2 Underlying concepts

Much of particle physics is concerned with the high-energy interactions of relativistic particles. Therefore the calculation of interaction and decay rates requires a relativistic formulation of quantum mechanics. Relativistic quantum mechanics (RQM) is founded on the two pillars of "modern" physics, Einstein's theory of special relativity and the wave mechanics developed in the early part of the twentieth century. It is assumed that you are already familiar with special relativity and non-relativistic quantum mechanics. The purpose of this chapter is to review the specific aspects of special relativity and quantum mechanics used in the subsequent development of relativistic quantum mechanics. Before discussing these important topics, the system of units commonly used in particle physics is introduced.

2.1 Units in particle physics

The system of S.I. units [kg, m, s] forms a natural basis for the measurements of mass, length and time for everyday objects and macroscopic phenomena. However, it is not a natural choice for the description of the properties of particles, where we are almost always dealing with very small quantities, such as the mass of the electron, which in S.I. units is 9.1×10^{-31} kg. One way to avoid carrying around large exponents is to use S.I. based units. For example, interaction cross sections (which have the dimension of area) are usually quoted in *barns*, where

$$1\,\text{barn} \equiv 10^{-28}\,\text{m}^2.$$

The cross sections for the more interesting physical processes at the highest energies are typically in the picobarn (pb) to femtobarn (fb) range, where $1\,\text{pb} = 10^{-12}$ barn and $1\,\text{fb} = 10^{-15}$ barn. The use of derived S.I. units solves the problem of large exponents, nevertheless, it is more convenient to work with a system of units that from the outset reflects the natural length and time scales encountered in particle physics.

2.1.1 Natural units

The system of units used in particle physics is known as natural units. It is based on the fundamental constants of quantum mechanics and special relativity. In natural units, [kg, m, s] are replaced by [$\hbar$, c, GeV], where $\hbar = 1.055 \times 10^{-34}$ J s is the unit of action in quantum mechanics, $c = 2.998 \times 10^{8}$ m s^{-1} is the speed of light in vacuum, and $1\,\text{GeV} = 10^{9}\,\text{eV} = 1.602 \times 10^{-10}$ J, which is very approximately the rest mass energy of the proton. Table 2.1 lists the units used for a number of commonly encountered quantities expressed in terms of both [kg, m, s] and [$\hbar$, c, GeV], where the conversion can be obtained from dimensional analysis.

Natural units provide a well-motived basis for expressing quantities in particle physics and can be simplified by *choosing*

$$\hbar = c = 1.$$

In this way, all quantities are expressed in powers of GeV, as shown in the rightmost column of Table 2.1. Setting $\hbar = c = 1$ has the advantage of simplifying algebraic expressions as there is no longer the need to carry around (possibly large) powers of $\hbar$ and c. For example, the Einstein energy–momentum relation

$$E^2 = p^2c^2 + m^2c^4 \quad \text{becomes} \quad E^2 = p^2 + m^2.$$

At first sight it might appear that information has been lost in setting $\hbar = c = 1$. However, the factors of $\hbar$ and c have not simply vanished; they are still present in the dimensions of quantities. The conversion back to S.I. units is simply a question of reinserting the necessary missing factors of $\hbar$ and c, which can be identified from dimensional analysis. For example, the result of a calculation using natural units might determine the root-mean-square charge radius of the proton to be

$$\langle r^2 \rangle^{1/2} = 4.1\,\text{GeV}^{-1}.$$

Table 2.1 Relationship between S.I. and natural units.

Quantity	[kg, m, s]	[$\hbar$, c, GeV]	$\hbar = c = 1$
Energy	kg m^2 s^{-2}	GeV	GeV
Momentum	kg m s^{-1}	GeV/c	GeV
Mass	kg	GeV/c^2	GeV
Time	s	(GeV/$\hbar$)$^{-1}$	GeV^{-1}
Length	m	(GeV/$\hbar c$)$^{-1}$	GeV^{-1}
Area	m^2	(GeV/$\hbar c$)$^{-2}$	GeV^{-2}

To convert this to back into S.I. units the correct dimensions are obtained by multiplying by $\hbar c$, giving

$$\langle r^2 \rangle^{1/2} = 4.1 \times \frac{1.055 \times 10^{-34} \times 2.998 \times 10^{8}}{1.602 \times 10^{-10}} \text{ m}$$
$$= 4.1 \times (0.197 \times 10^{-15}) \text{ m} = 0.8 \times 10^{-15} \text{ m}.$$

In converting from natural units to S.I. units, it is useful to remember the conversion factor

$$\hbar c = 0.197 \text{ GeV fm},$$

where one femtometre (fm) = 10^{-15} m.

Heaviside–Lorentz units

The equations of classical electromagnetism can be simplified by adopting Heaviside–Lorentz units. The value of the electron charge is defined by the magnitude of the Coulomb force between two electrons separated by a distance r,

$$\text{F} = \frac{e^2}{4\pi\varepsilon_0 \text{r}^2},$$

where ε_0 is the permittivity of free space. In Heaviside–Lorentz units ε_0 is set to unity, and the expression for the Coulomb force becomes

$$\text{F} = \frac{e^2}{4\pi \text{r}^2}.$$

Effectively ε_0 has been absorbed into the definition of the electron charge. Because $1/(\varepsilon_0\mu_0) = c^2$, choosing $\varepsilon_0 = 1$ and $c = 1$ implies that the permeability of free space $\mu_0 = 1$. Hence, in the combined system of natural units and Heaviside–Lorentz units used in particle physics,

$$\hbar = c = \varepsilon_0 = \mu_0 = 1.$$

With $c = \varepsilon = \mu_0 = 1$, Maxwell's equations take the same form as with S.I. units.

The strength of the QED interaction is defined in terms of the dimensionless fine structure constant,

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c}. \tag{2.1}$$

Since α is dimensionless, it has the same numerical value regardless of the system of units used,

$$\alpha \approx \frac{1}{137}.$$

In natural units, the relationship between α and the electron charge (which is not dimensionless) is simply

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}.$$

2.2 Special relativity

This section gives a brief overview of the basic concepts of special relativity, with the emphasis on the definition and application of four-vectors and the concept of Lorentz invariance and Lorentz invariant quantities.

2.2.1 The Lorentz transformation

Special relativity is based on the space-time transformation properties of physical observables as measured in two or more inertial frames moving relative to each other. For example, Figure 2.1 shows a space-time event that occurs at $(t, \mathbf{r})$ in the inertial frame Σ and at $(t', \mathbf{r}')$ in the inertial frame Σ' that is moving with a velocity v in the z-direction relative to the frame Σ. For the case where $\mathrm{v} \ll c$ and the origins of two inertial frames coincide at $t = t' = 0$, the two sets of coordinates are related by the Galilean transformation

$$t' = t, \quad x' = x, \quad y' = y \quad \text{and} \quad z' = z - \mathrm{v}t.$$

Einstein postulated that the speed of light in the vacuum is the same in all inertial frames. This primary postulate of special relativity implies that a space-time point on the wavefront of a pulse of light emitted at $t = t' = 0$ satisfies both $x^2 + y^2 + z^2 = c^2t^2$ and $x'^2 + y'^2 + z'^2 = c^2t'^2$. Consequently the space-time interval,

$$c^2t^2 - x^2 - y^2 - z^2 = c^2t'^2 - x'^2 - y'^2 - z'^2, \tag{2.2}$$

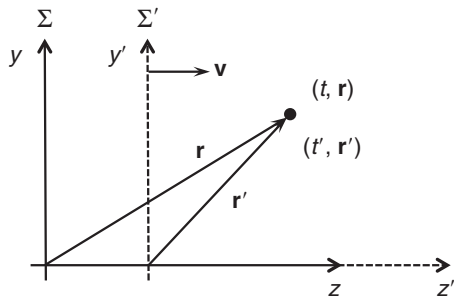

Fig. 2.1 A space-time event as seen into two inertial frames. The frame Σ' moves with a velocity $\mathbf{v}$ in the z direction relative to frame Σ.

is an invariant quantity; it is observed to be the same in all reference frames. Equation (2.2) is satisfied if the coordinates in Σ and Σ' are related by the Lorentz transformation

$$t' = \gamma\left(t - \frac{\mathrm{v}}{c^2}z\right), \quad x' = x, \quad y' = y \quad \text{and} \quad z' = \gamma(z - \mathrm{v}t),$$

where the Lorentz factor γ is given by

$$\gamma = (1 - \beta^2)^{-\frac{1}{2}},$$

and $\beta = \mathrm{v}/c$. In the low velocity limit $\mathrm{v} \ll c$, the Lorentz factor reduces to unity and the Galilean transformations are recovered. In natural units, where $c = 1$, the Lorentz transformation of the space-time coordinates becomes

$$t' = \gamma\,(t - \beta z)\,, \quad x' = x, \quad y' = y \quad \text{and} \quad z' = \gamma(z - \beta t). \tag{2.3}$$

This can be written in matrix form as $\mathbf{X}' = \mathbf{\Lambda X}$,

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \tag{2.4}$$

where $\mathbf{X}$ is the four-component vector $\{t, \mathbf{x}\}$. The inverse Lorentz transformation, from Σ' to Σ, is obtained by reversing the sign of the velocity in (2.3) such that

$$t = \gamma\,(t' + \beta z')\,, \quad x = x', \quad y = y' \quad \text{and} \quad z = \gamma(z' + \beta t'). \tag{2.5}$$

In matrix form this can be written $\mathbf{X} = \mathbf{\Lambda^{-1}X'}$,

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & +\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}. \tag{2.6}$$

It is straightforward to confirm that the matrices appearing in (2.4) and (2.6) are the inverse of each other, $\mathbf{\Lambda\Lambda^{-1}} = \mathbf{I}$. The matrix equations of (2.4) and (2.6) define the Lorentz transformation between the space-time coordinates measured in two inertial frames with relative motion in the z-direction.

2.2.2 Four-vectors and Lorentz invariance

Throughout particle physics it is highly desirable to express physical predictions, such as interaction cross sections and decay rates, in an explicitly Lorentz-invariant form that can be applied directly in all inertial frames. Although the Lorentz transformation forms the basis of special relativity, Lorentz invariance is the more important concept for much that follows. Lorentz invariance is best expressed in terms of four-vectors. A *contravariant* four-vector is defined to be a set of quantities that

when measured in two inertial frames are related by the Lorentz transformation of (2.4). For example, the contravariant four-vector x^μ is defined as

$$x^\mu = (t, x, y, z),$$

where the indices $\mu = \{0, 1, 2, 3\}$ label the space-time coordinates with the zeroth component representing time. In tensor form, the Lorentz transformation of (2.4) now can be expressed as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{2.7}$$

where $\Lambda^\mu{}_\nu$ can be thought of as the elements of the matrix $\mathbf{\Lambda}$ and Einstein's summation convention for repeated indices is used to express the matrix multiplication.

The magnitude of a normal three-vector, which is given by the three-vector scalar product $\mathbf{x} \cdot \mathbf{x}$, is invariant under rotations. The Lorentz invariance of the space-time interval, $t^2 - x^2 - y^2 - z^2$, can be expressed as a four-vector scalar product by defining the *covariant* space-time four-vector,

$$x_\mu = (t, -x, -y, -z).$$

With this notation, the Lorentz-invariant space-time interval can be written as the four-vector scalar product

$$x^\mu x_\mu = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3 = t^2 - x^2 - y^2 - z^2.$$

The main reason for introducing covariant four-vectors, which are denoted with a "downstairs" index to distinguish them from the corresponding contravariant four-vectors, is to keep account of the minus signs in Lorentz-invariant products. The Lorentz transformation of the space-time coordinates (2.3) can be written in terms of the components of the covariant four-vector as

$$\begin{pmatrix} t' \\ -x' \\ -y' \\ -z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & +\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} t \\ -x \\ -y \\ -z \end{pmatrix}. \tag{2.8}$$

The sign changes in this matrix relative to that of (2.4) compensate for the changes of sign in the definition of x_μ relative to x^μ. Both (2.8) and (2.4) are *equivalent* expressions of the same Lorentz transformation originally defined in (2.3). The transformation matrix appearing in (2.8) is the inverse of that of (2.4). To make this distinction explicit in tensor notation the transformation of a covariant four-vector is written as

$$x'_\mu = \Lambda_\mu{}^\nu x_\nu, \tag{2.9}$$

where the downstairs index appears first in $\Lambda_\mu{}^\nu$ which represents the elements of $\mathbf{\Lambda^{-1}}$.

In tensor notation, the relationship between covariant and contravariant four-vectors in special relativity can be expressed as

$$x_\mu = g_{\mu\nu} x^\nu,$$

where summation over repeated indices is again implicit and the diagonal *metric tensor* $g_{\mu\nu}$ is defined as

$$g_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.10}$$

By definition, only quantities with the Lorentz transformation properties of (2.4) are written as contravariant four-vectors. For such a set of quantities a^μ the scalar product with the corresponding covariant four-vector a_μ is guaranteed to be Lorentz invariant. Furthermore, if a^μ and b^μ are both (contravariant) four-vectors, then the scalar product

$$a^\mu b_\mu = a_\mu b^\mu = g_{\mu\nu} a^\mu b^\nu,$$

is automatically Lorentz invariant. Again this follows directly from the form of the Lorentz transformation for contravariant and covariant four-vectors. Hence *any* expression that can be written in terms of four-vector scalar products is guaranteed to be Lorentz invariant. From the linearity of the Lorentz transformation, it also follows that the sum of any number of contravariant four-vectors also transforms according to (2.4) and therefore is itself a four-vector.

Four-momentum

The relativistic expressions for the energy and momentum of a particle of mass m can be identified as $E = \gamma mc^2$ and $\mathbf{p} = \gamma m\mathbf{v}$, which when expressed in natural units are

$$E = \gamma m \quad \text{and} \quad \mathbf{p} = \gamma m\boldsymbol{\beta}. \tag{2.11}$$

By considering the transformation properties of velocity, $\mathrm{d}\mathbf{x}/\mathrm{d}t$, it can be shown that relativistic energy and momentum, defined in this way, transform according to (2.4) and therefore form a contravariant four-vector,

$$p^\mu = (E, p_x, p_y, p_z),$$

referred to as four-momentum. Because momentum and energy are separately conserved, four-momentum is also conserved. Furthermore, since four-momentum is a four-vector, the scalar product

$$p^\mu p_\mu = E^2 - \mathbf{p}^2,$$

is a Lorentz-invariant quantity.

From (2.11) it can be seen that a single particle at rest has four-momentum $p^\mu = (m, 0, 0, 0)$ and therefore $p^\mu p_\mu = m^2$. Since $p^\mu p_\mu$ is Lorentz invariant, the relation

$$E^2 - \mathbf{p}^2 = m^2$$

holds in *all* inertial frames. This is, of course, just the Einstein energy–momentum relationship. For a system of n particles, the total energy and momentum

$$p^\mu = \sum_{i=1}^{n} p_i^\mu$$

is also a four-vector. Therefore for a system of particles the quantity

$$p^\mu p_\mu = \left(\sum_{i=1}^{n} E_i\right)^2 - \left(\sum_{i=1}^{n} \mathbf{p}_i\right)^2$$

is a Lorentz-invariant quantity, which gives the squared *invariant mass* of the system. In a particle decay $a \to 1+2$, the invariant mass of the decay products is equal to the mass of the decaying particle,

$$(p_1 + p_2)^\mu (p_1 + p_2)_\mu = p_a^\mu p_{a\mu} = m_a^2.$$

Four-derivative

The transformation properties of the space-time derivatives can be found by using the Lorentz transformation of (2.3) to express the coordinates of an event in the frame Σ' as functions of the coordinates measured in the frame Σ, for example $z'(t, x, y, z)$ and $t'(t, x, y, z)$. Hence, for a Lorentz transformation in the z-direction, the derivatives in the primed-frame can be expressed as

$$\frac{\partial}{\partial z'} = \left(\frac{\partial z}{\partial z'}\right)\frac{\partial}{\partial z} + \left(\frac{\partial t}{\partial z'}\right)\frac{\partial}{\partial t} \quad \text{and} \quad \frac{\partial}{\partial t'} = \left(\frac{\partial z}{\partial t'}\right)\frac{\partial}{\partial z} + \left(\frac{\partial t}{\partial t'}\right)\frac{\partial}{\partial t}.$$

From (2.5), the relevant partial derivatives are

$$\left(\frac{\partial z}{\partial z'}\right) = \gamma, \quad \left(\frac{\partial t}{\partial z'}\right) = +\gamma\beta, \quad \left(\frac{\partial z}{\partial t'}\right) = +\gamma\beta \quad \text{and} \quad \left(\frac{\partial t}{\partial t'}\right) = \gamma,$$

and therefore,

$$\frac{\partial}{\partial z'} = \gamma\frac{\partial}{\partial z} + \gamma\beta\frac{\partial}{\partial t} \quad \text{and} \quad \frac{\partial}{\partial t'} = \gamma\beta\frac{\partial}{\partial z} + \gamma\frac{\partial}{\partial t}. \tag{2.12}$$

From (2.12) it can be seen that Lorentz transformation properties of the partial derivatives are

$$\begin{pmatrix} \partial/\partial t' \\ \partial/\partial x' \\ \partial/\partial y' \\ \partial/\partial z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & +\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} \partial/\partial t \\ \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix},$$

and comparison with (2.8) shows that

$$\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

transforms as a *covariant* four-vector, which is written as

$$\partial_\mu = \frac{\partial}{\partial x^\mu},$$

and has components

$$\partial_0 = \frac{\partial}{\partial t}, \quad \partial_1 = +\frac{\partial}{\partial x}, \quad \partial_2 = +\frac{\partial}{\partial y} \quad \text{and} \quad \partial_3 = +\frac{\partial}{\partial z}.$$

The corresponding contravariant four-derivative is therefore

$$\partial^\mu = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right),$$

and it should be noted that here the space-like coordinates enter with minus signs. The equivalent of the Laplacian for the four-derivative, which is known as the d'Alembertian, is therefore

$$\Box = \partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}.$$

In this book the symbol $\Box$ is used to represent the d'Alembertian, in some textbooks you may see it written as $\Box^2$.

Vector and four-vector notation

This is a convenient place to introduce the notation used in this book. Unless otherwise stated, quantities written simply as x and p always should be interpreted as four-vectors. Three-vectors, such as the three-momentum of a particle, are always written in boldface, for example $\mathbf{p}$, with three-vector scalar products written as

$$\mathbf{p}_1 \cdot \mathbf{p}_2.$$

The magnitude of a three-vector is written either as $|\mathbf{p}|$ or simply p. Four-vector scalar products are written either as $a^\mu b_\mu$ or $a \cdot b$, with

$$a \cdot b \equiv a^\mu b_\mu \equiv g_{\mu\nu} a^\mu b^\nu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3.$$

Just as $\mathbf{p}^2$ is shorthand for $\mathbf{p} \cdot \mathbf{p}$, then for a four-vector a, the expression a^2 is shorthand for the four-vector scalar product $a \cdot a$. For example, the Einstein energy–momentum relationship for a single particle can be expressed as $p^2 = m^2$, since $p^2 = p \cdot p = E^2 - \mathbf{p}^2$. Finally, it will sometimes be convenient to work with quantities measured in the centre-of-mass frame of a system of particles, and such quantities are denoted by a star. For example, p^* is the magnitude of the three-momentum of

a particle evaluated in the centre-of-mass frame, which for a system of particles is the inertial frame in which there is no net three-momentum.

2.2.3 Mandelstam variables

Feynman diagrams, involving the exchange of a single force mediating particle, can be placed in the three categories shown in Figure 2.2. The first two diagrams represent the s-channel annihilation process and the t-channel scattering process. The third diagram represents u-channel scattering and is only relevant when there are identical particles in the final state. In Chapter 5 it will be shown that four-momentum is conserved at each vertex in a Feynman diagram. In a process involving two initial-state and two final-state particles, the Mandelstam variables

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2, \\ t &= (p_1 - p_3)^2 = (p_2 - p_4)^2, \\ u &= (p_1 - p_4)^2 = (p_2 - p_3)^2, \end{aligned}$$

are equivalent to the four-momentum squared q^2 of the exchanged boson in the respective class of diagram. For identical final-state particles the distinction between u- and t-channel diagrams is necessary because the final-state particle with four-momentum p_3 can originate from either interaction vertex, and the four-momentum q of the virtual particle is different for the two cases.

Since the Mandelstam variables are four-vector scalar products, they are manifestly Lorentz invariant and can be evaluated in any frame. For example, in the centre-of-mass frame where there is no net momentum, the four-momenta of two colliding particles are $p_1 = (E_1^*, \mathbf{p}^*)$ and $p_2 = (E_2^*, -\mathbf{p}^*)$, from which

$$s = (p_1 + p_2)^2 = (E_1^* + E_2^*)^2 - (\mathbf{p}^* - \mathbf{p}^*)^2 = (E_1^* + E_2^*)^2. \tag{2.13}$$

Hence, the Lorentz-invariant quantity $\sqrt{s}$ can be identified as the total energy available in the centre-of-mass frame. It is worth noting that for the process

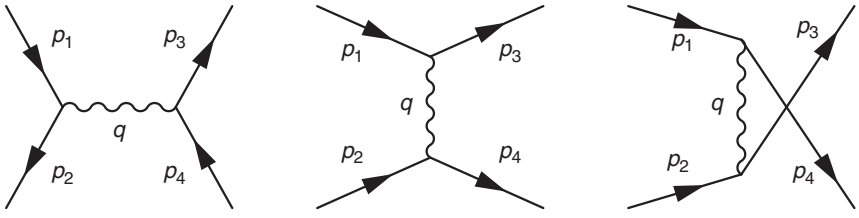

Fig. 2.2 The Feynman diagrams for s-channel, t-channel and u-channel processes. The u-channel diagram applies only when there are identical particles in the final state.

$1+2\to 3+4$, the sum of $s+u+t$ can be shown to be equal to the sum of the squares of the masses of the four particles (see Problem 2.12),

$$s+u+t=m_1^2+m_2^2+m_3^2+m_4^2. \tag{2.14}$$

2.3 Non-relativistic quantum mechanics

This section gives a brief overview of topics in non-relativistic quantum mechanics which are of direct relevance to the development of the relativistic treatment of spin-half particles in Chapter 4. It also reviews of the algebraic treatment of angular momentum that serves as an introduction to the algebra of the SU(2) symmetry group.

2.3.1 Wave mechanics and the Schrödinger equation

In quantum mechanics it is postulated that free particles are described by wave packets which can be decomposed into a Fourier integral of plane waves of the form

$$\psi(\mathbf{x},t)\propto\exp\{i(\mathbf{k}\cdot\mathbf{x}-\omega t)\}. \tag{2.15}$$

Following the de Broglie hypothesis for wave–particle duality, the wavelength of a particle in quantum mechanics can be related to its momentum by $\lambda=h/\mathrm{p}$, or equivalently, the wave vector $\mathbf{k}$ is given by $\mathbf{k}=\mathbf{p}/\hbar$. The angular frequency of the plane wave describing a particle is given by the Planck–Einstein postulate, $E=\hbar\omega$. In natural units with $\hbar=1$, the de Broglie hypothesis and Planck–Einstein postulate imply $\mathbf{k}=\mathbf{p}$ and $\omega=E$, and thus the plane wave of (2.15) becomes

$$\psi(\mathbf{x},t)=N\exp\{i(\mathbf{p}\cdot\mathbf{x}-Et)\}, \tag{2.16}$$

where N is the normalisation constant.

In classical physics, the energy and momentum of a particle are dynamical variables represented by time-dependent real numbers. In the Schrödinger picture of quantum mechanics, the wavefunction is postulated to contain all the information about a particular state. Dynamical variables of a quantum state, such as the energy and momentum, are obtained from the wavefunction. Consequently, in the Schrödinger picture of quantum mechanics, the time-dependent variables of classical dynamics are replaced by time-independent operators acting on the time-dependent wavefunction. Because the wavefunction is postulated to contain all the information about a system, a physical observable quantity A corresponds to the action of a quantum mechanical *operator* $\hat{A}$ on the wavefunction. A further

postulate of quantum mechanics is that the result of the measurement of the observable A will be one of the eigenvalues of the operator equation

$$\hat{A}\psi = a\psi.$$

For A to correspond to a physical observable, the eigenvalues of the corresponding operator must be real, which implies that the operator is Hermitian. This is formally defined by the requirement

$$\int \psi_1^* \hat{A}\psi_2 \, \mathrm{d}\tau = \int \left[\hat{A}\psi_1\right]^* \psi_2 \, \mathrm{d}\tau.$$

Because the plane wave of (2.16) is intended to represent a free particle with energy E and momentum $\mathbf{p}$, it is reasonable to identify the momentum and energy operators, $\hat{\mathbf{p}}$ and $\hat{E}$, as

$$\hat{\mathbf{p}} = -i\boldsymbol{\nabla} \quad \text{and} \quad \hat{E} = i\frac{\partial}{\partial t}, \tag{2.17}$$

such that $\hat{\mathbf{p}}$ and $\hat{E}$ acting on the plane wave of (2.16) give the required eigenvalues,

$$\hat{\mathbf{p}}\psi = -i\boldsymbol{\nabla}\psi = \mathbf{p}\psi,$$
$$\hat{E}\psi = i\frac{\partial \psi}{\partial t} = E\psi.$$

In classical dynamics, the total energy of a non-relativistic particle can be expressed as the sum of its kinetic and potential energy terms,

$$E = H = T + V = \frac{\mathbf{p}^2}{2m} + V,$$

where $H = T + V$ is the Hamiltonian. The equivalent quantum mechanical expression is obtained by replacing each of the terms with the corresponding operators defined in (2.17) acting on the wavefunction. This gives rise to the time-dependent Schrödinger equation,

$$i\frac{\partial \psi(\mathbf{x}, t)}{\partial t} = \hat{H}\psi(\mathbf{x}, t), \tag{2.18}$$

where, for a non-relativistic particle, the Hamiltonian operator is

$$\hat{H}_{NR} = \frac{\hat{\mathbf{p}}^2}{2m} + \hat{V} = -\frac{1}{2m}\nabla^2 + \hat{V}. \tag{2.19}$$

For a one-dimensional system (2.18) and (2.19) reduce to the familiar one-dimensional time-dependent Schrödinger equation,

$$i\frac{\partial \psi(\mathbf{x}, t)}{\partial t} = -\frac{1}{2m}\frac{\partial^2 \psi(\mathbf{x}, t)}{\partial x^2} + \hat{V}\psi(\mathbf{x}, t).$$

2.3.2 Probability density and probability current

The physical interpretation of the wavefunction $\psi(\mathbf{x}, t)$ is that $\psi^*\psi\, \mathrm{d}^3\mathbf{x}$ is the probability of finding the particle represented by the wavefunction in the volume element $\mathrm{d}^3\mathbf{x}$. This is equivalent to identifying the probability density $\rho(\mathbf{x}, t)$ as

$$\rho(\mathbf{x}, t) = \psi^*(\mathbf{x}, t)\, \psi(\mathbf{x}, t).$$

Assuming the particle does not decay or interact, its associated total probability will be constant. This conservation of probability can be expressed in terms of a continuity equation by defining the probability current density (sometimes referred to as the probability flux density), denoted $\mathbf{j}(\mathbf{x}, t)$, such that the flux of probability across an elemental surface $\mathrm{d}\mathbf{S}$ is given by $\mathbf{j} \cdot \mathrm{d}\mathbf{S}$. The rate of change of the total probability contained within a volume V, shown in Figure 2.3, is related to the net flux leaving the surface by

$$\frac{\partial}{\partial t} \int_V \rho \, \mathrm{d}V = - \int_S \mathbf{j} \cdot \mathrm{d}\mathbf{S}.$$

Using the divergence theorem this can be written as

$$\frac{\partial}{\partial t} \int_V \rho \, \mathrm{d}V = - \int_V \boldsymbol{\nabla} \cdot \mathbf{j} \, \mathrm{d}V.$$

Because this holds for an arbitrary volume, the continuity equation for the conservation of quantum mechanical probability can be written

$$\boldsymbol{\nabla} \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \tag{2.20}$$

In non-relativistic quantum mechanics, the expression for the probability current can be obtained from the free particle time-dependent Schrödinger equation,

$$i\frac{\partial \psi}{\partial t} = -\frac{1}{2m}\nabla^2\psi, \tag{2.21}$$

and the corresponding equation for the complex conjugate of ψ,

$$-i\frac{\partial \psi^*}{\partial t} = -\frac{1}{2m}\nabla^2\psi^*. \tag{2.22}$$

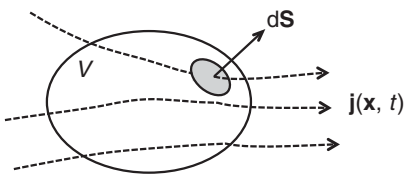

Fig. 2.3 The net flux of probability leaving a volume V.

Taking $\psi^* \times$ (2.21) $- \psi \times$ (2.22) then gives

$$-\frac{1}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) = i\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right)$$
$$\Rightarrow \quad -\frac{1}{2m}\boldsymbol{\nabla}\cdot\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right) = i\frac{\partial}{\partial t}(\psi^*\psi) = i\frac{\partial\rho}{\partial t}. \tag{2.23}$$

Comparing (2.23) with the general form of the continuity equation of (2.20) leads to the identification of probability current as

$$\mathbf{j} = \frac{1}{2im}\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right). \tag{2.24}$$

The plane wave

$$\psi(\mathbf{x}, t) = Ne^{i(\mathbf{p}\cdot\mathbf{x} - Et)},$$

is therefore associated with a constant probability density of $\psi\psi^* = |N|^2$ and can be interpreted as representing a region of space with a number density of particles $n = |N|^2$. The corresponding expression for the probability current density of (2.24) gives

$$\mathbf{j} = |N|^2\frac{\mathbf{p}}{m} \equiv n\mathbf{v},$$

where $\mathbf{v}$ is the (non-relativistic) velocity. Thus, the plane wave $\psi(\mathbf{x}, t)$ represents a region of space with number density of $n = |N|^2$ particles per unit volume moving with velocity $\mathbf{v}$, such that flux of particles passing through a unit area per unit time is $\mathbf{j} = n\mathbf{v}$.

2.3.3 Time dependence and conserved quantities

The time evolution of a quantum mechanical state is given by the time-dependent Schrödinger equation of (2.18). If ψ_i is an eigenstate of the Hamiltonian $\hat{H}$ with energy E_i such that

$$\hat{H}\psi_i(\mathbf{x}, t) = E_i\psi_i(\mathbf{x}, t),$$

then, from (2.18), the time evolution of the wavefunction is given by

$$i\frac{\partial\psi_i(\mathbf{x}, t)}{\partial t} = E_i\psi_i(\mathbf{x}, t).$$

Hence, the time dependence of an eigenstate of the Hamiltonian is given by

$$\psi_i(\mathbf{x}, t) = \phi_i(\mathbf{x})e^{-iE_it}. \tag{2.25}$$

For a system in a quantum mechanical state[1] $|\psi(\mathbf{x}, t)\rangle$, the expectation value of an operator $\hat{A}$ is given by

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \int \psi^\dagger \hat{A} \psi \, \mathrm{d}^3\mathbf{x},$$

where the complex conjugate used up to this point has been replaced by the Hermitian conjugate, $\psi^\dagger = (\psi^*)^T$. The time evolution of the expectation value $\langle \hat{A} \rangle$ can therefore be expressed as

$$\frac{\mathrm{d}\langle \hat{A} \rangle}{\mathrm{d}t} = \int \left[\frac{\partial \psi^\dagger}{\partial t} \hat{A} \psi + \psi^\dagger \hat{A} \frac{\partial \psi}{\partial t} \right] \mathrm{d}^3\mathbf{x}, \tag{2.26}$$

where it has been assumed that there is no explicit time dependence in the operator itself, i.e. $\partial \hat{A}/\partial t = 0$. The time derivatives in (2.26) can be expressed using (2.18) and its Hermitian conjugate, giving

$$\begin{aligned} \frac{\mathrm{d}\langle \hat{A} \rangle}{\mathrm{d}t} &= \int \left[\left\{ \tfrac{1}{i} \hat{H} \psi \right\}^\dagger \hat{A} \psi + \psi^\dagger \hat{A} \left\{ \tfrac{1}{i} \hat{H} \psi \right\} \right] \mathrm{d}^3\mathbf{x} && (2.27) \\ &= i \int \left[\psi^\dagger \hat{H}^\dagger \hat{A} \psi - \psi^\dagger \hat{A} \hat{H} \psi \right] \mathrm{d}^3\mathbf{x} \\ &= i \int \psi^\dagger (\hat{H}\hat{A} - \hat{A}\hat{H}) \psi \, \mathrm{d}^3\mathbf{x}. && (2.28) \end{aligned}$$

The last step follows from the fact that the Hamiltonian is Hermitian (which must be the case for it to have real eigenvalues). The relation of (2.28) implies that for *any* state

$$\frac{\mathrm{d}\langle \hat{A} \rangle}{\mathrm{d}t} = i \left\langle [\hat{H}, \hat{A}] \right\rangle, \tag{2.29}$$

where $[\hat{H}, \hat{A}] = \hat{H}\hat{A} - \hat{A}\hat{H}$ is the commutator of the Hamiltonian and the operator $\hat{A}$. Hence, if the operator $\hat{A}$ commutes with the Hamiltonian, the corresponding observable A does not change with time and therefore corresponds to a conserved quantity. Furthermore if ψ_i is an eigenstate of the Hamiltonian, then (2.27) immediately reduces to

$$\frac{\mathrm{d}\langle \hat{A} \rangle}{\mathrm{d}t} = \int \left[[iE_i \psi_i^\dagger] \hat{A} \psi_i + \psi_i^\dagger \hat{A} [-iE_i \psi_i] \right] \mathrm{d}^3\mathbf{x} = 0.$$

Therefore, for an eigenstate of the Hamiltonian, the expectation value of *any* operator is constant. For this reason, the eigenstates of the Hamiltonian are known as the *stationary states* of the system.

[1] The wavefunction ψ has been replaced by the more general state $|\psi\rangle$ written in Dirac ket notation which may have a number of degrees of freedom, for example spin.

In general, a state $|\varphi\rangle$ can be expressed in terms of the complete set of states formed from the eigenstates of the Hamiltonian $|\psi_i\rangle$,

$$|\varphi\rangle = \sum_i c_i |\psi_i\rangle,$$

and the time dependence of the system is determined by the evolution of the stationary states according to (2.25). If at time $t = 0$, a system is in the state $|\varphi(\mathbf{x})\rangle$, then the time evolution of the system is determined by the time evolution of the component stationary states

$$|\varphi(\mathbf{x}, t)\rangle = \sum_i c_i \, |\phi_i(\mathbf{x})\rangle e^{-iE_i t}. \tag{2.30}$$

This relationship between the time evolution of a state and the time dependence of the stationary states will be used extensively in the discussion of neutrino and strangeness oscillations in Chapters 13 and 14.

2.3.4 Commutation relations and compatible observables

The commutation relation between the operators for different observables determines whether they can be known simultaneously. Consider two observables corresponding to operators $\hat{A}$ and $\hat{B}$ which commute,

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0.$$

If $|\phi\rangle$ is an non-degenerate eigenstate of $\hat{A}$ with eigenvalue a, such that

$$\hat{A}|\phi\rangle = a|\phi\rangle,$$

then

$$\hat{A}\hat{B}\,|\phi\rangle = \hat{B}\hat{A}\,|\phi\rangle = a\hat{B}\,|\phi\rangle.$$

Therefore the state $\hat{B}\,|\phi\rangle$ is also an eigenstate of $\hat{A}$ with eigenvalue a. For this to be true, $\hat{B}\,|\phi\rangle \propto |\phi\rangle$, which implies that $|\phi\rangle$ satisfies

$$\hat{B}|\phi\rangle = b|\phi\rangle.$$

Hence $|\phi\rangle$ is a simultaneous eigenstate of both $\hat{A}$ and $\hat{B}$ and the state corresponds to well-defined values of the two observables, a and b. The same conclusion is obtained even if the states are degenerate. If $\hat{A}$ and $\hat{B}$ commute, the corresponding observables are referred to as compatible. In general, a quantum mechanical state can be labelled by the *quantum numbers* specifying the complete set of compatible observables. In the above example $|\phi\rangle$ can be labelled by $|a, b\rangle$. If there is a further operator $\hat{C}$ that commutes with both $\hat{A}$ and $\hat{B}$, the state is labelled by the quantum numbers $|a, b, c\rangle$. In the quantum mechanical description of angular momentum, described in Section 2.3.5, the states are labelled in terms of the eigenvalues of angular momentum squared and the z-component of angular momentum, $|\ell, m\rangle$.

Similar arguments can be applied to show that if $\hat{A}$ and $\hat{B}$ do not commute,

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0,$$

then it is not in general possible to define a simultaneous eigenstate of the two operators. In this case, it is not possible to know simultaneously the exact values of the physical observables A and B and the limit to which A and B can be known is given by the generalised uncertainty principle

$$\Delta A \, \Delta B \geq \tfrac{1}{2} \left| \langle i[\hat{A}, \hat{B}] \rangle \right|, \tag{2.31}$$

where $(\Delta A)^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$.

Position–momentum uncertainty relation

An important example of incompatible variables is that of the position and momentum uncertainty principal. The operators corresponding to the x position of a particle and the x component of its momentum are respectively given by

$$\hat{x}\psi = x\psi \quad \text{and} \quad \hat{p}_x \psi = -i\frac{\partial}{\partial x}\psi.$$

The commutator $[\hat{x}, \hat{p}_x]$ can be evaluated from its action on a wavefunction ψ,

$$\begin{aligned}[\hat{x}, \hat{p}_x]\psi &= -ix\frac{\partial}{\partial x}\psi + i\frac{\partial}{\partial x}(x\psi) \\ &= -ix\frac{\partial \psi}{\partial x} + i\psi + ix\frac{\partial \psi}{\partial x} = +i\psi,\end{aligned}$$

giving

$$[\hat{x}, \hat{p}_x] = +i.$$

The usual expression of the Heisenberg uncertainty principle for position and momentum is then obtained by substituting this commutation relation into (2.31) giving (after reinserting the hidden factor of $\hbar$)

$$\Delta x \, \Delta p_x \geq \frac{\hbar}{2}.$$

2.3.5 Angular momentum in quantum mechanics

The concept of angular momentum and its quantum mechanical treatment plays an important role in particle physics. In classical dynamics, the angular momentum $\mathbf{L}$ of a body is defined by the moment of its momentum,

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = (yp_z - zp_y, zp_x - xp_z, xp_y - yp_x).$$

The corresponding quantum mechanical operator $\hat{\mathbf{L}}$ is obtained by replacing the position and momentum coordinates by their operator equivalents. Hence, in quantum mechanics, the components of angular momentum operator are given by

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \quad \text{and} \quad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x.$$

Because the position operator does not commute with the corresponding component of momentum,

$$[\hat{x}, \hat{p}_x] = \left[\hat{y}, \hat{p}_y\right] = [\hat{z}, \hat{p}_z] = +i,$$

the angular momentum operators do not commute with each other and it is straightforward to show that

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hat{L}_z, \quad \left[\hat{L}_y, \hat{L}_z\right] = i\hat{L}_x \quad \text{and} \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hat{L}_y. \tag{2.32}$$

It is important to realise that the commutation relations of (2.32) are sufficient to fully define the algebra of angular momentum in quantum mechanics. This is significant because exactly the same commutation relations arise naturally in the discussion of other symmetries, such as flavour symmetry which is described in Chapter 9. For this reason, the development of the algebra defined by (2.32) and the subsequent identification of the angular momentum states is directly applicable to the more abstract symmetry concepts encountered in context of the quark model and QCD.

Because $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ do not commute, they correspond to incompatible observables and (unless the state has zero angular momentum) it is not possible to define a simultaneous eigenstate of more than one of the components of angular momentum. However, it is relatively straightforward to show (see Problem 2.15) that the operator for the total squared angular momentum defined by

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2,$$

commutes with each of the components of angular momentum,

$$\left[\hat{L}^2, \hat{L}_x\right] = \left[\hat{L}^2, \hat{L}_y\right] = \left[\hat{L}^2, \hat{L}_z\right] = 0.$$

Hence it is possible to express the angular momentum states in terms of the simultaneous eigenstates of $\hat{L}^2$ and any *one* of the components of angular momentum which, by convention, is chosen to be $\hat{L}_z$.

It is also useful to define angular momentum raising and lowering ladder operators,

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y,$$
$$\hat{L}_- = \hat{L}_x - i\hat{L}_y,$$

for which $\hat{L}_+^\dagger = \hat{L}_-$ and $\hat{L}_-^\dagger = \hat{L}_+$. Because $\hat{L}^2$ commutes with both $\hat{L}_x$ and $\hat{L}_y$, it also commutes with both the ladder operators,

$$\left[\hat{L}^2, \hat{L}_\pm\right] = 0.$$

The commutator of the ladder operators with $\hat{L}_z$ is given by

$$\begin{aligned}\left[\hat{L}_z, \hat{L}_\pm\right] &= \left[\hat{L}_z, \hat{L}_x\right] \pm i\left[\hat{L}_z, \hat{L}_y\right] \\ &= i\hat{L}_y \pm \hat{L}_x,\end{aligned}$$

and therefore

$$\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hat{L}_\pm. \tag{2.33}$$

Furthermore, using $\left[\hat{L}_x, \hat{L}_y\right] = i\hat{L}_z$, it can be shown that $\hat{L}^2$ can be expressed as (see Problem 2.15)

$$\hat{L}^2 = \hat{L}_-\hat{L}_+ + \hat{L}_z + \hat{L}_z^2. \tag{2.34}$$

The simultaneous eigenstates of $\hat{L}^2$ and $\hat{L}_z$ can be obtained using the relations of (2.33) and (2.34). Suppose the state $|\lambda, m\rangle$ is a simultaneous eigenstate of both $\hat{L}_z$ and $\hat{L}^2$, with eigenvalues given by

$$\hat{L}_z|\lambda, m\rangle = m|\lambda, m\rangle \quad \text{and} \quad \hat{L}^2|\lambda, m\rangle = \lambda|\lambda, m\rangle. \tag{2.35}$$

Now consider the state $\psi = \hat{L}_+|\lambda, m\rangle$, defined by the action of the angular momentum raising operator on the original state. Because $\hat{L}^2$ commutes with $\hat{L}_+$,

$$\hat{L}^2\psi = \hat{L}^2\hat{L}_+|\lambda, m\rangle = \hat{L}_+\hat{L}^2|\lambda, m\rangle = \lambda\hat{L}_+|\lambda, m\rangle = \lambda\psi.$$

Furthermore from (2.33), $\hat{L}_z\hat{L}_+ = \hat{L}_+\hat{L}_z + \hat{L}_+$ and therefore

$$\begin{aligned}\hat{L}_z\psi = \hat{L}_z\left[\hat{L}_+|\lambda, m\rangle\right] &= (\hat{L}_+\hat{L}_z + \hat{L}_+)|\lambda, m\rangle \\ &= (m+1)\left[\hat{L}_+|\lambda, m\rangle\right] = (m+1)\psi.\end{aligned}$$

Hence, the state $\psi = \hat{L}_+|\lambda, m\rangle$ is also a simultaneous eigenstate of $\hat{L}^2$ and $\hat{L}_z$, with respective eigenvalues of λ and $m+1$. Therefore, the effect of the angular momentum raising operators is to step along states with the same value of total angular momentum squared but with one unit more of the z-component of angular momentum. The angular momentum lowering operator has the opposite effect, lowering the z-component of angular momentum by one unit.

The magnitude of the z-component of angular momentum can be no greater than the total angular momentum itself,

$$\left\langle\hat{L}_z^2\right\rangle \leq \left\langle\hat{L}^2\right\rangle.$$

This implies that, for a particular value of λ, there must be maximum and minimum values of m and that the action of $\hat{L}_+$ on the state with the largest value of m gives

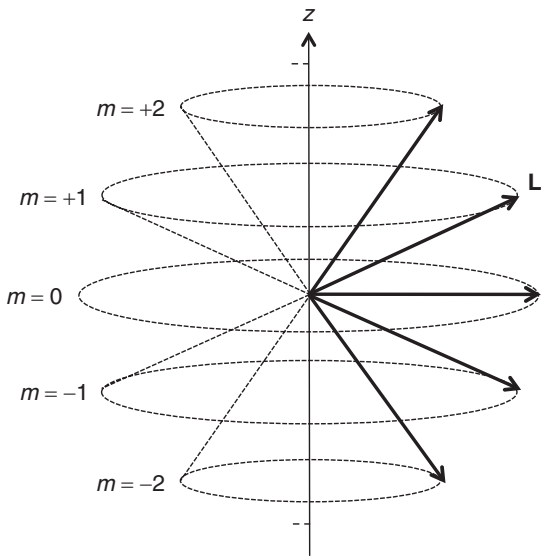

Fig. 2.4 A pictorial representation of the $2\ell + 1$ states for $\ell = 2$.

zero. Suppose the state with the largest z-component of angular momentum has $m = \ell$ such that

$$\hat{L}_+ |\lambda, \ell\rangle = 0,$$

then total angular momentum squared of this state is

$$\hat{L}^2 |\lambda, \ell\rangle = \left(\hat{L}_- \hat{L}_+ + \hat{L}_z + \hat{L}_z^2\right) |\lambda, \ell\rangle$$
$$\lambda |\lambda, \ell\rangle = (0 + \ell + \ell^2) |\lambda, \ell\rangle .$$

Hence, for the $m = \ell$ extreme state, the eigenvalue of $\hat{L}^2$ is $\lambda = \ell(\ell + 1)$. The same arguments can be applied to show at the other extreme, $m = -\ell$. Hence, for each value of λ (or equivalently for each value of ℓ), there are $2\ell + 1$ states (see Figure 2.4), differing by one unit of the z-component of angular momentum,

$$m = -\ell, \ -\ell + 1, \ldots, +\ell - 1, \ +\ell.$$

This implies that ℓ is quantised, and can take only integer or half-integer values. Expressing the states in terms of the quantum number ℓ rather than λ, the eigenvalue equations of (2.35) can be written as

$$\hat{L}_z |\ell, m\rangle = m |\ell, m\rangle \quad \text{and} \quad \hat{L}^2 |\ell, m\rangle = \ell(\ell + 1) |\ell, m\rangle .$$

The effect of the angular momentum raising operator on the state $|\ell, m\rangle$ is to generate the state $|\ell, m + 1\rangle$ with a coefficient $\alpha_{\ell,m}$ which still needs to be determined,

$$\hat{L}_+ |\ell, m\rangle = \alpha_{\ell,m} |\ell, m + 1\rangle . \tag{2.36}$$

Since $\hat{L}_+^\dagger = \hat{L}_-$, the Hermitian conjugate of (2.36) is

$$\left[\hat{L}_+ |\ell, m\rangle\right]^\dagger = \langle \ell, m|\hat{L}_- = \alpha^*_{\ell,m}\langle \ell, m+1|. \tag{2.37}$$

The coefficient $\alpha_{\ell,m}$ can be obtained by taking the product of (2.36) and (2.37) giving

$$\langle \ell, m|\, \hat{L}_-\hat{L}_+ \,|\ell, m\rangle = |\alpha_{\ell,m}|^2 \,\langle \ell, m+1|\ell, m+1\rangle.$$

Hence, for the normalised states $|\ell, m\rangle$ and $|\ell, m+1\rangle$,

$$\begin{aligned} |\alpha_{\ell,m}|^2 &= \langle \ell, m|\, \hat{L}_-\hat{L}_+ \,|\ell, m\rangle \\ &= \langle \ell, m|\, \hat{L}^2 - \hat{L}_z - \hat{L}_z^2 \,|\ell, m\rangle \\ &= (\ell(\ell+1) - m - m^2)\,\langle \ell, m|\ell, m\rangle \\ &= \ell(\ell+1) - m(m+1), \end{aligned}$$

and therefore,

$$\hat{L}_+ |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m+1)}\,|\ell, m+1\rangle. \tag{2.38}$$

The corresponding relation for the angular momentum lowering operator, which can be obtained in the same manner, is

$$\hat{L}_- |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m-1)}\,|\ell, m-1\rangle. \tag{2.39}$$

The relations given in (2.38) and (2.39) will be used to construct the angular momentum (and flavour states) formed from the combination of more than one particle.

2.3.6 Fermi's golden rule

Particle physics is mainly concerned with decay rates and scattering cross sections, which in quantum mechanics correspond to transitions between states. In non-relativistic quantum mechanics, calculations of transition rates are obtained from Fermi's golden rule. The derivation of Fermi's golden rule is far from trivial, but is included here for completeness.

Let $\phi_k(\mathbf{x}, t)$ be the normalised solutions to the Schrödinger equation for the unperturbed time-independent Hamiltonian $\hat{H}_0$, where

$$\hat{H}_0\phi_k = E_k\phi_k \quad \text{and} \quad \langle \phi_j|\phi_k\rangle = \delta_{jk}.$$

In the presence of an interaction Hamiltonian $\hat{H}'(\mathbf{x}, t)$, which can induce transitions between states, the time-dependent Schrödinger equation becomes

$$i\frac{\mathrm{d}\psi}{\mathrm{d}t} = \left[\hat{H}_0 + \hat{H}'(\mathbf{x}, t)\right]\psi. \tag{2.40}$$

The wavefunction $\psi(\mathbf{x}, t)$ can be expressed in terms of complete set of states of the unperturbed Hamiltonian as

$$\psi(\mathbf{x}, t) = \sum_k c_k(t)\phi_k(\mathbf{x})e^{-iE_k t}, \tag{2.41}$$

where the time-dependent coefficients $c_k(t)$ allow for transitions between states. Substituting (2.41) into (2.40) gives a set of differential equations for the coefficients $c_k(t)$,

$$i\sum_k \left[\frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_k t} - iE_k c_k \phi_k e^{-iE_k t}\right] = \sum_k c_k \hat{H}_0 \phi_k e^{-iE_k t} + \sum_k \hat{H}' c_k \phi_k e^{-iE_k t}$$

$$\Rightarrow \qquad i\sum_k \frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_k t} = \sum_k \hat{H}' c_k(t)\phi_k e^{-iE_k t}. \tag{2.42}$$

Suppose at time $t = 0$, the initial-state wavefunction is $|i\rangle = \phi_i$ and the coefficients are $c_k(0) = \delta_{ik}$. If the perturbing Hamiltonian, which is constant for $t > 0$, is sufficiently small that at all times $c_i(t) \approx 1$ and $c_{k\neq i}(t) \approx 0$, then to a first approximation (2.42) can be written

$$i\sum_k \frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_k t} \approx \hat{H}'\phi_i e^{-iE_i t}. \tag{2.43}$$

The differential equation for the coefficient $c_f(t)$, corresponding to transitions to a particular final state $|f\rangle = \phi_f$, is obtained by taking the inner product of both the LHS and RHS of (2.43) with $\phi_f(\mathbf{x})$ and using $\langle\phi_f|\phi_k\rangle = \delta_{fk}$ to give

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} = -i\langle f|\hat{H}'|i\rangle e^{i(E_f - E_i)t}, \tag{2.44}$$

where

$$\langle f|\hat{H}'|i\rangle = \int_V \phi_f^*(\mathbf{x})\hat{H}'\phi_i(\mathbf{x})\,\mathrm{d}^3\mathbf{x}.$$

The *transition matrix element* $T_{fi} = \langle f|\hat{H}'|i\rangle$ has dimensions of energy because both ϕ_i and ϕ_f are normalised by a volume integral. At time $t = T$, the amplitude for transitions to the state $|f\rangle$ is given by the integral of (2.44)

$$c_f(T) = -i\int_0^T T_{fi}\, e^{i(E_f - E_i)t}\,\mathrm{d}t.$$

If the perturbing Hamiltonian is time-independent, so is the term $\langle f|\hat{H}'|i\rangle$ and thus

$$c_f(T) = -iT_{fi}\int_0^T e^{i(E_f - E_i)t}\,\mathrm{d}t. \tag{2.45}$$

The probability for a transition to the state $|f\rangle$ is given by

$$P_{fi} = c_f(T)c_f^*(T) = |T_{fi}|^2 \int_0^T \int_0^T e^{i(E_f - E_i)t} e^{-i(E_f - E_i)t'}\,\mathrm{d}t\,\mathrm{d}t'.$$

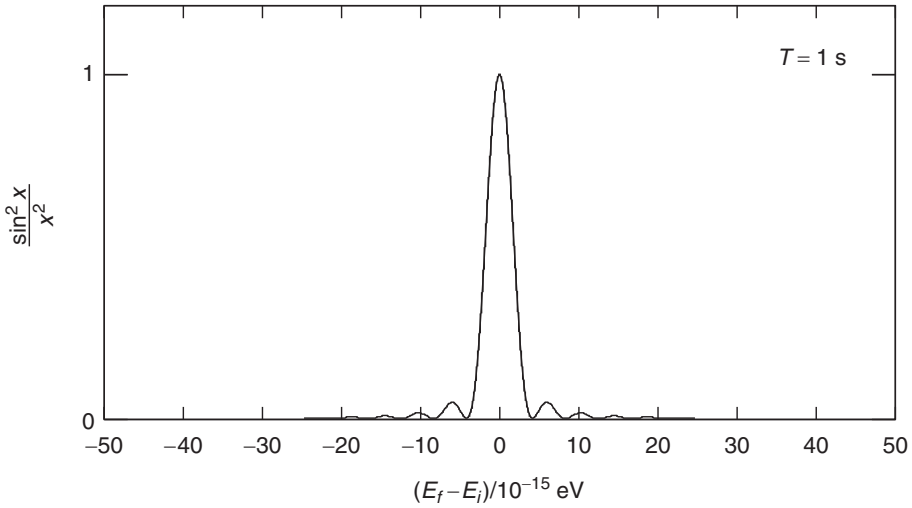

Fig. 2.5 The functional form of the integral of Equation (2.46) for $T = 1\,\mathrm{s}$.

The transition *rate* $\mathrm{d}\Gamma_{fi}$ from the initial state $|i\rangle$ to the single final state $|f\rangle$ is therefore

$$\mathrm{d}\Gamma_{fi} = \frac{P_{fi}}{T} = \frac{1}{T}|T_{fi}|^2 \int_{-\frac{T}{2}}^{+\frac{T}{2}} \int_{-\frac{T}{2}}^{+\frac{T}{2}} e^{i(E_f - E_i)t} e^{-i(E_f - E_i)t'}\,\mathrm{d}t\,\mathrm{d}t', \tag{2.46}$$

where the limits of integration are obtained by the substitutions $t \to t + T/2$ and $t' \to t' + T/2$. The exact solution to the integral in (2.46) has the form

$$\frac{\sin^2 x}{x^2} \quad \text{with} \quad x = \frac{(E_f - E_i)T}{2\hbar},$$

where the factor of $\hbar$ is included for clarity. This solution is shown for $T = 1\,\mathrm{s}$ in Figure 2.5, from which it can be seen that the transition rate is only significant for final states where $E_f \approx E_i$ and that energy is conserved within the limits of the energy–time uncertainty relation

$$\Delta E \Delta t \sim \hbar. \tag{2.47}$$

The narrowness of the functional form of (2.46) means that for all practical purposes, it can be written as

$$\mathrm{d}\Gamma_{fi} = |T_{fi}|^2 \lim_{T\to\infty} \left\{ \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \int_{-\frac{T}{2}}^{+\frac{T}{2}} e^{i(E_f - E_i)t} e^{-i(E_f - E_i)t'}\,\mathrm{d}t\,\mathrm{d}t' \right\}.$$

Using the definition of the Dirac delta-function given by (A.4) in Appendix A, the integral over dt' can be replaced by $2\pi\,\delta(E_f - E_i)$ and thus

$$d\Gamma_{fi} = 2\pi\,|T_{fi}|^2 \lim_{T\to\infty}\left\{\frac{1}{T}\int_{-\frac{T}{2}}^{+\frac{T}{2}} e^{i(E_f-E_i)t}\delta(E_f - E_i)\,dt\right\}.$$

If there are dn accessible final states in the energy range $E_f \to E_f + dE_f$, then the total transition rate Γ_{fi} is given by

$$\Gamma_{fi} = 2\pi \int |T_{fi}|^2 \frac{dn}{dE_f} \lim_{T\to\infty}\left\{\frac{1}{T}\int_{-\frac{T}{2}}^{+\frac{T}{2}} e^{-i(E_f-E_i)t}\delta(E_f - E_i)\,dt\right\} dE_f. \tag{2.48}$$

The delta-function in the integral implies that $E_f = E_i$ and therefore (2.48) can be written

$$\begin{aligned}\Gamma_{fi} &= 2\pi \int |T_{fi}|^2 \frac{dn}{dE_f}\delta(E_f - E_i) \lim_{T\to\infty}\left\{\frac{1}{T}\int_{-\frac{T}{2}}^{+\frac{T}{2}} dt\right\} dE_f \\ &= 2\pi \int |T_{fi}|^2 \frac{dn}{dE_f}\delta(E_f - E_i)\,dE_f \\ &= 2\pi\,|T_{fi}|^2 \left|\frac{dn}{dE_f}\right|_{E_i}.\end{aligned} \tag{2.49}$$

The term $\left|\frac{dn}{dE_f}\right|_{E_i}$ is referred to as the *density of states*, and is often written as $\rho(E_i)$ where

$$\rho(E_i) = \left|\frac{dn}{dE_f}\right|_{E_i}.$$

Fermi's golden rule for the total transition rate is therefore

$$\Gamma_{fi} = 2\pi\,|T_{fi}|^2 \rho(E_i),$$

where, to *first order*, $T_{fi} = \langle f|\hat{H}'|i\rangle$.

In the above derivation, it was assumed that $c_{k\neq i}(t) \approx 0$. An improved approximation can be obtained by again taking $c_i(t) \approx 1$ and substituting the expression for $c_{k\neq i}(t)$ from (2.45) back into (2.42), which after taking the inner product with a particular final state $\phi_f(\mathbf{x})$ gives

$$\frac{dc_f}{dt} \approx -i\langle f|\hat{H}|i\rangle e^{i(E_f-E_i)t} + (-i)^2 \sum_{k\neq i}\langle f|\hat{H}'|k\rangle e^{i(E_f-E_k)t}\int_0^t \langle k|\hat{H}'|i\rangle e^{i(E_k-E_i)t'}\,dt'. \tag{2.50}$$

Because the perturbation is not present at $t=0$, and for $t>0$ it is constant, the integral in (2.50) can be written

$$\int_0^t \langle k|\hat{H}'|i\rangle e^{i(E_k-E_i)t'}\,\mathrm{d}t' = \langle k|\hat{H}'|i\rangle \frac{e^{i(E_k-E_i)t}}{i(E_k-E_i)}.$$

Therefore, the improved approximation for the evolution of the coefficients $c_f(t)$ is given by

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} = -i\left(\langle f|\hat{H}|i\rangle + \sum_{k\neq i}\frac{\langle f|\hat{H}'|k\rangle\langle k|\hat{H}'|i\rangle}{E_i-E_k}\right)e^{i(E_f-E_i)t}.$$

Comparison with (2.44) shows that, to second order, the transition matrix element T_{fi} is given by

$$T_{fi} = \langle f|\hat{H}|i\rangle + \sum_{k\neq i}\frac{\langle f|\hat{H}'|k\rangle\langle k|\hat{H}'|i\rangle}{E_i-E_k}.$$

The second-order term corresponds to the transition occurring via some intermediate state $|k\rangle$. The full perturbation expansion can be obtained by successive substitutions. Provided the perturbation is sufficiently small, the successive terms in the perturbation expansion decrease rapidly, and it is possible to obtain accurate predictions using only the lowest-order term that contributes to a particular process.

Summary

Three main topics have been presented in this chapter. Firstly, the system of natural units with

$$\hbar = c = \varepsilon_0 = \mu_0 = 1$$

was introduced. It is used widely in particle physics and is adopted throughout this book. You should be comfortable with the concept of natural units and should be able to convert between natural units and S.I. units.

Because almost all of particle physics deals with relativistic particles, a sound understanding of special relativity and, in particular, the use of four-vectors is essential for much of what follows. Four-vector notation is used throughout this book with the conventions that the metric tensor is

$$g_{\mu\nu} = g^{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

such that zeroth component of a four-vector is the time-like quantity, for example

$$x^\mu = (t, x, y, z) \quad \text{and} \quad p^\mu = (E, p_x, p_y, p_z).$$

The scalar product of *any* two four-vectors,

$$a{\cdot}b \equiv a^\mu b_\mu \equiv g_{\mu\nu} a^\mu b^\mu \equiv a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = \text{invariant},$$

forms a Lorentz-invariant quantity that does not depend on the frame of reference. The results of the calculations that follow are usually presented in a frame independent manner using Lorentz invariant quantities.

A number of concepts in quantum mechanics are central to the theoretical ideas developed in the following chapters and it is important that you are familiar with the material reviewed in this chapter. Here the four most important concepts are: (i) the operator formulation of quantum mechanics, where physical observables are described by time-independent operators acting on time-dependent wavefunctions; (ii) the idea of stationary states of the Hamiltonian and the time development of a quantum mechanical system; (iii) the treatment of angular momentum in quantum mechanics and the algebra defined by the commutation relations between the angular momentum operators; and (iv) Fermi's golden rule to describe transition rates.

Problems

2.1 When expressed in natural units the lifetime of the W boson is approximately $\tau \approx 0.5\ \text{GeV}^{-1}$. What is the corresponding value in S.I. units?

2.2 A cross section is measured to be 1 pb; convert this to natural units.

2.3 Show that the process $\gamma \rightarrow e^+e^-$ can not occur in the vacuum.

2.4 A particle of mass 3 GeV is travelling in the positive z-direction with momentum 4 GeV; what are its energy and velocity?

2.5 In the laboratory frame, denoted Σ, a particle travelling in the z-direction has momentum $\mathbf{p} = \mathrm{p}_z\hat{\mathbf{z}}$ and energy E.

(a) Use the Lorentz transformation to find expressions for the momentum p'_z and energy E' of the particle in a frame Σ', which is moving in a velocity $\mathbf{v} = +\mathrm{v}\hat{\mathbf{z}}$ relative to Σ, and show that $E^2 - \mathrm{p}_z^2 = (E')^2 - (\mathrm{p}'_z)^2$.

(b) For a system of particles, prove that the total four-momentum squared,

$$p^\mu p_\mu \equiv \left(\sum_i E_i\right)^2 - \left(\sum_i \mathbf{p}_i\right)^2,$$

is invariant under Lorentz transformations.