitted more readily as Q^2 increases, the

[26] This is analogous to the situation in electromagnetism. In QED one finds that the more an electron is accelerated in a given process, the more likely it is that the process will be accompanied by photon emission. Similarly in QCD, the more a quark is accelerated in a given process, the more likely it is that the process will be accompanied by gluon emission.

[27] See also Problem 7.7.

sea distribution of quark–antiquark pairs is also found to increase with Q^2, leading to a slow rise in the structure functions with Q^2 at small x values, where the sea is important.

We shall not attempt to discuss these deviations from exact scaling any further, because they are quite complicated. Instead, we shall merely note that QCD correctly predicts their form at all x values and that their magnitude is proportional to the strong coupling constant $\alpha_s(Q^2)$, which is a measure of the probability of gluon emission. The measurement of these deviations from scaling is one of the classic means of determining the strong coupling constant, and consistent results have been obtained from many different experiments in both charged lepton and neutrino scattering.

★7.5 INELASTIC NEUTRINO SCATTERING

We finally turn to inelastic neutrino reactions such as

$$\nu_\mu + p \to \mu^- + X^{++} \tag{7.60a}$$

and

$$\nu_\mu + n \to \mu^- + X^{+}, \tag{7.60b}$$

where X^+ and X^{++} are as usual any hadronic states allowed by the conservation laws. At high energies these reactions are well described by the parton model, and they enable the contributions of different types of partons to be separated, as we noted in the previous section. Firstly, however, we shall use the parton model to predict the energy dependence of the cross-sections for reactions (7.60) at centre-of-mass energies that are much larger than the proton mass. This leads to one of the most striking experimental tests of the parton model, as we shall immediately see.

In the parton model, the reactions (7.60) are assumed to proceed via two-step processes like that shown in Figure 7.24. In the first step, the neutrino scatters from one of the quark constituents in a weak interaction like

$$\nu_\mu + d \to \mu^- + u, \tag{7.61}$$

while the second step is a complicated fragmentation process that converts the recoil quark and the remaining constituents of the nucleons into hadrons. This mechanism is similar to that shown in Figure 7.20 for inelastic electron and muon scattering. As in that case, we shall only discuss experiments in which the hadrons are not identified, and we will not need to consider the fragmentation process further.

In Section 2.2.1 we saw that at low energies the weak interaction could be approximated by a zero-range interaction whose strength is characterized by the Fermi coupling constant G_F given in (2.16). In this approximation, the neutrino reaction (7.61) is represented by the point interaction of Figure 7.25(a), in the same way that inverse muon decay was represented in Figure 2.7. This zero-range approximation has been used for the neutrino scattering process (7.61) in Figure 7.24, and it remains a

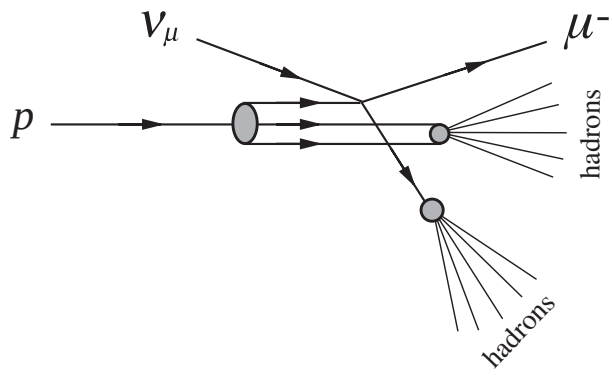


Figure 7.24 Dominant contribution to deep inelastic neutrino scattering in the parton model.

good approximation for the experiments we shall discuss.[28] The energy dependence of the cross-sections for reactions (7.60) then follows from a simple dimensional argument that rests on three points. The first is that the cross-sections are proportional to G_{F}^2, since the amplitudes for weak interactions like (7.61) are proportional to G_{F}. The second is that in the parton model these basic scattering processes only involve point-like particles (i.e. leptons and quarks) with no length scales associated with them. The third point is that, if we work in the centre-of-mass frame, it is reasonable to neglect the proton mass as well as the lepton masses, since the proton will also have a high energy in this frame. Hence the only energy scale available is the centre-of-mass energy E_{CM}, and a simple dimensional argument gives

$$\sigma_\nu = K\, G_F^2\, E_{CM}^2 \tag{7.62}$$

for the total cross-section. Here K is a dimensionless constant and the power of energy is determined by the fact that both σ_ν and G_{F} have natural dimensions[29] $[E^{-2}]$. This result is conventionally rewritten in terms of the neutrino energy E_ν in the laboratory frame, which is given by[30]

$$4\,E_{CM}^2 = 2\,ME_L + M^2 \approx 2\,ME_L$$

at high energies $E_L \gg M$. Hence we finally obtain

$$\sigma \propto E_L \tag{7.63}$$

for the energy dependence of the total cross-section in the parton model, and a similar result holds for antineutrinos.

[28] It is expected to break down for total centre-of-mass energies of order 80 GeV, for reasons which will become clear in the next chapter. Here we restrict ourselves to centre-of-mass energies $E_{\mathrm{CM}} < 25$ GeV, which are smaller than this, but still large compared with the proton mass.

[29] See Table 1.1.

[30] This follows directly from Equation (A.12) of Appendix A.

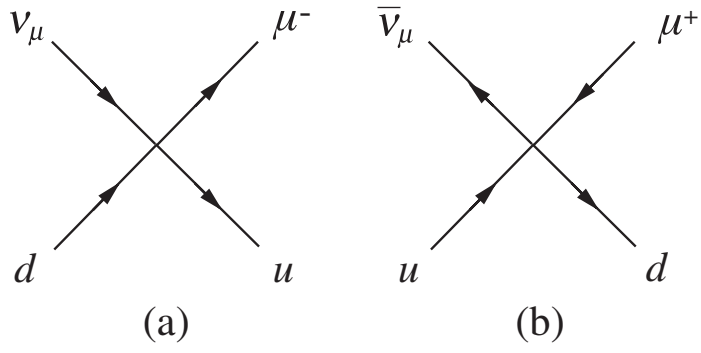


Figure 7.25 Allowed neutrino quark scattering processes (7.61) and (7.68).

In order to verify the predicted behaviour (7.63), accurate measurements of the neutrino cross-sections are required. In spite of their small size, this has been done by exploiting the intense neutrino beams available at a number of accelerators[31] together with very large target/detector systems. An example of such a detector (WA1 at CERN) is shown in Figure 7.26.

This 20 m long, 1400 ton device was a modular assembly consisting of a succession of magnetized iron toroids interspersed with drift chambers and scintillation counters. The iron toroids served as the target and also produced the magnetic field

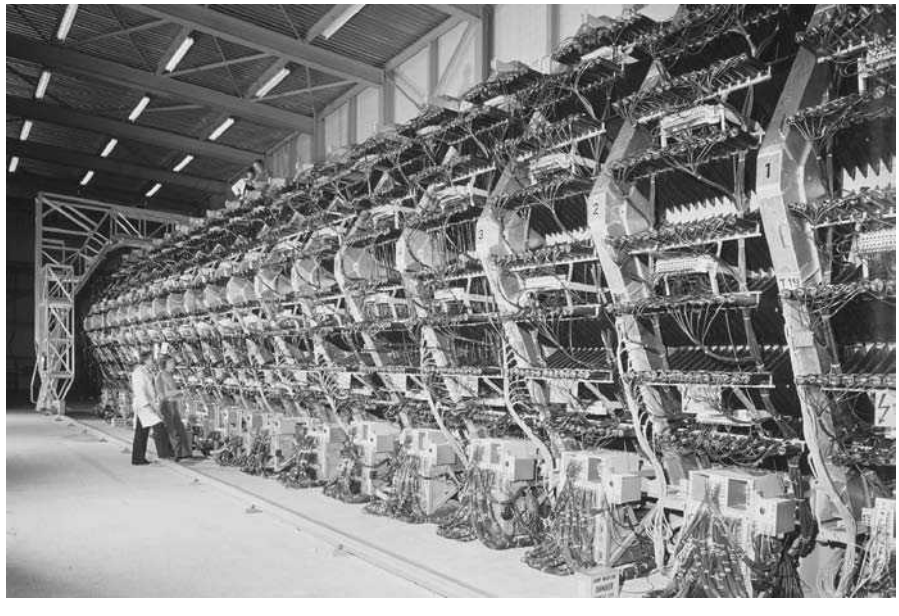


Figure 7.26 The electronic detector used in the WA1 neutrino experiment at CERN. (Reproduced by permission of CERN.)

[31] The production of such beams was discussed briefly in Section 4.2.4.

in the detector. The muons produced in the reactions (7.60) penetrated many modules of the detector and their momenta were determined from the curvatures of their trajectories in the magnetic field. Finally, the whole detector acted as a hadron calorimeter, enabling the total energy of the unidentified hadron state X to be measured. Because the targets were iron nuclei, the experiment essentially measured the sum of the cross-sections for the reactions (7.60a) and (7.60b) for incident neutrinos and of the cross-sections for the corresponding reactions

$$\bar{\nu}_\mu + p \to \mu^+ + X^0 \tag{7.64a}$$

and

$$\bar{\nu}_\mu + n \to \mu^+ + X^- \tag{7.64b}$$

for incident antineutrinos. The results for the total cross-sections are shown in Figure 7.27. As can be seen, they are in excellent agreement with the prediction of the parton model (7.63).

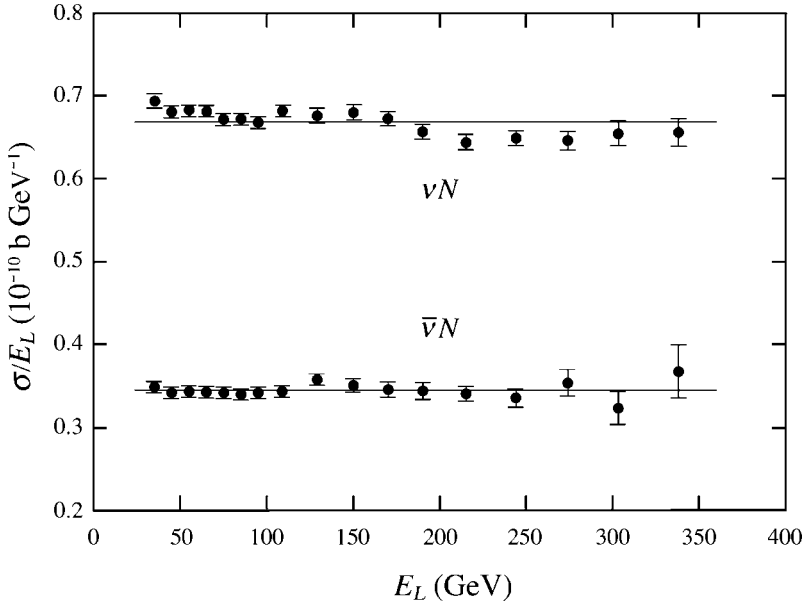


Figure 7.27 Neutrino and antineutrino total cross-sections. (Data from Seligman *et al.*, 1997.)

★7.5.1 Quark identification and quark charges

The above measurements strikingly confirm the idea of scattering from point-like constituents within the nucleons. However, to extract the different quark and antiquark distributions from neutrino data, it is necessary to measure differential cross-sections analogous to those given in (7.53) for charged lepton scattering. The corresponding

formulas for neutrino and antineutrino scattering are somewhat similar to (7.53), but are more complicated because in addition to structure functions $F_1^{\nu N}$, $F_2^{\bar{\nu}N}$ analogous to F_1 and F_2 in (7.53), there is a third structure function $F_3^{\nu N}$ that occurs because of the existence of parity violation in weak interactions. We will not discuss this here, but simply cite without proof the results obtained for the neutrino and antineutrino structure functions analogous to F_2. In addition, we shall confine ourselves for simplicity to the region $x \geq 0.3$, where the sea of quark–antiquark pairs can to a good approximation be neglected. In this region, the only nonvanishing quark distributions are

$$u(x) \equiv f_u(x) \quad \text{and} \quad d(x) \equiv f_d(x) \tag{7.65}$$

arising from the valence quarks *uud* within the proton, and the proton structure function (7.58) measured in electron or muon scattering reduces to

$$F_2^p(x) = e_u^2\, x\, u(x) + e_d^2\, x\, d(x) \quad (x \geq 0.3). \tag{7.66}$$

The corresponding results for neutrino and antineutrino scattering can be shown to be

$$F_2^{\nu p}(x) = 2x\, d(x) \quad \text{and} \quad F_2^{\bar{\nu}p}(x) = 2x\, u(x) \quad (x \geq 0.3), \tag{7.67}$$

and can be understood, apart from their normalization, from the conservation of electric charge, baryon number and muon number. The reaction (7.61) shown in Figure 7.25(a), and the corresponding reaction

$$\bar{\nu}_\mu + u \to \mu^+ + d \tag{7.68}$$

shown in Figure 7.25(b), satisfy all these conservation laws. Other hypothetical reactions like

$$\nu_\mu + u \to \mu^+ + d \quad \text{and} \quad \nu_\mu + u \to \mu^+ + d$$

that satisfy charge and baryon number conservation, violate the conservation of muon number (2.3b), and there are no $\nu_\mu u$ or $\bar{\nu}_\mu d$ scattering reactions that satisfy all three conservation laws. Hence neutrinos can scatter from d quarks but not from u quarks, while antineutrinos can scatter from u quarks but not from d quarks, as implied by Equations (7.67). It is for this reason that neutrino and/or antineutrino beams are an ideal tool for separating u and d quark contributions, among others.

When the contributions of the quark–antiquark sea at small x are included, the separation of the various distributions becomes more complicated, but we shall not pursue this further.[32] Instead, we will continue to work at $x > 0.3$ and show how the comparison of neutrino with electron or muon scattering on nuclear targets leads to a determination of the sum of quark charges $(e_u^2 + e_d^2)$. To do this, we first note that, according to isospin symmetry, a neutron $n(940) = ddu$ differs from a proton

[32] See, for example, Section 9.3 of Halzen and Martin (1984).

$p(938) = uud$ in that u and d quarks are interchanged. Hence the quark distributions $u^n(x)$ and $d^n(x)$ of the neutron are related to those of the proton by

$$u^n(x) = d^p(x) \equiv d(x) \quad \text{and} \quad d^n(x) = u^p(x) \equiv u(x)$$

to a very good approximation. The neutron structure function F_2^n measured in electron or muon scattering is then given by

$$F_2^n(x) = e_u^2\, x\, u^n(x) + e_d^2\, x\, d^n(x) = e_u^2\, x\, d(x) + e_d^2\, x\, u(x) \quad (x > 0.3) \tag{7.69}$$

in analogy to (7.66), and the corresponding neutrino structure functions are given by

$$F_2^{\nu n}(x) = 2x d^n(x) = 2x\, u(x) \quad \text{and} \quad F_2^{\bar{\nu} n}(x) = 2x u^n(x) = 2x\, d(x) \tag{7.70}$$

in analogy to (7.67). From these results, one can easily show that

$$\frac{F_2^p(x) + F_2^n(x)}{F_2^{\nu p}(x) + F_2^{\nu n}(x)} = \frac{e_u^2 + e_d^2}{2} = \frac{5}{18}, \tag{7.71}$$

where we have substituted the values of the quark charges. This prediction can be tested by taking data on nuclei that contain equal numbers of protons and neutrons. It is verified within experimental errors of order 10 % and was historically important in confirming that the charges of the 'partons' found in deep inelastic scattering were indeed the same as the charges of the quarks with which they are now identified.

PROBLEMS 7

7.1 Estimate the cross-section ratio R, defined by (7.18), at total centre-of-mass energies $E_{CM} = 2.8$, 5 and 15 GeV, assuming that there are no resonance peaks at these energies and that (7.22) remains approximately valid even at energies as low as 2.8 GeV. How would you expect R to change when the energy becomes large enough to produce $t\bar{t}$ pairs, where t is the top quark?

7.2 Perform the angular integration in (7.27) and use the result (a) to show that $G_E(q^2)$ is indeed a function of q^2 only and (b) to derive (7.31) for the root-mean-square charge radius r_E.

7.3 Verify (7.43) by explicitly evaluating the integral.

7.4 Show that the scattering variable ν defined by (7.50) reduces to (7.52) in the rest frame of the proton. Hence show that the scaling variable (7.51) lies in the range $0 \leq x \leq 1$ if the lepton masses are neglected.

7.5 Elastic lepton scattering (7.23) can be regarded as special case of inelastic scattering (7.48) in which the final hadronic state X is a single proton. Find the corresponding value of the scaling variable (7.51). Hence show that in the rest frame of the proton, the initial (E) and final (E') lepton energies are related by

$$M(E - E') = EE'(1 - \cos\theta),$$

where θ is the lepton scattering angle, M is the proton mass and lepton masses are assumed to be negligible compared with their energies.

7.6 According to (7.55), the fraction z of the proton momentum carried by the struck parton in Figure 7.20 is equal to the scaling variable x if certain approximations are made. These are: the proton momentum $\mathbf{P}$ is much larger than its mass, i.e. $|\mathbf{P}| \gg M$; the masses and

transverse momenta of the partons are negligible compared with $|\mathbf{P}|$ and Q^2; and $Q^2 \gg M^2$. Verify this starting from the relation

$$\omega^2 - \mathbf{k}^2 = \tilde{m}^2 \approx 0,$$

where $\mathbf{k}$ is the final momentum of the struck parton, as shown in Figure 7.20, and ω and $\tilde{m}$ are its energy and mass, respectively. (*This problem should not be attempted before Problem 7.4.*)

7.7 Use the quark distributions of Figure 7.23 to make a rough estimate of the fraction of the proton momentum in a high-momentum frame that is carried by its quark and antiquark constituents. To what would you ascribe the 'missing momentum'?

8

Weak Interactions: Quarks and Leptons

We turn now to the third fundamental force of nature, the weak interaction. Like the strong and electromagnetic interactions, the weak interaction is also associated with elementary spin-1 bosons that act as 'force carriers' between quarks and/or leptons. However, in contrast to photons and gluons, these are very massive particles and the resulting interactions are consequently of very short range. There are three such 'intermediate vector bosons': the charged bosons W^+ and W^- and the neutral Z^0. All were discovered at CERN in 1983.[1] Their masses are measured to be

$$\mathrm{M}_W = 80.40\,\mathrm{GeV/c^2}, \quad M_Z = 91.19\,\mathrm{GeV/c^2}, \tag{8.1}$$

which from (1.26) give ranges

$$R_W \approx R_Z \approx 2 \times 10^{-3}\,\mathrm{fm} \tag{8.2}$$

for the weak interaction. These are very small distances, even when compared with the size of the nucleon, and at low energies the weak interaction can be treated as a zero-range interaction as discussed in Section 2.2.1 (cf. Figures 2.7 and 2.8). In this and the following chapter, we shall extend the boson-exchange theory to include quarks and also consider high-energy weak interactions, where the zero-range approximation is no longer appropriate. Firstly, however, it will be useful to make a few remarks on the historical development of the modern theory of weak interactions.

[1] One of the experiments (UA1) is described in Section 4.5.1.

Particle Physics Third Edition B. R. Martin and G. Shaw

The idea that weak interactions are due to the exchange of massive charged bosons seems to have been first proposed by Klein in 1938.[2] These particles were later called $W^{\pm}$ bosons, and until 1973 all observed weak interactions were consistent with the hypothesis that they were mediated by the exchange of heavy charged bosons $W^{\pm}$ only. However, in the 1960s Glashow, Salam and Weinberg had developed a theory that unified electromagnetic and weak interactions in a way that is often compared to the unification of electric and magnetic interactions by Faraday and Maxwell a century earlier. This new theory made several remarkable predictions, including the existence of a neutral vector boson Z^0 and of weak reactions arising from its exchange. These processes are called *neutral current* reactions to distinguish them from the so-called *charged current* reactions arising from charged $W^{\pm}$ boson exchange. In particular, neutral current reactions of the type

$$\nu_\mu + N \to \nu_\mu + X \tag{8.3}$$

were predicted to occur via the mechanism of Figure 8.1, where N is a nucleon and the hadrons are any allowed by the conservation laws. Although difficult to detect, such reactions were first observed in an experiment at CERN in 1973 using a heavy-liquid bubble chamber, and one of the events obtained is shown in Figure 8.2. In this event, an incoming muon neutrino interacts with a nucleon inside the chamber, producing several hadrons, but no muon. The presence of a neutrino in the final state is inferred from the law of conservation of muon number (2.3b) discussed in Chapter 2.

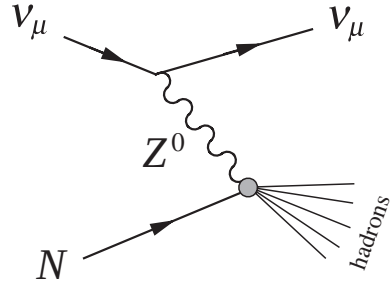


Figure 8.1 Mechanism for neutral current reactions of the type (8.3).

The prediction of the existence and properties of neutral currents, prior to their discovery, is only one of many spectacular successes of the unified theory of electromagnetic and weak interactions. Others include the prediction of the existence of the charmed quark, prior to its discovery in 1974, and the prediction of the masses of the $W^{\pm}$ and Z^0 bosons, prior to the long-awaited detection of these particles in 1983. In general, the theory is in complete agreement with all the data on both weak and electromagnetic interactions, which are now usually referred to collectively as the *electroweak interaction*, in the same way that electric and magnetic interactions

[2] Klein (1938), and see also Klein (1948).

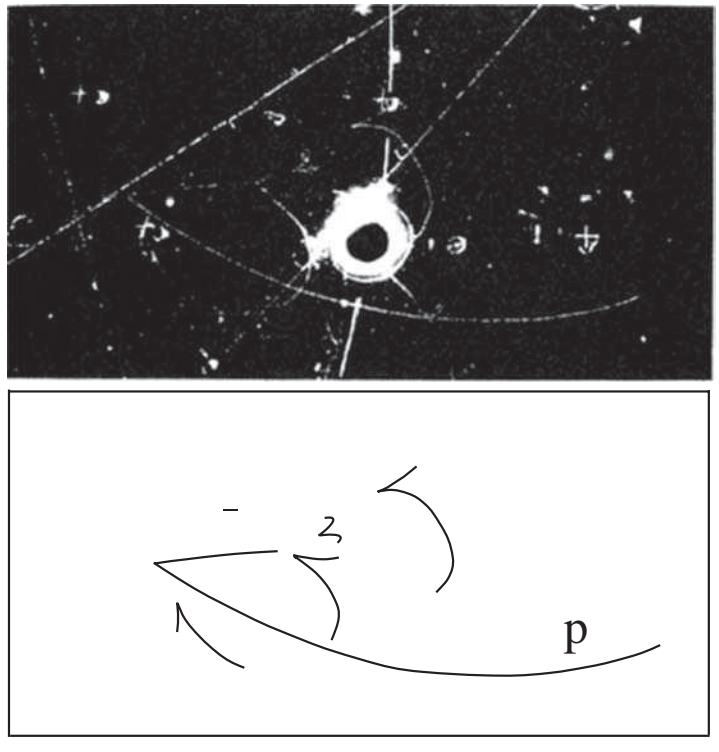


Figure 8.2 Neutral current reactions of the type (8.3) observed in a heavy-liquid bubble chamber at CERN in 1973. The neutrino is incident from the left and interacts with an atomic nucleus inside the chamber to produce a negative pion and a proton. Four photons are also detected via their conversion to e^+e^- pairs. These photons almost certainly arise from the decays of neutral pions produced in the initial neutrino interaction and in the subsequent interaction of the negative pion with another atomic nucleus. (Reproduced by permission of CERN.)

are referred to collectively as electromagnetic interactions. However, the new unification only becomes manifest at very high energies, and at lower energies weak and electromagnetic interactions can still be clearly separated, as we shall see.

In this chapter we will discuss only the charged-current reactions, i.e. those that involve the emission, absorption or exchange of $W^{\pm}$ bosons, leaving neutral current reactions, which involve Z^0 bosons, and the unified theory to Chapter 9.

8.1 CHARGED CURRENT REACTIONS

The simplest charged current reactions are purely leptonic processes like muon decay

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \tag{8.4}$$

and we discuss these first. We then discuss purely hadronic processes, like lambda decay

$$\Lambda \rightarrow p + \pi^- \tag{8.5}$$

and *semileptonic* reactions, such as neutron decay

$$n \to p + e^- + \bar{\nu}_e, \tag{8.6}$$

which involve both hadrons and leptons. As usual, the interactions of the hadrons are interpreted in terms of their quark constituents, and we will see that there is a remarkable symmetry between the weak interactions of leptons and quarks.

8.1.1 $W^{\pm}$–lepton interactions

In Chapter 1 we saw that all the electromagnetic interactions of electrons and positrons could be built from eight basic interactions in which a photon is either emitted or absorbed. These processes were derived from hole theory in Section 1.3.1, and they are conveniently summarized by the vertex of Figure 8.3(a), from which they can be obtained by associating different particle lines with the initial and final states. This is illustrated in Figures 8.3(b) and (c), while the full set of eight processes corresponding to the vertex is shown in Figure 1.4. Leptonic weak interaction processes like muon decay (8.4) and inverse muon decay

$$\nu_\mu + e^- \to \mu^- + \nu_e \tag{8.7}$$

can also be built from a limited number of basic vertices in the same way. In the case of the charged current reactions, it can be shown that for each lepton type $\ell = e, \mu, \tau$ there are 16 such basic reactions corresponding to the two vertices of Figure 8.4. The eight processes corresponding to Figure 8.4(a) are shown explicitly in Figure 8.5. The eight corresponding to Figure 8.4(b) are not shown explicitly since they can be obtained from those of Figure 8.5 by replacing all the particles by their antiparticles, as illustrated in Figure 8.6.

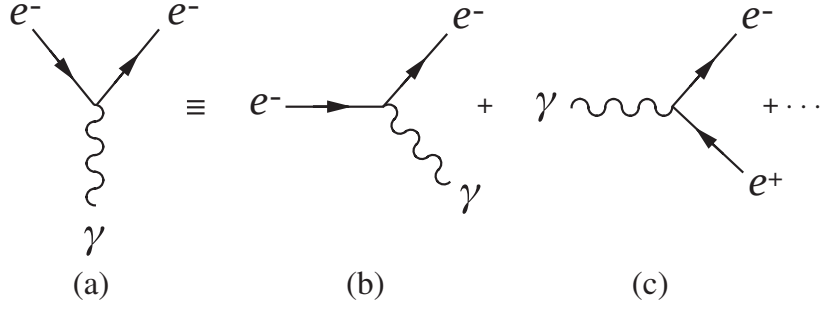


Figure 8.3 The basic vertex for electron–photon interactions, together with two of the eight basic processes derived from it.

One of the most important properties of the weak interaction is that it conserves the lepton numbers L_e, L_μ and L_τ defined in Equations (2.3a), (2.3b) and (2.3c). This is guaranteed by the fact that in each of the vertices of Figure 8.4 there is one arrow pointing into the vertex and one pointing out of it, and the lepton label ℓ is the same on

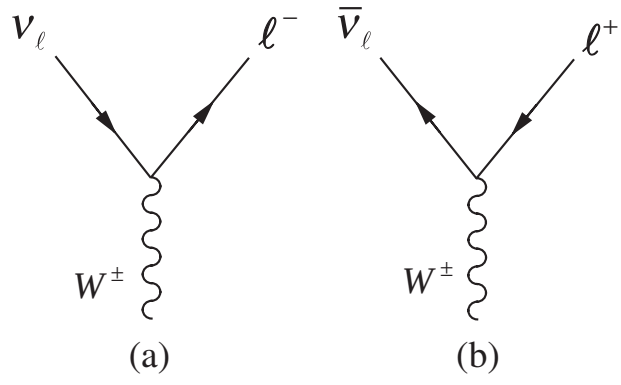


Figure 8.4 The two basic vertices for $W^\pm$–lepton interactions.

both lines. In contrast, vertices like those in Figure 8.7(a), which give rise to processes that violate lepton number, like those shown in Figures 8.7(b) and (c), are excluded from the scheme.[3]

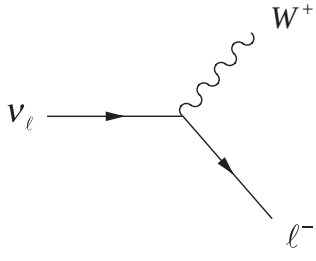
Although the processes of Figure 8.5 conserve lepton numbers, this does not mean that they can occur as isolated reactions in free space. As we saw in Section 1.3.2, the basic electromagnetic processes of Figure 1.4 violate energy conservation if momentum conservation is assumed, and they can only contribute to physical processes if two or more are combined in such a way that energy is conserved overall. The same applies to most of the processes of Figure 8.5, but in this case there are some important exceptions. Specifically, by working in the rest frame of the W boson one can show that Figures 8.5(e) and 8.5(f) are compatible with both energy and momentum conservation provided that

$$M_W > M_\ell + M_{\nu_\ell} \quad (\ell = e, \mu, \tau),$$

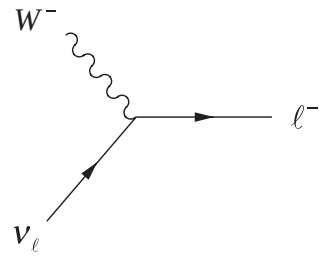
which is obviously the case. In particular, Figure 8.5(f) and its 'antiparticle' reaction, which are shown together in Figure 8.6, are the dominant mechanisms for the leptonic decays (4.28a) and (4.28b), which were used to detect the $W^\pm$ bosons in the experiments discussed in Section 4.5.1.

We turn now to the strength of the reaction. In Section 2.2.2 we showed that the weak and electromagnetic interactions of electrons, muons and tauons are identical, provided their mass differences are taken into account. Correspondingly, each of the vertices of Figure 8.4 is characterized by the same dimensionless strength parameter α_W, which is analogous to the fine structure constant α in QED, independent of which lepton type is involved. We can obtain an order-of-magnitude estimate for α_W by applying the method of dimensions to the rate for the leptonic decays shown in

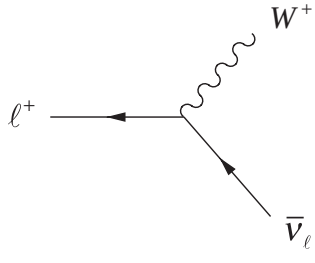
[3] In fact the diagrams of Figure 8.4 are the only possible $W^\pm$–lepton vertices that satisfy both charge and lepton number conservation (see Problem 8.1).



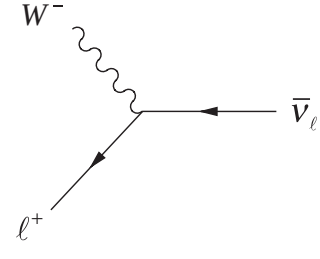
(a) $\nu_\ell \to \ell^- + W^+$



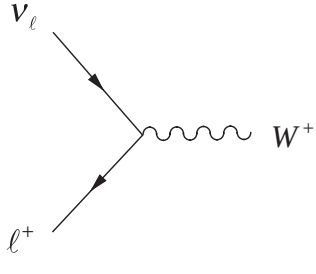
(b) $W^- + \nu_\ell \to \ell^-$



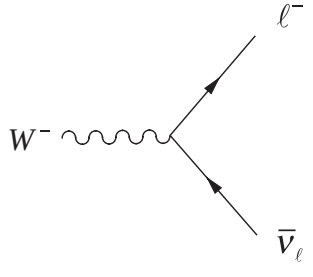
(c) $\ell^+ \to \overline{\nu}_\ell + W^+$



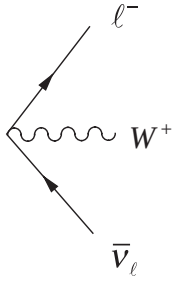
(d) $W^- + \ell^+ \to \overline{\nu}_\ell$



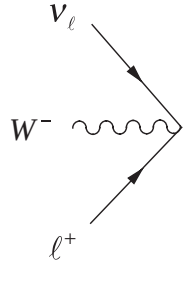
(e) $\nu_\ell + \ell^+ \to W^+$



(f) $W^- \to \ell^- + \overline{\nu}_\ell$



(g) vacuum $\to \ell^- + \overline{\nu}_\ell + W^+$



(h) $W^- + \ell^+ + \nu_\ell \to$ vacuum

Figure 8.5 The eight basic processes derived from the vertex of Figure 8.4(a).

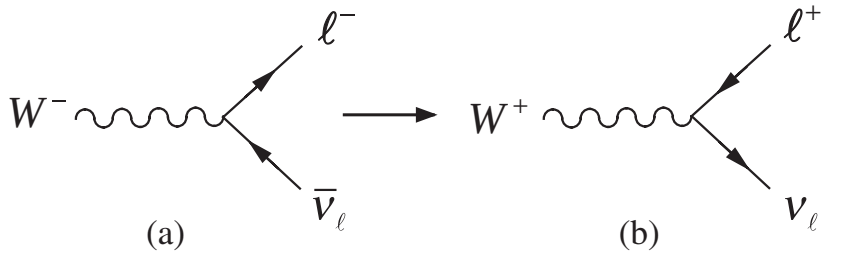


Figure 8.6 A pair of weak interaction processes, which are related by replacing all particles by their antiparticles. The two diagrams (a) and (b) arise from the vertices of Figures 8.4(a) and (b), respectively.

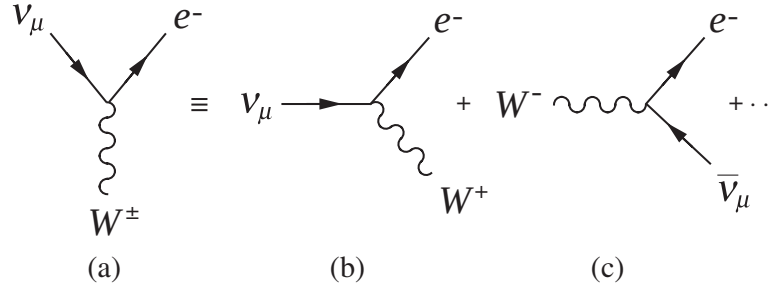


Figure 8.7 Example of a vertex that violates lepton number conservation, together with two of the forbidden processes to which it would give rise.

Figures 8.5(a) and (b) for $\ell = e$. This decay rate has a natural dimension [E], since in natural units time has a dimension $[E^{-1}]$[4] and is measured to be

$$\Gamma(W \to e\nu) \approx 0.230 \pm 0.008\,\text{GeV}. \tag{8.8}$$

Since it involves only a single basic process, Figure 8.6(a) or (b), this decay rate is proportional to α_W. If we neglect the lepton masses, the only other parameter is the W^- boson mass. Hence a simple dimensional estimate gives

$$\Gamma(W \to e\nu) \approx \alpha_W M_W \approx 80\alpha_W\,\text{GeV},$$

which combined with (8.8) gives

$$\alpha_W \approx 1/350 = O(\alpha).$$

[4] See Table 1.1.

Alternatively, a more detailed calculation, which is beyond the scope of this book,[5] gives

$$\Gamma(W \to ev) = 2\alpha_W M_W/3,$$

resulting in the value

$$\alpha_W = 0.0043 \pm 0.0002, \tag{8.9}$$

in good agreement with the value (2.19) obtained from the precise values of the Fermi coupling constant and the W-mass using (2.17). As already noted in Section 2.2.1, the weak and electromagnetic interactions are actually of comparable intrinsic strength, and weak interaction rates are only small at low energies because of the large mass of the W boson, which enters the low-energy effective coupling constant (2.17) as the inverse square.

8.1.2 Lepton–quark symmetry

The weak interactions of hadrons are understood in terms of basic processes in which $W^{\pm}$ bosons are emitted or absorbed by their constituent quarks. These can give rise to semileptonic processes like neutron decay (8.6), for which the dominant diagram is shown in Figure 8.8, and purely hadronic decays such as Λ decay (8.5), for which the dominant diagrams are shown in Figure 8.9. The purely hadronic decays are not nearly as well understood as the semileptonic ones, because the final state particles interact strongly with each other, leading to effects that are very difficult to calculate. We shall therefore concentrate mainly on the semileptonic interactions in what follows, and we shall initially restrict ourselves to the first two generations of quarks

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c \\ s \end{pmatrix}, \tag{8.10}$$

of which the overwhelming majority of known hadrons are composed.

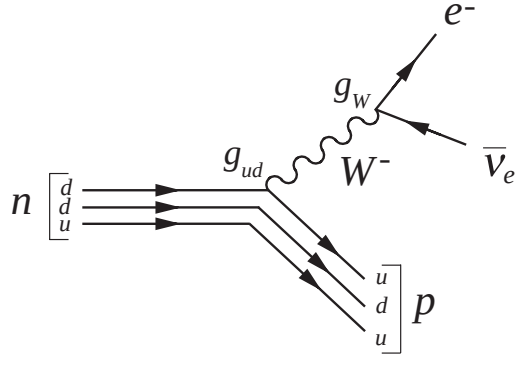


Figure 8.8 Dominant diagram for neutron decay (8.6).

[5] See, for example, Section 11.6.3 of Mandl and Shaw (1993).

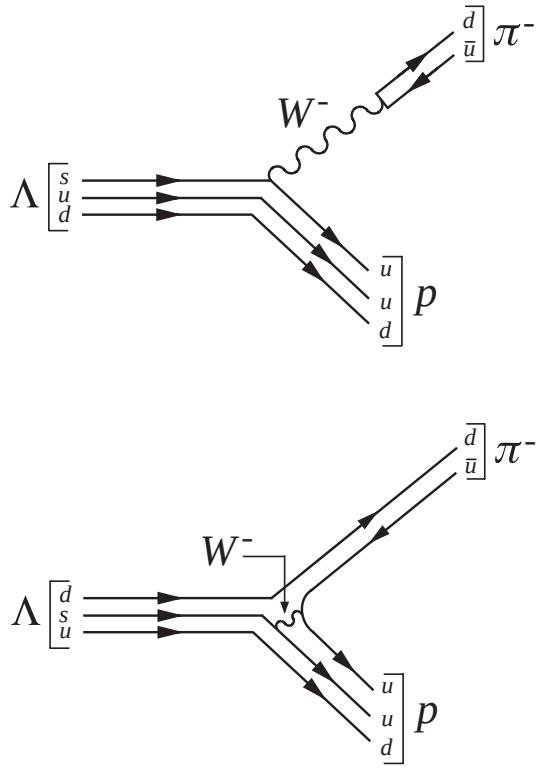


Figure 8.9 Dominant diagrams for Λ decay (8.5).

The weak interactions of quarks are best understood in terms of two ideas: *lepton–quark symmetry* and *quark mixing*. Here we shall introduce these ideas and use them to deduce the form of the $W^{\pm}$ boson–quark interactions. They may also be used to deduce the form of the Z^0 boson–quark interaction,[6] and their justification lies in the fact that the interactions so deduced successfully explain the observed weak interactions of hadrons in both cases.

We start with lepton–quark symmetry. In its simplest form this asserts that the two generations of quarks (8.10) and the two generations of leptons

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \tag{8.11}$$

have identical weak interactions; i.e. one obtains the basic $W^{\pm}$–quark vertices by making the replacements $\nu_e \to u$, $e^- \to d$, $\nu_\mu \to c$, $\mu^- \to s$ in the basic $W^{\pm}$–lepton vertices in Figure 8.4, leaving the coupling constant g_{w} unchanged. In this way one obtains the vertices of Figure 8.10, where the coupling constants

$$g_{ud} = g_{cs} = g_W. \tag{8.12}$$

[6] This is done in Section 9.1.1.

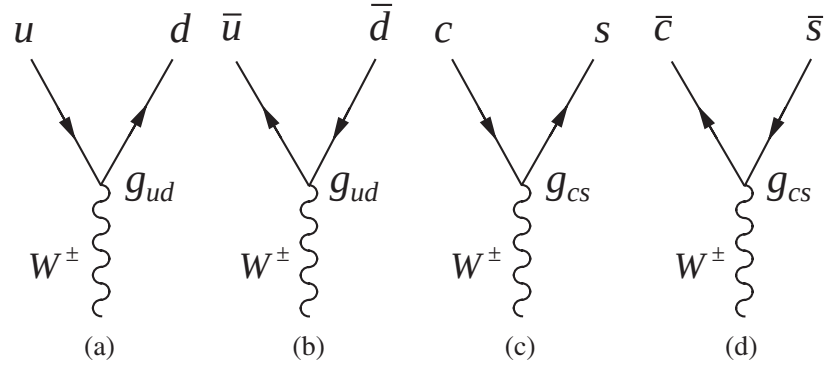


Figure 8.10 The $W^{\pm}$ quark vertices obtained from lepton–quark symmetry when quark mixing is ignored.

This works quite well for many reactions, like the pion decay

$$\pi^- \to \mu^- + \bar{\nu}_\mu, \tag{8.13a}$$

which corresponds to

$$d\bar{u} \to \mu^- + \bar{\nu}_\mu \tag{8.13b}$$

at the quark level, and is interpreted by the Feynman diagram of Figure 8.11(a).

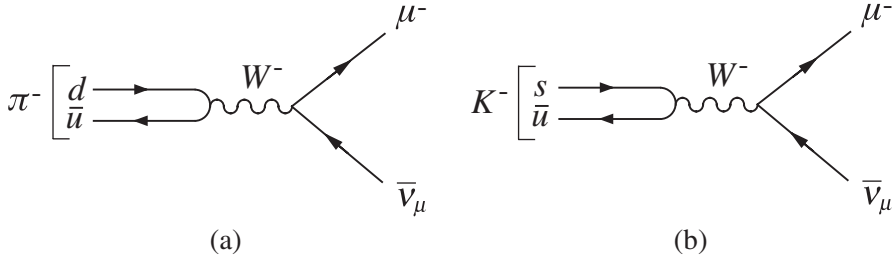


Figure 8.11 Feynman diagrams for the semileptonic decays: (a) $\pi^- \to \mu^- + \bar{\nu}_\mu$ and (b) $K^- \to \mu^- + \bar{\nu}_\mu$.

However, many other decays which are experimentally observed are forbidden in this simple scheme. An example is the kaon decay

$$K^- \to \mu^- + \bar{\nu}_\mu, \tag{8.14a}$$

which corresponds to

$$s\bar{u} \to \mu^- + \bar{\nu}_\mu \tag{8.14b}$$

at the quark level. This decay is naturally explained by the Feynman diagram of Figure 8.11(b), which includes a usW vertex, which is not included in the vertices of Figure 8.10. It can be incorporated, however, by introducing an hypothesis due to Cabibbo called quark mixing, mentioned above. According to this idea, the d and s quarks participate in the weak interactions via the linear combinations

$$d' = d\cos\theta_C + s\sin\theta_C \tag{8.15a}$$

and

$$s' = -d\sin\theta_C + s\cos\theta_C, \tag{8.15b}$$

where the parameter θ_C is called the *Cabibbo angle*. Therefore lepton–quark symmetry is assumed to apply to the doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c \\ s' \end{pmatrix}. \tag{8.16}$$

This is illustrated, for the first of these, in Figure 8.12. As can be seen, the coupling of the previously allowed udW vertex is suppressed by a factor $\cos\theta_C$ relative to (8.12), while the previously forbidden usW vertex is now allowed with a coupling $g_W \sin\theta_C$. The same argument applies to the other three vertices in Figure 8.10, so that in addition to the four vertices of Figure 8.10 with the couplings

$$g_{ud} = g_{cs} = g_W \cos\theta_C \tag{8.17a}$$

one has the vertices of Figure 8.13 with the couplings

$$g_{us} = -g_{cd} = g_W \sin\theta_C. \tag{8.17b}$$

The set of $W^{\pm}$–quark couplings of (8.17a) and (8.17b) successfully accounts for the charged current interactions of hadrons. Since the lepton coupling g_W is already

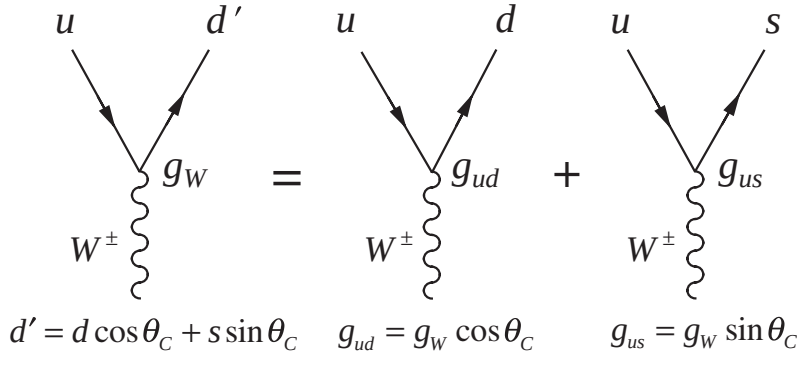


Figure 8.12 The $ud'W$ vertex and its interpretation in terms of udW and usW vertices.

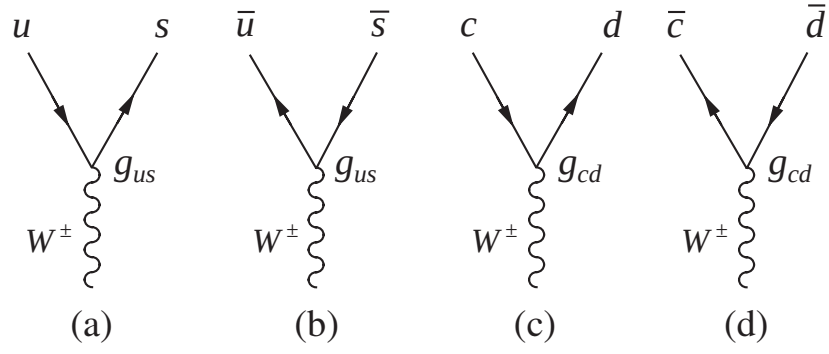


Figure 8.13 The additional vertices arising from lepton–quark symmetry when quark mixing is taken into account.

fixed, these couplings are specified in terms of a single free parameter, the Cabibbo angle θ_C. This is determined by deducing the values of the couplings g_{ud} and g_{us} from the measured rates of various hadron decays. One way of doing this is to compare the rates of decays like those shown in Figure 8.11. As can be seen, they differ in that a d quark has been replaced by an s quark in the initial state and a coupling g_{ud} has been replaced by a coupling g_{us}, so that the ratio of their rates

$$\frac{\Gamma(K^- \to \mu^- \bar{\nu}_\mu)}{\Gamma(\pi^- \to \mu^- \bar{\nu}_\mu)} \propto \frac{g_{us}^2}{g_{ud}^2} = \tan^2 \theta_C,$$

by (8.17). Of course, the difference in the d and s quark masses will also have an effect on the rates, and this must be taken into account. We omit the details and merely quote the mean value obtained from this and other similar determinations, which is

$$g_{us}/g_{ud} = \tan \theta_C = 0.232 \pm 0.002, \tag{8.18}$$

corresponding to a Cabibbo angle of

$$\theta_C = 13.1 \pm 0.1 \text{ degrees}. \tag{8.19}$$

A similar value is obtained by comparing the rates for neutron and muon decay, which depends on the ratio

$$(g_{ud}/g_W)^2 = \cos^2 \theta_C,$$

as can be seen by comparing the quark diagrams of Figures 8.8 and 2.5.

It remains to consider the charmed quark couplings g_{cd} and g_{cs}. These are measured most accurately in neutrino scattering experiments and yield a value consistent with (8.19), but with a larger error. However, the most striking result is for charmed particle decays, which almost always yield a strange particle in the final state. In order to

understand this we first note that decays that involve the couplings (8.17b) are called *Cabibbo-suppressed* because their rates are typically reduced by a factor of order

$$g_{us}^2/g_{ud}^2 = g_{cd}^2/g_{cs}^2 = \tan^2\theta_C = 1/20 \tag{8.20}$$

compared with similar *Cabibbo-allowed* decays, which involve the couplings (8.17a). The Cabibbo-allowed decays

$$c \to s + \ell^+ + \nu_\ell (\ell = e, \mu) \tag{8.21a}$$

and

$$c \to s + u + \bar{d} \tag{8.21b}$$

of a charmed quark to lighter quarks and leptons are shown in Figure 8.14 and they necessarily produce an s quark in the final state, as can be seen. Of course, both these quarks must be bound into hadrons, and the implication of this result is that the decay of a charmed hadron will almost always lead to a strange hadron in the final state. Experimentally, this is indeed the case.

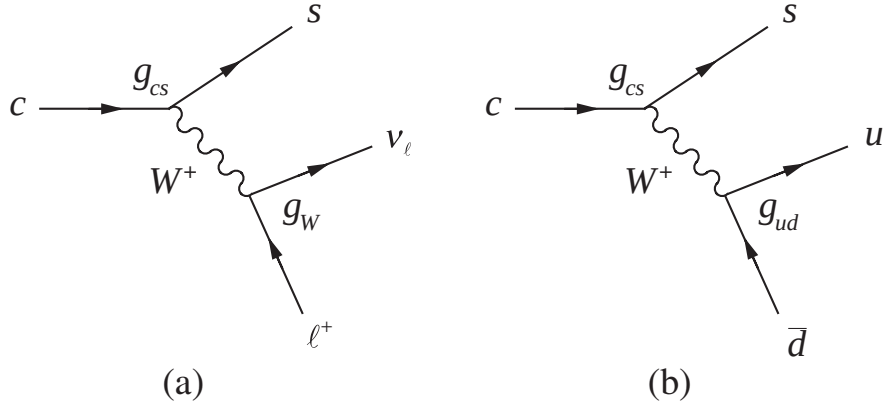


Figure 8.14 Cabibbo-allowed decays (8.21) of a charmed quark.

8.1.3 *W* boson decays

In Section 8.1.1, we saw that the W^+ boson can decay to lepton pairs $\ell^+\nu_\ell$ ($\ell = e, \mu, \tau$), by the mechanism of Figure 8.6(a). Lepton–quark symmetry implies that it can also decay to quark–antiquark pairs

$$W^+ \to \bar{u}d', \bar{c}s' \tag{8.22}$$

by the corresponding mechanism of Figure 8.15 with the same coupling strength α_W. Here d' and s' are given by (8.15). Thus (8.22) implies, for example,

$$\bar{u}d' = \bar{u}d\cos\theta_C + \bar{u}s\sin\theta_C,$$

so that $\bar{u}d$ and $\bar{u}s$ will occur with relative probabilities $\cos^2\theta_C$ and $\sin^2\theta_C$, respectively. In the rest frame of the decaying W^+, the quark and antiquark are emitted back-to-back with high energies $E \approx M_W/2 \approx 40\,\text{GeV}$. They fragment into hadron jets, like those shown in Figure 4.12, so that the observed decay reaction

$$W^+ \to \text{hadrons} \tag{8.23}$$

is dominated by the two-step mechanism of Figure 8.16. However, the rate for hadron decays is determined by the initial decays (8.22), in the same way that the rate for electron–positron annihilation into hadrons is determined by the rate for the initial quark reaction $e^+e^- \to q\bar{q}$ (cf. Section 7.2.1).

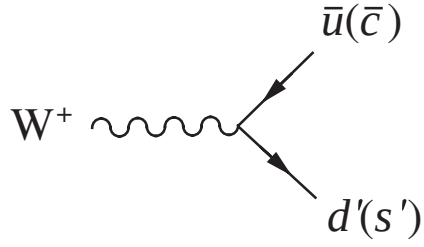


Figure 8.15 Feynman diagram for the decays $W^+ \to \bar{u}d'$, $\bar{c}s'$.

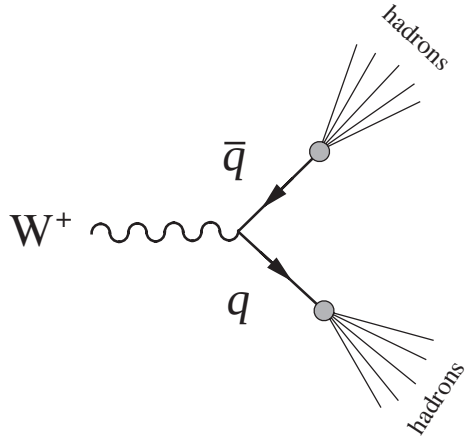


Figure 8.16 The dominant mechanism for the decay of $W^\pm$ bosons into hadrons.

The relative rates for the quark decays (8.22) and the leptonic decays

$$W^+ \to \ell^+ + \nu_\ell \quad (\ell = e, \mu, \tau) \tag{8.24}$$

are easily estimated in the good approximation that the final state lepton and quark masses are neglected. We then have

$$\Gamma(W^+ \to \bar{u}d') = \Gamma(W^+ \to \bar{c}s') = 3\Gamma(W^+ \to e^+\nu_e), \tag{8.25a}$$

since the mechanisms of these reactions are identical, but the $q\bar{q}$ pairs can be produced in three colours states with equal probabilities, while universality gives

$$\Gamma(W^+ \to e^+ \nu_e) = \Gamma(W^+ \to \mu^+ \nu_\mu) = \Gamma(W^+ \to \tau^+ \nu_\tau). \tag{8.25b}$$

Since these are the only first-order weak decays possible and there are two quark combinations (8.22) contributing to the hadron decays (8.23), we immediately arrive at the predicted branching ratios

$$B(W^+ \to \text{hadrons}) \approx 2/3, \tag{8.26a}$$

together with

$$B(W^+ \to \ell^+ \nu_\ell) \approx 1/9 \qquad (\ell = e, \mu, \tau), \tag{8.26b}$$

for each of the three leptonic decay modes. This simple approximation is in remarkably good agreement with experiment.

★8.1.4 Selection rules in weak decays

Many observations about the weak decays of hadrons are explained by $W^\pm$ exchange without the need for detailed calculation. For example, the decays

$$\Sigma^- \to n + e^- + \bar{\nu}_e \tag{8.27}$$

and

$$\Sigma^+ \to n + e^+ + \nu_e \tag{8.28}$$

seem very similar, where $\Sigma^+(1189) = uus$ and $\Sigma^-(1197) = dds$ are the charged Σ baryons discussed in Section 6.1.3. However, while reaction (8.27) is observed, reaction (8.28) is not, and the experimental upper limit on its rate relative to the observed decay is

$$\frac{\Gamma(\Sigma^+ \to n + e^+ + \nu_e)}{\Gamma(\Sigma^- \to n + e^- + \bar{\nu}_e)} < 5 \times 10^{-3}.$$

The reason for this is that reaction (8.27) is allowed via the mechanism of Figure 8.17, whereas no diagram with single W exchange can be drawn for reaction (8.28), which at the quark level is

$$uus \to udd + e^+ + \nu_e$$

and would require two separate quark transitions. It can go via a mechanism involving the emission and absorption of two W bosons, but this contribution is higher order in the weak interaction and is negligibly small.

Reaction (8.28) is just one of many that cannot proceed via single $W^\pm$ exchange and are therefore not observed despite the fact that they satisfy all the appropriate

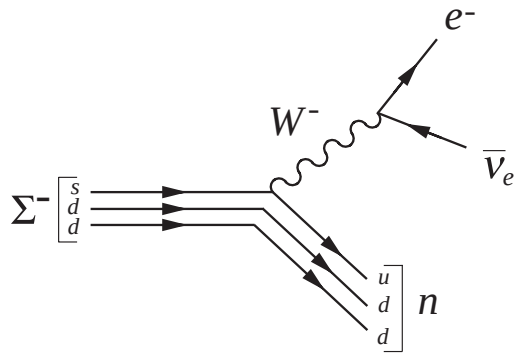


Figure 8.17 Mechanism for the decay $\Sigma^- \to n + e^- + \bar{\nu}_e$.

conservation laws for weak interactions. These 'forbidden' reactions can be identified systematically using a number of selection rules for single $W^{\pm}$-exchange processes, which can be deduced from the vertices of Figures 8.10 and 8.13. We shall illustrate this by considering the allowed changes of strangeness S.

We consider firstly semileptonic decays, like those of Figures 8.11 and 8.17. Since these involve a single $W^{\pm}$–quark vertex, the changes in the strangeness and electric charge of the hadrons are given by the possible changes in S and Q at this vertex. There are just two possibilities. If no strange quarks are involved at the vertex, as in Figures 8.10(a) and (b) and Figures 8.13(c) and (d), there is obviously no change in strangeness, while the quark charge changes by ± 1 depending on the charge of the W boson. Hence the changes ΔS and ΔQ in the strangeness and the electric charge of the hadrons satisfy

$$\Delta S = 0, \quad \Delta Q = \pm 1. \tag{8.29}$$

On the other hand, those vertices like Figures 8.10(c) and (d) and Figures 8.13(a) and (b) that do involve a strange quark give rise to processes like

$$u \to s + W^+ \quad \text{or} \quad W^- \to s + \bar{c},$$

in which the total quark charge and strangeness[7] both decrease, giving $\Delta S = \Delta Q = -1$, or processes like

$$s \to u + W^- \quad \text{or} \quad W^+ \to \bar{s} + c$$

in which the total quark charge and strangeness both increase, giving $\Delta S = \Delta Q = 1$. Thus the allowed semileptonic decays are characterized by the selection rules (8.29) and

$$\Delta S = \Delta Q = \pm 1, \tag{8.30}$$

[7] The charge, strangeness, and other quantum numbers of the quarks are listed for convenience in Appendix E, Table E.3.

where ΔQ is the change in the charge of the hadrons only. The latter is called the $\Delta S = \Delta Q$ *rule* for strangeness-changing decays, and decays with

$$\Delta S = -\Delta Q = \pm 1 \tag{8.31}$$

are forbidden. Reaction (8.28) is a typical example of a forbidden $\Delta S = -\Delta Q$ reaction requiring changes (8.31) since the Σ^+ has strangeness $S = -1$ and $Q = +1$ while the neutron has $S = 0$ and $Q = 0$. Other examples are discussed in Problem 8.4.

Interesting results are also obtained for purely hadronic decays. In such decays, the exchanged W boson must be both emitted and absorbed at $W^{\pm}$-quark vertices, as illustrated in Figure 8.9. Hence the selection rules for strangeness can be inferred by applying the selection rules (8.29) and (8.30) to each individual vertex, subject to the constraint that the change in the hadron charge must now be $\Delta Q = 0$ overall since no leptons are involved and the total charge must of course be conserved. If two vertices satisfying (8.29) are involved, strangeness is conserved and $\Delta S = 0$, while if one satisfies (8.29) and the other (8.30) we obviously have $\Delta S = \pm 1$. Finally, if two vertices satisfying (8.30) are involved, we again obtain $\Delta S = 0$ overall because of the charge conservation condition $\Delta Q = 0$. We thus obtain the selection rule

$$\Delta S = 0, \pm 1 \tag{8.32}$$

for hadronic weak decays, and the same selection rule also holds for semileptonic decays since it is guaranteed by the semileptonic selection rules (8.29) and (8.30).

The selection rule (8.32), which holds for all weak decays, has its most important implications for the decays of the so-called *cascade* particles

$$\Xi^0(1532) = ssu, \quad \Xi^-(1535) = ssd \quad (S = -2) \tag{8.33}$$

and the *omega-minus* baryon

$$\Omega^-(1672) = sss \quad (S = -3), \tag{8.34}$$

which we met briefly in Section 6.2.2.[8] Because the baryon number is conserved, these must ultimately decay to give a proton in the final state, since this is the only stable baryon. However, this cannot occur directly because of the selection rule (8.32), but must proceed via a series of successive decays, in which strangeness changes by at most one unit in each step. Thus the Ξ^- particle frequently decays by the sequence

$$\begin{aligned} \Xi^- &\to \Lambda^0 + \pi^- \\ &\quad \lfloor\!\to p + \pi^-, \end{aligned} \tag{8.35}$$

giving rise to a characteristic pattern of charged particle tracks. In contrast, decay processes like

$$\Xi^- \to n + \pi^-$$

[8] See Tables 6.5 and 6.6, and Table E.4 in Appendix E.

that satisfy all known conservation laws, but violate the selection rule (8.32), are never observed. Similar phenomena occur in Ω^- decays, and it is a remarkable fact that when the Ω^- particle was first observed, both its production and the whole decay sequence

$$
\begin{array}{l}
K^- + p \to \Omega^- + K^+ + K^0 \\
\quad \llcorner\!\!\rightarrow \Xi^0 + \pi^- (\Delta S = 1 \text{ weak decay}) \\
\quad\quad \llcorner\!\!\rightarrow \pi^0 + \Lambda (\Delta S = 1 \text{ weak decay}) \\
\quad\quad\quad \llcorner\!\!\rightarrow \pi^- + p \ (\Delta S = 1 \text{ weak decay}) \\
\quad\quad\quad \llcorner\!\!\rightarrow \gamma + \gamma \text{ (electromagnetic decay)} \\
\quad\quad\quad\quad \downarrow \quad \downarrow \\
\quad\quad\quad\quad e^+e^- \ e^+e^-
\end{array}
$$

was captured in a single bubble chamber picture, as shown in Figure 8.18.

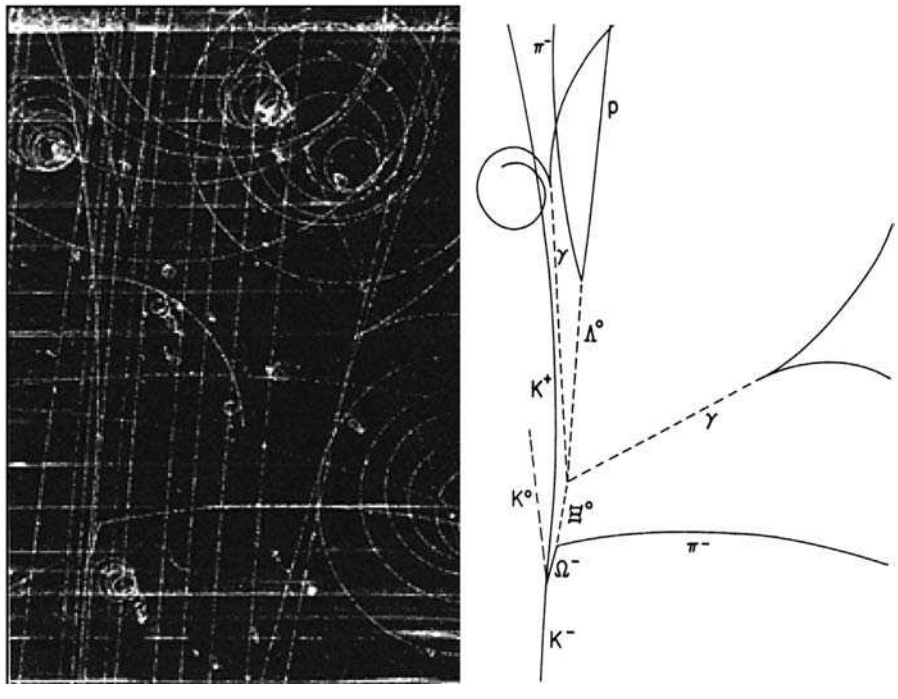


Figure 8.18 Characteristic pattern of tracks produced by the production and decay of the Ω^-. (From Barnes *et al.*, 1964, photo courtesy of Brookhaven National Laboratory.)

8.2 THE THIRD GENERATION

In Section 8.1.2 we deduced the weak couplings of the first two generations of quarks (u, d), (c, s) using the ideas of lepton–quark symmetry and quark mixing.

Historically, the most remarkable thing about these ideas is that they were formulated before the discovery of the charmed quark. In 1971 only seven fundamental fermions were known: the four leptons $\nu_e, e^-, \nu_\mu, \mu^-$ and the three quarks u, d, s. This led Glashow, Iliopoulos and Maiani to propose the existence of a fourth quark c to complete the lepton–quark symmetry and to solve problems associated with neutral currents.[9] The charmed quark was subsequently discovered in 1974 and its measured weak couplings are consistent with the predictions of lepton–quark symmetry, as we have seen.

Since 1974, events have moved on. By 1975 there were six known leptons

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}, \tag{8.36}$$

and by 1977 five known quarks

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} \\ b \end{pmatrix},$$

so that once again an extra quark – the top quark – was needed to restore lepton–quark symmetry. By 1994, the mass of this quark had been predicted to be

$$m_t = 170 \pm 30\,\mathrm{GeV/c^2}, \tag{8.37}$$

by arguments based on small effects in the unified theory of electroweak interactions (see Section 9.1.2). The top quark was finally detected at Fermilab in 1995 with a mass

$$m_t \approx 180\,\mathrm{GeV/c^2}, \tag{8.38}$$

compatible with the prediction of (8.37).

The rest of this section is devoted to the third generation of quarks b and t, and to the experiments that first established the existence of the latter.

8.2.1 More quark mixing

Lepton–quark symmetry was applied to the first two generations (8.10) and (8.11) in Section 8.1.2. In doing so, we allowed for mixing between the d and s quarks according to (8.15), which is conveniently rewritten in the matrix form[10]

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{8.39}$$

[9] Neutral currents are discussed in Section 9.1, and these problems in particular in Section 9.1.1.
[10] This is the most general form of mixing possible between two generations (see Problem 8.6).

When the third generation of quarks (t, b) is included, we must allow for the possibility of mixing between all three 'lower' quarks d, s and b instead of just the first two. This is done by generalizing (8.39) to

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{8.40}$$

where the so-called *CKM matrix*[11] $V_{\alpha\beta}(\alpha = u, c, t;\ \beta = d, s, b)$ must be unitary to ensure that d', s' and b' are orthonormal single-quark states, like d, s and b. Lepton–quark symmetry is then applied to the doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \quad \begin{pmatrix} c \\ s' \end{pmatrix}, \quad \begin{pmatrix} t \\ b' \end{pmatrix} \tag{8.41}$$

in order to obtain values for the $\alpha\beta W$ couplings $g_{\alpha\beta}$, where again $\alpha = u, c, t$ and $\beta = d, s, b$. The argument is analogous to that used earlier following Equation (8.16) and gives

$$g_{\alpha\beta} = g_W V_{\alpha\beta} \quad (\alpha = u, c, t; \beta = d, s, b). \tag{8.42}$$

To progress further, we must consider the values of the CKM matrix elements $V_{\alpha\beta}$. The general form of this matrix, analogous to (8.39) for two generations, is quite complicated and will be considered later, in Section 10.4. However, in the limit that the mixing between the b quark and the (d, s) quarks can be neglected, it reduces to

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} \cos\theta_C & \sin\theta_C & 0 \\ -\sin\theta_C & \cos\theta_C & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{8.43}$$

which is equivalent to our previous result (8.39) together with

$$b' = b. \tag{8.44}$$

This must be at least a good first approximation, since its predictions (8.22) for the couplings of the u, d, c and s quarks agree with the experimental data within errors,[12] as we saw in Section 8.1.2.

The accuracy of the approximation (8.43) can be very sensitively tested by considering decays of the b quark and of hadrons containing it. The b quark has a mass of approximately 4.5 GeV/c^2. It can decay to lighter quarks by the mechanism of

[11] The initials stand for Cabibbo, Kobayashi and Maskawa. The last two authors extended the original Cabibbo scheme to include the third generation of quarks.

[12] If the predictions (8.17) were exact, then (8.43) would follow exactly since the CKM matrix $V_{\alpha\beta}$ is a unitary matrix (cf. Problem 8.7).

Figure 8.19. As can be seen, all of these decay modes have rates that are proportional to one of the squared couplings

$$|g_{ub}|^2 = |V_{ub}|^2 g_W^2 \quad \text{or} \quad |g_{cb}|^2 = |V_{cb}|^2 g_W^2. \tag{8.45}$$

Hence the b quark is stable in the approximation (8.43), since both these couplings vanish.

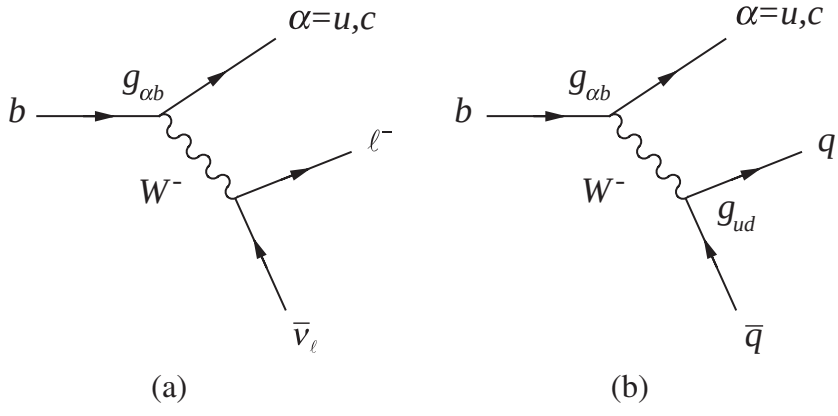


Figure 8.19 The dominant decays of the b quark to lighter quarks and leptons. Here $\ell = e, \mu$ or τ and $q\bar{q} = d\bar{u}, d\bar{c}, s\bar{u}$ or $s\bar{c}$.

Experimentally, b decays occur at rates that can be inferred from the decays of hadrons containing b quarks and imply

$$\tau_b \approx 10^{-12}\,\text{s} \tag{8.46}$$

for the b quark lifetime. This lifetime is very long compared with what one would expect if the couplings (8.45) were not suppressed, i.e. if

$$g_{ub} = g_{cb} = g_W. \tag{8.47}$$

To see this, we compare the decays of the b quark shown in Figure 8.19 with the decays of the lighter τ lepton shown in Figure 8.20. As can be seen, all the mechanisms are identical in the approximation (8.47) and the difference in the total decay rates of the τ and b can only arise from the different masses of the particles involved. This contributes in two ways. Firstly, by counting the numbers of different final states in Figures 8.19 and 8.20, it can be seen that for each τ lepton decay there are N decays of the b quark possible, where $N = 3$ and 4 for diagrams (a) and (b), respectively. Secondly, and more importantly, weak decay rates increase rapidly with the energy released in the decay. If the masses of the final state particles can be neglected, the decay rates increase as the fifth power of the mass of the decaying particle, as shown

for leptonic τ-decay in Section 2.2.2. In this simple but rough approximation[13] we obtain

$$\tau_b \approx \frac{1}{N}\left(\frac{m_\tau}{m_b}\right)^5 \tau_\tau \approx 10^{-15}\,\mathrm{s}$$

on substituting the experimental τ lifetime $\tau_\tau \approx 3 \times 10^{-13}$ s. This is much shorter than the observed lifetime (8.46), implying that the couplings (8.45) are heavily suppressed relative to (8.47), as stated. Furthermore, this qualitative conclusion is confirmed by a more thorough analysis of the various decay modes, which yields

$$|V_{ub}|^2 \approx 2 \times 10^{-5} \quad \text{and} \quad |V_{cb}|^2 \approx 2 \times 10^{-3}. \tag{8.48}$$

In other words, while the neglect of mixing between the b quark and the (d, s) quarks is not exact, it is a very good first approximation.

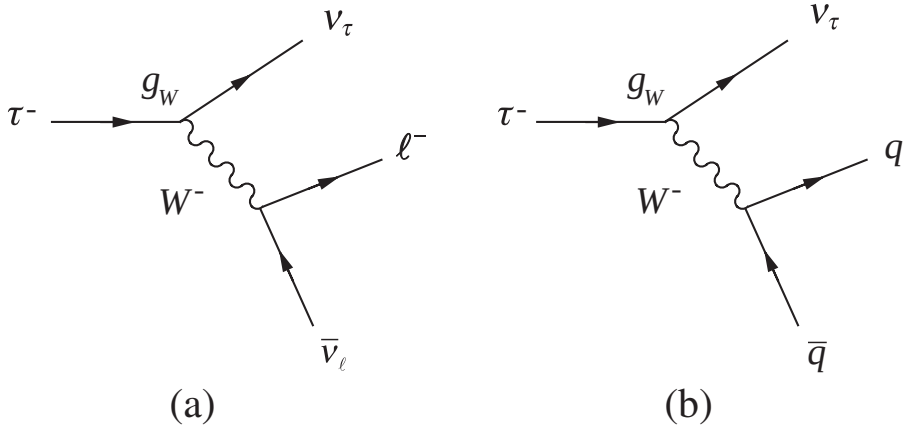


Figure 8.20 The dominant decays of the τ lepton to quarks and leptons. Here $\ell = e$ or μ and $q\bar{q} = d\bar{u}$ or $s\bar{u}$.

8.2.2 Properties of the top quark

The properties of the top quark differ markedly from those of the other quarks because it is so much heavier. In particular, it is much heavier than the $W^\pm$ bosons. Hence it can decay by the first-order weak interaction

$$t \to q + W^+ \quad (q = d, s, b),$$

[13] A more accurate treatment of the leptonic decay modes is given in Problem 8.8.

with rates proportional to the squared couplings $\left|g_{tq}\right|^2$, as shown in Figure 8.21. In the approximation (8.43), these couplings are predicted by (8.47) to be

$$g_{td} = 0, \quad g_{ts} = 0, \quad g_{tb} = g_W.$$

Hence the only significant decay mode is[14]

$$t \to b + W^+, \tag{8.49}$$

with a rate proportional to α_W, given by (8.9). A crude dimensional estimate of the decay width $\Gamma \sim \alpha_W m_t \sim 1\,\text{GeV}$ is enough to establish that the top quark is highly unstable. A full calculation of Figure 8.21 for $m_t = 180\,\text{GeV/c}^2$ leads to the prediction $\Gamma \approx 1.7\,\text{GeV}$, with a corresponding lifetime

$$\tau = \Gamma^{-1} \approx 4 \times 10^{-25}\,\text{s}. \tag{8.50}$$

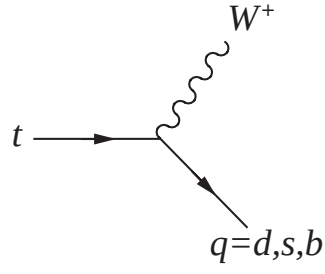


Figure 8.21 The mechanism for t quark decays. The decays that lead to b quarks are overwhelmingly the most important.

The predicted lifetime (8.50) is at the crux of top quark physics. By relativity, a hadron state of diameter $d \approx 1\,\text{fm}$[15] cannot form in a time less than $t \approx d/c = O(10^{-23}\,\text{s})$. The other five quarks u, d, s, c and b have lifetimes of order $10^{-12}\,\text{s}$ or more, and there is plenty of time for them to form hadrons, which can be observed in the laboratory. In contrast, when top quarks are created they decay too rapidly to form observable hadrons. Instead they decay by (8.49) to give a b quark[16] and a W boson, which in turn decays predominantly to either light quarks

$$W^+ \to q_1 + \bar{q}_2 \quad (q_1\bar{q}_2 = u\bar{d}, u\bar{s}, c\bar{d} \text{ or } c\bar{s}) \tag{8.51a}$$

[14] When deviations from the approximation (8.43) are taken into account, about 2 decays in 1000 lead to a d or s quark instead.

[15] A state formed of top quarks only (e.g. $t\bar{t}$) would be much smaller than this. This case is dealt with separately in Problem 8.9.

[16] In identifying the dominant decays, we neglect mixing between the b and (d, s) quarks.

or leptons

$$W^+ \to \ell^+ + \nu_\ell \quad (\ell = e, \mu, \tau). \tag{8.51b}$$

Furthermore, the quarks released in these decays are not seen directly, but 'fragment' into jets of hadrons. This is illustrated in Figure 8.22, which shows the observable final states resulting from top quark decay.

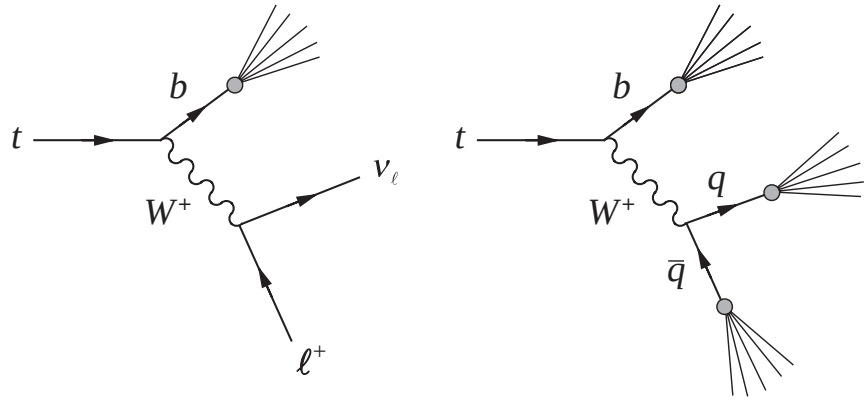


Figure 8.22 Production of hadrons jets from the decay $t \to b + W^+$, where the W boson decays to give either leptons or hadrons.

★8.2.3 Discovery of the top quark

Top quarks were first produced in pairs in the reaction

$$p + \bar{p} \to t + \bar{t} + X^0, \tag{8.52}$$

where X^0 is an arbitrary hadronic state allowed by the conservation laws. These pairs were identified by their subsequent decay products. The dominant decays of the t quark are shown in Figure 8.22, while the $\bar{t}$ quark decays by the corresponding antiparticle reactions. Clearly the final state resulting from the initial $t\bar{t}$ pair is very complicated and difficult to identify in the presence of backgrounds from other processes. In addition, very high energies are required if such pairs are to be produced at a reasonable rate in the reaction (8.52). The dominant mechanism for this is shown in Figure 8.23 and involves the quark–antiquark annihilation process[17]

$$q + \bar{q} \to t + \bar{t}. \tag{8.53}$$

[17] This mechanism is very similar to the mechanism for $W^\pm$ and Z^0 production shown in Figure 4.18. Consequently, there are several similarities between the present discussion and that of the discovery of the $W^\pm$ and Z^0 bosons in Section 4.5.1.

This can only occur if the total energy of the $q\bar{q}$ pair is at least $2m_t \approx 360\,\text{GeV}$, corresponding to the top quark and antiquark being produced at rest. Since each initial quark (or antiquark) carries only a fraction of the initial proton (or antiproton) energy, the energy of the $p\bar{p}$ system must be much higher if a reasonable reaction rate is to be obtained.

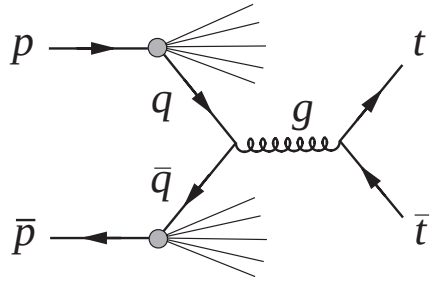


Figure 8.23 The dominant mechanism for top quark production in proton–antiproton collisions at 1.8 TeV, the energy of the $p\bar{p}$ collider at Fermilab.

These problems were first overcome by two experiments at Fermilab in 1995. In both cases, proton and antiproton colliding beams were brought together at the centre of very large and complex detectors. The proton and antiproton beams each had an energy of 900 GeV, corresponding to a total centre-of-mass energy of 1.8 TeV overall. Like the UA1 detector described in Section 4.5.1, both detectors could reconstruct very complicated multiparticle events, and both could detect and identify all long-lived particles except neutrinos. One of them (the CDF detector) is shown in Figure 4.24 (see also Plate 5). Here we describe its use to identify a particular class of events. In these, top quarks decay by (8.49) to give

$$t + \bar{t} \to b + W^+ + \bar{b} + W^-. \tag{8.54}$$

One of the W bosons then decays by (8.51a) to give light quarks, while the other decays by (8.51b) to give either an electron or a muon.[18] The net result is therefore either

$$t + \bar{t} \to \ell^+ + \nu_l + q + \bar{q} + b + \bar{b} \quad (\ell = e, \mu) \tag{8.55a}$$

or

$$t + \bar{t} \to \ell^- + \bar{\nu}_l + q + \bar{q} + b + \bar{b} \quad (\ell = e, \mu), \tag{8.55b}$$

[18] A top quark signal was also identified corresponding to events in which both W bosons decayed leptonically. However, there were fewer events in this case and they did not lead to a good determination for the top quark mass. Events in which both W bosons decay to give light quarks are much more difficult to separate from 'background' processes that give rise to a similar experimental signal.

where the quarks manifest themselves as hadron jets like those in Figure 7.14. The reconstruction of a typical event seen in the CDF detector is shown in the coloured Plate 6.

We shall proceed as follows: firstly, we identify an initial experimental signal corresponding to the desired events; then we consider what background processes could give rise to a similar signal and how they can be eliminated; finally, we present the results.

★ *8.2.3(a) Initial event selection*[19]

For each top quark event of the type (8.52) there are more than 10^9 events in which hadrons alone are produced. The extraction of a signal in the presence of this background is only possible because the top quarks are heavy and are produced with relatively low momenta. Because of this, their decay products are often emitted with large momenta at large angles to the initial beam direction. It is extremely rare for the hadrons produced in proton–antiproton collisions to be emitted with these characteristics. Hence the overwhelming majority of background events can be eliminated if events are selected that contain the combinations

$$\ell^+ + \nu_l + N \text{ jets} \quad (\ell = e, \mu) \tag{8.56a}$$

and

$$\ell^- + \bar{\nu}_l + N \text{ jets} \quad (\ell = e, \mu), \tag{8.56b}$$

where $N \geq 3$ and the leptons and jets are all required to have large momenta transverse to the incoming beams.[20] Of course, neutrinos cannot be observed directly. However, since they are the only long-lived particles that are not detected, their presence can be revealed by summing the transverse momenta p_T of all the observed particles. If this sum is not zero within errors, as required by momentum conservation, the 'missing transverse momentum' is ascribed to a neutrino.

Two further comments on the initial trigger (8.56) are required before we go on to see whether it can be produced in other ways. The first is that while (8.55) gives rise to four quarks, they will not always give rise to four distinct jets with high p_T. Sometimes one or more jets will be emitted close to the beam direction, where there are many hadrons from other sources. Sometimes two jets will be emitted in more or less the same direction and mimic a single jet in the detector. The trigger specifies events in which at least three distinct jets emerge at high transverse momentum.[21] The

[19] The method of selection described here is very similar to that used to select events in which a $W^{\pm}$ boson is produced in reaction (4.27) and decays to leptons by (4.28) (cf. Section 4.5.1). The essential difference is that we now also require the presence of hadron jets.

[20] In the CDF experiment, for example, the leptons were required to have transverse energies greater than 20 GeV. The jets were required to have transverse momenta greater than 15 GeV/c and their longitudinal momenta in the beam direction were restricted to ensure that they were emitted at a reasonably large angle to the beam direction.

[21] In complex reactions there is always a balance to be struck between a very precise trigger (e.g. four jets observed) with a smaller number of events and a looser trigger (e.g. at least two jets observed), giving rise to more severe background problems.

second point is that a high-momentum lepton could arise from the decay $Z^0 \to \ell^+\ell^-$ of a produced Z^0 boson, rather than a W boson decay. In this case, the lepton would be part of an $\ell^+\ell^-$ pair with invariant mass equal to the Z^0 mass. Such events are removed from the sample.

★ *8.2.3(b) Background events*

So far we have identified a distinctive class of events that can arise from the production and decay of top quarks. As in all experiments, it is necessary to consider whether such events could arise from other 'background' processes. In the present case, the most important backgrounds arise when the subprocess (8.53)

$$q + \bar{q} \to t + \bar{t}$$

is replaced by a subprocess of the type

$$q + \bar{q} \to W^{\pm} + (N \geq 3) \text{ jets.} \tag{8.57}$$

Examples of such processes, corresponding to

$$q + \bar{q} \to W + q + \bar{q} + g + g \tag{8.58a}$$

and

$$q + \bar{q} \to W + g + g + g, \tag{8.58b}$$

are shown in Figure 8.24. If the W boson decays to leptons, such processes can give rise to events that satisfy the trigger (8.56). Theoretical calculations and experimental observations both indicate that the ratio of events corresponding to this background to those for the 'signal' of top quarks is about 4:1.

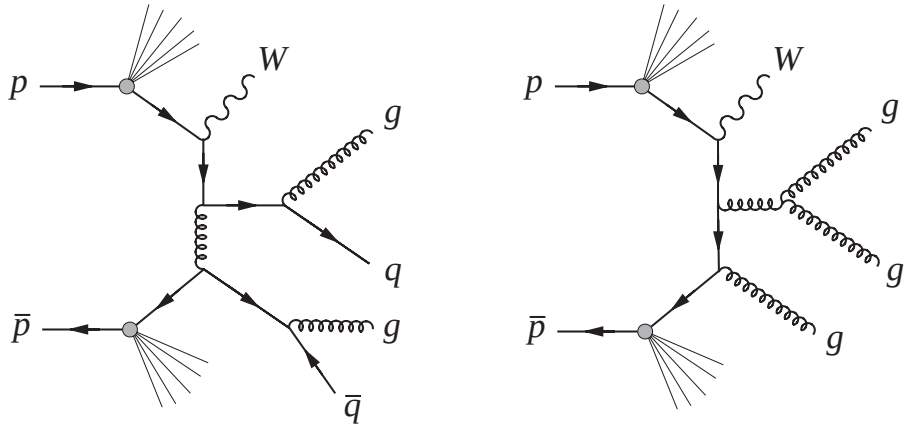


Figure 8.24 Reactions involving the subprocesses of Equations (8.57) that contribute to the background for top quark production. The quarks and gluons are observed as jets.

★ 8.2.3(c) *b-jet tagging*

Background events of the type (8.57) do not usually contain any jets associated with b quarks. This is illustrated by the examples of Figure 8.24, where the jets arise either from gluons or from quarks or antiquarks that originate from the proton or antiproton. In contrast, the top quark reaction (8.55) also invariably gives rise to b quarks. Hence the signal can be considerably enhanced relative to the background if 'b quark jets' can be distinguished from jets arising from other quarks and gluons. This is called 'b-jet tagging'.

One of the most successful methods of b-jet tagging relies on the fact that b-jets nearly always contain a fast-moving hadron with nonzero bottom $\tilde{B} \neq 0$. Such particles are characterized by decays to many-particle final states, with lifetimes of about 10^{-12} s. Hence b-jets typically contain multiprong decay vertices close to the production vertex. Other jets do not usually contain such vertices.[22] It is not always possible to resolve such vertices from the production vertices, even with silicon vertex detectors of the type discussed in Section 4.4.3. In the CDF experiment, however, it was possible to identify at least one of the b-jets in a $t\bar{t}$ event with about 40 % efficiency. Demanding such a trigger therefore reduces the signal by a factor of order 2, but reduces the background by a factor of order 20.

★ 8.2.3(d) *Final results*

The number of events of the type (8.56) obtained by CDF, with and without the b-jet tag, is shown in Table 8.1 for various numbers of observed hadron jets. As can be seen, there is a large signal compared with the expected background in the $N = 3, 4$ jet cases with a b-jet tag. Since the observed dependence of the background on the various triggers is in good agreement with theoretical expectation, this is compelling evidence for the existence of the top quark. Furthermore, since four-jet events correspond to all the decay products of the top quarks being observed, it is possible to reconstruct the top quark mass from these events. The resulting mass distribution for four-jet events with a b-jet tag is shown[23] in Figure 8.25. As can be seen, there is a clear peak, corresponding to a top quark mass

$$m_t = 176 \pm 8 \pm 10\,\mathrm{GeV/c^2},$$

TABLE 8.1 The number of lepton + N-jet events of the type (8.56) observed with and without b-jets. (Data from Abe *et al.*, 1995.)

N	Observed events	Observed b-jet tags	Background tags expected
1	6578	40	50 ± 12
2	1026	34	5021 ± 6.5
3	164	17	5.2 ± 1.7
4	39	10	1.5 ± 0.4

[22] Charmed particles decay with a similar lifetime, but to fewer particles.

[23] In this data sample, the cuts on the fourth jet were less stringent than those used in Table 8.1. This resulted in 19 events, of which 7 ± 2 are expected to be background events.

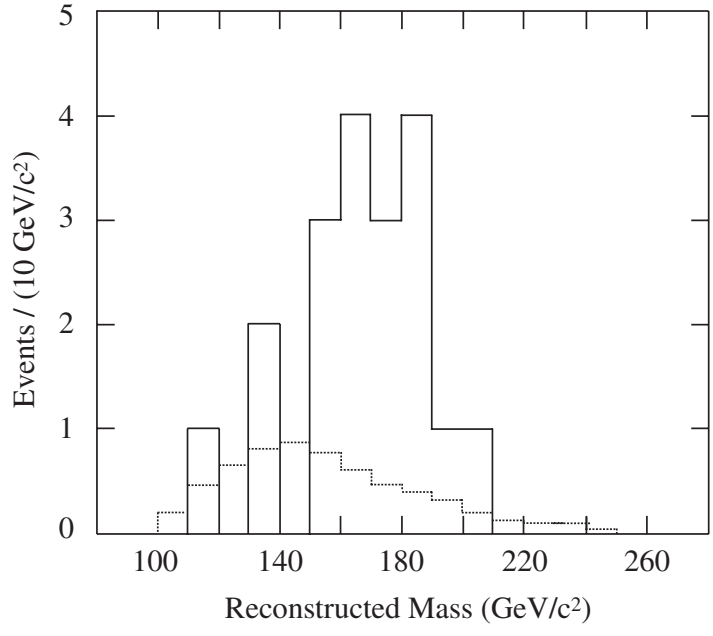


Figure 8.25 The reconstructed t quark mass distribution for the b-tagged 'four-jet plus lepton' events of type (8.56). The shape expected for background events (dotted line) is also shown. (Reprinted Figure 3 with permission from K. Abe *et al.*, *Phys. Rev. Lett.*, **74**, 2626. Copyright 1995 American Physical Society.)

where the first error is statistical and the second is systematic. A similar result was obtained by the DØ experiment, also performed at Fermilab. The current best average value using all data is

$$m_t = 174.2 \pm 3.3\,\mathrm{GeV/c^2}. \tag{8.59}$$

PROBLEMS 8

8.1 Show that the vertices of Figure 8.4 are the only possible abW vertices allowed by charge and lepton number conservation, where a and b can be any lepton or antilepton.

8.2 The mechanisms of muon decay (8.4) and the semileptonic decay of the charmed quark (8.21a) are essentially the same. Use this to estimate the rate for the charmed quark decay, assuming that its mass is $1.5 \pm 0.1\,\mathrm{GeV/c^2}$ and the masses of final state particles may be neglected in both decays. (*Hint*. Since m_c, $m_\mu \ll M_W$, a zero-range approximation is appropriate in both cases.)

8.3 The charmed particle decays

$$D^+(1869) \to e^+ + \text{anything}$$
$$D^0(1869) \to e^+ + \text{anything}$$
$$\Lambda_c^+(2285) \to e^+ + \text{anything}$$

are usually assumed to result from the charmed quark decay $c \to s + e^+ + \nu_e$ discussed in the previous problem. Use the data given in Tables E.4 and E.5 to test this assumption by

comparing the measured rates of these decays with each other, and with the estimated rate for the charmed quark decay obtained in the previous problem.

8.4 Which of the following six decays are allowed in lowest-order weak interactions?

(a) $K^+ \to \pi^+ + \pi^+ + e^- + \bar{\nu}_e$
(b) $K^- \to \pi^+ + \pi^- + e^- + \bar{\nu}_e$
(c) $\Xi^0 \to \Sigma^- + e^+ + \nu_e$
(d) $\Omega^- \to \Xi^0 + e^- + \bar{\nu}_e$
(e) $\Xi^0 \to p + \pi^- + \pi^0$
(f) $\Omega^- \to \Xi^- + \pi^+ + \pi^-$

8.5 Classify the following semileptonic decays of the $D^+(1869) = c\bar{d}$ meson as Cabibbo-allowed, Cabibbo-suppressed or forbidden in lowest-order weak interactions, by finding selection rules for the changes in strangeness, charm and electric charge in such decays:

(a) $D^+ \to K^- + \pi^+ + e^+ + \nu_e$
(b) $D^+ \to K^+ + \pi^- + e^+ + \nu_e$
(c) $D^+ \to \pi^+ + \pi^+ + e^- + \bar{\nu}_e$
(d) $D^+ \to \pi^+ + \pi^- + e^+ + \nu_e$

8.6 Show that an arbitrary $n \times n$ unitary matrix has n^2 real parameters, and hence that

$$\mathbf{U} = \mathrm{e}^{-i\alpha} \begin{pmatrix} \cos\theta_C \mathrm{e}^{i\beta} & \sin\theta_C \mathrm{e}^{i\gamma} \\ -\sin\theta_C \mathrm{e}^{-i\gamma} & \cos\theta_C \mathrm{e}^{-i\beta} \end{pmatrix} \tag{8.60}$$

is the most general form of a 2×2 unitary matrix.

The most general form of (d, s) mixing is

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \mathbf{U} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{8.61}$$

where $\mathbf{U}$ is an arbitrary 2×2 unitary matrix, $\mathbf{U}^\dagger \mathbf{U} = 1$. Show that this can be reduced to the form (8.39) by adjusting the arbitrary phases of the quark states s, s' and d.

8.7 Show that if (8.17) were exact, then (8.43) would follow exactly, since the CKM matrix $V_{\alpha\beta}$ is a unitary matrix.

8.8 Use the method of dimensions to show that

$$\frac{\Gamma(b \to q + e^- + \bar{\nu}_e)}{\Gamma(\tau^- \to \nu_\tau + e^- + \bar{\nu}_e)} = \left|V_{qb}\right|^2 \frac{m_b^5}{m_\tau^5} \quad (q = u, c), \tag{8.62}$$

if W exchange is approximated by a zero-range interaction and the masses of all final state particles are neglected.

If quark mass corrections are taken into account, this result is modified to

$$\frac{\Gamma(b \to q + e^- + \bar{\nu}_e)}{\Gamma(\tau^- \to \nu_\tau + e^- + \bar{\nu}_e)} = \left|V_{qb}\right|^2 \frac{m_b^5}{m_\tau^5} f(m_q/m_b), \tag{8.63}$$

where

$$f(x) = 1 - 8x^2 - 24x^4 \ln(x) + 8x^6 - x^8.$$

Use this, together with the experimental result

$$\Gamma(b \to c + e^- + \bar{\nu}_e) + \Gamma(b \to u + e^- + \bar{\nu}_e) = (6.8 \pm 0.5) \times 10^{10}\,\mathrm{s}^{-1}, \tag{8.64}$$

deduced from the decay $B^- \to e^- + \bar{\nu}_e +$ hadrons, to obtain upper bounds on the magnitudes of V_{ub} and V_{cb}.

8.9 If the top quark were stable, the low-lying states of toponium $t\bar{t}$ could be approximated by nonrelativistic motion in a Coulomb potential (cf. Equation (7.2a))

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r}$$

with $\alpha_s \approx 0.1$. Use the simple Bohr model to calculate the radius of the ground state and the time taken to complete a single Bohr orbit in the ground state. Compare this with the expected lifetime of the top quark.

9

Weak Interactions: Electroweak Unification

In this chapter we continue our discussion of weak interactions by turning our attention to the unified theory of weak and electromagnetic processes. This theory was first proposed in order to solve problems associated with Feynman diagrams in which more than one W boson was exchanged, like that shown in Figure 9.1(a), which contributes to the reaction

$$e^+ + e^- \to \mu^+ + \mu^-. \tag{9.1}$$

Such contributions are expected to be small because they are higher order in the weak interaction. This appears to be confirmed by experimental data, which are in good agreement with theoretical predictions that neglect them entirely. However, when these contributions are explicitly calculated, they are found to be proportional to divergent integrals; i.e. they are infinite.[1] In the unified theory, this problem is automatically solved when other higher-order diagrams, like those shown in Figures 9.1(b) and (c), are taken into account. These also give infinite contributions, but when all the diagrams of a given order are added together the divergences cancel, giving a well-defined and finite contribution overall. This cancellation is not accidental, but is a consequence of relations between the various coupling constants associated with the γ, $W^\pm$ and Z^0 vertices in the unified theory, as we shall see in Section 9.1.2.

[1] These infinites can be removed by modifying the theory in an ad hoc manner, but at the price of introducing extra free parameters, which is almost as bad.

Particle Physics Third Edition B. R. Martin and G. Shaw

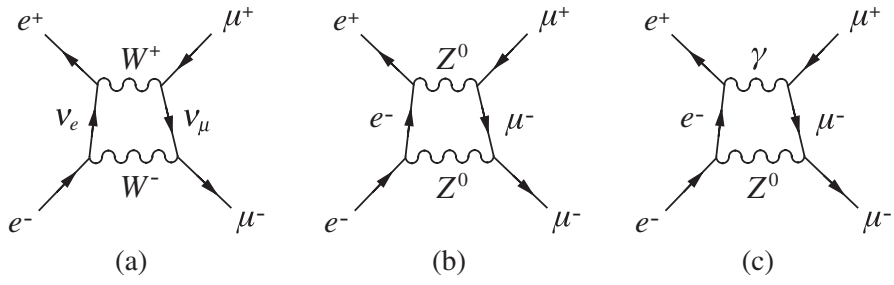


Figure 9.1 Some of the higher-order diagrams that contribute to the reaction $e^+ + e^- \to \mu^+ + \mu^-$ in the unified theory of the electroweak interaction.

We shall say little about the net contributions of higher-order diagrams like those of Figure 9.1 because they are not only finite but very small compared with the dominant lowest-order diagrams. Rather, we shall concentrate initially on lowest-order diagrams, like those shown for the same reaction (9.1) in Figure 9.2, and then go on to other aspects of the unified theory, including the elusive Higgs boson, in Section 9.2.

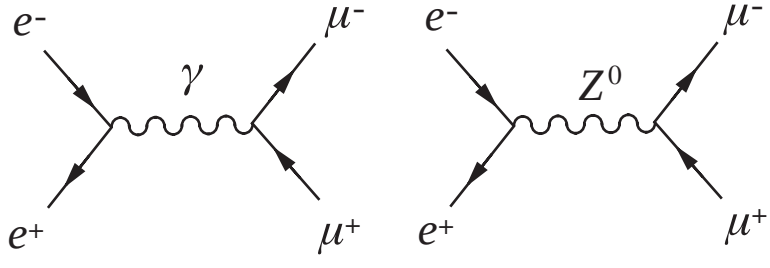


Figure 9.2 The two dominant contributions to the reaction $e^+ + e^- \to \mu^+ + \mu^-$ in the unified theory.

9.1 NEUTRAL CURRENTS AND THE UNIFIED THEORY

This section is devoted to weak neutral current reactions, which are those involving the emission, absorption or exchange of Z^0 bosons. The existence of such reactions was predicted by the unified theory in 1968, several years before their experimental detection. They play a central role in the theory, as we shall see.

9.1.1 The basic vertices

In Section 8.2 we saw that all the charged current interactions of leptons could be understood in terms of the basic $W^\pm$–lepton vertices of Figure 8.4. In the same way, all known neutral current interactions can be accounted for in terms of the basic Z^0–lepton vertices of Figure 9.3. Like the $W^\pm$–lepton vertices, these conserve the lepton numbers L_e, L_μ and L_τ in addition to the electric charge Q, and indeed they are the only possible vertices that can be written down which satisfy these conservation laws. The corresponding quark vertices can be obtained from the lepton vertices

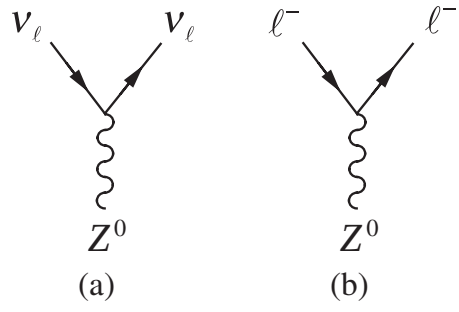


Figure 9.3 The basic Z^0–lepton vertices, where $\ell = e, \mu, \tau$.

using lepton–quark symmetry and quark mixing, in the same way that the $W^\pm$–quark vertices were obtained from the $W^\pm$–lepton vertices in Section 8.1.2. In other words, if we again confine ourselves to the first two generations, the quark vertices are obtained from the lepton vertices in Figure 9.3 by making the replacements

$$\nu_e \to u, \quad \nu_\mu \to c, \quad e^- \to d', \quad \mu^- \to s', \tag{9.2}$$

where d' and s' are the mixtures (8.15). For $\ell = e$ or μ, the lepton vertices of Figure 9.3 can be denoted

$$\nu_e \nu_e Z^0, \quad \nu_\mu \nu_\mu Z^0, \quad e^- e^- Z^0, \quad \mu^- \mu^- Z^0$$

in an obvious notation. The corresponding quark vertices are therefore

$$uuZ^0, \quad ccZ^0, \quad d'd'Z^0, \quad s's'Z \tag{9.3}$$

in the same notation, where the latter two must be interpreted in terms of the physical states d and s. In Section 8.1.2 we had a vertex

$$ud'W = udW\cos\theta_C + usW\sin\theta_C,$$

which we interpreted as udW and usW vertices with relative strengths $\cos\theta_C$ and $\sin\theta_C$, as illustrated in Figure 8.12. In the same way we have

$$\begin{aligned} d'd'Z^0 &= (d\cos\theta_C + s\sin\theta_C)(d\cos\theta_C + s\sin\theta_C)Z^0 \\ &= ddZ^0\cos^2\theta_C + ssZ^0\sin^2\theta_C + (dsZ^0 + sdZ^0)\sin\theta_C\cos\theta_C \end{aligned} \tag{9.4a}$$

and

$$\begin{aligned} s's'Z^0 &= (-d\sin\theta_C + s\cos\theta_C)\,(-d\sin\theta_C + s\cos\theta_C)Z^0 \\ &= ddZ^0\sin^2\theta_C + ssZ^0\cos^2\theta_C - (dsZ^0 + sdZ^0)\sin\theta_C\cos\theta_C, \end{aligned} \tag{9.4b}$$

so that the primed vertices $d'd'Z^0$ and $s's'Z^0$ both contribute to the vertices ddZ^0, ssZ^0, sdZ^0 and dsZ^0 for the real particles d and s. However, the interesting

point is that when both sets of contributions (9.4a) and (9.4b) are combined, one obtains

$$d'd'Z^0 + s's'Z^0 = ddZ^0 + ssZ^0,$$

so that the four vertices (9.3) can be replaced by the equivalent vertices

$$uuZ^0, \quad ccZ^0, \quad ddZ^0, \quad ssZ^0, \tag{9.5}$$

which are shown in Figure 9.4.

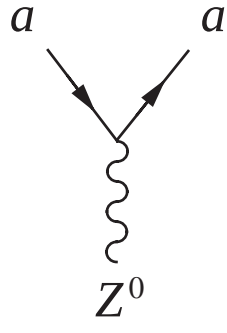


Figure 9.4 The basic Z^0–quark vertices, where $a = u, d, s, \ldots$

The four vertices of Figure 9.4 conserve both strangeness and charm, whereas the 'flavour-changing' vertices ucZ^0 and dsZ^0 do not occur. Thus we arrive at the conclusion that neutral current interactions conserve strangeness and charm, in contrast to the charged current interactions that do not in general conserve these quantum numbers. This result remains valid when mixing between all three generations is considered, and is confirmed by experiment. For example, the strangeness-changing decays

$$K^+ \to \pi^0 + \mu^+ + \nu_\mu \tag{9.6a}$$

and

$$K^+ \to \pi^+ + \nu_\ell + \bar{\nu}_\ell, \tag{9.6b}$$

shown in Figure 9.5, differ in that the first involves the allowed charged current vertex usW of Figure 8.13(a) whereas the second involves the forbidden neutral current vertex dsZ^0. Experimentally, no forbidden decays (9.6b) have been observed, and the upper limit on their possible decay rates is

$$\frac{\sum_\ell K^+ \to \pi^+ + \nu_\ell + \bar{\nu}_\ell}{K^+ \to \pi^0 + \mu^+ + \nu_\mu} < 10^{-7}, \tag{9.7}$$

in impressive agreement with the above scheme.

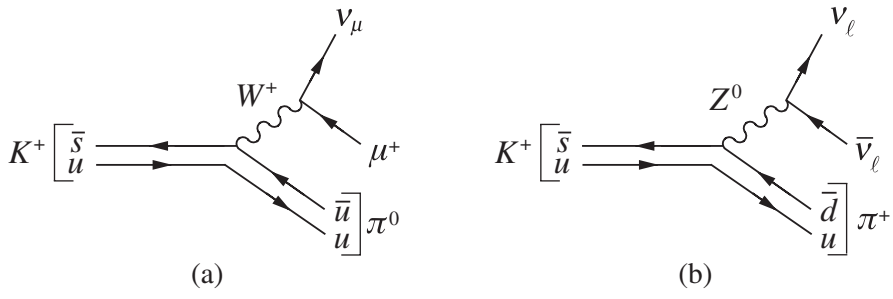


Figure 9.5 (a) The allowed decay $K^+ \to \pi^0 + \mu^+ + \nu_\mu$ and (b) the forbidden decay $K^+ \to \pi^+ + \nu_\ell + \bar{\nu}_\ell$.

9.1.2 The unification condition and the $W^\pm$ and Z^0 masses

The various coupling constants that occur in electroweak interactions are not independent, but are related by the requirement that the infinities that occur in higher-order diagrams, like those of Figure 9.1, should exactly cancel in all possible processes. This is guaranteed in the unified theory provided that two equations, called the *unification condition* and the *anomaly condition*, are satisfied. The proof that this is so is formidable and relies on a fundamental symmetry of the theory called gauge invariance.[2] However, the form of the equations is simple. The unification condition is

$$\frac{e}{2(2\varepsilon_0)^{1/2}} = g_W \sin\theta_W = g_Z \cos\theta_W, \tag{9.8}$$

where the *weak mixing angle* θ_W is given by

$$\cos\theta_W = M_W/M_Z \quad (0 < \theta_W < \pi/2) \tag{9.9}$$

and g_Z is a coupling constant that characterizes the strength of the neutral current vertices[3] of Figures 9.3 and 9.4. This condition explicitly relates the weak and electromagnetic coupling constants. In contrast, the anomaly condition relates the electric charges Q_ℓ and Q_a of the leptons ℓ and quarks a, and is

$$\sum_\ell Q_\ell + 3\sum_a Q_a = 0. \tag{9.10}$$

The sums extend over all leptons ℓ and all quark flavours $a = u, d, s, \ldots$ and the factor 3 arises because there is also a sum over the three quark colour states, which has been explicitly carried out. On substituting for the quark and lepton charges, one easily finds that the anomaly condition is satisfied by the six known leptons and the six known quarks.

---

[2] This is discussed in Section 9.2.1 and also in Appendix D.

[3] The strengths are not all equal, but are given by g_Z multiplied by known constants, which depend on θ_W and the vertex in question. Their explicit form for leptons is derived in Appendix D, Section D.7.4.

The unification condition (9.8) relates the strengths of the various interactions to the W and Z masses and historically was used to predict the latter from the former before the $W^{\pm}$ and Z^0 bosons were discovered. In the low-energy limit, the charged current reactions are characterized by the Fermi constant, and on substituting for g_W from (9.8) one obtains

$$M_W^2 = \frac{\sqrt{2}g_W^2}{G_{\mathrm{F}}} = \frac{\pi\alpha}{\sqrt{2}G_{\mathrm{F}}\sin^2\theta_W}, \tag{9.11a}$$

which together with (9.9) implies

$$M_Z^2 = \frac{\pi\alpha}{\sqrt{2}G_{\mathrm{F}}\sin^2\theta_W\cos^2\theta_W} \tag{9.11b}$$

for the Z^0 mass. The weak mixing angle itself can be determined by comparing neutral and charged current processes at low energies $E \ll M_W, M_Z$. In this regime, neutral and charged current processes are characterized by an effective zero-range coupling constant G_Z in exactly the same way that charged current reactions are characterized by the Fermi coupling constant G_F.[4] The neutral current coupling is given by

$$\frac{G_Z}{\sqrt{2}} = \frac{g_Z^2}{M_Z^2}, \tag{9.12}$$

by analogy with (2.17) for the Fermi constant G_F, and the ratio of these low-energy couplings can be expressed in the form

$$\frac{G_Z}{G_{\mathrm{F}}} = \frac{g_Z^2}{g_W^2}\cdot\frac{M_W^2}{M_Z^2} = \sin^2\theta_W, \tag{9.13}$$

using (9.8) and (9.9). Hence the weak mixing angle θ_W may be found by comparing the measured rates of charged and neutral current reactions at low energies, and by 1981 its value was determined in this way to be

$$\sin^2\theta_W = 0.227 \pm 0.014.$$

This value was used to predict the $W^{\pm}$ and Z^0 masses by substitution into (9.11a) and (9.11b) and using the values of α and G_F given in (1.16) and (2.16) respectively. The resulting values were

$$M_W = 78.3 \pm 2.4\,\mathrm{GeV/c^2} \text{ and } M_Z = 89.0 \pm 2.0\,\mathrm{GeV/c^2}$$

[4] This was discussed in Section 2.2.1.

and the subsequent discovery of the $W^{\pm}$ and Z^0 bosons with masses compatible with these predictions (as described in Section 4.5.1) is perhaps the greatest triumph of the unified theory.

Nowadays, the best value of the weak mixing angle,

$$\sin^2 \theta_W = 0.2315 \pm 0.0001 \tag{9.14}$$

is obtained by comparing the predictions of the unified theory with a wide range of measurements on different neutral current reactions. However, on substituting into (9.11) this gives

$$M_W = 77.50 \pm 0.03\,\mathrm{GeV/c^2} \text{ and } M_Z = 88.41 \pm 0.04\,\mathrm{GeV/c^2}, \tag{9.15}$$

which are not in very good agreement with the best experimental values (8.1). The reason for this is well understood. It arises because in deriving (9.11) we used the relation (2.17) for the Fermi constant G_F. This latter relation was obtained by taking the low-energy limit of single $W^{\pm}$ exchange only, as shown in Figure 2.7, whereas strictly speaking we should have also included the small contributions arising from higher-order diagrams. Two of the most important of these are shown in Figure 9.6. One of these involves an interaction between the $W^{\pm}$ and Z^0 bosons which is predicted by the unified theory, whereas the other involves the t quark. Hence the magnitude of the higher-order corrections to (9.11), and also to other predictions obtained using just lowest-order diagrams involving single $W^{\pm}$ and Z^0 exchange, depends on the mass of the t quark. We shall not discuss this in detail, but merely state the important result that when higher-order corrections are taken into account, the predictions of the theory agree with experiment in all cases.

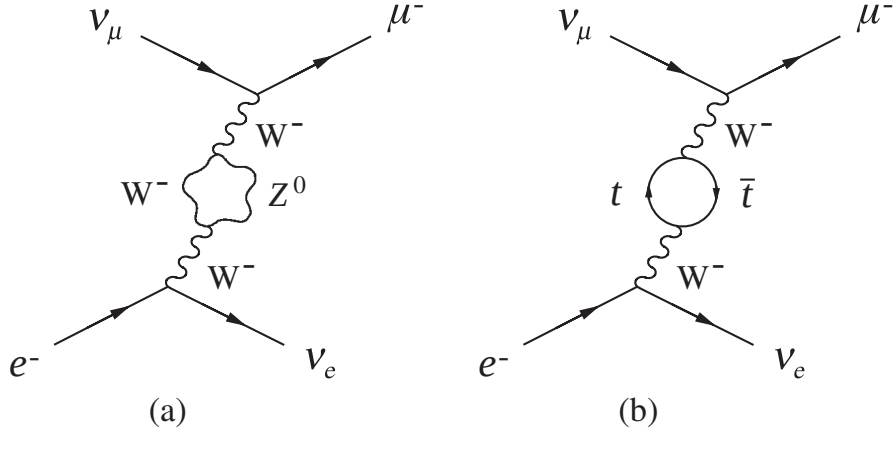


Figure 9.6 Two of the higher-order contributions to inverse muon decay which were neglected in obtaining the Fermi coupling constant G_F as a low-energy limit in Figure 2.7

9.1.3 Electroweak reactions

In any process in which a photon is exchanged, a Z^0 boson can be exchanged as well. This is illustrated for elastic electron–proton scattering in Figure 9.7, and follows in general from the fact that to each of the basic electromagnetic vertices there is a corresponding Z^0 vertex, as illustrated in Figure 9.8. At energy and momentum transfers that are small compared with the Z^0 mass, the Z^0-exchange contribution can be neglected compared with the corresponding photon exchange contributions and the reactions regarded as purely electromagnetic to a high degree of accuracy. However, at very high energy and momentum transfers, Z^0-exchange contributions become comparable with photon exchange, and we are therefore dealing with genuinely electroweak processes that involve both weak and electromagnetic interactions to a comparable degree.

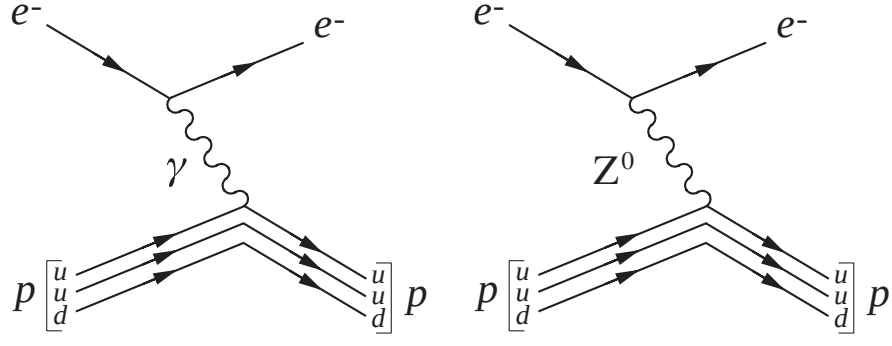


Figure 9.7 Photon and Z^0 boson exchange contributions to electron–proton scattering.

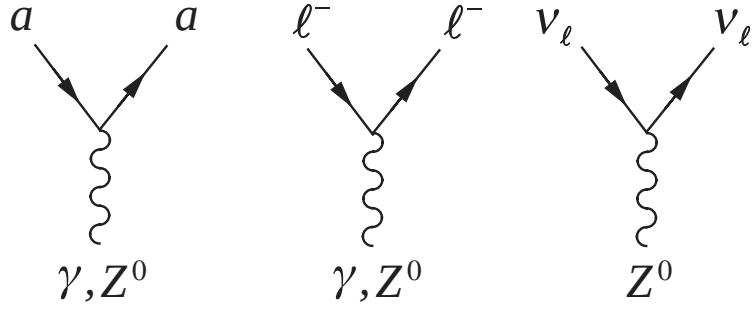


Figure 9.8 Summary of the Z^0 and γ couplings to leptons and quarks in the unified electroweak theory, where $\ell = e, \mu, \tau$ and $a = u, d, s, \ldots$

These points are beautifully illustrated by the cross-section for the muon pair production reaction,

$$e^+ + e^- \to \mu^+ + \mu^-, \tag{9.16}$$

which is shown in Figure 9.9. This is because the main features of this cross-section may be understood in terms of the dominant lowest-order diagrams of Figure 9.2

without detailed calculation, if we neglect the lepton masses compared with the beam energies E in the overall centre-of-mass frame. A simple dimensional estimate of the one-photon exchange contribution of Figure 9.2(a) then gives

$$\sigma_\gamma \approx \alpha^2/E^2, \tag{9.17}$$

where the factor α^2 follows because the diagram is second-order, and the energy dependence follows because the cross-section is an area with natural dimension $[E^{-2}]$[5] and the beam energy E is the only dimensional parameter available if we neglect the lepton masses. The same argument gives

$$\sigma_Z \approx G_Z^2 E^2 \tag{9.18}$$

for the contribution of the Z^0-exchange diagram of Figure 9.2(b) at energies $E^2 \ll M_Z^2$, where G_Z is the effective low-energy constant (9.12) which has natural dimensions $[E^{-2}]$. Finally, the ratio of (9.18) and (9.17) gives

$$\frac{\sigma_Z}{\sigma_\gamma} \approx \frac{G_Z^2 E^4}{\alpha^2} \approx \frac{E^4}{M_Z^4} \tag{9.19}$$

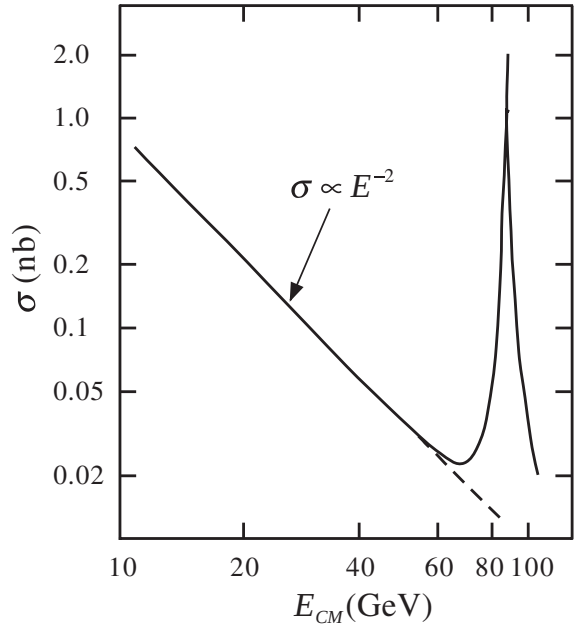


Figure 9.9 Total cross-section for the reaction $e^+ + e^- \to \mu^+ + \mu^-$ as a function of the total centre-of-mass energy (9.20). The dashed line shows the extrapolation of the low-energy behaviour (9.17) in the region of the Z^0 peak.

[5] See Table 1.1.

for the relative importance of the two diagrams, where we have used (9.13) and (9.11b) and neglected factors of order unity.[6]

We thus see that the one-photon exchange diagram dominates at low energies, and the cross-section falls as E^{-2} by (9.17). This is in agreement with the observed behaviour shown in Figure 9.9 and justifies our neglect of the Z^0-exchange contribution in our previous discussion of this process in Section 7.2. However, the relative importance of the Z^0- exchange contribution of Figure 9.2(b) increases rapidly with energy, as can be seen from (9.19), and at beam energies of about 25 GeV it begins to make a significant contribution to the total cross-section. At still higher energies, the cross-section is dominated by a very large peak at a centre-of-mass energy

$$E_{\mathrm{CM}} = 2E = M_Z, \tag{9.20}$$

corresponding to the Z^0 mass. At this energy the low-energy approximation (9.18) is irrelevant and Figure 9.2(b) corresponds to the formation of physical Z^0 bosons in the process

$$e^+ + e^- \to Z^0, \tag{9.21}$$

followed by the subsequent decay

$$Z^0 \to \mu^+ + \mu^-, \tag{9.22}$$

to give the final state muons. This is exactly like the formation and decay of a resonance, as shown, for example, in Figure B.4 except that we are now dealing with weak rather than strong interactions. As in the latter case, there is a peak in the total cross-section at the particle mass, with a width equal to the total decay rate; i.e.

$$\Gamma_Z = 1/\tau_Z = 2.49\,\mathrm{GeV}, \tag{9.23}$$

where τ_Z is the lifetime of the Z^0 boson. This peak is an enormous boon to experimental physics, as we shall now discuss.

9.1.4 Z^0 formation: how many neutrinos are there?

The Z^0 bosons produced in the reaction (9.21) will not only decay to muon pairs, as in (9.22), but also to other final states allowed by the appropriate conservation laws. Hence an electron–positron colliding-beam machine tuned to a centre-of-mass energy corresponding to the Z^0 mass is an ideal 'Z^0 factory', in which these particles can be copiously produced and studied. Two such colliders have been built: the Stanford linear collider (SLC), which was the first to produce results, and the 'large electron–positron' (LEP) collider at CERN, which produced Z^0 bosons much more copiously, at a rate of order 10^3 per hour. This enabled the mass, width and decay modes of the

[6] A full calculation, which is beyond the scope of this book, gives the same energy dependence as (9.19), but is about a factor of 4 smaller.

Z^0 to be measured with great precision, and in so doing verified that there are only three neutrino types in the sequence $\nu_e, \nu_\mu, \nu_\tau, \ldots$, as we shall see.

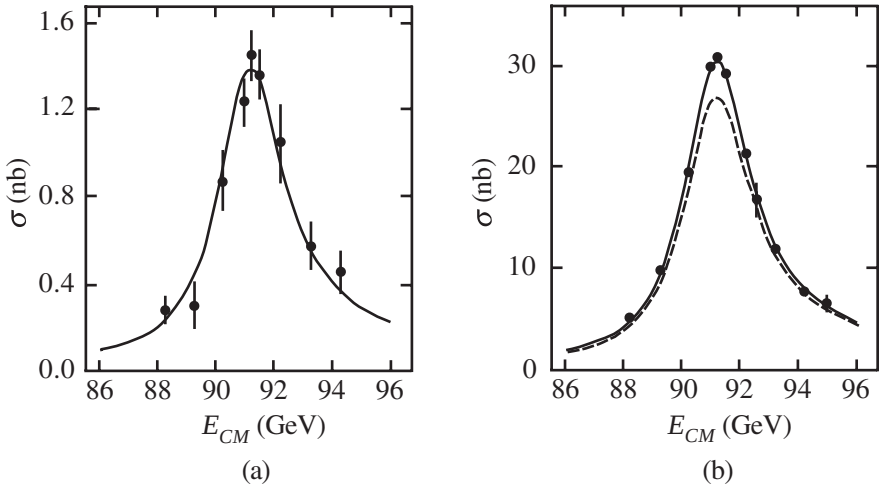


Figure 9.10 Measured cross-sections for (a) $e^+ + e^- \to \mu^+ + \mu^-$ and (b) $e^+ + e^- \to$ hadrons, in the region of the Z^0 peak. The solid and dashed lines show the predictions of the standard model on the assumptions that there are three and four types of light neutrinos, respectively. (Reprinted from Akrawy, M. Z., *et al.*, *Physics Letters B*, **240**, 497. Copyright 1990, with permission from Elsevier.)

The processes observed at the SLC and LEP collider are

$$e^+ + e^- \to \ell^+ + \ell^- \quad (\ell = e, \mu, \tau) \tag{9.24}$$

and

$$e^+ + e^- \to \text{hadrons}, \tag{9.25}$$

where in the latter case a sum over all possible hadron states is implied. In all cases there is a clear peak in the cross-section, as illustrated by the data from one experiment in Figures 9.10 (a) and (b). If this peak is interpreted as due to formation and decay of Z^0 bosons, as shown in Figure 9.11, then it can be shown that the cross-section is described by the Breit–Wigner formula (cf. Section B.5.2, Equation (B.46b))

$$\sigma(e^+ + e^- \to X) = \frac{12\pi M_Z^2}{E_{\text{CM}}^2}\left[\frac{\Gamma(Z^0 \to e^+ + e^-)\Gamma(Z^0 \to X)}{(E_{\text{CM}}^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2}\right], \tag{9.26}$$

where $\Gamma(Z^0 \to X)$ is the decay rate of the Z^0 to the observed final state X and Γ_Z is the total decay rate (9.23). The factor outside the square brackets is essentially kinematic in origin and we will not discuss it further. The denominator of the term in square brackets is the Breit–Wigner factor that was used in Section 3.5; the factor $\Gamma(Z^0 \to X)$ arises because the cross-section for a particular final state X will obviously

be proportional to the decay rate to that state; and the factor $\Gamma(Z^0 \to e^+ + e^-)$ occurs because it can be related by time-reversal invariance to the rate for the initial formation process (9.21).

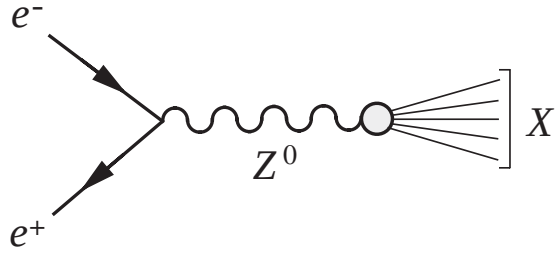


Figure 9.11 Formation of a Z^0 and its decay to an arbitrary state X.

We therefore see that, while the positions and widths of the peaks in (9.26) are determined by the Z^0 mass and total width Γ_Z, their heights are proportional to the products of branching ratios

$$B(Z^0 \to e^+ + e^-)B(Z^0 \to X) \equiv \frac{\Gamma(Z^0 \to e^+ + e^-)}{\Gamma_Z} \frac{\Gamma(Z^0 \to X)}{\Gamma_Z}. \tag{9.27}$$

Hence, by fitting the observed data, one obtains not only precise values

$$M_Z = 91.188 \pm 0.002\,\mathrm{GeV/c^2} \tag{9.28a}$$

and

$$\Gamma_Z = 2.495 \pm 0.002\,\mathrm{GeV} \tag{9.28b}$$

for the mass and total decay rate[7] but also for the various partial decay rates $\Gamma(Z^0 \to X)$. In this way the hadronic decay rate is found to be

$$\Gamma(Z^0 \to \mathrm{hadrons}) = 1.744 \pm 0.002\,\mathrm{GeV}, \tag{9.29a}$$

while the leptonic decay rates are all consistent with the value

$$\Gamma(Z^0 \to \ell^+ + \ell^-) = 0.0840 \pm 0.0009\,\mathrm{GeV}, \tag{9.29b}$$

independent of the identity of the lepton type $\ell = e, \mu, \tau$.[8]

The final states observed in the reactions (9.24) and (9.25) account for just over 80 % of all Z^0 decays, as can be seen by comparing the observed decay rates (9.29a)

[7] The values quoted in (9.28a) and (9.28b) are averages of several different determinations.

[8] This is exactly what is expected from the universality of leptonic interactions if the masses of the final state leptons can be neglected compared with the Z^0 mass. This is in fact a very good approximation (cf. the discussion of lepton decays in Section 2.2.2 and $W^\pm$ decays in Section 8.1.3).

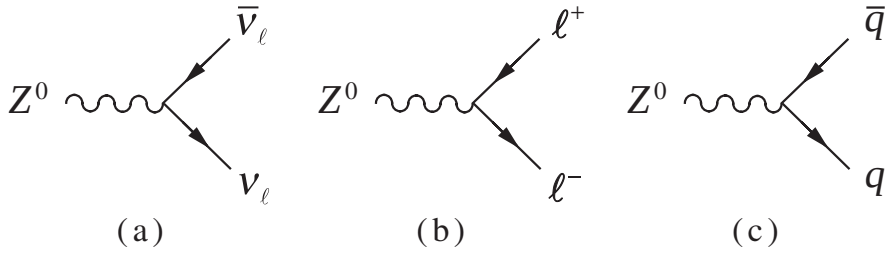


Figure 9.12 Feynman diagrams for the Z^0 decays (9.30a), (9.30b) and (9.30c), respectively.

and (9.29b) with the total decay rate (9.28b). The remaining decays are ascribed to final states containing only neutrinos. To see this, we note that in the unified theory, the only Z^0 decays that can arise from the basic vertices of Figures 9.3 and 9.4 are

$$Z^0 \to \nu_\ell + \bar{\nu}_\ell \quad (\ell = e, \mu, \tau), \tag{9.30a}$$

$$Z^0 \to \ell + \bar{\ell} \quad (\ell = e, \mu, \tau) \tag{9.30b}$$

and

$$Z^0 \to q + \bar{q}, \tag{9.30c}$$

corresponding to the Feynman diagrams of Figures 9.12(a), (b) and (c), respectively. In the last of these cases, the quark pair is not seen directly, but 'fragments' into two or more jets of hadrons that are observed in the final state, as shown in Figure 9.13. On adding up the various decays we thus obtain

$$\Gamma_Z = \Gamma(Z^0 \to \text{hadrons}) + 3\Gamma(Z^0 \to \ell^+\ell^-) + N_\nu \Gamma(Z^0 \to \nu_\ell \bar{\nu}_\ell) \tag{9.31}$$

for the total decay rate, where we have allowed for an arbitrary number N_ν of neutrino types in the sequence $\nu_e, \nu_\mu, \nu_\tau, \ldots$, but only three charged leptons e, μ and τ, for reasons to be discussed shortly. The total neutrino contribution is therefore

$$\begin{aligned} N_\nu \Gamma(Z^0 \to \nu_\ell \bar{\nu}_\ell) &= \Gamma_Z - \Gamma(Z^0 \to \text{ hadrons}) - 3\Gamma(Z^0 \to \ell^+\ell^-) \\ &= 0.499 \pm 0.004\,\text{GeV} \end{aligned} \tag{9.32}$$

where we have substituted the experimental values (9.28), (9.29a) and (9.29b). The decay rate to neutrino pairs cannot, of course, be measured directly, but can be calculated from the diagram of Figure 9.12(a) and is found to be[9]

[9] This result is given in, for example, p. 322 of Mandl and Shaw (1993). The decay rates $\Gamma(Z^0 \to \ell^+\ell^-)$ and $\Gamma(Z^0 \to \text{hadrons})$ can also be calculated from the Feynman diagrams of Figures 9.12(b) and 9.13, respectively, giving results in agreement with the measured values (9.29).

$$\Gamma(Z^0 \to \nu_\ell \bar{\nu}_\ell) = 0.166\,\text{GeV}. \tag{9.33}$$

This is compatible with the measured value (9.32) if, and only if, the number of neutrino types $N_\nu = 3$.

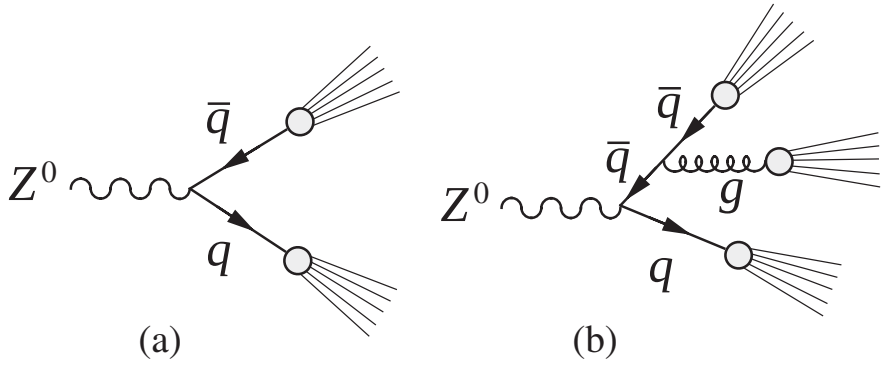


Figure 9.13 Decay of the Z^0 into (a) two and (b) three hadron jets.

This is a very significant result. Three generations of fermions (e^-, ν_e, u, d), (μ^-, ν_μ, c, s) and (τ^-, ν_τ, t, b) are known to exist, and each individually satisfies the anomaly condition (9.10); i.e.

$$Q_e + Q_{\nu_e} + 3(Q_u + Q_d) = 0,$$

with similar results for the second and third generations. Within the standard model, there are therefore no restrictions on adding more generations with the same properties. It was for a long time a matter of speculation whether such extra generations existed, since neutrinos are difficult to detect and additional charged leptons and quarks would have escaped detection if they were very heavy. In particular, any such charged leptons must be too heavy to contribute to Z^0 decays of the type (9.30b) or they would have been detected. However, if additional neutrinos existed in the sequence $\nu_e, \nu_\mu, \nu_\tau, \ldots$, with masses much less than M_Z, then the Z^0 would decay to them with a rate given by (9.33) according to the unified electroweak theory. Thus, in the framework of this theory, the result (9.32) restricts the number of lepton generations to three if we assume that neutrino masses are not very large. The same result can be illustrated graphically by simply comparing the predictions of the unified theory directly with the data of Figure 9.10 for the two cases $N_\nu = 3$ and 4. From (9.31) and (9.33) we see that the total width Γ_Z will be increased by 0.166 GeV if $N_\nu = 4$, and the product of branching ratios (9.27) will be correspondingly reduced. Consequently, the peaks in the cross-sections (9.26) are predicted to be broader and lower if $N_\nu = 4$. This is in clear disagreement with the data, as shown in Figure 9.10.

We thus see that, within the standard model, of which the unified electroweak theory is a part, there can only be three generations of leptons and quarks if neutrinos

are assumed to be light compared with the Z^0 mass; if this is so, no fundamental fermion remains to be discovered.

9.2 GAUGE INVARIANCE AND THE HIGGS BOSON

Gauge invariance is a fundamental symmetry associated with theories in which the force carriers are spin-1 bosons. It plays an important role in the unified electroweak theory, where it is needed to ensure the cancellation of the divergencies that occur in individual Feynman diagrams, as mentioned in Section 9.1. In addition, because the $W^{\pm}$ and Z^0 bosons have nonzero masses, it leads to the prediction of a new spin-0 boson – the *Higgs boson* – which has not yet been detected. This section is devoted to a qualitative introduction to gauge invariance and its consequences for weak interactions and the unified theory. A fuller account, including mathematical details, is given in Appendix D.

There are in fact different forms of gauge invariance, corresponding to the different interactions of particle physics, as we shall see. Their common feature is that the parameters of the corresponding symmetry transformations are allowed to vary with position ($\mathbf{r}$,t) in space and time.[10] For example, in QED one takes it as a fundamental assumption that the theory must be invariant under gauge transformations of the form

$$\psi(\mathbf{r},t) \rightarrow \psi'(\mathbf{r},t) = \exp[-iqf(\mathbf{r},t)]\psi(\mathbf{r},t), \tag{9.34}$$

where $\psi(\mathbf{r},t)$ is the wavefunction of a particle of charge[11] $\sqrt{\varepsilon_0}q$ and $f(\mathbf{r},t)$ is an arbitrary continuous function. In other words, one requires that if $\psi(\mathbf{r},t)$ is a solution of the equation of motion, $\psi'(\mathbf{r},t)$ must also be a solution. This condition is not satisfied by a free or noninteracting particle. In this case, the equation of motion is the free particle Schrödinger equation

$$i\frac{\partial\psi(\mathbf{r},t)}{\partial t} = -\frac{1}{2m}\nabla^2\psi(\mathbf{r},t) \tag{9.35}$$

in the nonrelativistic limit, and it is easy to see that if $\psi(\mathbf{r},t)$ is a solution, $\psi'(\mathbf{r},t)$ is not a solution for arbitrary $f(\mathbf{r},t)$. Hence gauge invariance requires the existence of an interaction. Furthermore, if one adds the minimal interaction required to make (9.35) invariant under (9.34), it can be shown that one arrives at the correct equation of motion for a charge particle interacting with an electromagnetic field. This approach, in which the form of the interaction is inferred by adding the minimal interaction terms needed to make the equation of motion gauge-invariant, is called the *principle of minimal gauge invariance*, or the *gauge principle* for short. It is discussed in some detail for the interested reader in Appendix D. Here we shall confine ourselves to a

[10] This is in contrast, for example, to our discussion of translational invariance in Section 5.1. In that case the parameter **a** that characterized the transformation (5.1) was assumed to be independent of position and time. Hence translational invariance is not an example of gauge invariance.

[11] We remind the reader that we are working in units that correspond to standard SI units when factors of $\hbar$ and c are restored. It is more usual to discuss this topic in Heavyside–Lorentz, or rationalized Gaussian units, in which case $\varepsilon_0 = 1$ and q is just the electric charge.

qualitative discussion of its application to weak interactions, omitting all technical details. In doing so, we shall restrict ourselves for simplicity to the first generation of leptons (e^-, ν_e) since the extension to other leptons, and to quarks using lepton–quark symmetry, is straightforward and involves no new issues of principle.

9.2.1 Unification and the gauge principle

The gauge principle described above begs the question: what is the appropriate form of the gauge transformation? In particular, what are the appropriate forms of the gauge transformations for the weak interactions? In QED, one considers phase transformations (9.34) that change an electron state to a new electron state, symbolically $e^- \to e^-$. The gauge principle then leads to interactions $e^- \to e^-\gamma$ in which a gauge boson (the photon) is emitted or absorbed. More generally, one can define gauge transformations in which not only the phase of the wavefunction but also the nature of the particle changes. Specifically, one can define a set of gauge transformations in which electrons and electron neutrinos can transform into themselves, or each other;[12] i.e. they incorporate the transformations

$$e^- \to \nu_e, \quad \nu_e \to e^-, \quad e^- \to e^-, \quad \nu_e \to \nu_e$$

and lead, via the gauge principle, to interactions

$$e^- \to \nu_e W^-, \quad \nu_e \to e^- W^+, \quad e^- \to e^- W^0, \quad \nu_e \to \nu_e W^0$$

in which gauge bosons of the appropriate charge are emitted or absorbed. As implied by our notation, the charged bosons can be identified with the observed $W^\pm$ bosons, leading to charged current processes of the type observed. However, we also have a W^0 boson, leading to neutral current processes of the same form and strength as the charge current interactions. This is not what is observed experimentally.

The resolution of this problem lies in unification of the weak interactions with electromagnetism. If electromagnetism were introduced directly, as described above following (9.34), we would simply add the photon to the list of gauge bosons W^+, W^-, W^0 and the problem of the W^0 would remain unchanged. Instead, we regard both the photon γ and the neutral boson Z^0 as mixtures of the W^0 with another neutral boson B^0; i.e. we write

$$\gamma = B^0 \cos\theta_W + W^0 \sin\theta_W \tag{9.36a}$$

and

$$Z^0 = -B^0 \sin\theta_W + W^0 \cos\theta_W, \tag{9.36b}$$

where θ_W is the weak mixing angle and the particle B^0 is such that the combination (9.36a) has all the properties of the observed photon. This can be achieved by requiring gauge invariance under the transformations

[12] The precise form of these transformations is discussed in Appendix D, Section D.7.

$$\psi_\ell(\mathbf{r},t) \to \psi'_\ell(\mathbf{r},t) = \exp[-ig_Z y_\ell f(\mathbf{r},t)]\psi_\ell(\mathbf{r},t) \quad (\ell = e, \nu_e), \tag{9.37}$$

where g_Z is the weak coupling constant introduced in Equation (9.8) and $y_\ell \equiv y(e^-)$, $y(\nu_e)$ are constants to be determined.

The transformations (9.37) are similar to (9.34), so the gauge principle leads to vertices

$$e^- \to e^- B^0, \quad \nu_e \to \nu_e B^0,$$

analogous to $e^- \to e^-\gamma$ but with the charge replaced by the couplings $g_Z y(e^-)$, $g_Z y(\nu_e)$. An explicit calculation[13] then shows that these can be chosen so that the combination (9.36a) has exactly the couplings of a photon, provided that the condition

$$\frac{e}{2(2\varepsilon_0)^{1/2}} = g_W \sin\theta_W = g_Z \cos\theta_W$$

is satisfied. This is just the unification condition (9.8).

In this way, one arrives at a unified theory of electroweak interactions that is gauge-invariant, as required. The assumed form of gauge invariance is complicated and is chosen to ensure that the resulting electromagnetic and charged current interactions have the form dictated by experiment. Once that has been done, the properties of the weak neutral current interactions arising from the exchange of the Z^0 boson (9.36b) are predicted in detail. The precise agreement between these predictions and experiment is one of the greatest triumphs of the unified theory, as we have seen. However, there is one remaining problem, to which we now turn.

9.2.2 Particle masses and the Higgs field

The *Higgs boson* is a neutral spin-0 particle whose existence is predicted by the standard model, but which has not yet been observed. It is required because gauge invariance can be shown[14] to imply that the spin-1 gauge bosons have zero masses. This is acceptable for QED and QCD, where the gauge bosons are the photons and the gluons that do indeed have zero masses. However, the $W^\pm$ and Z^0 bosons are very heavy and not massless, as they would be if gauge invariance were exact. This problem is overcome by assuming that the various particles interact with a new type of scalar field, called the *Higgs field*. The interactions of the Higgs field with the gauge bosons are gauge invariant, but it differs from other fields in its behaviour in the so-called *vacuum state*, which contains no particles of any kind. Other fields, such as the electromagnetic field, are assumed to be zero in the vacuum state, as one would naively expect. However, the Higgs field has a nonzero value η_0 in the vacuum, and this value is not invariant under a gauge transformation. Because of this, the theory is no longer gauge-invariant and the gauge bosons are no longer required to have zero mass. This form of symmetry breaking, in which the gauge invariance

[13] This calculation is done in Appendix D, Section D.7.3.
[14] This is done explicitly for the electromagnetic case in Appendix D, Section D.3.

of the interaction (as opposed to the gauge invariance of the vacuum) remains exact, is called *spontaneously symmetry breaking*.

Spontaneous symmetry breaking of this type is known in other branches of physics. It occurs whenever the vacuum state, defined as the state of lowest energy, is not unique. A familiar example is ferromagnetism. In a ferromagnetic material, the forces that couple the electronic spins are rotationally invariant and at high temperatures the net magnetization **M** is zero. However, below the Curie temperatures the spins are aligned in some definite direction in the ground state, resulting in a magnetization that breaks the rotational invariance. However, **M** could equally well point in any other direction and the rotational invariance of the interaction manifests itself in that all the properties, other than the direction of **M**, are independent of the actual direction.

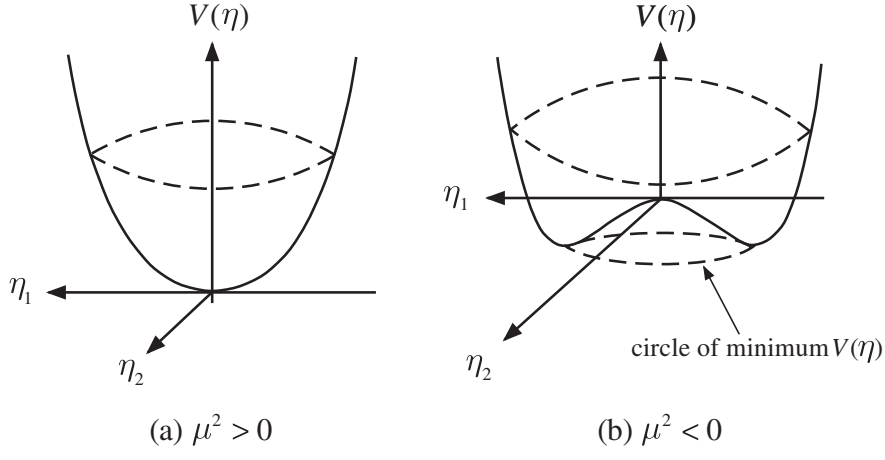


Figure 9.14 The potential energy density $V(\eta)$, as given by Equation (9.39), for $\lambda > 0$.

To see how such effects can occur in particle physics, we consider the simple case of a complex scalar field $\eta(\mathbf{r},t)$, which we write in terms of two real fields $\eta_1(\mathbf{r},t)$ and $\eta_2(\mathbf{r},t)$, i.e.

$$\eta(\mathbf{r},t) = \eta_1(\mathbf{r},t) + i\eta_2(\mathbf{r},t). \tag{9.38}$$

Let us assume that the potential energy density of $\eta(\mathbf{r},t)$ is given by

$$V(\eta) = \mu^2 |\eta(\mathbf{r},t)|^2 + \lambda |\eta(\mathbf{r},t)|^4 . \tag{9.39}$$

Here λ and μ^2 are real parameters, and the interactions are invariant under the transformation

$$\eta(\mathbf{r},t) \to \eta(\mathbf{r},t)e^{i\beta}, \tag{9.40}$$

where β is an arbitrary phase parameter. The vacuum state can be identified by minimizing the potential energy density, since the kinetic energy density is either positive or zero. There are two cases to consider, and in both we require $\lambda > 0$ so that the potential energy density is bounded from below:

(a) $\mu^2 > 0$. In this case, both terms in $V(\eta)$ are positive definite and the corresponding potential energy density surface $V(\eta)$ is sketched as a function of $\eta_1(\mathbf{r}, t)$ and $\eta_2(\mathbf{r}, t)$ in Figure 9.14(a). As can be seen, $V(\eta)$ has a unique minimum at $\eta(\mathbf{r}, t) = 0$ and spontaneous symmetry breaking does not occur.

(b) $\mu^2 < 0$. The potential energy surface for this case is shown in Figure 9.14(b). In this case, there is a circle of absolute minima at

$$\eta(\mathbf{r}, t) = \eta_0 = \left(-\frac{\mu^2}{2\lambda}\right)^{1/2} e^{i\theta}, \quad 0 \leq \theta \leq 2\pi, \tag{9.41}$$

where the phase angle θ specifies a direction in the complex η-plane. This arbitrariness in the direction of θ is analogous to that in the direction of the magnetization $\mathbf{M}$ of a ferromagnet. By analogy with the latter case, spontaneous symmetry breaking corresponds to taking a particular value of θ to represent the ground state. The actual value chosen is not significant, because of the invariance of the interactions under a phase transformation (9.40), and it is conventional to choose $\theta = 0$, so that the vacuum state corresponds to

$$\eta_0 = \left(-\frac{\mu^2}{2\lambda}\right)^{1/2}. \tag{9.42}$$

This is obviously not invariant under the transformation (9.40) and the symmetry is spontaneously broken.

To see how this affects gauge theories, consider the introduction of electromagnetic interactions into the above model in such a way that the equations of motion remain invariant under a gauge transformation of the standard form (9.34), i.e.

$$\eta(\mathbf{r}, t) \to \eta'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)]\eta(\mathbf{r}, t). \tag{9.43}$$

Since the electromagnetic field vanishes in the vacuum state, this has no effect on the above discussion of the scalar field value in the vacuum. Again, there are two cases to consider:

(a) $\mu^2 > 0$. In this case, we again have $\eta(\mathbf{r}, t) = 0$ in the vacuum, so that there is no spontaneous symmetry breaking. Gauge invariance remains exact and the photon must be massless, as observed in nature.

(b) $\mu^2 < 0$. In this case, spontaneous symmetry breaking does occur and we can again choose to represent the vacuum field by (9.42). This clearly is not invariant under

the gauge transformaton (9.43), so that gauge invariance is broken and the photon is no longer required to be massless. Furthermore, as first shown by Higgs, the photon automatically becomes massive[15] due to its interaction with the vacuum field η_0. This mechanism, whereby a gauge boson acquires mass as a result of its interaction with a nonvanishing vacuum field, is called the *Higgs mechanism*. It is in some ways analogous to the way a conduction electron in a semiconductor acquires an effective mass different from its mass *in vacuo* as a result of its interactions with the crystal lattice.

Of course the unified electroweak theory is more complicated than the above simple model. We will omit all details here and simply note that applying the Higgs mechanism to the unified theory leads to three main consequences.

The first is that the $W^{\pm}$ and Z^0 bosons acquire masses in the ratio

$$M_W/M_Z = \cos\theta_W. \tag{9.44}$$

These masses arise from the interactions of the gauge fields with the nonzero vacuum expectation value of the Higgs field.

The second is that there are electrically neutral quanta H^0 associated with the Higgs field, called Higgs bosons, in the same way that there are quanta associated with the electromagnetic field, i.e. photons.

The third is that the Higgs field throws light on the origin of quark and lepton masses. In the absence of a Higgs field, gauge invariance requires that the masses of spin-$\frac{1}{2}$ fermions with parity-violating interactions are zero.[16] Parity is conserved in strong and electromagnetic interactions, but is violated in weak interactions, so that quarks and leptons would be massless in this case. However, interactions with the Higgs field can generate fermion masses arising from the nonzero vacuum expectation value η_0 of the Higgs field, as well as interactions with the Higgs bosons. These interactions are of a form shown in Figure 9.15, with a dimensionless coupling constant g_{Hff} related to the fermion mass m_f by

$$g_{Hff} = \sqrt{2}\, g_W \left(\frac{m_f}{M_W}\right). \tag{9.45}$$

This theory of fermion masses – that they are generated by interactions with the Higgs field – does not at present make any predictions for their values m_f. However, it can be tested in the future by measuring the Higgs boson couplings g_{Hff} and verifying the predictions (9.45).

---

[15] This is discussed briefly in Section D.3 of Appendix D, and more completely in Sections 13.1 and 13.2 of Mandl and Shaw (1993).

[16] The reason for this is discussed in Appendix D.7.4.

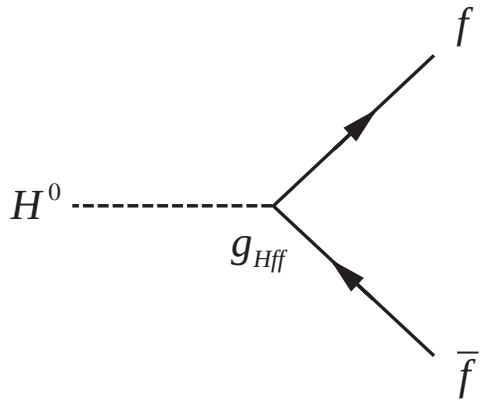


Figure 9.15 The basic vertices for Higgs boson–fermion interactions. The fermion f can be any quark, charged lepton or neutrino.

9.2.3 Higgs boson decays

A problem in designing suitable experiments to detect the Higgs boson is that its mass is not predicted by the theory. However, its couplings to other particles are predicted, and are essentially proportional to the mass of the particle to which it couples. The Higgs boson therefore couples very weakly to light particles like neutrinos, electrons, muons and u, d and s quarks, and much more strongly to heavy particles like $W^{\pm}$ and Z^0 bosons and t quarks.

This is reflected in the decay properties of the Higgs boson, which are predicted to decay predominantly to the heaviest particle–antiparticle pairs allowed by energy conservation. The resulting decay pattern for Higgs masses above $50\,\mathrm{GeV/c^2}$ are summarized in Figures 9.16 and Figures 9.17. The main features arise as follows.

For $M_H < 2M_W$, the heaviest particle available is the b quark. The dominant decay is predicted to be

$$H \to b + \overline{b}, \tag{9.46}$$

where the b quarks would manifest themselves as jets of hadrons. Furthermore, if we neglect the b-quark mass compared to the Higgs boson mass, then since the coupling g_{Hff} is dimensionless, a simple dimensional estimate gives

$$\Gamma(H \to b\overline{b}) = O(g^2_{Hff} M_H) = O(10^{-4} M_H), \tag{9.47}$$

where we have used (9.45) and (8.9) to evaluate g_{Hff}. In other words, the predicted Higgs boson decay width is quite small in this region, and it is much longer lived than the Z^0, for example. However, as the Higgs mass increases beyond the W^+W^- threshold, the decays

$$H^0 \to W^+ + W^- \quad (M_H > 2M_W) \tag{9.48a}$$

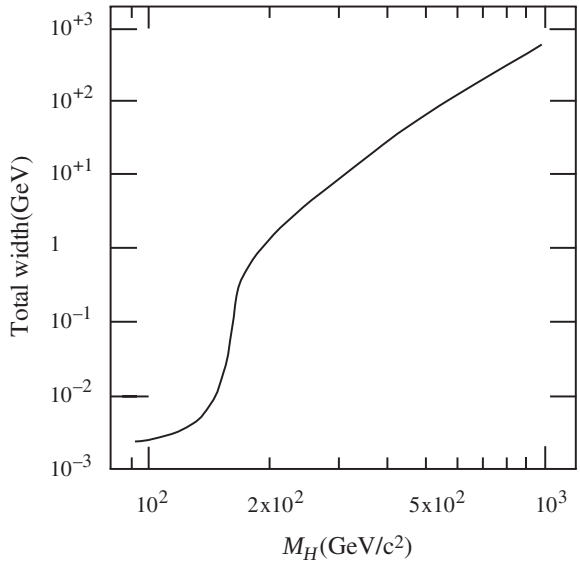


Figure 9.16 Width of the Higgs boson as a function of its mass. (With kind permission from Springer Science and Business Media, Kunszt *et al*., *Zeitschrift für Physik C*, **74**, 1997, 479)

and

$$H^0 \to Z^0 + Z^0 \quad (M_H > 2M_Z) \tag{9.48b}$$

are allowed, as shown in Figure 9.18. Furthermore, because the Higgs couplings are roughly proportional to the mass of the particle to which they couple, these decays completely dominate the decay mode (9.46). This leads to a rapid increase in the total decay width as the Higgs mass increases above the W^+W^- threshold energy, as shown in Figure 9.16. Finally, as the Higgs mass increases beyond $360\,\mathrm{GeV/c^2}$, the decay to $t\bar{t}$ pairs also becomes important.

So far we have only considered the dominant decay modes of the Higgs boson. There are also many other relatively rare decay modes predicted, as shown in Figure 9.17. For example, the relatively rare decay mode

$$H^0 \to \gamma + \gamma \tag{9.49}$$

can proceed by the mechanisms of Figure 9.19 with a branching ratio of order 10^{-3}. Nonetheless, it may well be important in detecting the Higgs boson if its mass is between 110 and $150\,\mathrm{GeV/c^2}$, as we shall see below. Another important class of rare decay modes is

$$H^0 \to Z^0 + f + \bar{f}, \tag{9.50}$$

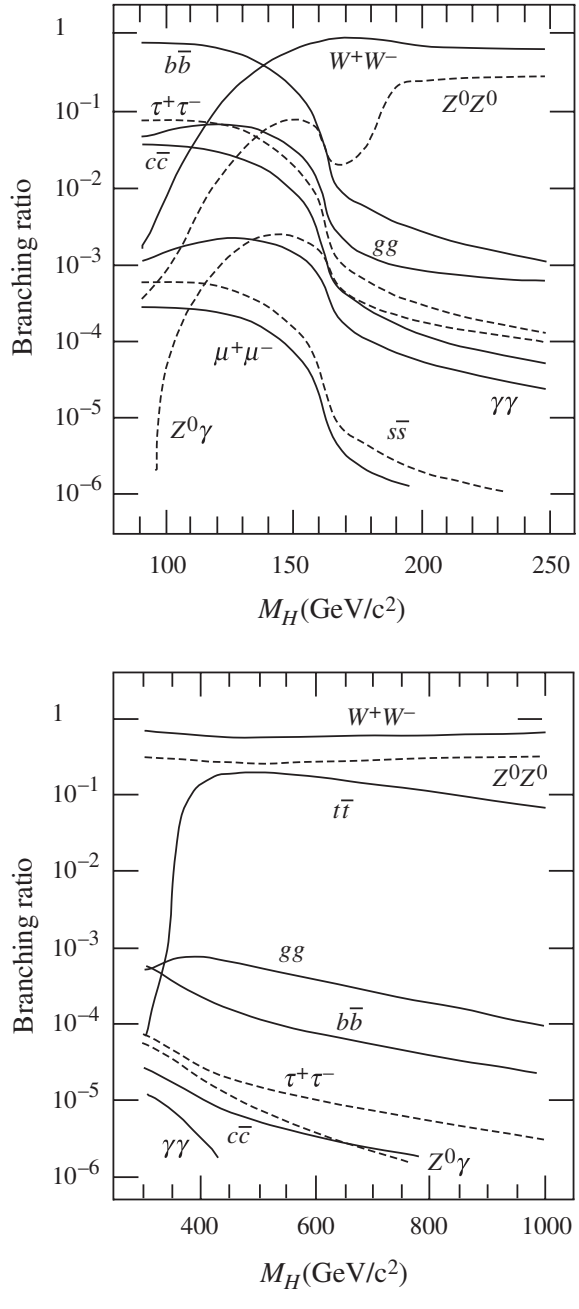


Figure 9.17 The branching ratios of a standard-model Higgs boson H^0 for all decays with branching ratios greater than 10^{-6}. (With kind permission from Springer Science and Business Media, Kunszt *et al*., *Zeitschrift für Physik C*, **74**, 1997, 479)

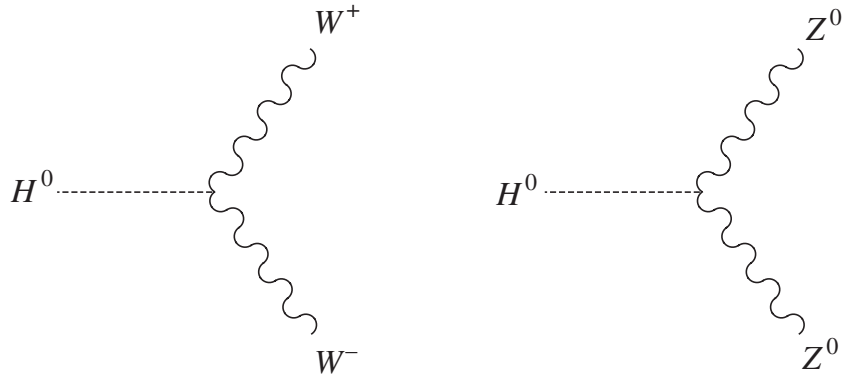


Figure 9.18 Decays of the Higgs boson to (a) W^+W^- and (b) Z^0Z^0 pairs.

where f can be any quark or lepton. These decays proceed by the mechanism of Figure 9.20, involving a real Z^0 and a virtual Z^0 as an intermediate state, and for this reason they are often classed as Z^0Z^0 decays. Thus the nonzero branching ratio for Z^0Z^0 decays shown in Figure 9.17 for $M_H < 2M_Z$ really refers to the decays (9.50).

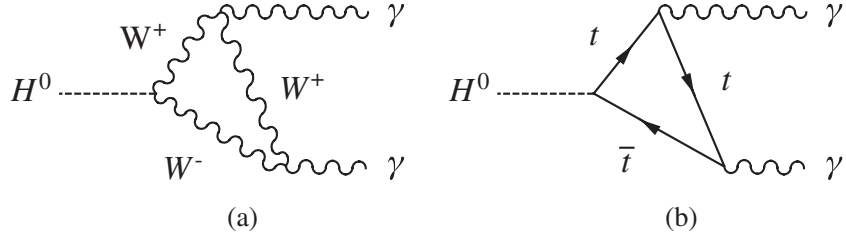


Figure 9.19 The dominant mechanisms for the decay $H^0 \to \gamma + \gamma$.

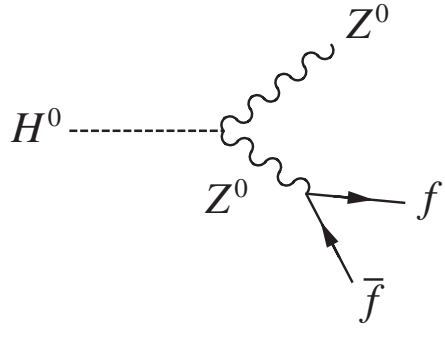


Figure 9.20 Mechanism of the rare H^0 decays (9.50).

Finally, we note that, just as we can predict the decay modes of the Higgs boson for any given mass, we can also calculate the contribution to any given

electroweak process from higher-order diagrams in which Higgs bosons are emitted and absorbed. These diagrams, together with other higher-order diagrams, give very small corrections to the leading order processes that we have considered throughout this chapter. Nonetheless, these corrections can be measured, and the requirement that the measured and calculated corrections agree lead to the restricted range

$$28 < M_H < 144\,\mathrm{GeV/c^2} \quad (95\%\,\text{confidence level}). \tag{9.51}$$

In fact current experiment already rules out most of this range, as we shall immediately see.

9.2.4 The search for the Higgs boson

The existence of the Higgs boson is the most important prediction of the standard model that has not been verified by experiment, and searches for it are a high priority at most high-energy accelerators. However, attempts to produce Higgs bosons are made more difficult by the need to first produce the very heavy particles to which they couple.

The failure to observe Higgs bosons in present experiments leads to limits on their mass. The best results to date come from the large electron–positron (LEPII) accelerator at CERN. This machine had a maximum energy of 208 GeV at the time of its closure in November 2002. This energy is high enough to produce Higgs bosons with masses up to almost $120\,\mathrm{GeV/c^2}$ in the reaction

$$e^+ + e^- \rightarrow H^0 + Z^0, \tag{9.52}$$

which is expected to occur by the dominant mechanism of Figure 9.21. Attempts were made to detect Higgs bosons by their decays to $b\bar{b}$ pairs, where the quarks would be observed as jets containing short-lived hadrons with nonzero bottom. The failure to observe such events implied a lower limit

$$M_H > 113.5\,\mathrm{GeV/c^2} \tag{9.53}$$

on the mass of the Higgs boson, assuming that it exists. Tantalizingly, some evidence was obtained for the existence of a Higgs boson with a mass of $115\,\mathrm{GeV/c^2}$, which is very close to the upper limit of masses that were accessible by LEPII. Unfortunately,

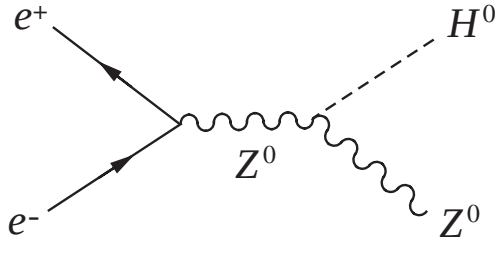


Figure 9.21 Dominant mechanism for Higgs boson production in e^+e^- annihilation.

while this signal was statistically more likely to be a genuine result rather than a statistical fluctuation, the latter could not be ruled out.

Higgs boson searches have also been carried out at the Tevatron $p\bar{p}$ collider, mentioned in Chapter 4. At the Tevatron, the searches concentrate on the associated production of a Higgs boson, i.e. $p + \bar{p} \rightarrow V + H + X$, where X is any hadron state consistent with conservation of the appropriate quantum numbers and where the vector boson $V \equiv W^{\pm}, Z^0$ decays into charged leptons and/or neutrinos. For masses below about 130 GeV/c^2, the decay $H^0 \rightarrow b\bar{b}$ provides the most sensitive channel, as was used at LEPII. With presently available data, the sensitivity of the two experiments at the Tevatron, CDF and DØ is still rather limited, but with increasing sample sizes the sensitivity may eventually exceed the LEP range and so it is possible that the Higgs boson will first be seen at the Tevatron. However, the greatest probability is that if the Higgs boson exists it will be found first at the LHC (also discussed in Chapter 4), where two of the experiments, ATLAS and CMS, have been optimized to search for Higgs bosons with mass up to 1 TeV/c^2.

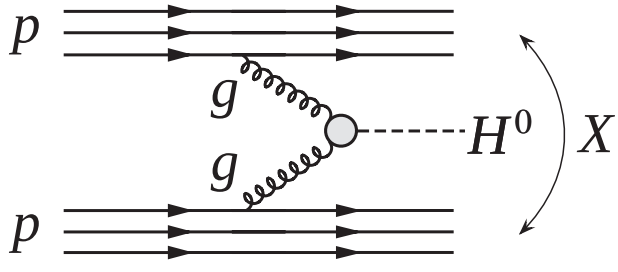


Figure 9.22 Dominant gluon fusion production mechanism for the reaction $p + p \rightarrow H^0 + X$ at the LHC.

At the LHC the reaction of interest will be

$$p + p \rightarrow H^0 + X, \tag{9.54}$$

where X is a hadron state consistent with conservation of the appropriate quantum numbers. The dominant Higgs production mechanism for this reaction at the LHC is 'gluon fusion' at all Higgs boson masses. This is shown in Figure 9.22, where the ggH^0 vertex is dominated by a loop involving top quarks because of the strong coupling of H^0 to the very heavy $t\bar{t}$ state, as shown in Figure 9.23(a). Other production processes are shown in Figure 9.23(b), (c) and (d). These are also of interest because of the different experimental signatures they provide for identifying the Higgs boson. In these diagrams, for simplicity the spectator particles are not shown. The contributions of these processes to the production cross-sections at the LHC, at an energy of 14 TeV, are shown in Figure 9.24.

From Equations (9.50) and (9.52), we see that the most likely mass range for a standard model Higgs boson is 113.5 GeV/c^2 $< M_H <$ 144 GeV/c^2. Referring to Figure 9.17, we see that the obvious mode to study would be the dominant $H^0 \rightarrow b\bar{b}$ channel with the quarks fragmenting to jets, as used in the LEP experiment. Unfortunately, it is very difficult to distinguish these jets from those produced by other means. Because of this

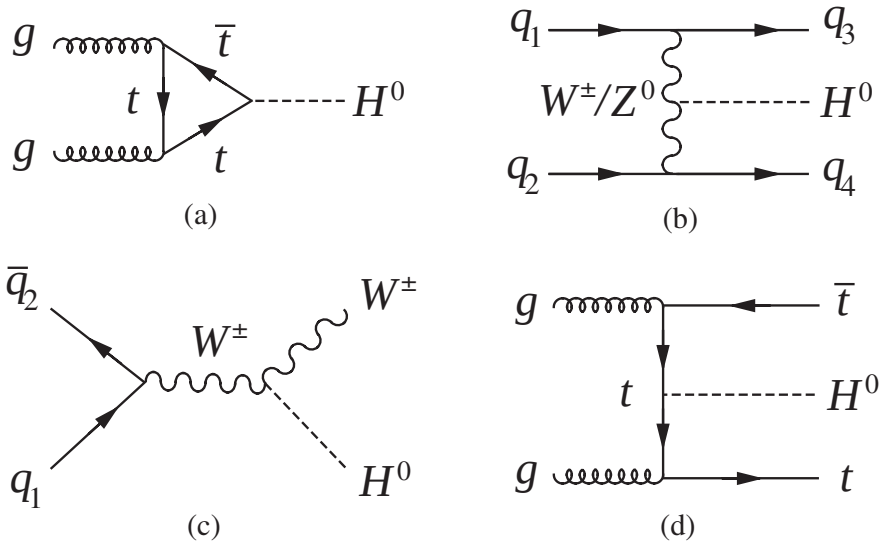


Figure 9.23 Production mechanisms for a standard model Higgs boson at the LHC.

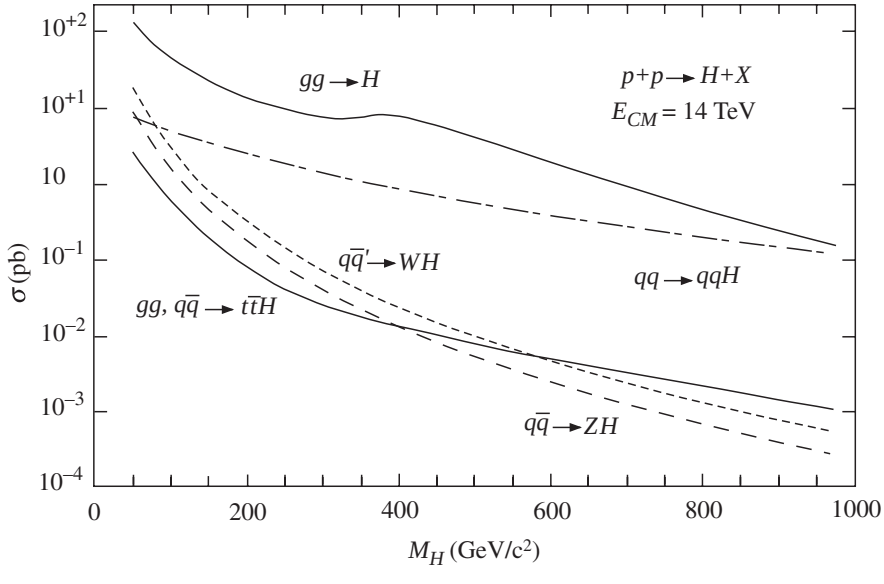


Figure 9.24 Production cross-sections of the standard model Higgs boson at the LHC at $E_{CM} = 14$ TeV. (With kind permission from Springer Science and Business Media, Kunszt *et al.*, *Zeitschrift für Physik C*, **74**, 1997, 479)

it is likely that other decay modes will be more useful, even though they have much smaller branching ratios. For example, the decay $H^0 \to \gamma\gamma$ mentioned above, with a branching ratio of order 10^{-3}, would have a signature of two isolated electromagnetic clusters and much lower backgrounds than the $H^0 \to b\bar{b}$ channel.

Above the W^+W^- threshold, the decays are almost exclusively to the W^+W^- and Z^0Z^0 channels, except in the mass range near the $t\bar{t}$ decay threshold. For masses in the range $2M_Z < M_H < 650\,\mathrm{GeV/c^2}$, detection of the Higgs boson will be straightforward via the decays

$$H^0 \to Z^0 + Z^0 \to \ell^+ + \ell^- + \ell^+ + \ell^- \quad (\ell = e, \mu), \tag{9.55}$$

where the $\ell^+\ell^-$ pairs have the same invariant mass as the Z^0. A computer simulation of an event of this type in the CMS detector is shown in coloured Plate 4. A disadvantage as one approaches the end of this range is the increased width of the H^0 and the reduced production rate.

If, against expectations, $M_H > 650\,\mathrm{GeV/c^2}$, production rates drop dramatically and the decay is dominated by vector boson channels. To obtain a significant counting rate, one will need to search for events in which one vector boson decays to jets, e.g. $H^0 \to W^+W^- \to \ell\nu jj$, where j is a jet. This has a branching ratio at least 50 times greater than the decay mode $H^0 \to Z^0Z^0 \to 4$ leptons. For example, fewer than 200 Higgs particles with $M_H = 700\,\mathrm{GeV/c^2}$ would decay in the $H^0 \to Z^0Z^0 \to \ell^+\ell^-\ell^+\ell^-$ channel in a year at high luminosity, and there would still be considerable backgrounds to contend with.

Finally, we note that in the above discussion we have considered only the simplest case of a single, electrically neutral, Higgs boson, as assumed in the standard model. However, some interesting extensions of the standard model, which we discuss in Chapter 11, require more than one Higgs particle, including electrically charged varieties. Whatever the truth, it is clear that experimental investigation of the Higgs sector will play a central role in future developments in particle physics.

PROBLEMS 9

9.1 Show that the vertices of Figure 9.3 are the only possible vertices abZ^0 allowed by charge and lepton number conservation, where a and b can be any lepton or antilepton.

9.2 Which of the following processes are allowed in electromagnetic interactions and which are allowed in weak interactions via the exchange of a single $W^\pm$ or Z^0?

(a) $K^+ \to \pi^0 + e^+ + \nu_e$
(b) $K^+ \to \pi^+ + e^+ + e^-$
(c) $\Sigma^0 \to \Lambda^0 + e^+ + e^-$
(d) $\Sigma^0 \to n + \nu_e + \bar{\nu}_e$

9.3 The reaction (9.6b) is forbidden to occur via lowest-order weak interactions like Figure 9.5(b). However, it can proceed by higher-order diagrams involving the exchange of two or more bosons. Draw examples of such diagrams. Make a simple dimensional estimate of the ratio of decay rates (9.7).

9.4 Draw the dominant Feynman diagrams for the reaction $e^+ + e^- \to \nu_e + \bar{\nu}_e$. Estimate the magnitude and energy dependence of the corresponding contributions to the total cross-section for centre-of-mass beam energies $m_e \ll E \ll M_W, M_Z$.

9.5 Identify the dominant decay mode of a Higgs boson of mass $M_H = M_Z$ and estimate its branching ratios into lepton pairs.

9.6 In our discussion of the reaction $e^+ + e^- \to \mu^+ + \mu^-$, we completely neglected the Higgs exchange diagram of Figure 9.25 compared with the dominant diagrams of Figures 9.2(a) and (b). Justify this approximation for the case $M_H = M_Z$ by using the results of the

previous question to estimate the total cross-section at $E_{CM} = M_Z$, which would arise from Figure 9.25 alone, and by comparing with estimates of the cross-sections, which would arise from Figure 9.2(a) alone and Figure 9.2(b) alone.

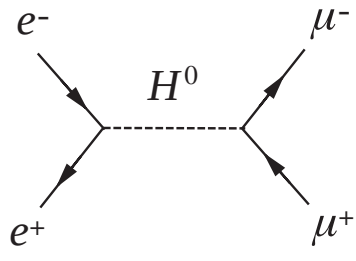


Figure 9.25 Feynman diagram for the reaction $e^+e^- \rightarrow \mu^+\mu^-$ via the exchange of a single Higgs boson.

10

Discrete Symmetries: *C, P, CP* and *CPT*

In Chapter 5 we introduced charge conjugation (C) and parity (P) as exact symmetries of the strong and electromagnetic interactions. Here we discuss them in the context of the weak interaction, where they are not conserved. In the course of this, two interconnected themes will emerge. The first is that these effects have their origin in the spin dependence of weak interactions, which is quite remarkable, as we shall see. The second is that while parity violation and C-parity violation are large effects, there is a weaker combined symmetry, called CP invariance, which seems to be exactly conserved in the weak interaction of leptons. However, while extending this symmetry to the weak interactions of hadrons initially leads to some striking experimental successes, it also leads to experimental evidence for deviations from CP invariance. This topic is discussed in Section 10.2, while in Section 10.3 we briefly discuss the more fundamental symmetry of CPT invariance, which is believed to be respected in all interactions. Finally in Section 10.4, we discuss the important question of whether CP violation can be fully accounted for within the standard model, or whether some new interaction is required.

10.1 *P* VIOLATION, *C* VIOLATION AND *CP* CONSERVATION

The concept of parity was first introduced in the context of atomic physics by Wigner in 1927 and for a long time afterwards its conservation was believed to be a universal law of nature, like angular momentum conservation. The first indications

Particle Physics Third Edition B. R. Martin and G. Shaw

that this was not the case came from observations on the pionic decays of K mesons,[1] and these led Lee and Yang in 1956 to make a thorough study of previous experiments in which parity conservation had been assumed. They came to the startling conclusion that while there was strong evidence for parity conservation in strong and electromagnetic interactions, there was no evidence for its conservation in weak interactions, and they suggested experiments in which the latter assumption could be tested. This led directly to the classic demonstration of parity violation from a study of the β-decay of polarized cobalt-60 nuclei. This was done in 1957 by Wu and co-workers, who placed a sample of cobalt-60 inside a solenoid and cooled it to a temperature of 0.01 K. At such temperatures, the interaction of the magnetic moments of the nuclei with the magnetic field overcomes the tendency to thermal disorder, and the nuclear spins align parallel to the field direction. The polarized cobalt-60 nuclei produced in this way decay to an excited state of nickel-60 by the process

$$^{60}\mathrm{Co} \rightarrow {}^{60}\mathrm{Ni}^{*} + e^{-} + \bar{\nu}_{e}. \tag{10.1}$$

Parity violation is established by the observation of a 'forward–backward decay asymmetry', i.e. the fact that fewer electrons are emitted in the forward hemisphere than in the backward hemisphere with respect to the spins of the decaying nuclei. This follows because the parity transformation (5.30) reverses all particle momenta $\mathbf{p}$ while leaving their orbital angular momenta $\mathbf{r} \times \mathbf{p}$, and by analogy their spin angular momenta, unchanged. Hence, in the rest frame of the decaying nuclei its effect is to reverse the electron velocity while leaving the nuclear spins unchanged, as shown in Figure 10.1. Parity invariance would then require that the rates for the two processes of Figures 10.1(a) and (b) were equal, so that equal numbers of electrons would be

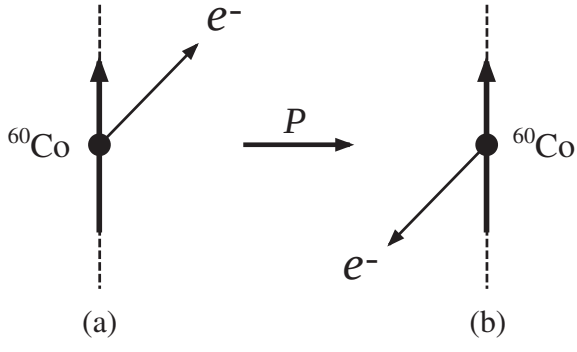


Figure 10.1 Effect of a parity transformation on ^{60}Co decay (10.1). The short thick arrows indicate the direction of the spin of the ^{60}Co nucleus, while the long arrows show the direction of the electron's momentum.

[1] Two particles, called at that time τ and θ, were observed to decay via the weak interaction to $\pi\pi$ and $\pi\pi\pi$ final states, respectively, which necessarily had different final state parities, as we shall show in Section 10.2.1. However, the τ and θ particles had properties, including the near equality of their masses, that strongly suggested that they were in fact the same particle, which we now know to be the K meson. Analysis of the '$\tau - \theta$ puzzle' suggested that parity was not conserved in the decays. See Problem 10.1.

emitted in the forward and backward hemispheres with respect to the nuclear spins, in contradiction to what was observed.

The discovery of parity violation was a watershed in the history of weak interactions because the effect is large, and the remarkable spin structure of the weak interactions cannot be understood unless it is taken into account. This will be illustrated in Sections 10.1.2 and 10.1.3 below. Firstly, however, we turn to the evidence for *C* violation in muon decays.

10.1.1 Muon decay symmetries

The existence of *C* violation and its close relationship to *P* violation in leptonic weak interactions are both conveniently illustrated by considering the angular distributions of the electrons and positrons emitted in the decays

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu \quad \text{and} \quad \mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu \tag{10.2}$$

of polarized muons.[2] In the rest frame of the decaying particle these were found to be of the form[3]

$$\Gamma_{\mu^\pm}(\cos\theta) = \frac{1}{2}\Gamma_\pm \left(1 - \frac{\xi_\pm}{3}\cos\theta\right), \tag{10.3}$$

where θ is the angle between the muon spin direction and the direction of the outgoing electron or positron, as shown in Figure 10.2(a). The quantities $\xi_\pm$ are called the

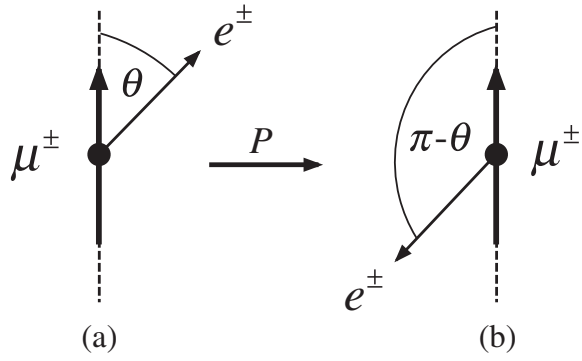


Figure 10.2 Effect of a parity transformation on the muon decays (10.2). The short thick arrows indicate the direction of the muon spin, while the long arrows show the direction of the electron's momentum.

[2] The muons produced in the decays of charged pions are naturally polarized, as we shall see in Section 10.1.3 below.

[3] An account of early experiments on the decay of positive muons, including experimental details, which are omitted below, is given in Chapter 10 of Trigg (1975), where the cobalt-60 experiment is also discussed.

asymmetry parameters and $\Gamma_\pm$ are equal to the total decay rates, or equivalently the inverse lifetimes, i.e.

$$\tau_\pm{}^{-1} \equiv \int_{-1}^{+1} \mathrm{d}\cos\theta \Gamma_{\mu^\pm}(\cos\theta) = \Gamma_\pm, \tag{10.4}$$

as may easily be checked by direct substitution.

We consider now the consequences of assuming parity and charge conjugation for these decays, starting with the latter as it is the simpler. Charge conjugation transforms all particles into their antiparticles, so that μ^- decay converts into μ^+ decay. *C* invariance then implies that the rates and angular distributions for these decays should be the same, i.e.

$$\Gamma_+ = \Gamma_- \quad \text{(C invariance)} \tag{10.5a}$$

and

$$\xi_+ = \xi_- \quad \text{(C invariance)}. \tag{10.5b}$$

The parity transformation (5.30) preserves the identity of the particles, but reverses their momenta, while leaving their spins unchanged. Its effect on muon decay is shown in Figure 10.2, where we see that it changes the angle θ to $\pi - \theta$ so that $\cos\theta$ changes sign. Hence *P* invariance implies

$$\Gamma_{\mu^\pm}(\cos\theta) = \Gamma_{\mu^\pm}(-\cos\theta) \quad \text{(P invariance)}, \tag{10.6a}$$

and substituting (10.3) into (10.6a) then gives the result that the asymmetry parameters vanish,

$$\xi_\pm = 0 \quad \text{(P invariance)}. \tag{10.6b}$$

Experimentally, the $\mu^\pm$ lifetimes are equal to very high precision, so that (10.5a) is satisfied, but the measured values of the asymmetry parameters are

$$\xi_- = -\xi_+ = 1.00 \pm 0.04, \tag{10.7}$$

which shows that both *C* invariance (10.5b) and *P* invariance (10.6b) are violated.

In view of these results, a question that arises is: why do the μ^+ and μ^- have the same lifetime if *C* invariance is violated? The answer lies in the principle of *CP* conservation, which states that the weak interaction is invariant under the combined operation *CP*, even though both *C* and *P* are separately violated. The *CP* operator transforms particles at rest to their corresponding antiparticles at rest, and *CP* invariance requires that these states should have identical properties. Thus, in particular, the masses of particles and antiparticles are predicted to be the same. More specifically, if we apply the *CP* operator to muon decays, the parity operator changes θ to $\pi - \theta$ as

before, while the *C* operator changes particles to antiparticles. Hence *CP* invariance alone implies that the condition (10.6a) obtained from *P* invariance is replaced by the weaker condition

$$\Gamma_{\mu^+}(\cos\theta) = \Gamma_{\mu^-}(-\cos\theta)\,. \tag{10.8}$$

Substituting (10.3) into this equation gives

$$\Gamma_+ = \Gamma_- \quad (CP \text{ invariance}), \tag{10.9a}$$

implying equal lifetimes by (10.4), as well as

$$\xi_+ = -\xi_- \quad (CP \text{ invariance}), \tag{10.9b}$$

in agreement with the experimental results (10.7). Thus *CP* invariance retains the symmetry between particles and antiparticles as observed by experiment, at least for μ decays. In fact *CP* invariance has been verified in a wide variety of experiments involving weak interactions, and the data are consistent with exact *CP* conservation in the weak interactions of leptons. The weak interactions of quarks do not respect *CP* invariance, although it is often a very good approximation, and the observed violations are so far confined to the decays of neutral kaons and *B* mesons. This will be discussed in Sections 10.2 to 10.4. For the moment, we will concentrate on the properties of the leptons and assume that *CP* conservation is exact.

10.1.2 Left-handed neutrinos and right-handed antineutrinos

We turn now to the spin structure of the weak interactions, which is closely related to the symmetry properties discussed above. As this spin structure takes its simplest form for neutrinos and antineutrinos, we will discuss these first. In doing so, it is convenient to use the so-called *helicity states*, in which the spin is quantized along the direction of motion of the particle, rather than along some arbitrarily chosen 'z direction'. For a spin-$\frac{1}{2}$ particle, the spin component along the direction of its motion can be either $+\frac{1}{2}$ or $-\frac{1}{2}$, as illustrated in Figure 10.3, corresponding to positive or negative helicity

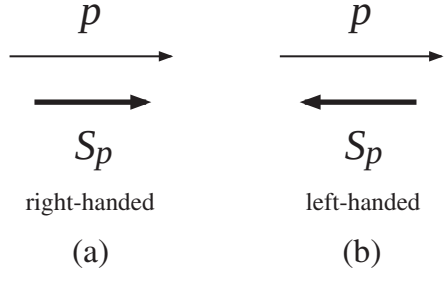


Figure 10.3 Helicity states of a spin-$\frac{1}{2}$ particle. The long thin arrows represent the momenta of the particles and the short thick arrows their spins.

respectively.[4] These states are called *right-handed* or *left-handed* respectively, since the spin direction corresponds to rotational motion in a right-handed or left-handed sense when viewed along the momentum direction. We will denote these states by a subscript R or L, so that ν_L means a left-handed neutrino, e_R^- a right-handed electron and so on.

The remarkable fact about neutrinos and antineutrinos, which only interact via the weak interaction, is that only left-handed neutrinos ν_L and right-handed antineutrinos $\bar{\nu}_R$ have been observed in nature. This obviously violates C invariance, which requires neutrinos and antineutrinos to have identical weak interactions. It also violates P invariance, which requires the states ν_L and ν_R to have identical weak interactions, since the parity operator reverses the momentum while leaving the spin unchanged, and so converts a left-handed neutrino into a right-handed neutrino. It is, however, compatible with CP invariance, since the CP operator converts a left-handed neutrino to a right-handed antineutrino, as illustrated in Figure 10.4.

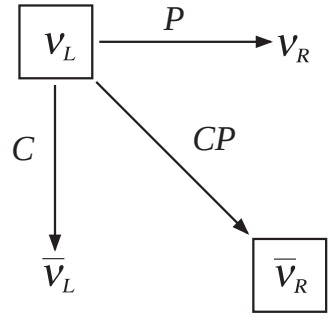


Figure 10.4 Effect of C, P and CP transformations on a left-handed neutrino ν_L. Only the states shown in boxes have been observed in nature.

The helicity of the neutrino was first measured in an ingenious experiment by M. Goldhaber and co-workers in 1958. They studied electron capture in europium-152, i.e.

$$e^- + {}^{152}\mathrm{Eu}(J=0) \rightarrow {}^{152}\mathrm{Sm}^*(J=1) + \nu_e, \tag{10.10}$$

where the spins of the nuclei are shown in brackets. The excited state of samarium that is formed decays to the ground state by γ emission

$$^{152}\mathrm{Sm}^*(J=1) \rightarrow {}^{152}\mathrm{Sm}\,(J=0) + \gamma, \tag{10.11}$$

and it is these γ rays that were detected in the experiment. In the reaction (10.10) the electrons are captured from the K shell and the initial state has zero momentum, so that the neutrino and the $^{152}\mathrm{Sm}^*$ nucleus recoil in opposite directions. The experiment

[4] More precisely, the helicity of a particle is defined as the spin projection along its direction of motion, divided by the magnitude of the spin.

selected events in which the photon was emitted in the direction of motion of the decaying $^{152}\mathrm{Sm}^{*}$ nucleus,[5] so that overall the observed reaction was

$$e^{-} + {}^{152}\mathrm{Eu}\,(J=0) \rightarrow {}^{152}\mathrm{Sm}\,(J=0) + \nu_e + \gamma, \tag{10.12}$$

where the three final state particles were collinear and the neutrino and photon emerged in opposite directions, as shown in Figure 10.5. The helicity of the neutrino can then be deduced from the measured helicity of the photon by applying angular momentum conservation about the event axis to the overall reaction (10.12). In doing this, no orbital angular momentum is involved, because the initial electron is captured from the K shell and the final state particles all move along the event axis. Hence the spin components of the neutrino and photon, which can be $\pm\frac{1}{2}$ and ± 1, respectively,[6] must add to give the spin component of the initial electron, which can be $\pm\frac{1}{2}$. This gives two possible spin configurations, as shown in Figures 10.5(a) and (b), and in each case the helicities of the photon and neutrino are the same. In the actual experiment, the polarization of the photons was determined by studying their absorption in magnetized iron,[7] and the results obtained were consistent with the occurrence of left-handed neutrinos only, corresponding to Figure 10.5(a).

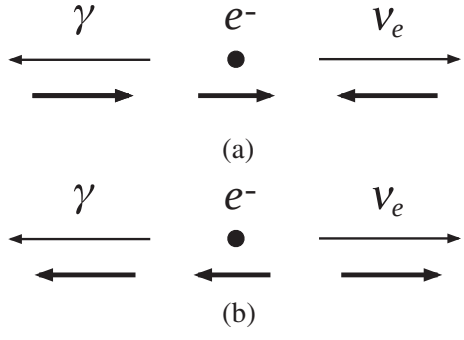


Figure 10.5 Possible helicities of the photons and neutrinos emitted in (10.12) for those events in which they are emitted in opposite directions. Measurements of the photon helicity show that only events corresponding to case (a), left-handed neutrinos, are observed.

A similar experiment for antineutrinos has been carried out involving the emission of a γ ray following the positron decay of the $^{203}\mathrm{Hg}$ nucleus. We will not discuss this, but note that the measured photon polarization in this case is indeed consistent with the occurrence of right-handed antineutrinos only, in agreement with the statement at the beginning of this section.

---

[5] This was done by resonant scattering from a second samarium target. It relies on the fact that those γ rays travelling in the opposite direction to the neutrino have slightly more energy than those emitted in other directions, and only the former have enough energy to excite the resonance level.

[6] There are only two possible spin states for photons because there are only two possible polarization states for electromagnetic waves. In fact, the spin states with components $+1$ and -1 in the direction of the photon momentum correspond to right and left circular polarization respectively.

[7] This is discussed further in Section 7.5 of Perkins (1987), where a brief description of the apparatus is also given.

10.1.3 Pion and muon decays revisited

In our discussion of the charged current interactions in Section 8.1, we completely ignored their spin dependence. This is of a special form, called a V–A interaction.[8] Here we shall consider only the most important characteristic of this spin dependence, which is that the results discussed for neutrinos in Section 10.1.2 hold for all fermions in the ultra-relativistic limit; i.e. in the limit that their velocities approach that of light, only left-handed fermions ν_L, e_L^-, etc., and right-handed antifermions $\bar{\nu}_R$, e_R^+ are emitted in charged current interactions. For neutrinos this is always a good approximation. For other particles it is only a good approximation for large energies E, when the contributions of the 'forbidden' helicity states e_R^-, e_L^+, etc., are suppressed by factors that are typically of order

$$\left(1 - \frac{\upsilon}{c}\right) \approx \frac{m^2}{2E^2}, \tag{10.13}$$

where m is the appropriate fermion mass and υ is its velocity.

The above spin properties can be verified most easily for the electrons and muons emitted in weak decays by directly measuring their spins. Here we shall assume them to hold and use them to understand some interesting features of pion and muon decays.

We start by considering the pion decay modes

$$\pi^+ \to \ell^+ + \nu_\ell \quad (\ell = e, \mu). \tag{10.14}$$

In the rest frame of the decaying pion, the charged lepton and the neutrino recoil in opposite directions, and because the pion has zero spin, their spins must be opposed to satisfy angular momentum conservation about the decay axis. Since the neutrino is left-handed, it follows that the charged lepton must also be left-handed, as shown in Figure 10.6, in contradiction to the expectations for a relativistic antilepton. For the case of a positive muon this is unimportant, since it is easy to check that it recoils nonrelativistically and so both helicity states are allowed. However, if a positron is emitted it does recoil relativistically, implying that this mode is suppressed by a factor that is estimated from (10.13) to be of order $\frac{1}{2}(m_e/m_\pi)^2 = 9.3 \times 10^{-5}$. Thus the positron decay mode is predicted to be much rarer than the muonic mode. This is indeed the case, and the measured ratio

$$\frac{\Gamma(\pi^+ \to e^+ \nu_e)}{\Gamma(\pi^+ \to \mu^+ \nu_\mu)} = (1.230 \pm 0.004) \times 10^{-4} \tag{10.15}$$

[8] This name is derived from the behaviour under a parity transformation of the weak interaction analogue of the electromagnetic current. The letter V denotes a proper vector, i.e. one whose direction is reversed by a parity transformation, such as momentum $\mathbf{p}$. The familiar electric current, to which photons couple, transforms as a proper vector under parity. Because parity is not conserved in weak interactions, the corresponding weak current, to which $W^\pm$ bosons couple, has in addition to a vector (V) component another component whose direction is unchanged by a parity transformation. Such a quantity is called an axial vector (A); an example is orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$.

is in excellent agreement with a full calculation that takes into account both the above suppression and the difference in the density of final states for the two reactions.

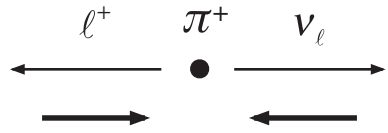
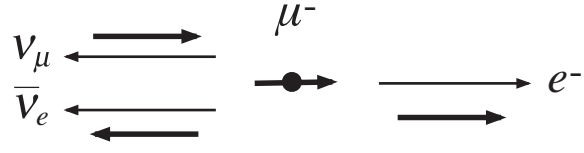


Figure 10.6 Helicities of the charged leptons and neutrinos emitted in the pion decays (10.14).

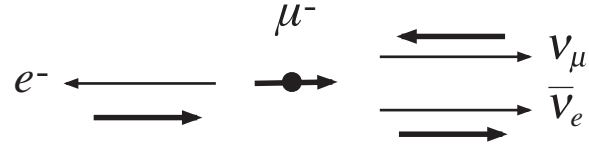
A second consequence of the helicity argument is that the muons emitted in pion decays are polarized (see Figure 10.6) and can therefore be used to measure the muon decay asymmetries discussed in Section 10.1.1. These have their origin in the spin structure of the interaction, as we shall illustrate for the highest-energy electrons emitted in the decays (10.2). These have energy

$$E = \frac{m_\mu}{2}\left(1 + \frac{m_e^2}{m_\mu^2}\right) \gg m_e \tag{10.16}$$

and correspond to decays in which the neutrino and antineutrino are both emitted in the direction opposite to the electron. This is illustrated in Figure 10.7 for the two simplest cases in which the electron is emitted in the muon spin direction, and opposite to it. Since the neutrino and antineutrino have opposite helicities, the muon and electron must have the same spin component along the event axis in order to conserve angular momentum, implying the electron helicities shown in Figure 10.7. When combined with the fact that the relativistic electrons emitted must be left-handed, this implies that electrons cannot be emitted in the muon spin direction. We thus see that the spin



(a) Forbidden



(b) Allowed

Figure 10.7 Muon decays in which electrons of the highest possible energy are emitted (a) in the muon spin direction and (b) opposite to the muon spin direction. Case (a) is heavily suppressed, because in the limit that its velocity approaches that of light, an electron emitted in a charge current reaction must necessary be left-handed.

structure of the interaction automatically gives rise to a forward–backward asymmetry in polarized muon decays. Of course, not all the electrons have the maximum energy (10.16), and the actual asymmetry, averaged over all electron energies, can only be calculated by using the full form of the *V–A* interaction. We will not pursue this further, but merely note that the resulting prediction is in excellent agreement with the measured values of Equation (10.7).

10.2 *CP* VIOLATION

We now turn to the evidence for *CP* violation in the weak interactions of hadrons, which so far is confined to the decays of neutral K mesons and of B mesons. Neutral kaons and neutral B mesons also exhibit the phenomenon of particle–antiparticle mixing, which is crucial for understanding the *CP* violation experiments. Here we first discuss the evidence for *CP* violation in kaon decays, where it was first discovered, and then go on to consider the B meson system, where recent discoveries have been of great importance in testing the standard model.

10.2.1 *CP* eigenstates

In Section 3.4 we saw that there are two neutral kaon states

$$K^0(498) = d\bar{s} \quad \text{and} \quad \bar{K}^0(498) = s\bar{d}, \tag{10.17}$$

which have strangeness $S = +1$ and $S = -1$, respectively. However, because strangeness is not conserved in weak interactions, these states can be converted into each other by higher-order weak processes like that shown in Figure 10.8. This is in marked contrast to most other particle–antiparticle systems, for which such transitions are forbidden because the particle and its antiparticle differ by quantum numbers that are conserved in all known interactions. For example, the π^+ and π^- have opposite electric charges and the neutron and antineutron have opposite baryon numbers. For neutral kaons, however, there is no conserved quantum number to distinguish the K^0 and $\bar{K}^0$ states when weak interactions are taken into account and the observed physical particles correspond not to the K^0 and $\bar{K}^0$ states themselves, but to linear combinations of them. This is the phenomenon of $K^0 - \bar{K}^0$ mixing, which we shall now describe. We shall see later that a similar phenomenon occurs for some other states, including neutral B mesons and neutral D mesons.

We start by assuming that *CP* invariance is exact and that neutral kaons are eigenstates of the combined *CP* operator. To find the effect of this operator, we note that the operator $\hat{C}$ changes a quark into its antiquark, and so interchanges the states (10.17), giving

$$\hat{C}\,|K^0, \mathbf{p}\rangle = -\,|\bar{K}^0, \mathbf{p}\rangle \quad \text{and} \quad \hat{C}\,|\bar{K}^0, \mathbf{p}\rangle = -\,|K^0, \mathbf{p}\rangle. \tag{10.18}$$

Here we have used the Dirac notation to denote a state with momentum $\mathbf{p}$, and the signs in (10.18) are chosen to accord with the standard convention used

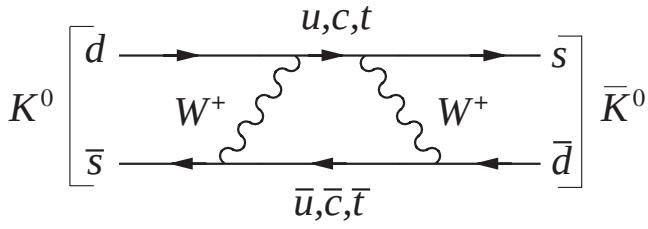


Figure 10.8 Example of a process that can convert a $K^0 \equiv \bar{s}d$ state into a $\bar{K}^0 \equiv s\bar{d}$ state. The labels on the internal lines identify the quark or antiquark going from left to right, i.e. not necessarily in the direction of the arrow.

to discuss neutral kaon decays.[9] Since kaons have negative intrinsic parity, we also have

$$\hat{P}|\,K^0, \mathbf{p}=\mathbf{0}\rangle = -|K^0, \mathbf{p}=\mathbf{0}\rangle \quad \text{and} \quad \hat{P}\,|\bar{K}^0, \mathbf{p}=\mathbf{0}\rangle = -\,|\bar{K}^0, \mathbf{p}=\mathbf{0}\rangle, \tag{10.19}$$

which together with (10.18) give

$$\hat{C}\hat{P}\,|\,K^0, \mathbf{p}=0\rangle = |\bar{K}^0, \mathbf{p}=\mathbf{0}\rangle \quad \text{and} \quad \hat{C}\hat{P}\,|\,\bar{K}^0, \mathbf{p}=0\rangle = |K^0, \mathbf{p}=\mathbf{0}\rangle \tag{10.20}$$

for the action of the combined operator $\hat{C}\hat{P}$. From these equations it is straightforward to verify that the *CP* eigenstates are

$$\left|K_1^0,\ \mathbf{p}=\mathbf{0}\right\rangle \equiv \frac{1}{\sqrt{2}}\left[\left|K^0,\ \mathbf{p}=\mathbf{0}\right\rangle + \left|\bar{K}^0,\ \mathbf{p}=\mathbf{0}\right\rangle\right] \tag{10.21a}$$

and

$$\left|K_2^0,\ \mathbf{p}=\mathbf{0}\right\rangle \equiv \frac{1}{\sqrt{2}}\left[\left|K^0,\ \mathbf{p}=\mathbf{0}\right\rangle - \left|\bar{K}^0,\ \mathbf{p}=\mathbf{0}\right\rangle\right], \tag{10.21b}$$

with

$$CP\left|K_1^0,\ \mathbf{p}=\mathbf{0}\right\rangle = \left|K_1^0, \mathbf{p}=\mathbf{0}\right\rangle \tag{10.22a}$$

and

$$CP\left|K_2^0,\ \mathbf{p}=\mathbf{0}\right\rangle = -\left|K_2^0,\ \mathbf{p}=\mathbf{0}\right\rangle, \tag{10.22b}$$

respectively. According to our initial assumptions, these are the states corresponding to the observed particles, and while the K_1^0 should decay entirely to states with $CP=1$,

[9] See Equations (5.50a) and (5.50b). If the negative signs were to be omitted in (10.18) then some sign changes would be necessary in the equations that follow, but the physical predictions would remain the same.

the K_2^0 should decay entirely into states with $CP = -1$. As we shall see, this leads to the predictions that the pionic decays

$$K_1^0 \to \pi^+\pi^-,\ \pi^0\pi^0 \tag{10.23a}$$

and

$$K_2^0 \to \pi^+\pi^-\pi^0,\ \pi^0\pi^0\pi^0 \tag{10.23b}$$

are allowed by *CP* conservation, while the decays

$$K_1^0 \to \pi^+\pi^-\pi^0,\ \pi^0\pi^0\pi^0 \tag{10.24a}$$

and

$$K_2^0 \to \pi^+\pi^-,\ \pi^0\pi^0 \tag{10.24b}$$

are forbidden.

The results (10.23) and (10.24) are central to the interpretation of the observed neutral kaon decays that follows. They are established by showing that the two-pion final states have $CP = +1$, while the three-pion final states have $CP = -1$. We will show this by considering each possible final state in turn.

10.2.1(a) Two-pion final states

Since the kaon has spin-0, the pion pair must have zero orbital angular momentum in the rest frame of the decaying particle, by angular momentum conservation. Its parity is therefore given by

$$P = P_\pi^2\,(-1)^L = 1, \tag{10.25}$$

where $P_\pi = -1$ is the intrinsic parity of the pion. For the $\pi^0\,\pi^0$ the C-parity is given by

$$C = (C_{\pi^0})^2 = 1$$

where C_{π^0} is the C-parity of the neutral pion. Combining these results gives $CP = +1$ for the $\pi^0\pi^0$ final state, as required. For the $\pi^+\pi^-$ state, the parity is again given by (10.25), while charge conjugation interchanges the two charged pions, giving

$$C = (-1)^L = 1, \tag{10.26}$$

by (5.52). Combining (10.25) and (10.26) again gives $CP = +1$ for the $\pi^+\pi^-$ state, as required.

10.2.1(b) Three-pion final states

The argument for three-pion final states is more complicated because there are two orbital angular momenta to consider, as shown in Figure 10.9. Here $\mathbf{L}_{12}$ is the orbital

angular momentum of the chosen pair in their mutual centre-of-mass frame and $\mathbf{L}_3$ is the orbital angular momentum of the third pion about the centre-of-mass of the pair in the overall centre-of-mass frame. The total orbital angular momentum

$$\mathbf{L} \equiv \mathbf{L}_{12} + \mathbf{L}_3 = \mathbf{0},$$

since the decaying particle has spin-0, and this can only be satisfied if $L_{12} = L_3$. This implies that the parity of the final state is

$$P = P_\pi^3 (-1)^{L_{12}} (-1)^{L_3} = -1. \tag{10.27}$$

For the state $\pi^0\pi^0\pi^0$, the C-parity is

$$C = (C_{\pi^0})^3 = 1,$$

and so combining this with (10.27) gives $CP = -1$ overall, as required. For the state $\pi^+\pi^-\pi^0$, the parity is again given by (10.27) but now charge conjugation interchanges the π^+ and π^- giving a factor $(-1)^{L_{12}}$ in analogy with (10.26), together with a factor for the C-parity of a neutral pion; i.e. we have

$$C = C_{\pi^0} (-1)^{L_{12}} = (-1)^{L_{12}},$$

and combining this with (10.27), we obtain

$$CP = (-1)^{L_{12}+1}.$$

The value of L_{12} can be determined experimentally by studying the angular distribution of the pions, leading to the results $L_{12} = 0$ and $CP = -1$ as required.

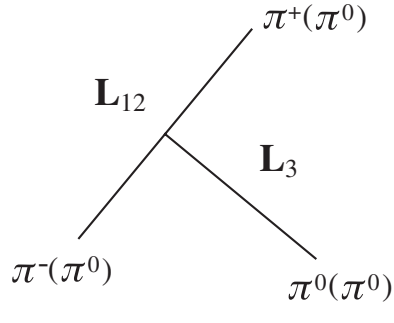


Figure 10.9 Angular momenta in the three-pion systems $\pi^+\pi^-\pi^0$ and (in brackets) $\pi^0\pi^0\pi^0$.

10.2.2 The discovery of *CP* violation

Two neutral kaons are observed experimentally, called K^0-*short* and K^0-*long*, denoted K_S^0 and K_L^0, respectively. They have almost equal masses[10] of about

[10] The tiny mass difference between them will be determined in Section 10.3.

498 MeV/c^2, but very different lifetimes and decay modes. The K_S^0 has a lifetime of $\sim 9 \times 10^{-11}$ s and principal decay modes

$$K_S^0 \to \pi^0 + \pi^0, (B = 0.31) \quad \text{and} \quad K_S^0 \to \pi^+ + \pi^-, \ (B = 0.69), \tag{10.28}$$

with the branching ratios shown, while the longer-lived K_L^0 has a lifetime of $\sim 5 \times 10^{-8}$ s and its principal decay modes are

$$K_L^0 \to \pi^0 + \pi^0 + \pi^0 \ (B = 0.20), \quad K_L^0 \to \pi^+ + \pi^- + \pi^0 \ (B = 0.13) \tag{10.29}$$

and

$$K_L^0 \to \pi^\pm + l^\mp + \nu_l \, (\bar{\nu}_l) \ \ (B = 0.67). \tag{10.30}$$

If *CP* is conserved, the selection rules (10.23) and (10.24) immediately suggest the identification

$$K_S^0 = K_1^0 \quad \text{and} \quad K_L^0 = K_2^0. \tag{10.31}$$

However, in 1964 Christenson, Cronin, Fitch and Turlay discovered that K_L^0 also decayed to two pions

$$K_L^0 \to \pi^+ + \pi^-, \tag{10.32}$$

with a branching ratio of order 10^{-3}. This result is clear evidence of *CP* violation since two-pion final states have $CP = 1$, whereas the dominant three-pion final states (10.29) have $CP = -1$ as shown above.

The first stage in demonstrating the existence of the decay mode (10.32) was to prepare a K_L^0 beam. This was done by colliding a 30 GeV proton beam with a metal target and forming a secondary beam from the produced particles using a lead collimator. Charged particles were swept out of the latter beam by a bending magnet, while most of the photons were removed by passing it through a 4 cm thick block of lead. At this point the beam contained both K_S^0 and K_L^0 mesons, but by the time it reached the detection apparatus 18 m away the short-lived K_S^0 component had completely decayed, leaving a beam that consisted primarily of K_L^0 particles and neutrons, together with a few photons that had penetrated the lead filter.

The detection apparatus used is shown schematically in Figure 10.10. It consisted of a second lead collimator to define the direction of the beam precisely, a helium-filled bag within which the observed decays occurred and a pair of symmetrically placed spectrometers to observe the pions produced in the decay. The main part of each spectrometer consisted of a pair of spark chambers separated by a bending magnet to measure the direction and momentum of any incident charged particles. However, these were only triggered when scintillation and water Čerenkov counters, also contained in the spectrometers, recorded the passage of two charged particles with velocities greater than or equal to $0.75c$. This automatically eliminated many background events involving slow-moving particles produced, for example, in collisions involving the residual neutrons in the beam. Even so, only some of the detected

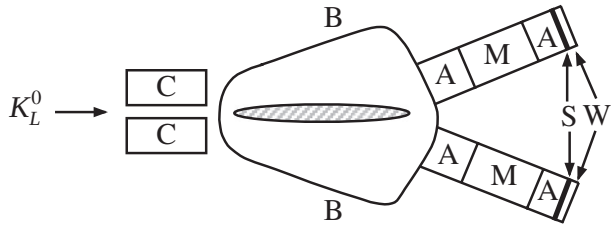


Figure 10.10 Schematic diagram of the apparatus used in the discovery of *CP* violation. The K_L^0 beam entered a helium-filled bag B through a lead collimator C, and those *CP*-violating decays that occurred in the shaded region were detected by the symmetrically spaced spectrometers. These each contained a pair of spark chambers A separated by a magnet M, followed by a scintillation counter S and a water Čerenkov counter W. (Reprinted Figure 1 with permission from J. H. Christenson *et al.*, *Phys. Rev. Lett.*, **13**, 138. Copyright 1964 American Physical Society.)

events corresponded to the sought after decays (10.32). These were identified using the facts that if the pair of charged particles detected arose from K_L^0 decays they should have opposite charges, their invariant mass should correspond to the K_L^0 mass of 498 MeV/c^2 and their net momentum should lie in the direction of the K_L^0 beam.

The results of the experiment are illustrated in Figure 10.11. This shows the observed distributions of oppositely charged pairs as a function of the angle θ between

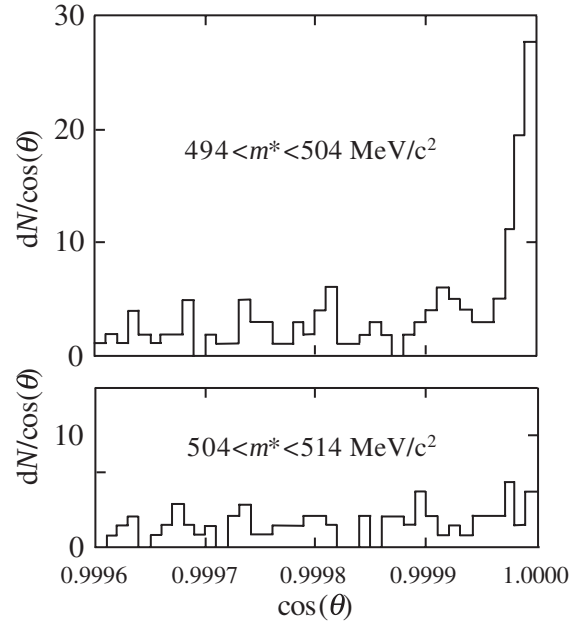


Figure 10.11 Angular distribution of the $\pi^+\pi^-$ pairs detected using the apparatus of Figure 10.10, where θ is the angle between the line-of-flight of the centre-of-mass of the pair and the initial beam direction. Results are shown for an invariant mass range including the K_L^0 mass (498 MeV/c^2) and for a neighbouring mass range. (Reprinted Figure 1 with permission from J. H. Christenson *et al.*, *Phys. Rev. Lett.*, **13**, 138. Copyright 1964 American Physical Society.)

the net momentum of the pair and the beam direction. The point here is that decay events should occur only at zero angles to the beam and at an invariant mass of 498 MeV/c^2, whereas background events from other sources will presumably have a smooth dependence on this angle and on the invariant mass. As can be seen, there is a clear peak in the beam direction $\theta = 0^0$ for the appropriate invariant mass range $494 < m^* < 504\,\mathrm{GeV/c^2}$, which is interpreted as arising from the *CP*-violating decay (10.32).

10.2.3 *CP*-violating K_L^0 decays

Because *CP* is not conserved, the physical states K_S^0 and K_L^0 need not correspond to the *CP* eigenstates K_1^0 and K_2^0 as assumed in (10.31), but can contain small components of states with the opposite *CP*; i.e.

$$\left| K_S^0,\, \mathbf{p}=\mathbf{0}\right\rangle = \frac{1}{(1+|\varepsilon|^2)^{1/2}}\left[\left| K_1^0,\, \mathbf{p}=\mathbf{0}\right\rangle - \varepsilon \left| K_2^0,\, \mathbf{p}=\mathbf{0}\right\rangle\right] \tag{10.33a}$$

and

$$\left|K_L^0,\, \mathbf{p}=\mathbf{0}\right\rangle = \frac{1}{(1+|\varepsilon|^2)^{1/2}}\left[\varepsilon\, \left|K_1^0,\, \mathbf{p}=\mathbf{0}\right\rangle + \left|K_2^0,\, \mathbf{p}=\mathbf{0}\right\rangle\right], \tag{10.33b}$$

where ε is a small complex parameter. This mixing can be verified rather directly by measurements on the semileptonic decays

$$K_L^0 \to \pi^- + e^+ + \nu_e \quad \text{and} \quad K_L^0 \to \pi^+ + e^- + \bar{\nu}_e. \tag{10.34}$$

To understand this, we note that the K^0 and $\bar{K}^0$ can decay by the semileptonic reactions

$$K^0 \to \pi^- + e^+ + \nu_e \quad \text{and} \quad \bar{K}^0 \to \pi^+ + e^- + \bar{\nu}_e, \tag{10.35}$$

whereas the corresponding reactions

$$K^0 \to \pi^+ + e^- + \bar{\nu}_e \quad \text{and} \quad \bar{K}^0 \to \pi^- + e^+ + \nu_e \tag{10.36}$$

are forbidden by the $\Delta S = \Delta Q$ selection rule (8.30). Thus the relative yields N^+ and N^- of positrons and electrons in the decays (10.34) yield a direct measure of the relative probabilities of finding a K^0 or $\bar{K}^0$, respectively, in the decaying K_L^0. From (10.33b) and (10.21) these are given by

$$N^{\pm} \propto |1 \pm \varepsilon|^2 \left(1+|\varepsilon|^2\right)^{-1}$$

for K^0 and $\bar{K}^0$, respectively, so if we neglect terms of order $|\varepsilon|^2$, the asymmetry

$$A \equiv (N^+ - N^-)/(N^+ + N^-) = 2\,\mathrm{Re}\varepsilon \tag{10.37}$$

for a pure K_L^0 beam.

Figure 10.12 shows data on the asymmetry (10.37) as a function of proper time, i.e. time in the rest frame of the decaying particles, for the case of a beam that is initially predominantly a K^0 state. After the initial oscillations, which occur as the K_S^0 decays more rapidly than the K_L^0, only the K_L^0 state survives and there is seen to be an asymmetry whose value is $2\mathrm{Re}\varepsilon \approx 3.3 \times 10^{-3}$.

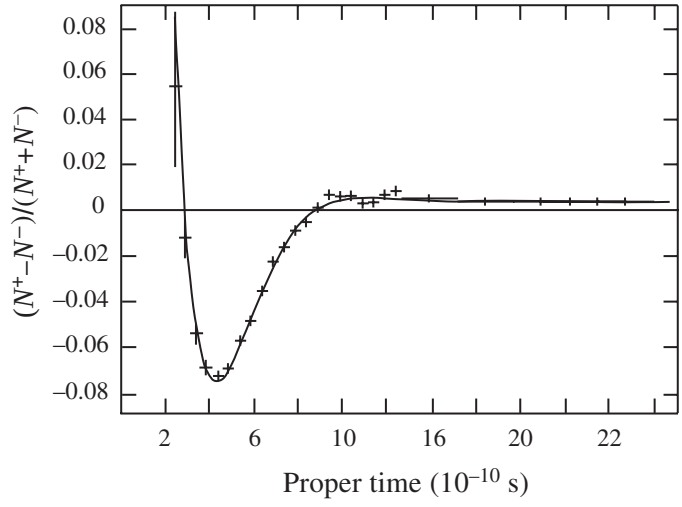


Figure 10.12 The charge asymmetry observed for $K^0 \to \pi^- e^+ \nu_e$ and $\bar{K}^0 \to \pi^+ e^- \bar{\nu}_e$ as a function of proper time, for a beam that is initially predominantly K^0. (Reprinted from *Physics Letters B*, **52**, 113 Gjesdal, S., *et al*. Copyright 1974, with permission from Elsevier.)

The magnitude of ε can be deduced from the measured rates of the two *CP*-violating decays

$$K_L^0 \to \pi^+ + \pi^- \ (B = 2.0 \times 10^{-3}) \quad \text{and} \quad K_L^0 \to \pi^0 + \pi^0 \ (B = 8.7 \times 10^{-4}), \qquad (10.38)$$

which have both been observed with the branching ratios shown. These can occur in two different ways, either by:

(a) *CP violation by mixing*, in which the *CP*-forbidden K_1^0 component in the K_L^0 decays via the *CP*-allowed process (10.23a), giving a contribution proportional to the probability

$$|\varepsilon|^2/(1+|\varepsilon|^2) \approx |\varepsilon|^2$$

of finding a K_1^0 component in the K_L^0, or

(b) *direct CP violation,* in which the *CP*-allowed K_2^0 component in the K_L^0 decays via the *CP*-violating reactions (10.24a).

In practice, *CP* violation by mixing dominates, as can be seen by comparing the ratios of probability amplitudes

$$\eta_{00} \equiv \frac{\mathcal{M}(K_L^0 \to \pi^0\pi^0)}{\mathcal{M}(K_S^0 \to \pi^0\pi^0)}, \quad \eta_{+-} \equiv \frac{\mathcal{M}(K_L^0 \to \pi^+\pi^-)}{\mathcal{M}(K_S^0 \to \pi^+\pi^-)}. \tag{10.39}$$

From (10.33) one easily sees that, if direct *CP* violation is completely neglected, these parameters are given by

$$\eta_{+-} = \eta_{00} = \varepsilon \left(1 + |\varepsilon|^2\right)^{-1/2}. \tag{10.40}$$

The measured values,

$$|\eta_{00}| = (2.225 \pm 0.007) \times 10^{-3}, \quad |\eta_{+-}| = (2.236 \pm 0.007) \times 10^{-3}$$

are in almost exact agreement with the prediction of *CP* violation by mixing alone, although a very small, but nonzero, contribution is required for full agreement. A detailed analysis confirms this interpretation with

$$|\varepsilon| = (2.232 \pm 0.007) \times 10^{-3}. \tag{10.41}$$

10.2.4 *CP* violation in *B* decays

Until 2001, the evidence for *CP* violation was confined solely to the $K^0 - \bar{K}^0$ system. However, since then many examples of *CP* violation have been discovered in the analogous $B^0 - \bar{B}^0$ system, where B^0 is the $B^0(5279) = d\bar{b}$ meson. $B^0 - \bar{B}^0$ mixing can then occur by diagrams like Figure 10.13, in analogy to $K^0 - \bar{K}^0$ mixing (Figure 10.8), and can be described by the same formalism. The physical particles analogous to the K_S^0 and K_L^0 mesons are in this case called B_L^0 and B_H^0, where L and H stand for 'light' and 'heavy', respectively.[11] They have almost identical lifetimes of approximately 1.5×10^{-12} s. Because this is very short compared to the lifetimes of

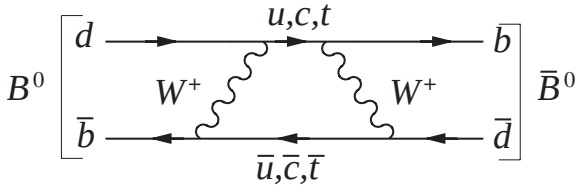


Figure 10.13 Example of a process that can convert a $B^0 \equiv d\bar{b}$ state into a $\bar{B}^0 \equiv \bar{d}b$ state. The labels on the internal lines identify the quark or antiquark going from left to right, i.e. not necessarily in the direction of the arrow.

[11] The masses are in fact almost identical, and the tiny difference between them will be determined in Section 10.3 below.

the neutral kaons, it is not possible to form well-defined beams of neutral B mesons, so some other means of studying their decay modes, which are numerous, must be found.

This problem has been overcome by the construction of so-called *B factories*. These facilities exploit the properties of the $\Upsilon(4S) = b\bar{b}$ resonance, which has a mass of 10.58 GeV and a width of only 20 MeV. This is a state of bottomium that is just heavy enough to decay to the lightest meson states with the nonzero bottom quantum number:

$$B^+(5279) = u\bar{b}, \quad B^0(5279) = d\bar{b} \quad (\tilde{B} = +1) \tag{10.42a}$$

and

$$B^-(5279) = b\bar{u}, \quad \bar{B}^0(5279) = b\bar{d} \quad (\tilde{B} = -1) \tag{10.42b}$$

by the mechanism of Figure 10.14, but not heavy enough to decay to any other final states by the same mechanism.[12] It therefore decays almost entirely to B^+B^- and $B^0\bar{B}^0$ pairs, in approximately equal numbers. In addition, it has the same quantum numbers $J^{PC} = 1^{--}$ as the photon and so can be produced in e^+e^- annihilation by the mechanism of Figure 6.3. There is correspondingly a peak in the e^+e^- annihilation cross-section, and tuning the beam energies to coincide with this peak results in a copious source of $B^\pm$, B^0 and $\bar{B}^0$ mesons.

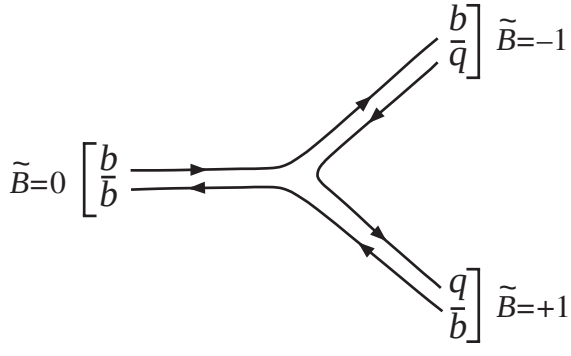


Figure 10.14 Mechanism for the decay of the 4S $\tilde{B} = 0$ state $\Upsilon(10.58)$; $q = u$ or d.

Two such B factories have been constructed to study *CP* violation in B decays: the PEP-II facility at SLAC, California, in which a 3.1 GeV positron beam is collided with a 9.0 GeV electron beam, and the KEK-B facility in Japan, in which a 3.5 GeV positron beam is collided with an 8 GeV electron beam. In each case, the centre-of-mass energy corresponds to the $\Upsilon(4S)$ mass, while the asymmetric beam energies ensure that B mesons are produced with enough momentum to travel a measurable distance before decaying. This is important, because studies of *CP* violation often require the

[12] Cf. the discussion of charmonium and bottomium in starred Sections 6.4.1 and 6.4.2.

measurement of the time between the production and decay of the B mesons. This is achieved using the dedicated detectors BaBar at PEP-II and Belle at KEK-B.

The BaBar detector is shown in Figure 10.15 and, like the Belle detector, has the typical structure of a 'layered' detector. It is made up of five subdetectors, which from the inside outwards are:

1. A silicon vertex tracker that provides precise position information on charged tracks, and is also the sole tracking device for very low-energy particles.
2. A tracking (drift) chamber that provides the main momentum measurements for charged particles and helps in particle identification through measurements of the rate of energy loss.
3. A detector of internally reflected Čerenkov radiation that provides charged particle identification.
4. An electromagnetic calorimeter that provides particle identification for electrons and photons.
5. An instrumented flux return that provides muon and neutral hadron identification.

In addition, there is a solenoid that provides the 1.5 T magnetic field required for charge and momentum measurements.

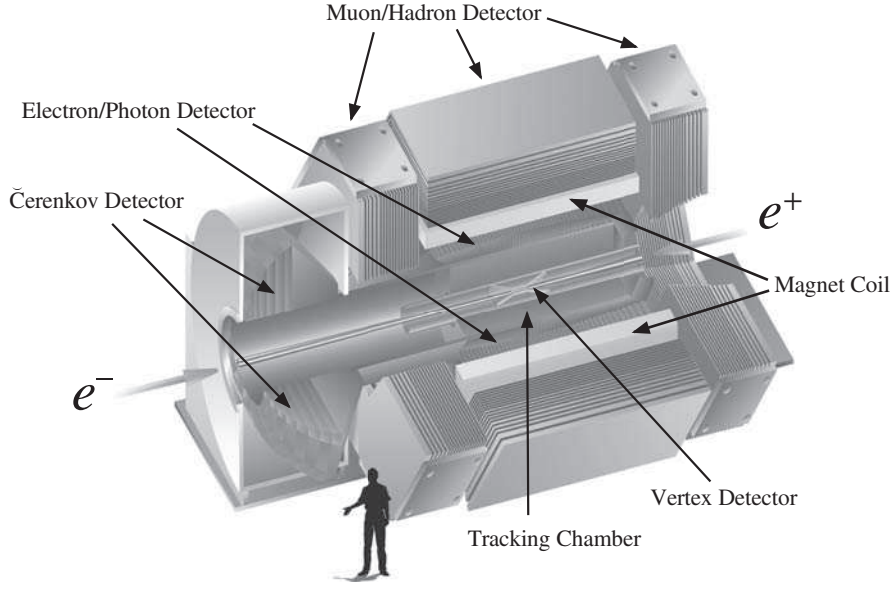


Figure 10.15 Schematic view of the BaBar detector at the PEP-II facility at SLAC, California. The components of the detector are described in the text. (Photo courtesy of Stanford Linear Accelerator Center.)

Because B_L^0 and B_H^0 have very similar lifetimes, it is not possible to observe a pure sample of one species analogous to the K_L^0 in neutral kaon decays, so one is always dealing with particle mixtures, making the analysis of the data more complicated

than in the neutral kaon case. We will not go into details, but merely note that many *CP*-violating effects have been observed. For example, direct *CP* violation has been observed in the $\bar{B}^0 \to K^-\pi^+$ mode, where the combined data on the asymmetry parameter

$$A_{K\pi} = \frac{\Gamma(\bar{B}^0 \to K^-\pi^+) - \Gamma(B^0 \to K^+\pi^-)}{\Gamma(\bar{B}^0 \to K^-\pi^+) + \Gamma(B^0 \to K^+\pi^-)} \tag{10.43}$$

yields the value

$$A_{K\pi} = -0.095 \pm 0.013.$$

Effects of similar magnitude, involving both direct *CP* violation and mixing, have been observed in other decay modes, including $J/\Psi K_S^0$, $K^+K^-K_S^0$ and $\eta' K_S^0$, where $\eta'(958)$ is the meson resonance listed in Table E.5. The effects are very big compared to those observed in neutral kaon decays, a result that was expected on the basis of the standard model. This will be discussed in Section 10.4, where we will see that large *CP*-violating effects are also expected for decays of B_S^0 and $\bar{B}_S^0$, but these are much more difficult to detect. In contrast, those predicted for the $D^0 - \bar{D}^0$ system are much smaller and have not been observed at the time of writing.

10.3 FLAVOUR OSCILLATIONS AND THE *CPT* THEOREM

Although *CP* is not conserved, there is good reason to believe that all interactions are invariant under the combined operation of charge conjugation C, parity P and time reversal T, taken in any order. This result is called the *CPT theorem*, and can be shown to hold in any relativistic quantum theory in which signals cannot propagate faster than the speed of light. Like *CP*, the combined operation of *CPT* converts particles at rest to antiparticles at rest, and invariance under this operation requires particles and antiparticles to have the same masses and lifetimes. This is in accord with experiment and is tested by observations on so-called 'flavour oscillations' in neutral meson systems. Here we will concentrate on the neutral kaons and B mesons, as these yield the most precise data.

When neutral kaons are produced in strong interaction processes, they are almost invariably produced with definite strangeness. For example, the neutral kaon produced in the reaction

$$\begin{aligned} &\pi^- + p \to K^0 + \Lambda^0 \\ S = 0 \quad &0 \quad 1 \quad -1 \end{aligned} \tag{10.44}$$

must necessarily be in a K^0 state with $S = 1$, in order to conserve strangeness. However, if the produced particle is allowed to travel through free space and its strangeness is measured, one finds that it no longer has a definite strangeness $S = 1$, but has components with both $S = 1$ and $S = -1$ whose intensities oscillate with time. These are called *strangeness oscillations* and enable the mass difference between K_S^0 and K_L^0 particles to be measured with extraordinary precision, as we will now show.

In what follows, we shall measure time in the rest frame of the produced particle and define $t = 0$ as the moment when it is produced. The initial state produced in the reaction (10.44) is

$$\left|K^0, \mathbf{p}\right\rangle = \frac{1}{\sqrt{2}}\left[\left|K_S^0, \mathbf{p}\right\rangle + \left|K_L^0, \mathbf{p}\right\rangle\right],$$

where for simplicity we ignore small corrections from *CP* violation throughout. At later times, however, this will become

$$\frac{1}{\sqrt{2}}\left[a_S(t)\ \left|K_S^0, \mathbf{p}\right\rangle + a_L(t)\ \left|K_L^0, \mathbf{p}\right\rangle\right], \tag{10.45}$$

where

$$a_\alpha\,(t) = e^{-im_\alpha t}\, e^{-\Gamma_\alpha t/2} \quad (\alpha = S, L), \tag{10.46}$$

and m_α and Γ_α are the mass and decay rate of the particle concerned. Here the first exponential factor is the usual oscillating time factor $\exp(-iEt)$ associated with any stationary state, evaluated in the rest frame of the particle. The second exponential factor reflects the fact that the particles decay, and it ensures that the probability

$$\left|\frac{1}{\sqrt{2}} a_\alpha\,(t)\right|^2 = \frac{1}{2} e^{-\Gamma_\alpha t} \quad (\alpha = S, L)$$

of finding a K_S^0 or K_L^0 decreases exponentially with a mean lifetime

$$\tau_\alpha = \Gamma_\alpha^{-1} \quad (\alpha = S, L),$$

whose value is given in Section 10.2.2. Because $\tau_S \ll \tau_L$, for times t such that $\tau_S \ll t \ll \tau_L$ only the K_L^0 component survives, implying equal intensities for the K^0 and $\bar{K}^0$ components. Here we are interested in the intensities of the K^0 and $\bar{K}^0$ components at shorter times, and to deduce these we use

$$\left|\,K_S^0, \mathbf{p}\right\rangle = \left|\,K_1^0, \mathbf{p}\right\rangle \equiv \frac{1}{\sqrt{2}}\left[\left|K^0, \mathbf{p}\right\rangle + \left|\bar{K}^0, \mathbf{p}\right\rangle\,\right] \tag{10.47a}$$

and

$$\left|\,K_L^0, \mathbf{p}\right\rangle = \left|\,K_2^0, \mathbf{p}\right\rangle \equiv \frac{1}{\sqrt{2}}\left[\left|K^0, \mathbf{p}\right\rangle - \left|\bar{K}^0, \mathbf{p}\right\rangle\,\right], \tag{10.47b}$$

where again we ignore small *CP*-violating effects. Equation (10.45) can then be written in the form

$$A_0\,(t)\ \left|K^0, \mathbf{p}\right\rangle + \bar{A}_0\,(t)\ \left|\bar{K}^0, \mathbf{p}\right\rangle, \tag{10.48}$$

where

$$A_0(t) = \frac{1}{2}\left[a_S(t) + a_L(t)\right] \tag{10.49a}$$

and

$$\bar{A}_0(t) = \frac{1}{2}\left[a_S(t) - a_L(t)\right]. \tag{10.49b}$$

The intensities of the two components are then given by

$$I(K^0) \equiv |A_0(t)|^2 = \frac{1}{4}\left[e^{-\Gamma_S t} + e^{-\Gamma_L t} + 2e^{-(\Gamma_S+\Gamma_L)t/2}\cos(\Delta m t)\right] \tag{10.50a}$$

and

$$I(\bar{K}^0) \equiv \left|\bar{A}_0(t)\right|^2 = \frac{1}{4}\left[e^{-\Gamma_S t} + e^{-\Gamma_L t} - 2e^{-(\Gamma_S+\Gamma_L)t/2}\cos(\Delta m t)\right], \tag{10.50b}$$

where

$$\Delta m = |m_S - m_L| \tag{10.51}$$

and we have used (10.47) to explicitly evaluate the amplitudes (10.49). The resulting predictions for $\Delta m \cdot \tau_S = 0.5$, where τ_S is the lifetime of the K_S^0, are shown in Figure 10.16, and the 'strangeness oscillations' associated with the cosine terms in (10.50) are clearly seen.

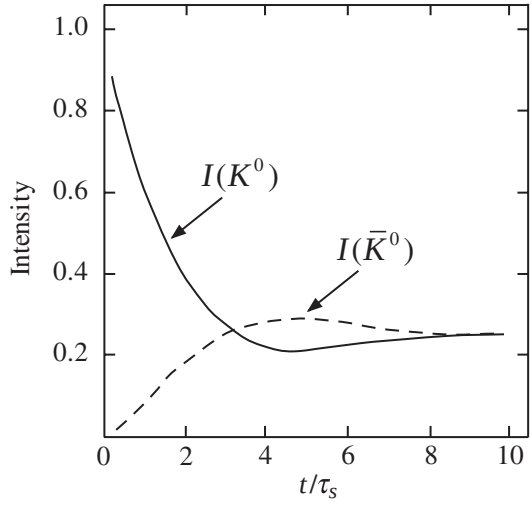


Figure 10.16 Predicted variation with time of the intensities $I(K^0)$ (solid line) and $I(\bar{K}^0)$ (dashed line) for an initial K^0 beam. The curves are calculated using (10.50) for $\Delta m \cdot \tau_S = 0.5$, where Δm is the mass difference (10.51) and τ_s is the K_S^0 lifetime.

The variation of the $\bar{K}^0$ intensity $I(\bar{K}^0)$ with time can be determined experimentally by measuring the rate of hyperon production in strangeness-conserving reactions like

$$\bar{K}^0 + p \to \pi^+ + \Lambda^0, \quad \pi^0 + \Sigma^+$$

as a function of the distance from the K^0 source. The results are in good agreement with the predictions (10.50) for a mass difference

$$\Delta m = (3.483 \pm 0.006)\, 10^{-12}\, \mathrm{MeV/c^2}, \tag{10.52}$$

which is close to the value assumed, for illustrative purposes, in Figure 10.16.

A similar phenomenon for neutral B mesons can be observed by using B meson 'tagging'. In B factories, as we have seen, pairs of B mesons result from $\Upsilon(4S)$ decays, and recoil against each other in their mutual centre-of-mass frame. Like kaons, the states produced will evolve in time, but which is which is unclear. Let us suppose, however, that one of the mesons, referred to as the tagged meson, is observed to decay by a $\bar{B}^0$ mode; the B meson that recoils against it must at that instant be thrown into a B^0 state to conserve the bottom quantum number. (This is an example of the quantum mechanical phenomenon known as 'collapse of the wavefunction'.) Away from this instant, the state will evolve into an oscillating superposition of B^0 and $\bar{B}^0$ states, whose frequency will be determined by the $m_H - m_L$ mass difference, in analogy to the oscillations described above for the neutral kaons. Hence, when both decays are observed, one obtains an oscillating likelihood of finding 'unmixed events', in which a B^0 decay is associated with a $\bar{B}^0$ decay, as a function of Δt, the measured time between the two decays. Figure 10.17 shows the asymmetry

$$A_m = \frac{N(\text{unmixed}) - N(\text{mixed})}{N(\text{unmixed}) + N(\text{mixed})} \tag{10.53}$$

obtained in one such experiment, where 'mixed' means that both mesons decay as B^0 (or $\bar{B}^0$) states, which would be impossible in the absence of $B^0 - \bar{B}^0$ mixing. The oscillations are clearly visible, and when combined with data from other experiments yield a best value of

$$m_H - m_L = (3.337 \pm 0.033) \times 10^{-10}\, \mathrm{MeV/c^2}. \tag{10.54}$$

Flavour oscillations have also been observed in the $D^0 - \bar{D}^0$, $D_S^0 - \bar{D}_S^0$ and $B_S^0 - \bar{B}_S^0$ meson systems.

The result (10.52) leads to a very precise confirmation that the masses of particles and antiparticles are equal. The K_S^0 and K_L^0 are not antiparticles, but the K^0 and $\bar{K}^0$ are, as can be seen from (10.18). The *CPT* theorem requires

$$m_{K^0} \equiv m_{\bar{K}^0}, \tag{10.55}$$

where the masses are defined as the expectation values of the energy for states of zero momentum. If this is so, the mass difference (10.51) can be shown to arise solely from the possibility of transitions $K^0 \leftrightarrow \bar{K}^0$, whose magnitude can be calculated from

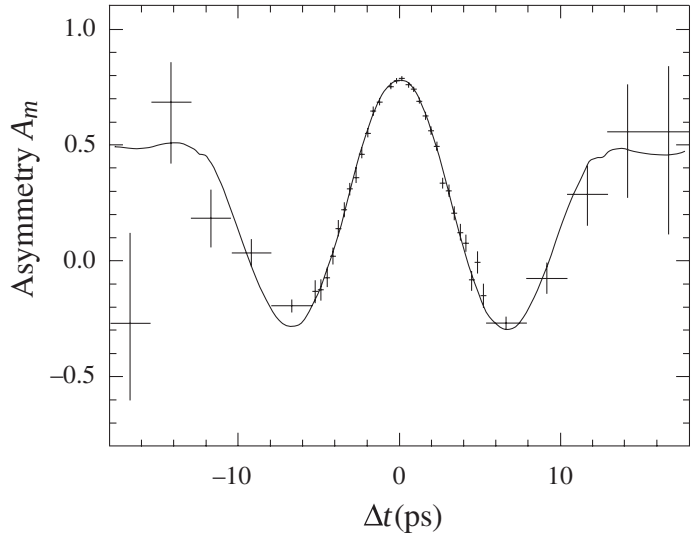


Figure 10.17 The measured asymmetry A_m, Equation (10.53), as a function of the observed time difference Δt between the two tagged neutral B decays. The error increases at large Δt as the mesons decay and fewer events are observed. (Reprinted Figure 11 with permission from B. Aubert *et al*., *Phys. Rev. D*, **73**, 012004. Copyright 2006 American Physical Society.)

diagrams like those shown in Figure 10.8. We shall not discuss this further, but merely note that the resulting agreement between the predicted and measured values confirms the identity (10.55) to better than one part in 10^{18}. In contrast, the particle–antiparticle mass relation that has been most precisely tested by direct measurement is

$$m_{e^+} = m_{e^-},$$

which is only verified to within an experimental error of order of one part in 10^8.

10.4 *CP* VIOLATION IN THE STANDARD MODEL

We conclude this chapter by reviewing whether the experimental observations on *CP* violation can be accommodated in the standard model or whether there is some other mechanism necessary. In the standard model, *CP* violation has its origins in the quark mixing described by (8.40), where the elements of the CKM matrix

$$V \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

determine the relative strengths of the various $\alpha\beta W$ couplings, where $\alpha = u, c, t$ and $\beta = d, s, b$. The CKM matrix in general contains nine complex elements. However, the unitary nature of the matrix implies that there are nine relations between the elements, such as

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0.$$

Using these, and exploiting the freedom to define the phases of the basic quark states, the CKM matrix may be parameterized by just four quantities, three mixing angles and one phase angle, and is conveniently written in the form

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix}. \tag{10.56}$$

Here $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$ and the angles $\theta_{ij} = \theta_{12}$, θ_{13} and θ_{23} can be chosen to lie in the first quadrant, so that s_{ij}, $c_{ij} \geq 0$. Since the time-reversal operator acting on a state involves complex conjugation (cf. Equations (5.71) and (5.72)), a nonzero phase δ implies T violation, and hence, if *CPT* invariance is assumed, *CP* violation. This is in contrast to the general form of the mixing matrix (8.43) for two generations only, which is purely real, so that *CP* violation could not arise from quark mixing if there were only two generations.

The mixing angles and the *CP*-violating phase must be determined from experiment, and the magnitude of a particular *CP*-violating effect depends not only on the magnitude of δ but also on the mixing angles. Experimentally, $s_{12} \gg s_{23} \gg s_{13}$, and it is instructive to approximate the resulting matrix by the so-called *Wolfenstein parameterization*:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3[1 - (1 - \frac{1}{2}\lambda^2)(\rho + i\eta)] & -A\lambda^2 + A\lambda^4[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + O(\lambda^6), \tag{10.57}$$

with parameters A, λ, ρ and η, with a nonzero value of η being indicative of *CP* violation. The quantity $\lambda = |V_{us}| \approx 0.23$ plays the role of an expansion parameter in this approximation and the other parameters are

$$A \approx 0.82, \quad \rho \approx 0.24, \quad \eta \approx 0.34.$$

From this one sees immediately that observable decays *CP*-violating effects will be largest in B decays, where V_{ub} plays an important role, as illustrated in Figure 10.18. For K and D decays, however, the main contributions to *CP* violation arise from V_{cd} and are of order λ^5, whereas V_{cd} itself is of order λ. Hence *CP*-violating effects are predicted to be much smaller. This is in general agreement with experiment, where *CP* violation in D decays is yet to be detected. Quantitative predictions are, however, more difficult to obtain, since one must take into account a variety of diagrams like those shown in Figure 10.18 for the decay $b \to du\bar{u}$ and, more problematically, the effects of the quarks being bound in mesons. We will not pursue this, but merely note that the success of the mixing model in accounting for all *CP*-violating data in terms of a single *CP*-violating phase is a major triumph of the standard model. However, this cannot be the complete story on *CP* violation, because we will see in Chapter 11 that suggested explanations of the matter–antimatter asymmetry observed in the universe require *CP* violation at a level several orders-of-magnitude larger that that needed to account for meson decays.

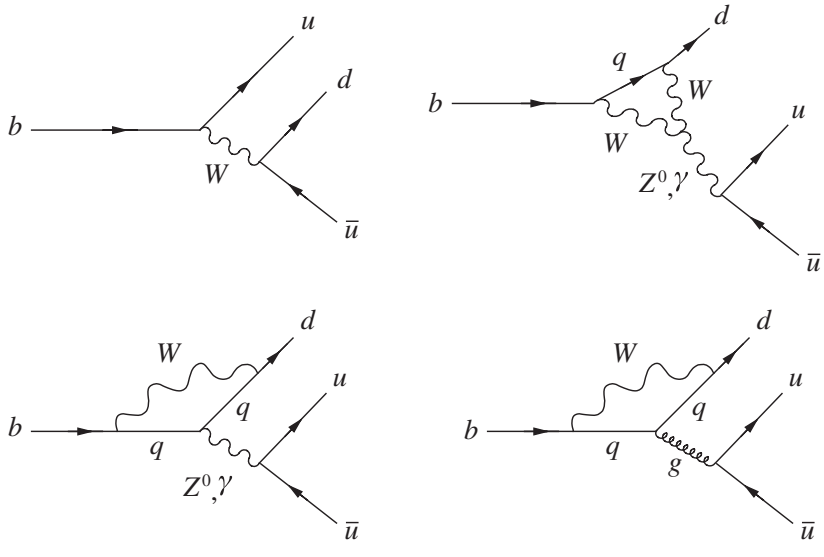


Figure 10.18 Quark diagrams that contribute to the decay $b \to du\bar{u}$, where $q = u, c, t$.

Finally, in Section 2.3.3, we noted that a complete discussion of neutrino oscillations must incorporate mixing between all three neutrino states, rather than just two. The most general form of this mixing is again (cf. (10.56))

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{10.58}$$

where the experimental values of the mixing angles θ_{12}, θ_{13} and θ_{23} are now given by (2.32), (2.38) and (2.39), respectively. The value of the phase δ is unknown, but it will hopefully be measured in future neutrino oscillation experiments, and a nonzero value will imply *CP* violation in the neutrino sector. However, such effects, if they exist, are unlikely to be detected in leptonic decays, since their contributions will be proportional to the very small neutrino masses.

PROBLEMS 10

10.1 Show that the existence of the decays $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ implies that parity is violated if the kaon is assumed to have spin zero.

10.2 The intensity of the electrons emitted in the decay (10.1) of polarized cobalt-60 nuclei is found to be consistent with the form

$$I(\upsilon, \theta) = 1 + \alpha \frac{\upsilon}{c} \cos\theta,$$

where υ is the magnitude of the electron velocity $\mathbf{v}$ and θ is the angle between its direction and the direction of the ^{60}Co spin. Deduce the value of the coefficient α by considering events in which the electron is emitted in the direction of the spin of the decaying nuclei. The spins of the ^{60}Co and ^{60}Ni nuclei are $J = 5$ and $J = 4$, respectively, and the energies of the emitted particles are sufficiently small that orbital angular momenta may be neglected.

10.3 Show that the total decay rates for the reactions $K^0 \to \pi^- e^+ \nu_e$ and $\bar{K}^0 \to \pi^+ e^- \bar{\nu}_e$ are equal if CP is conserved.

10.4 The lifetime τ_μ of the muon is given (in natural units) to a good approximation by

$$(\tau_\mu)^{-1} = G_F^2 m_\mu^5 \,/\, 192\pi^3 \,.$$

Use an analogous formula for B meson decay, ignoring phase space corrections and the fact that the quarks are bound in the hadrons, to estimate the appropriate element of the CKM matrix. Relevant decay data are given in Table E.5 of Appendix E.

10.5 The amplitudes for the decays $K^0_{S,L} \to \pi^0\pi^0$ may be written

$$\mathcal{M}(K^0_{S,L} \to \pi^0\pi^0) = \sqrt{\tfrac{2}{3}}\mathrm{e}^{i\delta_2}\mathcal{M}^{(2)}_{S,L} - \sqrt{\tfrac{1}{3}}\mathrm{e}^{i\delta_0}\mathcal{M}^0_{S,L},$$

where the phase factors are due to the strong interaction between the final state pions and the labels (0,2) refer to the isospin of the $\pi\pi$ state. ($I = 1$ is forbidden by the Bose statistic.) Use the relations

$$\left|K^0_L\right\rangle = N\left[(1+\varepsilon)\left|K^0\right\rangle - (1-\varepsilon)\left|\bar{K}^0\right\rangle\right]$$

and

$$\left|K^0_S\right\rangle = N\left[(1+\varepsilon)\left|K^0\right\rangle + (1-\varepsilon)\left|\bar{K}^0\right\rangle\right],$$

where $N = [2(1+|\varepsilon|^2]^{-1/2}$, to show that the ratio η_{00} defined in Equation (10.39) may be written

$$\eta_{00} \approx \varepsilon - i\sqrt{2}\frac{\mathrm{Im}A_2}{A_0}\exp[i(\delta_2 - \delta_0],$$

where

$$\mathcal{M}[K^0 \to (\pi\pi)_{0,2}] \equiv A_{0,2},$$

and by CPT invariance,

$$\mathcal{M}[\bar{K}^0 \to (\pi\pi)_{0,2}] = A^*_{0,2}.$$

with $A_2 \ll A_0$ and where A_0 may be taken as real. In deriving this result, you may neglect second-order terms in the small quantities ε and A_2.

11

Beyond the Standard Model

What is the path? There is no path. On into the unknown.

Goethe[1]

At the present time, provided nonzero neutrino masses are incorporated,[2] all experimental observations in particle physics are consistent with the standard model, in which strong interactions are mediated by gluons and the electroweak interactions are mediated by photons and the $W^{\pm}$ and Z^0 bosons. Nonetheless, many questions remain unanswered, and there could well be phenomena awaiting discovery, especially at higher energies, that are not described by the standard model. In this chapter we will briefly discuss a few of the possibilities currently being investigated. This material is necessarily highly conjectural and our aim is to give the flavour of current research without detailed argument. In doing so, we shall focus on three themes: the search for further unifications of the fundamental forces; the close relation between research in particle physics and important problems in astrophysics and cosmology; and whether neutrinos are their own antiparticles.

[1] This was a favourite quotation of Niels Bohr, from *Faust*. See, for example, p. 126 of Moore (1985).
[2] In the original formulation of the standard model, the neutrino mass was assumed to be zero for simplicity. However, nonzero masses are easily incorporated, as we have assumed in earlier chapters. See also Section 11.6 below.

Particle Physics Third Edition B. R. Martin and G. Shaw

11.1 GRAND UNIFICATION

With the success of the unified theory of electroweak interactions, it became natural to ask whether the strong interaction could also be united with the weak and electromagnetic interactions into a single so-called *grand unified theory* (GUT). The much greater strength of the strong interaction compared with the electroweak interaction at presently accessible energies would seem to make this a hopeless task. However, the strength of an interaction depends on the distance over which it acts, or more precisely on the magnitude of the associated energy–momentum scale μ^2, as discussed in Section 7.1.1. In particular, the strong interaction coupling decreases with μ^2, as shown in Figure 7.3, while the electroweak couplings vary much more slowly, and a naive extrapolation from their low-energy values suggests that the various couplings might become equal at an enormous value $\mu^2 = M_X^2$, where the so-called *unification mass* M_X is of order $10^{15}\,\mathrm{GeV/c^2}$. This is illustrated in Figure 11.1, where g_s is related to the QCD coupling α_s of Section 7.1 by

$$\alpha_s = g_s^2/4\pi, \tag{11.1a}$$

and the electroweak couplings are related to those defined in Sections 2.2 and 9.1, respectively, by[3]

$$g = 2\sqrt{2}g_W; \quad g' = 2\sqrt{2}g_Z. \tag{11.1b}$$

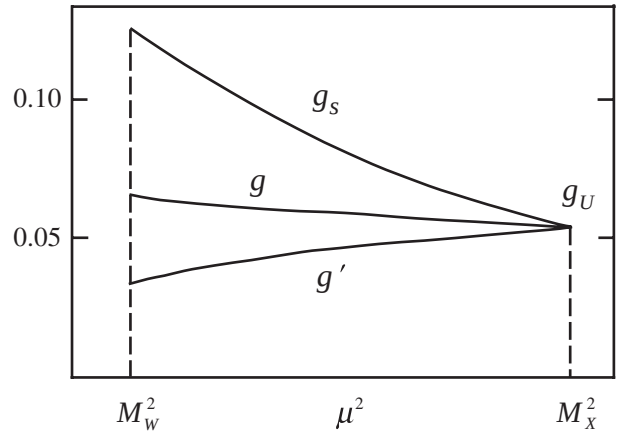


Figure 11.1 Typical behaviour of the strong and electroweak couplings (11.1) as functions of the squared energy–momentum transfer μ^2 in a typical grand unified theory.

In grand unified theories all three interactions are united into a single interaction, characterized by a single coupling constant, at the unification mass; differences

[3] The precise couplings which become equal at the unification mass depend on the particular grand unified theory in question. In the simplest case, discussed below, they are g_s, g and $\sqrt{5/3}g'$.

between them emerge as one interpolates downwards to currently available energies. Of course, this interpolation assumes that nothing totally unexpected will emerge between energies of order 10^2 GeV and 10^{15} GeV that could spoil the predictions. This is a tremendously optimistic assumption given the history of particle physics. Nevertheless, this assumption is central to grand unified theories, which have aroused considerable interest for reasons that we will describe shortly.

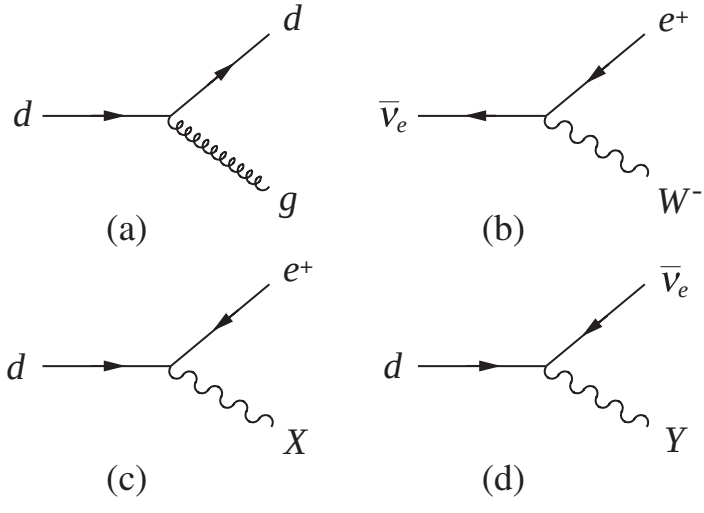


Figure 11.2 Two familiar processes (a, b) that can occur within the family of particles (11.3), together with two new processes (c, d) whose existence is predicted by grand unified theories.

There are many ways in which grand unified theories may be constructed so that they contain the predictions of both QCD and the unified electroweak theory at currently attainable energies. The earliest and simplest, due to Georgi and Glashow in 1974, incorporates the known quarks and leptons into common families. For example, in the standard model the three colour states (6.39) of the down quark, which are conveniently denoted

$$(d_r, d_g, d_b), \tag{11.2a}$$

can be converted into each other by gluon emissions as illustrated in Figure 11.2(a), while the positron and antineutrino

$$(e^+, \bar{\nu}_e) \tag{11.2b}$$

can be converted into each other by $W^\pm$ emissions as illustrated in Figure 11.2(b). In the Georgi–Glashow model both sets of particles are assigned to a single family[4]

[4] Strictly speaking, it is the 'right-handed' states defined in Section 10.1.2 that are associated in a single family, but we will ignore such spin complications in this simple account.

$$(d_r, d_g, d_b, e^+, \bar{\nu}_e) \tag{11.3}$$

and, in addition to the familiar process of Figures 11.2(a) and (b), the quarks and leptons can convert into each other by processes like those shown in Figures 11.2(c) and (d). These involve the emission of two new gauge bosons X and Y with electric charges $-\frac{4}{3}$ and $-\frac{1}{3}$, respectively, and masses of order $M_X \approx 10^{15}\,\mathrm{GeV/c^2}$. At the unification mass, all the processes of Figure 11.2 are characterized by a single 'grand unified coupling constant' g_U, whose value is such that the analogue of the fine structure constant is

$$\alpha_U \equiv \frac{g_U^2}{4\pi} \approx \frac{1}{42}. \tag{11.4}$$

However, at energies $E \ll M_X$, processes involving the exchange of the X and Y bosons are heavily suppressed because of their large masses, in the same way that $W^\pm$ exchange processes are suppressed relative to electromagnetic ones at energies $E \ll M_W$ in the unified electroweak theory.[5] Because of this, processes involving the exchange of X and Y particles are difficult, but perhaps not impossible, to observe at presently attainable energies, as we shall see in the next section. Firstly, however, we briefly comment on two other characteristics of this simplest of all grand unification schemes.

11.1.1 Quark and lepton charges

An attractive feature of the Georgi–Glashow model is that it offers an explanation of one of the longest-standing mysteries in physics–the equal magnitudes of the electric charges of the electron and proton. In the standard model this equality is guaranteed by assigning electric charges $2e/3$ and $-e/3$ to the u and d quarks, respectively, where $-e$ is the charge of the electron, but no explanation is given for these empirical assignments. However, in the Georgi–Glashow model it can be shown that the sum of the electric charges of all the particles in any given family must be zero. Applying this to (11.3) gives

$$3Q_d + e = 0,$$

where Q_d is the charge of the down quark. Hence Q_d is determined to be $-e/3$ in terms of the electron charge, and the mysterious factor of 3 is seen to be a consequence of the fact that the quarks have three distinct colour states. The charge of the up quark Q_u is shown to be $2e/3$ by a similar argument, and the equality of the proton charge

$$Q_p = 2Q_u + Q_d = e$$

and the positron charge follows from the usual quark assignment $p = uud$.

[5] This was discussed in Section 2.2.1.

11.1.2 The weak mixing angle

When the strong and electroweak interactions are extrapolated to the unification mass M_X, they are characterized by a single coupling constant g_U, as illustrated in Figure 11.1. Conversely, the three effective low-energy couplings (11.1) of the standard model can be expressed in terms of the two parameters g_U and M_X by reversing this extrapolation. Consequently, in grand unified theories one of the three low-energy coupling constants (11.1) can be predicted, given the values of the other two. It is conventional to convert this result into a prediction of the weak mixing angle θ_W, which is related to the coupling constants (11.1b) by (9.8), and in the Georgi–Glashow model this can be shown to give the value

$$\sin^2\theta_W = 0.21. \tag{11.5}$$

This value is encouragingly close to the measured value (9.14), but is not in precise agreement with it.[6] We shall return to this point in Section 11.2.

11.1.3 Proton decay

The most striking prediction of grand unified theories is that the proton is unstable, and in the Georgi–Glashow model it can decay by a variety of processes involving the exchange of X and Y bosons and their antiparticles $\bar{X}$ and $\bar{Y}$. These are built from a number of basic vertices, two of which are shown in Figures 11.2(c) and (d). If we restrict ourselves to the family (11.3) and their antiparticles, the model predicts three more such vertices, which are shown in Figure 11.3, and there are also another five, which can be obtained from Figures 11.2(c) and (d) and Figure 11.3 by replacing all particles by their antiparticles. These give rise to proton decays such as

$$p \to \pi^0 + e^+ \quad \text{and} \quad p \to \pi^+ + \bar{\nu}_e\,, \tag{11.6}$$

by mechanisms like those shown in Figure 11.4. In all such processes, both the baryon and lepton numbers are not conserved, but the combination

$$B - L \equiv B - \sum_\alpha L_\alpha \quad (\alpha = e, \mu, \tau) \tag{11.7}$$

is conserved, as may be verified directly from the vertices of Figure 11.2 and Figure 11.3.[7]

Although the proton is predicted to decay via processes like (11.6), its expected lifetime is extremely long, and before quoting detailed values it is useful to make a qualitative estimate to understand why this is so. The first step is to estimate the magnitude of the effective coupling constant G for processes involving the exchange of X and Y bosons at energies $E \ll M_X$. At these energies, X and Y exchanges may be approximated by point interactions as shown in Figure 11.5, where on the left-hand

[6] Another way of putting this is to say that if the measured value (9.14) is imposed the three curves in Figure 11.1 fail to meet at the same point.

[7] This is done explicitly in Problem 11.1.

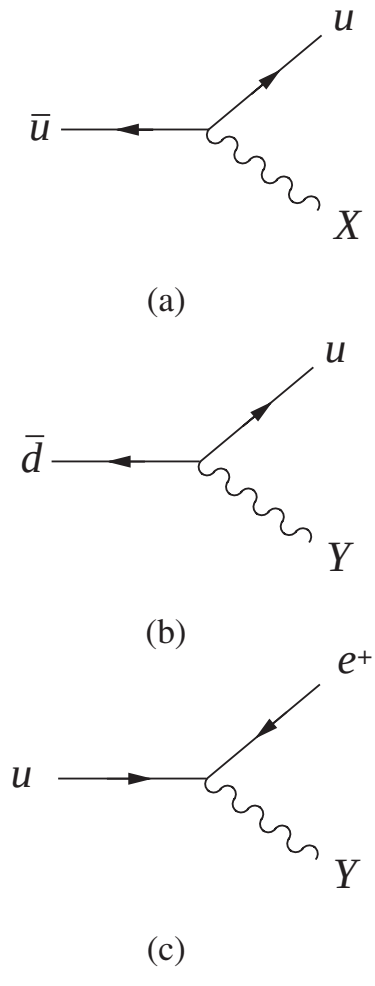


Figure 11.3 The three fundamental vertices involving X and Y bosons that are predicted by the Georgi–Glashow model, in addition to those shown in Figures 11.2 (c) and (d).

side we have approximated the couplings of the X and Y bosons to the lepton and quarks by their value g_U at the unification mass M_X, and ignored their dependence on the range of the interaction. The effective low-energy coupling is then given by

$$\frac{G}{\sqrt{2}} \approx \frac{g_U^2}{M_X^2} = \frac{4\pi\alpha_U}{M_X^2}, \tag{11.8}$$

by analogy with (2.17) for the Fermi coupling, and like it has dimension $[E^{-2}]$ in natural units. Since the rate Γ of any X or Y exchange process has a natural dimension $[E]$ and is proportional to G^2, a simple dimensional estimate gives

$$\Gamma \approx G^2 E^5 \approx \frac{g_U^2 E^5}{M_X^4},$$

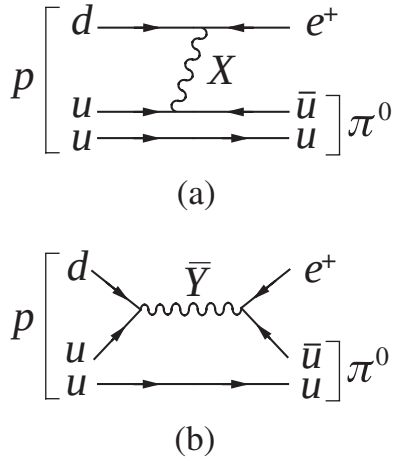


Figure 11.4 Examples of processes that contribute to the proton decay reaction (11.6).

where $E \ll M_X$ is some energy characteristic of the process. A reasonable estimate for proton decay is $E = M_p$, and so we finally arrive at

$$\tau = \Gamma^{-1} \approx \frac{M_X^4}{g_U^2 M_p^5} \tag{11.9}$$

for the lifetime of the proton in natural units. Taking M_X to be $10^{15}\,\mathrm{GeV/c^2}$ and g_U from (11.4) gives an estimate of 10^{29} years, in agreement with the range of values

$$\tau = 10^{29} - 10^{30}\,\text{years} \tag{11.10a}$$

obtained from more detailed and reliable calculations. This value is, however, very sensitive to the value chosen for the unification mass M_X, as can be seen from (11.9); in other grand unified theories in which M_X is somewhat larger, the lifetime can be as long as

$$\tau = 10^{32} - 10^{33}\,\text{years}. \tag{11.10b}$$

For comparison, the age of the universe is believed to be of order 10^{10} years.

To detect proton decays with lifetimes the size of those in (11.10) requires a very large mass of detector material. For example, 300 tons of iron would only yield about one proton decay per year if the lifetime were of order 10^{32} years. Several large detectors of various types have been built, including the SuperKamiokande detector described in Section 2.3.2, but no clear example of a proton decay event has been observed, enabling upper limits to be placed on the rates of various possible decay modes. For the $\pi^0 e^+$ mode this can be expressed in the form

$$\frac{\tau_p}{B(p \to \pi^0 e^+)} > 5 \times 10^{32}\,\text{years}, \tag{11.11}$$

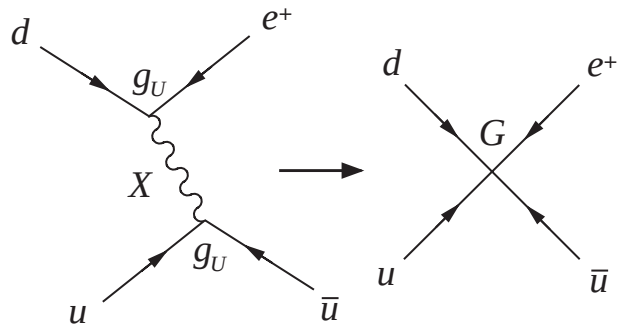


Figure 11.5 The zero-range approximation to an X-boson-exchange process.

where $B(p \to \pi^0 e^+)$ is the branching ratio. In the Georgi–Glashow model this branching ratio is predicted to be about 0.3, so that (11.11) is clearly incompatible with the predicted lifetime (11.10a). However, other models, which retain the successes of the Georgi–Glashow model, predict longer lifetimes (11.10b), and are compatible with the proton decay data, as we shall see in Section 11.2.

11.2 SUPERSYMMETRY

The problems with the predictions for the proton lifetime (11.10a) and the weak mixing angle (11.5) in grand unified theories can be resolved by incorporating a new proposed symmetry of nature called *supersymmetry*. According to this, every known elementary particle has a supersymmetric partner (called a *superpartner*), which is like it in all respects except for its spin. Leptons and quarks have spin-0 superpartners, while spin-1 bosons, like photons, have spin-$\frac{1}{2}$ superpartners. The supersymmetric partners of fermions are named by adding a prefix 's' to the name of the fermion, while the superpartners of the bosons are named by adding the ending 'ino' to the root of the normal name. This is illustrated in Table 11.1, where we list the various particles, together with their superpartners and the corresponding spins.

TABLE 11.1 The particles of the standard model and their superpartners

Particle	Symbol	Spin	Superparticle	Symbol	Spin
Quark	q	$\frac{1}{2}$	Squark	$\tilde{q}$	0
Electron	e	$\frac{1}{2}$	Selectron	$\tilde{e}$	0
Muon	μ	$\frac{1}{2}$	Smuon	$\tilde{\mu}$	0
Tauon	τ	$\frac{1}{2}$	Stauon	$\tilde{\tau}$	0
W boson	W	1	Wino	$\tilde{W}$	$\frac{1}{2}$
Z boson	Z	1	Zino	$\tilde{Z}$	$\frac{1}{2}$
Photon	γ	1	Photino	$\tilde{\gamma}$	$\frac{1}{2}$
Gluon	g	1	Gluino	$\tilde{g}$	$\frac{1}{2}$
Higgs	H	0	Higgsino	$\tilde{H}$	$\frac{1}{2}$

If supersymmetry were exact, a particle and its superpartner would have exactly the same mass. This is obviously not realized in nature or superparticles would have been detected long ago. Supersymmetry is at best only an approximate symmetry. However, the degree of symmetry-breaking is restricted by another motivation for supersymmetry, the so-called hierarchy problem. In the absence of supersymmetry, radiative corrections give rise to contributions to the mass of the Higgs boson that are of the same order as the unification scale, and a correspondingly huge Higgs mass can only be avoided by rather artificial 'fine tuning' of parameters. However, fine tuning is avoided in supersymmetric models by automatic cancellations between the contributions of particles and their superpartners, provided that their masses are not too different. This is assumed to be the case in most supersymmetric versions of grand unified theories in which the lightest supersymmetric particles have masses that are of the same order of magnitude as the $W^{\pm}$ and Z^0 masses. When these supersymmetric particles are taken into account, it can be shown that the extrapolation of Figure 11.1 is modified in such a way that the grand unification mass M_X is increased to a value of order $10^{16}\,\mathrm{GeV/c^2}$, while the value of the grand unified coupling constant g_U remains relatively constant. As a consequence of this, the proton lifetime increases, as expected from (11.9), to a value of order 10^{32}–10^{33} years, which is consistent with present experimental limits. At the same time, the prediction (11.5) of the weak mixing angle is slightly modified to yield a value in agreement with the experimental value (9.14).

11.2.1 *CP* violation and electric dipole moments

Supersymmetric theories, with their host of new particles, contain several new unknown mixing angles and *CP*-violating phases, and by adjusting these parameters it is possible to induce much larger *CP*-violating effects than those predicted in the standard model. One way to explore this experimentally is to search for nonzero electric dipole moments (EDMs) of leptons and hadrons.

Consider the nonrelativistic Hamiltonian H_{EDM} for the interaction of an EDM $\mathbf{d}$ with an electric field $\mathbf{E}$. For an elementary particle, or a hadron, its spin $\mathbf{J}$ is the only vector available to define a direction and so $\mathbf{d}$ must be collinear with $\mathbf{J}$. Thus we can write

$$H_{EDM} = -\mathbf{d}\cdot\mathbf{E} = -\frac{d}{J}\mathbf{J}\cdot\mathbf{E},$$

where $d = |\mathbf{d}|$ and $J = |\mathbf{J}|$. However, $\mathbf{E}$ is even under time reversal, while $\mathbf{J}$ is odd, and thus H_{EDM} is odd under time reversal. It is also odd under parity. Thus an EDM can only exist if both parity (P) and time-reversal (T) invariance are violated and by implication any nonzero EDM detected would imply *CP* violation, assuming that *CPT* invariance holds.

The magnitudes of predicted electric dipole moments depend upon the assumed mechanism that causes *CP* violation. For example, in the standard model a nonzero EDM could be generated by P, T violating weak radiative corrections to the P, C and T conserving electromagnetic interaction. For the neutron and the electron, which are the two most extensively studied, the estimates from the standard model are

$$d_n \sim 10^{-32}\,e\,\mathrm{cm} \quad \text{and} \quad d_e \leq 10^{-38}\,e\,\mathrm{cm}, \tag{11.12}$$

where $e = 4.8 \times 10^{-10}$ esu is the magnitude of the charge on the electron. For the electron EDM, for example, this is 16 orders of magnitude smaller than the electron magnetic moment and is far too small to be measured in the foreseeable future. However, in supersymmetric theories much larger *CP*-violating effects can be generated by diagrams involving virtual superparticles, like Figure 11.6(a) for the electron. In comparison, the corresponding particle diagram of Figure 11.6(b) does not violate *CP* invariance in the standard model, and a nonzero electron EDM can only arise in higher orders in perturbation theory and is consequently much smaller.

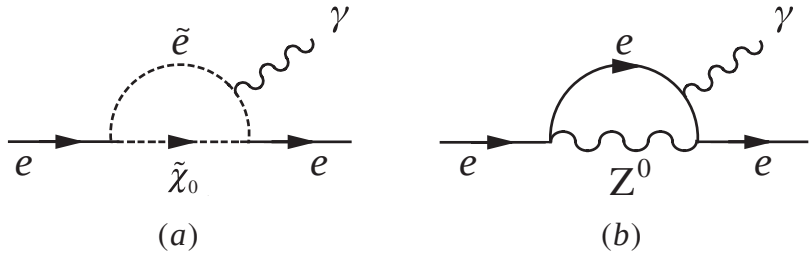


Figure 11.6 (a) An example of a diagram involving superparticles that can lead to a nonzero electron electric dipole moment in supersymmetric theories, together with (b) a corresponding diagram without superparticles, which is *CP*-conserving in the standard model.

For neutral spin-$\frac{1}{2}$ particles, nonzero EDMs can be detected by making use of the fact that in collinear **E** and **B** fields, the particle's spin will undergo a Larmor precession[8] about the field direction. The frequency f of this precession is given by

$$hf = (2\mu_n B \pm 2d_n E), \tag{11.13}$$

where μ_n is the magnetic moment, d_n is the electric dipole moment and the $\pm$ sign refers to E being parallel (antiparallel) to B. Thus, a nonzero d_n will be revealed by an electric field-dependent shift in f. The experiments are technically very demanding, because of the relative sizes of μ_n and d_n, and innovative techniques are required to detect the frequency shift. The same technique cannot be used directly for charged particles, such as the electron, because the interaction with the electric field will accelerate the particle out of the observed region, and in practice the best measurements of the EDM of the electron have been found from studies of paramagnetic atoms and molecules with unpaired electrons. For the neutron and the electron, the present experimental limits (at 95 % confidence level) are

$$|d_n| < 3 \times 10^{-26}\, e\,\text{cm} \quad \text{and} \quad |d_e| < 1.6 \times 10^{-27}\, e\,\text{cm}. \tag{11.14}$$

These values are much larger than those predicted by the standard model, but are getting very close to testing the predictions of some supersymmetric models.

[8] A particle with a magnetic moment $\boldsymbol{\mu}$ placed in an external magnetic field **B** will experience a torque T given by $\mathbf{T} = \boldsymbol{\mu} \times \mathbf{B} = \gamma \mathbf{J} \times \mathbf{B}$, where **J** is the angular momentum vector and γ is the gyromagnetic ratio. As a result, the angular momentum vector precesses about the external field axis with frequency $f = \mu B / Jh$, known as the *Larmor frequency*.

Furthermore, experiments planned and in progress should reduce the limits (11.14) by factors of order 100 and 10, respectively. Such experiments will either detect nonzero EDMs for the first time, implying a new source of *CP* violation, as required to explain the matter–antimatter asymmetry observed in the present universe (to be discussed in Section 11.4.2 below), or lead to more stringent upper bounds and hence useful constraints on theories beyond the standard model.

11.2.2 Detection of superparticles

To definitively verify supersymmetry, it will of course be necessary to detect the predicted superparticles. Unfortunately, it is difficult to say very much about their masses because the nature of supersymmetry breaking is unknown, and even the simplest and most popular supersymmetric model – the so-called minimal supersymmetric standard model (MSSM) – contains many unknown parameters. However, according to most versions of the theory, including the MSSM, superparticles can only be created or destroyed in pairs. Hence the decay of a superparticle must yield at least one superparticle in the final state, and the lightest such particle must be stable.[9]

There are several candidates for the identity of the lightest superparticle, but most models assume it is a neutralino $\tilde{\chi}^0$ which is a mixture of the photino, Higgsino and zino.[10] If we accept this, then a simple reaction that can be studied is

$$e^+ + e^- \rightarrow \tilde{e}^+ + \tilde{e}^-, \tag{11.15}$$

followed by the decays

$$\tilde{e}^\pm \rightarrow e^\pm + \tilde{\chi}^0, \tag{11.16}$$

giving an overall reaction

$$e^+ + e^- \rightarrow e^+ + e^- + \tilde{\chi}^0 + \tilde{\chi}^0. \tag{11.17}$$

Here $\tilde{\chi}^0$ denotes the neutralino, introduced above, and other superparticles are identified by placing tildes over the symbols for the corresponding particles; thus, for example, $\tilde{e}$ is a selectron. The theory makes three clear predictions for this reaction sequence:

1. The cross-section for producing selectron pairs via (11.15) should be comparable to that for producing ordinary charged particle pairs of the same mass.
2. The selectrons decay very rapidly by (11.16) before they can reach the detector.
3. The neutralinos should escape detection because they interact extremely weakly with ordinary matter.

[9] Because of this, the lightest superparticle is a possible contributor to the so-called dark matter in the universe, to be discussed in Section 11.4.1.

[10] The photino, Higgsino and zino are all neutral spin-$\frac{1}{2}$ particles that interact by electroweak forces only. They can mix together in just the same way as neutrinos.

Hence only the final state electron and positron should be detected in (11.17), and the events would be characterized by the fact that on average these will account for only half the energy of the initial state. Moreover, the electron and positron will not in general be emitted in opposite directions in the centre-of-mass frame, as they would be in the case of a two-body final state. Thus there are clearly defined criteria for the identification of selectron pair production events, provided of course that there is enough energy to produce them. A similar discussion applies to the production of other pairs of charged superparticles in electron–positron collisions, while both charged and neutral superparticles can also be produced at hadron colliders. We shall not discuss this further, but will merely note that the failure to detect any such events sets lower limits on the masses of the different superparticles, which vary somewhat but are typically in the range 40–100 GeV/c^2. This is of limited significance since their masses are expected to be order M_W or more, and the decisive search for supersymmetry will probably take place at the LHC from 2009 onwards.

Finally, superparticles are not the only new particles predicted by grand unified and supersymmetric theories, because the Higgs mechanism is more complicated in such theories and requires more than the single Higgs boson of the standard model. The details vary between models, but in the MSSM, for example, three neutral and two charged Higgs bosons are required, together with the corresponding spin–$\frac{1}{2}$ Higgsinos. Thus the search for Higgs bosons is itself an important part of the search for physics beyond the standard model.

11.3 STRINGS AND THINGS

Supersymmetry is an important component in even more ambitious schemes to unify gravity with the other forces of nature at an enormous *superunification* scale, where gravitational interactions are comparable in strength with those of the grand unified strong and electroweak interactions. The problems here are mathematically formidable, not the least of them being that the divergences encountered in trying to quantize gravity are far more severe than those in either QCD or the electroweak theory and there is at present no successful 'stand-alone' quantum theory of gravity analogous to the former two. In order to resolve this problem, the theories that have been proposed invariably replace the idea of point-like elementary particles with tiny quantized one-dimensional *strings*, and for reasons of mathematical consistency are formulated in many more dimensions (usually 10 but sometimes 11, including one time dimension) than we observe in nature. Such theories have a single free parameter – the string tension. However, we live in a four-dimensional world and so the 'extra' dimensions have to be 'compactified', i.e. reduced to an unobservably small size. It was originally hoped that in doing this the standard model, with its many free parameters (masses of quarks and leptons, coupling constants, mixing angles, etc.), would somehow emerge from string theory as a unique low-energy four-dimensional theory and thus the precise values of the parameters of the model would be explainable in terms of just a single parameter, the string tension.

Early optimism has not been sustained. In the particle picture, the structure of the corresponding quantum field theory (such as QED or QCD) is known and physical predictions may be obtained using the appropriate Feynman rules. However, in string theory the structure is not known, and there are five sets of possible Feynman rules,

each operating in a ten-dimensional space–time continuum. Furthermore, string theorists have discovered that, far from being unique, there is a vast 'landscape' of at least 10^{500} (!) possible low-energy theories that could result after compactification, each corresponding to a universe with a different set of fundamental particles, interactions and parameters. Unless there is a method of choosing between the vast possibilities offered by this 'landscape', string theories have little or no real predictive power. For this reason they have generated a lively philosophical debate as to whether they should even be considered as scientific theories, although their proponents claim that string theories are being judged by standards that historically have not been applied to other emergent theories.

One controversial approach to the question of choice has been to invoke the so-called 'anthropic principle'. There are various forms of this, but essentially it states that what we can expect to observe must be restricted by the conditions necessary for our presence as observers. In other words, the world is observed to be the way it is because that is the only way that humans could ever be here to consider such questions in the first place. This circular-sounding 'principle' has been invoked by cosmologists to explain the apparent improbable values of some cosmological constants, but it is by no means generally accepted as a way forward for string theories and other theorists believe that some form of dynamical selection will eventually be possible. All one can say at present is that there is no consensus on how the problem of choice is to be solved.

The self-consistency of string theories in 10 dimensions has been shown to imply the existence of higher-dimensional objects, called *branes* (short for membranes), and it has been conjectured that using these it will be possible to construct an even more fundamental theory in 11 dimensions in which all five supersymmetric string theories are unified. This theory even has a name – *M-theory* – although no one knows if the conjecture is true, or how to construct such a theory. Nevertheless, string theory has provided some powerful theoretical tools that have contributed to a better understanding of gauge theories and their relation to gravity.

Leaving aside the mathematical difficulties of string theories, a major practical problem is that they apply at an energy scale where gravitational effects are comparable to those of the gauge interactions, i.e. at energies defined by the so-called *Planck mass* M_P, given by

$$M_P = \left(\frac{\hbar c}{G}\right)^{1/2} = 1.2 \times 10^{19}\,\mathrm{GeV/c^2}, \tag{11.18}$$

where G is the gravitational constant.[11] This energy is so large that it is difficult to think of a way that the theories could be tested at currently accessible energies, or even indeed at energies accessible in the conceivable future, although some theorists believe that information produced even at the 'low' energies of the LHC (e.g. the discovery of superparticles) may help to test string theories. Despite this, the appeal of string theories at present is mainly the mathematical beauty and 'naturalness' that their sponsors claim for them. Needless to say, experimentalists will remain sceptical until definite experimental tests can be suggested and carried out.

[11] This implies that strings have dimensions of order $x_P \sim \hbar/M_p c = 1.6 \times 10^{-35}$ m.

11.4 PARTICLE COSMOLOGY

Modern theories of particle physics and cosmology are closely related. This is because the conditions in the early universe that are implied by the standard cosmological theory, known as the big bang model, can only be approached, even remotely, in high-energy particle collisions. At the same time, laboratory experiments cannot reach anything like the energies at which we might expect to see the full effects of unification, so the conditions that existed shortly after the formation of the universe constitute a unique laboratory for the study of unification models. Here we will concentrate on just two or three aspects of this connection.

11.4.1 Dark matter: WIMPs and neutrinos

The modern description of the universe is based on the observation that it is expanding and assumes that the origin of this is a sudden explosion at some time in the past. For this reason the description is called the *big bang model*. Because the universe appears isotropic at large distance scales, there can be no preferred points in space and so the big bang must have occurred everywhere at once, thus ensuring that the expansion appears the same to all observers irrespective of their locations. Two pieces of evidence for this model are the existence of a cosmic background radiation (CMB), now known to be very accurately represented by a black-body spectrum at an effective temperature of 2.7 K, and the cosmic abundance of light elements.[12] Whether the expansion will continue indefinitely depends on the average density of the universe ρ. The critical density ρ_c at present times, below which the expansion will continue indefinitely and above which it will eventually halt and the universe will start to contract, can be written

$$\rho_c = \frac{3H_0^2}{8\pi G} \sim 10^{-26}\,\mathrm{kg\,m^{-3}} \approx 5.1\,(\mathrm{GeV/c^2})\mathrm{m^{-3}}, \tag{11.19}$$

where G is the gravitational constant and we have used the current value for Hubble's constant H_0 to evaluate (11.19). In the most popular version of the model, called the *inflationary big bang model*, the relative density

$$\Omega \equiv \rho/\rho_c = 1. \tag{11.20}$$

The relative density is conveniently written as the sum of three components,

$$\Omega \equiv \Omega_{total} = \Omega_r + \Omega_m + \Omega_\Lambda, \tag{11.21}$$

where Ω_r is the contribution due to *radiation*, Ω_m is that due to *matter* and Ω_Λ, the so-called *vacuum* density, is related to a term in the equation governing the evolution of the universe that contains a so-called *cosmological constant* Λ. It is also referred to as *dark energy*. Of these terms, only Ω_r has been accurately directly measured, from an analysis of the precisely known form of the CMB. Numerically it is $\Omega_r \approx 5 \times 10^{-5}$, i.e.

[12] For an accessible discussion of the big bang model and other matters discussed in this section see, for example, Perkins (2003).

negligible. The total matter contribution Ω_m can be deduced from the gravitational energy needed for consistency with observations on the rotation of galaxies and the kinematics of large-scale structures in the universe. The values obtained from such analyses are in the range 0.24–0.30 and are consistent with the values found from measurements of distant Type Ia supernova. Some individual contributions to Ω_m can also be estimated. Thus the contribution of luminous baryonic matter is obtained from the observed matter in the form of stars and intergalactic gas and dust, and is about 0.01. The total baryonic contribution to Ω_m may be inferred from knowledge of how nuclei are formed in the universe (*nucleosynthesis*) and its value is $\Omega_b \approx 0.05$. Finally, the vacuum term can be estimated from various cosmological observations, including recently measured minute temperature fluctuations in the microwave background radiation. Its value is about 0.7 and is the largest contribution to Ω_{total}. Thus we see that the value of Ω_{total} is consistent with (11.20), although the uncertainties are considerable. This conclusion is supported by very detailed measurements of the fluctuations in the microwave background. Analysis of data from the WMAP satellite yield the value

$$\Omega_{total} = 1.003 \pm 0.017. \tag{11.22}$$

There are two unsatisfactory features of the decomposition (11.21). One is that the origin of the largest term, Ω_Λ, is totally unknown. The second is that from the above it follows that most matter is nonluminous and the proportion that is baryonic is only a small fraction, about (15–20) % of the total matter contribution. There could be other sources of nonluminous baryonic matter, e.g. in the form of brown dwarfs and small black holes the size of planets, and there is experimental evidence that such 'massive, compact halo objects' (MACHOs) do indeed exist in the halo of our galaxy. However, it is not thought that they alone can account for the 'missing' matter. Thus we are forced to conclude that the bulk of matter, as much as 85 %, is nonbaryonic. It is referred to collectively as *dark matter*.

There are several dark matter candidates. Massive neutrinos might be one possibility. Such particles would have to be heavy enough to have been nonrelativistic in the early stages of the universe (so-called *cold dark matter*), because if they were relativistic (*hot dark matter*) they would have rapidly dispersed, giving rise to a uniform energy distribution in space. Calculations suggest that in this case there would have been insufficient time for the observed galaxies to have formed. The contribution of massive neutrinos to the matter term can be calculated once the number of species and their masses are known. For masses in the range 5×10^{-4} eV to 1 MeV, the contribution of neutrinos to Ω_m is

$$\Omega_\nu = \sum m_\nu / 49\,\mathrm{eV/c^2}, \tag{11.23}$$

so using $\Omega_\nu < \Omega_m - \Omega_b$ gives

$$\sum m_\nu \leq (10 - 12)\,\mathrm{eV/c^2}. \tag{11.24}$$

This bound is not very useful. However, it can be greatly improved. This is because neutrinos with masses as small as 0.1 eV can have an observable effect on the formation of large-scale structure in the universe because free-streaming neutrinos dampen the growth of perturbations. This enables an upper limit to be placed on Ω_ν/Ω_m. Two major surveys of large-scale structure exist and detailed analyses of their data using the standard cosmological model with a small number of parameters give values for the sum of neutrino masses. Although some groups claim a value as low as $0.2\,\mathrm{eV/c^2}$, taking account of the different assumptions made about some of the fixed parameters in the model and the various data sets used, a reasonable range is

$$\sum m_\nu \leq (0.5-1.0)\,\mathrm{eV/c^2}. \tag{11.25}$$

Even allowing for the uncertainties, (11.25) is still lower than the upper bound from tritium decay. If the analyses leading to (11.25) are correct, then neutrinos play a minor role in contributing to the matter deficit.

The most popular hypothesis is that the bulk of dark matter consists of nonbaryonic cold dark matter in the form of 'weakly interacting massive particles', or WIMPs. The problem is that there are no known candidates for WIMPs, so new types of particle have to be postulated. One possibility is the lightest particle that appears in supersymmetric theories. In most such theories this is a stable neutralino which interacts only via electroweak interactions. Its mass is not precisely predicted, but must certainly be in excess of $10\,\mathrm{GeV/c^2}$. If such theories are correct, these particles would have been produced in large numbers following the big bang, and could well account for the bulk of dark matter.

If WIMPs are the dominant form of dark matter, their number can be inferred from the estimated density of dark matter. From estimates of the dark matter density in our own galaxy, it follows that of order 10^{13} WIMPs per second will pass through every kilogram of matter on Earth if their mass is assumed to be of order $15\,\mathrm{GeV/c^2}$. Their velocities relative to the galaxy will be similar to those of baryonic matter, which are of order 10^{-3} c, and when they interact with nuclei, the nuclei will recoil with energies up to about 50 keV. Several dedicated experiments have been mounted to detect such recoils, which can in principle be done in a number of ways. For example, in semiconductors such as GaAs, free charge will be produced that can be detected electronically; in a scintillator such as NaI the emission of photons can be detected using photomultipliers and in crystals at low temperatures the energy can be converted to phonons that can be detected by a very small rise in temperature. In practice the problems are formidable because of the very low expected event rate. For example, if WIMPs are identified with neutralinos, then expectations range from 1 to 10 events per kilogram of detector per week. This is very small compared to the event rate from naturally occurring radioactivity, including that in the materials of the detectors themselves. The former is minimized by working deep underground to shield the detector from cosmic rays and in areas with geological structures where radioactive rocks are absent; the latter is minimized by building detectors of extreme purity. Finally, any WIMP signal obtained is expected to have a very distinctive time dependence. This is a consequence of the Earth's motion around the Sun. The Sun moves in the galactic plane with a speed of order $2 \times 10^5\,\mathrm{m\,s^{-1}}$. The Earth moves

around the Sun with a speed order of $3 \times 10^4\,\mathrm{m\,s^{-1}}$, in an orbit that is inclined at about 60^0 to the galactic plane, as seen in Figure 11.7. Hence the velocity of the Earth through dark matter varies by about $\pm 15\,\%$ over the course of the year. This leads to a corresponding modulation in the flux and energy of WIMP collisions, resulting in a maximum event rate in June and a minimum in December. One experiment claims to have seen this variation. However, experiments are still at an early stage, although some versions of supersymmetric theories with low-mass neutralinos can probably already be ruled out.[13]

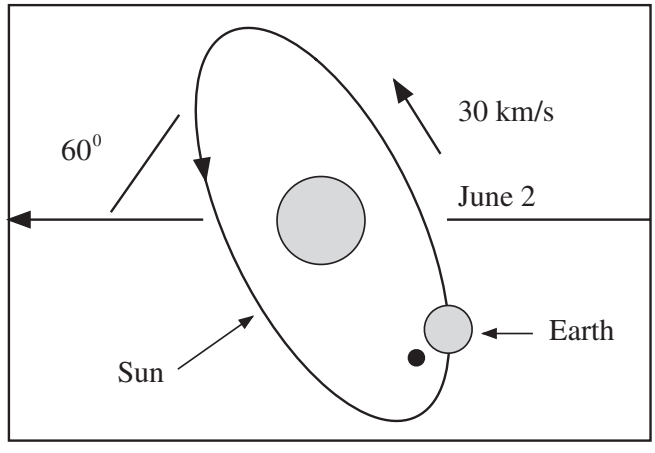


Figure 11.7 Motion of the Sun and Earth through the dark matter background. (N. Spooner, *Astronomy Now*, **7**, 40 (1993). With permission.)

11.4.2 Matter–antimatter asymmetry

One of the most striking facts about the universe is the paucity of antimatter compared to matter. There is ample evidence for this. For example, cosmic rays are overwhelmingly composed of matter and what little antimatter is present is compatible with its production in intergalactic collisions of matter with photons. Neither do we see intense outbursts of electromagnetic radiation that would accompany the annihilation of clouds of matter with similar clouds of antimatter. The absence of antimatter is very puzzling, because in the original big bang it would be natural to assume a total baryon number $B = 0$.[14] Then during the period when kT was large compared to hadron energies, baryons and antibaryons would be in equilibrium with photons via reversible reactions such as

$$p + \bar{p} \rightleftharpoons \gamma + \gamma. \tag{11.26}$$

[13] Reviews of the status of dark matter searches are given in Perkins (2003) and Yao *et al.* (2006).

[14] One could, of course, simply bypass the problem by arbitrarily assigning an initial nonzero baryon number to the universe, but it would have to be exceedingly large to accommodate the observed asymmetry, as well as being an unaesthetic solution.

This situation would continue until the temperature fell to a point where the photons no longer had sufficient energy to produce $p\bar{p}$ pairs and the expansion had proceeded to a point where the density of protons and antiprotons was such that their mutual annihilation became increasingly unlikely. The critical temperature is $kT \approx 20\,\text{MeV}$ and at this point the ratios of baryons and antibaryons to photons 'freezes' to values that can be calculated to be

$$N_B/N_\gamma = N_{\bar{B}}/N_\gamma \sim 10^{-18}, \tag{11.27}$$

with of course $N_{\bar{B}}/N_B = 1$. These ratios would then be maintained in time, whereas the ratios currently observed are

$$N_B/N_\gamma \approx 10^{-9}, \quad N_{\bar{B}}/N_\gamma \sim 10^{-13}, \tag{11.28}$$

with $N_{\bar{B}}/N_B \sim 10^{-4}$. The simple big bang model fails spectacularly.

The conditions whereby a baryon–antibaryon asymmetry could arise were first stated by Sakharov. It is necessary to have: (a) an interaction that violates charge conjugation C and the combined symmetry CP, (b) an interaction that violates baryon number and (c) a nonequilibrium situation that must exist at some point to 'seed' the process. The reason for the first condition is that a baryon excess cannot be generated if the production of any particular particle (e.g. the proton) is balanced by an equal production of the corresponding antiparticle (the antiproton), as required by C or CP conservation. CP violation has already been observed, as we saw in Chapter 10, and the other conditions can be realized within the context of supersymmetry and the big bang model. However, the source and size of CP violation in the standard model are not compatible with that required for the observed matter–antimatter asymmetry in the universe and we must conclude that there is another as yet undetected source of CP violation 'beyond the standard model' if this explanation is correct. Clearly, the observed matter–antimatter asymmetry remains a serious unsolved problem.

11.5 NEUTRINO ASTRONOMY

Neutrino astronomy, by which we mean the use of neutrinos rather than electromagnetic waves and cosmic rays to explore the universe, is of growing interest. One reason is that neutrinos travel to the Earth from their point of origin without significant further interaction, enabling such regions to be more directly explored. Studies began with the observation of solar neutrinos. This was discussed in Chapter 2, where we emphasized its important role in the study of neutrino mixing and neutrino masses. However, the energy distribution of the neutrino flux, after the oscillations have been unravelled, also serves as a direct test of the solar model itself.

Another and very different source of neutrinos is a Type II supernova. Despite the spectacular visual ouput from these cataclysmic events, it is believed that about 99 % of the total energy released is in the form of neutrinos. According to astrophysical theory, these consist of an intense burst of electron neutrinos with energies of a few MeV, which lasts for a few milliseconds, after which neutrinos of all flavours, with average energies of about 15 MeV, will be emitted in all directions over a period of 0.1–10 s. The first experiments that detected such neutrinos were Kamiokande, which

was an earlier version of SuperKamiokande described in Chapter 2, and the IMB collaboration, which also used a water Čerenkov detector. Both had been constructed to search for proton decay as predicted by grand unified theories, but by good fortune both detectors were 'live' in 1987 at the time of a spectacular supernova explosion (now named SN1987A) and both detected a small number of antineutrino events. The Kamiokande experiment detected 12 $\bar{\nu}_e$ events and the IMB experiment 8 events, both over a time interval of approximately 10 s and with energies in the range 0–40 MeV. The data are shown in Figure 11.8. When the energy acceptances of the detectors are taken into account, these values are consistent with the estimates for the energies of neutrinos that had been produced by the initial pulse in the supernova, and which then diffused out of the collapsed core and escaped.

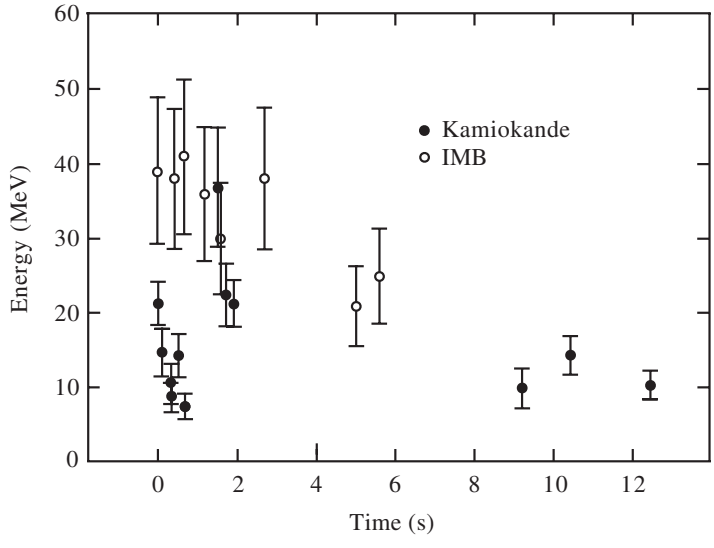


Figure 11.8 Data for neutrinos from SN1987A detected in the Kamiokande and IMB experiments. The threshold for detecting neutrinos in the experiments are 6 MeV (Kamiokande) and 20 MeV (IMB). In each case the first neutrino detected is assigned the time zero.

The neutrinos from SN1987A were of low energy, but there is also great interest in detecting ultra-high-energy neutrinos. For example, it is known that there exist point sources of γ rays with energies in the TeV range, many of which have their origin within so-called 'active galactic nuclei'. It is an open question whether this implies the existence of point sources of neutrinos with similar energies. The neutrinos to be detected would be those travelling upwards through the Earth, as upward-travelling muons originating from pion decay in the atmosphere would be absorbed on their passage through the Earth, whereas those travelling downwards would swamp the muon signal originating from cosmic neutrinos. Although ultra-high-energy neutrinos are also attenuated as they travel through the Earth, the signal from upward-travelling muons is largely background free, because the flux from atmospheric neutrinos at these energies is negligible. Like all weak interactions the intrinsic rate would be very low, especially so for such rare high-energy events, but this is partially compensated by the fact that the ν–nucleon cross-section increases with energy, as discussed in Section 7.5.

To detect neutrinos in the TeV energy range using the Čerenkov effect in water requires huge volumes, orders-of-magnitude larger than that used in the SuperKamiokande detector. An ingenious solution to this problem is to use the vast quantities of water available in liquid form in the oceans, or frozen in the form of ice at the South Pole, and several experiments have been built, or are being built, using these sources. The largest so far is the Antartic muon and neutrino detector array (AMANDA), which is sited at the geographical South Pole. A schematic diagram of this detector is shown in Figure 11.9.

The detector consists of strings of 'optical modules' containing photomultiplier tubes that convert the Čerenkov radiation to electrical signals. The enlarged inset in Figure 11.9 shows the details of an optical module. They are located in the ice at great depths by using a novel hot water boring device. The ice then refreezes around them. In the first phase of the experiment in 1993/94 (AMANDA-A) four detector strings were located at depths of between 800 and 1000 m. The ice at these depths is filled with air bubbles and so the detectors are not capable of precision measurements, but they proved the validity of the technique. In the next phase a few years later (AMANDA-B10), 10 more strings containing 320 optical modules were located at depths between 1.5 and 2.0 km, where the properties of ice are suitable for muon detection. Finally, the current version of the detector (AMANDA-II) has an additional nine strings

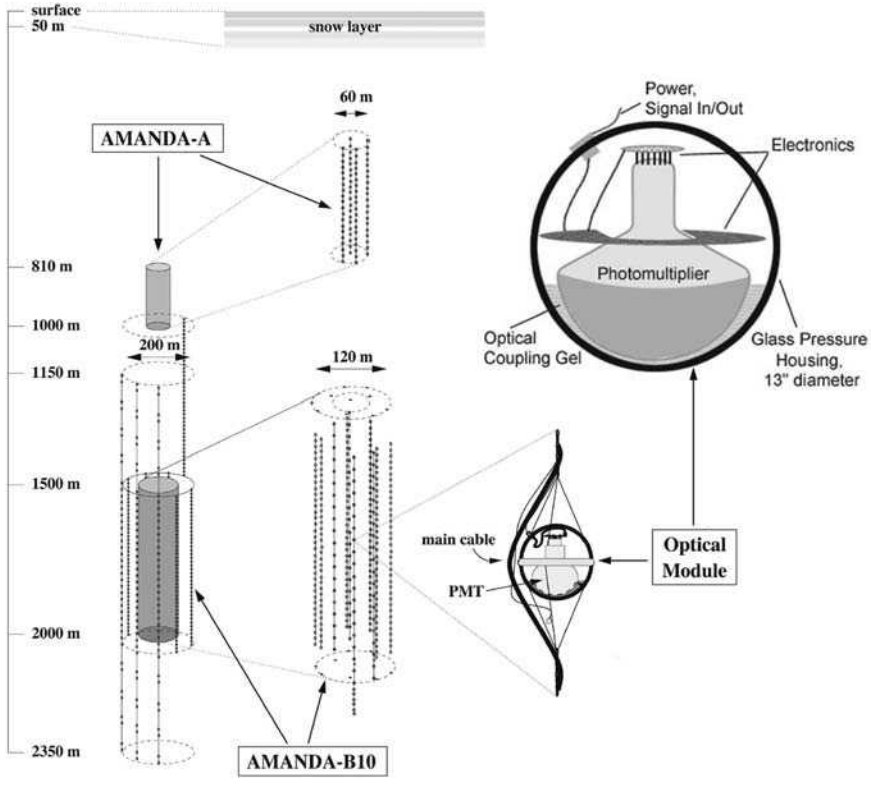


Figure 11.9 A schematic diagram of the AMANDA neutrino detector. (Courtesy of the AMANDA Collaboration.)

extending to a depth of 2.35 km. In total there are 680 optical modules covering a cylindrical volume with a cross-sectional diameter of 120 m. The AMANDA detector has successfully detected atmospheric neutrinos and has produced the most detailed map of the high-energy neutrino sky to date. However, no source of continuous emission has yet been observed that would be a candidate for a point source.

AMANDA can detect neutrinos with energies up to about 10^{15} eV, but an even bigger detector, called IceCube, is under construction at the South Pole. This will eventually use about 80 strings each containing 60 optical modules regularly spaced over an area of 1 km^2 and at depths of between 1.4 and 2.4 km. The volume covered by AMANDA is only 1.5 % of the volume to be covered by IceCube, which will be capable of detecting neutrinos with energies as high as 10^{18} eV. Completion is expected in 2011, although some experiments can be performed before this date.

Detection of even higher energy neutrinos is being pursued in several innovative experiments. One of these is called ANITA (Antarctic impulsive transient antenna). Its primary interest is to address the nature of ultra-high-energy cosmic rays ($E > 10^{19}$ eV) that have been observed over many years, but the origin of which is unknown. It seeks to do this by detecting the associated neutrinos produced by the interaction of the cosmic rays with the cosmic microwave photons that pervade the universe. To do this it exploits an effect similar to the Čerenkov effect that we have met in previous chapters. In this case, a particle travelling faster than the speed of light in a dense radio transparent medium produces a shower of charged particles that contain a charge anisotropy, and thus emits a cone of coherent radiation in the radio or microwave part of the electromagnetic spectrum. This is the Askaryan effect, predicted by Askaryan in 1965 and first confirmed, using sand, in 2000. It was confirmed in ice in 2006 as part of the preliminary work for ANITA. The neutrinos cascading through the Antarctic ice sheet lead to a strong electromagnetic pulse that propagates through the ice, because the latter is transparent to radio waves up to a frequency of 1.5 GHz. The ice sheet is thus a converter of neutrino energy to radio waves. The experiment consists of a detector system mounted on a balloon platform at a height of about 40 km about the Antarctic ice shelf. The balloon traverses a circumpolar flight path due to the continuous wind circulation around the South Pole and sees the ice below out to the horizon at about 700 km. Thus the effective telescope lens has an area of approximately 1.5×10^6 m^2! Experiments such as IceCube and ANITA will ensure that neutrino astronomy will be of great interest for many years to come.

11.6 DIRAC OR MAJORANA NEUTRINOS?

In Section 1.2.2 we noted that neutral particles might, or might not, have distinct antiparticles. Thus there are two possible descriptions of neutral leptons, i.e. neutrinos. Neutrinos that have distinct antiparticles are called *Dirac neutrinos*. There are then two neutrino states

$$\nu_L, \nu_R \quad (L = 1)$$

and two antineutrino states

$$\bar{\nu}_L, \bar{\nu}_R \quad (L = -1),$$

with opposite values of the lepton number L, which is conserved in the standard model. Neutrinos that do not have distinct antiparticles are called *Majorana neutrinos*, and there are only two states, denoted by

$$\nu_L,\ \nu_R.$$

In both cases, the subscripts L and R denote right- and left-handed helicity states as usual.

In the original formulation of the standard model, neutrino masses were assumed to be zero and it is not possible to distinguish experimentally between Dirac and Majorana neutrinos for two reasons. Firstly, for Dirac neutrinos with zero mass, the weak interaction only couples the left-handed neutrinos ν_L and right-handed anti-neutrinos $\bar{\nu}_R$, as we saw in Chapter 10. Secondly, in the zero-mass limit, the Dirac equation decouples into two two-component equations describing the states $(\nu_L, \bar{\nu}_R)$ and $(\nu_R, \bar{\nu}_L)$, respectively. The second pair can therefore be completely eliminated from the theory, which becomes indistinguishable from an analogous Majorana theory with the replacements $(\nu_L, \bar{\nu}_R) \to (\nu_L, \nu_R)$.

In contrast, as we saw in Section 10.1.3, the weak interaction couples to both helicity states of the electron, which is not massless. Similarly, for nonzero masses, the weak interaction would also couple to the other two Dirac neutrino states, albeit with relative couplings of order $(m_\nu/E)^2$, where E is the neutrino energy. Because of this, differences of the same order will emerge between the two descriptions of the neutrino, which can, at least in principle, be detected experimentally. In particular, in the Majorana formalism, lepton number conservation for reactions involving neutrinos is an 'accident' arising from the spin structure of the weak interaction for zero mass neutrinos, and tiny deviations from it are expected for nonzero neutrino masses. In most cases these will be too small to observe, and the most promising prospect for detecting them, if they exist, is by detecting neutrinoless double-beta decay. This will be discussed in Section 11.6.2. First, however, we digress to mention a way in which very heavy Majorana neutrinos may play a role.

11.6.1 The seesaw mechanism

At present we have no theoretical understanding of why quarks and leptons have the particular masses they do. However, the three known neutrinos have much lower masses than the other fundamental fermions, and a possible explanation for this has been suggested in the context of grand unified theories. In these, it is possible for both types of neutrino to co-exist, and for a very small neutrino mass to emerge from the mixing of a zero-mass neutrino with a very heavy Majorana neutrino. The corresponding mass matrix for the right-handed neutrino is essentially of the form

$$\mathbf{M} = \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix}, \tag{11.29}$$

where the off-diagonal term m_D, called the Dirac mass, is of the order of the electroweak scale and m_M, the Majorana mass, is of the order of the GUT scale. Since $m_M \gg m_D$, the eigenvalues of $\mathbf{M}$ are

$$|\lambda_+| \approx m_M \quad \text{and} \quad |\lambda_-| \approx m_D^2/m_M. \tag{11.30}$$

If λ_- is associated with the observed neutrinos, we have a natural explanation for a very small mass of order 1 eV or less, while the other eigenvalue implies a very heavy neutrino, yet to be discovered. This is called the 'seesaw mechanism', because from (11.30) we see that as one mass goes up, the other goes down.

11.6.2 Double-beta decay

The double-β decay process

$$(Z,A) \to (Z+2,A) + 2e^- + 2\bar{\nu}_e \tag{11.31}$$

(denoted by the label $\beta\beta 2\nu$), in which two neutrons decay to give two electrons and two antineutrinos, is allowed in the standard model, as illustrated in Figure 11.10(a). In contrast, the neutrinoless double-β decay process (labelled $\beta\beta 0\nu$)

$$(Z,A) \to (Z+2,A) + 2e^- \tag{11.32}$$

is not allowed because it violates lepton number. However, it can occur for Majorana neutrinos of nonzero mass, which are their own antiparticles, by the mechanism illustrated in Figure 11.10(b). Observation of the reaction (11.32) would therefore be strong evidence for the existence of Majorana neutrinos, although in principle the reaction could also proceed by some other unknown mechanism beyond the standard model. On the other hand, the $\beta\beta 2\nu$ process of (11.31) is allowed for both Dirac and Majorana neutrinos.

Because the double-β decay process (11.31) is second-order in the weak interaction, with a rate of order G_F^4, it can in practice only be observed if the single-β decay process

$$(Z,A) \to (Z+1,A) + e^- + \bar{\nu}_e \tag{11.33}$$

is forbidden. For an even-Z, even-$N = A - Z$ nucleus, double-β decay leads to another even–even nucleus $(Z+2,A)$, while single-β decay leads to an odd–odd nucleus $(Z+1,A)$. Since odd–odd nuclei are less stable than even–even nuclei, due to the pairing term in the semi-empirical mass formula,[15] a sequence of states can result in which single-β decay is forbidden by energy conservation, while double-β decay is allowed. If the parent nucleus (A,Z) is also stable against α and γ decay, it will decay by double-β decay with a mean lifetime in the range $10^{18} - 10^{24}$ year, depending on the nucleus concerned. Such decays were first directly observed in 1987, and have now been established for the 10 isotopes: ^{48}Ca, ^{76}Ge, ^{82}Se, ^{96}Zr, ^{100}Mo, ^{116}Cd, ^{128}Te, ^{130}Te, ^{150}Nd, and ^{238}U.

Neutrinoless double-β decay (11.32), which has not yet been observed, can be distinguished from double-β decay (11.31) by measuring the energies of the emitted electrons. In $\beta\beta 2\nu$ decays, energy is carried away by the undetected neutrinos, resulting in a continuous spectrum for the combined energy of the electrons, whereas in

[15] See, for example, Section 2.6 of Martin (2006).

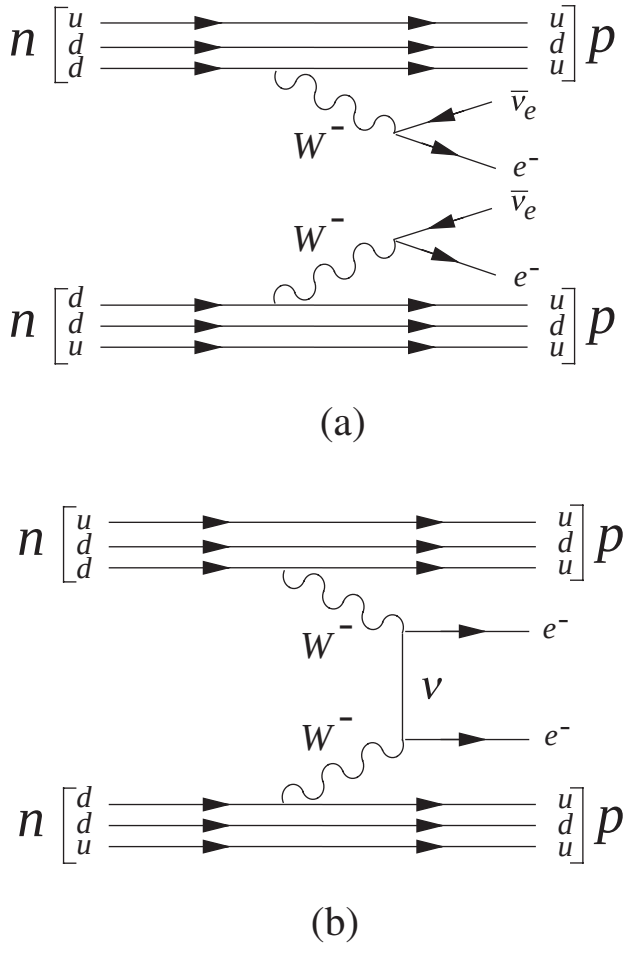


Figure 11.10 (a) Double-β decay $\beta\beta 2\nu$, as allowed in the standard model. (b) Neutrinoless double-βdecay $\beta\beta 0\nu$, forbidden in the standard model.

$\beta\beta 0\nu$ decays the electrons carry off all the available energy, resulting in a sharp peak in their combined energy, as shown in Figure 11.11. However, the big problem is that the rate for $\beta\beta 0\nu$ decay is expected to be much smaller that that for $\beta\beta 2\nu$ decays, even for Majorana neutrinos, and for zero-mass neutrinos it would actually vanish. For nonzero mass Majorana neutrinos, the decay rate Γ is given by

$$\Gamma = a \langle m \rangle^2 , \tag{11.34}$$

where the constant of proportionality a is itself proportional to the product of the squared nuclear matrix element and the coupling constant G_F^4. The quantity $\langle m \rangle$ is the square of the 'effective Majorana mass', which reduces to

$$\langle m \rangle = |U_{e1}|^2 m_1 + |U_{e2}|^2 m_2 + |U_{e3}|^2 m_3$$

if, for simplicity, we neglect *CP*-violating phases in the neutrino sector. Here U_{e1}, U_{e2} and U_{e3} are the elements of the neutrino mixing matrix (10.58) that defines the electron–neutrino state and m_1, m_2 and m_3 refer to the mass eigenstates ν_1, ν_2 and ν_3, respectively.

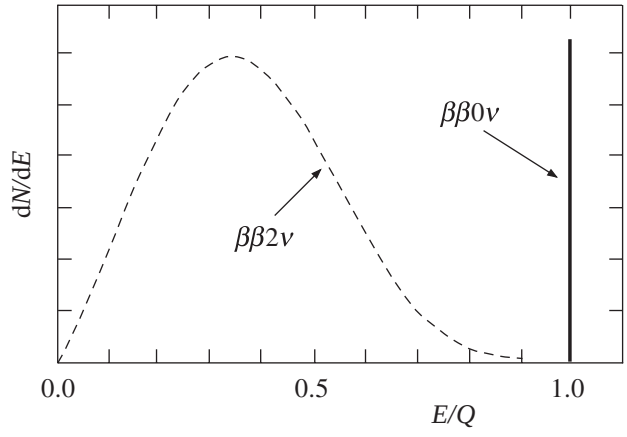


Figure 11.11 Energy spectra for the two electrons in $\beta\beta 2\nu$ (dashed line) and $\beta\beta 0\nu$ (solid line) as a function of E/Q, where E is their combined energy and Q is the energy released.

The constants of proportionality a in (11.34) for $\beta\beta 0\nu$ are somewhat uncertain, due to uncertainties in the nuclear structure, but for $\langle m \rangle \sim 1\,\text{eV}$, they lead to decay rates of order one per year, or less, per kg of unstable isotope. For Dirac neutrinos, the predicted decay rate is of course zero, irrespective of the mass. Because the background counting rate must be even lower than the predicted very low counting rates for a signal, $\beta\beta 0\nu$ experiments are ultra-low background experiments. For this reason they are invariably located deep underground, to shield them from cosmic rays, and they must also be shielded to eliminate background arising from ambient radiation in the surroundings. In addition, the sample of decaying isotope must be extremely pure, since even a very small contamination of a β-decaying impurity would swamp the signal from double-β decay. Since several kilograms of an isotope are needed to obtain a detectable counting rate, this is a highly nontrivial requirement.

The experiments can be roughly divided into three types. If the decaying isotope is a semiconductor, such as ^{76}Ge, the isotope can be both the sample and a solid-state detector that will measure the energy released in a given decay. The GERDA experiment and the planned COBRA experiment, both located at the Gran Sasso Laboratory in Italy, are of this type. In another type of experiment in which the sample and detector are bolometers, the energy released in the decay is turned into heat and detected. The CUORE experiment, also at Gran Sasso, is of this type. Finally, the decaying sample may be surrounded by a separate detector that can observe and identify the electron tracks from the decays, as well as measure their energies. An example of this type is the NEMO3 detector, shown in Figure 11.12, which is located in the Fréjus Tunnel beneath Mont Blanc in the French Alps. In contrast to the combined sample–detector experiments, this experiment can study a range of isotopes, in the form of thin sheets located in a central tower, including approximately 7 kg of ^{100}Mo

and 1 kg of ^{82}Se. The isotopes are surrounded by multiwire drift chambers to record the electron tracks, electromagnetic calorimeters to measure their energy and a magnetic coil to provide a field for charge information. This in turn is surrounded by pure iron shielding to eliminate γ rays, and wood and pure borated water shielding to eliminate neutrons. A much larger experiment based on the same technology, called SuperNEMO, will hopefully start taking data in 2010 on the $\beta\beta 2\nu$ and $\beta\beta 0\nu$ decays of ^{82}Se and ^{150}Nd, with isotope samples in the range 100–200 kg.

At the time of writing, there is no confirmed evidence for neutrinoless double-β decay, and current experiments can be interpreted as setting a limit of 0.5 eV = 500 meV on the effective Majorana mass. Experiments currently under construction or in progress, such as CUORE, GERDA and SuperNEMO, will reduce this limit by a factor of order 10, to approximately 50 meV, if no events are detected. On the other hand, a positive result would be unambiguous evidence for new physics beyond the standard model. It would also give an indication of the scale of neutrino masses, as opposed to the mass differences, which are measured in neutrino oscillations.

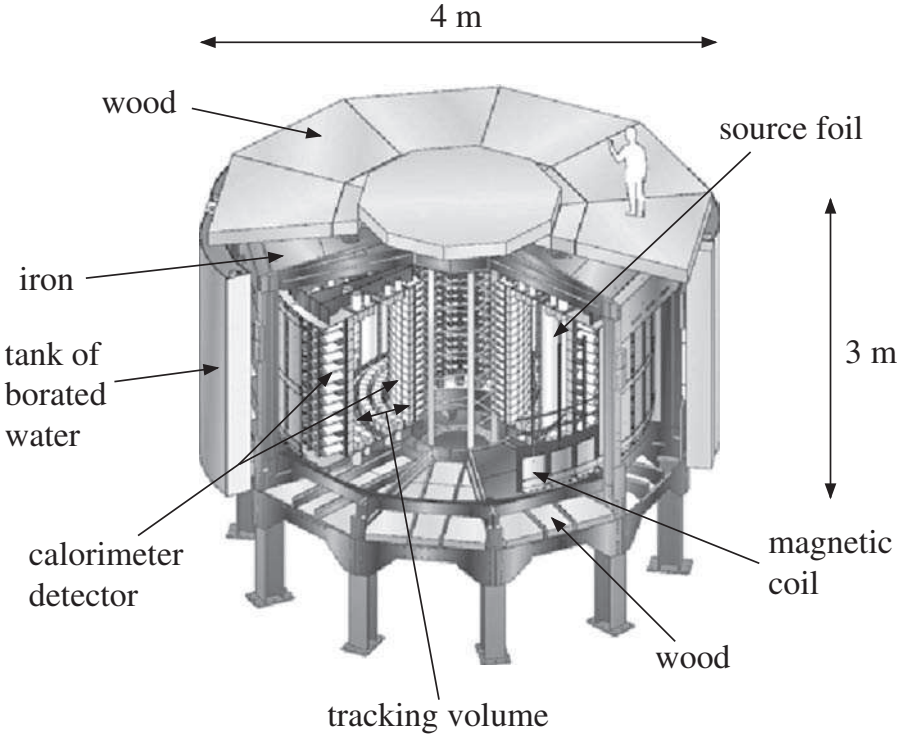


Figure 11.12 Schematic diagram of the NEMO3 detector: neutron and gamma shielding is provided by a layer of iron, 40 cm of wood top and bottom and a 30 cm thick cylindrical tank of borated water around the circumference; the tracking detector consists of 6180 wire drift chamber cells operating in Geiger mode in a helium and alcohol gas mixture; the calorimeter detector consists of 1940 plastic scintillator blocks coupled to low-radioactivity photomultiplier tubes; the magnetic coil produces a vertical field of 25 Gauss; and the source contains 10 kg of double beta isotopes. (Courtesy of the NEMO3 Collaboration.)

PROBLEMS 11

11.1 Show that it is possible to assign values of $B - L$ (B is the baryon number and L is the lepton number) to the X and Y bosons in such a way that $B - L$ is conserved for all the fundamental vertices of Figures 11.2(c) and (d) and Figure 11.3. What are the values of $B - L$ required?

11.2 Following the supernova explosion 1987A at a distance $d = 1.5 \times 10^{21}$ m, a burst of $\bar{\nu}_e$ interactions was observed in a terrestrial detector. The interactions occurred over a time interval of 7 s and the antineutrinos had a range of energies between 7 and 11 MeV. Estimate a reasonable upper limit on the $\bar{\nu}_e$ mass.

11.3 Find two possible mechanisms for the decay $p \to \pi^+ + \bar{\nu}_e$ by suitably relabelling the lines in Figures 11.4(a) and (b). Check that electric charge Q and the difference $B - L$ are both conserved at the vertices in the resulting diagrams, using the values for the latter quantity deduced in Problem 11.1.

11.4 Consider a star moving with velocity v in a circular orbit of radius R about the centre of a spiral galaxy. Calculate the dependence of v on R for the following extreme cases:

(a) The total mass of the galaxy, like the luminous mass, is concentrated almost entirely at the centre of the galaxy.
(b) The mass of the galaxy is distributed almost entirely in a spherical halo of dark matter, assumed to be of constant density and radius $R_h > R$.

Compare your results with the observation that v is almost independent of R for stars observed in the outer arms of spiral galaxies.

11.5 Show that the electric field **E** is odd under parity but even under time reversal, as stated in the discussion of electric dipole moments in Section 11.2.1. How does the magnetic field **B** behave under P and T?

11.6 Consider a collision between a cosmic ray proton and a photon in the cosmic microwave background whose energy corresponds to the ambient temperature of 2.7 K. Estimate the minimum proton energy required to produce (a) an e^+e^- pair and (b) a proton–antiproton pair.

A

Relativistic Kinematics

Particle physics utilizes relativistic kinematics in many ways: to relate energies, momenta and scattering angles in different frames of reference; to deduce the masses of unstable particles from measurements on their decay products; to work out threshold energies for the production of new particles; and in a variety of other applications. In this appendix we briefly illustrate such calculations and introduce the concept of *invariant mass*, which can often be used to simplify them. The reader is assumed to be familiar with the basic theory of special relativity, as discussed, for example, in Chapters 11 to 13 of Kleppner and Kolenkow (1973).

A.1 THE LORENTZ TRANSFORMATION FOR ENERGY AND MOMENTUM

If a particle has coordinates $(\mathbf{r}, t) \equiv (x, y, z, t)$ in an inertial frame S, then its coordinates $(\mathbf{r}', t')$ in a second inertial frame S' moving with uniform velocity relative to S are given by a Lorentz transformation. If the two frames coincide at $t = 0$ and frame S' is moving with uniform speed $v = |\mathbf{v}|$ in the positive x direction relative to S, this takes the form[1]

$$x' = \gamma(v)\,(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(v)\,(t - vx/c^2), \tag{A.1}$$

where $\gamma(v) \equiv (1 - \beta^2)^{-1/2}$ and $\beta \equiv v/c$ as usual. The transformation laws for energy and momentum are then obtained as follows.

[1] In this appendix we retain factors of c throughout. The corresponding equations in natural units are trivially obtained by setting $c = 1$.

Particle Physics Third Edition B. R. Martin and G. Shaw

Consider a particle i of rest mass m_i moving with uniform velocity $\mathbf{u}_i$ in S, so that its energy E_i and momentum $\mathbf{p}_i$ are given by[2]

$$E_i = m_i c^2 \gamma(u_i), \quad \mathbf{p}_i m_i \gamma(u_i), \tag{A.2}$$

where $u_i = |\mathbf{u}_i|$. For simplicity, we will take $\mathbf{u}_i$ to be in the x direction, i.e. in the same direction as $\mathbf{v}$. Then it is straightforward to show from the definition of velocity and the transformation (A.1) that the velocity of the particle in S' is[3]

$$u_i' = \frac{u_i - \upsilon}{1 - u_i \upsilon / c^2}. \tag{A.3}$$

Hence

$$\gamma(u_i') \equiv \left(1 - u_i'^2/c^2\right)^{-1/2} = \gamma(u_i)\gamma(\upsilon)\left(1 - u_i \upsilon / c^2\right)$$

and the energy and momentum in the S' frame are given by

$$E_i' = m_i c^2 \gamma(u_i') = \gamma(\upsilon)\,(E_i - \upsilon p_i) \tag{A.4a}$$

and

$$p_i' = m_i u_i' \gamma(u_i') = \gamma(\upsilon)\,(p_i - \upsilon E_i / c^2), \tag{A.4b}$$

using (A.2) and (A.3). These results are easily generalized to an arbitrary direction for the particle's velocity and momentum, when they become

$$(p_i')_x = \gamma(\upsilon)[(p_i)_x - \upsilon E_i / c^2], \tag{A.5a}$$

$$(p_i')_y = (p_i)_y,\ (p_i')_z = (p_i)_z \tag{A.5b}$$

and

$$E_i' = \gamma(\upsilon)[E_i - \upsilon (p_i)_x]. \tag{A.5c}$$

Furthermore, since Equations (A.5) are linear in both energy and momentum, they imply the corresponding results

$$p_x' = \gamma(\upsilon)(p_x - \upsilon E / c^2), \tag{A.6a}$$

$$p_y' = p_y, \quad p_z' = p_z \tag{A.6b}$$

---

[2] Energy always means the total energy, including the rest energy, as distinct from the kinetic energy $T_i \equiv E_i - m_i c^2$.

[3] This is just the usual transformation law for velocities; see, for example, Section 12.4 of Kleppner and Kolenkow (1973).

and

$$E' = \gamma(\upsilon)(E - \upsilon p_x), \tag{A.6c}$$

for the total energy and momentum

$$E = E_1 + E_2 + \cdots + E_N, \quad \mathbf{p} = \mathbf{p}_1 + \mathrm{p}_2 + \cdots + \mathbf{p}_N, \tag{A.7}$$

of N particles that are sufficiently far apart for their mutual interaction energies to be neglected.

A.2 THE INVARIANT MASS

The *invariant mass* W of a set of N particles is defined by

$$W^2c^4 \equiv E^2 - \mathbf{p}^2c^2, \tag{A.8}$$

where E and $\mathbf{p}$ are the total energy and momentum given by (A.7). As its name implies, W has the same value in any reference frame, since

$$(E')^2 - (\mathbf{p}')^2c^2 = (E)^2 - (\mathbf{p})^2c,$$

as may be verified from (A.5). It is most easily evaluated in the centre-of-mass (CM) frame where

$$\mathbf{p} = \sum_i \mathbf{p}_i = \mathbf{0}, \tag{A.9}$$

so that

$$Wc^2 = E_{CM} \tag{A.10a}$$

is the total energy in the CM frame. For a single particle i, the invariant mass W_i is identical with the rest mass m_i, i.e.

$$W_i = m_i, \tag{A.10b}$$

since the CM frame is identical with the rest frame and E_{CM} is just m_ic^2.

Many applications of relativistic kinematics can be simplified by evaluating the invariant mass in the frame of interest using (A.7) and (A.8), and equating the results to the values obtained in the CM frame (A.10a) and (A.10b). We will illustrate this by some examples.

A.2.1 Beam energies and thresholds

The reference frames most commonly met in particle physics are those of the laboratory and the centre-of-mass, and as the first example we will relate the energies in these two frames. Specifically, we shall consider a beam of particles b with momentum

$\mathbf{p}_L$ incident upon target particles t, which are at rest in the laboratory, so that $\mathbf{p}_t = \mathbf{0}$. The energies of the particles in the laboratory frame are

$$E_L = (m_b^2c^4 + \mathbf{p}_L^2c^2)^{1/2}, \quad E_t = m_tc^2 \tag{A.11}$$

and the invariant mass W is given by

$$W^2c^4 = (E_L + m_tc^2)^2 - \mathbf{p}_L^2c^2 = m_b^2c^4 + m_t^2c^4 + 2m_tc^2E_L,$$

using Equations (A.7) and (A.8). On comparing with Equation (A.10a) we immediately obtain

$$E_{CM} = (m_b^2c^4 + m_t^2c^4 + 2m_tc^2E_L)^{1/2}, \tag{A.12}$$

which is the result of Equation (4.1) quoted for the total centre-of-mass energy E_{CM} in Chapter 4.

The total centre-of-mass energy is important, because it determines the energy available for particle production. For example, the antiproton production reaction

$$p + p \rightarrow p + p + p + \bar{p} \tag{A.13}$$

can only occur if

$$E_{CM} \geq 4m_Pc^2. \tag{A.14}$$

Here the minimum value corresponds to all four final state particles produced at rest, which is only possible in the centre-of-mass frame (A.9). In the laboratory frame, the beam energy E_{min} corresponding to the minimum centre-of-mass energy $4m_Pc^2$ is given by

$$4m_Pc^2 = (2m_P^2c^4 + 2m_Pc^2E_{min})^{1/2}$$

from (A.12) with $m_b = m_t = m_P$, leading to

$$E_L \geq E_{min} = 7m_Pc^2 = 6.6\,\text{GeV} \tag{A.15}$$

as the range of beam energies corresponding to (A.14). Consequently, in an experiment in which a beam of protons is incident upon a static proton target, the beam energy must be at least 6.6 GeV if antiprotons are to be produced. Here $E_{min} = 6.6\,\text{GeV}$ is called the threshold energy for antiproton production,[4] and the corresponding thresholds for the production of other particles may be calculated in a similar way. In doing this, account must be taken of the appropriate conservation laws in identifying the lightest

[4] This energy is somewhat reduced if a nuclear target is used, owing to the motion of the target nucleons within the nucleus (see Problem A.5). Antiprotons were first detected by Chamberlain, Segré, Wiegland and Ypsilantis in 1955 using a proton beam incident upon a copper target. A simple account of this classic experiment may be found, for example, in Section 3.9 of Hughes (1985).

final state containing the desired particle. For example, for positive and negative pion production in proton–proton collisions, the lightest final states that conserve baryon number and electric charge are

$$p + p \rightarrow p + n + \pi^+ \tag{A.16a}$$

and

$$p + p \rightarrow p + p + \pi^+ + \pi^-. \tag{A.16b}$$

These lead to threshold energies of 1.23 and 1.54 GeV for π^+ and π^- mesons, respectively, corresponding to beam momenta of 0.79 and 1.22 GeV/c. In practice, to get observable counting rates, one must work at energies somewhat above threshold.

A.2.2 Masses of unstable particles

A simple but important use of the invariant mass is in determining the masses of very short-lived particles by measurements on their decay products. If a particle A decays to N particles $i = 1, 2, \ldots, N$ in the final state, i.e.

$$A \rightarrow 1 + 2 + \ldots + N,$$

then the invariant mass (A.8) of the final state particles is given by

$$W^2c^4 = \left(\sum_i E_i\right)^2 - \left(\sum_i \mathbf{p}_i c\right)^2 = E_A^2 - (\mathbf{p}_A c)^2 = m_A^2 c^4 \tag{A.17}$$

using energy–momentum conservation. Hence the mass of the decaying particle is equal to the invariant mass of its decay products, and the latter can be used to determine the former if the particle is too short-lived for its mass to be measured directly.

A.3 TRANSFORMATION OF THE SCATTERING ANGLE

As a final example of relativistic kinematics, we consider the transformation of scattering angles between the laboratory frame, where they are measured in fixed-target experiments, and the centre-of-mass frame, which is almost invariably used in theoretical discussions. In doing this we will find that for very high beam energies the scattering angles in the laboratory tend to be very small, and most of the particles produced in the interaction emerge in a narrow cone about the incoming beam direction.

We consider a reaction of the general form

$$b(E_L, \mathbf{p}_L) + t(m_t c^2, \mathbf{0}) \rightarrow P(E, \mathbf{q}) + \ldots, \tag{A.18}$$

where

$$E = (m_p^2 c^4 + \mathbf{q}^2 c^2)^{1/2} \tag{A.19}$$

and the dots indicate that one or more particles are produced along with the particle P. Without loss of generality, we can choose axes so that

$$\mathbf{p}_L = (p_L, 0, 0), \quad \mathbf{q} = (q\cos\theta_L, q\sin\theta_L, 0), \tag{A.20}$$

i.e. so that the beam is in the x direction and the momentum $\mathbf{q}$ lies in the xy plane. The angle θ_L between the beam direction and the direction of the particle P is called the *production angle*, or *scattering angle*, and we want to find its value in the centre-of-mass frame.

The first step is to find the velocity υ of the Lorentz transformation that takes us from the laboratory frame to the centre-of-mass frame. For any Lorentz transformation along the beam direction (i.e. the x axis) the beam and target momenta become

$$p_b' = \gamma(\upsilon)\left(p_L - \upsilon E_L/c^2\right), \quad p_t' = -m_t\upsilon\gamma(\upsilon)$$

by (A.4), while in the centre-of-mass frame

$$p_b' + p_t' = 0$$

by definition. Combining these results gives

$$\upsilon = \frac{p_L c^2}{E_L + m_t c^2} \tag{A.21}$$

as the velocity of the transformation.

The next step is to work out the momentum $\mathbf{q}'$ of the produced particle P in the centre-of-mass frame and hence the production or scattering angle between it and the beam direction. From (A.5) the components of the momentum are

$$q_x' = \gamma(\upsilon)\,(q\cos\theta_L - \upsilon E/c^2), \quad q_y' = q\sin\theta_L, \quad q_z' = 0, \tag{A.22}$$

where υ is given by (A.21). Writing

$$\mathbf{q}' = (q'\cos\theta_C,\, q'\sin\theta_C,\, 0), \tag{A.23}$$

where $q' = |\mathbf{q}'|$ and θ_C is the angle between $\mathbf{q}'$ and the x axis, we immediately obtain

$$\tan\theta_C = \frac{q_y'}{q_x'} = \frac{1}{\gamma(\upsilon)}\frac{q\sin\theta_L}{q\cos\theta_L - \upsilon E/c^2}. \tag{A.24}$$

Since the beam direction is along the x axis, this is the desired result, which relates the centre-of-mass scattering angle θ_C to measured quantities in the laboratory frame.

The dominance of small laboratory scattering angles θ_L at high energies follows from the inverse of (A.24),

$$\tan\theta_L = \frac{1}{\gamma(\upsilon)}\frac{q'\sin\theta_C}{q'\cos\theta_C + \upsilon E'/c^2}, \tag{A.25}$$

where

$$E' = (m_P^2 c^4 + q'^2 c^2)^{1/2}$$

is the energy of the produced particle in the centre-of-mass frame. This result follows from (A.24), since to go from the centre-of-mass to the laboratory frame requires the inverse transformation to (A.22) with the velocity reversed, i.e. $\upsilon \to -\upsilon$. At high energies

$$E_L \approx p_L c \gg m_b c^2,\, m_t c^2,$$

so that

$$\upsilon \approx c(1 - m_t c/p_L) \approx c$$

by (A.21) and

$$\gamma(\upsilon) = (1 - \upsilon^2/c^2)^{-1/2} \approx (p_L/2m_t c)^{1/2}.$$

On substituting these values into (A.25), together with

$$E' = m_P c^2 \gamma(u), \quad q' = m_P u \gamma(u),$$

where u is the magnitude of the particle's velocity in the centre-of-mass frame, we obtain

$$\tan\theta_L \approx \left(\frac{2m_t c}{p_L}\right)^{1/2} \frac{u \sin\theta_C}{u\cos\theta_C + c} \tag{A.26}$$

as the appropriate form at high energies. The presence of the factor $p_L^{-1/2}$ then implies that θ_L will be very small, unless $u \approx c$ and $\cos\theta_C \approx -1$. Because these conditions are not usually satisfied, most particles are produced in a narrow cone about the beam direction in the laboratory frame, as we have asserted. A similar result applies to decay angles: when a high-energy particle decays, the decay products emerge predominantly at very small angles to the initial beam. This is demonstrated by a very similar argument to the above, and is left as an exercise for the reader (see Problem A.7).

PROBLEMS A

A.1 Derive an expression for the muon energy in the decay $\pi^+ \to \mu^+ + \nu_\mu$ in terms of the pion, muon and neutrino masses for the case of pions at rest.

A.2 A particle of mass m decays in flight to give two photons of energies E_1 and E_2. Show that the opening angle θ between the photon directions is given by

$$\cos\theta = 1 - m^2 c^4/2E_1 E_2.$$

(Several short-lived particles – e.g. $\pi^0(140)$, $\eta^0(547)$– can decay in this way, and since they have different masses the above relation gives a simple means of identifying them.)

A.3 Calculate the lowest energy at which $\Lambda(1116)$ baryons can be produced in the strong interaction of negative pions with protons at rest.

A.4 A neutral particle X^0 decays to two charged particles A^+ and B^-. The momentum components of the decay products in GeV/c are measured to be:

	p_x	p_y	p_z
A^+	−0.488	−0.018	2.109
B^-	−0.255	−0.050	0.486

Is the decay $K_S^0 \to \pi^+\pi^-$ or $\Lambda \to p\pi^-$?

A.5 Show that when the motion of nucleons within the nucleus is taken into account, the threshold energy E'_{min} for producing antiprotons in the collision of protons with a nuclear target is given *approximately* by

$$E'_{min} \approx (1 - p/m_P c)E_{min} \approx 5.2\,\text{GeV},$$

where $E_{min} \approx 6.6\,\text{GeV}$is the threshold energy for producing antiprotons on a proton target, and $p = |\mathbf{p}|$, the magnitude of the internal momenta of the nucleons, is taken to be typically of order 0.2 GeV/c.

A.6 (a) Find the maximum values of velocity υ and γ if the nonrelativistic approximation for energy, $E \approx mc^2 + p^2/2m$, is to be used with an error less than $\varepsilon p^2/2m$. Hence show that for an error of 1 %, the momentum satisfies $p \approx 0.2mc$.

(b) Find the minimum values of velocity υ and γ if the relativistic approximation for energy, $E \approx pc$, is to be used with an error of less than εpc. Hence show that for an error of 1 %, the momentum satisfies $p \approx 7mc$.

A.7 When high-energy particles decay, the decay products emerge predominantly at very small angles θ_L to the initial beam direction. Verify this by showing that for decays of the form

$$b(E_L, \mathbf{p}_L) \to P(E, \mathbf{q}) + \ldots$$

the decay angle θ_L, defined by

$$\mathbf{p}_L \cdot \mathbf{q} = |\mathbf{p}_L|\,|\mathbf{q}| \cos\theta_L,$$

is given at high beam energies $E \gg m_b c^2$ by

$$\tan\theta_L \approx \frac{m_b c^2}{\sqrt{2}E_L} \cdot \frac{u \sin\theta_C}{u\cos\theta_C + c},$$

where u is the magnitude of the velocity of particle P in the laboratory frame and θ_C is the decay angle in the rest frame of the decaying particle.

B

Amplitudes and Cross-sections

In this appendix we define the various cross-sections and scattering amplitudes used in the book and derive the Breit–Wigner resonance formulas.

B.1 RATES AND CROSS-SECTIONS

In a typical scattering experiment, a beam of particles (e.g. negative pions), ideally monoenergetic, is directed on to a target (e.g. liquid hydrogen) and the rates of production of various particles are measured. These can be regarded as the sums of contributions resulting from the interactions of individual beam particles with individual target particles (i.e. individual $\pi^- p$ interactions in our example) provided the following conditions are satisfied:

(a) The target particles are separated by distances that are much greater than the de Broglie wavelength of the incident particles, so that interference effects between waves scattered from different target particles can be neglected.
(b) The target is small enough, or of sufficiently low density, for multiple scattering to be neglected, so that any particles produced in the interaction will almost certainly leave the target without further interactions.
(c) The collision energy is sufficiently high that the binding energies of the particles in the target can be neglected.
(d) The beam density is low enough that mutual interactions of its particles can be neglected.

These conditions are usually satisfied in particle physics experiments, and we will focus our attention on interactions between individual beam and target particles.

Particle Physics Third Edition B. R. Martin and G. Shaw

For each initial state, a number of final states may be possible. For example, in $\pi^- p$ collisions the reactions

$$\pi^- + p \to \pi^- + p, \tag{B.1a}$$

$$\pi^- + p \to K^0 + \Lambda, \tag{B.1b}$$

$$\pi^- + p \to \pi^+ + \pi^- + n \tag{B.1c}$$

and

$$\pi^- + p \to \pi^+ + \pi^- + \pi^0 + n, \tag{B.1d}$$

as well as many others compatible with the conservation laws, can all be observed. However, whatever reaction we choose, the rate will be proportional to:

(a) the number N of particles in the target illuminated by the beam;
(b) the rate per unit area at which beam particles cross a small surface placed in the beam at rest with respect to the target and perpendicular to the beam direction.

The latter rate is called the *flux*, and is given by

$$J = n_b \upsilon_i, \tag{B.2}$$

where n_b is the density of particles in the beam and υ_i their velocity in the rest frame of the target. Hence the rate $W_r$at which a specific reaction r occurs in a particular experiment can be written in the form

$$W_r = JN\sigma_r, \tag{B.3}$$

where σ_r, the constant of proportionality, is called the *cross-section* for reaction r. This decomposition is very convenient, because the dependence of the observed rate on the densities and geometries of the beam and target is absorbed into the factor

$$L \equiv J\,N, \tag{B.4}$$

called the *luminosity* L, while the cross-section is independent of these factors, but characteristic of the particular reaction r. It follows from (B.3) that σ_r has the dimensions of an area, and the rate per target particle $J\sigma_r$ at which the reaction occurs is equal to the rate at which beam particles would hit a surface of area σ_r placed in the beam at rest with respect to the target and perpendicular to the beam direction. Since the area of such a surface is unchanged by a Lorentz transformation in the beam direction, the cross-section has the same value in both the laboratory and centre-of-mass frames, which are related by just such a transformation.

B.2 THE TOTAL CROSS-SECTION

The quantity σ_r is better named the *partial cross-section*, because it is the cross-section for a particular reaction r. The *total cross-section* σ is defined by

$$\sigma \equiv \sum_r \sigma_r, \tag{B.5}$$

where the sum extends over all reactions r that satisfy the relevant conservation laws, such as the reactions listed in (B.1) for $\pi^- p$ scattering. The total reaction rate is then

$$W \equiv \sum_r W_r = JN\sigma = L\sigma, \tag{B.6}$$

by (B.3) and (B.4). Since every reaction removes a particle from the incoming beam, it is the total cross-section that determines how the beam intensity is reduced on traversing a given target.

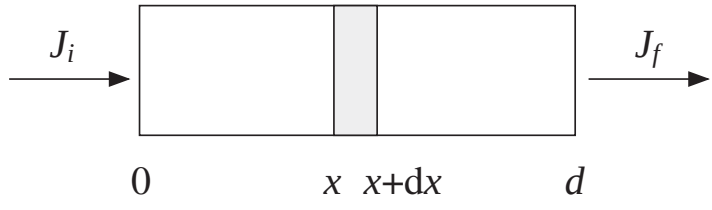


Figure B.1 Beam of particles with initial flux J_i being transmitted through a target of length d with reduced flux $J_f < J_i$.

To explore this further, we consider a beam of cross-sectional area A traversing a stationary target of length d, and focus our attention on a thin segment of the target of thickness dx at a distance x within it, as shown in Figure B.1. The reduction in the beam intensity on crossing this segment is equal to the reaction rate dW within the segment, i.e.

$$A\mathrm{d}J(x) = -\mathrm{d}W, \tag{B.7a}$$

where dJ is the change in the flux $J(x)$ between x and $x + \mathrm{d}x$. By (B.6) the rate is also given by

$$\mathrm{d}W = J(x)\sigma n_t A\mathrm{d}x, \tag{B.7b}$$

where n_t is the density of particles in the target. Combining (B.7a) and (B.7b) gives

$$\mathrm{d}J(x) = -n_t\sigma J(x)\mathrm{d}x$$

and if the incident flux is $J(0) \equiv J_i$, a simple integration gives

$$J(x) = J_i \exp(-x/l_c) \quad (0 \le x \le d), \tag{B.8}$$

where the collision length

$$l_c \equiv 1/n_t\sigma \tag{B.9}$$

is the mean free path travelled by a beam particle in an infinitely long target ($d = \infty$) before it interacts. Beyond the target, the transmitted flux

$$J_f \equiv J(x \geq d) = J_i \exp(-d/l_c). \tag{B.10}$$

These results have two important consequences. The first is that multiple scattering within the target can only be neglected if the target thickness d is much less than the collision length l_c. The basic cross-section formulas (B.3) and (B.6) obviously depend on this, since they treat the flux J as a constant. The second consequence is that the measured depletion of a beam flux on passing through a long target ($d \sim l_c$) gives a relatively simple determination of the total cross using (B.10).

B.3 DIFFERENTIAL CROSS-SECTIONS

We next consider the angular distributions of the particles produced in a scattering reaction. These are described by writing the cross-section σ_r as a sum of contributions corresponding to different directions of emission of the final state particles. For processes with two-body final states, such as (B.1a) and (B.1b), only the direction of one particle needs to be specified, since the direction of the other follows uniquely from energy and momentum conservation.[1] The direction of the chosen particle, e.g. the π^- in (B.1a) or the K^0 in (B.1b), is specified by a polar angle θ and an azimuthal angle ϕ defined in a convenient reference frame in which the z axis coincides with the beam direction. The angle θ between the particle direction and the beam direction is called the *scattering angle* in an elastic reaction like (B.1a) or the *production angle* in an inelastic reaction like (B.1b).

The angular distribution of the chosen particle (e.g. π^- or K^0) produced in a two-body reaction r is described by the *differential cross-section*

$$\frac{\mathrm{d}\sigma_r(\theta,\phi)}{\mathrm{d}\Omega}, \tag{B.11}$$

defined by

$$\mathrm{d}W_r \equiv JN \frac{\mathrm{d}\sigma_r(\theta,\phi)}{\mathrm{d}\Omega}\mathrm{d}\Omega, \tag{B.12}$$

where $\mathrm{d}W_r$ is the measured rate for the particle to be emitted into an element of solid angle $\mathrm{d}\Omega = \mathrm{d}\cos\theta\mathrm{d}\phi$ in the direction (θ,ϕ), as shown in Figure B.2. The reaction

[1] For example, in the centre-of-mass frame the particles are emitted in opposite directions, since the sum of their momenta must be zero.

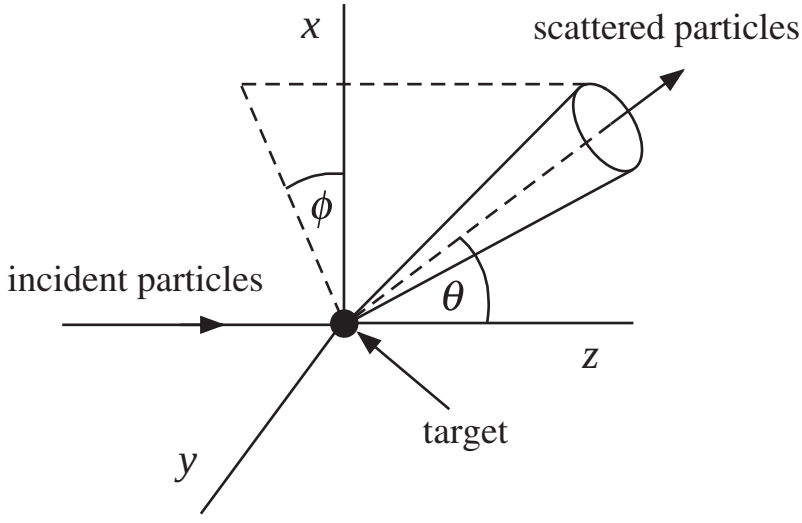


Figure B.2 Geometry of the differential cross-section: a beam of particles is incident along the z direction and collides with a stationary target at the origin. The differential cross-section is proportional to the rate for particles to be scattered into a small solid angle $\mathrm{d}\Omega$ in the direction (θ, ϕ).

cross-section σ_r can then be recovered by integrating the differential cross-section over all solid angles,

$$\sigma_r = \int_0^{2\pi} \mathrm{d}\phi \int_{-1}^{1} \mathrm{d}\cos\theta \frac{\mathrm{d}\sigma_r(\theta,\phi)}{\mathrm{d}\Omega}, \tag{B.13}$$

as can be seen by integrating (B.12) to give the total rate W_r and comparing with (B.3).

The differential cross-section (B.11) depends not only on the energy (which is not explicitly shown in (B.11)) and the angles but also on the spin states of the particles involved. In practice, most experiments are done using unpolarized beams and targets, and the spins of the particles in the final state are not measured. The cross-sections measured in such experiments are called *unpolarized cross-sections*, and for them the question of spin dependence does not arise. In addition, there is complete cylindrical symmetry about the beam direction (i.e. the z direction), with nothing to distinguish physically between different directions in the xy plane, and hence between different values of the azimuthal angle ϕ. Thus unpolarized cross-sections are independent of ϕ, i.e.

$$\frac{\mathrm{d}\sigma_r(\theta,\phi)}{\mathrm{d}\Omega} \to \frac{\mathrm{d}\sigma_r(\theta)}{\mathrm{d}\Omega},$$

and it is customary to define an unpolarized differential cross-section

$$\frac{\mathrm{d}\sigma_r}{\mathrm{d}\cos\theta} \equiv \int_0^{2\pi} \mathrm{d}\phi \frac{\mathrm{d}\sigma_r}{\mathrm{d}\Omega} = 2\pi \frac{\mathrm{d}\sigma_r}{\mathrm{d}\Omega}, \tag{B.14}$$

with

$$\sigma_r = \int_{-1}^{1} \mathrm{d}\cos\theta \frac{\mathrm{d}\sigma_r}{\mathrm{d}\cos\theta}. \tag{B.15}$$

The cross-sections discussed in this book are almost exclusively unpolarized cross-sections.

★B.4 THE SCATTERING AMPLITUDE

In this section we will relate the experimentally defined cross-sections to theoretical calculations of the probability of a transition from an initial quantum state i to a final quantum state f. In doing so, it is convenient to consider a single beam particle interacting with a single target particle and to confine the whole system in a large cube defined by

$$-L/2 \leq x, y, z \leq L/2, \tag{B.16}$$

so that the volume $V = L^3$. The incident flux (B.2) is then given by

$$J = n_b \upsilon_i = \upsilon_i / V,$$

and since the number of target particles $N = 1$, Equation (B.12) becomes

$$\mathrm{d}W_r = \frac{\upsilon_i}{V} \frac{\mathrm{d}\sigma_r(\theta, \phi)}{\mathrm{d}\Omega} \mathrm{d}\Omega. \tag{B.17}$$

Expressions for the differential cross-sections are found by comparing (B.17) with theoretical predictions for $\mathrm{d}W_r$. In doing this, we shall impose the periodic boundary conditions

$$\begin{aligned} \psi(-L/2, y, z) &= \psi(L/2, y, z), \\ \psi(x, -L/2, z) &= \psi(x, L/2, z), \\ \psi(x, y, -L/2) &= \psi(x, y, L/2) \end{aligned} \tag{B.18}$$

on the wavefunctions, and take the limit $V = L^3 \to \infty$ at the end of the calculation. The results are independent of V in this limit and also of the detailed form of the boundary conditions chosen.

We start by considering a particularly simple case: the scattering of a nonrelativistic particle by a potential $V(\mathbf{r})$, using lowest-order perturbation theory. We are then interested in transitions from an initial state described by a wavefunction

$$\psi_i = \frac{1}{\sqrt{V}} \exp(i\mathbf{q}_i \cdot \mathbf{r}) \tag{B.19a}$$

to final states described by wavefunctions

$$\psi_f = \frac{1}{\sqrt{V}} \exp(i\mathbf{q}_f \cdot \mathbf{r}), \tag{B.19b}$$

where the final momentum $\mathbf{q}_f$ lies within a small solid angle $\mathrm{d}\Omega$ located in the direction (θ, ϕ). In nonrelativistic quantum mechanics, the transition rate for any process is given in first-order perturbation theory by the Born approximation[2]

$$\mathrm{d}W_r = 2\pi \left| \int \mathrm{d}^3\mathbf{r} \psi_f^* V(\mathrm{r}) \psi_i \right|^2 \rho(E_f), \tag{B.20}$$

where the *density of states factor* $\rho(E_f)$ is specified below. Substituting (B.19) then gives

$$\mathrm{d}W_r = \frac{2\pi}{V} \rho(E_f) \left| \mathcal{M}_{if} \right|^2, \tag{B.21}$$

where the *scattering amplitude* M_{if} is

$$\mathcal{M}_{if} = \int d^3\mathbf{r} V(\mathbf{r}) \exp[(i(\mathbf{q}_i - \mathbf{q}_f) \cdot \mathbf{r}]. \tag{B.22}$$

The density of states factor $\rho(E)$ is defined by setting $\rho(E)\mathrm{d}E$ equal to the number of possible quantum states of the final state particles that have a total energy between E and $E + \mathrm{d}E$. It is found by first evaluating $\rho(q)$, where $\rho(q)\mathrm{d}q$ is the number of possible final states with $q = |\mathbf{q}|$ lying between q and $q + \mathrm{d}q$, and then changing variables using

$$\rho(q) \frac{\mathrm{d}q}{\mathrm{d}E} \mathrm{d}E = \rho(E)\mathrm{d}E. \tag{B.23}$$

The possible values of the momentum $\mathbf{q}$ are restricted by the boundary conditions (B.18) to be

$$q_x = \left(\frac{2\pi}{L}\right) n_x, \quad q_y = \left(\frac{2\pi}{L}\right) n_y, \quad q_z = \left(\frac{2\pi}{L}\right) n_z,$$

where n_x, n_y and n_z are integers. Hence the number of final states with momenta lying in the momentum space volume

$$\mathrm{d}^3\mathbf{q} = q^2 \mathrm{d}q \, \mathrm{d}\Omega,$$

[2] For a discussion of this formula see, for example, Section 10.2.2 of Mandl (1992) or pp. 397–399 of Gasiorowicz (1974). Note that we are using natural units, so that $\hbar = 1$.

corresponding to momenta pointing into the solid angle $\mathrm{d}\Omega$ with magnitude between q and $q + \mathrm{dq}$, is given by

$$\rho(q)\mathrm{d}q = \left(\frac{\mathrm{L}}{2\pi}\right)^3 \mathrm{d}^3\mathbf{q} = \frac{V}{8\pi^3} q^2 \mathrm{d}q\, \mathrm{d}\Omega. \tag{B.24}$$

The derivative

$$\frac{\mathrm{d}q}{\mathrm{d}E} = \frac{1}{\upsilon}, \tag{B.25}$$

where $q = m\upsilon$ for a nonrelativistic particle, and substituting (B.24) and (B.25) in (B.23) gives

$$\rho(E_f) = \frac{V}{8\pi^3} \frac{q_f^2}{\upsilon_f} \mathrm{d}\Omega \tag{B.26}$$

for the density of states evaluated in the final energy E_f.The desired expression for the differential cross-section can now be found by substituting (B.26) into (B.21) and comparing with (B.17). On doing this, and remembering that for elastic scattering energy conservation requires $\upsilon_i = \upsilon_f = \upsilon$ and $q_i = q_f = m\upsilon$, we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{m^2}{4\pi^2} \left|\mathcal{M}_{if}\right|^2, \tag{B.27}$$

where $\mathcal{M}_{if}$ is given by the Fourier transform of the potential, (B.22).

The result (B.27) has been obtained for nonrelativistic elastic scattering in lowest-order perturbation theory. Nonetheless, it requires very little modification to extend it to any reaction of the form

$$a(\mathbf{q}_a, m_a) + b(\mathbf{q}_b, m_b) \rightarrow c(\mathbf{q}_c, m_c) + d(\mathbf{q}_d, m_d), \tag{B.28}$$

where m_a, etc., specify the spin states of the particles, and we have indicated their momenta in the centre-of-mass frame. To see this, we first note that any theoretically calculated reaction rate involves a factor proportional to the density of final states $\rho(E_f)$; hence (B.21) may be taken as a *definition* of the scattering amplitude $\mathcal{M}_{if}$, apart from an overall phase factor, which is irrelevant in what follows. The density of states must of course be evaluated using relativistic kinematics. The only place where we have used nonrelativistic kinematics in the evaluation above was in finding the derivative (B.25). It is straightforward to show that (B.25) also holds if relativistic kinematics are used, and thus the same expression (B.26) results for the density of states. Substituting into (B.21) and comparing with (B.17) then gives

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{4\pi^2} \frac{q_f^2}{\upsilon_i \upsilon_f} \left|\mathcal{M}_{if}\right|^2 \tag{B.29}$$

as the general expression for the differential cross-section for a reaction of the type (B.28), evaluated in the centre-of-mass frame. It reduces to our previous result (B.27) for nonrelativistic elastic scattering on setting $\upsilon_i = \upsilon_f = \upsilon$ and $q_i = q_f = m\upsilon$.

The result (B.29) applies to the scattering of polarized particles as symbolized in (B.28), and the amplitude

$$\mathcal{M}_{if}(\theta, \phi) \equiv \mathcal{M}_{if}(\theta, \phi; m_a, m_b, m_c, m_d)$$

depends on the particular spin states involved. If the beam and target are unpolarized, one must *average* over the initial spin states, whereas if all final state particles are detected, independent of their spin projections, one must *sum* over the final spin states. Hence in an unpolarized scattering experiment, the spin-averaged reaction rates must be multiplied by a factor $(2s_f + 1)$ for each final state particle of spin s_f. In addition, since the cross-section only depends on the angle θ in this case, we may integrate over the azimuthal angle ϕ using (B.14) to give

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta} = \frac{(2S_c + 1)\,(2S_d + 1)}{2\pi} \frac{q_f^2}{\upsilon_i \upsilon_f} \overline{\left|\mathcal{M}_{if}\right|^2} \tag{B.30}$$

for the unpolarized differential cross-section, where the final factor is the spin average of the squared matrix element $\left|\mathcal{M}_{if}\right|^2$.

★B.5 THE BREIT–WIGNER FORMULA

An unstable particle may be characterized by its lifetime at rest τ, or equivalently by its natural decay width $\Gamma = \hbar/\tau$. Since we use natural units $\hbar = 1$, the decay width is equal to the decay rate

$$\Gamma = 1/\tau \tag{B.31a}$$

and we shall not distinguish between them in what follows. In general, the particle will decay to a number of different channels f (for example $Z^0 \to e^+e^-$, $\mu^+\mu^-$, ...) and the total decay width/rate Γ is the sum of the partial decay widths/rates Γ_f for the various channels; i.e.

$$\Gamma = \sum_f \Gamma_f, \tag{B.31b}$$

while the *branching ratio*, the fraction of decays leading to a particular channel f, is given by

$$B_f \equiv \Gamma_f \,/\, \Gamma\ . \tag{B.31c}$$

The lifetimes and branching ratios for 'long-lived' particles with $\tau \geq 10^{-16}$ s can be measured directly. However, some very heavy states, and particles that decay by strong interactions, have lifetimes that are much shorter than this. In such cases, the decay widths are inferred from the observation of peaks in cross-sections and mass

distributions, as we saw in Chapter 3, Section 3.5. In the present section we give a derivation of the Breit–Wigner formula, which was used in Section 3.5 to describe the shape of such peaks.

B.5.1 Decay distributions

We begin by discussing how unstable states are described in quantum mechanics. If we denote the wavefunction of the unstable state by ψ_0 and the wavefunctions of the possible final states by $\psi_n (n \geq 1)$, then we can always choose these to be orthonormal, i.e.

$$\int \psi_m^* \psi_n \mathrm{d}x = 1, \quad \text{if } m = n,$$
$$= 0, \quad \text{if } m \neq n,$$

where x denotes all relevant variables. The wavefunction ψ_0 describes the system at $t = 0$. At a later time t the wavefunction Ψ will be a linear superposition of the ψ_n, which is conveniently written in the form

$$\Psi(\mathbf{r}, t) = \sum_{n=0}^{\infty} a_n(t) \mathrm{e}^{-iE_n t} \psi_n(\mathbf{r}), \tag{B.32}$$

where

$$a_0(0) = 1, \quad a_n(0) = 0 \, (n \geq 1) \tag{B.33}$$

and

$$E_n \equiv H_{nn} = \int \psi_n^* H \psi_n \mathrm{d}x.$$

The exponential factors in (B.32) are inserted purely for later convenience, since the coefficients $a_n(t)$ remain to be determined. In addition, the energy of the initial state $E_0 \pm \Delta E$ has an uncertainty $\Delta E \approx 1/\tau$ by the energy–time uncertainty principle, so that $E_n \neq E_0$ in general. Because of this, there are always many quantum states ψ_n contributing to any decay channel i, and the individual coefficients $a_n(t)$ $(n \neq 0)$ remain very small even when $t \gg \tau$, the lifetime of the decaying particle.

The time dependence of the coefficients $a_n(t)$ is determined by substituting (B.32) into the Schrödinger equation

$$i \frac{\partial \Psi}{\partial t} = H \Psi$$

to give

$$\sum_m \left\{ i \frac{\mathrm{d}a_m}{\mathrm{d}t} \mathrm{e}^{-iE_m t} \psi_m + E_m a_m \mathrm{e}^{-iE_m t} \psi_m \right\} = \sum_m a_m \mathrm{e}^{-iE_m t} H \psi_m.$$

Multiplying this equation by ψ_n^* and integrating using the orthonormality relations above projects out the term $\mathrm{d}a_n/\mathrm{d}t$:

$$i\frac{\mathrm{d}a_n}{\mathrm{d}t} = \sum_{m\neq n} H_{nm}\mathrm{e}^{-i(E_m-E_n)t}a_m. \tag{B.34}$$

So far we have made no approximations, and the rigorous way forward would be to solve the set of differential equations (B.34) with the initial conditions (B.33) and an appropriate Hamiltonian, as originally done by Breit and Wigner. Here, following Fermi,[3] we will obtain the same results in an approximate way by working to first order in H_{nm}, which are assumed to be small when $n \neq m$. In this case we have, for $n \geq 1$,

$$i\frac{\mathrm{d}a_n}{\mathrm{d}t} = H_{n0}\mathrm{e}^{-i(E_0-E_n)t}a_0, \tag{B.35}$$

since a_m is of order H_{0m} for $m \neq 0$. Also, we specialize to the rest frame of the decaying particle so that $E_0 = M$, where M is its rest mass, and we assume $a_0(t) = \mathrm{e}^{-\Gamma t/2}$, consistent with the exponential decay law

$$|a_0(t)|^2 = \mathrm{e}^{-\Gamma t} = \mathrm{e}^{-t/\tau},$$

where τ is the lifetime at rest. Substituting this form for $a_0(t)$into (B.35) and solving with the boundary conditions (B.33) then gives

$$ia_n = -iH_{n0}\left\{\frac{\exp\{-i[(M-E_n)-i\Gamma/2]t\}-1}{(E_n-M)+i\Gamma/2}\right\}.$$

For large values of $t \gg 1/\Gamma$, the exponential term tends to zero, so that the probability P_n of finding a final quantum state n is given by

$$P_n = |a_n(\infty)|^2 = \frac{|H_{n0}|^2}{(E_n-M)^2+\Gamma^2/4}.$$

This is conveniently rewritten in the form

$$P_n = \frac{2\pi}{\Gamma}|H_{n0}|^2 P(E_n-M), \tag{B.36}$$

where the distribution function

$$P(E-M) = \frac{\Gamma/2\pi}{(E-M)^2+\Gamma^2/4} \tag{B.37a}$$

[3] See pp. 152–157 of Fermi (1950).

peaks sharply at $E = M$ with width Γ, and is normalized to unity:

$$\int_{-\infty}^{\infty} P(E-M)\mathrm{d}E = 1. \tag{B.37b}$$

Equation (B.36) is the probability for decay to a specific quantum state n. In general, a particle can decay to several channels f, and to obtain the probability for any given channel we must sum over the many quantum states n contained in that channel. If we ignore spin, the probability $P_f(E)\,\mathrm{d}E$ for decay to a particular channel f with energy in the interval E to $E + \mathrm{d}E$ is therefore

$$P_f(E)\,\mathrm{d}E = \frac{2\pi}{\Gamma}\left|H_{f0}\right|^2 P(E-M)\rho_f(E)\mathrm{d}E,$$

where the *density of states* $\rho_f(E)$ is the number of quantum states with energy between E and $E + \mathrm{d}E$ in channel f. If Γ is small, $\rho_f(E)$ is sharply peaked at $E = M$, because of the factor $P(E-M)$, and it is reasonable to approximate H_{f0} and $\rho_f(E)$ by their values at $E = M$. We then obtain

$$P_f(E)\,\mathrm{d}E = \frac{2\pi}{\Gamma}\left|H_{f0}\right|^2 P(E-M)\rho_f(M)\mathrm{d}E \tag{B.38}$$

for the energy distribution of the decays of a particle of mass M into a particular channel f. The total probability for decay to channel f is equal to the branching ratio (B.31c) and can be evaluated by integrating (B.38) over all energies using (B.37b). In this way one obtains

$$\Gamma_f = 2\pi\left|H_{f0}\right|^2 \rho_f(M), \tag{B.39a}$$

so that (B.38) becomes

$$P_f(E) = \frac{1}{2\pi}\cdot\frac{\Gamma_f}{(E-M)^2 + \Gamma^2/4} \tag{B.40}$$

on substituting (B.37a). This is the *Breit–Wigner formula for the decay distribution* of a spin-0 particle. It also holds in the general case where the initial particle is unpolarized but has spin S_i, provided the decay width (B.39a) is replaced by

$$\Gamma_f = \frac{2\pi}{2S_i+1}\overline{\left|H_{f0}\right|^2}\rho_f(M), \tag{B.39b}$$

where $\overline{\left|H_{f0}\right|^2}$ is the spin average of the squared matrix element and the density of states $\rho_f(M)$ includes a sum over all final particle spin states.

The Breit–Wigner formula (B.40) represents the final result of this section, and holds in the rest frame of the decaying particle. Since this is identical with the

centre-of-mass frame of the decay products, their invariant mass $W = E$, and (B.40) can be written in the form

$$P_f(W) = \frac{1}{2\pi} \cdot \frac{\Gamma_f}{(W-M)^2 + \Gamma^2/4}, \tag{B.41}$$

which holds in any frame of reference. This is just the Breit–Wigner formula (3.27) used to extract resonance masses and widths in Section 3.5. Its characteristic shape is plotted in Figure B.3.

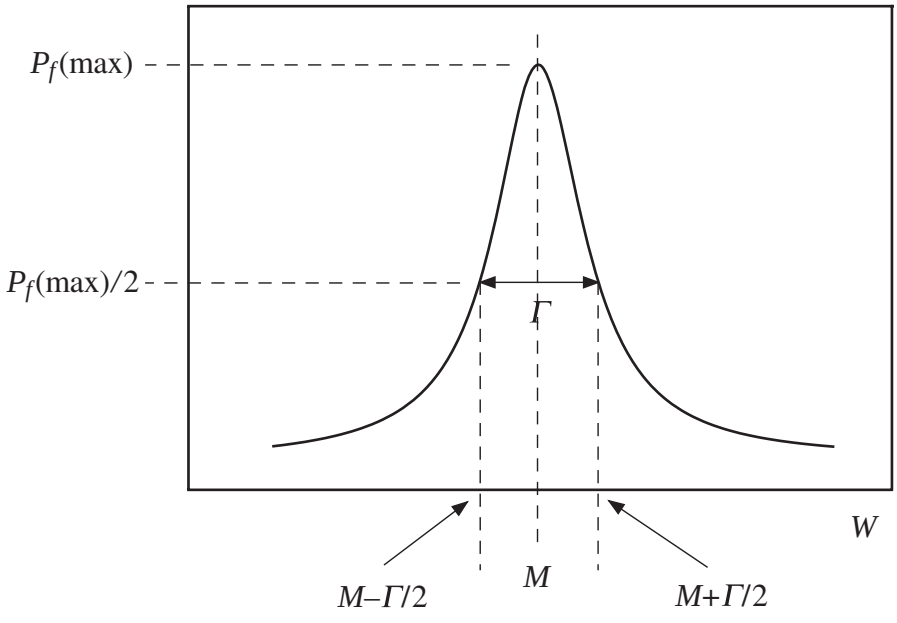


Figure B.3 The Breit–Wigner formula (B.41).

B.5.2 Resonant cross-sections

We now turn to the case where an isolated unstable particle (or resonance) is formed in a scattering process and subsequently decays to a final channel f, as shown in Figure B.4. For steady incoming beams, the resonant state will be formed at a constant rate. However, this will eventually be balanced by its subsequent decay, leading to a constant average number of resonances present, independent of time. The rate of production of the final channel f is then determined by the partial decay rate Γ_f of the resonance to that channel.

To solve this problem exactly, we should solve (B.34) with the boundary conditions

$$a_1(0) = 1, \quad a_n(0) = 0 \ (n \neq 1) \tag{B.42}$$

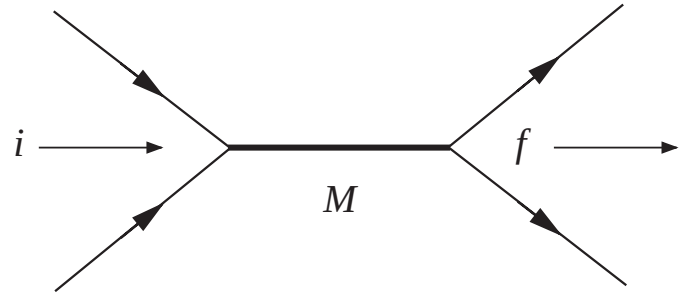


Figure B.4 Formation and decay of an unstable particle of mass M in a two-body scattering process, where the initial channel i and the final channel f may, or may not, be the same.

at $t = 0$, where $n = 1$ refers to the initial state and $n = 0$ to the resonant state, as usual. In practice, we shall again obtain the results in an approximate way, following Fermi. To first order in H_{nm}, (B.34) gives

$$i\frac{\mathrm{d}a_0(t)}{\mathrm{d}t} = H_{01}\mathrm{e}^{-i(E_1-E_0)t}, \tag{B.43}$$

since $a_n = \mathrm{O}(H_{1n})$ for $n \neq 1$, and

$$|a_1|^2 = 1 - \sum_{n\neq 1} |a_n|^2.$$

Equation (B.43) describes the 'feeding' of the resonance state from the initial state. However, it ignores the decays of the resonance, which involve an additional power of H_{0n}, but can involve a very large number of quantum states n. These can be taken into account by modifying (B.43) to

$$i\frac{\mathrm{d}a_0(t)}{\mathrm{d}t} = H_{01}\mathrm{e}^{-i(E_1-E_0)t} - i\left(\frac{\Gamma}{2}\right)a_0(t), \tag{B.44}$$

where the additional term ensures that $|a_0(t)|^2$ would decay exponentially if the first term were switched off by, for example, stopping the beam. The solution of (B.44) with $a_0 = 0$ at $t = 0$ may be found by rewriting it as

$$i\frac{d}{dt}\left(a_0\mathrm{e}^{\Gamma t/2}\right) = H_{01}\mathrm{e}^{-i(E_1-E_0+i\Gamma/2)t}$$

and integrating over time. For times $t \gg 1/\Gamma$, this gives

$$a_0(t) = \frac{H_{01}\mathrm{e}^{-i(E_1-E_0)t}}{E_1 - E_0 + i\Gamma/2}$$

and the partial decay rate into the final channel f is therefore

$$W_f = \Gamma_f \, |a_0(t)|^2 = \frac{|H_{01}|^2 \, \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}, \tag{B.45}$$

where we have set the initial energy E_1 identically equal to E as usual.

It remains to relate (B.45) to the centre-of-mass cross-section. From (B.17) we have

$$W_f = \frac{\upsilon_i}{V} \sigma_{if},$$

where υ_i is the relative velocity of the particles in the initial state. In addition, from (B.39a) and (B.26) we have

$$\Gamma_i = 2\pi \, |H_{10}|^2 \, \rho_i(E_0) = \frac{V}{\pi} \frac{q_i^2}{\upsilon_i} \, |H_{10}|^2 \, ,$$

where q_i is the initial momentum and the density of states (B.26) has been integrated over angles to accord with the density of states used in (B.39a). Substituting these equations into (B.45) gives

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{\Gamma_i \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}, \tag{B.46a}$$

where $E_0 = M$, the resonance mass, if we work in the centre-of-mass frame. This is the *Breit–Wigner resonance formula*. In the general case where the initial particles have spins S_1 and S_2 and the unstable particle has spin j, (B.46a) is generalized to

$$\sigma_{if} = \frac{\pi}{q_i^2} \frac{2j + 1}{(2S_1 + 1)\,(2S_2 + 1)} \frac{\Gamma_i \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}. \tag{B.46b}$$

It is often written in slightly different forms, which are all equivalent when $\Gamma^2 \ll E_0^2 = M^2$, and we leave it as an exercise to show that the form (9.26) used in Section 9.1.4 is identical to (B.46b) in this approximation.

The Breit–Wigner formula is a good representation of data when an isolated state dominates the cross-section, but in many cases nonresonant scattering also occurs and other terms have to be added to the amplitude in order to extract reliable resonance parameters from experimental data.

PROBLEMS B

B.1 A liquid hydrogen target of volume $10^{-4}\,\mathrm{m}^3$ and density $71\,\mathrm{kg\,m}^{-3}$ is exposed to a wide uniform monoenergetic beam of negative pions of flux 10^7 particles $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ and the reaction $\pi^- p \to K^0 \Lambda$ is observed. If the cross-section for this reaction is 0.4 mb, what is the rate of production of Λ particles?

B.2 A high-energy beam of neutrons of intensity $10^6\,\mathrm{s}^{-1}$ traverses a target of ^{238}U in the form of a thin foil whose density per unit area is $10^{-1}\,\mathrm{kg\,m}^{-2}$. If the elastic and inelastic cross-sections are 1.4 and 2.0 b, respectively, calculate: (a) the attenuation of the beam, (b) the rate of inelastic reactions and (c) the flux of elastically scattered neutrons 5 m from the target, averaged over all scattering angles. (In practice the neutrons are emitted at small angles relative to the beam direction.)

B.3 Show that Equation (9.26) reduces to (B.46b) in the region of the peak if $\Gamma^2 \ll M_Z^2$.

C

★The Isospin Formalism

The discussion of isospin in Section 6.1 was based on the assumed analogy with ordinary angular momentum and spin. In this appendix we give a more formal treatment in order to justify this analogy and to derive a number of other results that were cited in the text. Our approach will be based on the fact that many properties of angular momentum can be derived from the commutation relations

$$\left[\hat{J}_x, \hat{J}_y\right] = i\hat{J}_z \quad \left[\hat{J}_y, \hat{J}_z\right] = i\hat{J}_x, \quad \left[\hat{J}_z, \hat{J}_x\right] = i\hat{J}_y, \tag{C.1}$$

where $\hat{J}_{x,y,z}$ are the components of any angular momentum operator (e.g. total or orbital angular momentum or spin.)[1] An equivalent, and in some ways more convenient, set of commutation relations that follows from (C.1) is

$$\left[\hat{J}_+, \hat{J}_-\right] = 2\hat{J}_3, \quad \left[\hat{J}_3, \hat{J}_\pm\right] = \pm\hat{J}_\pm \tag{C.2}$$

where $\hat{J}_\pm$ and $\hat{J}_3$ are the linear combinations

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y, \quad \hat{J}_3 = \hat{J}_z. \tag{C.3}$$

Hence, if we can identify the appropriate isospin operators

$$\hat{I}_+, \quad \hat{I}_- \quad \text{and} \quad \hat{I}_3 \tag{C.4}$$

[1] This approach to angular momentum is summarized in, for example, Sections 5.2, 5.4 and 5.8 of Mandl (1992).

Particle Physics Third Edition B. R. Martin and G. Shaw

that satisfy the commutation relations

$$\left[\hat{I}_+, \hat{I}_-\right] = 2\hat{I}_3, \quad \left[\hat{I}_3, \hat{I}_\pm\right] = \pm\hat{I}_\pm, \tag{C.5}$$

then the same properties will follow for isospin. These include the multiplet structure of (6.7) and (6.8), and the rules of the addition of isospin quoted in Section 6.1.1.

C.1 ISOSPIN OPERATORS

The isospin operators $\hat{I}_\pm$ and $\hat{I}_3$ are defined by specifying their actions on a complete set of quantum states. This is simplest for the operator $\hat{I}_3$, since in Section 6.1.1 we showed that the corresponding quantum number

$$I_3 \equiv Q - Y/2 \tag{C.6}$$

was equivalent to

$$I_3 = \tfrac{1}{2}\left(N_u - N_d\right), \tag{C.7}$$

where N_u and N_d are the quark numbers defined in Equation (3.3).[2] Hence for a u quark state we have, for example,

$$\hat{I}_3 \left|u, \Psi\right\rangle = \tfrac{1}{2} \left|u, \Psi\right\rangle$$

for an arbitrary space and spin wavefunction Ψ. The definition of the operators $\hat{I}_+$ and $\hat{I}_-$ is less obvious, but can be inferred from the analogy between isospin and ordinary spin.[3] Here we shall just write them down and justify them afterwards by showing that they do indeed lead to the desired commutation relations (C.5). Specifically, we define

$$\hat{I}_+ u = 0, \quad \hat{I}_- u = d, \qquad \hat{I}_3 u = \tfrac{1}{2} u, \tag{C.8a}$$

$$\hat{I}_+ d = u, \quad \hat{I}_- d = 0, \qquad \hat{I}_3 d = -\tfrac{1}{2} d, \tag{C.8b}$$

$$\hat{I}_+ \bar{d} = 0, \quad \hat{I}_- \bar{d} = -\bar{u}, \quad \hat{I}_3 \bar{d} = \tfrac{1}{2} \bar{d} \tag{C.8c}$$

and

$$\hat{I}_+ \bar{u} = -\bar{d}, \quad \hat{I}_- \bar{u} = 0, \quad \hat{I}_3 \bar{u} = -\tfrac{1}{2} \bar{u}, \tag{C.8d}$$

---

[2] Compare the discussion leading to Equation (6.6).

[3] Specifically, the quark states (u, d) are assumed to be analogous to the spin up and spin down states α and β of an electron, and the isospin operators $\hat{I}_3$, $\hat{I}_\pm$ are assumed to be analogous to the spin operators $\hat{S}_3$ and $\hat{S}_\pm = \hat{S}_x \pm \hat{S}_y$. The result $\hat{I}_+ d = u$, for example, then follows from the analogous relation $\hat{S}_+ \beta = \alpha$, which can be easily derived using (5.21) and (5.22).

together with

$$\hat{I}_+ r = \hat{I}_- r = \hat{I}_3 r = 0 \tag{C.8e}$$

for all other quarks and antiquarks, collectively denoted r. In all these relations we have suppressed the space and spin variables, which are assumed to be left unchanged. For example, $\hat{I}_- u = d$ is really shorthand for

$$\hat{I}_- \, |u, \Psi\rangle = |d, \Psi\rangle \,,$$

which tells us that the operator $\hat{I}_-$ replaces the u quark by a d quark with the same combined space and spin wavefunction Ψ.

To complete the definition of the isospin operators we need to generalize (C.8) to an arbitrary state $|A\rangle$ composed of quarks and antiquarks. Such a state can be written as a linear combination

$$|A\rangle = \sum_i \beta_i \, |A_i\rangle \tag{C.9}$$

of product states

$$|A_i\rangle = a, \, ab, \, abc, \, \ldots, \tag{C.10}$$

where $a, b, c, \ldots$ can be any quark or antiquark $u, \bar{u}, d, \bar{d}, \ldots$ and β_i are arbitrary constants. The action of the operators on the product states $|A_i\rangle$ is fixed by the distributive laws:

$$\hat{I}_\alpha (ab) = \left(\hat{I}_\alpha a\right) b + a \left(\hat{I}_\alpha b\right), \tag{C.11a}$$

$$\hat{I}_\alpha (abc) = \left(\hat{I}_\alpha a\right) bc + a \left(\hat{I}_\alpha b\right) c + ab \left(\hat{I}_\alpha c\right) \tag{C.11b}$$

and so on, where $\alpha = +, -$ or 3. The definition is finally completed by requiring that the $\hat{I}_\alpha$ are linear operators, so that

$$\hat{I}_\alpha \, |A\rangle = \hat{I}_\alpha \sum_i \beta_i \, |A_i\rangle = \sum_i \beta_i \hat{I}_\alpha \, |A_i\rangle, \tag{C.12}$$

where $|A\rangle$ is the arbitrary state (C.9).

The above relations completely define the isospin operators, but we have not shown that they satisfy the crucial commutation relations (C.5). To do this, we note that if, for example, the relation

$$\left[\hat{I}_+, \hat{I}_-\right] a = 2\hat{I}_3 a \tag{C.13}$$

holds for an arbitrary single-particle state $a = u, \bar{u}, d, \ldots$, then it follows from (C.11) and (C.12) that

$$\left[\hat{I}_+, \hat{I}_-\right] |A\rangle = 2\hat{I}_3 \, |A\rangle$$

holds for an arbitrary state (C.9). Hence, to verify the first of the commutators (C.5), it is only necessary to verify (C.13). For the case $a \equiv u$,we have

$$\left[\hat{I}_+, \hat{I}_-\right] u = \hat{I}_+ \hat{I}_- u - \hat{I}_- \hat{I}_+ u = \hat{I}_+ d = u$$

and

$$2\hat{I}_3 u = u$$

by (C.8a) and (C.8b). Hence (C.13) is satisfied for $a \equiv u$, and it is equally easy to verify it for the other possible cases. This establishes the first of the commutation relations (C.5) and the others follow in a similar manner.

C.2 ISOSPIN STATES

The properties of isospin states can be derived using the commutation relations (C.5) in exactly the same way that the properties of angular momentum states can be derived from the corresponding commutation relations (C.2). In particular, it can be shown[4] that there exist families of $2I + 1$ states $|I, I_3\rangle$ with

$$I_3 = I, I - 1, \ldots, -I,$$

which are simultaneous eigenstates of the operators and

$$\hat{I}^2 \equiv \frac{1}{2}\left[\hat{I}_+ \hat{I}_- + \hat{I}_- \hat{I}_+\right] + \hat{I}_3^2, \tag{C.14}$$

with eigenvalues given by

$$\hat{I}^2 \, |I, I_3\rangle = I\,(I+1)\,|I, I_3\rangle \tag{C.15a}$$

and

$$\hat{I}_3 \, |I, I_3\rangle = I_3 \, |I, I_3\rangle \,. \tag{C.15b}$$

The action of the operators $\hat{I}_\pm$ on these states is also fixed from the commutation relations, and is given by

$$\hat{I}_\pm \, |I, I_3\rangle = C_\pm(I, I_3) \, |I, I_3 \pm 1\rangle \,, \tag{C.16a}$$

[4] The proofs of the following results are identical with those of the corresponding results for angular momentum given in, for example, Sections 5.2 and 5.8 of Mandl (1992).

where the coefficients are

$$C_{\pm}(I,I_3) = +[(I \mp I_3)(I \pm I_3 + 1)]^{1/2}. \tag{C.16b}$$

As the action of $\hat{I}_{\pm}$ shifts the values of I_3 by ± 1, the operators $\hat{I}_{\pm}$ are known as *shift* or *ladder operators*. These equations are extremely useful in applications, and in particular they can be used to show (cf. Problem C.2) that the $(2I+1)$ states $|I,I_3\rangle$ have the same energy if isospin is conserved, i.e. if

$$\left[H,\hat{I}_{\pm}\right] = \left[H,\hat{I}_3\right] = 0. \tag{C.17}$$

C.3 ISOSPIN MULTIPLETS

The practical applications of isospin symmetry rest on the assignment of the observed hadrons to isospin multiplets. Here we shall discuss this a little more fully than in the main text, with a view, in particular, to establishing the quark content of the neutral pion (6.23). In doing so we will write the members of isospin multiplets inside 'column vectors' in the order

$$\begin{pmatrix} |a;I,I_3 = I\rangle \\ |a;I,I_3 = I-1\rangle \\ \vdots \\ |a;I,I_3 = -I\rangle \end{pmatrix}, \tag{C.18}$$

where a is some convenient label used to identify a particular multiplet. In this notation, the basic quark and antiquark isodoublets are defined by

$$\begin{pmatrix} u \\ d \end{pmatrix}, \qquad \begin{pmatrix} \bar{d} \\ -\bar{u} \end{pmatrix}, \tag{C.19}$$

which tells us that the quarks and antiquarks are assigned to the isospin states $|q,I,I_3\rangle$ and $|\bar{q},I,I_3\rangle$ according to

$$u = \left|q;\tfrac{1}{2},\tfrac{1}{2}\right\rangle, \quad d = \left|q;\tfrac{1}{2},-\tfrac{1}{2}\right\rangle \tag{C.20}$$

and

$$\bar{d} = \left|\bar{q};\tfrac{1}{2},\tfrac{1}{2}\right\rangle, \quad \bar{u} = -\left|\bar{q};\tfrac{1}{2},-\tfrac{1}{2}\right\rangle, \tag{C.21}$$

respectively. These assignments are, of course, chosen to be consistent with both the definitions (C.8) and the fundamental results (C.15) and (C.16) (cf. Problem C.3). The remaining quarks and antiquarks

$$r = s, \bar{s}, c, \bar{c}, \ldots \tag{C.22}$$

are assigned to isosinglets, with $I = I_3 = 0$, as in Table 6.2.

C.3.1 Hadron states

The isospin assignments of the hadron states $q\bar{q}$ and qqq can be deduced from the basic quark assignments using the rules for the addition of isospin, which are identical to those for angular momentum. For example, since strange quarks and antiquarks are isosinglets with $I = I_3 = 0$, adding them to any state does not change its isospin properties. Hence from (C.19) we can infer that the meson states

$$\begin{pmatrix} u\bar{s} \\ d\bar{s} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} s\bar{d} \\ -s\bar{u} \end{pmatrix}$$

form isodoublets. With the usual quark assignments (3.16), this implies that the kaons in turn form isodoublets

$$\begin{pmatrix} K^+ \\ K^0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \bar{K}^0 \\ -K^- \end{pmatrix}.$$

A less trivial example concerns mesons which are bound states of a u or d quark with a $\bar{u}$ or $\bar{d}$ antiquark. Since these quarks and antiquarks belong to isodoublet states (C.19) with $I = \frac{1}{2}$, we expect to find three isotriplet states with $I = 1$ and $I_3 = 1, 0$ and -1, and an isosinglet state with $I = 0$ and $I_3 = 0$. The problem is to identify which isospin states $|I, I_3\rangle$ correspond to which combinations $q\bar{q}$. In such problems, it is often convenient to start with the state with the largest possible values of I and I_3. This has $I = I_3 = 1$ in this case, and since only the u quark and d antiquark have $I_3 = \frac{1}{2}$, the corresponding $q\bar{q}$ combination must be $u\bar{d}$. We therefore start from the identification

$$|I = 1, I_3 = 1\rangle = -u\bar{d}, \tag{C.24a}$$

where the sign is arbitrary and is chosen to agree with standard conventions. The other $I = 1$ states are obtained by using the operator $\hat{I}_-$, which decreases I_3 by one unit while leaving I unchanged. Explicitly, we have

$$\hat{I}_- \, |I = 1, I_3 = 1\rangle = \sqrt{2}\, |I = 1, I_3 = 0\rangle$$

by (C.16) and

$$\hat{I}_- \left(-u\bar{d}\right) = -\left(\hat{I}_- u\right)\bar{d} - u\left(\hat{I}_- \bar{d}\right) = -d\bar{d} + u\bar{u}$$

by (C.10) and (C.8). Comparing these equations gives

$$|I = 1, I_3 = 0\rangle = \frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right) \tag{C.24b}$$

and acting on both sides with $\hat{I}_-$ again leads to

$$|I = 1, I_3 = -1\rangle = d\bar{u}. \tag{C.24c}$$

The only remaining state is the isosinglet, which must be of the form

$$|I=0,I_3=0\rangle=\alpha\, u\bar{u}+\beta\, d\bar{d} \tag{C.25}$$

in order to ensure that $I_3 = 0$. The normalization condition requires

$$\alpha^2+\beta^2=1,$$

so that one more condition is required to determine (C.25) up to an arbitrary overall phase factor. From (C.16) we have

$$\hat{I}_-\,|I=0,I_3=0\rangle=0=\hat{I}_+\,|I=0,I_3=0\rangle \tag{C.25}$$

and either of these leads to

$$|I=0,I_3=0\rangle=\frac{1}{\sqrt{2}}\left(u\bar{u}+d\bar{d}\right), \tag{C.26}$$

where the overall phase has again been chosen to coincide with the most popular convention. Alternatively, the same result can be obtained by requiring that the new state is orthogonal to (C.24b), since the latter is already known. Equations (C.24b) and (C.26) were used in Chapter 6 (where they appeared as (6.23) and (6.24a), respectively) to determine the quark content of the neutral pion. Specifically, on identifying pions with the states (C.24), we obtain an isotriplet

$$\begin{pmatrix} -\pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}, \tag{C.27}$$

with the quark assignments

$$\pi^+=u\bar{d},\quad \pi^0=\tfrac{1}{\sqrt{2}}(u\bar{u}-d\bar{d}),\quad \pi^-=d\bar{u}. \tag{C.28}$$

Similar arguments can be applied to qqq states, leading to baryon multiplets like

$$(\Lambda),\quad \begin{pmatrix} p \\ n \end{pmatrix},\quad \begin{pmatrix} \Sigma^+ \\ \Sigma^0 \\ \Sigma^- \end{pmatrix},\quad \begin{pmatrix} \Delta^{++} \\ \Delta^+ \\ \Delta^0 \\ \Delta^- \end{pmatrix}, \tag{C.29}$$

corresponding to $\Lambda(1116)$, $N(940)$, $\Sigma(1193)$ and $\Delta(1232)$ states with $I=0,\frac{1}{2},1,\frac{3}{2}$, respectively. We shall not pursue this because the quark content of baryons follows unambiguously from their quantum numbers B, Q, S, C, B and T alone, as we saw in Section 6.2.2, and because there are no minus signs in baryon multiplets like those that occurred in the meson multiplets (C.23) and (C.27), since for baryons there are no antiquarks involved.

C.4 BRANCHING RATIOS

Up to now we have concentrated on the use of isospin as a symmetry to classify hadron states. It also leads to many useful relations between the rates of reactions involving members of the same isospin multiplets. Such applications are especially useful in strong interactions, where isospin is conserved, and in particular they lead to predictions for branching ratios for resonance decays.

To illustrate this, we start by calculating the branching ratios of the $\Delta^+(1232)$ resonance, which has two decay modes

$$\Delta^+ \to p\pi^0, \quad n\pi^+. \tag{C.30}$$

From (C.29) we see that the Δ^+ has $I = \frac{3}{2}$ and $I_3 = \frac{1}{2}$, and so by isospin conservation must decay to a pion–nucleon state with the same quantum numbers, which we denote

$$\left| N\pi; \frac{3}{2}, \frac{1}{2} \right\rangle. \tag{C.31}$$

To predict the branching ratios we must relate this state to specific charge states like $p\pi^0$ and $n\pi^+$. Since nucleons have $I = \frac{1}{2}$ and pions have $I = 1$, possible $N\pi$ states have $I = \frac{3}{2}$ or $\frac{1}{2}$. As usual, the easiest state to identify is the one with the maximum values $I = I_3 = \frac{3}{2}$, since the only possible combination, which gives $I = \frac{3}{2}$, is

$$\left| N\pi; \frac{3}{2}, \frac{3}{2} \right\rangle = \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 1\rangle . \tag{C.32}$$

The other possible states can be deduced from this by using the shift operators and orthogonality. For example, to obtain (C.31) we act on both sides of (C.32) with $\hat{I}_-$ using (C.16). This gives

$$\hat{I}_- \left| N\pi; \frac{3}{2}, \frac{3}{2} \right\rangle = \sqrt{3} \left| N\pi; \frac{3}{2}, \frac{1}{2} \right\rangle$$

for the left-hand side and

$$\hat{I}_- \left[\left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 1\rangle \right] = \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle + \sqrt{2} \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle$$

for the right-hand side. Comparing these, we obtain the desired results

$$\begin{aligned} \left| N; \frac{3}{2}, \frac{1}{2} \right\rangle &= \frac{1}{\sqrt{3}} \left| N; \frac{1}{2}, -\frac{1}{2} \right\rangle |\pi; 1, 1\rangle + \sqrt{\frac{2}{3}} \left| N; \frac{1}{2}, \frac{1}{2} \right\rangle |\pi; 1, 0\rangle \\ &= -\frac{1}{\sqrt{3}} \left| n\pi^+ \right\rangle + \sqrt{\frac{2}{3}} \left| p\pi^0 \right\rangle, \end{aligned} \tag{C.33}$$

where in the second equation we have identified the pion and nucleon charge states. Thus we see that isospin conservation predicts that the Δ^+ has a probability of 1/3

to decay into $n\pi^+$ and a probability of 2/3 to decay into $p\pi^0$. This is in excellent agreement with experiment.

We next turn to the N^+ resonances, which are excited states of the proton with the isospin quantum numbers $I = \frac{1}{2}, I_3 = \frac{1}{2}$, and can decay to pion–nucleon final states via

$$N^+ \to p\pi^0, \quad n\pi^+. \tag{C.34}$$

In addition, they can decay to more complicated final states allowed by conservation laws, like $p\pi^+\pi^-$, $n\pi^+\pi^0$, etc. We will not consider these here, but just use isospin conservation to predict the relative decay rates for the two reactions (C.34).

To do this, we must construct a pion–nucleon final state with the same quantum numbers $I = \frac{1}{2}, I_3 = \frac{1}{2}$ as the decaying N^+ state. This must be of the general form

$$\left|N\pi; \frac{1}{2}, \frac{1}{2}\right\rangle = \alpha \left|N; \frac{1}{2}, -\frac{1}{2}\right\rangle |\pi; 1, 1\rangle + \beta \left|N; \frac{1}{2}, \frac{1}{2}\right\rangle |\pi; 1, 0\rangle ,$$

since there are no other πN combinations with $I_3 = \frac{1}{2}$. The coefficients α, β are then determined, up to an arbitrary phase factor, by requiring that the state be normalized and orthogonal to the $I = \frac{3}{2}, I = \frac{1}{2}$ state (C.33). In this way we obtain

$$\left|N\pi; \frac{1}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}} \left|N; \frac{1}{2}, -\frac{1}{2}\right\rangle |\pi; 1, 1\rangle - \sqrt{\frac{1}{3}} \left|N; \frac{1}{2}, \frac{1}{2}\right\rangle |\pi; 1, 0\rangle , \tag{C.35}$$

where we have chosen the overall phase in accordance with popular convention. On identifying the pion and nucleon charge states, (C.27) and (C.29), this becomes

$$\left|N\pi; \frac{1}{2}, \frac{1}{2}\right\rangle = -\sqrt{\frac{2}{3}} \left|n\pi^+\right\rangle - \sqrt{\frac{1}{3}} \left|p\pi^0\right\rangle,$$

implying

$$\Gamma(N^+ \to n\pi^+) \big/ \Gamma(N^+ \to p\pi^0) = 2$$

for the ratio of the decay rates. This is in contrast to the ratio

$$\Gamma(\Delta^+ \to n\pi^+) \big/ \Gamma(\Delta^+ \to p\pi^0) = 1/2,$$

which follows for Δ^+ resonances from (C.33) and is one of the means by which N^+ and Δ^+ resonances can be experimentally distinguished.

C.5 SPIN STATES

Finally, we stress again that the isospin formalism described above has been constructed in exact analogy to the theory of angular momentum. Conversely, all the results that we have derived for isospin states $|A : I, I_3\rangle$ apply equally well to ordinary spin states $|A : S, S_3\rangle$. In particular, Equations (C.31), (C.33) and (C.35) apply equally

well to the combined spin states of a spin-$\frac{1}{2}$ particle A and a spin-1 particle B, when they become

$$\left|AB;\frac{3}{2},\frac{3}{2}\right\rangle \equiv \left|AB;S=\frac{3}{2},S_z=\frac{3}{2}\right\rangle = \left|A;\frac{1}{2},\frac{1}{2}\right\rangle |B;1,1\rangle\,, \tag{C.36a}$$

$$\left|AB;\frac{3}{2},\frac{1}{2}\right\rangle = \sqrt{\frac{1}{3}}\left|A;\frac{1}{2},-\frac{1}{2}\right\rangle |B;1,1\rangle + \sqrt{\frac{2}{3}}\left|A;\frac{1}{2},\frac{1}{2}\right\rangle |B;1,0\rangle \tag{C.36b}$$

and

$$\left|AB;\frac{1}{2},\frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\left|A;\frac{1}{2},-\frac{1}{2}\right\rangle |B;1,1\rangle - \sqrt{\frac{1}{3}}\left|A;\frac{1}{2},\frac{1}{2}\right\rangle |B;1,0\rangle\,. \tag{C.36c}$$

This last equation was used in Section 6.2.4 to derive the magnetic moments of the $J^P = \frac{1}{2}^+$ baryons aab, with the identification $A = b$ and $B = aa$.

PROBLEMS C

C.1 Use the relations (C.1) and (C.3) to verify the commutation relations (C.2).

C.2 Use the relations (C.16) and (C.17) to show that the $2I + 1$ members of an isomultiplet $|I, I_3\rangle$ all have the same energy if isospin is conserved.

C.3 Use the definitions (C.8) to verify that Equations (C.15) and (C.16) are satisfied for the quark and antiquark assignments (C.19).

C.4 A resonance $X^0(1520)$ decays via the strong interaction to the final states $n\pi^0$ and $p\pi^-$ with branching ratios of approximately 18 and 36 %, respectively. What is its isospin?

C.5 The $\Lambda(1405)$ resonance decays by the strong interaction to $\Sigma\pi$ final states with a branching ratio of 42 %. What are the individual branching ratios for the charged modes $\Sigma^+\pi^-$, $\Sigma^0\pi^0$ and $\Sigma^-\pi^+$ if the $\Lambda(1405)$ is an isosinglet state?

D

★Gauge Theories

Gauge theories have a particular type of symmetry, called *gauge invariance*. In this appendix we shall explain what is meant by gauge invariance and discuss its consequences. This is important because QED, QCD and the unified electroweak theory are all examples of gauge theories, with somewhat different forms of gauge invariance in each case. The simplest type of gauge invariance is that found in electromagnetism, and it is on this case that we initially concentrate in Sections D.1 to D.5. In particular, in Section D.4, we shall use it to introduce the so-called *gauge principle*, which makes gauge invariance the fundamental requirement from which the detailed properties of the interaction are deduced. This principle is applied to the strong interactions (QCD) and the unified electroweak theory in Sections D.6 and D.7, respectively.

Another thing that will emerge in Section D.3 is that gauge invariance implies that the photon is strictly massless. While this is not a problem for electromagnetic interactions, it is a problem for the analogous case of the electroweak theory, where the $W^{\pm}$ and Z^0 are massive bosons. We have already mentioned briefly in Section 9.2.2 that nonzero masses can be generated by spontaneous symmetry breaking. This is illustrated in Section D.5 for the simple case of hypothetical massive photons.

The topics of this appendix are usually discussed in more advanced books, where advantage can be taken of students' greater knowledge to use explicitly Lorentz covariant Lagrangian field theory.[1] However, in keeping with the level of the rest of this book, we have avoided four-vector notation and based our discussion directly on the relevant equations of motion.

[1] This approach may be found in many more advanced books, e.g. Chapters 12 and 13 of Mandl and Shaw (1993) or Chapter 14 of Halzen and Martin (1984).

Particle Physics Third Edition B. R. Martin and G. Shaw

D.1 ELECTROMAGNETIC INTERACTIONS

The concept of gauge invariance has its origin in the study of electromagnetic interactions, and in order to introduce it we first consider the wave equation for a particle of mass m and charge q moving nonrelativistically in an electromagnetic field.[2] In the case of an electrostatic field $\phi(\mathbf{r})$ this is obtained by making the substitution

$$\mathbf{p} \to -i\boldsymbol{\nabla} \tag{D.1}$$

in the classical Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + q\,\phi(\mathbf{r})$$

to give the familiar Schrödinger equation

$$H\,\Psi(\mathbf{r},t) = \left[-\frac{1}{2m}\boldsymbol{\nabla}^2 + q\,\phi(\mathbf{r})\right]\Psi(\mathbf{r},t) = i\frac{\partial\Psi(\mathbf{r},t)}{\partial t}. \tag{D.2}$$

The wave equation for a general electromagnetic field is obtained in a similar way from the classical Hamiltonian

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\,\phi, \tag{D.3}$$

where $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the scalar and vector potentials defined by[3]

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}, \quad \mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}. \tag{D.4}$$

The Hamiltonian (D.3) is justified by the fact that it can be shown[4] to lead to the familiar classical equation of motion

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = q\,(\mathbf{E} + \mathbf{v}\times\mathbf{B}) \tag{D.5}$$

for a charged particle moving with velocity $\mathbf{v}$ in electric and magnetic fields given by (D.4). On making the substitution (D.1), the Hamiltonian (D.3) also leads to the desired wave equation

$$H\,\Psi = \left[-\frac{1}{2m}(\boldsymbol{\nabla} - iq\mathbf{A})^2 + q\,\phi\right]\Psi = i\frac{\partial\Psi}{\partial t},$$

---

[2] In this appendix, we work in units that coincide with the Heavyside–Lorentz or rationalized Gaussian electromagnetic units when factors of $\hbar$ and c are restored. These units, in which $\varepsilon_0 = 1$, are almost invariably used when this topic is discussed. To convert to the units used elsewhere in the book – which coincide with SI units when factors of $\hbar$ and c are restored – make the replacements $q \to q/\sqrt{\varepsilon_0}$, $\phi \to \sqrt{\varepsilon_0}\phi$, $\mathbf{A} \to \sqrt{\varepsilon_0}\mathbf{A}$, $\mathbf{E} \to \sqrt{\varepsilon_0}\,\mathbf{E}$ and $\mathbf{B} \to \sqrt{\varepsilon_0}\,\mathbf{B}$.

[3] See, for example, p. 229 of Grant and Phillips (1990).

[4] See Problem D.1.

which is conveniently written in the form

$$i\left(\frac{\partial}{\partial t}+iq\phi\right)\Psi=-\frac{1}{2m}\left(\boldsymbol{\nabla}-iq\mathbf{A}\right)^2\Psi \tag{D.6}$$

for what follows.

D.2 GAUGE TRANSFORMATIONS

The wave equation (D.6) determines the motion of a charged particle in an arbitrary electromagnetic field described by the potentials $(\phi,\mathbf{A})$. However, these potentials are not unique, since for any choice $(\phi,\mathbf{A})$ there are many other choices $(\phi',\mathbf{A}')$ that give rise to the same values for the electric and magnetic fields. In particular, one easily shows that substituting

$$\phi\rightarrow\phi'=\phi+\frac{\partial f}{\partial t},\quad \mathbf{A}\rightarrow\mathbf{A}'=\mathbf{A}-\boldsymbol{\nabla}f \tag{D.7}$$

in (D.4) leaves the electric and magnetic fields unchanged, where $f(\mathbf{r},t)$ is an arbitrary scalar function.

The replacement of $(\phi,\mathbf{A})$ by $(\phi',\mathbf{A}')$ in accordance with (D.7) is called a *gauge transformation*, and a theory whose physical predictions are unaltered by such a transformation is said to be *gauge-invariant*. Because observable quantities depend only on the fields $\mathbf{E}$ and $\mathbf{B}$, and not on the particular choice of potentials used to describe them, it is a fundamental requirement that any theory formulated in terms of these potentials should be gauge-invariant. At first sight the wave equation (D.6) seems to contradict this requirement, since (D.6) and (D.7) together lead to a completely different wave equation,

$$i\left(\frac{\partial}{\partial t}+iq\phi'-iq\frac{\partial f}{\partial t}\right)\Psi=-\frac{1}{2m}\left(\boldsymbol{\nabla}-iq\mathbf{A}'-iq\boldsymbol{\nabla}f\right)^2\Psi,$$

in terms of the potentials $(\phi',\mathbf{A}')$. This problem is resolved by noting that in obtaining this equation we assumed that the wavefunction Ψ remained unchanged when the electromagnetic potentials underwent the gauge transformation (D.7). If instead we assume that Ψ simultaneously transforms as

$$\Psi(\mathbf{r},t)\rightarrow\Psi'(\mathbf{r},t)=\exp[-iqf(\mathbf{r},t)]\,\Psi(\mathbf{r},t), \tag{D.8}$$

then from (D.7) and (D.8) we can show that

$$\left(\frac{\partial}{\partial t}+i\,q\phi'\right)\Psi'=\mathrm{e}^{-iqf}\left(\frac{\partial}{\partial t}+i\,q\phi\right)\Psi$$

and

$$(\boldsymbol{\nabla}-iq\mathbf{A}')^2\Psi'=(\boldsymbol{\nabla}-iq\mathbf{A}')\mathrm{e}^{-iqf}(\boldsymbol{\nabla}-iq\mathbf{A})\Psi=\mathrm{e}^{-iqf}(\boldsymbol{\nabla}-iq\mathbf{A})^2\Psi.$$

The right-hand sides of these equations are related by (D.6), and thus the left-hand sides lead to an identical equation

$$i\left(\frac{\partial}{\partial t}+iq\phi'\right)\Psi'=-\frac{1}{2m}\left(\boldsymbol{\nabla}-iq\mathbf{A}'\right)^2\Psi'$$

for the transformed potentials and wavefunction defined by (D.7) and (D.8). We thus arrive at the conclusion that the fundamental equation of motion is indeed gauge-invariant, provided that we extend our definition of a gauge transformation to include (D.8) in addition to (D.7).

D.3 GAUGE INVARIANCE AND THE PHOTON MASS

So far we have discussed the equation of motion of a charged particle in the electromagnetic potentials $(\phi,\mathbf{A})$, but we have not discussed the equations of motion of the potentials themselves. In this section we shall consider these equations and show that they are only gauge-invariant if free photons have zero mass.

The equations of motion of the electromagnetic field in free space are Maxwell's equations

$$\text{div}\,\mathbf{B}=0,\quad \text{curl}\,\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} \tag{D.9a}$$

and

$$\text{div}\,\mathbf{E}=0,\quad \text{curl}\,\mathbf{B}=\frac{\partial\mathbf{E}}{\partial t}. \tag{D.9b}$$

These can be re-expressed in terms of the potentials by substituting (D.4) for $\mathbf{E}$ and $\mathbf{B}$. The first two equations (D.9a) are automatically satisfied for any choice of potentials $(\phi,\mathbf{A})$. The second two equations (D.9b) lead to Maxwell's equations for the potentials themselves, which are conveniently written, for what follows, in the form

$$\left(\frac{\partial^2}{\partial t^2}-\nabla^2\right)\phi-\frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t}+\boldsymbol{\nabla}\cdot\mathbf{A}\right)=0 \tag{D.10a}$$

and

$$\left(\frac{\partial^2}{\partial t^2}-\boldsymbol{\nabla}^2\right)\mathbf{A}+\boldsymbol{\nabla}\left(\frac{\partial\phi}{\partial t}+\boldsymbol{\nabla}\cdot\mathbf{A}\right)=\mathbf{0}. \tag{D.10b}$$

These equations are gauge-invariant and their interpretation is facilitated by 'choosing a gauge'; i.e. we take advantage of the ambiguity (D.7) in the choice of electromagnetic potentials to impose a convenient subsidiary condition. In particular, for any set of fields $(\tilde{\phi},\tilde{\mathbf{A}})$ satisfying (D.10), we can always find a function $f(\mathbf{r},t)$ such that the transformed fields

$$\phi = \tilde{\phi} + \frac{\partial f}{\partial t}, \quad \mathbf{A} = \tilde{\mathbf{A}} - \boldsymbol{\nabla} f$$

satisfy the so-called *Lorentz condition*

$$\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} = 0. \tag{D.11}$$

Imposing this condition is called 'choosing a Lorentz gauge' and we are free to make this choice without restricting or altering the physical content of the theory in any way. When we do so, Equations (D.10) become

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = 0, \quad \left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \mathbf{0}. \tag{D.12}$$

Their interpretation follows directly from comparison with the Klein–Gordon equation (1.5) for a particle of mass m, which is

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \Psi + m^2 \Psi = 0 \tag{D.13}$$

in natural units. Equations (D.12) are of exactly the same form as (D.13), except that the associated particles – photons – have zero mass.

Photons with nonzero masses (or any other massive spin-1 particles) would be described by the *Proca equations*

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi - \frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) + m_\gamma^2 \phi = 0 \tag{D.14a}$$

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} + \boldsymbol{\nabla} \left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) + m_\gamma^2 \mathbf{A} = \mathbf{0}, \tag{D.14b}$$

as we shall shortly demonstrate. Firstly, we note that the Proca equations differ from the gauge-invariant Maxwell equations (D.10) by the addition of 'mass terms' $m_\gamma^2 \phi$ and $m_\gamma^2 \mathbf{A}$, which are obviously not invariant under (D.7). Hence the Proca equations are only gauge-invariant in the case $m_\gamma = 0$, when they reduce to Maxwell's equations. In other words, gauge invariance requires the photon to be massless.

It remains to show that Equations (D.14) really do describe particles of mass $m_\gamma \neq 0$ and to consider how this would affect the behaviour of the electric and magnetic fields (D.4). To do this we differentiate (D.14a) with respect to time, take the divergence of (D.14b) and add the two together. After a short calculation, this gives

$$m_\gamma^2 \left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \right) = 0.$$

Consequently, if $m_\gamma \neq 0$ the Lorentz condition (D.11) necessarily holds (there is no choice!)[5] and the Proca equations reduce to the Klein–Gordon equations[6]

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi + m_\gamma^2\phi = 0 \tag{D.15a}$$

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} + m_\gamma^2\mathbf{A} = \mathbf{0} \tag{D.15b}$$

for particles of mass $m_\gamma \neq 0$. Such a mass could modify the characteristic long-range behaviour of the electromagnetic fields derived from the potentials. For example, in the case of electrostatics, $\phi(\mathbf{r}, t) \to \phi(\mathbf{r})$ and (D.15a) becomes

$$\nabla^2\phi(\mathbf{r}) - m_\gamma^2\phi(\mathbf{r}) = 0,$$

which, following the discussion of Section 1.4.2, leads to a Coulomb potential of the Yukawa form

$$\phi(\mathbf{r}) = -\frac{e^2}{4\pi}\cdot\frac{e^{-r/R}}{r},$$

where $R = 1/m_\gamma$, or $\hbar/m_\gamma c$ in practical units. Laboratory tests of the Coulomb law provide the upper limit $m_\gamma \leq 10^{-14}\,\mathrm{eV/c^2}$, while similar studies on the variation of the magnetic field of Jupiter over large distances give $m_\gamma \leq 6 \times 10^{-17}\,\mathrm{eV/c^2}$. In short, the prediction of zero photon mass, as required by gauge invariance, has been confirmed experimentally to quite extraordinary accuracy.

D.4 THE GAUGE PRINCIPLE

In Section D.2, invariance of the wave equation (D.6) under a gauge transformation (D.7) of the electromagnetic potentials was shown to require that the wavefunction also underwent a gauge transformation (D.8). With the development of new gauge theories – QCD and the unified electroweak theory – it has become fashionable to reverse this argument for reasons that will become clear shortly. In this approach, invariance of the wave equation under the gauge transformation (D.8) is taken as the fundamental requirement and used to infer the form of the interactions. This is called the *principle of minimal gauge invariance*, or the *gauge principle* for short, and we shall introduce it by considering the electromagnetic interactions of relativistic spin-$\frac{1}{2}$ particles.[7]

---

[5] For a gauge-invariant theory there is always a choice, since a gauge transformation (D.7) can change a set of potentials that satisfy the Lorentz condition into one that does not, or vice versa.

[6] In essence, the Proca equations combine the Klein–Gordon equations (D.15) and the Lorentz condition (D.11) into a single set of equations, provided the mass is nonzero. The Lorentz condition effectively reduces the four functions $(\phi, \mathbf{A}) \equiv (\phi, A_x, A_y, A_z)$ to three independent components, corresponding to the three possible spin states of a massive spin-1 particle.

[7] We start with the relativistic case because the nonrelativistic case is algebraically more complicated. (We will comment on the nonrelativistic case later.)

In Chapter 1, we saw that in the absence of interactions a relativistic spin-$\frac{1}{2}$ particle obeys the Dirac equation (1.10), i.e.

$$i\frac{\partial \Psi}{\partial t} = -i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\Psi + \beta m\Psi, \tag{D.16}$$

where we have introduced the notation

$$\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} \equiv \sum_{i=1}^{3} \alpha_i \frac{\partial}{\partial x_i}$$

and set $\hbar = 1$ and $c = 1$ as usual. However, this equation is not invariant under the fundamental gauge transformation (D.8), since the transformed wavefunction Ψ' satisfies

$$i\left(\frac{\partial}{\partial t} + iq\frac{\partial f}{\partial t}\right)\Psi' = -i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla} + iq\boldsymbol{\nabla}f)\,\Psi' + \beta m\Psi', \tag{D.17}$$

as is easily verified by substituting (D.8) into (D.16). This problem is remedied by adding just those terms involving the electromagnetic potentials that are needed to make the Dirac equation (D.16) gauge-invariant. To see how to do this, we note that the troublesome terms in (D.17) involve $\partial f/\partial t$ and $\boldsymbol{\nabla}f$, and originate from the derivatives $\partial\Psi/\partial t$ and $\boldsymbol{\nabla}\Psi$ in (D.16), and that the same derivatives of $f(\mathbf{r},t)$ occur in the gauge transformations (D.7) of ϕ and **A,** respectively. Hence it should be possible to add one suitably chosen term involving a potential to each of the derivatives in (D.16) in such a way that the troublesome terms cancel when one makes a gauge transformation. This leads to the equation

$$i\left(\frac{\partial}{\partial t} + iq\phi\right)\Psi = -i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla} - iq\mathbf{A})\,\Psi + \beta m\Psi, \tag{D.18}$$

which differs from (D.16) by the so-called minimal substitutions

$$\frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + iq\phi, \quad \boldsymbol{\nabla} \to \boldsymbol{\nabla} - iq\mathbf{A}. \tag{D.19}$$

We leave it as an exercise for the reader to verify that (D.18) is indeed gauge-invariant. In the familiar case of an electron, $q = -e$ and (D.18) becomes

$$i\left(\frac{\partial}{\partial t} - ie\phi\right)\Psi = -i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla} + ie\mathbf{A})\,\Psi + \beta m\Psi. \tag{D.20}$$

This is precisely the equation of motion that is used for the electron in quantum electrodynamics with such spectacular success.[8]

D.5 THE HIGGS MECHANISM

In Section D.3, we showed that Maxwell's equations in free space imply that photons have zero mass and that, in the absence of interactions, introducing a hypothetical nonzero photon mass necessarily violates gauge invariance. However, in Section 9.2.2, we stated that when the electromagnetic field interacts with a complex scalar field, spontaneous symmetry breaking can generate a nonzero mass without violating the gauge invariance of the interactions. In this section we will give a simple account of how this arises.[9] We will do this by, firstly, in Sections D.5.1 and D.5.2, exploring the equations of motion of the electromagnetic fields in the presence of interactions, and then discuss, in Section D.5.3, the result of spontaneous symmetry breaking itself.

D.5.1 Charge and current densities

In classical physics, the sources of the electromagnetic field are the electric charge and current densities $\rho(\mathbf{r},t)$ and $\mathbf{j}(\mathbf{r},t)$, respectively, which satisfy the continuity condition

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0 \tag{D.21}$$

in order to guarantee electric charge conservation. When these sources are included, the first pair of Maxwell's equations, (D.9a), remains unchanged, while the second pair, (D.9b), becomes

$$\operatorname{div}\mathbf{E} = \rho, \quad \operatorname{curl}\mathbf{B} = \mathbf{j} + \frac{\partial\mathbf{E}}{\partial t}, \tag{D.9c}$$

in our units.[10] On expressing the $\mathbf{E}$ and $\mathbf{B}$ field in terms of the potentials using (D.4), Equations (D.9a) are again automatically satisfied, while Equations (D.9c) lead to Maxwell's equations for the potentials in the presence of sources. These are conveniently written in the forms

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi - \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) = \rho \tag{D.22a}$$

---

[8] If we had applied our present argument to a nonrelativistic particle, we would have started from the Schrödinger equation

$$i\frac{\partial\Psi}{\partial t} = -\frac{1}{2m}\nabla^2\Psi,$$

which is also not invariant under (D.8). The same minimal substitution (D.19) would then have led us to the wave equation (D.6), which we have already shown to be gauge invariant.

[9] Our discussion completes that of Section 9.2.2. For a more rigorous account, which requires the formalism of quantum field theory, see, for example, Section 13.2 of Mandl (1992).

[10] Cf. Footnote 2.

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) = \mathbf{j}. \tag{D.22b}$$

They differ from our previous Maxwell equations (D.10a) and (D.10b) solely by the addition of the source terms, and since the values of the electric charges and currents are unaffected by gauge transformations (D.7) on the potentials, Equations (D.22) are, like Equations (D.10), explicitly gauge-invariant.

In quantum physics, the classical charge and current densities are replaced by charge and current densities expressed in terms of fields of the associated charged particles. For point-like spin-$\frac{1}{2}$ fermions, these take the form

$$\rho = q\Psi^\dagger\Psi, \quad \mathbf{j} = q\Psi^\dagger\boldsymbol{\alpha}\Psi, \tag{D.23}$$

where $\boldsymbol{\alpha}$ are the Dirac matrices (1.9) and $\Psi^\dagger$ is the 'Hermitian conjugate spinor'

$$\Psi^\dagger \equiv (\Psi_1^*, \Psi_2^*, \Psi_3^*, \Psi_4^*). \tag{D.24}$$

It is straightforward to show that the charge and current densities (D.23) satisfy the continuity condition (D.21) by substituting (D.23) into (D.21) and using the Dirac equation (D.16) together with the equation

$$-i\frac{\partial \Psi^\dagger}{\partial t} = i\boldsymbol{\nabla}\Psi^\dagger\cdot\boldsymbol{\alpha} + \beta m\Psi^\dagger, \tag{D.25}$$

obtained by taking the Hermitian conjugate of (D.16) and using the fact that the matrices $\boldsymbol{\alpha}$ and β must be Hermitian to ensure that the Hamiltonian (1.8) is Hermitian. In addition, the charge and current densities are both invariant under the combined gauge transformations (D.7) and (D.8). Hence, substituting (D.23) into (D.22) gives an equation of motion for the electromagnetic field that is gauge-invariant as required. This is precisely the equation of motion used for the electromagnetic field in QED, where the charge $q = -e$ for the electron.

D.5.2 Spin-0 bosons

In order to understand the Higgs mechanism, we first need to extend the above formalism to include charged spin-0 bosons rather than spin-$\frac{1}{2}$ fermions. Such particles are described by a complex field $\eta(\mathbf{r}, t)$, which satisfies the Klein–Gordon equation (1.5), i.e.

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\eta(\mathbf{r}, t) = 0. \tag{D.26}$$

In this case, charge and current densities that satisfy the continuity condition (D.21) are given by

$$\rho = iq\left(\eta^*\frac{\partial\eta}{\partial t} - \frac{\partial\eta^*}{\partial t}\eta\right), \quad \mathbf{j} = -iq\left(\eta^*\boldsymbol{\nabla}\eta - \eta\boldsymbol{\nabla}\eta^*\right), \tag{D.27}$$

as can be verified by substituting (D.27) into (D.21) and using the Klein–Gordon equation (D.26). Implicit in this is that, for particles of charge q, the probability amplitude for position, ρ_P, analogous to $\Psi^*\Psi$ for the Schrödinger equation, or $\Psi^\dagger\Psi$ for the Dirac equation, is given by

$$\rho_P = -i\left(\eta^*\frac{\partial\eta}{\partial t} - \frac{\partial\eta^*}{\partial t}\eta\right). \tag{D.28}$$

This requires that $\rho_P(\mathbf{r},t)$ must not only satisfy a continuity equation of the form (D.21) but must also be positive definite. One can easily verify that this is true for positive energy solutions of the form (1.6a), but not for negative energy solutions of the form (1.6b), for which $\rho_P(\mathbf{r},t) < 0$. Here we shall simply confine ourselves to positive energy solutions and note that a more rigorous treatment using quantum field theory, which incorporates antiparticles of charge $-q$ associated with the negative energy solutions, yields the same expressions (D.27) for the electric charge and current densities.

The most striking thing about the densities (D.27) is that they contain derivatives, essentially because the Klein–Gordon equation contains second-order derivatives, in contrast to the Dirac equation, which is first-order. Because of these derivatives, the charge and current densities are not invariant under the gauge transformation (cf. Equation (9.43))

$$\eta(\mathbf{r},t) \to \exp[-iqf(\mathbf{r},t)]\eta(\mathbf{r},t), \tag{D.29}$$

but rather transform as

$$\rho \to \rho + 2q^2\eta^*\eta\frac{\partial f}{\partial t}, \quad \mathbf{j} \to \mathbf{j} - 2q^2\eta^*\eta\boldsymbol{\nabla}f. \tag{D.30}$$

Thus, in this case, the proposed equations of motion for the electromagnetic fields (D.22a) and (D.22b) do not satisfy the fundamental requirement of gauge invariance under the combined gauge transformations (D.7) and (D.29). To resolve this problem, we invoke the gauge principle discussed in Section D.4 and add suitably chosen interaction terms to (D.22a) and (D.22b) to ensure that the resulting equations are gauge-invariant. This is achieved by modifying (D.22a) and (D.22b) to be

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi - \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) = \rho - 2q^2\eta^*\eta\phi \tag{D.31a}$$

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) = \mathbf{j} - 2q^2\eta^*\eta\mathbf{A}. \tag{D.31b}$$

We leave it to the reader to verify that the right-hand sides of (D.31a) and (D.31b), and hence the equations themselves, are gauge-invariant. These are precisely the equations of motion for the electromagnetic field used to describe the quantum electrodynamics of spin-0 bosons with charge q. However, if spontaneous symmetry breaking occurs, it profoundly alters the physical interpretation of these equations, as we shall immediately see.

D.5.3 Spontaneous symmetry breaking

In Section 9.2.2, we saw that when spontaneous symmetry breaking occurs the field $\eta(\mathbf{r}, t)$ has a nonzero value in the vacuum state that can, without loss of generality, be chosen to be (cf. Equation (9.42))

$$\eta_0 = \frac{v}{\sqrt{2}} \equiv \left(-\frac{\mu^2}{2\lambda}\right)^{1/2} \geq 0. \tag{D.32}$$

This clearly violates gauge invariance and to understand its consequences, we rewrite the theory in terms of a new field $\sigma(\mathbf{r}, t)$ that does vanish in the vacuum state, defined by

$$\eta(\mathbf{r}, t) \equiv \frac{1}{\sqrt{2}}\left[v + \sigma(\mathbf{r}, t)\right]. \tag{D.33}$$

In addition, as in Section (D.3), we again 'choose a gauge' that makes the physical interpretation more transparent. In this case, however, instead of imposing the Lorentz condition (D.11), we note that for any set of fields $(\hat{\phi}, \hat{\mathbf{A}}, \hat{\eta})$ we can always find a gauge function $f(\mathbf{r}, t)$ such that the transformed field

$$\eta(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)]\hat{\eta}(\mathbf{r}, t),$$

and hence the corresponding field $\sigma(\mathbf{r}, t)$, is real, and therefore

$$\eta^*(\mathbf{r}, t) = \eta(\mathbf{r}, t). \tag{D.34}$$

Imposing (D.34) defines the 'unitary gauge', in which the charge and current densities (D.27) vanish:

$$\rho(\mathbf{r}, t) = 0, \quad \mathbf{j}(\mathbf{r}, t) = \mathbf{0}. \tag{D.35}$$

Substituting (D.33) and (D.35) into (D.31a) and (D.31b) then gives

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi - \frac{\partial}{\partial t}\left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) + q^2v^2\phi = -2q^2v\sigma\phi - q^2\sigma^2\phi \tag{D.36a}$$

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) + q^2\nu^2\mathbf{A} = -2q^2\nu\sigma\mathbf{A} - q^2\sigma^2\mathbf{A} \qquad \text{(D.36b)}$$

as the new equations of motion for the electromagnetic field. On comparing with (D.14a) and (D.14b), we see that they are identical with the Proca equations for 'photons' with nonzero mass

$$m_\gamma = q\nu, \qquad \text{(D.37)}$$

whose interactions with the real scalar field $\sigma(\mathbf{r}, t)$ are specified by the terms on the right-hand side of (D.36). In addition, the vanishing of the electric charge and current densities implies that the quanta of the field $\sigma(\mathbf{r}, t)$ – the Higgs bosons – are electrically neutral.

This is the essence of the Higgs mechanism, whereby a theory of charged bosons interacting with massless photons is converted by spontaneous symmetry breaking into a theory of massive photons interacting with neutral Higgs bosons. This mechanism is not realized in nature for photons, which are assumed to be exactly massless. An analogous mechanism is, however, used in the more complicated unified electroweak theory to generate masses for the $W^{\pm}$ and Z^0 bosons, as discussed in Section 9.2.

D.6 QUANTUM CHROMODYNAMICS

In QED, the gauge principle approach we have outlined in Section D.4 is an interesting way of looking at an interaction whose form was originally determined in other ways. In contrast, in QCD it was used to infer the detailed form of interactions that were previously unknown. To do this, one starts from new types of gauge transformations that are similar to, but more complicated than, the gauge transformations we have discussed previously. In particular, they not only change the phase of the wavefunction as in (D.8), but also the colour state. These colours states have been discussed in Section 6.3, but we repeat the salient points here for convenience.

There are three independent colour wavefunctions $\chi^C = r, g, b$, which are conveniently represented by the colour spinors (6.39):

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \qquad \text{(D.38)}$$

These are acted upon by eight independent colour charge operators $\hat{F}_1, \hat{F}_2, \ldots, \hat{F}_8$, which are represented by the matrices (6.40) and are assumed to be conserved:

$$[\hat{F}_i, H] = 0 \quad (i = 1, 2, \ldots, 8).$$

The commutation relations among the $\hat{F}_i$ were not stated in Section 6.3, but are required below. They follow directly from (6.40) and are

$$[\hat{F}_i, \hat{F}_j] = i \sum_k f_{ijk} \hat{F}_k, \tag{D.39}$$

where some of the nonzero coefficients f_{ijk} are given in Table D.1 and the rest follow from the fact that they are totally antisymmetric:

$$f_{ijk} = -f_{jik} = f_{jki} = -f_{kji} = f_{kij} = -f_{ikj}. \tag{D.40}$$

TABLE D.1 Some of the coefficients in the commutation relations (D.39) of the colour charges $\hat{F}_i (1 = 1, 2, \ldots, 8)$. All the other nonzero coefficients may be deduced from these by using the fact that f_{ijk} is totally antisymmetric.

ijk	123	147	156	246	257	345	367	458	678
f_{ijk}	1	1/2	−1/2	1/2	1/2	1/2	−1/2	$\sqrt{3}/2$	$\sqrt{3}/2$

Because the charges $\hat{F}_i$ do not commute with each other, they cannot all have definite values simultaneously. In fact the colour states $\chi^C = r, g, b$ are chosen to be eigenstates of the charges $\hat{F}_3$ and $\hat{F}_8$, with eigenvalues given by Equation (6.41) and Table 6.8. The remaining charges $\hat{F}_1, \hat{F}_2, \hat{F}_4, \ldots, \hat{F}_7$, change the colour state, as illustrated, for example, by Equation (6.43):

$$\hat{F}_1 r = \frac{1}{2} g, \quad \hat{F}_1 g = \frac{1}{2} r, \quad \hat{F}_1 b = 0. \tag{D.41}$$

Finally, the overall state of any relativistic quark is given by the product of a Dirac spinor $\psi(\mathbf{r}, t)$ with a colour wavefunction

$$\Psi \equiv \psi(\mathbf{r}, t) \chi^C, \tag{D.42}$$

or a linear combination of such wavefunctions.

The basic idea behind QCD is that the colour charges described above act as the sources of the strong interaction, in the same way that the electric charge acts as the source of the electromagnetic interaction. This idea is implemented by modifying the form of gauge invariance imposed. In QED we demanded invariance under the gauge transformation

$$\Psi(\mathbf{r}, t) \to \Psi'(\mathbf{r}, t) = \exp[-iqf(\mathbf{r}, t)] \Psi(\mathbf{r}, t) \tag{D.8}$$

on the wavefunctions of quarks or leptons of charge q. In QCD we demand invariance under the gauge transformations

$$\Psi \to \Psi' = \exp\left[-ig_s \sum_i \hat{F}_i \omega_i(\mathbf{r}, t) \right] \Psi \tag{D.43}$$

on the wavefunctions (D.42) of coloured quarks, where g_s is a coupling constant that will determine the strength of the interaction,[11] and there are now eight arbitrary gauge functions $\omega_i(\mathbf{r},t)(1,2,\ldots,8)$, corresponding to the eight colour charges[12] $\hat{F}_1,\hat{F}_2,\ldots,\hat{F}_8$.

We now use invariance under (D.43) to determine the form of the interaction, following a similar argument to that given for QED in Section D.4. Indeed, if we pretend for a moment that the charges $\hat{F}_i$ commute (i.e. that $f_{ijk}=0$), the argument leading from (D.16) to (D.20) goes through with only trivial modifications. This gives

$$i\left(\frac{\partial}{\partial t}+ig_s\sum_i\hat{F}_i\phi_i\right)\Psi=-i\alpha\cdot\left(\boldsymbol{\nabla}-ig_s\sum_i\hat{F}_i\mathbf{A}_i\right)\Psi+\beta m\Psi \tag{D.44}$$

as the equation of motion analogous to (D.20), where $(\phi_i,\mathbf{A}_i)(i=1,2,\ldots,8)$ are eight new gluon fields, which transform as

$$\phi_i\rightarrow\phi_i'=\phi_i+\frac{\partial\omega_i}{\partial t},\quad \mathbf{A}_i\rightarrow\mathbf{A}'_i=\mathbf{A}_i-\boldsymbol{\nabla}\omega_i\quad(f_{ijk}=0) \tag{D.45}$$

in analogy with (D.7) for the electromagnetic field $(\phi,\mathbf{A})$. However, $f_{ijk}\neq 0$ and the transformation law (D.45) has to be modified if (D.44) is to remain invariant when this is taken into account. In practice, it is sufficient to restrict ourselves to infinitesimal transformations, corresponding to small $\omega_i(\mathbf{r},t)$[13] when (D.45) becomes

$$\phi_i\rightarrow\phi_i'=\phi_i+\frac{\partial\omega_i}{\partial t}+g_s\sum_{jk}f_{ijk}\omega_j\phi_k \tag{D.46a}$$

and

$$\mathbf{A}_i\rightarrow\mathbf{A}'_i=\mathbf{A}_i-\boldsymbol{\nabla}\omega_i+g_s\sum_{jk}f_{ijk}\omega_j\mathbf{A}_k. \tag{D.46b}$$

The proof that this is correct will be given shortly. Firstly, we consider its physical interpretation.

The key to the interpretation of (D.46) lies in the form of the interaction terms in (D.44). For example, for a red quark, $\Psi=\psi(\mathbf{r},t)\,r$, Equation (D.44) contains interaction terms like

$$-g_s\hat{F}_1\phi_1\psi\,r=-\frac{g_s}{2}\phi_1\psi\,g,$$

---

[11] It is related to α_s used in Section 7.1 by $\alpha_s=g_s^2/4\pi$, in analogy to $\alpha=e^2/4\pi\varepsilon_0$ for the electromagnetic interaction, where $\varepsilon_0=1$ in this appendix.

[12] The parallel between (D.8) and (D.43) would be even more transparent if we wrote $q=eQ$, where e characterizes the strength of the electromagnetic interaction and $Q=0,-\frac{1}{3},\frac{2}{3}$ or -1, the charge of the quark or lepton concerned.

[13] This is because finite transformations can be treated as the limit as $N\rightarrow\infty$ of N successive small transformations.

in which the colour charge of the quark changes. By colour charge conservation, this is only possible if the gluon fields $(\phi_i, \mathbf{A}_i)$ can carry away colour charge; i.e. the gluons are coloured. Finally, if the colour charges $\hat{F}_i$ are the sources of the gluon fields $(\phi_i, \mathbf{A}_i)$, and gluons are coloured, this implies that gluons interact with other gluons. This has already been discussed qualitatively in Section 7.1.[14] Here we note that the existence of such interactions and their detailed form are determined by the transformation law (D.46). To see this, we note that free gluons, like free photons, would be described by equations of the form (D.10)

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi_i - \frac{\partial}{\partial t}\left(\frac{\partial \phi_i}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}_i\right) = 0 \tag{D.47a}$$

and

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A}_i + \boldsymbol{\nabla}\left(\frac{\partial \phi_i}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}_i\right) = \mathbf{0}. \tag{D.47b}$$

These equations are invariant under the transformations (D.45), but they are not invariant under the required gauge transformation (D.46) because of the final terms proportional to f_{ijk}. So once again one is forced to add the minimal interaction terms necessary to make a free-particle wave equation gauge-invariant. These terms can be shown to lead to gluon–gluon interactions of the type shown in Figure 7.2. However, we shall not pursue this further, because the algebra is complicated and we have not used the detailed results. Rather, we reiterate that the principle of minimal gauge invariance predicts the existence and form of both quark–gluon and gluon–gluon interactions as outlined above, and that these predictions are in agreement with a large body of data and are not contradicted by any known experiment.

We conclude this section by proving that (D.44) is invariant under the gauge transformations (D.43) and (D.46) for infinitesimal $\omega_i(\mathbf{r}, t)$ as stated earlier.[15] For infinitesimal $\omega_i(\mathbf{r}, t)$, (D.43) reduces to

$$\Psi \to \Psi' = \left(1 - ig_s \sum_j \hat{F}_j \omega_j\right)\Psi, \tag{D.48}$$

where we have relabelled the summation variable $i \to j$ for later convenience. By inspection, this implies that (D.44) will be invariant if

$$\left(\frac{\partial}{\partial t} + ig_s \sum_i \hat{F}_i \phi_i'\right)\Psi' = \left(1 - ig_s \sum_j \hat{F}_j \omega_j\right)\left(\frac{\partial}{\partial t} + ig_s \sum_i \hat{F}_i \phi_i\right)\Psi \tag{D.49a}$$

[14] See Figure 7.1, Equations (7.1) and the accompanying discussion.

[15] This proof is included for completeness and is not required in what follows. The reader who wishes to omit it should proceed directly to Section D.7.

and

$$\boldsymbol{\alpha}\cdot\left(\boldsymbol{\nabla}-ig_s\sum_i\hat{F}_i\mathbf{A}'_i\right)\Psi'=\left(1-ig_s\sum_j\hat{F}_j\omega_j\right)\boldsymbol{\alpha}\cdot\left(\boldsymbol{\nabla}-ig_s\sum_i\hat{F}_i\mathbf{A}_i\right)\Psi. \quad \text{(D.49b)}$$

The next step is to write the gauge field transformations in the general form

$$\phi_i\rightarrow\phi'_i=\phi_i+\frac{\partial\omega_i}{\partial t}+\Delta\phi_i \quad \text{(D.50a)}$$

and

$$\mathbf{A}_i\rightarrow\mathbf{A}'_i=\mathbf{A}_i-\boldsymbol{\nabla}\omega_i+\Delta\mathbf{A}_i \quad \text{(D.50b)}$$

and use (D.49) to identify the unknown functions $\Delta\phi_i$ and $\Delta\mathbf{A}_i$. On substituting (D.48) and (D.50a) into (D.49a), we obtain

$$ig_s\sum_i\hat{F}_i\Delta\phi_i=(ig_s)^2\sum_{ij}\left[\hat{F}_i,\hat{F}_j\right]\omega_j\phi_i,$$

where we have neglected terms that are second order in the infinitesimal quantities ω_i and $\Delta\phi_i$. Using the commutation relations (D.39) gives

$$\sum_i\hat{F}_i\Delta\phi_i=-g_s\sum_{ijk}f_{ijk}\hat{F}_k\omega_j\phi_i=g_s\sum_{ijk}f_{ijk}\hat{F}_i\omega_j\phi_k,$$

where we have interchanged the summation labels $i\rightarrow k$ in the last line, using the antisymmetry property (D.40). This implies

$$\Delta\phi_i=g_s\sum_{jk}f_{ijk}\omega_j\phi_k, \quad \text{(D.51)}$$

which, together with (D.50a), is equivalent to the desired transformation law (D.46a). A similar argument applied to (D.49b) establishes the transformation law (D.46b).

D.7 ELECTROWEAK INTERACTIONS

The application of the gauge principle to weak interactions and its role in their unification with the electromagnetic interaction have been discussed qualitatively, for massless electrons and neutrinos, in Section 9.2.1. Here we will supply some of the details that were omitted in that account. In particular, we will specify in more detail the form of the gauge invariance assumed and we will explicitly derive the crucial unification condition (9.8).

The discussion proceeds as follows. We first set up a combined description of electron and neutrino states, which is similar to the description of coloured quark states used in Section D.6. This is then exploited to define a set of gauge transformations,

which lead, via the gauge principle, to the known charged current weak interactions. However, the corresponding neutral current interactions are incorrect, and this problem is resolved by unification with the electromagnetic interaction, leading to the unification condition (9.8). In the final section, we indicate how the spin structure of the interaction and parity violation, which are ignored in the previous discussion, are incorporated into the picture.

Finally, although our discussion is confined to leptons, it is easily extended to quarks using lepton–quark symmetry and quark mixing. This is left as an exercise for the reader (cf. Problems D.6 and D.7).

D.7.1 Weak isospin

To formulate a gauge theory of weak interactions, we require a set of gauge transformations that transforms electrons and electron neutrinos into themselves or into each other, as outlined in Section 9.2.1. In order to do this, we recall that in QCD we saw how to define gauge transformations that transform different quark colour states into each other. We therefore begin by setting up a combined description of electrons and neutrinos, which is closely analogous to the description of coloured quark states used in Section D.6. Specifically, we introduce two independent flavour wavefunctions $\chi^f = \nu, e$, which are conveniently represented by the flavour spinors

$$\nu = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{D.52}$$

in analogy to the quark colour spinors (D.38). Just as the colour spinors are acted on by colour charge operators, the flavour spinors are acted upon by analogous operators, called *weak isospin* operators. These are represented by two-dimensional matrices in the same way that the colour charges are represented by the three-dimensional matrices (6.40). There are three such independent weak isospin operators

$$\hat{I}_i^W = \frac{1}{2}\tau_i \quad (i = 1, 2, 3), \tag{D.53a}$$

where the matrices[16]

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{D.53b}$$

The commutation relations among the $\hat{I}_i^W$ follow directly from (D.53) and are

$$\left[\hat{I}_i^W, \hat{I}_j^W\right] = i \sum_k \varepsilon_{ijk} \hat{I}_k^W, \tag{D.54}$$

[16] The weak isospin operators, like the colour charge operators, must be Hermitian if they are to correspond to physical observables. The most general Hermitian matrices in two or three dimensions can be written as linear combinations of the unit matrix and the τ_i or λ_i, respectively.

where the coefficients ε_{ijk} are totally antisymmetric:

$$\begin{aligned} &\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = 1, \\ &\varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1, \\ &\varepsilon_{ijk} = 0 \text{ if } i = j,\ j = k \text{ or } i = k. \end{aligned} \tag{D.55}$$

The states $\chi^f = \nu, e$ are eigenstates of $\hat{I}_3^W$:

$$\hat{I}_3^W \nu = \frac{1}{2}\nu, \quad \hat{I}_3^W e = -\frac{1}{2}e, \tag{D.56}$$

while $\hat{I}_1^W$ and $\hat{I}_2^W$ interchange electron and neutrino states:

$$\hat{I}_1^W \nu = \frac{1}{2}e, \quad \hat{I}_1^W e = \frac{1}{2}\nu, \quad \hat{I}_2^W \nu = \frac{i}{2}e, \quad \hat{I}_2^W e = -\frac{i}{2}\nu, \tag{D.57}$$

as one easily verifies from (D.52) and (D.53). Finally, the overall state of any relativistic lepton is given by the product of a Dirac spinor[17] $\psi(\mathbf{r}, t)$ with a flavour spinor χ^f:

$$\Psi = \psi(\mathbf{r}, t)\chi^f,$$

or by a linear combination of such states, where we restrict ourselves to the first generation of leptons ($\ell = \nu_e, e$) only.

D.7.2 Gauge invariance and charged currents

We now have a description of electron and electron neutrino states that differs from our description of coloured states (D.38) to (D.42) by the replacements

$$\Psi = \psi(\mathbf{r}, t)\chi^C \to \Psi = \psi(\mathbf{r}, t)\chi^f,$$

$$\hat{F}_i(1, 2, \ldots, 8) \to \hat{I}_i^W(1,\ 2,\ 3)$$

and

$$f_{ijk} \to \varepsilon_{ijk}.$$

[17] We remind the reader that we are for the moment ignoring the spin structure of the interaction. When this is taken into account, the present formalism is applied to 'left-handed' spin states, as we shall see in Section D.7.4.

These can be used to adapt the results derived for QCD to the case of weak interactions. Specifically, we assume gauge invariance under a set of transformations

$$\Psi \to \Psi' \exp\left(-ig\sum_i \hat{I}_i^W f_i(\mathbf{r},t)\right)\Psi \tag{D.58}$$

analogous to (D.43), where g is an arbitrary coupling constant and $f_i(\mathbf{r},t)$ are a set of arbitrary functions analogous to the $\omega_i(\mathbf{r},t)$ of (D.43). The gauge principle then leads to the equation of motion

$$i\left(\frac{\partial}{\partial t}+ig\sum_i \hat{I}_i^W \phi_i^W\right)\Psi = -i\boldsymbol{\alpha}\cdot\left(\boldsymbol{\nabla}-ig\sum_i \hat{I}_i^W \mathbf{W}_i\right)\Psi \tag{D.59}$$

analogous to (D.44), where the leptons are assumed to have zero mass and the gauge fields transform as

$$\phi_i^W \to \phi_i'^W = \phi_i^W + \frac{\partial f_i}{\partial t} + g\sum_{jk}\varepsilon_{ijk} f_j \phi_k^W \tag{D.60a}$$

and

$$\mathbf{W}_i \to \mathbf{W}'_i = \mathbf{W}_i - \boldsymbol{\nabla} f_i + g\sum_{jk}\varepsilon_{ijk} f_j \mathbf{W}_k \tag{D.60b}$$

for infinitesimal $f_i(\mathbf{r},t)$, by analogy with (D.46).

We have now arrived at a set of gauge-invariant equations of motion incorporating interaction terms of a specific form. Their interpretation is best brought out by considering neutrino and electron states separately. For neutrino states $\Psi \equiv \psi(\mathbf{r},t)\nu$, Equation (D.59) becomes

$$\begin{aligned} i\left(\frac{\partial}{\partial t}+i\frac{g}{2}\phi_3^W\right)\psi(\mathbf{r},t)\nu - \frac{g}{2}\left(\phi_1^W + i\phi_2^W\right)\psi(\mathbf{r},t)e \\ = -i\boldsymbol{\alpha}\cdot\left(\boldsymbol{\nabla}-i\frac{g}{2}\mathbf{W}_3\right)\psi(\mathbf{r},t)\nu - \frac{g}{2}\boldsymbol{\alpha}\cdot(\mathbf{W}_1+i\mathbf{W}_2)\psi(\mathbf{r},t)e\,, \end{aligned}$$

or equivalently

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\nu = \frac{g}{2}\left(\phi_3^W - \boldsymbol{\alpha}\cdot\mathbf{W}_3\right)\psi(\mathbf{r},t)\nu + 2g_W\left(\phi^+ - \boldsymbol{\alpha}\cdot\mathbf{W}^+\right)\psi(\mathbf{r},t)e, \tag{D.61a}$$

where we have separated out the interaction terms and introduced the linear combinations

$$\phi^\pm = \frac{1}{\sqrt{2}}(\phi_1^W \pm i\phi_2^W), \quad \mathbf{W}^\pm = \frac{1}{\sqrt{2}}(\mathbf{W}_1^W \pm i\mathbf{W}_2^W) \tag{D.62}$$

and the coupling constant

$$g_W = g/(2\sqrt{2}) \tag{D.63}$$

for later convenience. The corresponding equation for the electron case is

$$i\left(\frac{\partial}{\partial t} + \boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)e = -\frac{g}{2}\left(\phi_3^W - \boldsymbol{\alpha}\cdot\mathbf{W}_3\right)\psi(\mathbf{r},t)e + 2g_W\left(\phi^-\boldsymbol{\alpha}\cdot\mathbf{W}^-\right)\psi(\mathbf{r},t)\nu. \tag{D.61b}$$

The interpretation of these equations is now relatively straightforward. If we consider Equation (D.61a) for the neutrino, we see that the term proportional to g_W couples it to an electron, while the corresponding term in (D.61b) for the electron couples it to a neutrino. These are just the charged current interactions required by experiment, involving the transitions $\nu \to e^- W^+, e^- \to \nu W^-$, etc. In addition, Equations (D.61) contain neutral current interactions involving $(\phi_3, \mathbf{W}_3)$, which couple the electron to an electron and a neutrino to a neutrino, corresponding to transitions $e^- \to e^- W^0$ and $\nu \to \nu W^0$. These terms are of the same form and strength (apart from a factor of $\sqrt{2}$) as the charged current interactions. This is not what is required by experiment, as emphasized in Section 9.2.1.

D.7.3 The unification condition

The problem of the erroneous neutral current terms in (D.61) is resolved by unification with the electromagnetic interaction. To do this, we incorporate the electric charge $q = Qe$ into the theory by introducing the *weak hypercharge* Y^W defined by

$$Q = I_3^W + Y^W, \tag{D.64}$$

which is similar to the definition of ordinary hypercharge Y in Equation (6.5). (Note, however, that the definitions of Y and Y^W differ by a factor of 2.) We then treat the weak hypercharge as a source; i.e. we require gauge invariance under the transformations

$$\Psi \to \Psi' = \exp\left[-ig'\hat{Y}^W\omega(\mathbf{r},t)\right]\Psi, \tag{D.65}$$

where g' is a coupling constant to be determined and $\omega(\mathbf{r},t)$ is an arbitrary function. For the neutrino and electron (ν_e, e^-), we have $I_3^W = (\frac{1}{2}, -\frac{1}{2})$ by (D.56), so that $Y^W = (-\frac{1}{2}, -\frac{1}{2})$ and (D.65) reduces to the transformation

$$\Psi \to \Psi' = \exp\left[ig'\omega(\mathbf{r},t)\,/2\right]\Psi \tag{D.66}$$

in both cases. This is similar in form to the electromagnetic gauge transformation[18] (D.8) that led, via the gauge principle, to the equation of motion (D.18).

[18] In the unified theory, the electromagnetic gauge transformations (D.8) are not imposed directly, but emerge indirectly as an appropriate combination of (D.58) and (D.66). This is because the electric charge is not introduced directly, but indirectly as the combination (D.64) of Y^W and I_3^W.

Equation (D.66) leads to an analogous equation of motion

$$i\left(\frac{\partial}{\partial t}-i\frac{g'}{2}\phi^B\right)\Psi=-i\boldsymbol{\alpha}\cdot\left(\boldsymbol{\nabla}+i\frac{g'}{2}\mathbf{B}\right)\Psi, \tag{D.67a}$$

or more conveniently

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\Psi=-i\frac{g'}{2}\left(\phi^B-i\boldsymbol{\alpha}\cdot\mathbf{B}\right)\Psi, \tag{D.67b}$$

where the leptons are again assumed to have zero mass and the new gauge fields $\phi^B(\mathbf{r},t),\mathbf{B}(\mathbf{r},t)$ transform as

$$\phi^B\rightarrow\phi'^B=\phi^B+\frac{\partial\omega}{\partial t},\quad \mathbf{B}\rightarrow\mathbf{B}'=\mathbf{B}-\boldsymbol{\nabla}\omega, \tag{D.68}$$

by analogy with (D.7).

The next step is to identify the electromagnetic fields $(\phi,\mathbf{A})$ with appropriate linear combinations of the fields $(\phi_3^W,\mathbf{W}_3)$ and $(\phi^B,\mathbf{B})$, as outlined in Section 9.2.1. Since we require gauge invariance under both (D.58) and (D.66) – so far we have treated them separately – we must add the interaction terms in (D.67) resulting from (D.66) to the interaction terms in (D.61a) and (D.61b) resulting from (D.58). For the electron case $\Psi=\psi(\mathbf{r},t)e$ this gives

$$\begin{aligned}i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)e=&\,2g_W\left(\phi^-\mathbf{1}-\boldsymbol{\alpha}\cdot\mathbf{W}^-\right)\psi(\mathbf{r},t)\nu\\&-\frac{1}{2}\left(g\phi_3^W+g'\phi^B\right)\psi(\mathbf{r},t)e+\frac{1}{2}\boldsymbol{\alpha}\cdot\left(g\mathbf{W}_3+g'\mathbf{B}\right)\psi(\mathbf{r},t)e.\end{aligned} \tag{D.69}$$

If we write the electromagnetic field $\mathbf{A}$ and the Z^0 field $\mathbf{Z}$ as arbitrary linear combinations,

$$\mathbf{A}(\mathbf{r},t)=\mathbf{B}(\mathbf{r},t)\cos\theta_W+\mathbf{W}_3(\mathbf{r},t)\sin\theta_W \tag{D.70a}$$

and

$$\mathbf{Z}(\mathbf{r},t)=-\mathbf{B}(\mathbf{r},t)\sin\theta_W+\mathbf{W}_3(\mathbf{r},t)\cos\theta_W,$$

with analogous equations for the scalar fields $(\phi,\,\phi^Z)$,

$$\phi(\mathbf{r},t)=\phi^B(\mathbf{r},t)\cos\theta_W+\phi_3^W(\mathbf{r},t)\sin\theta_W \tag{D.70b}$$

and

$$\phi^Z(\mathbf{r},t)=-\phi^B(\mathbf{r},t)\sin\theta_W+\phi_3^W(\mathbf{r},t)\cos\theta_W,$$

then Equation (D.69) becomes

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)e$$
$$=2g_W\left(\phi^- -\boldsymbol{\alpha}\cdot\mathbf{W}^-\right)\psi(\mathbf{r},t)\nu-\frac{1}{2}\left(g\sin\theta_W+g'\cos\theta_W\right)(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})\,\psi(\mathbf{r},t)e$$
$$-\frac{1}{2}\left(g\cos_W-g'\sin\theta_W\right)\left(\phi^Z-\boldsymbol{\alpha}\cdot\mathbf{Z}\right)\psi(\mathbf{r},t)e. \tag{D.71a}$$

The same argument for the neutrino case $\Psi=\psi(\mathbf{r},t)\nu$, using (D.61a), (D.67) and (D.70), gives

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\nu$$
$$=2g_W\left(\phi^+-\boldsymbol{\alpha}\cdot\mathbf{W}^+\right)\psi(\mathbf{r},t)e+\frac{1}{2}\left(g\sin\theta_W-g'\cos\theta_W\right)(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})\,\psi(\mathbf{r},t)\nu$$
$$+\frac{1}{2}\left(g\cos_W+g'\sin\theta_W\right)\left(\phi^Z-\boldsymbol{\alpha}\cdot\mathbf{Z}\right)\psi(\mathbf{r},t)\nu. \tag{D.71b}$$

We can now identify the conditions whereby the electromagnetic interactions in (D.71a) and (D.71b), proportional to the electromagnetic field combinations $(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})$, coincide with those of quantum electrodynamics. They are

$$g\sin\theta_W=g'\cos\theta_W,$$

which guarantees that the electromagnetic field does not couple directly to neutrinos, and

$$\frac{1}{2}(g\sin\theta_W+g'\cos\theta_W)=e,$$

which ensures that the coupling of the electromagnetic field has the strength required in QED, Equation (D.20). These equations are equivalent to the unification condition

$$e=g\sin\theta_W=g'\cos\theta_W \tag{D.72}$$

or alternatively

$$\frac{e}{2\sqrt{2}}=g_W\sin\theta_W=g_Z\cos\theta_W,$$

where

$$g_W\equiv\frac{g}{2\sqrt{2}},\quad \mathrm{g_Z}\equiv\frac{g'}{2\sqrt{2}}. \tag{D.73}$$

This is identical to the unification condition (9.8), except that we are using Heavyside–Lorentz rather than SI electromagnetic units. When it is imposed, the electromagnetic interactions coincide with those used in QED with such spectacular success, while the neutral current interactions, proportional to $(\phi^Z - \boldsymbol{\alpha} \cdot \mathbf{Z})$, are uniquely predicted in terms of the electric charge e and the weak mixing angle θ_W. However, before discussing this further we must take the spin structure of the interaction into account.

D.7.4 Spin structure and parity violation

For zero-mass fermions, the Dirac equation (D.16) can be shown to separate into two independent equations[19]

$$i\frac{\partial \psi^L}{\partial t} = -i\boldsymbol{\alpha} \cdot \boldsymbol{\nabla}\psi^L, \quad i\frac{\partial \psi^R}{\partial t} = -i\boldsymbol{\alpha} \cdot \boldsymbol{\nabla}\psi^R, \tag{D.74}$$

where $\psi^{L,R}$ are Dirac spinors corresponding to the left- and right-handed spin states of Figure 10.3, respectively. Hence in discussing the spin structure of weak interactions we can treat the left- and right-handed spin states completely separately.

In Chapter 10 we saw that when fermion masses can be neglected, the charged weak current interactions involve left-handed spin states only, in violation of parity conservation. We therefore assume that the preceding sections D.1, D.2 and D.3 apply to left-handed spin states only; i.e. (D.71) is assumed to hold for left-handed spin states only. It now remains to incorporate the right-handed spin states. To do this, we note that (D.71) has been obtained by combining the left-handed spin states of electrons and neutrinos into weak isodoublets

$$\Psi^L = \psi^L(\mathbf{x}, t)\chi^f,$$

with $I_3^W = \frac{1}{2}, -\frac{1}{2}$ for neutrinos and electrons, respectively, and $Y^W = -\frac{1}{2}$ in both cases by (D.64). Right-handed states are instead assigned to weak isosinglets with $I_i^W = 0 (i = 1, 2, 3)$. In this case, both electrons and neutrinos are described by separate Dirac spinors $\psi_e^R(\mathbf{r}, t)$, $\psi_\nu^R(\mathbf{r}, t)$ and the gauge transformations (D.58) reduce to

$$\psi_e^R \to \psi_e'^R = \psi_e^R, \quad \psi_\nu^R \to \psi_\nu'^R = \psi_\nu^R.$$

In other words, the wavefunctions remain unchanged, and no interactions with W bosons of the type discussed in Section D.7.2 are generated by the gauge principle. However, by (D.64), $Y^W = (0, -1)$ for the neutrino and the electron, respectively, so that the gauge transformations (D.65) become

$$\psi_\nu^R \to \psi_\nu'^R = \psi_\nu^R, \quad \psi_{\mathrm{e}}^R \to \psi_{\mathrm{e}}'^R = \mathrm{e}^{ig'\omega}\psi_{\mathrm{e}}^R.$$

[19] See, for example, Section 5.6 of Kane (1987).

The argument of Section D.7.3 then leads to the equations

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi_{\nu}^{R}(\mathbf{r},t)=0 \tag{D.75a}$$

and

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi_{e}^{R}(\mathbf{r},t)=-g'\left(\phi^{B}-\boldsymbol{\alpha}\cdot\mathbf{B}\right)\psi_{e}^{R}(\mathbf{r},t)$$

in analogy to (D.67). The second of these equations becomes

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi_{e}^{R}=-e\left(\phi-\boldsymbol{\alpha}\cdot\mathbf{A}\right)\psi_{e}^{R}+e\tan\theta_{W}\left(\phi^{Z}-\boldsymbol{\alpha}\cdot\mathbf{Z}\right)\psi_{e}^{R}, \tag{D.75b}$$

using the mixing equations (D.70) and the unification condition (D.72). The first term on the right-hand side of (D.75b) corresponds to the usual electromagnetic interaction, while the second term is a weak neutral current interaction. In other words, for electrons (but not neutrinos) both right-handed and left-handed spin states are predicted to have neutral current interactions.

Equations (D.71) and (D.75) represent the final predictions for the electroweak interactions of electrons and neutrinos. Essentially, the gauge transformations have been chosen throughout to reproduce the known properties of the electromagnetic and weak charged current interactions. The form of the neutral current interactions, including their spin structure, is predicted uniquely in terms of the mixing angle θ_W. The agreement between these predictions and the measured properties of neutral current interactions is one of the great triumphs of the unified theory.

Finally, we note that this analysis is based on the fact that the Dirac equation can be separated into two equations (D.74) in the limit of zero fermion mass. This enabled us to ascribe different gauge transformation properties to the left- and right-handed states, which are treated completely separately. If finite masses are assumed, then the left- and right-handed states do not decouple, and the analysis breaks down. This implies that, in the absence of a Higgs field, different interactions for the left- and right-handed states (i.e. parity violation) can only be incorporated into a gauge theory having massless fermions. This result can, however, be avoided if the fermion masses, like the $W^{\pm}$ and Z^0 masses, arise from interactions with the Higgs field.

PROBLEMS D

D.1 Show that Hamilton's equations of motion

$$\dot{x}_i=\frac{\partial H}{\partial p_i},\quad \dot{p}_i=-\frac{\partial H}{\partial x_i}\quad (i=1,2,3)$$

lead to the equations of motion (D.5) for a charged particle in an electromagnetic field if the Hamiltonian is of the form (D.3), where $\mathbf{r}=(x_1,x_2,x_3)\equiv(x,y,z)$.

D.2 Show that Maxwell's equations in free space (D.22) reduce to (D.10) when expressed in terms of the potentials (D.4).

D.3 Show that for any electromagnetic potentials $(\phi, \mathbf{A})$ there are many possible choices of gauge-transformed potentials

$$\phi = \tilde{\phi} + \frac{\partial f}{\partial t}, \quad \mathbf{A} = \tilde{\mathbf{A}} - \boldsymbol{\nabla} f,$$

such that the Lorentz condition (D.11) is satisfied.

D.4 Verify that the Dirac equation in the presence of an electromagnetic field (D.18) is gauge-invariant, as stated in the text.

D.5 Show that the gauge transformation (D.8) used in our discussion of QED can be written as two successive electroweak gauge transformations of the form (D.58) and (D.66).

D.6 Our treatment of the electroweak interactions of massless leptons may be extended to massless quarks by invoking lepton–quark symmetry between the doublets (ν_e, e^-) and (u, d) where we neglect quark mixing.

(a) What are the appropriate weak isospin and hypercharge quantum numbers (I_3^W, Y^W) for u and d quarks in both left-handed and right-handed states?

(b) Show that with these assignments, the unification condition (D.72) leads to the appropriate electromagnetic interactions for (u, d) quarks. (*Hint.* This does not require complicated algebra, but just appropriate modifications of the results of sections D.7.3 and D.7.4 to take account of the different (I_3^W, Y^W) values.)

D.7 How would you modify your answer to Problem D.6 to take account of mixing between the quark doublets (u, d) and (c, s)? In particular, obtain the equations analogous to (D.71a) and (D.71b) for left-handed d and s quarks, instead of electrons. Hence confirm that the charged current couplings are consistent with (8.17) and that there are no terms corresponding to strangeness-changing neutral currents, in agreement with the discussion of Section 9.1.1.

E

Tables of Particle Properties

The tables below summarize the main properties of the gauge bosons, leptons, quarks and the well-established low-lying hadrons. Errors shown in brackets refer to the last significant figures of the values given. For unstable particles, in the decay column, X stands for any state allowed by the appropriate conservation laws. In general, only those decay modes that have a branching ratio of greater than about 3 % are included. In the case of charge conjugate particles, decay modes are only shown for one partner, since those of the other are just the corresponding charge-conjugated reactions.

A comprehensive compilation of particle physics data may be found in the biannual publication of the Particle Data Group (PDG), which is the source of the data below. The 2006 edition of their definitive 'Review of particle properties' is W.-M. Yao *et al.*, *Journal of Physics*, **G33**, 1 (2006). The PDG Review is available online at http://pdg.lbl.gov and this site also contains links to other sites where compilations of particle data may be found.

E.1 GAUGE BOSONS

The gauge bosons all have $J^P = 1^-$.

			Decay	
Particle	Mass	Full width	Mode	Fraction (%)
g	0 (assumed)	Stable		
γ	$<6 \times 10^{-17}$ eV	Stable		
$W^{\pm}$	80.403(±29) GeV/c^2	2.141(±41) GeV	Hadrons	67.60(±27)
			$\tau^+\nu_\tau$	11.25(±20)
			$e^+\nu_e$	10.75(±13)
			$\mu^+\nu_\mu$	10.57(±15)

Particle Physics Third Edition B. R. Martin and G. Shaw

(Continued).

Particle	Mass	Full width	Decay Mode	Decay Fraction (%)
Z^0	$91.1876(\pm 21)\,\mathrm{GeV/c^2}$	$2.4952(\pm 23)\,\mathrm{GeV}$	Hadrons	69.91(±6)
			$\nu_\ell\bar{\nu}_\ell$ (all ℓ)	20.00(±6)
			$\tau^+\tau^-$	3.370(±8)
			$\mu^+\mu^-$	3.366(±7)
			e^+e^-	3.363(±4)

E.2 LEPTONS

All leptons are believed to have $J^P = \frac{1}{2}^+$. The neutrinos are shown as stable; neutrino oscillations are discussed in Section 2.3.

Particle	Mass ($\mathrm{MeV/c^2}$)	Mean lifetime (s)	Decay Mode	Decay Fraction (%)
ν_e	$<2\,\mathrm{eV/c^2}$	Stable		
ν_μ	<0.19	Stable		
ν_τ	<18.2	Stable		
$e^\pm$	0.511[a]	Stable		
$\mu^\pm$	105.66[b]	2.197×10^{-6}[c]	$e^+\nu_e\bar{\nu}_\nu$	100
$\tau^\pm$	1776.99(±27)	$(2.906\pm10)\times10^{-13}$	Hadrons $+\nu_\tau$	∼64
			$e^+\nu_e\bar{\nu}_\tau$	17.84(±5)
			$\mu^+\nu_\mu\bar{\nu}_\tau$	17.36(±5)

[a] The error on the $e^\pm$ mass is $4\times10^{-8}\,\mathrm{MeV/c^2}$; [b] The error on the $\mu^\pm$mass is $9\times10^{-6}\,\mathrm{MeV/c^2}$.
[c] The error on the $\mu^\pm$ lifetime is 4×10^{-11} s.

E.3 QUARKS

The quarks all have $J^P = \frac{1}{2}^+$. The table shows their approximate masses (see footnote to table) and the values of the quantum numbers: baryon number B, isospin I, charge Q, strangeness S, charm C, bottom $\tilde{B}$ and top T. For antiquarks, the signs of all these quantum numbers are reversed.

Name	Symbol	Mass ($\mathrm{GeV/c^2}$)[a]	Q	S	C	$\tilde{B}$	T	B	I
Down	d	∼0.35	−1/3	0	0	0	0	1/3	1/2
Up	u	$m_u \approx m_d$	2/3	0	0	0	0	1/3	1/2
Strange	s	∼0.5	−1/3	−1	0	0	0	1/3	0
Charmed	c	∼1.5	2/3	0	1	0	0	1/3	0
Bottom	b	∼4.5	−1/3	0	0	−1	0	1/3	0
Top	t	174.2 ± 3.3	2/3	0	0	0	1	1/3	0

[a] The definition of quark masses is complicated by the fact that quarks are not observed as free particles. So-called 'current' quark masses are those that appear in the QCD Hamiltonian. The masses shown in the table (with the exception of the top quark) are the effective values obtained from the quark model of the hadron spectrum; they include contributions from interactions between the quarks and associated gluons. These are the so-called 'constituent' quark masses. The mass of the top quark is inferred directly from its decay products.

E.4 LOW-LYING BARYONS

Particle	I, J^P	Mass ($\mathrm{MeV/c^2}$)	Mean lifetime or width	Decay Mode	Decay Fraction (%)
Unflavoured states of light quarks ($S = C = \tilde{B} = 0$) Quark content: $N = (p,n) : p = uud$, $n = udd$; $\Delta^{++} = uuu$, $\Delta^{+} = uud$, $\Delta^{0} = udd$, $\Delta^{-} = ddd$					
p	$\frac{1}{2}, \frac{1}{2}^+$	938.27203(±8)	$>2.1 \times 10^{29}$ yr[a]		
n	$\frac{1}{2}, \frac{1}{2}^+$	939.56536(±8)	$8.857(\pm8) \times 10^{2}$ s	$pe^-\bar{\nu}_e$	100
Δ	$\frac{3}{2}, \frac{3}{2}^+$	1232(±1)	118(±2) MeV	$N\pi$	100
Strange baryons ($S = -1$, $C = \tilde{B} = 0$) Quark content: $\Lambda = uds$; $\Sigma^+ = uus$, $\Sigma^0 = uds$, $\Sigma^- = dds$, similarly for $\Sigma^* s$.					
Λ	$0, \frac{1}{2}^+$	1115.683(±6)	$2.631(\pm20) \times 10^{-10}$	$p\pi^-$	63.9(±5)
				$n\pi^0$	35.8(±5)
Σ^+	$1, \frac{1}{2}^+$	1189.37(±7)	$8.018(\pm26) \times 10^{-11}$	$p\pi^0$	51.57(±30)
				$n\pi^+$	48.31(±30)
Σ^0	$1, \frac{1}{2}^+$	1192.642(±24)	$7.4(\pm7) \times 10^{-20}$	$\Lambda\gamma$	100
Σ^-	$1, \frac{1}{2}^+$	1197.449(±30)	$1.479(\pm11) \times 10^{-10}$	$n\pi^-$	99.848(±5)
Σ^{*+}	$1, \frac{3}{2}^+$	1382.8(±4)	35.8(±8) MeV	$\Lambda\pi$	87.0(±15)
				$\Sigma\pi$	11.7(±15)
Σ^{*0}	$1, \frac{3}{2}^+$	1383.7(±10)	36(±5) MeV	As above	
Σ^{*-}	$1, \frac{3}{2}^+$	1387.2(±5)	39.4(±21) MeV	As above	
Strange baryons ($S = -2$, $C = \tilde{B} = 0$) Quark content: $\Xi^0 = uss$, $\Xi^- = dss$, similarly for $\Xi^* s$					
Ξ^0	$\frac{1}{2}, \frac{1}{2}^+$	1314.83(±20)	$2.90(\pm9) \times 10^{-10}$ s	$\Lambda\pi^0$	99.523(±13)
Ξ^-	$\frac{1}{2}, \frac{1}{2}^+$	1321.31(±13)	$1.639(\pm15) \times 10^{-10}$ s	$\Lambda\pi^-$	99.887(±35)
Ξ^{*0}	$\frac{1}{2}, \frac{3}{2}^+$	1531.80(±32)	9.1(±5) MeV	$\Lambda\bar{K}$, $\Sigma\bar{K}$, $\Xi\pi$	Seen
Ξ^{*-}	$\frac{1}{2}, \frac{3}{2}^+$	1535.0(±6)	9.9(±18) MeV	As above	

(Continued).

Particle	I, J^P	Mass (MeV/c^2)	Mean lifetime or width	Decay Mode	Decay Fraction (%)
Strange baryons ($S=-3$, $C=\tilde{B}=0$) Quark content: $\Omega^- = sss$					
Ω^-	$0, \frac{3}{2}^+$	1672.45(±29)	$8.21(\pm11)\times10^{-11}$ s	ΛK^-	67.8(±7)
				$\Xi^0\pi^-$	23.6(±7)
				$\Xi^-\pi^0$	8.6(±4)
Charmed baryons ($S=0$, $C=+1$, $\tilde{B}=0$) Quark content: $\Lambda_c^+ = udc$; $\Sigma_c^{++} = uuc$, $\Sigma_c^+ = udc$, $\Sigma_c^0 = ddc$, similarly for $\Sigma_c^* s$					
Λ_c^+	$0, \frac{1}{2}^+$	2286.46(±14)	$2.00(\pm6)\times10^{-13}$ s	$n+X$	50(±16)
				$p+X$	50(±16)
				$\Lambda+X$	35(±11)
				$\Sigma^\pm+X$	10(±5)
				e^++X	4.5(±17)
Σ_c^{++}	$1, \frac{1}{2}^+$	2454.02(±18)	2.23(±30) MeV	$\Lambda_c^+\pi$	Seen
Σ_c^+	$1, \frac{1}{2}^+$	2452.9(±4)	< 4.6 MeV		
Σ_c^0	$1, \frac{1}{2}^+$	2453.76(±18)	2.2(±4) MeV		
Charmed strange baryons ($S=-1,-2$, $C=1$, $\tilde{B}=0$) Quark content: $\Xi_c^+ = usc$, $\Xi_c^0 = dsc$, $\Omega_c^0 = ssc$					
Ξ_c^+	$\frac{1}{2}, \frac{1}{2}^+$	2467.9(±4)	$4.42(\pm26)\times10^{-13}$ s	Several seen	
Ξ_c^0	$\frac{1}{2}, \frac{1}{2}^+$	2471.0(±4)	$1.12(\pm4)\times10^{-13}$ s	Several seen	
Ω_c^0	$\frac{1}{2}, \frac{1}{2}^+$	2697.5(±26)	$6.9(\pm1.2)\times10^{-14}$ s	Several seen	
Bottom baryons ($S=C=0$, $\tilde{B}=-1$) Quark content: $\Lambda_b^0 = udb$					
Λ_b^0	$0, \frac{1}{2}^+$	5624(±9)	$1.230(\pm74)\times10^{-12}$ s	$\Lambda_c^+ + X$	9.1(±2.3)
Ξ_b^0	$0, \frac{1}{2}^+$	5792(±3)	$1.42(\pm28)\times10^{-12}$ s		
Ξ_b^-	$0, \frac{1}{2}^+$	5792(±3)	$1.42(\pm28)\times10^{-12}$ s		

[a] This is the limit obtained from experiment without making any assumption about the nature of the final state. For specific assumed final states, the limit is $>10^{31}-10^{33}$ yr.

E.5 LOW-LYING MESONS

In the J^{PC} column, the C quantum number applies to just the neutral states of an isospin multiplet.

Particle	I, J^{PC}	Mass (MeV/c^2)	Mean lifetime or width	Decay Mode	Decay Fraction (%)
Unflavoured states of light quarks ($S = C = \tilde{B} = 0$)					
Quark content:					
$I = 1$ states, $u\bar{d}$, $\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$, $d\bar{u}$; $I = 0$ states, $c_1(u\bar{u} - d\bar{d}) + c_2 s\bar{s}$ ($c_{1,2}$ are constants)					
$\pi^{\pm}$	$1, 0^-$	139.57018(±35)	$2.6033(\pm5) \times 10^{-8}$ s	$\pi^+ \nu_\mu$	99.98770(±4)
π^0	$1, 0^{-+}$	134.9766(±6)	$8.4(\pm6) \times 10^{-17}$ s	$\gamma\gamma$	98.798(±32)
η	$0, 0^{-+}$	547.51(±18)	1.30(±7) keV	$\gamma\gamma$	39.38(±26)
				$\pi^0\pi^0\pi^0$	32.51(±28)
				$\pi^+\pi^-\pi^0$	22.7(±4)
				$\pi^+\pi^-\gamma$	4.69(±11)
ρ	$1, 1^{--}$	775.5(±4)	149.4(±10) MeV	$\pi\pi$	∼100
ω^0	$0, 1^{--}$	782.65(±12)	8.49(±9) MeV	$\pi^+\pi^-\pi^0$	89.1(±7)
				$\pi^0\gamma$	8.90(±25)
η'	$0, 0^{-+}$	957.78(±14)	0.203(±16) MeV	$\pi^+\pi^-\eta$	44.5(±14)
				$\pi^+\pi^-\gamma$	29.4(±9)
				$\pi^0\pi^0\eta$	20.8(±12)
				$\omega\gamma$	3.03(±31)
ϕ	$0, 1^{--}$	1019.460(±19)	4.26(±5) MeV	K^+K^-	49.2(±6)
				$K_L^0 K_S^0$	34.0(±5)
				$\rho\pi + \pi^+\pi^-\pi^0$	15.3(±4)
Strange mesons ($S = \pm1$, $C = \tilde{B} = 0$)					
Quark content: $K^+ = u\bar{s}$, $K^0 = d\bar{s}$, $\bar{K}^0 = s\bar{d}$, $K^- = s\bar{u}$, similarly for K^*					
$K^{\pm}$	$\frac{1}{2}, 0^-$	493.667(±16)	$1.2385(\pm24) \times 10^{-8}$ s	$\mu^+ \nu_\mu$	63.44(±14)
				$\pi^+\pi^0$	20.92(±12)
				$\pi^+\pi^+\pi^-$	5.590(±31)
				$\pi^0 e^+ \nu_e$	4.98(±7)
				$\pi^0 \mu^+ \nu_\mu$	3.32(±6)
K^0, $\bar{K}^0$	$\frac{1}{2}, 0^-$	497.648(±22)			
K_S^0	See note *a*		$8.953(\pm5) \times 10^{-11}$ s	$\pi^+\pi^-$	69.20(±5)
				$\pi^0\pi^0$	30.69(±5)
K_L^0	See note *a*		$5.114(\pm21) \times 10^{-8}$ s	$\pi^{\pm} e^{\mp} \nu_e(\bar{\nu}_e)$	40.53(±15)
				$\pi^{\pm} \mu^{\mp} \nu_\mu(\bar{\nu}_\mu)$	27.02(±7)
				$\pi^0\pi^0\pi^0$	19.56(±14)
				$\pi^+\pi^-\pi^0$	12.56(±5)
$K^{*\pm}$	$\frac{1}{2}, 1^-$	891.66(±26)	50.8(±9) MeV	$K\pi$	∼100
K^{*0}	$\frac{1}{2}, 1^-$	896.00(±25)	50.3(±6) MeV	$K\pi$	∼100

(Continued).

Particle	I, J^{PC}	Mass (MeV/c^2)	Mean lifetime or width	Decay Mode	Decay Fraction (%)
Charmed mesons ($S=0$, $C=\pm1$, $\tilde{B}=0$) Quark content: $D^+=c\bar{d}$, $D^0=c\bar{u}$, $\bar{D}^0=u\bar{c}$, $D^-=d\bar{c}$, similarly for D^*s					
$D^\pm$	$\frac{1}{2}, 0^-$	1869.3(±4)	$1.040(\pm7)\times10^{-12}$ s	K^0+X	
				plus	
				$\bar{K}^0+X$	61 (±8))
				K^-+X	27.5(±24)
				$\bar{K}^{*0}+X$	23(±5)
				e^++X	17.2(±19)
				K^++X	5.5(±16)
D^0, $\bar{D}^0$	$\frac{1}{2}, 0^-$	1864.5(±4)	$4.101(\pm2)\times10^{-13}$ s	K^-+X	53(±4)
				K^0+X	
				plus	
				$\bar{K}^0+X$	42 (±5))
				$\bar{K}^{*0}+X$	9(±4)
				e^++X	6.71(±29)
				K^++X	3.4(±5)
$D^{*\pm}$	$\frac{1}{2}, 0^-$	2010.0(±4)	96(±21) keV	$D^0\pi^+$	67.7(±5)
				$D^+\pi^0$	30.7(±5)
D^{*0}, $\bar{D}^{*0}$	$\frac{1}{2}, 1^-$	2006.7(±4)	<2.1 MeV	$D^0\pi^0$	61.9(±29)
				$D^0\gamma$	38.1(±29)
Charmed strange mesons ($S=C=\pm1$, $\tilde{B}=0$) Quark content: $D_s^+=c\bar{s}$, $D_s^-=s\bar{c}$, similarly for D_s^*s					
$D_s^\pm$	$0, 0^-$	1968.2(±5)	$5.00(\pm7)\times10^{-13}$ s	K^0+X	
				plus	
				$\bar{K}^0+X$	39(±28)
				K^++X	20(±16)
				$\phi+X$	18(±13)
				K^-+X	13(±13)
				e^++X	8(±7)
				$\tau^+\nu_\tau$	6.4(±15)
$D_s^{*\pm}$	$0, 1^-$	2112.0(±6)	< 1.9 MeV	$D_s^+\gamma$	94.2(±7)
				$D_s^+\pi^0$	5.8(±7)
Bottom mesons ($S=C=0$, $\tilde{B}=\pm1$) Quark content: $B^+=u\bar{b}$, $B^0=d\bar{b}$, $\bar{B}^0=b\bar{d}$, $B^-=b\bar{u}$, similarly for B^*s					
$B^\pm$	$\frac{1}{2}, 0^-$	5279.0(±5)	$1.638(\pm11)\times10^{-12}$ s	$\bar{c}X$ (see note *b*)	98(±6)
				cX (see note *b*)	33(±5)
				$\ell^+\nu_\ell+X$	10.9(±4)

$B^0, \bar{B}^0$	$\frac{1}{2}, 0^-$	5279.4(±5)	$1.530(\pm 9) \times 10^{-12}$ s	$\bar{c}X$ (see note *b*)	104(±8)
				cX (see note *b*)	24(±5)
				$\ell^+ \nu_\ell + X$	10.4(±4)
Bottom strange mesons ($S = \mp 1$, $C = 0$, $\tilde{B} = \pm 1$) Quark content: $B_s^0 = s\bar{b}$, $\bar{B}_s^0 = b\bar{s}$, similarly for $B^* s$					
$B_s^0, \bar{B}_s^0$	$0, 0^-$	5367.5(±18)	$1.466(\pm 59) \times 10^{-12}$ s	$D_s^- + X$	94(±30)
				$D_s^- \ell^+ \nu_\ell + X$	7.9(±24)
Bottom charmed mesons ($S = 0$, $\tilde{B} = C = \pm 1$) Quark content: $B_c^+ = c\bar{b}$, $B_c^- = \bar{c}b$					
$B_c^\pm$	$0, 0^-$	6286(±5)	$4.6(\pm 17) \times 10^{-13}$ s		Several seen
$c\bar{c}$ mesons					
$\eta_c(1S)$	$0, 0^{-+}$	2980.4(±12)	25.5(±34) MeV	$K\bar{K}\pi$	7.2(±12)
				$\eta\pi\pi$	4.9(±18)
				$\eta'\pi\pi$	4.1(±17)
$J/\psi(1S)$	$0, 1^{--}$	3096.916(11)	93.4(±21) keV	Hadrons	87.7(±5)
				e^+e^-	5.94(±6)
				$\mu^+\mu^-$	5.93(±6)
$b\bar{b}$ mesons					
$\Upsilon(1S)$	$0, 1^{--}$	9460.30(±26)	54.02(±125) keV	$\eta' + X$	2.8(±4)
				$\ell^+\ell^-$ all ℓ	7.95(±43)

[a] These states are discussed in Sections 10.2.1 and 10.2.2.
[b] The $\bar{c}$ stands for any state containing a $\bar{c}$ quark and the c stands for any state containing a c quark.

F

Solutions to Problems

PROBLEMS 1

1.1 (a) $\overline{\nu}_e + e^+ \rightarrow \overline{\nu}_e + e^+$
(b) $\overline{p} + n \rightarrow \pi^- + \pi^+ + \pi^-$ or $\pi^- + \pi^0 + \pi^0$

1.2 The topologically distinct diagrams for reactions (a) are shown in Figure F.1, those for (b) in Figure 5.3 of Section 5.5.1 and those for (c) in Figure F.2.

1.3 An example of a fourth-order diagram for the reaction is given in Figure F.3.

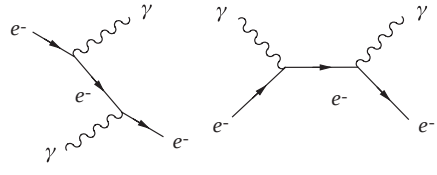


Figure F.1

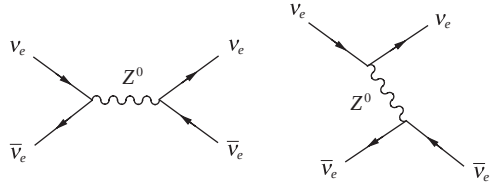


Figure F.2

Particle Physics Third Edition B. R. Martin and G. Shaw

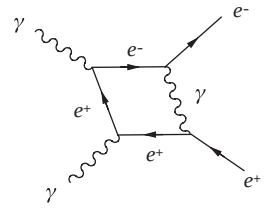


Figure F.3

1.4 For a spherically symmetric static solution we can set $\Psi(\mathbf{r},t)=\phi(r)=\phi$, where $r=|\mathbf{r}|$, and use

$$\nabla^2=\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r},$$

giving

$$\nabla^2\phi=\frac{\partial^2\phi}{\partial r^2}+\frac{2}{r}\frac{\partial\phi}{\partial r}=\left(\frac{mc}{\hbar}\right)^2\phi.$$

Substituting $\phi=u(r)/r$ gives

$$\frac{\mathrm{d}^2u(r)}{\mathrm{d}r^2}=\left(\frac{mc}{\hbar}\right)^2u(r)$$

and the result follows by solving for u and imposing $\phi\to 0$ as $r\to\infty$.

1.5 If we impose momentum conservation and neglect the momenta of the initial $e^{\pm}$, we have

$$e^+(m,\mathbf{0})+e^-(m,\mathbf{0})\to\gamma(E_k,\mathbf{k})+\gamma(E_k,-\mathbf{k}),$$

where $E_k=|\mathbf{k}|=m$. Hence in Figure 1.9(a), the initial virtual process is

$$e^-(m,\mathbf{0})\to e^-(E,\mathbf{k})+\gamma(m,-\mathbf{k}),$$

where the energy of the virtual electron is

$$E=(k^2+m^2)^{1/2}=\sqrt{2}m\approx m.$$

Hence $\Delta E=E\approx m$ and, from the uncertainty principle, $r\approx 1/\Delta E\approx 1/m$. Restoring factors of $\hbar$ and c gives $r\approx\hbar c/mc^2=368\,\mathrm{fm}$.

1.6 The distance between the two vertices is given by

$$r\approx c\tau\approx\hbar c/\Delta E=1/\Delta E,$$

where ΔE is the energy violation at the first vertex. On evaluating ΔE in each case, one obtains (a) $r\approx m^{-1}$ and (b) $r'\approx E^{-1}$, respectively, where E is the initial electron energy in the centre-of-mass frame. These are related by a Lorentz contraction $r'=r/\gamma$, where $\gamma=E/m$. (A resumé of special relativity is given in Appendix A.)

1.7 Set $\tau=(2/m\alpha^5)\hbar^a c^b$ and demand that τ has the dimensions of time. This gives $a=1$ and $b=-2$ and hence $\tau=1.245\times10^{-10}\,\mathrm{s}$.

PROBLEMS 2

2.1 Weak interactions conserve baryon number B, charge Q and lepton numbers L_e, L_μ and L_τ. They need not conserve the quantum numbers S, C or $\tilde{B}$. Of the decays given, (a) violates L_μ conservation and (d) violates both L_μ and L_τ conservation. They are therefore both forbidden. Reactions (b) and (c) satisfy all the conservation laws and are allowed.

2.2 The lowest-order diagram in Figure F.4 involves the exchange of a single Z^0 boson.

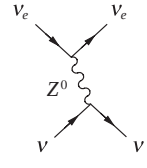


Figure F.4

2.3 The two diagrams are shown in Figure F.5.

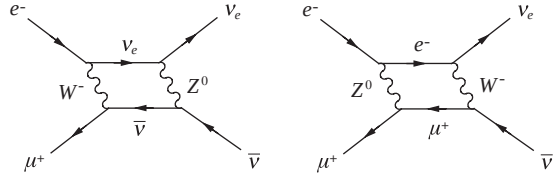


Figure F.5

2.4 The electron neutrino may interact with electrons via both Z^0 and W^- exchange, as shown in Figure F.6, whereas, because of lepton number conservation, the muon neutrino can

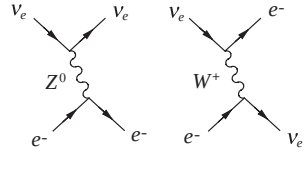


Figure F.6

only interact via Z^0 exchange, as shown in Figure F.7.

2.5 Firstly, restore factors of $\hbar$ and c by writing $L_0 = 4E\hbar^a c^b/\Delta m_{ij}^2$ and find a and b by demanding that the right-hand side has the dimensions of a length. This gives $a = 1$ and $b = -3$, so that $L_0 = 4E(hc)/\Delta m_{ij}^2 c^4$. Then if L_0 is expressed in km, E in GeV and Δm_{ij}^2 in $(\mathrm{eV/c^2})^2$, we have

$$L_0 = \frac{4E \times (1.97 \times 10^{-13}) \times 10^{18}}{\Delta m_{ij}^2} = \frac{E}{1.27\Delta m_{ij}^2}\ \mathrm{km}.$$

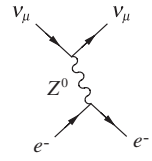


Figure F.7

2.6 From Equation (2.30) and Problem 2.5, we have, for maximal mixing ($\theta = \pi/4$), $P(\overline{\nu}_e \to \overline{\nu}_x) = \sin^2[1.27\Delta(m^2c^4)L/E]$, where L is measured in metres, E in MeV, $\Delta(m^2c^4)$ in $(\text{eV})^2$ and $\Delta m^2 \equiv m^2(\overline{\nu}_e) - m^2(\overline{\nu}_x)$. If $P(\overline{\nu}_e \to \overline{\nu}_e) = 0.9 \pm 0.1$, then at 95 % confidence level, $0.3 \geq P(\overline{\nu}_e \to \overline{\nu}_x) \geq 0$ and hence $0 \leq \Delta(m^2c^4) \leq 3.8 \times 10^{-3}\ (\text{eV})^2$.

2.7 From the data given, the total number of nucleons is given by

$$N = \frac{2 \times 10^{30}}{(0.98 \times 10^9)(1.78 \times 10^{-36})} = 1.1 \times 10^{57}$$

and hence $n = 7.7 \times 10^{38}\,\text{km}^{-3}$. Also, the mean energy of the neutrinos from reaction (2.35) is 0.26 MeV, so the cross-section is $\sigma = 1.8 \times 10^{-46}\,\text{m}^2$. Thus, finally, $\lambda \approx 7 \times 10^6$ km, i.e. about 10 times the solar radius.

PROBLEMS 3

3.1 (a) Involves neutrinos and is therefore a weak interaction. (b) Involves photons and is therefore electromagnetic. (c) Conserves all quantum numbers and is therefore a strong interaction. (d) Violates both strangeness and charm and is therefore a weak interaction. (e) Conserves all quantum numbers and is a strong interaction. (f) Involves electrons and positrons and is therefore an electromagnetic interaction.

3.2 The quantum numbers are

$$X^0 : B = 1, \quad S = -1, \quad C = 0, \quad \tilde{B} = 0; \quad Y^- : B = 1, \quad S = -2, \quad C = 0, \quad \tilde{B} = 0.$$

From their charges and the definitions of $\tilde{B}, S, C$ and B, it follows that $X^0 = uds$ and $Y^- = dss$.

The decay $X^0 \to \Lambda + \gamma$ is electromagnetic, since it involves a photon and conserves S, so we expect that $\tau = 10^{-16} - 10^{-21}$ s. The decay $Y^- \to \Lambda + \pi^-$ violates strangeness conservation and is a weak interaction, so we expect that $\tau = 10^{-6} - 10^{-13}$ s. (The X^0 and the Y^- are in fact the $\Sigma^0(1193)$ and the $\Xi^-(1321)$ states listed in Appendix E, Table E.4.)

3.3 The quantum number combination (2, 1, 0, 1, 0) corresponds to a baryon qqq. It has $S = 0, C = 1$ and $B = 0$, so must be of the form cxy, where x and y are u or d quarks. The charge $Q = 2$ requires both x and y to be u quarks, i.e. cuu. The others are established by similar arguments and so the full set is

$$(2,1,0,1,0) = cuu, \qquad (0,1,-2,1,0) = css, \qquad (0,0,1,0,-1) = b\bar{s},$$

$$(0,-1,1,0,0) = \bar{s}\bar{d}\bar{u}, \quad (0,1,-1,1,0) = csd, \quad (-1,1,-3,0,0) = sss.$$

They are called Σ_c^{++}, Ω_c^0, $\overline{B}_s^0$, $\overline{\Lambda}$, Σ_c^0 and Ω^-, respectively.

3.4 (a) The quark compositions are $D^- = d\bar{c}$; $K^0 = d\bar{s}$; $\pi^- = d\bar{u}$ and since the dominant decay of a c quark is $c \to s$, we can have either of the diagrams shown in Figure F.8.

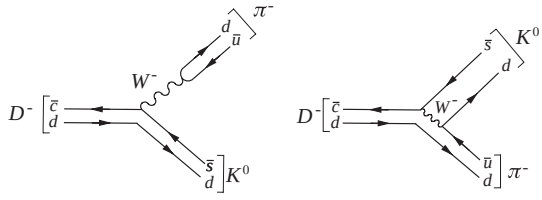


Figure F.8

(b) The quark compositions are $\Lambda = sud$; $p = uud$ and since the dominant decay of an s quark is $s \to u$, we have the result shown in Figure F.9.

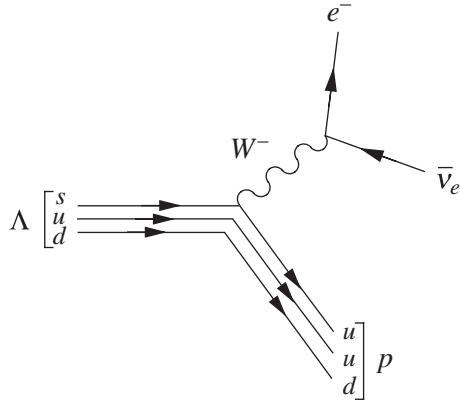


Figure F.9

3.5 For $\gamma \approx 10, \tau \approx c$ the average distance $d \approx c\gamma\tau \approx 3 \times 10^{-14}\,\text{m} = 30\,\text{fm}$, if we assume a lifetime for the particle at rest of 10^{-23} s. This is much smaller than the best experimental resolution given, but larger than the range of the strong interaction. This is important if the decay of, for example, the X^- produced in reaction (3.22a) is to be treated as the decay of a free particle, since this requires that it should be sufficiently far from any other hadrons that are present (e.g. the proton in (3.22a)).

3.6 An argument similar to that given in Section 3.6 gives the following allowed combinations:

Baryons		*Mesons*	
C	Q	C	Q
3	2	1	1,0
2	2,1	0	1,0, −1
1	2,1,0	−1	0, −1
0	2,1,0, −1		

3.7 The first two quantum number combinations are compatible with the assignments

$$(1,0,0,1,1) = c\bar{b}, \quad (-1,1,-2,0,-1) = ssb$$

and can exist within the quark model. There are no combinations $q\bar{q}$ or qqq that are compatible with the second two combinations, so these cannot exist within the quark model. The combination $(0,0,1,0,1)$ must be a meson $q\bar{q}$ because $B = 0$, but must contain both an $\bar{s}$ antiquark and a $\bar{b}$ antiquark since $S = \tilde{B} = 1$. These are incompatible requirements.

The combination $(-1, 1, 0, 1, -1)$ must be a baryon qqq of the form xcb (where $x = u$ or d) since $B = 1$, $S = 0$, $C = 1$ and $\tilde{B} = -1$. The possible charges are $Q = 1$ and 0, corresponding to ucd and dcb, respectively, which are incompatible with the requirement $Q = -1$.

PROBLEMS 4

4.1 For constant acceleration, the proton must travel the length of the drift tube in half a cycle of the r.f. field. Thus, $L = \upsilon/2f$, where υ is the velocity of the proton. Since the energy is far less than the rest mass of the proton, we can use nonrelativistic kinematics to find υ, i.e. $\upsilon = c\sqrt{600/938.3} = 2.40 \times 10^8$ m s^{-1} and finally $L = 2.4$ m.

4.2 The equations of motion are $\mathrm{d}\mathbf{p}/\mathrm{d}t = e\mathbf{v} \times \mathbf{B}$ and $\mathrm{d}E/\mathrm{d}t = 0$, where the second follows from the fact that the Lorentz force is perpendicular to the direction of motion and so does no work on the particle. The magnitude of the velocity $\upsilon = |\mathbf{v}|$ and $\gamma(\upsilon)$ are therefore constant, so that the first equation becomes

$$m\gamma \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = e\mathbf{v} \times \mathbf{B}.$$

which reduces to $m\gamma\upsilon^2/\rho = e\upsilon B$ for a circular orbit. Hence $p = m\gamma\upsilon = e\rho B$, and the desired result follows from

$$pc = ec\rho B = 4.8 \times 10^{-11} \rho B$$

in SI units (i.e. joules, where 1 joule $= 0.62 \times 10^{19}$ eV).

4.3 The total centre-of-mass energy is given by

$$E_{CM}^2 = W^2c^4 = (E_A + E_B)^2 - (\mathbf{p}_A c + \mathbf{p}_B c)^2$$

by (A.8) and (A.10a). On neglecting the particle masses this reduces to

$$E_{CM}^2 = 2E_A E_B(1 - \cos\phi),$$

where ϕ is the angle between $\mathbf{p}_A$ and $\mathbf{p}_B$. Here $\phi = \pi - \theta$ (since the zero crossing angle corresponds to a head-on collision $\mathbf{p}_A = -\mathbf{p}_B$) and the desired result

$$E_{CM}^2 = 2E_A E_B(1 + \cos\theta)$$

follows directly.

The resulting centre-of-mass energy at HERA was 310 GeV. For a fixed-target machine, the centre-of-mass energy is given by (4.1), which reduces to

$$E_{CM} = \left(2m_P c^2 E_L\right)^{1/2}$$

if we neglect the squared particle masses. Hence

$$E_L = E_{CM}^2/2m \approx 5 \times 10^4 \text{ GeV}$$

for a centre-of-mass energy of 310 GeV.

4.4 From Problem (4.2), the momentum in GeV/c is given by $p = 0.3B\rho$, (where B is in Tesla and ρ is in metres) and to a good approximation $\theta \approx L/\rho = 0.3LB/p$, where the various quantities are defined in Figure F.10.

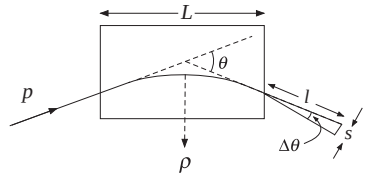


Figure F.10

Thus

$$\Delta\theta = \frac{s}{l} = \frac{0.3BL}{p^2}\Delta p,$$

where $\Delta p/p = 0.02$ if a slit of width s is used to select particles within 1 % of the central value. Substituting the data given gives $l = 8.3$ m.

4.5 The average distance between collisions of a neutrino and an iron nucleus is the mean free path $\lambda = 1/n\sigma_\nu$, where $n \approx \rho/m_P$ is the number of nucleons per cm^3. Using the data given we obtain $n \approx 5 \times 10^{24}$ cm^{-3} and $\sigma_\nu \approx 2 \times 10^{-36}$ cm^2, so that $\lambda \approx 10^{11}$ cm. Thus if 1 in 10^9 neutrinos is to interact, the thickness of iron required is 1 m.

4.6 Using $E = \gamma Mc^2$ and $\gamma = (1-\beta^2)^{-1/2}$, where M is the mass of the projectile, we have

$$\frac{\mathrm{d}E}{\mathrm{d}\beta} = Mc^2\frac{\beta}{(1-\beta^2)^{3/2}}.$$

Therefore, using this and Equation (4.11) in (4.12) gives

$$R = \frac{M}{q^2 n_e}\int_0^{\beta_i} f(\beta)\mathrm{d}\beta,$$

where $f(\beta)$ is a function of β. The result follows directly.

4.7 A particle with velocity υ will take time $t = L/\upsilon$ to pass between the scintillation counters. Relativistically, $p = m\gamma\upsilon$, which gives $\upsilon = c(1 + m^2c^2/p^2)^{-1/2}$ on solving for υ. Thus, the difference in flight times is (taking $m_1 > m_2$)

$$\Delta t = \frac{L}{c}\left[\left(1 + \frac{m_1^2c^2}{p^2}\right)^{1/2} - \left(1 + \frac{m_2^2c^2}{p^2}\right)^{1/2}\right].$$

For large momenta such that $p^2 \gg m_1^2c^2 > m_2^2c^2$, we can expand the brackets to give

$$\Delta t \approx \frac{L(m_1^2c^2 - m_2^2c^2)}{2p^2c},$$

which decreases like p^{-2}. For pions and kaons with momentum 3 GeV/c, we can use the approximate formula and the minimum value of the flight path is $L \approx 4.8$ m.

4.8 For a pion to give a signal, but not a kaon, we must have $\beta_\pi n > 1 > \beta_K n$, where $\beta = \upsilon/c = (1 + m^2c^2/p^2)^{-1/2}$ from $p = m\gamma\upsilon$. For $p = 20$ GeV/c, this gives $\beta_\pi = 0.99997$ and $\beta_K = 0.99970$, so that the condition on the refractive index n becomes

$$3 \times 10^{-4} \geq n - 1 \geq 3 \times 10^{-5}.$$

Using the largest value of $n = 1.0003$ and taking the visible spectrum to lie between $\lambda_1 = 400\,\text{nm}$ and $\lambda_2 = 700\,\text{nm}$ gives

$$2\pi\alpha\left(1 - \frac{1}{\beta_\pi^2 n^2}\right)\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) = 26.5 \text{ photons/metre}$$

by (4.23). Therefore to obtain 200 photons requires a detector of length 7.5 m.

4.9 The number of particles with energy exceeding E is

$$N(>E) = \sum_{t=0}^{t(E)} N(t) \approx \int_0^{t(E)} N(t)\,\mathrm{d}t,$$

where $N(t) = 2^t$ and $t(E) = \ln(E_0/E)/\ln 2$, since after t radiation lengths the average energy is $E_0/2^t$. Thus,

$$N \approx \int_0^{t(E)} \exp(t\ln 2)\,\mathrm{d}t = \frac{1}{\ln 2}\{\exp[t(E)\ln 2] - 1\} \approx \frac{1}{\ln 2}\frac{E_0}{E},$$

implying $\mathrm{d}N/\mathrm{d}E \propto E^{-2}$ in this approximation.

Since all the particles travel one radiation length and two-thirds of them are charged, the total track length of all charged secondaries (in radiation lengths) is approximately

$$L = \frac{2}{3}\int_0^{t_{\max}} N(t)\mathrm{d}t \quad \text{where} \quad t_{\max} = \frac{\ln(E_0/E_C)}{\ln 2}.$$

Thus,

$$L = \frac{2}{3}\frac{1}{\ln 2}\left[\exp(t_{\max}\ln 2) - 1\right] \approx \frac{2}{3\ln 2}\frac{E_0}{E_C}$$

and hence $L \propto E_0$.

4.10 To be detected, the event must have $155^\circ < \theta < 25^\circ$, i.e. $|\cos\theta| < 0.906$. Setting $x = \cos\theta$, the fraction of events in this range is

$$f = \int_{-0.906}^{+0.906} \frac{\mathrm{d}\sigma}{\mathrm{d}x}\mathrm{d}x \Bigg/ \int_{-1.0}^{+1.0} \frac{\mathrm{d}\sigma}{\mathrm{d}x}\mathrm{d}x = \left(x + x^3/3\right)_{-0.906}^{+0.906} \Big/ \left(x + x^3/3\right)_{-1}^{+1} = 0.864.$$

The total cross-section is given by

$$\sigma = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_{-1}^{+1} \mathrm{d}\cos\theta \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = 2\pi\frac{\alpha^2\hbar^2c^2}{4E_{CM}^2}\int_{-1}^{+1}(1 + \cos^2\theta)\,\mathrm{d}\cos\theta.$$

Using $E_{CM} = 6\,GeV$ gives $\sigma = 4\pi\alpha^2\hbar^2c^2/3E_{CM}^2 = 2.406\,\text{nb}$. The rate of production of events is given by $L\sigma$ and since L is a constant, the total number of events produced will be $L\sigma\, t = 2.41 \times 10^5$.

PROBLEMS 5

5.1 Since each particle is displaced by the same amount $\delta\mathbf{r}$ under a translation, the appropriate generalization of (5.4) is

$$\hat{D}\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) \equiv \Psi(\mathbf{r}_1 + \delta\mathbf{r}, \mathbf{r}_2 + \delta\mathbf{r}, \ldots, \mathbf{r}_N + \delta\mathbf{r}).$$

The argument then follows the same steps as for a single-particle system.

5.2 This is done by verifying (5.26). For example, in the case of $\hat{\mathbf{L}}^2$,

$$\begin{aligned}\left[\hat{\mathbf{L}}^2, H\right] &= \alpha\left[\hat{\mathbf{L}}^2, \hat{\mathbf{L}}\cdot\hat{\mathbf{S}}\right] \\ &= \alpha\left[\hat{\mathbf{L}}^2, \hat{L}_x\right]\hat{S}_x + \alpha\left[\hat{\mathbf{L}}^2, \hat{L}_y\right]\hat{S}_y + \alpha\left[\hat{\mathbf{L}}^2, \hat{L}_z\right]\hat{S}_z = 0\end{aligned}$$

where we have used the standard commutation relations $\left[\hat{\mathbf{L}}^2, \hat{L}_x\right] = 0$, etc.

5.3 For a *pn* state with orbital angular momentum L, the parity is given by

$$P = P_n P_p (-1)^L = (-1)^L,$$

where the intrinsic parities $P_n = P_p = 1$ and the total angular momentum J can be $J = L+1, L$ or $L-1$, since the total spin S can be 0 or 1. For $J = 1$ and $P = 1$, the only possible solutions are $L = 0$ and $L = 2$.

5.4 The initial state has $L = 0$, and hence using (5.36) its parity is

$$P_i = P_p P_{\bar{p}} (-1)^L = -1.$$

In the final state, the neutral pions are identical bosons and so the wavefunction must be symmetric under their interchange. This implies even orbital angular momentum L, and hence

$$P_f = P_\pi^2 (-1)^L = 1 \neq P_i.$$

The reaction therefore violates parity conservation, and is thus forbidden as either a strong or electromagnetic interaction.

5.5 By (5.29b), $J = \frac{1}{2}, \frac{3}{2}$ and $\frac{5}{2}$, while by (5.39b), $P = -1$ in all cases.

5.6 Since $\hat{C}^2 = 1$, we must have

$$\hat{C}^2 |a, \Psi\rangle = C_a \hat{C} |\bar{a}, \Psi\rangle = |a, \Psi\rangle$$

implying

$$\hat{C} |\bar{a}, \Psi\rangle = C_{\bar{a}} |a, \Psi\rangle$$

with $C_a C_{\bar{a}} = 1$ independent of C_a. The desired result follows because an eigenstate of $\hat{C}$ must contain only particle–antiparticle pairs $a\bar{a}$, leading to intrinsic parity factors $C_a C_{\bar{a}} = 1$, independent of C_a.

5.7 Since parity is conserved, we need to calculate the parity P_f of the three-pion final state in its centre-of-mass frame, which is identical with the rest frame of the η. The parity of the η is then given by

$$P_\eta = P_f = P_\pi^3 (-1)^{L_{12}} (-1)^{L_3},$$

where $P_\pi = -1$ is the intrinsic parity of the pion and L_{12}, L_3 are orbital angular momenta defined in analogy with Figure 5.1 for a three-quark state (cf. Equations (5.39b) and (5.39c)). The final state must also have zero angular momentum since the η has spin-0, implying

$$\mathbf{L} = \mathbf{L}_{12} + \mathbf{L}_3 = \mathbf{0}$$

and hence $L_{12} = L_3$. We thus obtain the final result $P_\eta = P_\pi^3 = -1$ for the parity of the η.

By angular momentum conservation, the final states in the decays $\eta \to \pi^+ + \pi^-$ and $\eta \to \pi^0 + \pi^0$ must have $L = 0$ and hence parity $P_f = P_\pi^2 (-1)^L = 1$. This is inconsistent with parity conservation since $P_\eta = -1$, as we have seen, so that these decays are forbidden in both strong and electromagnetic interactions. They can occur in weak interactions, but the branching ratios are negligibly small.

PROBLEMS 6

6.1 The quantum numbers are $B = 1, Q = 1, S = 0, C = 1, \tilde{B} = 0$ and $T = 0$, implying the quark content $\Sigma_c^+(2455) = udc$. Since the π^0 has $I = 1, I_3 = 0$ and the Λ_c^+ has $I = I_3 = 0$, the Σ_c^+ has $I = 1, I_3 = 0$. It therefore has partners that have $I = 1$, $I_3 = \pm 1$, and hence charge $Q = 2$ or 0 by (6.5). They are the $\Sigma_c^{++} = uuc$ and the $\Sigma_c^0 = ddc$. All three states have been observed.

6.2 From Equations (6.19) one obtains

$$\delta = M(\Sigma^+) + M(\Sigma^-) - 2M(\Sigma^0)$$

while the Coulomb approximation gives

$$\delta = \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r}.$$

Converting this expression to practical units and substituting the measured Σ masses from Table E.4 gives $\delta = 1.5 \pm 0.1\,\mathrm{MeV/c^2}$ and a distance r of order 1 fm. (In a full treatment one must average properly over interquark distances and take the full interaction into account, including the magnetic interaction, which is not negligible. Nevertheless, one obtains a similar estimate for the average interquark distance, which is comparable to the typical size of a hadron.)

6.3 The parity follows from

$$J = L;\ P = P_\pi^2 (-1)^L = (-1)^J,$$

and the C-parity follows from Equation (5.52). The decay $f_2^0 \to \pi^0\gamma$ violates $C-$parity and is forbidden as an electromagnetic interaction. (The $\rho^0 \to \pi^0\gamma$ decay is allowed and has a branching ratio of order 10^{-3}.) The decay $\rho^0 \to \pi^0\pi^0$ is absolutely forbidden because it is incompatible with angular momentum conservation and the Pauli principle. The first of these requires $L = 1$, while the second requires L to be even in order that the final state wavefunction is symmetric under the interchange of identical bosons. (The decay $f_2^0 \to \pi^0\pi^0$ is allowed as a strong decay and has a branching ratio comparable with that for $f_2^0 \to \pi^+\pi^-$.)

6.4 The possible spectroscopic states and their J^P values are the same as those of positronium listed in Table 5.1 or charmonium listed in Table 6.9. The states shown in Figure 3.12 are

0^-:	$^1S_0$	with $n=1$
1^-:	$^3S_1$	with $n=1$ and $n=2$
0^+:	$^3P_0$	with $n=2$
1^+:	$^3P_1$ and $^1P_1$	with $n=2$
2^+:	$^3P_2$	with $n=2$

The lighter of the two 1^+ states is in fact the $^1P_1$ state, but this cannot be deduced from the information given.

6.5 For an arbitrary value of L, the possibilities are $S=0, J=L$, leading to $P=(-1)^{J+1}$, $C=(-1)^J$ by (5.39a) and (5.53); $S=1, J=L\pm 1$, leading to $P=(-1)^J$, $C=(-1)^J$; and $S=1, J=L$, leading to $P=(-1)^{J+1}$, $C=(-1)^{J+1}$. This exhausts the possibilities, so that the combination $P=(-1)^J$, $C=(-1)^{J+1}$ does not occur, and mesons with $J^{PC}=0^{+-}, 1^{-+}, 2^{+-}, 3^{-+}$, etc., are forbidden in the simple quark model. Such mesons are sometimes called *exotics of the second kind* and none have been observed.

6.6 There would still be an octet of $\frac{1}{2}^+$ baryons as before, but only a single $\frac{3}{2}^+$ state $\Lambda=uds$ instead of the observed decuplet. This can be seen by the argument of Section 6.2.2 with the appropriate modifications. Thus, the parity of baryon states $B=abc$ is $P=P_aP_bP_c=1$ by (5.39b) and the baryon spin is the sum of the quark spins. If we assume that for $L=0$ the lowest-lying states have symmetric spatial wavefunctions, then the spin wavefunctions will be antisymmetric and any like pair of quarks must have spin-0 corresponding to the antisymmetric spin wavefunction (5.44). Hence in the six combinations (uud, uus, ddu, dds, ssu, ssd) the identical pair of quarks have spin-0, and adding the spin of the unlike quark gives $J=S=\frac{1}{2}$ for all six combinations. The three combinations (uuu, ddd, sss) are forbidden because there is no way to have an antisymmetric spin wavefunction with respect to all pairs. There remains the case (uds) with all quarks different. Here the uds pair can have either spin-0 or spin-1. In the former case, adding the spin of the third quark gives a single $\frac{1}{2}^+$ baryon. In the spin-1 case, adding the third spin yields a $\frac{1}{2}^+$ state and a $\frac{3}{2}^+$ state. Thus, overall, there are eight $J^P=\frac{1}{2}^+$ baryons and one with $J^P=\frac{3}{2}^+$.

6.7 The key point is that for quark combinations cab with $a\neq b$, the ab pair can have spin-0 or spin-1, whereas for the combinations caa the aa pair must have spin-1 to give a spin wavefunction that is symmetric under the exchange of the light quarks. On adding the spin of the c quark in each case one obtains:

(a) A triplet of spin-$\frac{1}{2}$ states csu, csd and cdu in which the light quark pairs have spin-0. These states have quantum numbers $(C,S,I_3)=(1,-1,\frac{1}{2}),(1,-1,-\frac{1}{2}),(1,0,0)$, respectively, and constitute an isodoublet Ξ_c^+, $\Xi_c^0\equiv csu, csd$ and an isosinglet $\Lambda_c^+\equiv cud$, where we introduce the conventional symbols for each type of multiplet.

(b) A sextet of spin-$\frac{1}{2}$ states css, csu, csd, cuu, cud and cdd in which the light quark pair has spin-1. These constitute an isotriplet Σ_c^{++}, Σ_c^+, $\Sigma_c^0\equiv cuu$, cud, cdd an isodoublet Ξ_c^+, $\Xi_c^0\equiv csu$, csd and an isosinglet state $\Omega_c^0\equiv css$.

(c) A sextet of spin-$\frac{3}{2}$ states with the same quark and isospin compositions as the spin-$\frac{1}{2}$ sextet.

(The reader may find it instructive to plot these states on weight diagrams similar to those of Figures 6.1 and 6.2.)

6.8 The quantum numbers Q, C and S establish the quark compositions

$$Y^+, Y^0 = csu,\ csd \equiv \Xi_c^+, \Xi_c^0$$

$$Z^{++}, Z^+, Z^0 = cuu,\ cud,\ cdd \equiv \Sigma_c^{++}, \Sigma_c^+, \Sigma_c^0$$

$$X^+ = cud \equiv \Lambda_c^+$$

$$V^0 = css \equiv \Omega_c^0$$

The Λ_c^+ can only belong to the triplet (a) of the previous question and the Ω_c^0 and $\Sigma_c^{++}, \Sigma_c^+, \Sigma_c^0$ can only belong to the sextet (b), but the Ξ_c^+, Ξ_c^0 could belong to either. However, it is reasonable to assume that the main contribution to the mass differences in a baryon supermultiplet arises from the differences in the quark masses, as was the case for noncharmed baryons. The observed mass differences

$$M_Y - M_{\Lambda_c} \approx 184\,\mathrm{MeV/c^2} \approx m_s - m_{u,d}$$

and

$$M_Y - M_{\Sigma_c} \approx 15\,\mathrm{MeV/c^2} \ll m_s - m_{u,d}$$

suggest that we identify the Y^+, Y^0 with the Ξ_c^+, Ξ_c^0 states in the triplet (a). The same argument leads to the prediction

$$M_{\Omega_c} - M_{Y_c} \approx M_{Y_c} - M_{\Xi_c} \approx m_s - m_{u,d}$$

for the states of the sextet. Using the values $M_{\Omega_c} = 2700\,\mathrm{MeV/c^2}$ and $M_{\Sigma_c} = 2455\,\mathrm{MeV/c^2}$ gives the prediction

$$M_{\Xi_c} \approx \tfrac{1}{2}(M_{\Omega_c} + M_{\Sigma_c}) = 2579\,\mathrm{MeV/c^2}$$

for the $C = 1$, $S = -1$ isodoublet state of the sextet. Such a state has been seen with a mass of $2576\,\mathrm{MeV/c^2}$.

6.9 In the $\Sigma^0(1193) = uds$, the ud pair is in a spin-1 state, as discussed in Section 6.2.2. This pair has a magnetic moment $\mu_u + \mu_d$ since the u and d spins are parallel. An argument similar to that given for the magnetic moments of baryons $B = aab$ in Section 6.2.4 then gives

$$\mu_{\Sigma^0} = \tfrac{2}{3}(\mu_u + \mu_d) - \tfrac{1}{3}\mu_s = 0.83\,\mu_N,$$

where μ_N is the nuclear magneton and we have used the quark masses given in (6.32). (Unfortunately, this prediction cannot easily be checked experimentally because the Σ^0 is too short-lived.)

6.10 The conditions $\hat{F}_i^C \chi_B = 0$ are automatically satisfied for $i = 3$ and 8, as shown in the text. It was also shown that for $i = 1$ the conditions

$$\alpha_1 = -\alpha_2, \quad \alpha_3 = -\alpha_4, \quad \alpha_5 = -\alpha_6$$

resulted. A similar argument for $i = 2$, 4, 5, 6 and 7 gives

$$\begin{array}{llll} \alpha_1 = -\alpha_2, & \alpha_3 = -\alpha_4, & \alpha_5 = -\alpha_6 & (i = 1, 2) \\ \alpha_1 = -\alpha_4, & \alpha_2 = -\alpha_5, & \alpha_3 = -\alpha_6 & (i = 4, 5) \\ \alpha_1 = -\alpha_6, & \alpha_2 = -\alpha_3, & \alpha_4 = -\alpha_5 & (i = 6, 7) \end{array}$$

which are satisfied if, and only if,

$$\alpha_1 = -\alpha_2 = \alpha_3 = -\alpha_4 = \alpha_5 = -\alpha_6,$$

giving (6.36) up to an overall normalization factor.

PROBLEMS 7

7.1 Substituting (7.21) into (7.22) and setting $N_C = 3$ gives

$$R = 3(1 + \alpha_s/\pi) \sum e_q^2$$

where α_s is given by (7.6) evaluated at $\mu^2 = E_{CM}^2$ and where the sum is over those quarks that can be produced in pairs at the energy in question. Using the quark masses given in Table E.3 we find that u, d and s quarks can contribute at 2.8 GeV; u, d, s and c quarks at 5 GeV; and *u, d, s, c* and b quarks at 15 GeV. Using the quark charges and the values of α_s from Figure 7.3 then gives $R \approx 2.17$, 3.58 and 3.89 at $E_{CM} = 2.8$, 5 and 15 GeV, respectively. When E_{CM} is above the $t\bar{t}$ threshold, t quarks also contribute to the sum and R rises to $R = 5(1 + \alpha_s/\pi)\, R$. (This ignores weak interactions.)

7.2 By choosing the z axis in the $\mathbf{q}$ direction and using spherical polar coordinates so that $\mathbf{q} \cdot \mathbf{r} = qr\cos\theta$, we obtain

$$G_E(q^2) = 4\pi \int_0^\infty \mathrm{d}r\; r^2 \rho(r) \left(\frac{\sin qr}{qr} \right).$$

(a) The integral is an even function of q, so that an expansion can contain only even powers $q^{2n} = (q^2)^n$. Thus G_E is a function of q^2 only.

(b) On expanding the sine function, we obtain

$$G_E(q^2) = 4\pi \int_0^\infty \mathrm{d}r\; r^2 \rho(r) \left[1 - \frac{q^2 r^2}{3!} + \frac{q^4 r^4}{5!} + \ldots \right]$$

so that

$$\left. \frac{\mathrm{d}G_E(q^2)}{\mathrm{d}q^2} \right|_{q^2=0} = -\frac{4\pi}{6} \int_0^\infty \mathrm{d}r\, r^4\, \rho(r) = -\frac{r_E^2}{6},$$

which is the desired result.

7.3 By choosing the z axis in the $\mathbf{q}$ direction and using spherical polar coordinates so that $\mathbf{q} \cdot \mathbf{r} = qr\cos\theta$, the angular integrations may be easily done to yield

$$\mathcal{M}_C(\mathbf{q}) = -\frac{4\pi\alpha}{q} \lim_{\mu \to 0} \int_0^\infty \mathrm{d}r (\mathrm{e}^{-\mu r} \sin qr).$$

The latter integral may be solved by integrating twice by parts and yields a value $1/q$.

7.4 By applying energy–momentum conservation to Figure 7.19 we find

$$\begin{aligned} W^2 &= [(E - E') + E_P]^2 - [(\mathbf{p} - \mathbf{p}') + \mathbf{P}]^2 \\ &= -Q^2 + M^2 + 2E_P(E - E') - 2\mathbf{P} \cdot (\mathbf{p} - \mathbf{p}'), \end{aligned}$$

where we have used (7.32) and $E_P^2 - \mathbf{P}^2 = M^2$. Substituting in the definition (7.50) then gives

$$2Mv = 2E_P(E - E') - 2\mathbf{P} \cdot (\mathbf{p} - \mathbf{p}')$$

in an arbitrary frame. The desired result is obtained in the proton rest frame by setting $E_P = M$ and $\mathbf{P} = \mathbf{0}$.

The lightest possible hadron state X is the proton, so that $W^2 \geq M^2$, and $x \leq 1$ follows directly from (7.50) and (7.51). The lower bound $x \geq 0$ follows from the fact that both v and Q^2 are positive in the approximation when the lepton masses are neglected. For v this follows from the proton rest frame result (7.52), since some of the initial lepton energy E must be converted to proton recoil energy; for Q^2 it follows from

$$Q^2 = 2EE'(1 - \cos\theta) \geq 0,$$

which is easily derived from the definition (7.32) when the lepton masses are neglected.

7.5 For elastic scattering, $W^2 = M^2$, since the final hadronic state is a single proton. Substituting into (7.50) gives $2Mv = Q^2$ and hence $x = 1$ by (7.51). To show the relation between E and E', we substitute (7.52), together with the result

$$Q^2 = 2EE'(1 - \cos\theta) \geq 0$$

(see the solution to Problem 7.4), into the above equation for v.

7.6 Before the scattering occurs, the momentum of the parton is $z\mathbf{P}$ and its energy is zE_P, as discussed just before Equation (7.55). Hence, from energy and momentum conservation, we have

$$k = (\mathbf{p} - \mathbf{p}') + z\mathbf{P}, \quad \omega = (E - E') + zE_P,$$

as can be seen from Figure 7.20. Substituting into $\omega^2 - \mathbf{k}^2 = \tilde{m}^2 \approx 0$ leads to $2Mvz - Q^2 = 0$ and hence the desired result, where we have neglected the squared proton, parton and lepton masses compared with Q^2, and have used (7.32) and the relation

$$2Mv = 2E_P(E - E') - 2\mathbf{P} \cdot (\mathbf{p} - \mathbf{p}')$$

(see the solution to Problem 7.4).

7.7 Since $q(x)\mathrm{d}x$ and $\overline{q}(x)\mathrm{d}x$ are the probabilities of finding a quark or antiquark, respectively, with fractional momentum between x and $x + \mathrm{d}x$, the total fractional momentum f carried by the quarks and antiquarks is

$$f = \int_0^1 \mathrm{d}x[xq(x) + x\overline{q}(x)].$$

This is just the average of $q(x)\mathrm{d}x$ plus the average of $\overline{q}(x)\mathrm{d}x$, and from Figure 7.23 we can estimate f to be approximately one-half. The 'missing momentum' is assumed to be carried by gluons produced in processes like Figure 7.22(b).

PROBLEMS 8

8.2 In Section 2.2.2 we used a simple dimensional argument to show that

$$\Gamma(\ell^- \to e^- + \overline{\nu}_e + \nu_\ell) = KG_F^2 m_\ell^5 \quad (\ell = \tau, \mu),$$

when the masses of the final particles are neglected. The same argument gives

$$\Gamma(c \to s + e^+ + \nu_e) = KG_F^2 m_c^5 \cos^2\theta_C,$$

where the constant K is the same because the mechanism is the same except for the replacement of the $\mu\nu W$ coupling of strength g_W by the csW coupling of strength $g_{cs} = g_W \cos\theta_C$ (cf. Equation (8.17a)). From the above equations and setting $\Gamma(\mu^- \to e^- \overline{\nu}_e \nu_\mu) = \tau_\mu^{-1}$, where τ_μ is the muon lifetime, one obtains

$$\Gamma(c \to s e^+ \nu_e) = \frac{m_c^5 \cos^2\theta_C}{m_\mu^5 \tau_\mu} = (1.7 - 3.4) \times 10^{11}\ \mathrm{s}^{-1},$$

where the range reflects the uncertainty in the c quark mass. There will be an additional error of order 10 % from the neglect of the s quark mass (cf. the discussion in the second paragraph of Section 2.2.2).

8.3 For any decay, the rate is given by $\Gamma = B/\tau$, where B is the branching ratio and τ is the lifetime of the decaying particle. Substituting the measured values for these gives $\Gamma = (1.6 \pm 0.2) \times 10^{11}\ \mathrm{s}^{-1}$, $(1.9 \pm 0.2) \times 10^{11}\ \mathrm{s}^{-1}$ and $(2.4 \pm 0.1) \times 10^{11}\ \mathrm{s}^{-1}$ for the D^+, D^0 and Λ_c^+ decays, respectively. These values are in good agreement with each other and the estimated rate obtained in the previous problem.

8.4 The decays (a) and (c) are forbidden by the $\Delta S = \Delta Q$ rule (8.30) and (e) is forbidden by the $\Delta S = 0, \pm 1$ rule (8.32). Decays (b), (d) and (f) are allowed decays and have all been observed experimentally.

8.5 Cabibbo-allowed decays involve the csW vertex of Figure 8.10(c), giving rise to the selection rule $\Delta C = \Delta S = \Delta Q = \pm 1$. Cabibbo-suppressed decays involve the cdW vertex of Figure 8.13(c), giving rise to the selection rules $\Delta C = \Delta Q = \pm 1, \Delta S = 0$. Using these rules one sees that the decays are: (a) Cabibbo-allowed, (b) forbidden, (c) forbidden and (d) Cabibbo-suppressed.

8.6 An arbitrary complex $n \times n$ matrix $\mathbf{U}$ has $2n^2$ real parameters. The matrix $\mathbf{F} \equiv \mathbf{U}^\dagger \mathbf{U}$ is Hermitian by construction, so $F_{ij} = F_{ji}^*$ and it has n^2 real parameters. Hence, the condition $\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}$ imposes n^2 conditions on $\mathbf{U}$, leaving n^2 real parameters undetermined. Since (8.60) has $n^2 = 4$ real parameters and satisfies $\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}$, it is the most general 2×2 unitary matrix.

Substituting (8.60) into (8.61) gives

$$d' = \mathrm{e}^{-i\alpha}(\mathrm{e}^{i\beta} d \cos\theta_C + \mathrm{e}^{i\gamma} s \sin\theta_C),$$
$$s' = \mathrm{e}^{-i\alpha}(-\mathrm{e}^{-i\gamma} d \sin\theta_C + \mathrm{e}^{-i\beta} s \cos\theta_C),$$

which can be written

$$\mathrm{e}^{i(\alpha-\beta)} d' = d \cos\theta_C + \mathrm{e}^{i(\gamma-\beta)} s \sin\theta_C,$$
$$\mathrm{e}^{i(\alpha+\gamma)} s' = -d \sin\theta_C + \mathrm{e}^{i(\gamma-\beta)} s \cos\theta_C.$$

Redefining the phases of the quark states by

$$e^{i(\alpha-\beta)} d' \to d', \quad e^{i(\alpha+\gamma)} s' \to s', \quad e^{i(\gamma-\beta)} s \to s$$

gives (8.39) as required.

8.7 If (8.17) is exact, then

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C & V_{ub} \\ -\sin\theta_C & \cos\theta_C & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

by (8.42). The remaining elements are then determined by exploiting the fact that the sum of the squared moduli of the elements in any row or column of a unitary matrix must be unity. (This follows directly from $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$ and $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$, respectively.) For example,

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = \cos^2\theta_C + \sin^2\theta_C + |V_{ub}|^2 = 1,$$

implying $V_{ub} = 0$, and so on.

8.8 For τ decay, the argument of Section 2.2.2 gives

$$\Gamma(\tau^- \to \nu_\tau + e^- + \overline{\nu}_e) = KG_F^2 m_\tau^5, \tag{2.20b}$$

where K is a dimensionless constant. On comparing Figures 8.20 and 8.22, we see that b decay differs by the replacement

$$g_W \to g_{\alpha b} = V_{\alpha b} g_W$$

at one vertex, where $\alpha = u, c$. Hence (2.20b) is modified to

$$\Gamma(b \to q + e^- + \overline{\nu}_e) = K \left|V_{qb}\right|^2 G_F^2 m_b^5$$

and (8.62) follows directly.

When quark masses are taken into account, (8.62) is replaced by (8.63). Substituting data from Appendix E in (8.63) gives

$$\Gamma(b \to u + e^- + \overline{\nu}_e) = 5.9 \times 10^{13} \, |V_{ub}|^2 \ \mathrm{s}^{-1}$$

and

$$\Gamma(b \to c + e^- + \overline{\nu}_e) = 2.8 \times 10^{13} \, |V_{cb}|^2 \ \mathrm{s}^{-1}.$$

On comparing with (8.64), this implies

$$5.9 \times 10^3 \, |V_{ub}|^2 + 2.8 \times 10^3 \, |V_{cb}|^2 = 6.8 \pm 0.5$$

and hence the upper bounds

$$|V_{ub}|^2 \leq (3.4 \pm 0.1) \times 10^{-2} \quad \text{and} \quad |V_{cb}|^2 \leq (4.9 \pm 0.2) \times 10^{-2}.$$

8.9 In the simple Bohr model, the strong attraction between the two t quarks is balanced by the centripetal force, i.e.

$$\frac{\mu v^2}{r} = \frac{4}{3}\frac{\alpha_s}{r^2},$$

where r is the radius of the orbit, υ the quark velocity and μ the reduced mass ($\mu = m_t/2$). In addition, the angular momentum is quantized, i.e. $\mu\upsilon r = n$. From these two equations, the radius of the ground state ($n = 1$) is given by

$$a = \frac{3}{4}\frac{1}{\mu\alpha_s}.$$

Restoring factors of $\hbar$ and c, and using $m_t = 180\,\mathrm{GeV/c^2}$, gives

$$a = \frac{3(\hbar c)}{4(\mu c^2)\alpha_s} = 1.6 \times 10^{-2}\,\mathrm{fm}.$$

The time taken to traverse a single orbit of the ground state is

$$t = 2\pi a/\upsilon = 2\pi\mu a^2.$$

Restoring factors of $\hbar$ and $c$gives

$$t = \frac{2\pi(\mu c^2)a^2}{(\hbar c)c} = 2.8 \times 10^{-24}\,\mathrm{s}.$$

This is much longer than the expected lifetime of the top quark, which is about 4×10^{-25} s.

PROBLEMS 9

9.2 The decay (a) is allowed by W exchange (i.e. it is a weak interaction) while (c) is an allowed electromagnetic process. Decays (b) and (d) are both forbidden as electromagnetic interactions because $\Delta S \neq 0$, and are also forbidden as weak interactions because there are no strangeness-changing weak neutral currents.

9.3 From Figure 9.5(b), we see that the process required is $\bar{s} \to \bar{d} + \nu_\ell + \bar{\nu}_\ell$. This is allowed in second order by diagrams such as Figure F.11, so that the rate is proportional to $G_F^4 \sin^2\theta_C$, where we have taken $G_W \approx G_Z \approx G_F$, whereas the 'allowed' decay in Figure 9.5(a) is proportional to $G_F^2 \sin^2\theta_C$. Using the fact that G_F has dimensions $[E]^{-2}$ in natural units (see Table 1.1), and inserting sufficient powers of the kaon mass to give the correct dimension for the decay rate, gives

$$\frac{\sum_\ell \Gamma(K^+ \to \pi^+ \nu_\ell \bar{\nu}_\ell)}{\Gamma(K^+ \to \pi^0 \mu^+ \nu_\mu)} \sim \frac{3G_F^4 \sin^2\theta_C m_K^9}{G_F^2 \sin^2\theta_C m_K^5} = O(10^{-11}),$$

which is well below the experimental limit (9.7).

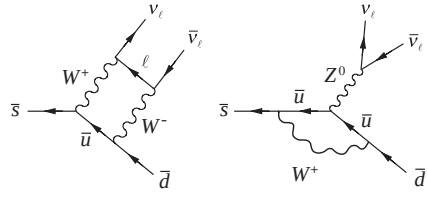


Figure F.11

9.4 See Figure F.12. At low energies, both diagrams reduce to the zero-range approximation diagram of Figure F.13, with couplings G_F and G_Z, respectively, where $G_Z \approx G_F \sin^2\theta_W$ by (9.13). By dimensional arguments, the cross-section will be of order $\sigma \sim G_F^2 E^2$ in natural units. Restoring factors of $\hbar$ and c in the usual way gives $\sigma \sim 5 \times 10^{-8} E^2$ mb, where E is measured in GeV.

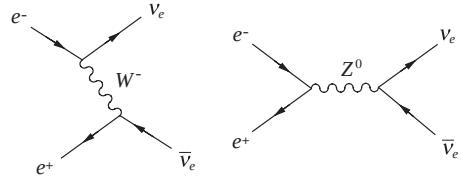


Figure F.12

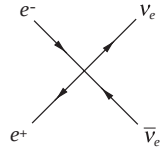


Figure F.13

9.5 Since $M_H < 2M_W$, the Higgs boson will decay to fermion–antifermion pairs by the mechanism of Figure 9.15, with a rate proportional to

$$\alpha_{Hff} \equiv \frac{g_{Hff}^2}{4\pi} = \frac{G_F m_f^2}{2\pi\sqrt{2}}$$

by (9.45), where we introduce α_{Hff} by analogy with the fine structure constant.

The dominant decay will be $H^0 \to b\bar{b}$, since these are the heaviest fermions available, where the b and $\bar{b}$ will manifest themselves as hadron jets. Using the above equation for α_{Hff}, the relative decay rates will be

$$\Gamma(H^0 \to \ell^+\ell^-) : \Gamma(H^0 \to b\bar{b}) = m_\ell^2 : 3m_b^2,$$

where the factor 3 arises from the three colour states of the b quark. Since the $b\bar{b}$ decay dominates, we have

$$B \approx \frac{\Gamma(H^0 \to \ell^+\ell^-)}{\Gamma(H^0 \to b\bar{b})} = \frac{m_\ell^2}{3m_b^2} \approx 4 \times 10^{-9}; 2 \times 10^{-4}; 5 \times 10^{-2}$$

for $\ell = e, \mu, \tau$, respectively.

Finally, by dimensional arguments, the total decay rate will be of order

$$\Gamma_{total} \approx \Gamma(H^0 \to b\bar{b}) \sim 3\alpha_{Hbb} M_H = O(10\,\text{MeV}).$$

9.6 From (9.17) at $E_{CM} = 2E = M_Z^2$, we have

$$\sigma_\gamma(e^+e^- \to \mu^+\mu^-) \approx \frac{4\alpha^2}{E_{CM}^2} = \frac{4\alpha^2}{M_Z^2} = \frac{2 \times 10^{-4}}{M_Z^2}.$$

From the Breit–Wigner formula (9.26) and (9.27) we have

$$\sigma_{Z^0}(e^+e^- \to \mu^+\mu^-) = \frac{12\pi}{M_Z^2} B(Z^0 \to e^+e^-)B(Z^0 \to \mu^+\mu^-) \approx \frac{4 \times 10^{-2}}{M_Z^2}$$

at $E_{CM} = M_Z$, where we have substituted the experimental branching ratios from Table E.1.

The Breit–Wigner formula for the corresponding Higgs contribution is

$$\sigma_{H^0}(e^+e^- \to \mu^+\mu^-) = \frac{4\pi}{M_Z^2} B(H^0 \to e^+e^-)B(Z^0 \to \mu^+\mu^-) \approx \frac{1 \times 10^{-11}}{M_Z^2},$$

using the branching ratios from the previous question. The amplitudes are proportional to the square root of the above hypothetical cross-sections, but the Higgs contribution is still negligible.

PROBLEMS 10

10.1 The argument used in Sections 10.2.1(a) and 10.2.1(c) to deduce the parities of the $\pi^0\pi^0$ and $\pi^0\pi^0\pi^0$ states shows that the $\pi^+\pi^0$ state has $P = 1$ and that the $\pi^+\pi^+\pi^-$ state has $P = -1$. Hence parity must be violated in one of the reactions irrespective of which value is assigned to the kaon.

10.2 The initial spin $J = 5$ of the ^{60}Co nuclei can only be conserved if the spins of all the final state particles point in the same direction. For $\theta = 0$, the electrons are emitted in this direction so they must have positive helicity; i.e. they must be in right-handed states e_R. For relativistic electrons, when $\upsilon \to c$, this is forbidden by the $\mathbf{V} - \mathbf{A}$-interaction, as discussed in Section 10.1.3. Hence $I(\upsilon = c, \theta = 0) = 0$, implying $\alpha = -1$, which is the observed value.

10.3 Consider the decay of a K^0 at rest, i.e.

$$K^0(\mathbf{p} = \mathbf{0}) \to \pi^-(\mathbf{p}_1) + e^+(\mathbf{p}_2, s_2) + \nu_e(\mathbf{p}_3, s_3)$$

where the momenta satisfy $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = \mathbf{0}$ and s_2 and s_3 are the electron and neutrino spins. *CP* changes particles into antiparticles, reverses the signs of the momenta and leaves the spins unchanged. Hence, by *CP* invariance the rates for the above decay and for

$$\overline{K}^0(\mathbf{p} = \mathbf{0}) \to \pi^+(-\mathbf{p}_1) + e^-(-\mathbf{p}_2, s_2) + \overline{\nu}_e(-\mathbf{p}_3, s_3)$$

are equal. The desired result follows on summing over all possible final state momenta and spins.

10.4 The decay in question is $B \to D\ell\nu_\ell$ and the appropriate matrix element is V_{cb}. The analogous formula (in natural units) for the lifetime τ_B is therefore

$$\frac{1}{\tau_B} = \frac{G_F^2 m_b^5}{192\pi^3} \frac{|V_{cb}|^2}{B(B \to D\ell\nu)},$$

where m_b is the mass of the bottom quark and $B(B \to D\ell\nu)$ is the branching ratio. Restoring factors of $\hbar$ and c gives

$$|V_{cb}|^2 = \frac{192\pi^3\hbar(\hbar c)^6 B(B \to D\ell\nu)}{\tau_B (m_b c^2)^5 G_F^2}.$$

Using $B = 0.11$, $\tau_B = 1.6 \times 10^{-12}\,\mathrm{s}$ and $m_b \approx 4.5\,\mathrm{GeV/c^2}$ from Table E.5, and $G_F/(\hbar c)^3 = 1.17 \times 10^{-5}\,\mathrm{GeV}^{-2}$, gives $|V_{cb}| = 0.032$. (The actual value is 0.04.)

10.5 Direct substitution, with A_0 real, gives

$$\mathcal{M}(K_S^0 \to \pi^0\pi^0) = \frac{2N}{\sqrt{3}}\left[\sqrt{2}\mathrm{e}^{i\delta_2}\,(\mathrm{Re}A_2 + i\varepsilon \mathrm{Im}A_2) - \mathrm{e}^{i\delta_0}A_0\right]$$
$$\approx \frac{2N}{\sqrt{3}}\left[\sqrt{2}\mathrm{e}^{i\delta_2}\mathrm{Re}A_2 - \mathrm{e}^{i\delta_0}A_0\right]$$

and

$$\mathcal{M}(K_L^0 \to \pi^0\pi^0) = \frac{2N}{\sqrt{3}}\left[\sqrt{2}\mathrm{e}^{i\delta_2}\,(\varepsilon \mathrm{Re}A_2 + i\ \mathrm{Im}\ A_2) - \varepsilon\mathrm{e}^{i\delta_0}A_0\right]$$
$$\approx \frac{2N}{\sqrt{3}}\left[i\sqrt{2}\,\mathrm{e}^{i\delta_2}\ \mathrm{Im}A_2 - \varepsilon\mathrm{e}^{i\delta_0}A_0\right]$$

where second-order terms have been neglected. Thus

$$\eta_{00} = \frac{\varepsilon - i\sqrt{2}\,\mathrm{e}^{i\Delta}\mathrm{Im}A_2/A_0}{1 - \sqrt{2}\,\mathrm{e}^{i\Delta}\mathrm{Re}A_2/A_0} \approx \varepsilon - i\sqrt{2}\exp(i\Delta)\frac{\mathrm{Im}A_2}{A_0},$$

where $\Delta \equiv \delta_2 - \delta_0$ and again second-order terms have been neglected.

PROBLEMS 11

11.1 The quantity $B - L$ is conserved provided the value $-2/3$ is assigned to both X and Y bosons.

11.2 Consider a single $\overline{\nu}_e$ with energy E and velocity υ. Then from $E = mc^2\gamma$, we have

$$\upsilon = c\left(1 - m^2c^4/E^2\right)^{1/2}$$

and the time taken to travel to Earth is

$$t = \frac{d}{\upsilon} = \frac{d}{c}\left(1 - \frac{m^2c^4}{E^2}\right)^{-1/2}.$$

The difference in travel times for two neutrinos with energies E_1 and E_2 is, for $E_2 > E_1 \gg mc^2$,

$$\Delta t = t_1 - t_2 \approx \frac{d(mc^2)^2}{2c}\left(\frac{1}{E_1^2} - \frac{1}{E_2^2}\right),$$

and if we assume that the observed time interval is due to the different energies of the neutrinos, then the data given lead to the upper bound $m < 15\,\mathrm{eV/c^2}$.

11.3 The appropriate relabellings are

$$\text{(a) } e^+ \to \overline{\nu}_e, \quad X \to Y, \quad \overline{u} \to \overline{d}; \quad \text{(b) } e^+ \to \overline{\nu}_e, \quad \overline{u} \to \overline{d}$$

for Figures 11.4(a) and (b), respectively.

11.4 For spherical mass distributions,

$$\frac{mv^2}{R} = \frac{GmM_R}{R^2},$$

where M_R is the galactic mass enclosed in a sphere of radius R.

(a) M_R is independent of R, giving $v \propto R^{-1/2}$.
(b) M_R increases like R^3, giving $v \propto R$.

The observed behaviour lies between these cases, implying that M_R is roughly proportional to R for the range of radii at which observations are made.

11.5 The results follow from the invariance of Maxwell's equations

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{1}{\varepsilon_0}\rho, \quad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

giving $\mathbf{B} \to \mathbf{B}$ and $\mathbf{B} \to -\mathbf{B}$ under P and T, respectively.

11.6 The energy of the photon is $k = \frac{3}{2}k_B T = 1.29 \times 10^{-10}$ MeV. The maximum centre-of-mass energy corresponds to 'head-on' collisions, so the invariant mass is given by

$$W^2 = (E + k)^2 - (p - k)^2 \approx m^2 + 4kE,$$

where (E, p) are the proton energy and momentum and we have assumed $E \gg m$, where m is the proton mass.

(a) For e^+e^- production, $W^2 \geq (m + 2m_e)^2 \approx m^2 + 4mm_e$, and hence

$$E \geq mm_e/k \approx 0.4 \times 10^{10}\,\text{GeV}.$$

(b) For $p\overline{p}$ production, $W^2 \geq 9m^2$ and hence

$$E \geq 2m^2/k \approx 1.4 \times 10^{13}\,\text{GeV}.$$

PROBLEMS A

A.1 For pions at rest, energy conservation gives $E_\mu + E_{\nu_\mu} = m_\pi^2 c^2$, and hence

$$m_\pi^2 c^4 = E_\mu^2 + E_\nu^2 + 2E_\mu(m_\pi^2 c^2 - E_\mu).$$

Using $E^2 = p^2c^2 + m^2c^4$ gives

$$m_\pi^2 c^2 = -m_\mu^2 c^2 + m_\nu^2 c^2 + 2E_\mu m_\pi^2 + (p_\nu^2 - p_\mu^2).$$

However, momentum conservation is $\mathbf{p}_\mu + \mathbf{p}_\nu = \mathbf{0}$; i.e. $\mathbf{p}_\mu^2 = \mathbf{p}_\nu^2$. Thus,

$$E_\mu = \frac{(m_\pi^2 + m_\mu^2 - m_\nu^2)c^2}{2m_\pi}.$$

(This result has been used to obtain an upper limit on the mass of the muon neutrino from a measurement of E_μ.)

A.2 If the momenta of the photons are $\mathbf{p}_{1,2}$, then the invariant mass W of the initial state is given by

$$W^2c^4 = E^2 - p^2c^2 = m^2c^4$$

and the same quantity for the final state is

$$W^2c^4 = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2c^2 = 2E_1E_2(1 - \cos\theta)$$

since

$$\mathbf{p}_1 \cdot \mathbf{p}_2 = p_1p_2\cos\theta = (E_1E_2\cos\theta)/c^2$$

for zero-mass photons. Equating the two expressions for the invariant mass gives

$$\cos\theta = 1 - m^2c^4/2E_1E_2.$$

A.3 The lightest final state that conserves strangeness, electric charge and baryon number is $K^0\Lambda$, so the reaction is $\pi^- + p \rightarrow K^0 + \Lambda$. This leads to a threshold energy of 0.908 GeV, corresponding to a threshold momentum of 0.897 GeV/c, by an argument similar to that used in Section A.2 to deduce the threshold energy for antiproton production.

A.4 Use momentum conservation to find the momentum of X^0. Then test the two hypotheses by calculating the invariant masses of the initial and final states in each case using the masses given in Appendix E. The decay is $\Lambda \rightarrow p\pi^-$.

A.5 For a free proton target, the energies and momenta of the initial state can be written $p(E_L, p_L)$ and $p(m, 0)$ for the beam and target protons, respectively. For a bound proton target the corresponding quantities are denoted $p(E'_L, p'_L)$ and $p(E, -p)$. Here, the momenta p_L and p'_L are in the beam direction and we have chosen the internal momentum of the bound proton to be in the opposite direction since this gives the maximum invariant mass for a given energy E'_L. The invariant mass W for the free proton case is then given by

$$W^2c^4 = (m_pc^2 + E_L)^2 - p_L^2c^2 = 2m_p^2c^4 + 2m_pc^2E_L$$

and the invariant mass W' for the bound proton case is given by

$$W'^2c^4 = (E'_L + E_L)^2 - (p'_L - p)^2c^2 = 2m_p^2c^4 + 2E_LE'_L + 2pp'_Lc^2.$$

Since the thresholds $E_L = E_{min}$ and $E'_L = E'_{min}$, $p'_L = p'_{min}$ correspond to the same invariant mass $W = W' = 4m_p$, we obtain

$$2m_pc^2E_{min} = 2EE'_{min} + 2pp'_{min}c^2$$

on equating the two expressions. For the nonrelativistic protons in the nucleus $E \approx m_pc^2$, while for the incident relativistic protons $p_{min} \approx E_{min}/c$, so that this relation becomes

$$2m_pc^2E_{min} = 2m_pc^2E'_{min} + 2pcE'_{min},$$

which gives the desired result.

A.6 (a) We require that

$$\left(mc^2 + \frac{p^2}{2m}\right) - mc^2\gamma < \varepsilon\frac{p^2}{2m}.$$

This can be expressed in terms of γ by using $p = m\upsilon\gamma$ and $\upsilon = c\left(1 - 1/\gamma^2\right)^{1/2}$, leading to the condition $\gamma < (1+\varepsilon)/(1-\varepsilon)$ and hence $\upsilon < 2c\sqrt{\varepsilon}/(1+\varepsilon)$. From $p = m\upsilon\gamma$ we have

$$p < \frac{2mc\sqrt{\varepsilon}}{1-\varepsilon} \approx 0.2\,mc$$

for $\varepsilon = 0.01$.

(b) We require that $mc^2\gamma - pc < \varepsilon pc$, which can be re-expressed in terms of υ to give $\upsilon > c/(1+\varepsilon)$, and hence

$$\gamma > \frac{1+\varepsilon}{\sqrt{[\varepsilon(2+\varepsilon)]}}.$$

Thus, for $\varepsilon = 0.01$,

$$p > \frac{mc}{\sqrt{[\varepsilon(2+\varepsilon)]}} \approx 7\,mc.$$

A.7 The argument is identical to that leading from (A.18) to (A.26), except that the transformation velocity υ that is required to bring the particle a to rest is given by

$$\upsilon = \frac{p_L c^2}{E_L} \approx c\left(1 - \frac{m_a^2 c^4}{2E_L^2}\right) \approx c.$$

PROBLEMS B

B.1 The values given for the volume of the target and the density of liquid hydrogen yield a target mass of 7.1×10^{-3} kg, and hence it contains 4.24×10^{24} protons. Since the beam is wide, we may assume that all of these protons are illuminated by the beam, so that from Equation (B.3)

$$W_{K\Lambda} = (10^7\,\mathrm{m^{-2}s^{-1}}) \times (4.24 \times 10^{24}) \times (0.4 \times 10^{-31}\,\mathrm{m^2}) = 1.70\,\mathrm{s^{-1}}.$$

B.2 (a) If the incident and final fluxes are J_i and J_f, respectively, then after traversing a thickness d of the target $J_f = J_i \exp(-d/l_c)$. The collision length l_c is given by $l_c = 1/n_t\sigma$, where n_t is the density of particles in the target and σ is the total cross-section:

$$\sigma = \sigma_{el} + \sigma_{inel} = 3.4\,\mathrm{b} = 3.4 \times 10^{-28}\,\mathrm{m^2}.$$

To find n_t we use $n_t = N_A\rho/A$, where N_A is Avogadro's number, ρ is the density of the target and A is its atomic weight in kg. Thus $n_t/\rho = 2.53 \times 10^{24}\,\mathrm{m^{-3}}$. Also, $d = 10^{-1}/\rho$ m, so that

$$\frac{d}{l_c} = \left(\frac{10^{-1}}{\rho}\,\mathrm{m}\right) \times (2.53 \times 10^{24}\,\rho\,\mathrm{m}^{-3}) \times (3.4 \times 10^{-28}\,\mathrm{m}^2) = 8.6 \times 10^{-5} \ll 1.$$

Thus the attenuation is $(J_j - J_i)/J_i = 8.6 \times 10^{-5}$.

(b) From Equation (B.3), the rate of inelastic reactions is $W_{inel} = J_i N \sigma_{inel}$, where $J_i = I_i/a$. and $N = n_t a d$. Here, I_i is the initial intensity of the beam and a is the area of the target illuminated by the beam, so that

$$W_{inel} = d n_t I_i \sigma_{inel} = 10^{-1} \frac{N_A}{A} I_i \sigma_{inel}.$$

Using the numerical values given, yields $W_{inel} = 51\,\mathrm{s}^{-1}$.

(c) Proceeding as in (b) we have

$$W_{el} = 10^{-1} I_i \frac{N_A}{A} \sigma_{el}.$$

Now at a distance r from the target, the flux is distributed over an area $4\pi r^2$. Thus, the average elastic flux at a distance $r$is

$$J_{el} = \frac{10^{-1}}{4\pi r^2} I_i \frac{N_A}{A} \sigma_{el},$$

which for $r = 5\,\mathrm{m}$, gives $J_{el} = 0.11\,\mathrm{m}^{-2}\mathrm{s}^{-1}$.

B.3 If we set $i = e^+e^-$, $f = X$, $E = E_{CM}$, $E_0 = M_Z$ and $\Gamma = \Gamma_Z$, then (9.26) may be written

$$\sigma_{if} = \frac{12\pi E_0^2}{E^2} \left[\frac{\Gamma_i \Gamma_f}{(E^2 - E_0^2)^2 + E_0^2 \Gamma^2}\right].$$

However, near the peak, $|E - E_0| \leq \Gamma$, so

$$(E^2 - E_0^2) = (E - E_0)\,(E + E_0) \approx 2E_0(E - E_0)$$

and

$$E = 2(q_i^2 + m_e^2)^{1/2} \approx 2q_i$$

since $E \gg m_e$, so that

$$\sigma_{ij} = \frac{3\pi}{4q_i^2} \left[\frac{\Gamma_i \Gamma_f}{(E - E_0)^2 + \Gamma^2/4}\right],$$

which is the same as (B.46b) since $S_1 = S_2 = \frac{1}{2}$ and $j = S_z = 1$.

PROBLEMS C

C.4 We need to identify the isospin of the final state. This is either $I = \frac{3}{2}$, $I_3 = -\frac{1}{2}$ or $I = \frac{1}{2}$, $I_3 = -\frac{1}{2}$. The former may be obtained from (C.33) by using the lowering operator $\hat{I}_-$ and is

$$\left|N\pi;\, \frac{3}{2}, -\frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}} \left|n\pi^0\right\rangle + \sqrt{\frac{1}{3}} \left|p\pi^-\right\rangle,$$

while the latter must be orthogonal to this and is therefore

$$\left|N\pi;\,\frac{1}{2},-\frac{1}{2}\right\rangle=\sqrt{\frac{1}{3}}\left|n\pi^0\right\rangle+\sqrt{\frac{2}{3}}\left|p\pi^-\right\rangle,$$

up to an irrelevant overall phase factor. The branching fractions for $X^0\to n\pi^0$ and $X^0\to p\pi^-$ are therefore in the ratio $2:1$ for $I=\frac{3}{2}$ and $1:2$ for $I=\frac{1}{2}$, since isospin is conserved in the strong interaction. From the data given, $I=\frac{1}{2}$. (The X^0 is actually the $N^0(1520)$ and is an excited state of the neutron.)

C.5 We need to identify the $\Sigma\pi$ state with $I=I_3=0$. Since $I_3=0$, this must be of the form

$$|\Sigma\pi;0,0\rangle=\alpha\,|\Sigma;\,1,\,1\rangle\,|\pi;\,1,-1\rangle+\beta\,|\Sigma;\,1,\,0\rangle\,|\pi;\,1,\,0\rangle+\gamma\,|\Sigma;\,1,\,-1\rangle\,|\pi;\,1,\,1\rangle\,,$$

and because it has $I=0$, it must satisfy

$$I_+\,|\Sigma\pi;\,0,\,0\rangle=I_-\,|\Sigma\pi;\,0,\,0\rangle=0$$

by (C.17). Either of these conditions is sufficient to determine β and γ in terms of α, and the resulting normalized state is

$$|\Sigma\pi;\,0,\,0\rangle=\sqrt{\frac{1}{3}}\left\{\left|\Sigma^+\pi^-\right\rangle-\left|\Sigma^0\pi^0\right\rangle-\left|\Sigma^-\pi^+\right\rangle\right\}$$

up to an irrelevant overall phase factor. Thus the branching ratios to each charged state are equal; i.e. they must each be 14 % to give a total $\Sigma\pi$ branching ratio of 42 % as stated.

PROBLEMS D

D.1 From (D.3) we obtain (recall that $\mathbf{r}=x_1,\,x_2,\,x_3$)

$$\dot{x}_i=\frac{\partial H}{\partial p_i}=\frac{(p_i-qA_i)}{m},\tag{1}$$

so that

$$p_i=m\dot{x}_i+qA_i(\mathbf{r},t)\tag{2}$$

and

$$\begin{aligned}\dot{p}_i&=-\frac{\partial H}{\partial x_i}=-q\frac{\partial\phi}{\partial x_i}+\frac{q}{m}\sum_j(p_j-qA_j)\frac{\partial A_j}{\partial x_i}\\&=-q\frac{\partial\phi}{\partial x_i}+q\sum_j\dot{x}_j\frac{\partial A_j}{\partial x_i}.\end{aligned}\tag{3}$$

However, from (2),

$$\dot{p}_i=m\ddot{x}_i+q\frac{\partial A_i}{\partial t}+q\sum_j\frac{\partial A_i}{\partial x_j}\dot{x}_j,$$

which combined with (3) and (D.4) gives

$$m\ddot{x}_i = qE_i + q\sum_j \dot{x}_j \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right).$$

This is identical with (D.5) as required if

$$(\mathbf{v}\times\mathbf{B})_i = (\dot{\mathbf{x}}\times\mathbf{B})_i = \sum_j \dot{x}_j \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right). \tag{4}$$

This is easily checked by inspecting components; e.g. in the x direction, using $\mathbf{B} = \operatorname{curl}\mathbf{A}$, we have

$$(\dot{\mathbf{x}}\times\mathbf{B})_1 = \dot{x}_2 B_3 - \dot{x}_3 B_2 = \dot{x}_2\left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right) - \dot{x}_3\left(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\right),$$

which is identical with the right-hand side of (4) for $i = 1$.

D.2 The first two Maxwell equations (D.9a) are automatically satisfied on substituting (D.4) since div curl $\mathbf{F} = 0$ (any $\mathbf{F}$) and curl grad $f = \mathbf{0}$ (any f). On substituting (D.4), div $\mathbf{E} = 0$ becomes

$$-\nabla^2\phi - \boldsymbol{\nabla}\cdot\frac{\partial\mathbf{A}}{\partial\tau} = 0,$$

which becomes

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi - \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A}\right) = 0, \tag{D.10a}$$

on adding and subtracting $\partial^2\phi/\partial t^2$. The remaining equation $\operatorname{curl}\mathbf{B} = \partial\mathbf{E}\big/\partial t$ reduces to (D.10b) using the identity

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{A} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \nabla^2\mathbf{A}.$$

D.3 Suppose

$$\frac{\partial\tilde{\phi}}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A} = -\alpha(\mathbf{r},t) \neq 0,$$

where $\alpha(\mathbf{r},t) \to 0$ as $|\mathbf{r}| \to \infty$ if the potentials vanish at infinity. Then

$$\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{A} = \frac{\partial\tilde{\phi}}{\partial t} + \boldsymbol{\nabla}\cdot\tilde{\mathbf{A}} + \left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)f = 0,$$

provided

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)f(\mathbf{r},t) = \alpha(\mathbf{r},t).$$

This is just a wave equation with a source $\alpha(\mathbf{r},t)$, and can be solved by standard methods.[1] In addition, if f_0 is a solution, then $f_0+\delta f$ is also a solution, provided that

$$\left(\frac{\partial^2}{\partial t^2}-\nabla^2\right)\delta f(\mathbf{r},t)=0$$

which is satisfied by any

$$\delta f=\int \mathrm{d}^3\mathbf{k}\, a(\mathbf{k})\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)],$$

where $\omega=|\mathbf{k}|$. In other words, there are an infinite number of gauge choices that satisfy the Lorentz condition.

D.4 Under a gauge transformation (D.7) and (D.8), $(\phi,\ \mathbf{A},\ \Psi)\rightarrow(\phi',\ \mathbf{A}',\ \Psi')$ and the terms in (D.18) become

$$\beta m\Psi'=\mathrm{e}^{-iqf}(\beta m\Psi),$$

$$i\left(\frac{\partial}{\partial t}+iq\phi'\right)\Psi'=i\left(\frac{\partial}{\partial t}+iq\phi+iq\frac{\partial f}{\partial t}\right)\mathrm{e}^{-iqf}\Psi=\mathrm{e}^{-iqf}\left[i\left(\frac{\partial}{\partial t}+iq\phi\right)\Psi\right]$$

and

$$-i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla}-iq\mathbf{A}')\Psi'=-i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla}-iq\mathbf{A}+iq\boldsymbol{\nabla}f)\,\mathrm{e}^{-iqf}\Psi=\mathrm{e}^{-iqf}\,[-i\boldsymbol{\alpha}\cdot(\boldsymbol{\nabla}-iq\mathbf{A})\Psi]\,.$$

Hence, if the Dirac equation (D.18) holds for $(\phi,\mathbf{A},\Psi)$, it also holds for $(\phi',\ \mathbf{A}',\ \Psi')$.

D.5 The transformation (D.8) can be rewritten using (D.64) as

$$\Psi\rightarrow\Psi'=\mathrm{e}^{-iqf}\Psi=\mathrm{e}^{-ieQf}\Psi=\exp(-ieI_3^Wf)\exp(-ieY^Wf)\Psi=\exp(-ieI_3^Wf)\Psi'',$$

where

$$\Psi''=\exp(-ieY^Wf)\Psi.$$

In other words, it can be written as a gauge transformation $\Psi\rightarrow\Psi''$ of the form (D.65) with $g'\omega(\mathbf{r},t)=ef(\mathbf{r},t)$, followed by a transformation $\Psi''\rightarrow\Psi'$ of the form (D.58) with $gf_3=ef$ and $f_i=0\,(i\neq3)$.

D.6 (a) Left-handed $(u,\ d)$ states have charged current interactions like (ν_e,e^-), so that $I_3(u,\ d)=(\frac{1}{2},-\frac{1}{2})$. Since $Q=I_3^W+Y^W$, this implies $Y^W=(\frac{1}{6},\frac{1}{6})$. Right-handed $(u,\ d)$ states have no charged current interactions, so that $I_3(u,\ d)=(0,0)$, implying $Y^W=(\frac{2}{3},\ -\frac{1}{3})$.

(b) For left-handed states, the difference between leptons (ν_e,e^-) and quarks (u,d) is that $Y^W=-\frac{1}{2}$ is replaced by $Y^W=\frac{1}{6}$, so that in (D.66), etc., $g'Y^W=-g'/2$ is replaced by $g'Y^W=g'/6$. This can be achieved by replacing $g'\rightarrow-g'/3$ in Section D.7.3. Hence, if we just retain the terms proportional to the electromagnetic field on the right-hand side of (D.71a), it becomes

[1] See, for example, Section 6.6 of Jackson (1975).

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,d=-\frac{1}{2}\left(g\sin\theta_W-\frac{g'}{3}\cos\theta_W\right)(\phi-\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})\psi(\mathbf{r},t)\,d+\cdots$$
$$=-\frac{1}{3}e(\phi-\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})\psi(\mathbf{r},t)\,d+\cdots$$

using the unification condition (D.72), while (D.71b) becomes

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,u=\frac{1}{2}\left(g\sin\theta_W+\frac{g'}{3}\cos\theta_W\right)(\phi-\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})\psi(\mathbf{r},t)\,u+\cdots$$
$$=\frac{2}{3}e(\phi-\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})\psi(\mathbf{r},t)\,u+\cdots$$

These are the appropriate equations for particles of charge $(u,\,d)=(\frac{2}{3},-\frac{1}{3})$ in units of e, as required. A similar argument establishes the same results for right-handed spin states.

D.7 Apply the arguments of the last question to the doublets $(u,\,d')$ and $(c,\,s')$, where d' and s' are given by (8.15). In this way, the analogues of (D.71a) (with $g'\to -g'/3$ as in Problem D.6) become

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,d'=2g_W(\phi^-\boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,u$$
$$-\frac{e}{3}(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})\psi(\mathbf{r},t)\,d'-g_n(\phi^Z-\boldsymbol{\alpha}\cdot\mathbf{Z})\psi(\mathbf{r},t)\,d'$$

and

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,s'=2g_W(\phi^-\boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,c$$
$$-\frac{e}{3}(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})\psi(\mathbf{r},t)\,s'-g_n(\phi^Z-\boldsymbol{\alpha}\cdot\mathbf{Z})\psi(\mathbf{r},t)s',$$

where we have used the unification condition (D.72) and introduced

$$g_n=\frac{1}{2}(g\cos\theta_W+\frac{1}{3}g'\sin\theta_W).$$

Using (8.15) these give

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,d=2g_W\cos\theta_C(\phi^-\boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,u$$
$$-2g_W\sin\theta_C(\phi^-\boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,c$$
$$-\frac{e}{3}(\phi-\boldsymbol{\alpha}\cdot\mathbf{A})\psi(\mathbf{r},t)\,d-g_n(\phi^Z-\boldsymbol{\alpha}\cdot\mathbf{Z})\psi(\mathbf{r},t)\,d$$

and

$$i\left(\frac{\partial}{\partial t}+\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r},t)\,s = 2g_W\sin\theta_C(\phi^- - \boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,u$$
$$+ 2g_W\cos\theta_C(\phi^- - \boldsymbol{\alpha}\cdot\mathbf{W}^-)\psi(\mathbf{r},t)\,c$$
$$-\frac{e}{3}(\phi - \boldsymbol{\alpha}\cdot\mathbf{A})\psi(\mathbf{r},t)\,s - g_n(\phi^Z - \boldsymbol{\alpha}\cdot\mathbf{Z})\psi(\mathbf{r},t)\,s.$$

We thus see that the weak coupling g_W in (D.71a) is replaced by (8.17) as appropriate, and the Z boson terms couple $d \leftrightarrow d$ and $s \leftrightarrow s$, but not $s \leftrightarrow d$.

References

K. Abe *et al*. (1982) *Physical Review Letters*, **48**, 1526.
K. Abe *et al*. (1995) *Physical Review Letters*, **74**, 2626.
M. Z. Akrawy *et al*. (1990) *Physics Letters B*, **240**, 497.
M. H. Alston *et al*. (1961) *Physical Review Letters*, **6**, 520.
C. D. Anderson (1933) *Physical Review*, **43**, 491.
E. N. da C. Andrade (1961) *Sir Isaac Newton*, Fontana Books.
Y. Ashie, J. Hosaka, Y. Itow *et al*. (2004) *Physical Review Letters*, **93**, 101801.
B. Aubert *et al.* (2006) *Physical Review D*, **73**, 012004.
V. Barnes *et al.* (1964) *Physical Review Letters*, **12**, 204.
H. J. Behrend *et al*. (1987) *Physics Letters B*, **183**, 400.
R. Brandelik *et al*. (1978) *Physics Letters B*, **76**, 361.
R. Brandelik *et al.* (1980) *Physics Letters*, *B*, **97**, 453.
M. Campbell (1989) *Proceedings of the 1989 International Symposium on Lepton and Photon Interactions at High Energies,* Stanford University.
J. H. Christenson *et al*. (1964) *Physical Review Letters*, **13**, 138.
F. E. Close (1997) *Contemporary Physics*, **38**, 1.
F. E. Close (2007) In *Electromagnetic Interactions and Hadron Structure*, eds F. E. Close, A. Donnachie and G. Shaw, Cambridge University Press, Section 1.9.
E. Fermi (1950) *Nuclear Physics*, University of Chicago Press.
R. Fernow (1986) *Introduction to Experimental Particle Physics*, Cambridge University Press.
S. Gasiorowicz (1974) *Quantum Mechanics,* John Wiley & Sons Ltd.
S. Gjesdal *et al.* (1974) *Physics Letters B*, **52**, 113.
K. Gottfried and V. F. Weisskopf (1986) *Concepts of Particle Physics,* vol. 2, Oxford University Press.
I. S. Grant and W. R. Phillips (1990) *Electromagnetism*, 2nd edition, John Wiley and Sons, Ltd.
C. Grupen (1996) *Particle Detectors*, Cambridge University Press.
F. Halzen and A. D. Martin (1984) *Quarks and Leptons*, John Wiley and Sons, Ltd.
I. S. Hughes (1985) *Elementary Particles*, 2nd edition, Cambridge University Press.
J. D. Jackson (1975) *Classical Electrodynamics*, 2nd edition, John Wiley and Sons, Inc., New York.
G. L. Kane (1987) *Modern Elementary Particle Physics*, Addison-Wesley.
O. Klein (1938) *Proceedings of the Symposium on Les Nouvelles Theories de la Physique,* Warsaw.
O. Klein (1948) *Nature*, **161**, 897.
K. Kleinknecht (1986) *Detectors for Particle Radiation*, Cambridge University Press.
D. Kleppner and R. J. Kolenkow (1973) *Introduction to Mechanics*, McGraw-Hill.
K. S. Krane (1988) *Introductory Nuclear Physics*, John Wiley and Sons, Ltd.

Z. Kunszt, S. Moretti and W. J. Stirling (1997) *Zeitschrift für Physik C*, **74**, 479.
C. M. G. Lattes *et al*. (1947) *Nature*, **159**, 694.
W. R. Leo (1994) *Techniques for Nuclear and Particle Physics,* Springer-Verlag.
F. Mandl, (1992) *Quantum Mechanics,* 2nd edition, John Wiley and Sons, Ltd.
F. Mandl and G. Shaw (1993) *Quantum Field Theory*, revised edition, John Wiley and Sons, Ltd, London.
B. R. Martin (2006) *Nuclear and Particle Physics*, John Wiley and Sons, Ltd.
R. Moore (1985) *Niels Bohr*, MIT Press.
L. Montanet *et al.* (1994) *Physical Rev. D*, **50**, 1173.
National Research Council (1999) *Nuclear Physics: The Core of Matter*, *The Fuel of Stars*, Report of the Board of Physics and Astronomy of the National Research Council, USA, National Academies Press, Washington, DC.
S. Orito (1979) *Proceedings of the 1979 International Symposium on Lepton and Photon Interactions at High Energies,* Fermilab.
W. Panofsky (1968) *International Conference on High Energy Physics*, Vienna.
W. Pauli (1933) *Handbuch der Physik*, **24**, 246.
D. H. Perkins (1987) *Introduction to High Energy Physics*, 3rd edition, Addison-Wesley.
D. H. Perkins (2003) *Particle Astrophysics*, Oxford University Press.
A. C. Phillips (1994) *The Physics of Stars*, John Wiley & Sons, Ltd, Chichester.
B. Povh, K. Rith, C. Scholz and F. Zetsche (1999) *Particles and Nuclei. An Introduction to the Physical Concepts*, revised 2nd edition, Springer.
G. D. Rochester and C. C. Butler (1947) *Nature*, **160**, 855.
B. Sadoulet (1983) *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies*, Cornell University, p. 27.
L. I. Schiff (1968) *Quantum Mechanics*, 3rd edition, McGraw-Hill.
R. H. Schindler *et al.* (1980) *Physical Review D*, **21**, 2716.
W. G. Seligman *et al.* (1997) *Physical Review Letters B*, **79**, 1213.
M. Spiro (1983) *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies,* Cornell University.
N. Spooner (1993) *Astronomy Now*, **7**, 40.
G. L. Trigg (1995) *Landmark Experiments in 20th Century Physics,* Crane, Russak and Co. Ltd, New York.
D. H. Weingarten (1996) *Scientific American*, February 1996, p. 104.
G. H. Wichmann (1967) *Quantum Physics*, McGraw-Hill.
S.-L. Wu (1984) *Physics Reports*, **107**, 59.
W.-M. Yao *et al*. (2006) *Journal of Physics*, **G33**,1.

Index

Particle Physics Third Edition B. R. Martin and G. Shaw

Plate 1 LHC tunnel and beam line. The blue cylinders contain dipole magnets and the liquid helium system required to cool them to superconducting temperatures. (Reproduced by permission of CERN.)

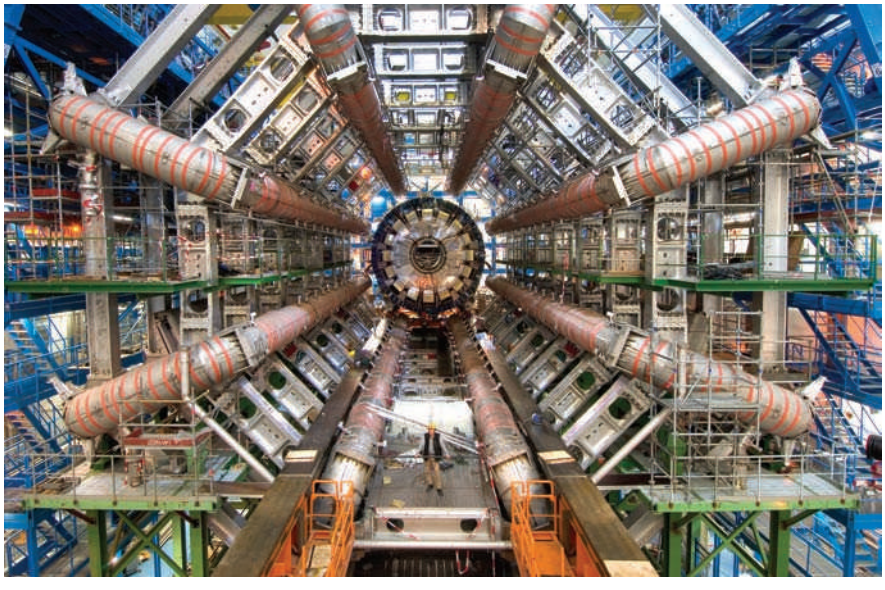


Plate 2 ATLAS detector: view along the beam direction during the construction phase showing the eight barrel toroids installed, with a calorimeter at the end before it is moved into the middle of the detector. In the completed detector (see Figure 4.26), the central cavern is filled with subdetectors. (Reproduced by permission of CERN.)

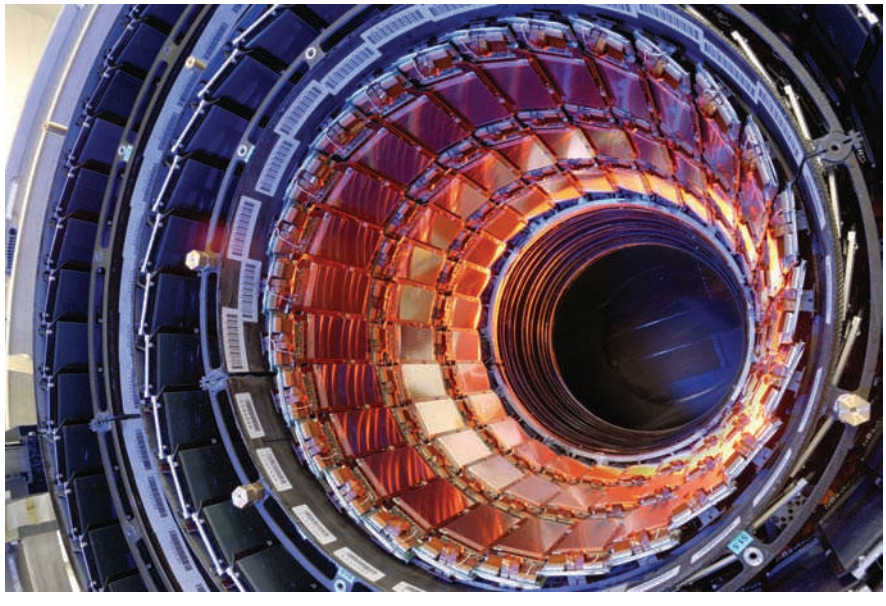


Plate 3 The first half of the CMS inner tracker barrel, consisting of three layers of silicon modules, is installed at the centre of the experiment, close to the interaction point of the collisions. (Reproduced by permission of CERN.)

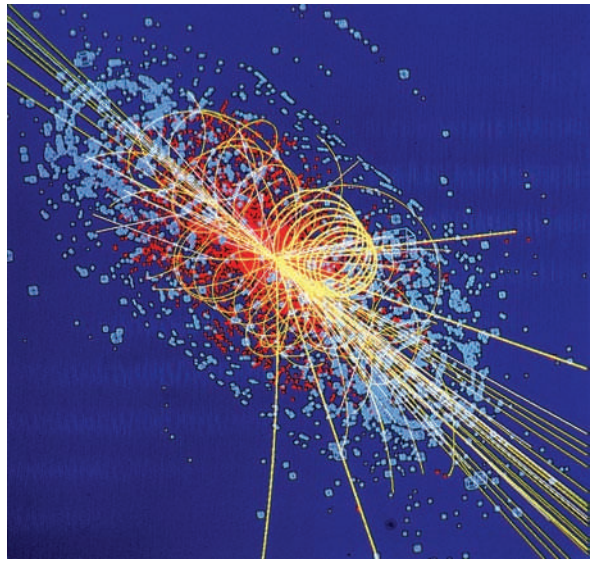


Plate 4 Computer simulation of a Higgs boson produced in the CMS detector from the collision of two protons at 14 TeV. The tracks of the collision products are shown by lines and the energy deposited in the detector is shown in light blue. (Reproduced by permission of CERN.)

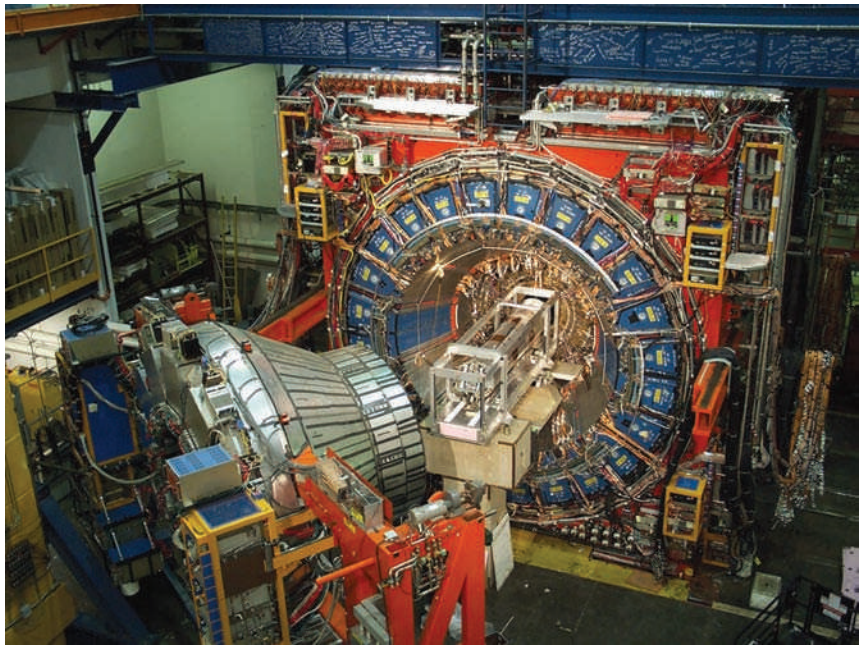


Plate 5 Installation of the silicon vertex detector in the CDF detector (cf. Figure 4.24). (Courtesy of Fermilab and Brookhaven National Laboratory.)

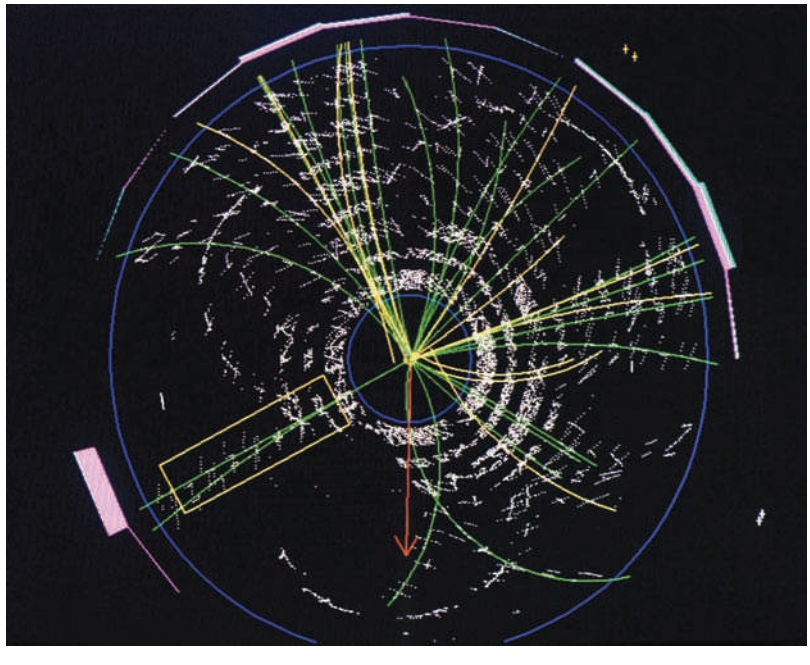


Plate 6 A polar plot of a $t\bar{t}$ event in the CDF detector. The yellow and green lines represent charged particles reconstructed in the central drift chamber. The event consists of four jets, two of them from b quarks, a positron (outlined by a yellow rectangle) and a neutrino (direction denoted by a red arrow). (Courtesy of Fermilab and Brookhaven National Laboratory.)

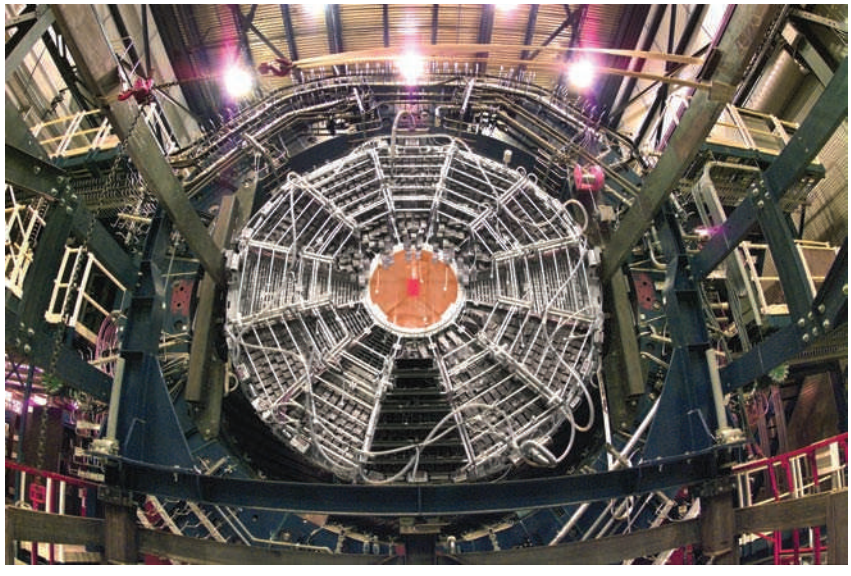


Plate 7 The STAR detector, showing an end cap of the time projection chamber during construction (cf. Figure 4.25). (Courtesy of Fermilab and Brookhaven National Laboratory.)

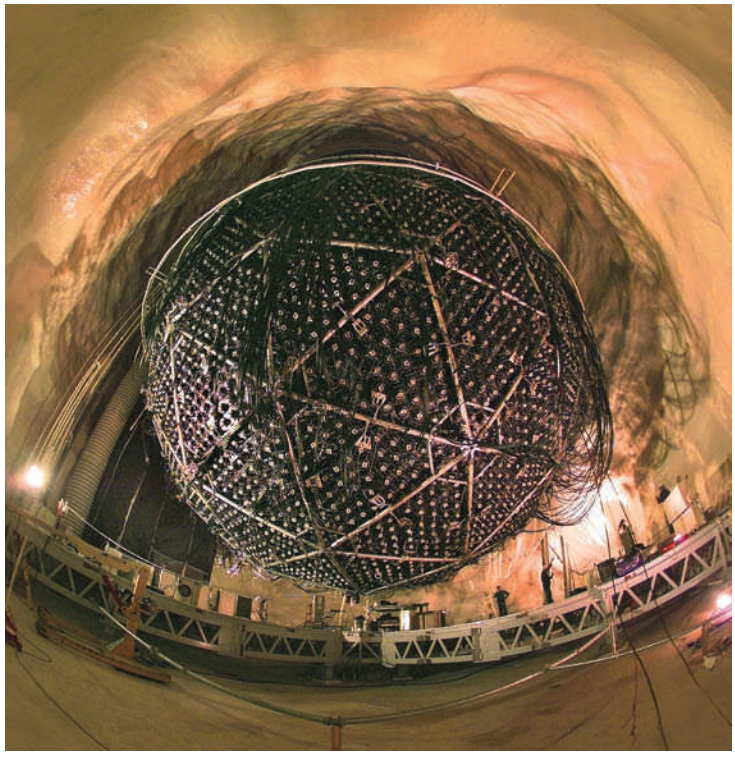


Plate 8 The Sudbury neutrino detector during construction, showing the large number of photomultiplier tubes surrounding the sphere. (Courtesy of Lawrence Berkeley National Laboratory – Roy Kaltschmidt, photographer.)

PHYSICAL CONSTANTS

Quantity	*Symbol*	*Value*
Speed of light in vacuum	c	$2.998 \times 10^{8}\,\mathrm{ms}^{-1}$
Planck's constant	h	$4.136 \times 10^{-24}\,\mathrm{GeV\,s}$
Planck's constant reduced	$\hbar \equiv h/2\pi$	$6.582 \times 10^{-25}\,\mathrm{GeV\,s}$
	$\hbar c$	$1.973 \times 10^{-16}\,\mathrm{GeV\,m}$
	$(\hbar c)^2$	$3.894 \times 10^{-32}\,\mathrm{GeV^2\,m^2}$
Electron charge (magnitude)	e	$1.602 \times 10^{-19}\,\mathrm{C}$
Fine structure constant	$\alpha \equiv e^2/4\pi\varepsilon_0\hbar c$	$7.297 \times 10^{-3} = 1/137.04$
Electron mass	m_e	$0.511\,\mathrm{MeV/c^2}$
Proton mass	m_p	$0.9383\,\mathrm{GeV/c^2}$
W boson mass	M_W	$80.40\,\mathrm{GeV/c^2}$
Z boson mass	M_Z	$91.19\,\mathrm{GeV/c^2}$
Fermi coupling constant	$G_F/(\hbar c)^3$	$1.166 \times 10^{-5}\,\mathrm{GeV^{-2}}$
Strong coupling constant	$\alpha_s(M_Z c)$	0.118

CONVERSION FACTORS

$1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}$ $\qquad$ $1\,\mathrm{eV/c^2} = 1.783 \times 10^{-36}\,\mathrm{kg}$

$1\ \text{fermi} = 1\ \text{fm} \equiv 10^{-15}\,\mathrm{m}$ $\qquad$ $1\ \text{barn} = 1\ \text{b} \equiv 10^{-28}\,\mathrm{m^2}$

NATURAL UNITS

Throughout the book we have used natural units, $\hbar = c = 1$. They can be converted back to mks units by the method discussed in Section 1.5. The following conversion factors are often useful.

Time	$1\,\mathrm{GeV^{-1}} = 6.58 \times 10^{-25}\,\mathrm{s}$
Length	$1\,\mathrm{GeV^{-1}} = 0.197\,\mathrm{fm}$
Cross-section	$1\,\mathrm{GeV^{-2}} = 0.389\,\mathrm{mb}$