11 Weak interactions

11.1 *Introduction*

The two lightest charged leptons, the electron and the muon, have been known to physicists for roughly a century and a half-century respectively. In weak interactions each charged lepton has a neutrino associated with it. In section 5.2.2 evidence was presented to show that the neutrino associated with the electron, ν_e, is distinct from its antiparticle, the antineutrino $\bar{\nu}_e$. In the early 1960s, experiments performed at Brookhaven and CERN produced conclusive evidence that a neutrino ν_μ, quite distinct from the electron neutrino, is associated with the muon. In essence, this assertion was based on the observation that muon neutrinos ν_μ, produced in the decay of π mesons

$$\pi^+ \to \mu^+ + \nu_\mu$$

subsequently gave rise to the production of muons via the reaction

$$\nu_\mu + n \to p + \mu^-$$

but no reactions occurred in which electrons were observed, i.e. there were no interactions of the type

$$\nu_\mu + n \to p + e^-.$$

These observations are formalized by the introduction of a lepton number L_e for the electron family and L_μ for the muon family and the requirement that L_e and L_μ are separately conserved. The lepton numbers are assigned as follows:

$$\begin{array}{lcccc} & e^- & \nu_e & e^+ & \bar{\nu}_e \\ L_e = & +1 & +1 & -1 & -1 \\ & \multicolumn{4}{l}{= 0 \text{ for all other particles}} \end{array}$$

$$\begin{array}{lcccc} & \mu^- & \nu_\mu & \mu^+ & \bar{\nu}_\mu \\ L_\mu = & +1 & +1 & -1 & -1 \\ & \multicolumn{4}{l}{= 0 \text{ for all other particles}} \end{array}$$

These conservation laws explain why the decays

$$\mu^\pm \to e^\pm + \gamma$$

and

$$\mu^\pm \to e^\pm + e^+ + e^-$$

for instance, have never been observed.

The newest member of the lepton family, the τ lepton, was first observed in 1975 by Perl *et al.*[1,2] at the SLAC e^+e^- collider, SPEAR. The analysis of the e^+e^- annihilations was complicated by the presence of a large hadronic background but 24 events were found in which the only charged particles produced were a high energy electron and a high energy muon with a substantial 'missing' or undetected energy. These events were attributed to the process

$$\begin{array}{l} e^+ + e^- \to \tau^+ + \tau^- \\ \qquad\qquad\quad \tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau \\ \qquad\qquad\quad \tau^+ \to e^+ + \nu_e + \bar{\nu}_\tau \end{array}$$

with the neutrinos carrying away the missing energy. The reaction occurred only above the threshold energy of 3.56 GeV, corresponding to a mass of 1.78 GeV/c^2 for the τ lepton. The *direct* observation of τ neutrinos through their interactions with matter has not yet been observed. As in the case of the electron and muon families, a τ

lepton number L_τ, which again is separately conserved, is assigned as follows:

$$\begin{array}{lcccc} & \tau^- & \nu_\tau & \tau^+ & \bar{\nu}_\tau \\ L_\tau = & +1 & +1 & -1 & -1 \end{array}$$
$$= 0 \text{ for all other particles}$$

In section 11.3.2 it will be shown that neutrinos and antineutrinos differ in the value of their helicity, the component of spin along the direction of motion, but it should be noted that this distinction holds only if the neutrinos have zero mass. Indeed, if neutrinos have non-zero mass, lepton number conservation will no longer be valid: the phenomenon of neutrino 'oscillations' could then occur with the decay of one neutrino type into another (see section 15.3). In the absence of strong evidence to the contrary we assume in this chapter that the neutrino masses are zero.

Three different types of weak interaction are known experimentally – leptonic, semi-leptonic and non-leptonic. Examples of the former are muon decay

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

and the τ lepton decay given above; only leptons are involved. Semi-leptonic weak interactions involve the hadrons, strange and non-strange, as well as the leptons. The β decay of the neutron

$$n \to p + e^- + \bar{\nu}_e$$

is an example. Note that since the neutron is heavier than the proton, the decay

$$p \to n + e^+ + \nu_e$$

is energetically forbidden for free protons – it can occur of course if the proton is bound in a nucleus. The β decay of the Λ^0

$$\Lambda^0 \to p + e^- + \bar{\nu}_e$$

and pion decay

$$\pi^+ \to \mu^+ + \nu_\mu \qquad \pi^- \to \mu^- + \bar{\nu}_\mu$$

are further examples of semi-leptonic weak interactions. The non-leptonic weak interactions involve hadrons only. Examples are

$$K^+ \to \pi^+\pi^0 \qquad K^+ \to \pi^+\pi^+\pi^-$$

and

$$\Lambda^0 \to p\pi^- \qquad \Lambda^0 \to n\pi^0.$$

These non-leptonic decays of the Λ^0 account for essentially 100 per cent of the decay rate; the branching fraction for the decay $\Lambda^0 \to p + e^- + \bar{\nu}_e$ is $(8.35 \pm 0.14) \times 10^{-4}$. Note that there is a change of strangeness $|\Delta S| = 1$ in these decays but this does not imply that weak interactions between non-strange hadrons do not exist. Rather, only in those cases in which the strong interaction is forbidden – strangeness-changing interactions – can the non-leptonic interaction be observed experimentally with relative ease.

11.2 *Parity violation*

The most striking difference between the weak interactions and the strong and electromagnetic interactions is that the former violate parity conservation while the latter conserve parity. Until the early 1950s it was common belief that all interactions were invariant under spatial inversion but the so-called τ–θ puzzle* cast doubt on this assumption as far as the weak interactions are concerned. The τ and θ mesons were originally thought to be different particles with weak decay modes $\theta \to 2\pi$ and $\tau \to 3\pi$. Since the pion has spin-parity 0^- it is clear that the spin-parity of the θ must belong to the series $0^+, 1^-, 2^+, \ldots$ etc., if the decay conserves parity. The spin-parity of the τ meson was determined from an analysis of the decay Dalitz plot and found to be 0^- and thus the τ meson appeared to be a different particle from the θ meson. The puzzle was that measurements of the masses and lifetimes of the τ and θ showed them to be equal within experimental errors and this degeneracy made it unlikely that the τ and θ were in fact different particles.

In 1956 Lee and Yang[3] resolved the paradox by suggesting that the τ and θ were different decay modes of the same particle (the K meson) and that parity was *not* conserved in the weak interactions. These 2π and 3π decay modes of the K meson can in retrospect be regarded as evidence for the breakdown of parity conservation, but at the time so little was known about these strange particles that evidence for parity violation was

* τ meson is an old-fashioned name for one of the decay modes of what is now called the K meson and should not be confused with the more recently discovered τ lepton.

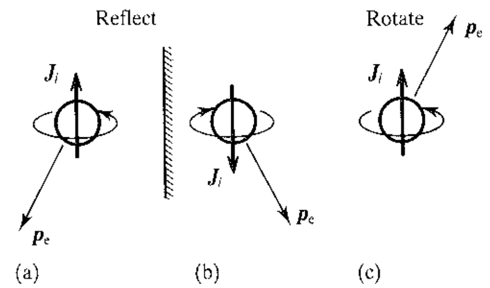


Figure 11.1 Schematic diagram showing parity violation in the β decay of aligned nuclei: (a) electrons emitted preferentially in a direction opposite to the nuclear spin J_i; (b) the system reflected in a plane; (c) rotation through 180° about an axis perpendicular to the plane and electrons now emitted preferentially in a direction along the nuclear spin.

sought in other weak interactions, in particular nuclear β decay. After their study of earlier work on β decay Lee and Yang concluded that no experiment had been performed which was sensitive to parity violating effects and they pointed out that in order to detect parity violation one would have to observe a *pseudoscalar* quantity.* In β decay experiments one suitable pseudoscalar is $\boldsymbol{J}_i \cdot \boldsymbol{p}_e$ where $\boldsymbol{J}_i$ (an axial vector) is the spin of the parent nucleus and $\boldsymbol{p}_e$ (a polar vector) is the electron momentum. An asymmetry in the angular distribution of the electrons measured with respect to the nuclear spin direction, i.e. a net value of the pseudoscalar $\boldsymbol{J}_i \cdot \boldsymbol{p}_e$, would constitute unequivocal proof of parity violation in the decay. This can be understood pictorially with the aid of figure 11.1. Imagine a β active source in which the spins of the parent nuclei are aligned and suppose that the electrons are emitted preferentially in a direction *opposite* to the direction of the nuclear spin as shown in figure 11.1(a). We recall that the parity operation can be envisaged as a reflection in a plane (figure 11.1(b)) followed by a 180° rotation about the normal to that plane (11.1(c)). Notice that in this parity transformation the polar vector (electron momentum) has changed sign whereas the axial vector (nuclear spin) has not. More importantly, we see in 11.1(c), the parity reflected system, that the electrons are emitted preferentially *along* the direction of the nuclear spin. The system is thus not invariant under the parity operation, i.e. parity is violated.

Another pseudoscalar quantity which has particular relevance to the study of β decay is the helicity of a particle; if parity is conserved the expectation value of the helicity must be zero (see example 11.1). Conversely, if a non-zero expectation value of helicity is observed in an interaction, parity must be violated in that interaction.

* A pseudoscalar is an observable which is invariant under rotation but which changes sign under spatial inversion (the parity operation); the scalar product of a polar vector and an axial vector is an example of a pseudoscalar. A scalar quantity on the other hand is invariant under both rotation and spatial inversion.

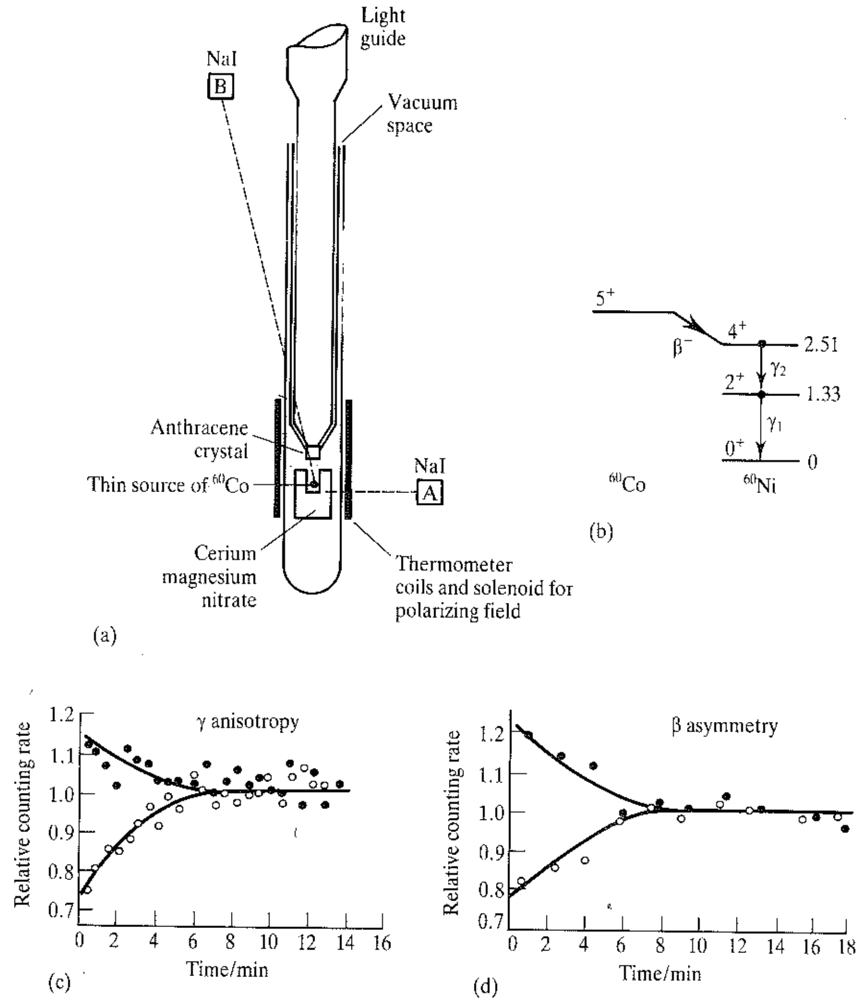


Figure 11.2 (a) Apparatus; (b) decay scheme of ^{60}Co (Gamow–Teller decay); (c) γ ray anisotropy obtained from counter A (●) and counter B (○) at different times as the crystal warms up; the difference between the curves measures the net polarization of the nuclei; (d) β ray asymmetry shown by the counting rate in the anthracene crystal for two directions of polarizing field (●, down ↓; ○ up ↑) (Wu C S *et al.* 1957 *Phys Rev* **105** (1413)).

11.3 Experimental verification of parity violation

11.3.1 The ^{60}Co experiment

Parity non-conservation was first established by measuring the angular distribution of electrons, emitted in the decay of ^{60}Co, relative to the orientation of the nuclear spin. Figure 11.2(a) shows the experimental arrangement used by Wu *et al.*[4] A source of ^{60}Co which decays according to the scheme shown in figure 11.2(b) was incorporated into a crystal of cerium magnesium nitrate. If a relatively small external magnetic field, ≈ 0.05 T, is applied to this paramagnetic salt the orientation of the *electronic* moments produces a strong local field of 10–100 T which, through the hyperfine coupling, will polarize the ^{60}Co nuclei if the sample is cooled to a temperature of about 0.01 K. The technique, first proposed by Gorter[5] and Rose,[6] achieves this low temperature through adiabatic demagnetization. The polarization was monitored by observing the anisotropy of the γ ray emission in the decay of ^{60}Ni to its ground state. For E2 transitions an angular distribution of the form $W(\theta) = \sum_{n=0}^{2} a_{2n} \times \cos^{2n}\theta$ is expected and a convenient measure of the nuclear polarization is the γ anisotropy coefficient $[W(\pi/2) - W(0)]/W(\pi/2)$. The γ rays were detected by the equatorial sodium iodide detector A and the polar detector B in figure 11.2(a) and the observed γ anisotropy is shown in figure 11.2(c). An anthracene scintillation counter was used to measure the intensity of β emission as a function of the direction of alignment (i.e. the direction of polarizing field), and the result shown in figure 11.2(d) was obtained. The asymmetry of β emission and the anisotropy of γ emission both disappeared as the crystal warmed up because of equalization of the population of magnetic substates.

The demonstration of β asymmetry in this experiment provided dramatic confirmation of the suggestion of Lee and Yang that parity might not be conserved in weak interactions. A similar confirmation for another weak decay process, the $\pi \to \mu \to e$ decay sequence, to which we will return in section 11.3.3, was almost simultaneously forthcoming. An asymmetric angular distribution always involves interference between amplitudes of opposite symmetry. The β asymmetry therefore means that the transition between the ground state of ^{60}Co and the second excited state of ^{60}Ni (both of definite parity) can take place by emission of an electron–antineutrino pair in both odd and even parity states and that the two corresponding amplitudes interfere, leading to an angular distribution of electron emission of the form $a + b\cos\theta$ with respect to the nuclear axis. That this asymmetry confirms the violation of parity conservation has already been explained in section 11.2.

11.3.2 *Measurement of the neutrino helicity*

The conjecture of Lee and Yang and its confirmation by Wu *et al.* was followed by a period of intense experimental and theoretical activity aimed at understanding the detailed nature of the weak interaction. Of paramount importance in this work was the measurement of the helicity of the neutrino by Goldhaber, Grodzins and Sunyar.[7] Here we give a simplified description of this classic experiment; for more details the reader is referred to the review article by Grodzins.[8]

The metastable state of ^{152}Eu decays about 24 per cent of the time by

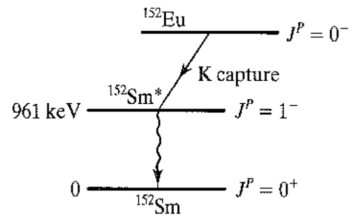


Figure 11.3
Simplified decay scheme of ^{152}Eu.

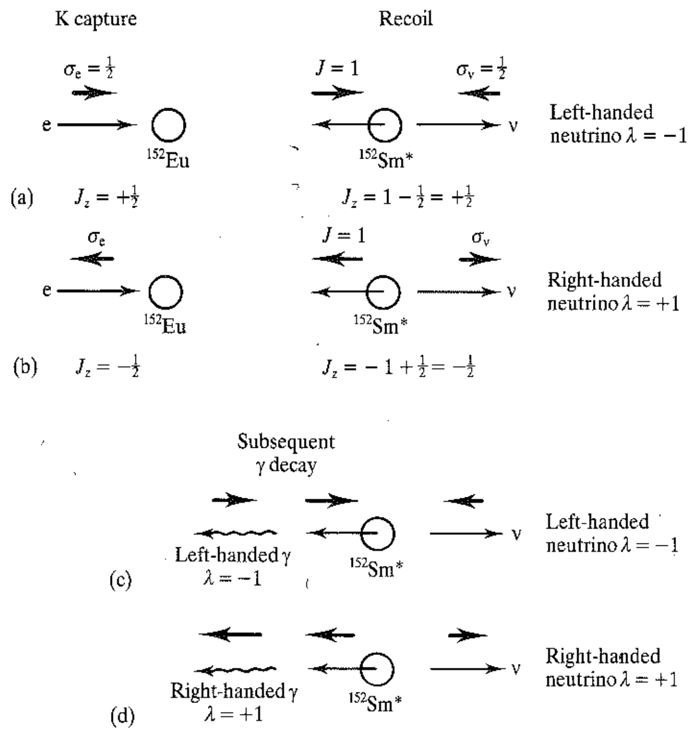


Figure 11.4 Schematic diagram, based on angular momentum conservation, showing the handedness of the particles involved in the K capture decay of ^{152}Eu: (a) production of left-handed neutrinos; (b) production of right-handed neutrinos. The subsequent decay of ^{152}Sm* illustrated in (c) and (d) shows that the forward-emitted γ rays have the same handedness as the neutrino.

K electron capture to an excited state of ^{152}Sm which in turn decays to the ground state of samarium by the emission of a 961 keV γ ray. (The lifetime of the excited state is extremely short, $\approx 10^{-14}$ s.) Furthermore, since K capture is a two-body process, the emitted neutrino has a unique energy, $E_\nu \approx 900$ keV in this case, i.e. close to the γ ray energy, an important consideration in this experiment. A simplified decay scheme of ^{152}Eu is shown in figure 11.3; note in particular the spin-parities of the levels involved. Conservation of angular momentum requires that the recoiling samarium nucleus (^{152}Sm*) has the same 'handedness' as the neutrino as depicted in figures 11.4(a) and 11.4(b), regardless of whether the neutrino has helicity $+1$ or -1. The problem of measuring the neutrino helicity (an extremely difficult one because of the very low interaction cross-section of neutrinos with matter) can thus be transferred to a measurement of the helicity of the recoiling $^{152}Sm^*$ nucleus. Since the lifetime of the 1^- level is only about 10^{-14} s the γ decay retains a knowledge of the nuclear recoil and again, by angular momentum conservation, γ rays emitted *in the direction of recoil* have the same helicity as the $^{152}Sm^*$ nucleus, and hence the same helicity as the neutrino as shown in figures 11.4(c) and 11.4(d). Thus the neutrino helicity can be inferred from a measurement of the helicity of the forward-produced γ rays, i.e. γ rays emitted in the direction of recoil of the $^{152}Sm^*$ nucleus. The experimental problem is thus to measure the γ ray helicity. This can be done relatively simply by examining the transmission of the γ rays through magnetized iron. The dominant interaction with matter of γ rays of energy 961 keV is the Compton effect and the method relies on the fact that the cross-section for Compton scattering is spin dependent. The formulae are somewhat complicated and can be found for instance in reference 8, but the main result is that the transmission is greatest when the photon spin is parallel to the electron spin. Of course, as indicated above, it is only those γ rays emitted in the opposite direction to the neutrino which have the same helicity as the neutrino and the ingenious method devised by Goldhaber *et al.* to select these γ rays was to make use of resonant scattering. In the emission of a γ ray from an excited state with energy of excitation E_0 a momentum E_0/c must be imparted to the emitting nucleus and consequently the energy of the γ ray is reduced by an amount $E_0^2/2Mc^2$ where M is the mass of the nucleus. Similarly, on absorption, an extra energy $E_0^2/2Mc^2$ must be supplied to counteract the nuclear recoil. This energy, $\Delta E = E_0^2/Mc^2$, lost by recoil in emission and absorption, is in general much greater than the level width so that resonant absorption will take place only if extra energy, equal to the energy lost, is supplied to the emitted γ rays. In the experiment under consideration, it is precisely those γ rays emitted in the direction of recoil of the $^{152}Sm^*$ nucleus, i.e. opposite to the neutrino direction, which have the correct energy to undergo resonant absorption. A simple calculation (see example 11.2) shows that the resonance condition is $E_\nu \cos\theta = E_0$ for a γ ray emitted at an angle θ with respect to the $^{152}Sm^*$ recoil direction. Thus for γ rays emitted opposite to the neutrino direction the requirement that $E_\nu \approx E_0$ becomes apparent. Thermal motion gives rise to a spread in energies so that the resonance condition can be met in practice.

The experimental arrangement used by Goldhaber *et al.* is shown in figure 11.5. The scatterer was in the form of a ring of samarium oxide and the γ rays were detected with an NaI(Tl) detector which was shielded from the direct radiation by 12 in. of lead. The helicity of the γ rays was analysed by transmission through fully magnetized iron. With the field direction as shown in figure 11.5 the γ ray energy spectrum shown in figure 11.6 was obtained. The presence of full energy peaks at 961 and

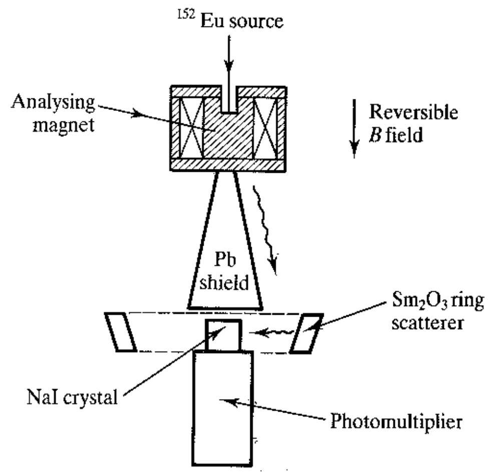


Figure 11.5 Experimental arrangement used by Goldhaber *et al.* to measure the circular polarization of resonantly scattered γ rays. With the field direction as shown left-handed γ rays are preferentially transmitted (Goldhaber M *et al.* 1958 *Phys Rev* **109** (1015)).

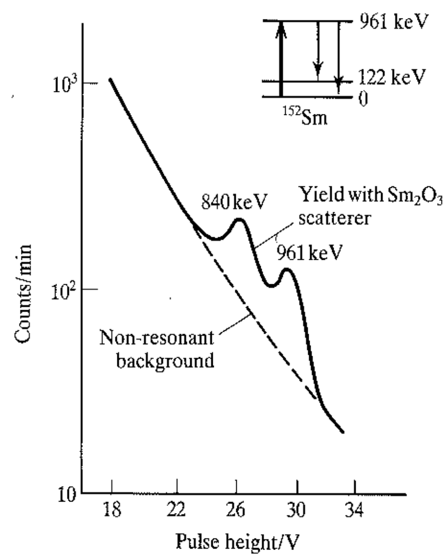


Figure 11.6 Pulse height spectrum of resonantly scattered γ rays from the decay of $^{152}Sm^*$ (upper curve). Non-resonant background is indicated by the lower curve. The energy levels of ^{152}Sm are shown in the inset.

837 keV indicates that resonant scattering has taken place and since, with the field direction indicated, γ rays with helicity $\lambda = -1$ are preferentially transmitted, the neutrino must have helicity -1. Two full energy peaks are observed in the energy spectrum of the scattered radiation because the decay scheme is slightly more complex than indicated in figure 11.3 and is shown as an inset in figure 11.6. After

resonant absorption de-excitation of the 961 keV level can take place either by a direct transition to the ground state or via an intermediate level at 122 keV.

We note that the observation of a definite helicity of the neutrino in K capture is evidence for parity violation in the decay as indicated in section 11.2.

The helicity of the antineutrino has been inferred to be positive from the asymmetry measurements in the decay of polarized neutrons,[9] a result which is supported by later measurements. The neutrino and antineutrino are thus distinguished from each other by their helicities: the neutrino has helicity -1 while the antineutrino has helicity $+1$; the parity violation is said to be maximal.

11.3.3 Parity violation in the $\pi \to \mu \to e$ *decay sequence*

Following the proposal of Lee and Yang,[3] Garwin, Lederman and Weinrich[10] confirmed that parity is violated in the successive decays $\pi^+ \to \mu^+ + \nu_\mu$, $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$. If parity violation is maximal, as in the case of β decay, then in the decay of stopped pions the muons will be polarized along the direction of motion as indicated in figure 11.7(a). Maximal parity violation implies that the muon neutrino ν_μ is left handed (helicity $= -1$) and since the pion has spin 0 and there can be no component of orbital angular momentum along the line of flight of the muon, conservation of angular momentum requires that the μ^+ must be polarized with helicity -1.

In the experiment of Garwin *et al.* the muons were brought to rest, without loss of polarization, in a carbon absorber. In the subsequent decay the positron energy spectrum was strongly peaked near the maximum allowed value ($m_\mu/2$) as shown schematically in figure 11.7(b). The most

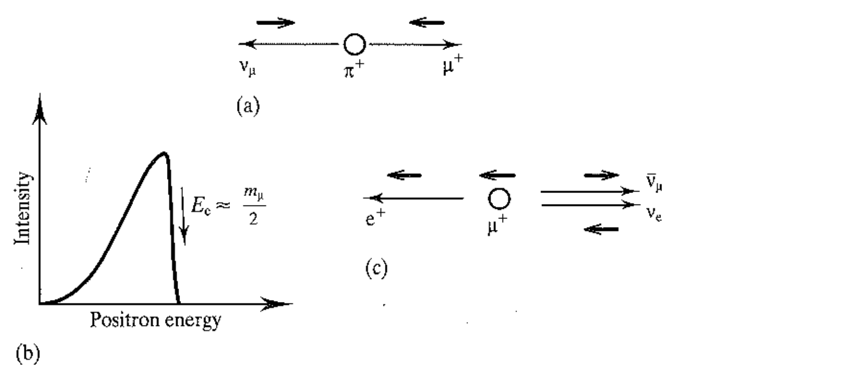


Figure 11.7 (a) Sketch showing polarization in the decay $\pi^+ \to \mu^+\nu_\mu$. Particle spins are shown by short arrows. (b) Sketch of the positron energy spectrum for the decay $\mu^+ \to e^+\nu_e\bar{\nu}_\mu$. The spectrum is strongly peaked towards the kinematically allowed maximum. (c) Particle polarizations in the most likely configuration in the decay $\mu^+ \to e^+\nu_e\bar{\nu}_\mu$.

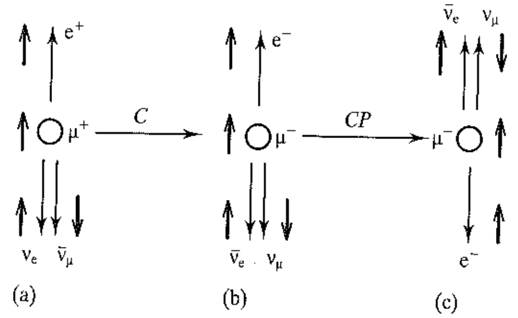


Figure 11.8
Sketch showing C violation and CP conservation in the decay $\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu$.

probable decay configuration is therefore that shown in figure 11.7(c) and, if parity is maximally violated, the particle helicities will be as indicated. One thus expects an asymmetry in the angular distribution of the positron with respect to the spin direction of the muon, with more positrons being emitted parallel to the muon spin direction, i.e. opposite to the incident muon direction. Experimentally the angular distribution was of the form $I(\theta) = 1 - \frac{1}{3}\cos\theta$ where θ was measured with respect to the line of flight of the incident muon. Not only does this asymmetry confirm parity violation in the decay processes but in addition the form of the distribution is precisely that required by the so-called V–A interaction discussed in section 11.5.5.

The weak interactions, in addition to violating parity also violate charge conjugation invariance, as can be seen by a consideration of the most probable configuration (figure 11.8(a)), in muon decay. Figure 11.8(b) shows the result of the charge conjugation operation on this configuration. The neutrinos ν_μ and $\bar{\nu}_e$ have helicities which do not occur in nature, thus the decay is not invariant under the charge conjugation operation. However, if this is followed by the parity transformation, the configuration in figure 11.8(c) results and in this case the particle helicities are precisely as expected. Thus, although the interaction separately violates P and C, it is invariant under the combined operation CP. This is true of most weak interactions but in the very important case of the decays of the neutral K mesons there is good evidence for the existence of a small CP-violating amplitude. We defer discussion of this phenomenon until section 11.13.4.

In this chapter so far we have learned two crucial experimental facts about the weak interactions, namely that parity is violated and that the neutrino (antineutrino) has negative (positive) helicity. In chapter 5 we outlined the Fermi theory of β decay and discussed the transition rate mainly in terms of the density of states factor. We now proceed to investigate what restrictions are imposed on the matrix element by the fact that parity is violated in the weak interactions and the violation is maximal, giving rise to neutrinos and antineutrinos of definite handedness. Since the particles involved often have relativistic velocities – neutrinos always travel with the velocity of light – our starting point must be a relativistic wave equation.

11.4 Relativisitic quantum mechanics

11.4.1 The Klein–Gordon equation

A relativistic wave equation may be constructed by making the usual association, $E = i\hbar(\partial/\partial t)$ and $\boldsymbol{p} = -i\hbar\nabla$, between the dynamical variables and operators and using the relativistic energy–momentum relation $E^2 = p^2 + m^2$ ($c = 1$). The resulting equation

$$\left(\nabla^2 - \frac{m^2}{\hbar^2}\right)\phi = \frac{\partial^2 \phi}{\partial t^2} \tag{11.1}$$

is known as the Klein–Gordon equation. By judicious manipulation of equation (11.1) we obtain (see example 11.4)

$$\frac{\partial}{\partial t}\left[i\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right)\right] + \boldsymbol{\nabla}\cdot\left[\frac{1}{i}(\phi^*\nabla\phi - \phi\nabla\phi^*)\right] = 0 \tag{11.2}$$

which has the form of the continuity equation $\partial\rho/\partial t + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$, if we define

$$\rho = i\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right) \qquad \text{and} \qquad \boldsymbol{j} = \frac{1}{i}(\phi^*\nabla\phi - \phi\nabla\phi^*).$$

The probability current density $\boldsymbol{j}$ is essentially the same as in the non-relativistic case, but a problem arises when we attempt to interpret ρ as a probability density. A probability density must be positive definite. If, as in example 11.4, we calculate ρ for a plane wave solution of the Klein–Gordon equation, we obtain $\rho = 2|N|^2E$, where E is the particle energy and N is a normalization constant. The problem arises because the energy eigenvalues of the Klein–Gordon equation are $E = \pm(p^2 + m^2)^{1/2}$. The negative energy solutions originally caused a problem but it was subsequently shown that they can be interpreted in terms of antiparticles, which certainly exist in nature, and may be regarded as support for this approach. However, since ρ has the same sign as E, the negative energy solutions present a problem in the interpretation of ρ as a (matter) probability density. We therefore follow the course of history and proceed to a discussion of the Dirac equation. In order to facilitate this discussion we first introduce the notation commonly used in relativistic quantum mechanics.

11.4.2 Relativistic notation

It is well known that space–time coordinates (ct, x, y, z) in a stationary reference frame are related to those, (ct', x', y', z'), in a frame moving with uniform velocity v in the direction of the positive z axis by the Lorentz transformation

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{11.3}$$

where $\beta = v/c$ and $\gamma = (1 - \beta^2)^{-1/2}$. The four-vector (ct, x, y, z) is often written as x^μ, with the understanding that $x^0 = ct$, $x^1 = x$, $x^2 = y$ and $x^3 = z$, and (11.3) can be written in the shorthand notation

$$x'^\mu = \sum_\nu \Lambda^\mu_\nu x^\nu \qquad (\mu, \nu = 0, 3). \tag{11.4}$$

Any set of four quantities which transform according to (11.4) is by definition a *contravariant* four-vector. A *covariant* four-vector, written with the index in the lower position, has the same time component as the contravariant vector but the space components have the opposite sign. Thus x_μ is a covariant vector with $x_0 = ct$, $x_1 = -x$, $x_2 = -y$ and $x_3 = -z$. The two are related through the metric tensor $g_{\mu\nu}$, with components $g_{00} = 1$, $g_{ii} = -1$ for $i = 1, 2, 3$ and $g_{\mu\nu} = 0$ for $\mu \neq \nu$, i.e.

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{11.5}$$

Then,

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

$$x^\mu = \sum_\nu g^{\mu\nu} x_\nu$$

with $g^{\mu\nu}$ defined by the relation

$$\sum_{\mu} g^{\mu\nu} g_{\mu\sigma} = \delta^{\nu}_{\sigma}$$

where the Kronecker symbol

$$\delta^{\nu}_{\sigma} = \begin{cases} 1 & \text{for } \sigma = \nu \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the raising or lowering of the indices of a four-vector changes the sign of the space components but leaves the time component unchanged.

A Lorentz invariant may be formed from the scalar product of two four-vectors and, in particular, the scalar product of x_μ with itself is

$$\sum_{\mu\nu} g^{\mu\nu} x_\nu x_\mu = \sum_{\mu} x^\mu x_\mu = c^2 t^2 - \boldsymbol{x}\cdot\boldsymbol{x}. \tag{11.6}$$

Similarly, for the energy–momentum four-vector of a free particle $p^\mu = (E/c, p_x, p_y, p_z)$ we have

$$\sum_{\mu\nu} g^{\mu\nu} p_\nu p_\mu = \sum_{\mu} p^\mu p_\mu = \frac{E^2}{c^2} - \boldsymbol{p}\cdot\boldsymbol{p} = m^2 c^2 \tag{11.7}$$

where m is the rest mass of the particle. Henceforth, we revert to natural units in which the speed of light $c = 1$. In addition, we will use Greek letters for indices which run from 0 to 3 and Latin letters for those which run from 1 to 3 and use the convention that repeated indices imply a summation over all possible values so that the Lorentz transformation for a contravariant vector (11.4) will simply be written

$$x'^\mu = \Lambda^\mu_\nu x^\nu. \tag{11.8}$$

The corresponding transformation law for a covariant four-vector is

$$\Lambda^\mu_\nu x'_\mu = x_\nu. \tag{11.9}$$

In what follows we shall need to differentiate with respect to the space–time coordinates. It can be shown that since x^μ is a contravariant vector $\partial/\partial x^\mu$ is a covariant vector and to stress this it is written with a lower index,

$$\partial_\mu \equiv \left(\frac{\partial}{\partial x^0}, \nabla\right) = \left(\frac{\partial}{\partial t}, \nabla\right) \tag{11.10}$$

where ∇ is the three-dimensional gradient operator. On the other hand,

$$\partial^\mu \equiv \left(\frac{\partial}{\partial x_0}, -\nabla\right) = \left(\frac{\partial}{\partial t}, -\nabla\right) \tag{11.11}$$

is a contravariant vector. It follows that, in this notation,

$$\partial_\mu \partial^\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 = \Box^2 \tag{11.12}$$

is the d'Alembertian operator. Note that in constructing a Lorentz invariant the upper and lower indices must balance. The condition that an equation be Lorentz covariant is that on the two sides of the equation the unrepeated upper and lower indices must balance separately and repeated indices must appear once as an upper and once as a lower index.

11.4.3 The Dirac equation

In 1928 Dirac[11] succeeded in finding a form of the equation

$$i\hbar \frac{\partial}{\partial t} \psi(\boldsymbol{x}, t) = H\psi(\boldsymbol{x}, t) \tag{11.13}$$

which satisfied Lorentz covariance and avoided the problem of negative probability densities encountered with the Klein–Gordon equation. Dirac argued that the symmetry between energy and momentum required by special relativity dictates that since the Hamiltonian H is linear in E it must be linear in $\boldsymbol{p}$. Equivalently, since (11.13) is linear in the time derivative, the Hamiltonian must be linear in the space derivatives. Equation (11.13) may therefore be written in the general form

$$i\frac{\partial\psi}{\partial t}(\boldsymbol{x}, t) = \left[-i\left(\alpha_1 \frac{\partial}{\partial x^1} + \alpha_2 \frac{\partial}{\partial x^2} + \alpha_3 \frac{\partial}{\partial x^3}\right) + \beta m\right]\psi(\boldsymbol{x}, t) \tag{11.14}$$

where we have used natural units with $\hbar = c = 1$. Here m is the rest energy of the particle.

We shall see that the coefficients α_i and β cannot simply be numbers and Dirac proposed that (11.14) be regarded as a matrix equation with the wavefunction ψ written as a column matrix with n components,

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{bmatrix}$$

and the coefficients α_i and β as $n \times n$ matrices. Equation (11.14) becomes in effect n coupled first-order equations. For (11.14) to be a suitable relativistic equation the correct energy–momentum relation, $E^2 = p^2 + m^2$, must hold for a free particle and therefore each component of ψ must satisfy the Klein–Gordon equation,

$$-\frac{\partial^2 \psi}{\partial t^2} = (-\nabla^2 + m^2)\psi .$$

With

$$\alpha_1 \frac{\partial}{\partial x^1} + \alpha_2 \frac{\partial}{\partial x^2} + \alpha_3 \frac{\partial}{\partial x^3}$$

written as $\boldsymbol{\alpha}\cdot\nabla$ the Dirac equation becomes

$$\mathrm{i}\frac{\partial \psi}{\partial t} = (-\mathrm{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi . \tag{11.15}$$

Then,

$$\begin{aligned}\left(\mathrm{i}\frac{\partial}{\partial t}\right)^2 \psi &= -\frac{\partial^2 \psi}{\partial t^2} \\ &= (-\mathrm{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)(-\mathrm{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \\ &= -\tfrac{1}{2}\sum_{i,\, j=1}^{3} (\alpha_i\alpha_j + \alpha_j\alpha_i)\frac{\partial^2 \psi}{\partial x^i\, \partial x^j} \\ &\quad - \mathrm{i}m \sum_{i=1}^{3} (\alpha_i\beta + \beta\alpha_i)\frac{\partial \psi}{\partial x^i} + \beta^2 m^2 \psi \\ &= (-\nabla^2 + m^2)\psi .\end{aligned}$$

These relations impose conditions on the coefficients α_i and β and inspection shows that in order that the last equality be satisfied

$$\alpha_i\alpha_j + \alpha_j\alpha_i = \{\alpha_i, \alpha_j\} = 2\delta_{ij} I \tag{11.16a}$$

$$\alpha_i\beta + \beta\alpha_i = \{\alpha_i, \beta\} = 0 \tag{11.16b}$$

$$\alpha_i^2 = \beta_i^2 = I \tag{11.16c}$$

where I is the unit matrix. It is now apparent, from the anticommutation properties (11.16a, b), that the α_i and β must be matrices and not simply numbers. In order that the Hamiltonian be a hermitian operator the α_i and β must be hermitian matrices. It can be shown that the smallest dimension in which the α_i and β can be realized is $n = 4$. One conventional choice for these 4 × 4 matrices is the Dirac–Pauli representation

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix} \qquad \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{11.17}$$

where the σ_i are the 2 × 2 Pauli spin matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{11.18}$$

and I is the 2 × 2 unit matrix. The wavefunction $\psi(\boldsymbol{x}, t)$ in (11.14) will then have four components and is known as a Dirac spinor,

$$\psi(\boldsymbol{x}, t) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix}.$$

We investigate immediately whether the Dirac equation (11.14) gives rise to a continuity equation with a positive definite probability density. We introduce the hermitian conjugate wavefunction, a row vector,

$$\psi^\dagger = [\psi_1^* \quad \psi_2^* \quad \psi_3^* \quad \psi_4^*]$$

and rewrite the Dirac equation as

$$\mathrm{i}\frac{\partial \psi}{\partial t} = -\mathrm{i}\sum_{k=1}^{3} \alpha_k \frac{\partial}{\partial x^k}\psi + m\beta\psi. \tag{11.19}$$

Multiplying from the left by $\psi^\dagger$ gives

$$\mathrm{i}\psi^\dagger \frac{\partial \psi}{\partial t} = -\mathrm{i}\sum_{k=1}^{3} \psi^\dagger \alpha_k \frac{\partial \psi}{\partial x^k} + m\psi^\dagger\beta\psi. \tag{11.20}$$

The hermitian conjugate of the Dirac equation is

$$-\mathrm{i}\frac{\partial \psi^\dagger}{\partial t} = \mathrm{i}\sum_{k=1}^{3} \frac{\partial \psi^\dagger}{\partial x^k}\alpha_k + m\psi^\dagger\beta \tag{11.21}$$

since $\alpha_k^\dagger = \alpha_k$ and $\beta^\dagger = \beta$ (α_k, β hermitian). On multiplying (11.21) from the right by ψ we obtain

$$-\mathrm{i}\frac{\partial \psi^\dagger}{\partial t}\psi = \mathrm{i}\sum_{k=1}^{3}\frac{\partial \psi^\dagger}{\partial x^k}\alpha_k\psi + m\psi^\dagger\beta\psi. \tag{11.22}$$

Subtraction of (11.22) from (11.20) gives

$$\frac{\partial}{\partial t}\psi^\dagger\psi + \sum_{k=1}^{3}\frac{\partial}{\partial x^k}(\psi^\dagger\alpha_k\psi) = 0 \tag{11.23}$$

which we recognize as an equation of continuity $\partial\rho/\partial t + \nabla\cdot\boldsymbol{j} = 0$ if we make the association $\rho = \psi^\dagger\psi$ and $\boldsymbol{j} = \psi^\dagger\boldsymbol{\alpha}\psi$. The probability current density $\boldsymbol{j}$ has three components $j^k = \psi^\dagger\alpha_k\psi$ with $k = 1, 2, 3$ and the probability density

$$\rho = \psi^\dagger\psi = \sum_{i=1}^{4}|\psi_i|^2 > 0$$

is now positive definite.

The Dirac equation can be cast into covariant form by introducing the Dirac γ matrices to replace the matrices α_i and β. They are defined as

$$\gamma^0 = \beta \qquad \gamma^k = \beta\alpha^k \qquad (k = 1, 2, 3)$$

or equivalently

$$\gamma^\mu = (\beta, \beta\boldsymbol{\alpha}) \qquad (\mu = 0, 1, 2, 3). \tag{11.24}$$

It is a simple matter to show from the anticommutation properties of the α_k and β matrices (11.16) that the γ matrices satisfy the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I. \tag{11.25}$$

In terms of these γ matrices and the covariant derivarive, ∂_μ, the Dirac equation takes on the covariant form

$$(\mathrm{i}\gamma^\mu\,\partial_\mu - m)\psi = 0 \tag{11.26}$$

and the equation of continuity becomes

$$\partial_\mu j^\mu = 0 \tag{11.27}$$

where $j^\mu = \bar{\psi}\gamma^\mu\psi$ is a four-vector current. Here, $\bar{\psi}$ is the row spinor defined by $\bar{\psi} = \psi^\dagger\gamma^0$. We identify the probability density ρ with the time component of j^μ;

$$\rho = \bar{\psi}\gamma^0\psi = \psi^\dagger(\gamma^0)^2\psi = \psi^\dagger\psi.$$

The space components are

$$j^k = \bar{\psi}\gamma^k\psi = \psi^\dagger\gamma^0\gamma^k\psi = \psi^\dagger\alpha_k\psi$$

in agreement with the result obtained earlier in this section.

11.4.4 The Dirac equation in the presence of an electromagnetic field: intrinsic spin of the electron

We can look for plane-wave solutions to the Dirac equation for a free particle in the form

$$\psi = u(\boldsymbol{p})\exp(-\mathrm{i}p\cdot x) \tag{11.28}$$

where $u(\boldsymbol{p})$ is a four-component spinor independent of the space–time coordinates x. On substituting (11.28) into the Dirac equation we get the Dirac equation for a free particle spinor,

$$(\gamma^\mu p_\mu - m)u(\boldsymbol{p}) = 0 \tag{11.29}$$

in which we note that p_μ is a covariant four-vector with components $(E, -\boldsymbol{p})$.

In non-relativistic quantum mechanics, the wave equation for an electron in the presence of an external electromagnetic field, specified by the scalar and vector potentials ϕ and $\boldsymbol{A}$, is set up by making the substitution

$$\mathrm{i}\frac{\partial}{\partial t} \to \mathrm{i}\frac{\partial}{\partial t} + e\phi \qquad -\mathrm{i}\nabla \to -\mathrm{i}\nabla + e\boldsymbol{A} \tag{11.30}$$

in the free particle Schrödinger equation. In (11.30) $-e$ is the charge on the electron. In terms of the four-vector potential $A^\mu = (\phi, \boldsymbol{A})$, the 'minimal substitution', as (11.30) is often called, takes on the form

$$\partial_\mu \to \partial_\mu - \mathrm{i}eA_\mu. \tag{11.31}$$

Then the Dirac equation becomes

$$[\gamma^\mu(p_\mu + eA_\mu) - m]u(\boldsymbol{p}) = 0. \tag{11.32}$$

It is convenient to write equation (11.32) in terms of the matrices $\boldsymbol{\alpha}$ and β. We recall that $\gamma^{\mu} \equiv (\beta, \beta\boldsymbol{\alpha})$, then

$$(\beta E - \beta\boldsymbol{\alpha}\cdot\boldsymbol{p} + e\beta\phi - e\beta\boldsymbol{\alpha}\cdot\boldsymbol{A} - m)u(\boldsymbol{p}) = 0.$$

If we multiply from the left by β and use the fact that $\beta^2 = I$ we have

$$[\boldsymbol{\alpha}\cdot(\boldsymbol{p} + e\boldsymbol{A}) + \beta m - e\phi]u(\boldsymbol{p}) = Eu(\boldsymbol{p}). \tag{11.33}$$

In attempting to solve (11.33) it is usual to express the four-component spinor $u(\boldsymbol{p})$ in terms of two two-component spinors,

$$u \equiv \begin{pmatrix} u_{\mathrm{A}} \\ u_{\mathrm{B}} \end{pmatrix}.$$

Then, using the explicit expressions for the matrices $\boldsymbol{\alpha}$ and β, (11.33) becomes

$$\begin{pmatrix} m - e\phi & \boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A}) \\ \boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A}) & -m - e\phi \end{pmatrix}\begin{pmatrix} u_{\mathrm{A}} \\ u_{\mathrm{B}} \end{pmatrix} = E\begin{pmatrix} u_{\mathrm{A}} \\ u_{\mathrm{B}} \end{pmatrix}. \tag{11.34}$$

This corresponds to the two coupled equations

$$[\boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A})]u_{\mathrm{B}} = [E - m + e\phi]u_{\mathrm{A}} \tag{11.35}$$

and

$$[\boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A})]u_{\mathrm{A}} = [E + m + e\phi]u_{\mathrm{B}}. \tag{11.36}$$

Eliminating u_{B} from these equations gives

$$[\boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A})(E + m + e\phi)^{-1}\boldsymbol{\sigma}\cdot(\boldsymbol{p} + e\boldsymbol{A})]u_{\mathrm{A}} = [E - m + e\phi]u_{\mathrm{A}}. \tag{11.37}$$

Now, in the non-relativistic limit, with kinetic energies and field interaction energies small compared with the rest energy m, we have

$$E + m + e\phi \rightarrow 2m$$
$$E - m \rightarrow E_{\mathrm{NR}}$$

where E_{NR} is the non-relativistic energy, so (11.37) approximates to

$$\left[\frac{1}{2m}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) - e\phi\right]u_{\mathrm{A}} = E_{\mathrm{NR}}u_{\mathrm{A}} \tag{11.38}$$

where $\boldsymbol{\pi} \equiv \boldsymbol{p} + e\boldsymbol{A}$.

Now, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are any two three-vectors it can be shown (see example 11.9) that

$$(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b}) = (\boldsymbol{a}\cdot\boldsymbol{b}) + \mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{a} \wedge \boldsymbol{b}).$$

Then, with $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{\pi}$ it follows that

$$\left\{\frac{1}{2m}\left[\boldsymbol{\pi}\cdot\boldsymbol{\pi} + \mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{\pi} \wedge \boldsymbol{\pi})\right] - e\phi\right\}u_{\mathrm{A}} = E_{\mathrm{NR}}u_{\mathrm{A}}. \tag{11.39}$$

In terms of $\boldsymbol{p}$ and $\boldsymbol{A}$ the vector product term is

$$\begin{aligned}(\boldsymbol{\pi} \wedge \boldsymbol{\pi})u_{\mathrm{A}} &= (\boldsymbol{p} + e\boldsymbol{A}) \wedge (\boldsymbol{p} + e\boldsymbol{A})u_{\mathrm{A}} \\ &= (\boldsymbol{p} \times \boldsymbol{p})u_{\mathrm{A}} + e(\boldsymbol{p} \wedge \boldsymbol{A} + \boldsymbol{A} \wedge \boldsymbol{p})u_{\mathrm{A}} + e^2(\boldsymbol{A} \wedge \boldsymbol{A})u_{\mathrm{A}}.\end{aligned}$$

The first and last terms are zero and, with $\boldsymbol{p} \equiv -\mathrm{i}\boldsymbol{\nabla}$, the middle term becomes

$$-e\mathrm{i}[\boldsymbol{\nabla} \wedge (\boldsymbol{A}u_{\mathrm{A}}) + \boldsymbol{A} \wedge (\boldsymbol{\nabla}u_{\mathrm{A}})] = -e\mathrm{i}(\boldsymbol{\nabla} \wedge \boldsymbol{A})u_{\mathrm{A}}$$

since $\boldsymbol{\nabla} \wedge (\boldsymbol{A}u_{\mathrm{A}}) = (\boldsymbol{\nabla} \wedge \boldsymbol{A})u_{\mathrm{A}} - \boldsymbol{A} \wedge (\boldsymbol{\nabla}u_{\mathrm{A}})$. Furthermore, $\boldsymbol{\nabla} \wedge \boldsymbol{A} = \boldsymbol{B}$, so that equation (11.39) becomes

$$\left[\frac{1}{2m}(\boldsymbol{p} + e\boldsymbol{A})^2 + \frac{e}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} - e\phi\right]u_{\mathrm{A}} = E_{\mathrm{NR}}u_{\mathrm{A}}. \tag{11.40}$$

The expression in square brackets is, apart from the term $(e/2m)\boldsymbol{\sigma}\cdot\boldsymbol{B}$, the same as the classical Hamiltonian for a slowly moving electron in an electromagnetic field. This extra term may be interpreted as arising from the interaction of the *magnetic moment of the electron* $-e\boldsymbol{\sigma}/2m$ with the magnetic field $\boldsymbol{B}$.

We wish to show explicitly that this corresponds to an intrinsic spin $\frac{1}{2}\boldsymbol{\sigma}$ for the electron. To do this we work with potentials ϕ and $\boldsymbol{A}$, such that no angular momentum is transferred to the electron, i.e. we consider the situation in which $\boldsymbol{A} = 0$ and ϕ is spherically symmetric – the field is central. With $\boldsymbol{A} = 0$ the Hamiltonian in (11.33) becomes

$$H = \boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m - e\phi. \tag{11.41}$$

We might expect that with a central potential the orbital angular momentum, $\boldsymbol{L} = \boldsymbol{x} \wedge \boldsymbol{p}$, would be conserved. Consider the component

$L_1 = x_2 p_3 - x_3 p_2$; the time rate of change of L_1 is given by

$$\begin{aligned} \mathrm{i}\frac{\mathrm{d}L_1}{\mathrm{d}t} &= [L_1, H] \\ &= \boldsymbol{\alpha}\cdot\{(x_2 p_3 - x_3 p_2)\boldsymbol{p} - \boldsymbol{p}(x_2 p_3 - x_3 p_2)\} \\ &= \mathrm{i}(\alpha_2 p_3 - \alpha_3 p_2) \end{aligned}$$

where we have used the basic commutation relations $[x_i, p_j] = \mathrm{i}\delta_{ij}$. Similar results hold for the components L_2 and L_3. We have, therefore,

$$\mathrm{i}\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \mathrm{i}(\boldsymbol{\alpha} \wedge \boldsymbol{p}) \tag{11.42}$$

and consequently the orbital angular momentum is not a conserved quantity. On physical grounds we might expect that there is some *total* angular momentum which is conserved in a central field and to this end we look for another operator whose commutation relation with the Hamiltonian gives the negative of the right-hand side of equation (11.42). Then the sum of the orbital angular momentum and this new quantity will be conserved. Consider the operator

$$\boldsymbol{\sigma}' = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \tag{11.43}$$

and specifically the component σ_1'. It is a simple matter, using the definitions of $\boldsymbol{\alpha}$ and β and the commutation relations for the Pauli spin matrices, to show that

$$[\sigma_1', \alpha_1] = [\sigma_1', \beta] = 0$$

and

$$\begin{aligned} [\sigma_1', \alpha_2] &= 2\mathrm{i}\alpha_3 \\ [\sigma_1', \alpha_3] &= -2\mathrm{i}\alpha_2. \end{aligned}$$

Then

$$\mathrm{i}\frac{\mathrm{d}\sigma_1'}{\mathrm{d}t} = [\sigma_1', H] = -2\mathrm{i}(\alpha_2 p_3 - \alpha_3 p_2)$$

with similar results for the components σ_2' and σ_3'. Then

$$\mathrm{i}\frac{\mathrm{d}\boldsymbol{\sigma}'}{\mathrm{d}t} = -2\mathrm{i}(\boldsymbol{\alpha} \wedge \boldsymbol{p}). \tag{11.44}$$

It follows that the operator

$$J = L + \tfrac{1}{2}\sigma' \tag{11.45}$$

commutes with the Hamiltonian and may be interpreted as a total angular momentum. The quantity $s = \frac{1}{2}\sigma'$ is interpreted as the intrinsic spin of the electron. We thus come to the important conclusion that according to equations (11.40) and (11.45) an electron has an intrinsic spin $\frac{1}{2}$, associated with which is an intrinsic magnetic moment

$$\mu = -\frac{e\sigma}{2m} = -g\frac{e}{2m}s. \tag{11.46}$$

The Dirac equation, therefore, predicts that the g factor of the electron is $g = 2$. Although experimentally $g - 2$ for the electron is not quite zero, the departure is small (and well understood) and the prediction $g = 2$ must be regarded as a great triumph of the Dirac equation.

11.4.5 Free particle solutions of the Dirac equation

Encouraged by the physical interpretation of the Dirac equation in the non-relativistic limit we turn to the problem of finding solutions in the relativistic case.

For a free particle the spinor equation (11.33) becomes

$$[\alpha\cdot p + \beta m]u(p) = Eu(p). \tag{11.47}$$

First, to enumerate the number of solutions we consider the particle to be at rest, then, recalling the explicit form of the matrix β, (11.47) becomes

$$\begin{pmatrix} mI & 0 \\ 0 & -mI \end{pmatrix} u = Eu \tag{11.48}$$

where as usual I is the 2×2 unit matrix and u is a four-component spinor. Inspection shows that there are four solutions to (11.48),

$$u^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad u^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

both with positive energy $E = m$, and

$$u^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad u^{(4)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

both with negative energy $E = -m$. This situation is reminiscent of that with the Klein–Gordon equation for which both $E > 0$ and $E < 0$ solutions arose. It remains to interpret these negative energy solutions. This we postpone for the time being but note with hindsight from the previous section that the two solutions for a given energy may somehow be connected with the two possible orientations of the spin vector of a spin $\frac{1}{2}$ particle.

For $\boldsymbol{p}$ non-zero (11.47) becomes

$$\begin{pmatrix} m & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -m \end{pmatrix}\begin{pmatrix} u_{\mathrm{A}} \\ u_{\mathrm{B}} \end{pmatrix} = E\begin{pmatrix} u_{\mathrm{A}} \\ u_{\mathrm{B}} \end{pmatrix} \tag{11.49}$$

where, as in the last section, we have written the four-component spinor in terms of the two-component spinors $u_{\mathrm{A}}(\boldsymbol{p})$ and $u_{\mathrm{B}}(\boldsymbol{p})$ and have used the explicit form of the matrices $\boldsymbol{\alpha}$. The matrix equation (11.49) corresponds to the two coupled equations

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}u_{\mathrm{B}} = (E - m)u_{\mathrm{A}} \tag{11.50a}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}u_{\mathrm{A}} = (E + m)u_{\mathrm{B}}. \tag{11.50b}$$

We want to find solutions $u(\boldsymbol{p})$ which reduce to the zero-momentum solutions obtained above as $\boldsymbol{p} \to 0$. We consider first the positive energy solutions and take $u_{\mathrm{A}}^{(s)} = \chi^{(s)}$ with

$$\chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

On substituting into (11.50b) for u_{A} we obtain

$$u_{\mathrm{B}} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi^{(s)}}{E + m}. \tag{11.51}$$

Then, the positive energy solutions are

$$u^{(1)}(p) = N\begin{pmatrix} 1 \\ 0 \\ \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m} \\ 0 \end{pmatrix} \qquad u^{(2)}(p) = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m} \end{pmatrix} \tag{11.52}$$

where N is a normalization constant. The negative energy solutions are obtained in a similar fashion. We take in this case $u_B^{(s)} = \chi^{(s)}$ and write the energy as $-|E|$; then, from (11.50a),

$$u_A = -\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi^{(s)}}{|E|+m} \tag{11.53}$$

and the negative energy solutions are

$$u^{(3)}(p) = N\begin{pmatrix} -\dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|E|+m} \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad u^{(4)}(p) = N\begin{pmatrix} 0 \\ -\dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|E|+m} \\ 0 \\ 1 \end{pmatrix}. \tag{11.54}$$

The reader may readily verify that the spinors $u^{(1)}, \ldots, u^{(4)}$ are orthogonal, i.e. $u^{(r)\dagger}u^{(s)} = 0$ for $r \neq s$. The normalization constant can be determined by setting $u^{(s)\dagger}u^{(s)} = 1$ in which case $N = [(E+m)/2E]^{1/2}$. Frequently, instead of this normalization a covariant normalization is used in which by convention there are considered to be $2E$ particles per unit volume. Then

$$\int \rho \, dV = \int \bar{\psi}\psi \, dV = 2E$$

so that $u^{(s)\dagger}u^{(s)} = 2E$, in which case $N = (E+m)^{1/2}$.

11.4.6 Interpretation of the Dirac spinors

Let us first turn our attention to the states with the same energy and consider further the suggestion in the last section that the internal degree of freedom which might distinguish the states is the spin orientation. One well-defined direction in the problem at hand is the direction of motion of the particle, so we consider first the spin projection in this direction.

It is straightforward to show that the operator

$$\Lambda = \begin{pmatrix} \boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}} \end{pmatrix} \tag{11.55}$$

where $\hat{\boldsymbol{p}} = \boldsymbol{p}/|p|$ is a unit vector in the direction of the momentum, commutes with the Hamiltonian and therefore has eigenvalues which are good quantum numbers and which may be used to distinguish the states. No other component can be found which commutes with the Hamiltonian. Furthermore, since $(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = \boldsymbol{p}^2$ we have $\boldsymbol{\sigma}\cdot\boldsymbol{p} = \pm|\boldsymbol{p}|$, i.e. $\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}} = \pm 1$, and the helicity eigenvalues of the operator Λ are ± 1. For simplicity, consider the case in which the particle moves along the 3 axis. Then,

$$\Lambda = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{11.56}$$

and it is a simple matter to show that

$$\Lambda u^{(1)} = +1 u^{(1)} \qquad \Lambda u^{(2)} = -1 u^{(2)}. \tag{11.57}$$

The positive energy spinors $u^{(1)}$ and $u^{(2)}$ thus describe states of helicity $+1$ and -1 respectively. Similarly,

$$\Lambda u^{(3)} = +1 u^{(3)} \qquad \Lambda u^{(4)} = -1 u^{(4)} \tag{11.58}$$

and the spinors $u^{(3)}$ and $u^{(4)}$ describe negative energy states with helicity $+1$ and -1 respectively. But what is the physical interpretation of these negative energy solutions?

When the Dirac equation is used to calculate, for instance, the energy levels of the hydrogen atom, the results agree well with experiments, but there remains the problem of the stability of the ground state. The existence of negative energy solutions might lead one to expect that radiative transitions could be made from the ground state to the negative energy states; indeed if the totality of negative energy states is included in a calculation of the transition rate the latter becomes infinite. The single-particle interpretation of the Dirac equation therefore needs some reappraisal.

In 1930 Dirac[12] formulated his 'hole' theory to overcome the difficulty. He invoked the Pauli exclusion principle and envisaged the negative energy states to be filled, two to each energy level, with electrons. In this picture the vacuum state is the state with all negative energy levels occupied and all positive energy levels unoccupied, as shown schematically in figure 11.9(a). If an energy $E \geqslant 2m$, where m is the rest energy of the electron, is available for instance in the form of electromagnetic radiation, an electron in the 'sea' of negative energy states may be elevated to a positive energy state leaving a 'hole' in the negative energy sea. This hole, being an *absence* of a negatively charged electron and an *absence* of a negative energy, was interpreted by Dirac as a positively charged 'electron' with a positive energy, i.e. as a positron, the antiparticle of the electron. The net result of the process is electron–positron pair production (figure 11.9(b)). The reverse process (figure 11.9(c)), in which a positive energy electron makes a transition to a vacancy or hole in the negative energy sea, results in the annihilation of the electron–positron pair with the emission of radiation. Dirac thus associated the negative energy solutions with the positron, the two positron solutions being distinguished by the helicity eigenvalues as shown above. Even though there are difficulties connected with this interpretation it was nevertheless a great triumph for the Dirac equation when Anderson[13] discovered the positron in 1932.

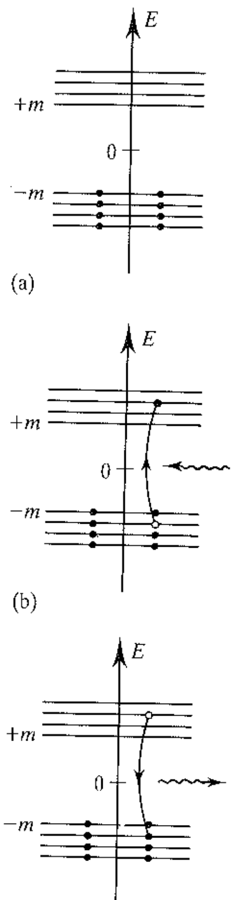


Figure 11.9
(a) The vacuum state in which all negative energy levels are filled with electrons; (b) e^+e^- pair production; (c) e^+e^- annihilation.

Following this success, Pauli and Weisskopf reconsidered the Klein–Gordon equation which we recall had been rejected as a possible relativistic wave equation on the grounds that there were negative energy solutions and the probability density was not positive definite. In relativistic notation the probability density and probability current density (section 11.4.1) become

$$j^\mu = \mathrm{i}(\phi^* \partial^\mu \phi - \phi\, \partial^\mu \phi^*). \tag{11.59}$$

The Pauli–Weisskopf proposal was that this should be regarded as a *charge* current density by incorporating into j^μ the electric charge: thus

$$j^\mu = -e\mathrm{i}(\phi^* \partial^\mu \phi - \phi\, \partial^\mu \phi^*) \tag{11.60}$$

and it becomes physically reasonable that this electron current could be negative. The Klein–Gordon equation does not of course possess the extra degrees of freedom embodied in the Dirac equation, nevertheless it is now regarded as a suitable relativistic equation for spin 0 bosons.

An alternative to the Dirac hole theory is the interpretation due to Stückelberg[14] and Feynman[15] which takes the view that the negative energy solutions describe a particle propagating backward in time or equivalently correspond to a *positive* energy *antiparticle* propagating forward in time. This prescription is commonly used in the evaluation of Feynman diagrams but since this will not concern us we do not develop the idea any further here. The interested reader is referred to the book by Halzen and Martin.[16]

11.5 Application of the Dirac theory to β decay

11.5.1 The Fermi theory

Having established a relativistic description of spin $\frac{1}{2}$ particles we are now in a position to investigate the form of the matrix element for β decay and to consider what restrictions might be imposed on it by the experimental results, discussed in section 11.3, that parity is violated and that neutrinos exist in nature only as 'left-handed' particles, i.e. having helicity -1.

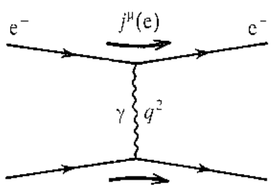


Figure 11.10
Feynman diagram for the electromagnetic interaction between an electron and a proton.

Fermi's original theory of β decay (section 5.2.3) used the well-understood electromagnetic interaction as an analogy. The electromagnetic interaction between an electron and a proton, for example, is described in lowest order by the Feynman diagram (figure 11.10). We can write down the matrix element $\mathcal{M}_{fi}$ for this process by considering the interaction between the electron current j^μ(e) with the electromagnetic field A^μ generated by the proton current j^μ(p):

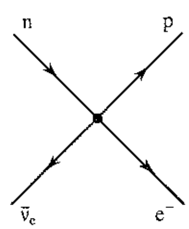
$$\mathcal{M}_{fi} \approx j_\mu(\mathrm{e})A^\mu \approx -\frac{1}{q^2}j_\mu(\mathrm{e})j^\mu(\mathrm{p})$$

where the identification of the electromagnetic potential A^μ with its source j^μ(p) is made through the Maxwell equations

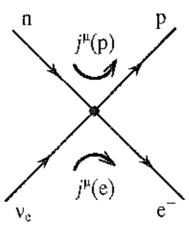
$$\Box^2 A^\mu = j^\mu(\mathrm{p}). \tag{11.61}$$

The term $-1/q^2$, where q^2 is the square of the four-momentum transferred to the photon, is essentially the photon propagator. Thus the interaction is viewed as a current–current interaction and, in terms of the Dirac spinors, the matrix element has the form

$$\mathcal{M}_{fi} \approx e^2\bar{u}_\mathrm{p}\gamma^\mu u_\mathrm{p}\left(-\frac{1}{q^2}\right)\bar{u}_\mathrm{e}\gamma_\mu u_\mathrm{e}. \tag{11.62}$$



(a)



(b)

Figure 11.11
(a) Point-like β decay of the neutron; (b) the crossed reaction $\mathrm{n} + \nu_e \to \mathrm{p} + \mathrm{e}^-$.

The β decay process $\mathrm{n} \to \mathrm{pe}^-\bar{\nu}_e$ was originally assumed to be a 'point' interaction as indicated in figure 11.11(a). The decay is conventionally analysed in terms of the 'crossed' process $\mathrm{n} + \nu_e \to \mathrm{p} + \mathrm{e}^-$ in which the outgoing antineutrino becomes an incoming neutrino (figure 11.11(b)). Fermi assumed that the interaction could be expressed as the product of two *weak* currents j^μ(h) and j^μ(l), where j^μ(h) is the weak hadron current and j^μ(l) the lepton current, and wrote the matrix element as

$$\mathcal{M}_{fi} \approx G(\bar{u}_\mathrm{p}\gamma^\mu u_\mathrm{n})(\bar{u}_\mathrm{e}\gamma_\mu u_\nu). \tag{11.63}$$

Note that since the interaction is regarded as taking place at a point there is no propagator term. In this expression the four-component spinors are interpreted as follows:

$\bar{u}_p$ creates a proton (destroys an antiproton in general)

u_n destroys a neutron (creates an antineutron in general)

$\bar{u}_e$ creates an electron (destroys a positron in general)

u_ν destroys a neutrino (creates an antineutrino in general).

The weak coupling constant G is known as the Fermi constant and, in analogy with (11.62) and in the spirit of the discussion in chapter 7, may be related to the square of the 'weak charge'.

As in the case of the electromagnetic interaction the currents in the matrix element (11.63) are four-vector (or simply vector) currents, i.e. they transform like a four-vector under Lorentz transformations. While the matrix element successfully describes some of the features of β decay it does not give rise either to parity violation or to left-handed neutrinos.

11.5.2 Generalization of the Fermi theory

There is no *a priori* reason why the weak currents should be vector currents; any bilinear covariant of the form $\bar{\psi}O_i\psi$, where the operators O_i are products of γ matrices, is in principle a possible candidate for the weak interactions. In specifying the operators O_i it is convenient to introduce a further γ matrix, γ_5, defined by

$$\gamma^5 \equiv \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{11.64}$$

In the Dirac–Pauli representation (11.17), γ_5 is given explicitly by

$$\gamma_5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{11.65}$$

and it is straightforward to verify that

$$\gamma_5^\dagger = \gamma_5, \qquad (\gamma_5)^2 = I \qquad \text{and} \qquad \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0.$$

The possible forms of O_i are distinguished by their transformation properties under proper Lorentz transformations and space inversion (the parity transformation) and are listed below. The number of components of each type of current is shown in parentheses. For brevity they are generally referred to simply as S, V, T, A and P.

$\bar{\psi}\psi$	scalar(1)	S
$\bar{\psi}\gamma^{\mu}\psi$	vector(4)	V
$\mathrm{i}\bar{\psi}\gamma^{\mu}\gamma^{\nu}\psi \equiv \bar{\psi}\sigma^{\mu\nu}\psi$	tensor(6)	T
$\mathrm{i}\bar{\psi}\gamma^{5}\gamma^{\mu}\psi$	axial vector(4)	A
$\bar{\psi}\gamma^{5}\psi$	pseudoscalar(1)	P

A Lorentz transformation may be regarded as a rotation in four-dimensional space so that $\bar{\psi}\psi$ for example transforms as a scalar (invariant) under Lorentz transformations *and* the parity transformation, whereas $\bar{\psi}\gamma^{5}\psi$ transforms as a scalar under Lorentz transformations but changes sign under the parity transformation and is therefore known as a pseudoscalar. On the other hand $\bar{\psi}\gamma^{\mu}\psi$ transforms as a vector under Lorentz transformations and changes sign under the parity transformation whereas $i\bar{\psi}\gamma^{5}\gamma^{\mu}\psi$ transforms as a vector under Lorentz transformations but is invariant under the parity transformation. It is therefore known as a pseudovector or axial vector.

Let us consider the parity transformation in more detail. Suppose $\psi(x_0, \boldsymbol{x})$ satisfies the Dirac equation (11.26). Under the parity transformation x_0 remains unchanged, $\boldsymbol{x} \to -\boldsymbol{x}$ and $\psi(x_0, \boldsymbol{x}) \to \psi(x_0, -\boldsymbol{x})$ which satisfies

$$\left(\mathrm{i}\gamma^{0}\frac{\partial}{\partial x^{0}} - \mathrm{i}\gamma^{k}\frac{\partial}{\partial x^{k}} - m\right)\psi(x_0, -\boldsymbol{x}) = 0 \tag{11.66}$$

which is *not* the Dirac equation. However, on multiplying from the left by γ^{0} and using the anticommutation relations for the γ matrices, equation (11.66) becomes

$$\left(\mathrm{i}\gamma^{\mu}\frac{\partial}{\partial x^{\mu}} - m\right)\gamma^{0}\psi(x_0, -\boldsymbol{x}) = 0. \tag{11.67}$$

Thus the correct spatially inverted state which satisfies the Dirac equation is $\gamma^{0}\psi(x_0, -\boldsymbol{x})$. If we consider the spinors $u^{(s)}$ in the particle rest frame (section 11.4.5) it can immediately be verified that

$$\gamma^{0}u^{(s)} = u^{(s)} \qquad (s = 1, 2)$$
$$\gamma^{0}u^{(s)} = -u^{(s)} \qquad (s = 3, 4).$$

It follows that in the zero-momentum limit the positive and negative energy spinors are eigenstates of parity with opposite eigenvalues, i.e. the electron and positron have opposite intrinsic parities.

11.5.3 *General form of the weak interaction Hamiltonian*

The matrix element suggested by Fermi for the β decay $n \to pe^- \bar{\nu}_e$, being a contraction of two vector currents, is a Lorentz scalar and therefore cannot account for parity violation. At the time of the Fermi theory there was no reason to suspect that parity was violated in the weak interactions. A more general scalar, or parity-conserving, Hamiltonian can be written down by contracting nucleon and lepton currents with the same properties under the parity transformation. Thus in general

$$\mathcal{M}_{fi} \approx \sum_i C_i(\bar{u}_p O_i u_n)(\bar{u}_e O_i u_\nu) \tag{11.68}$$

where the sum is over the possible forms of the bilinear covariants $i = \mathrm{S, V, T, A, P}$. In order to accommodate parity-violating effects one must add terms to the matrix element which are pseudoscalars, obtained by contracting two covariants which have the opposite behaviour under the parity transformation. The most general pseudoscalar is thus

$$\mathcal{M}_{fi} \approx \sum_i C_i'(\bar{u}_p O_i u_n)(\bar{u}_e O_i \gamma_5 u_\nu) \tag{11.69}$$

with different coefficients C_i' for the parity-violating terms. Combining (11.68) and (11.69) gives the most general β decay Hamiltonian:

$$\mathcal{M}_{fi} \propto \sum_i C_i(\bar{u}_p O_i u_n)\left[\bar{u}_e O_i\left(1 + \frac{C_i'}{C_i}\gamma_5\right)u_\nu\right]. \tag{11.70}$$

If, as the experimental evidence suggests, the β decay process is time-reversal invariant, the coefficients C_i and C_i' must be real. Furthermore, experiment shows that neutrinos and antineutrinos have a definite handedness – parity violation is *maximal* – and this implies that $C_i' = \pm C_i$; scalar and pseudoscalar terms appear in the Hamiltonian with equal magnitude. Thus

$$\mathcal{M}_{fi} = \frac{G}{\sqrt{2}}\sum_i C_i(\bar{u}_p O_i u_n)[u_e O_i(1 \pm \gamma_5)u_\nu] \tag{11.71}$$

is the most general form of the weak interaction Hamiltonian. The factor $\sqrt{2}$ is introduced so that G is defined as in the Fermi theory.

11.5.4 *The two-component neutrino theory*

We turn our attention to the terms $(1 \pm \gamma_5)u_\nu$ in (11.71) and our desire to show that these give rise to neutrinos of a definite handedness. If we persist with the Dirac–Pauli representation of the γ matrices, which we have used so far, it turns out that $(1 + \gamma_5)u_\nu$ represents left-handed neutrinos (right-handed antineutrinos) while $(1 - \gamma_5)u_\nu$ gives rise to right-handed neutrinos (left-handed antineutrinos). It is unfortunate that in the literature the present topic is most frequently discussed using a different representation of the γ matrices. Of course the physical content of the theory is independent of the particular choice of representation, the important requirement being the anticommutation relations of the matrices given in equations (11.25). An alternative to the Dirac–Pauli representation is one in which γ^5 is diagonal. With

$$\alpha_k = \begin{pmatrix} -\sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

we have

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \qquad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \qquad \gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}.$$

We use this representation in what follows.

If we assume that the mass of the neutrino is zero the Dirac equation may be written

$$H\psi = \boldsymbol{\alpha}\cdot\boldsymbol{p}\psi = E\psi. \tag{11.72}$$

Writing the four-component spinor as $\psi = \binom{\chi}{\phi}$ we have

$$\begin{pmatrix} -\boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \end{pmatrix}\begin{pmatrix} \chi \\ \phi \end{pmatrix} = E\begin{pmatrix} \chi \\ \phi \end{pmatrix} \tag{11.73}$$

corresponding to the two decoupled equations

$$-\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi = E\chi \tag{11.74}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\phi = E\phi. \tag{11.75}$$

Since, for a massless neutrino, $E^2 = p^2$, there are two solutions for each of these equations, $E > 0$ corresponding to neutrinos and $E < 0$ corresponding to antineutrinos.

For the positive energy solution to equation (11.74), with $E = |\boldsymbol{p}|$, we have

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|\boldsymbol{p}|}\chi = -\chi$$

which, since $(\boldsymbol{\sigma}\cdot\boldsymbol{p})/|\boldsymbol{p}|$ is the helicity operator, corresponds to a neutrino with helicity -1. For the negative energy solution, $E = -|\boldsymbol{p}|$, we have

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|\boldsymbol{p}|}\chi = +\chi \tag{11.76}$$

which corresponds to an antineutrino with helicity $+1$. Alternatively, the positive energy solution to (11.75) gives rise to right-handed neutrinos (helicity $+1$) and the negative energy solution to left-handed antineutrinos (helicity -1). In the representation of the γ matrices introduced in this section we note that

$$\tfrac{1}{2}(1-\gamma^5)u_\nu = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} \chi \\ 0 \end{pmatrix}, \tag{11.77}$$

i.e. the factor $\frac{1}{2}(1-\gamma^5)$ projects out left-handed neutrino and right-handed antineutrino solutions.

We saw in section 11.3.2 that the neutrino has helicity -1 and thus the appropriate form of the β decay matrix element (11.71) must be

$$\mathcal{M}_{fi} = \frac{G}{\sqrt{2}}\sum_i C_i(\bar{u}_p O_i u_n)[\bar{u}_e O_i(1-\gamma_5)u_\nu]. \tag{11.78}$$

11.5.5 The V–A interaction

It remains to determine the form of O_i in the matrix element (11.78), and to this end we must be guided by a comparison of experimental results with theoretical predictions for the possible forms of O_i. In β decay the pseudoscalar interaction is unimportant; it introduces a factor v/c in the matrix element, where v is the nucleon velocity and typically this is about 10^{-3}. We therefore consider the four possible couplings S, V, T and A. Neither the S nor V couplings can produce a nuclear spin change so they can contribute, in principle, only to Fermi transitions for which $\Delta J = 0$, J being the nuclear spin. If *both* S and V are present there will be an interference term, the so-called Fierz interference, which will affect the shape of the energy spectrum. Compared with that derived in section 5.2.3 the Fierz interference term gives rise to an enhancement in the region of low electron energies. Both the tensor and axial-vector couplings can give rise to a change in nuclear spin and can therefore contribute to Gamow–Teller transitions ($\Delta J = 0, 1$). Again, if both couplings are present the statistical spectrum is modified. Experimental studies of the energy spectra of allowed transitions show that no interference terms are required so we conclude that Fermi transitions are S *or* V coupling and Gamow–Teller transitions are T *or* A.

Which type of coupling is involved can in principle be determined by measuring either the electron–neutrino angular correlation or the helicity of the electron. Using the appropriate matrix element one finds, after summing over the electron and neutrino spin states, that the decay probability depends on the angle θ between the electron and neutrino through a factor of the form

$$I(\theta) = 1 + a\frac{v}{c}\cos\theta$$

where v is the electron velocity and a is a parameter which has the values -1, $+1$, $+\frac{1}{3}$, $-\frac{1}{3}$ for S, V, T and A couplings respectively. In the Gamow–Teller transitions (T or A) the factor $\frac{1}{3}$ arises from the three possible orientations of the total lepton spin ($s = 1$) and this dilutes the correlation.

Since neutrinos have a minute interaction cross-section they are extremely difficult to detect so that experimentally angular correlations between the electron and the recoil nucleus were measured. Even this is notoriously difficult; indeed for some time it was erroneously thought that the interactions were S and T whereas in fact it is now known that they are V and A. With reference to figure 11.12 one can see that for

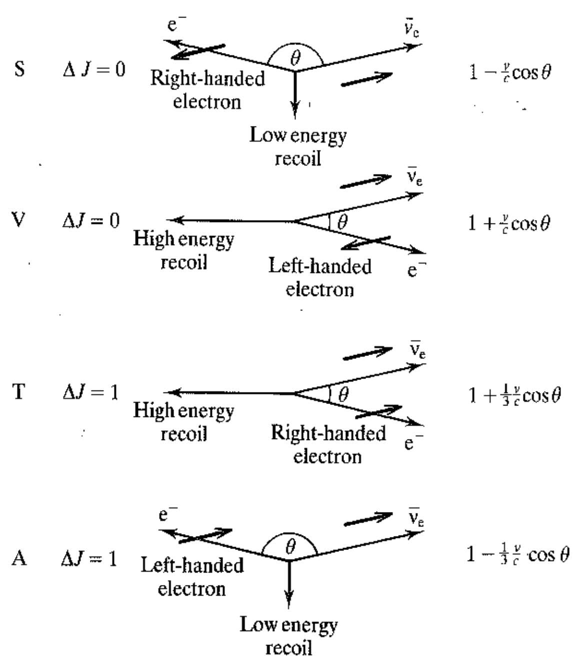


Figure 11.12 Electron–antineutrino angular correlations for the S, V, T and A interactions. The V and A interactions give rise to left-handed electrons while the S and T interactions yield right-handed electrons.

scalar and tensor interactions, assuming that antineutrinos are right handed, the electrons should be longitudinally polarized with, in the relativistic limit, helicity +1. The V and A interactions on the other hand predict an electron helicity −1. Thus the type of coupling in Fermi and Gamow–Teller transitions may be determined by measuring the longitudinal polarization of the electrons. Lack of space prevents us from giving a detailed description of the many techniques which have been used to do this; details may be found in the review article by Grodzins.[8] There is now overwhelming evidence that the electrons produced in nuclear β decay are longitudinally polarized with helicity $-v/c$, i.e. they are left handed; positrons have helicity $+v/c$ and are right handed. The β decay couplings are therefore V and A.

Then, with $O_i = \gamma^\mu$ for the vector coupling and $O_i = i\gamma^\mu\gamma^5$ for the axial vector coupling the matrix element (11.78) becomes

$$\begin{aligned}\mathscr{M}_{fi} &= \frac{G}{\sqrt{2}}\{C_V[\bar{u}_p\gamma^\mu u_n][\bar{u}_e\gamma^\mu(1-\gamma^5)u_\nu] \\ &\quad + C_A[\bar{u}_p i\gamma^\mu\gamma^5 u_n][\bar{u}_e i\gamma^\mu\gamma^5(1-\gamma^5)u_\nu]\} \\ &= \frac{G}{\sqrt{2}}[\bar{u}_p\gamma^\mu(C_V + C_A\gamma^5)u_n][\bar{u}_e\gamma^\mu(1-\gamma^5)u_\nu] \qquad (11.79)\end{aligned}$$

where C_A and C_V are coupling constants, in units of G, for the axial vector and vector interactions.

Let us investigate further the nucleon current, i.e. the first term in square brackets in (11.79). In a pure Fermi transition a measurement of the ft value determines the product GC_V. When this is compared with the value of G determined from a purely leptonic interaction such as μ decay (see section 11.6) there is very good agreement, so that $C_V = 1$. The ratio of the absolute values of C_A and C_V determines the relative strengths of the Fermi and Gamow–Teller couplings and can be determined for instance by comparing the ft values of a pure Fermi transition such as $^{14}O \to {}^{14}N + e^+ + \nu_e$ and a mixed transition such as $n \to p + e^- + \bar{\nu}_e$. In fact

$$\frac{(ft)_{^{14}O}}{(ft)_n} = \frac{C_V^2 + 3C_A^2}{2C_V^2} = \frac{3100 \pm 20}{1080 \pm 16}. \qquad (11.80)$$

The factor 3 arises in the numerator because of the three possible orientations of the total lepton spin in the Gamow–Teller transition (axial vector) and the factor 2 in the denominator arises because ^{14}O, $^{14}N^*$ and ^{14}C form an I spin triplet with $I_3 = 1$ for ^{14}O and $I_3 = 0$ for $^{14}N^*$. The I spin coupling rules give $\langle 1, 0|I_-|1, 1\rangle = \sqrt{2}$ where I_- is the I spin lowering operator.* Equation (11.80) gives

* Less formally, the decay $^{14}O \to {}^{14}N^*$ can arise from two equivalent protons outside the ^{12}C core.

$$\frac{C_A^2}{C_V^2} = 1.58 \pm 0.04 \qquad \left|\frac{C_A}{C_V}\right| = 1.26 \pm 0.02.$$

The sign of the ratio can in principle be determined from the longitudinal polarization of the proton in neutron decay; the factor $(1 \pm C_A/C_V)$ is related to the proton polarization in the same way that $1 \pm \gamma_5$ is related to the helicity of the neutrino. The velocity of recoil of the proton, however, is too small to allow such a measurement. Measurements of the angular correlation of the electrons emitted in the decay of *polarized* neutrons have detected a small anisotropy with respect to the neutron spin direction resulting from destructive interference between the Fermi and Gamow–Teller terms. This implies that the sign of the ratio is negative. With a value of $C_V \approx 1$ the axial vector coupling constant is then $C_A \approx -1.26$. If in (11.79) we put $C_A = -C_V = -1$, we have

$$\mathcal{M}_{fi} = \frac{G}{\sqrt{2}}[\bar{u}_p\gamma^\mu(1-\gamma^5)u_n][\bar{u}_e\gamma^\mu(1-\gamma^5)u_\nu] \qquad (11.81)$$

the so-called V–A interaction. It is interesting to note that, apart from the factors $1-\gamma^5$, the matrix element is the same as that originally proposed by Fermi.

11.6 The universal Fermi interaction

The matrix element for the purely leptonic decay $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$, considered as a four-fermion point interaction as shown in figure 11.13, is

$$\mathcal{M}_{fi} = \frac{G}{\sqrt{2}}[\bar{u}_{\nu_\mu}\gamma^\rho(1-\gamma^5)u_\mu][\bar{u}_e\gamma_\rho(1-\gamma^5)u_{\nu_e}] \qquad (11.82)$$

and is a pure V–A interaction, i.e. the purely leptonic vector and axial vector currents have opposite sign but equal magnitudes. On calculating the phase space factor for the decay $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$ and using the matrix element (11.82) in Fermi's 'golden rule', the muon lifetime is found to be

$$\tau_\mu = \frac{192\pi^3}{G^2 m_\mu^5}. \qquad (11.83)$$

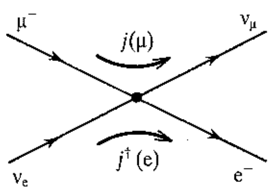


Figure 11.13
The purely leptonic decay $\mu^- \to e^-\bar{\nu}_e\nu_\mu$ viewed as a point interaction between the two leptonic currents.

The experimental values of the muon lifetime and mass when inserted into (11.83) give the weak coupling constant

$$G = (1.4358 \pm 0.0001) \times 10^{-62}\ \mathrm{J\,m^3}. \qquad (11.84)$$

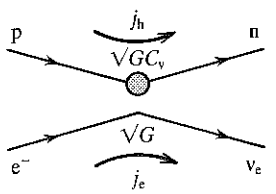


Figure 11.14
A pure Fermi nuclear β decay such as $^{14}O \rightarrow {}^{14}N + e^+ + \nu_e$.

Figure 11.14 represents a pure Fermi nuclear β decay such as $^{14}O \rightarrow {}^{14}N^* + e^+ + \nu_e$ and it is evident that the effective coupling constant which results from a determination of the ft value in such cases is GC_V. The directly observed value of GC_V has been obtained from the pure Fermi superallowed transitions of the mass $(4n + 2)$ nuclei ^{14}O, $^{26}Al^m$, ^{34}Cl, $^{38}K^m$, ^{42}Sc, ^{46}V, ^{50}Mn and ^{54}Co. All these decays are between states with $J^P = 0^+$ and have $I = 1$, $\Delta I = 0$. After the introduction of small charge-dependent corrections the $ft_{1/2}$ values are closely similar and have a mean of 3088 s. This gives a value

$$GC_V = (1.4116 \pm 0.0008) \times 10^{-62}\ \mathrm{J\,m^3}. \quad (11.85)$$

On inserting the value of G obtained from muon decay (equation (11.84)) into this expression we obtain a value 0.98 for C_V. Thus C_V is close but not quite equal to 1. We shall return to this point when we discuss the Cabibbo theory in section 11.10. The near equality of the Fermi constant obtained from an analysis of nuclear β decay, involving hadrons as well as leptons, and that derived from muon decay involving only leptons suggests a *universality* of the weak charge; the value of the weak charge is the same for all particles which possess it.* This is similar to the situation regarding the electromagnetic charge e; the interaction of a proton with an electromagnetic field is the same, apart from the sign, as that of an electron. The so-called universal Fermi interaction[17–20] assigns a single global coupling constant G for the coupling between any four fermion fields.

11.7 The conserved vector current hypothesis

Since the proton and neutron are strongly interacting composite objects and not structureless Dirac particles like the leptons one might expect that their weak couplings would be modified from the purely leptonic value G. We have seen above that indeed $C_A \approx -1.25$ but it is something of a surprise therefore that $C_V \approx 1$. The situation was clarified by Gerstein and Zel'dovitch[21] and independently by Feynman and Gell-Mann[22] in terms of the conserved vector current (CVC) hypothesis.

To gain some insight into this proposal we draw an analogy with electromagnetism, in which the currents are conserved vector currents. The interaction of a 'bare' proton with an electromagnetic field may be represented by the diagram figure 11.15(a). Since there is a finite probability

* Electron–muon universality is strictly true. In section 11.10 we shall see more precisely how universality extends to other particles.

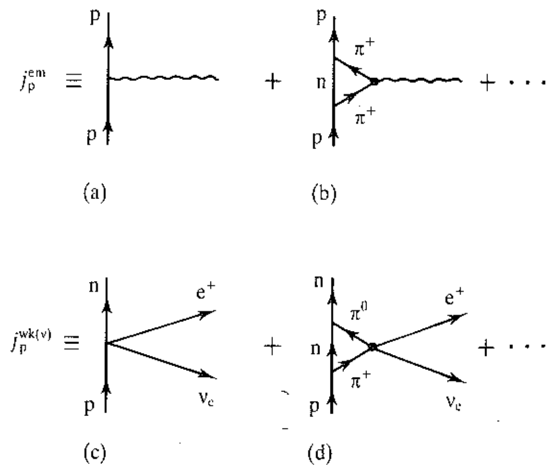


Figure 11.15
Coupling of a proton to the electromagnetic and weak interaction fields: (a) and (c) represent the 'bare' proton and (b) and (d) the modification due to virtual pion emission.

that the proton may emit a pion one might expect that the coupling of the proton to the electromagnetic field would be modified by diagrams such as figure 11.15(b). The electric charge, in fact, is not *renormalized* by the emission of virtual pions; the pion couples to the electromagnetic field with the same strength as the proton, with the result that the electromagnetic current is free of divergence, $\partial_\mu j^\mu(\text{em}) = 0$.

The CVC hypothesis asserts that the vector part of the hadronic weak current is *strictly* analogous to the electromagnetic current and is therefore also free of divergence: it is a conserved vector current. Physically, the situation is as shown in figures 11.15(c) and 11.15(d). One might expect the existence of the situation shown in figure 11.15(d) to alter the weak charge of the proton. As in the electromagnetic case the virtual pion couples to the weak field in such a way that the weak charge of the proton is not renormalized. The virtual pion itself can decay according to

$$\pi^+ \to \pi^0 + e^+ + \nu_e \tag{11.86}$$

for which $\Delta J = 0$ and $\Delta I = 0$, where J and I refer to the spins and isospins of the pions. It is therefore a superallowed Fermi transition in which the axial vector current plays no part. The decay (11.86) has been observed for real pions with the correct basic rate or $ft_{1/2}$ value; the net effect is that the overall decay probability is unaltered by the emission of virtual pions.

11.8 The current–current hypothesis of weak interactions

We have seen that neutron decay is described by the product of two currents, the nucleon current, which is approximately $j^\mu_\mathrm{N} = \bar{u}_\mathrm{p}\gamma^\mu(1-\gamma^5)u_\mathrm{n}$, and the leptonic current, $j_e^\mu = \bar{u}_e\gamma^\mu(1-\gamma^5)u_\nu$. Similarly, muon decay is described by the product of two leptonic currents $\bar{u}_\mu\gamma^\rho(1-\gamma^5)u_\nu$ and $\bar{u}_e\gamma^\rho(1-\gamma^5)u_\nu$. Each current is the sum of a vector and an axial vector current. This description was generalized by Feynman and Gell-Mann[22] to include *all** weak processes. They defined a weak current J^μ which is the sum of a leptonic current

$$J_e^\mu = \bar{u}_e\gamma^\mu(1-\gamma^5)u_{\nu_e} + \text{corresponding terms for other leptons (with equal amplitudes in keeping with universality)}$$

and a hadronic current

$$J_h^\mu = \bar{u}_p\gamma^\mu(1-\gamma^5)u_n + \text{corresponding terms for strange particles.}$$

The current–current hypothesis regards all weak processes as arising from the interaction of the current J^μ with itself:

$$\mathcal{M} \approx \frac{G}{\sqrt{2}} J^\mu J_\mu^\dagger. \tag{11.87}$$

The product (11.87) contains terms corresponding to μ decay, nuclear β decay, electron and muon capture, weak decays of strange particles either to leptons or hadrons and so on. Universality is a consequence of this self-interaction of the weak current but we will find in section 11.10 that important modifications have to be made to account for the different behaviour of the strange and non-strange hadrons.

11.9 *The intermediate boson*

We recall from section 11.5 that the matrix element for the electromagnetic interaction visualized in figure 11.10 is of the form

$$\mathcal{M} \approx j_\mu(\mathrm{e})\frac{1}{q^2}j^\mu(\mathrm{p})$$

where q^2 is the photon propagator. In the weak processes discussed so

* We note that all the weak currents we have discussed so far are charge-changing weak currents n → p, e → ν etc. In their current–current hypothesis Feynman and Gell-Mann deliberately ignored neutral currents for which there was no experimental evidence at the time.

far we have always considered the interaction as point-like, i.e. the four fermion fields interact at the same space–time point. In accordance with this the matrix elements have not included a propagator term. When used to calculate cross-sections for high energy processes serious difficulties are encountered. Consider as an example elastic neutrino–electron scattering, $\nu_e + e^- \rightarrow \nu_e + e^-$. The appropriate diagram, assuming a point-like interaction, is shown in figure 11.16 and is described by the matrix element

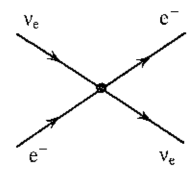


Figure 11.16
Point-like neutrino–electron scattering.

$$\mathcal{M} = \frac{G}{\sqrt{2}}[\bar{u}_e\gamma^\mu(1-\gamma^5)u_{\nu_e}][\bar{u}_{\nu_e}\gamma_\mu(1-\gamma^5)u_e]. \tag{11.88}$$

Using this matrix element and assuming that at very high energies the mass of the electron is negligible, the total cross-section is found to be

$$\sigma(\nu_e e \rightarrow \nu_e e) = \frac{G^2 s}{\pi} \tag{11.89}$$

where s is the square of the centre-of-mass energy.

It can be argued on purely dimensional grounds that the total cross-section for point-like ν_e–e scattering must vary as $G^2 s$. A total cross-section is an invariant quantity and is necessarily proportional to G^2. In natural units with $\hbar = c = 1$ a cross-section has dimensions $[M]^{-2}$ (see appendix F). Example 11.10 shows that in natural units G has dimensions $[M]^{-2}$ and thus to restore the correct dimensions for the total cross-section we must multiply G^2 by a suitable invariant quantity with dimensions $[M]^2$. In the present problem of point-like scattering at high energies the only available quantity is the square of the centre-of-mass energy, hence $\sigma \approx G^2 s$. Indeed the heart of the problem of the divergent cross-section (11.89) is the very fact that the weak coupling constant as defined in the Fermi interaction, unlike the electromagnetic coupling constant, is not a dimensionless quantity. Using arguments based on the partial wave formalism of chapter 9 it is possible to show that the cross-section for a point-like interaction violates unitarity at centre-of-mass energies of approximately 300 GeV. Since we are dealing here with particles which both possess spin the formulae of chapter 9 are not immediately applicable. Nevertheless, if we ignore spin effects equation (9.22) shows that for a given partial wave the maximum elastic cross-section allowed by unitarity is

$$\sigma_{\text{el}}^{\text{max}} = \frac{4\pi}{k^2}(2l+1) = \frac{4\pi}{k^2}$$

for point-like or S wave scattering. This *decreases* as the centre-of-mass energy increases and is clearly in conflict with (11.89). When spins are

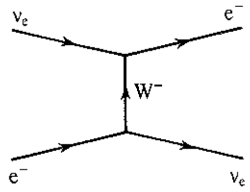


Figure 11.17
Elastic ν_e–e scattering via the exchange of a virtual intermediate boson W^-.

properly taken into account it is found that the Fermi cross-section violates unitarity when $\sqrt{s} \gtrsim G^{-1/2} \approx 300$ GeV.

The divergent behaviour of the point-like cross-section can be avoided if the analogy with the electromagnetic interaction is extended to include a virtual intermediate boson as the mediator of the weak interaction. The appropriate diagram for the scattering process $\nu_e + e^- \rightarrow \nu_e + e^-$ is that shown in figure 11.17 and the matrix element describing this process is

$$\mathcal{M} = \left[\frac{g}{\sqrt{2}}\,\bar{u}_e \gamma^{\mu}\tfrac{1}{2}(1-\gamma^5)u_{\nu_e}\right]\frac{1}{M_W^2 - q^2}\left[\frac{g}{\sqrt{2}}\,\bar{u}_{\nu_e}\gamma_{\mu}\tfrac{1}{2}(1-\gamma^5)u_e\right] \tag{11.90}$$

where $(M_W^2 - q^2)^{-1}$ is the boson propagator. M_W is the mass of the intermediate boson and q^2 is the square of the four-momentum carried by the intermediate boson. In (11.90) we have introduced a *dimensionless* weak coupling constant g and have included the factors $\sqrt{2}$ and $\frac{1}{2}$ in accordance with the conventional definition of g. Because the range of the weak interaction is extremely small the mass of the intermediate boson must be very large. For low energy weak processes such as muon decay and nuclear β decay the values of q^2 are small and for $q^2 \ll M_W^2$ the propagator term in (11.90) vanishes and the process is then essentially point-like. Comparison with (11.88) then gives

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8M_W^2}. \tag{11.91}$$

In chapter 14 we shall see that the measured mass of the W boson is about 80 GeV and equation (11.91) therefore suggests that the apparent weakness of the weak interaction compared with the electromagnetic interaction is due more to the very large mass of the intermediate boson than to the smallness of the weak coupling constant g compared with the electromagnetic coupling constant e.

11.10 *Quark–lepton universality and the Cabibbo theory*

In this section we return to a discussion of the slight discrepancy between the values for the Fermi coupling constant G obtained from measurements of nuclear β decay processes on the one hand and the purely leptonic decay of the muon on the other.

As pointed out already, the weak currents involved in muon decay (figure 11.18(a)) and β decay (figure 11.18(b)) are charge-changing weak currents; in the intermediate boson theory the exchanged boson is charged. In the β decay process the hadrons involved have the same strangeness ($S = 0$) and the hadron current is additionally described as a strangeness-conserving, or $\Delta S = 0$, current.

On a fundamental level the β decay process is $d \to ue^-\bar{\nu}_e$ (figure 11.19(a)); one of the d quarks in the neutron (*ddu*) transforms into a u quark with the remaining u and d quarks acting as 'spectators'. In contrast, in the β decay of the Λ^0, which has quark content *uds*, the strange quark in the Λ^0 transforms into a u quark (figure 11.19(b)). Again, this involves charge-changing weak currents but in this case there is also a change of strangeness at the baryon vertex and the hadronic weak current is therefore called a strangeness-changing, or $\Delta S = 1$, weak current. The quark currents have the same V–A structure encountered earlier, namely

$$J_\mu^{q \to q'} \approx \bar{u}_{q'}\gamma^\mu(1 - \gamma^5)u_q \tag{11.92}$$

and, if the notion of universality introduced in section 11.6 applies equally to the u, d and s quarks, the matrix elements for the neutron and Λ^0 β decay processes would be

$$\mathscr{M}_{d \to u} = \frac{g^2}{\sqrt{2}}\left[\bar{u}_u\gamma^\mu\tfrac{1}{2}(1 - \gamma^5)u_d\right]\left[\bar{u}_e\gamma^\mu\tfrac{1}{2}(1 - \gamma_5)u_{\nu_e}\right] \tag{11.93}$$

and

$$\mathscr{M}_{s \to u} = \frac{g^2}{\sqrt{2}}\left[\bar{u}_u\gamma^\mu\tfrac{1}{2}(1 - \gamma^5)u_s\right]\left[\bar{u}_e\gamma^\mu\tfrac{1}{2}(1 - \gamma_5)u_{\nu_e}\right] \tag{11.94}$$

respectively, with the *same* coupling constant g in each case. The decay rates for these processes, when the different phase space factors are taken into account, should then be equal.

Experimentally, the $\Delta S = 0$ hadronic current, which we have seen in section 11.5.5 to be very slightly weaker than the purely leptonic current, is some 20 times stronger than the $\Delta S = 1$ hadronic current and therefore universality appears to break down when extended to the quark states u, d and s.

In 1963 Cabibbo[23] suggested that universality could be resurrected if the weak interaction quark eigenstates are not the same as the mass or strong interaction eigenstates. He proposed that the weak interaction

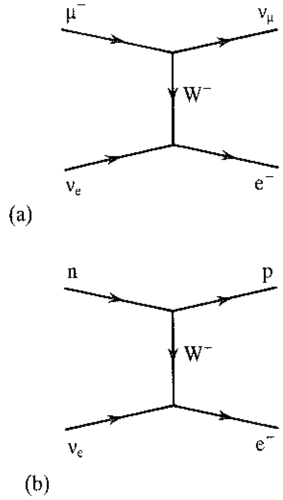


Figure 11.18
The decays
(a) $\mu^- \to e^-\bar{\nu}_e\nu_\mu$ and
(b) $n \to pe^-\bar{\nu}_e$ in the intermediate boson theory.

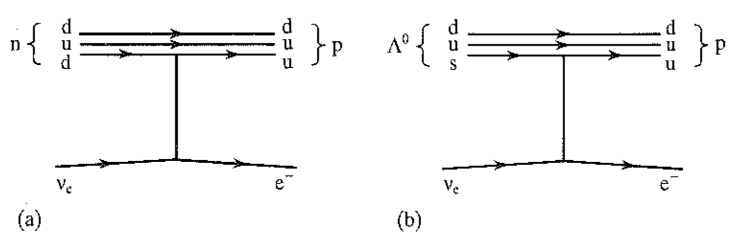


Figure 11.19
(a) Neutron β decay showing the quark transformation $d \to ue^-\bar{\nu}_e$ and (b) Λ^0 β decay, $s \to ue^-\bar{\nu}_e$.

eigenstates are linear superpositions of the strong interaction eigenstates and assigned the quarks to a 'weak isospin' doublet

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ d\cos\theta_C + s\sin\theta_C \end{pmatrix} \tag{11.95}$$

where θ_C is the Cabibbo angle and d' is the Cabibbo 'rotated' quark.* The Cabibbo structure of the charge-raising hadronic current is then

$$\begin{aligned} J_\mu^+(q) &\approx g(\bar{u}, \bar{d}\cos\theta_C + \bar{s}\sin\theta_C)\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u \\ d\cos\theta_C + s\sin\theta_C \end{pmatrix} \\ &= g(\bar{u}d\cos\theta_C + \bar{u}s\sin\theta_C). \end{aligned} \tag{11.96}$$

Note that the matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \tfrac{1}{2}(\tau_1 + \mathrm{i}\tau_2) = \tfrac{1}{2}\tau_+$$

is the (charge) raising operator for an isospin doublet, so that the charge-raising weak current may be written succinctly as

$$J^+(q) \approx g\bar{q}_L\tau_+ q_L \tag{11.97}$$

where

$$q_L = \begin{pmatrix} u \\ d' \end{pmatrix}$$

and the subscript L is included to indicate that, as in the case of the leptons, it is only the left-handed components of the quark currents which couple to the weak field.

In the Cabibbo theory, then, all particles – quarks as well as leptons – carry a weak charge g, but the quarks are mixed, with the consequence that

$$J_\mu^+(q) \approx g\cos\theta_C \text{ for } \Delta S = 0 \text{ currents}$$

and

$$J_\mu^+(q) \approx g\sin\theta_C \text{ for } \Delta S = 1 \text{ currents.}$$

* It is a matter of convention to express the quark mixing between the charge $-\frac{1}{3}$ quarks, leaving the charge $+\frac{2}{3}$ quark unmixed.

The transition rates for the three decays mentioned above then become

$$\begin{aligned} &\Gamma(\mu^- \to e^- \bar{\nu}_e \nu_\mu) \propto g^4 && \text{purely leptonic} \\ &\Gamma(n \to p e^- \bar{\nu}_e) \propto g^4 \cos^2\theta_C && \Delta S = 0 \text{ semi-leptonic} \\ &\Gamma(\Lambda^0 \to p e^- \bar{\nu}_e) \propto g^4 \sin^2\theta_C && \Delta S = 1 \text{ semi-leptonic.} \end{aligned}$$

Thus, for instance, by comparing the rate for nuclear β decay with that for μ decay and taking into account the kinematic factors due to the different masses involved, a value for $\cos\theta_C$ can be determined. Similarly

$$\frac{\Gamma(\Lambda^0 \to p e^- \bar{\nu}_e)}{\Gamma(n \to p e^- \bar{\nu}_e)} \approx \frac{g^4 \sin^2\theta_C}{g^4 \cos^2\theta_C} = \tan^2\theta_C.$$

Data on these and several other decays are consistent with a Cabibbo angle $\theta_C \approx 13°$. Because of the small value of θ_C, those decays which have amplitudes proportional to the $\cos\theta_C$ are known as 'Cabibbo favoured' decays while those with amplitudes proportional to $\sin\theta_C$ are 'Cabibbo suppressed'.

In summary, the Cabibbo theory establishes quark–lepton universality, removes the slight discrepancy between the vector coupling constant determined from nuclear β decay compared with that from μ decay and explains the suppression of $\Delta S = 1$ hadronic currents relative to $\Delta S = 0$ currents.

11.11 *The absence of strangeness-changing neutral currents and the need for charm*

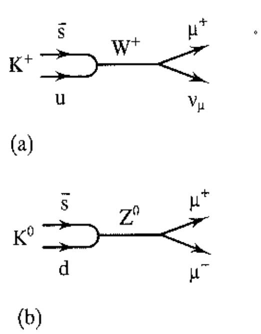


Figure 11.20
(a) Charged-current decay $K^+ \to \mu^+ \nu_\mu$; (b) the highly suppressed neutral-current decay $K^0 \to \mu^+ \mu^-$.

A major problem, the solution of which has far-reaching consequences, was the experimental observation that the decay $K_L^0 \to \mu^+\mu^-$* is suppressed by many orders of magnitude relative to the decay $K^+ \to \mu^+ \nu_\mu$. The quark content of the K^+ is $u\bar{s}$ while that of the K^0 is $d\bar{s}$ and one can visualize the decays taking place via the first-order diagrams shown in figure 11.20. The measured branching fraction[24] for the charged-current process $K^+ \to \mu^+ \nu_\mu$ is 63.51 ± 0.19 per cent while that for the neutral-current process $K_L^0 \to \mu^+\mu^-$ is only $(7.3 \pm 0.4) \times 10^{-7}$ per cent. We note that this neutral current is a *strangeness-changing* neutral current and will be of the form

$$J^0(q) \approx g\bar{q}\tau_3 q \tag{11.98}$$

* The neutral, long-lived kaon state, K_L^0, is defined in section 11.13.4..

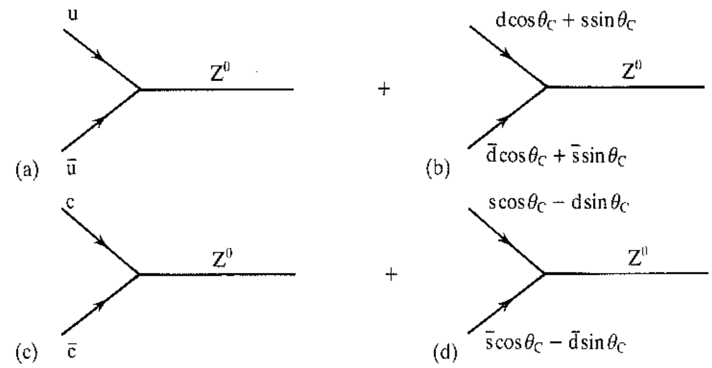


Figure 11.21
(a), (b) The coupling of the Z^0 to u, d and s quarks;
(c), (d) The coupling of the Z^0 to c, d and s quarks.

where τ_3 is the 2×2 matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Thus, according to the Cabibbo theory, the weak coupling of the quarks to the Z^0 boson shown schematically in figure 11.21 is of the form

$$\begin{aligned} J^0(q) &\approx u\bar{u} - d'\bar{d}' \\ &= \underbrace{u\bar{u} - d\bar{d}\cos^2\theta_C - s\bar{s}\sin^2\theta_C}_{\Delta S = 0} - \underbrace{(s\bar{d} + \bar{s}d)\sin\theta_C\cos\theta_C}_{\Delta S = 1}. \end{aligned} \tag{11.99}$$

For clarity we have omitted the factors $\gamma^\mu(1 - \gamma^5)$ in this expression. The last term in (11.99) is a strangeness-changing neutral current and might be expected to be responsible for the decay $K_L^0 \to \mu^+\mu^-$. We have indicated already, however, that this decay is highly suppressed with respect to the charged-current decay $K^+ \to \mu^+\nu_\mu$: the amplitude for the decay should be proportional to $\sin\theta_C\cos\theta_C$, but the calculated rate is many orders of magnitude greater than the actual rate. A way out of the dilemma was proposed in 1970 by Glashow, Iliopoulos and Maiani (GIM).[25] They introduced a new quark, the charmed quark c, with the same electric charge as the u quark, and suggested it belongs to a 'second generation' doublet

$$\begin{bmatrix} c \\ s' \end{bmatrix} = \begin{bmatrix} c \\ s\cos\theta_C - d\sin\theta_C \end{bmatrix} \tag{11.100}$$

in which the weak interaction eigenstate s' is orthogonal to d'. The relationship between the strong and weak quark eigenstates can be visualized as a rotation; the basis states are connected by a rotation

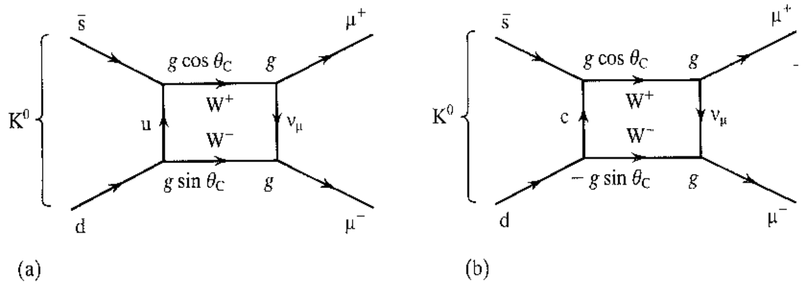


Figure 11.22
Box diagrams showing possible mechanisms for the decay $K_L^0 \to \mu^+\mu^-$: (a) a u-exchange graph; (b) a c-exchange graph.

matrix,

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{11.101}$$

According to the so-called GIM mechanism then, the complete neutral current has, in addition to the couplings shown in figures 11.21(a) and 11.21(b), those shown in figures 11.21(c) and 11.21(d). Explicitly,

$$\begin{aligned} J^0(q) &\approx u\bar{u} - d'\bar{d}' + c\bar{c} - s'\bar{s}' \\ &= \underbrace{u\bar{u} + c\bar{c} - (d\bar{d} + s\bar{s})\cos^2\theta_C - (s\bar{s} + d\bar{d})\sin^2\theta_C}_{\Delta S = 0} \\ &\quad + \underbrace{(s\bar{d} + \bar{s}d - \bar{s}d - s\bar{d})\cos\theta_C \sin\theta_C}_{\Delta S = 1} \\ &= u\bar{u} - d\bar{d} - s\bar{s} + c\bar{c}. \end{aligned} \tag{11.102}$$

Thus, by introducing a fourth quark, the unwanted strangeness-changing terms $s\bar{d}$ and $\bar{s}d$ have been explicitly cancelled. The Z^0 couples directly only to $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and $c\bar{c}$ states. There are no flavour-changing neutral currents.

Although the first-order diagram, involving a single virtual Z^0, figure 11.20(b), makes no contribution to the decay $K_L^0 \to \mu^+\mu^-$, a possible second-order contribution involving two intermediate W bosons is shown in figure 11.22. Again, in the absence of a charmed quark, the decay would proceed via the process shown in figure 11.22(a), the calculated rate for which is much greater than the measured value. The GIM mechanism involving the exchange of a c quark (figure 11.22(b)) suppresses the rate. As can be seen from the weak couplings shown at each vertex in the figure, the amplitudes are

$$\mathcal{M}_u \propto g^4 \cos\theta_C \sin\theta_C \qquad \text{for the u-exchange graph}$$

and

$$\mathcal{M}_c \propto -g^4 \cos\theta_C \sin\theta_C \qquad \text{for the c-exchange graph.}$$

If the masses of the u and c quarks were equal the two diagrams would exactly cancel. In order to obtain agreement with the experimental decay rate, Glashow, Iliopoulos and Maiani predicted that the mass of the charmed quark should lie in the range 1–3 GeV. As we have seen in section 10.7.1 this prediction was strikingly confirmed in 1974, some four years after the GIM mechanism was introduced, by the discovery of the J/ψ meson, a $c\bar{c}$ bound state with a mass $m_{J/\psi} = 3097$ MeV, and the subsequent discovery of open charm states such as the D^0 and $D^{\pm}$ mesons.

11.12 *A third generation of quarks*

Even before charmed quarks had been discovered Kobayashi and Maskawa[26] extended the Cabibbo–GIM scheme to include a *third* generation of quarks – the top t and bottom b quarks. The six quarks are arranged into three weak isospin doublets,

$$\begin{pmatrix} u \\ d' \end{pmatrix}\begin{pmatrix} c \\ s' \end{pmatrix}\begin{pmatrix} t \\ b' \end{pmatrix}$$

and these weak interaction eigenstates are related to the strong interaction eigenstates by the Kobayashi–Maskawa (K–M) mixing matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}\begin{pmatrix} d \\ s \\ b \end{pmatrix}. \qquad (11.103)$$

In this extension to three generations the 2×2 Cabibbo–GIM matrix in equation (11.101) is replaced by the 3×3 K–M 'rotation' matrix in which, for example, the element V_{ud} specifies the coupling of the u and d quarks: $d \to u + W^-$.

The K–M matrix is unitary and can be parametrized in various ways. The Particle Data Group[24] recommends the form

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\exp(-i\delta_{13}) \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\exp(i\delta_{13}) & c_{12}c_{23} - s_{12}s_{23}s_{13}\exp(i\delta_{13}) & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\exp(i\delta_{13}) & -c_{12}s_{23} - s_{12}c_{23}s_{13}\exp(i\delta_{13}) & c_{23}c_{13} \end{pmatrix}. \qquad (11.104)$$

In this expression $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ where $i, j = 1, 2, 3$ are generation labels. In the limit $\theta_{13} = \theta_{23} = 0$ the third generation decouples and if θ_{12} is identified with the Cabibbo angle θ_C the situation reduces to the original Cabibbo mixing of the first two generations. Kobayashi and Maskawa were motivated by a desire to explain *CP* violation within the Cabibbo–GIM scheme. To this end it was necessary to introduce a complex number into the Cabibbo rotation matrix, equation (11.101), but such a term can always be eliminated by a redefinition of the quark phases. They boldly introduced the third generation of quarks and a phase δ_{13} which lies in the range 0–2π with non-zero values giving rise to *CP* violation in the weak interactions. We return to this topic in section 14.3.10. The original appeal of the particular parametrization of the K–M matrix (11.104) is that it is in a form which can be readily generalized to an arbitrary number of generations. Recent results from LEP and SLC strongly suggest that the number of generations is in fact only three.

11.13 CP violation in kaon decay

11.13.1 Neutral kaons

The *I* spin doublet (K^0, K^+) and the antiparticle doublet (K^-, $\bar{K}^0$) should be regarded as strong interaction eigenstates; it is these states which are produced in strong interactions. For example, it is relatively easy to produce K^0s via the reaction $\pi^- p \to \Lambda^0 K^0$ but the production of $\bar{K}^0$ in $\pi^- p$ interactions requires a more exotic process such as $\pi^- p \to \bar{\Sigma}^+ \bar{K}^0 pn$. We note, for what follows, that since the threshold for the first reaction is lower than that for the second it is possible to produce a pure beam of K^0 mesons free from contamination from $\bar{K}^0$.

The kaons are unstable and, since they are the lightest strange particles, they cannot decay via the strangeness-conserving strong interactions; they can decay only via the weak interactions which do not conserve strangeness. In section 11.3.3 we saw that while the weak interactions violate parity and *C*-parity conservation separately they are invariant under the combined operation *CP*. It is therefore reasonable to assume that the states which participate in the weak interaction are *CP* eigenstates and not strangeness eigenstates. Although the K^0 and $\bar{K}^0$ are strangeness eigenstates they are not *CP* eigenstates, as can be seen as follows. Both the K^0 and $\bar{K}^0$ are eigenstates of parity with eigenvalue -1. Thus, we may write

$$P|K^0\rangle = -|K^0\rangle$$

and

$$P|\bar{K}^0\rangle = -|\bar{K}^0\rangle.$$

The charge conjugation operator simply changes particle to antiparticle so that

$$CP|K^0\rangle = -C|K^0\rangle = -|\bar{K}^0\rangle$$

and

$$CP|\bar{K}^0\rangle = -C|\bar{K}^0\rangle = -|K^0\rangle,$$

i.e. neither $|K^0\rangle$ nor $|\bar{K}^0\rangle$ are CP eigenstates. However, the linear combinations

$$|K_1^0\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle) \tag{11.105}$$

and

$$|K_2^0\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle) \tag{11.106}$$

are CP eigenstates with eigenvalues $+1$ and -1 respectively.

If we consider decays to pions both 2π and 3π decay modes are allowed kinematically, so, in order to determine the possible decay modes of the CP eigenstates $|K_1^0\rangle$ and $|K_2^0\rangle$, we must determine the CP eigenvalues of 2π and 3π systems. Consider first the two-pion systems $\pi^0\pi^0$ and $\pi^+\pi^-$. Suppose the relative orbital angular momentum between the two pions π_1 and π_2 is l.* Since the pions have negative intrinsic parity, $P(\pi_1\pi_2) = (-1)^l$. Since the π^0 and its antiparticle are indistinguishable ($B = L = S = Q = 0$) and the π^+ and π^- are particle and antiparticle, the charge conjugation operation is equivalent to the parity operation. Therefore, $C(\pi_1\pi_2) = (-1)^l$ and $CP(\pi_1\pi_2) = +1$.

Consider now the three-pion systems $\pi^0\pi^0\pi^0$ and $\pi^+\pi^-\pi^0$ (figure 11.23). We are free to couple the pions as we choose. Let π_1 and π_2 be either π^+ and π^- or $\pi^0\pi^0$. Kaons, like the pions, have spin $s = 0$ so conservation of angular momentum requires $l = l'$. The small Q values of the decays (≈ 80 MeV) suggest that $l = l' = 0$. Bose–Einstein statistics require that $l = l'$ is even for overall symmetry of the $3\pi^0$ system; values of $l = 2$ will be highly suppressed by angular momentum barrier effects. We assume therefore that in each case the three pions are in a relative S state. By the

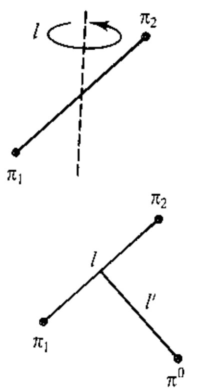


Figure 11.23
Relative orbital angular momenta in 2π and 3π systems.

* Bose–Einstein symmetry limits l to even values but we do not need to use this result.

previous argument $CP(\pi_1\pi_2) = +1$. The π^0 has $C = +1$ and $P = -1$ so when we combine the π^0 with the dipion system $\pi_1\pi_2$ we obtain $CP = -1$ for the overall 3π systems. Thus, in CP-conserving weak decays the state $|K_1^0\rangle$ must decay to two pions and $|K_2^0\rangle$ to three pions. The Q value for the 2π decay is much larger than that for the 3π decay so the phase space available, and hence the decay rate for $K_1 \rightarrow 2\pi$, is much larger than that for $K_2 \rightarrow 3\pi$. Equivalently, the K_2 lifetime is much greater than the K_1 lifetime.

To summarize, neutral kaons produced in strong interactions are the strangeness eigenstates $|K^0\rangle$ and $|\bar{K}^0\rangle$ with $S = +1$ and -1 respectively. The weak interaction eigenstates $|K_1^0\rangle$ with $CP = +1$ and $|K_2^0\rangle$ with $CP = -1$ are distinguished by their lifetimes and decay modes, the former decaying to 2π and the latter to 3π in CP-conserving weak decays.

11.13.2 Strangeness oscillations

A very interesting and important phenomenon, known as strangeness oscillations, occurs in the time evolution of the strong interaction eigenstates. Suppose, at time $t = 0$, we produce a beam of K^0 mesons, for example through the strong interaction process $\pi^- p \rightarrow \Lambda^0 K^0$.

From (11.105) and (11.106) we have

$$|K^0\rangle = \frac{1}{\sqrt{2}}(|K_1^0\rangle + |K_2^0\rangle)$$

and

$$|\bar{K}^0\rangle = \frac{1}{\sqrt{2}}(|K_2^0\rangle - |K_1^0\rangle).$$

At time t the wavefunction $\psi(t)$ is

$$|\psi(t)\rangle = |K^0(t)\rangle = \frac{1}{\sqrt{2}}[|K_1^0(t)\rangle + |K_2^0(t)\rangle].$$

Now, for an unstable particle with mass m and proper lifetime $\tau = 1/\Gamma$, the time dependence of the wavefunction, expressed in the particle rest system in which the total energy $E = m$, is

$$\psi(t) = \psi(0)\exp(-imt)\exp(-\Gamma t/2).$$

Note, this is consistent with the normal exponential decay law for unstable particles,

$$N(t) = |\psi(t)|^2 = |\psi(0)|^2\exp(-\Gamma t) = N(0)\exp(-\Gamma t).$$

Since they have different weak interactions the K_1^0 and K_2^0 will have different masses and decay widths which we denote m_1, Γ_1 and m_2, Γ_2 respectively. Thus,

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\,[|K_1^0(0)\rangle \exp(-\mathrm{i}m_1 t)\exp(-\tfrac{1}{2}\Gamma_1 t) + |K_2^0(0)\rangle \exp(-\mathrm{i}m_2 t)\exp(-\tfrac{1}{2}\Gamma_2 t)].$$

At time t the K^0 intensity in the beam is just $|\langle K^0|\psi(t)\rangle|^2$, so

$$I(K^0) = \tfrac{1}{4}[\exp(-\Gamma_1 t) + \exp(-\Gamma_2 t) + \exp[-\tfrac{1}{2}(\Gamma_1 + \Gamma_2)t]\cos(\Delta m t)$$

where the mass difference $\Delta m = m_2 - m_1$. Similarly the $\bar{K}^0$ intensity at time t is

$$I(\bar{K}^0) = \tfrac{1}{4}[\exp(-\Gamma_1 t) + \exp(-\Gamma_2 t) - \exp[-\tfrac{1}{2}(\Gamma_1 + \Gamma_2)t]\cos(\Delta m t)].$$

Thus the K^0 and $\bar{K}^0$ intensities *oscillate* with frequency $\Delta m/2\pi$.

One can determine the number of $\bar{K}^0$ mesons in a beam, which initially consists of 100 per cent K^0, by virtue of the very different strong interactions of the $S = -1$ $\bar{K}^0$ mesons compared with those of the $S = +1$ K^0 mesons. Since there are no $S = +1$ baryons, K^0 mesons can interact essentially only via elastic or charge exchange scattering whereas the $S = -1$ $\bar{K}^0$ mesons can produce hyperons in reactions such as $\bar{K}^0 p \to \Lambda^0\pi^+$, $\Sigma^0\pi^+$ etc. Thus, by measuring the hyperon yield as a function of the distance from the production target, the mass difference $|\Delta m|$ can be deduced. The current value is

$$|\Delta m|\tau_1 = 0.477 \pm 0.002.$$

The sign of the mass difference has been determined in separate regeneration experiments and it is found that $m_2 > m_1$. The actual mass difference is minute:

$$\Delta m = (0.535 \pm 0.002) \times 10^{10}\,\hbar\,\mathrm{s}^{-1} = (3.52 \pm 0.01) \times 10^{-6}\,\mathrm{eV}.$$

Figure 11.24 shows the variation in intensities $I(K^0)$ and $I(\bar{K}^0)$ as a function of time, in units of τ_1, for a value $\Delta m\tau_1 = 0.5$.

Physically, this strangeness oscillation arises because the K^0–$\bar{K}^0$ system is a coupled system; they are coupled through their common decays to virtual 2π and 3π intermediate states.

At the quark level the strangeness oscillation is visualized as arising from the Cabibbo-mixing of quark flavours. A detailed calculation

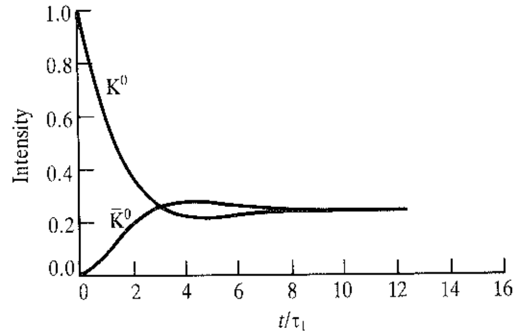


Figure 11.24
Strangeness oscillations. Variation in intensities of K^0 and $\bar{K}^0$ as a function of time in units of τ_1, for a value $\Delta m \tau_1 = 0.5$.

involving u, d, s and c quarks gives a mass difference of

$$\Delta m \approx \frac{G^2}{4\pi^2} f_K^2 m_K m_c^2 \cos^2 \theta_C \sin^2 \theta_C$$

where G is the Fermi constant, $f_k \approx 1.2 m_\pi$ is the kaon decay constant, m_c is the mass of the charmed quark and θ_C is the Cabibbo angle. This expression agrees well with the measured mass difference for $m_c \approx 1.5$ GeV. In this way the mass of the charmed quark was predicted by Gaillard, Lee and Rosner[27] before charmonium states were discovered.

11.13.3 K^0 regeneration

In 1955 Pais and Piccioni[28] pointed out that the existence of the $|K_1^0\rangle$ and $|K_2^0\rangle$ states should give rise to the phenomenon known as *regeneration.*

Suppose we produce a pure beam of K^0 mesons and allow it to coast in vacuum. Initially the beam consists of an equal mixture of the states $|K_1^0\rangle$ and $|K_2^0\rangle$ and

$$|\psi(0)\rangle = |K^0(0)\rangle = \frac{1}{\sqrt{2}}[|K_1^0(0)\rangle + |K_2^0(0)\rangle].$$

For times $t \gg \tau_1$ the short-lived component $|K_1^0\rangle$ will have decayed and the wavefunction will bey

$$|\psi(t)\rangle \approx \frac{1}{\sqrt{2}} |K_2^0(0)\rangle \exp(-\Gamma_2 t),$$

i.e. the beam consists purely of the long-lived component $|K_2^0\rangle$, which we

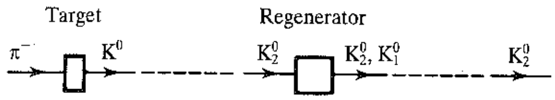


Figure 11.25
K^0 regeneration.

recall is given by

$$|K_2^0\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle).$$

Suppose now we place a block of material in the beam as shown in figure 11.25. As indicated in the last section, K^0 and $\bar{K}^0$ have very different strong interactions; in particular the $S = -1$ $\bar{K}^0$ mesons will be more strongly absorbed in the block. Let f and $\bar{f}$ be the fraction of K^0 and $\bar{K}^0$ mesons remaining in the beam after passing through the block. Since the $\bar{K}^0$ mesons have a higher interaction cross-section we have $\bar{f} < f < 1$. The strangeness content of the beam which emerges from the block is therefore, neglecting the time dependence,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(f|K^0\rangle + \bar{f}|\bar{K}^0\rangle).$$

In terms of the states $|K_1^0\rangle$ and $|K_2^0\rangle$ we have

$$\begin{aligned}|\psi\rangle &= \tfrac{1}{2}[f(|K_1^0\rangle + |K_2^0\rangle) + \bar{f}(|K_2^0\rangle - |K_1^0\rangle)] \\ &= \tfrac{1}{2}[(f-\bar{f})|K_1^0\rangle + (f+\bar{f})|K_2^0\rangle].\end{aligned}$$

Since $f \neq \bar{f}$, it follows that the short-lived state $|K_1^0\rangle$ has been regenerated. This regeneration phenomenon has been confirmed experimentally.

11.13.4 CP violation and the K_L^0–K_S^0 system

The K_2^0 state, which has $CP = -1$, cannot decay into two pions if CP is conserved in the weak interactions. The first experiment to show conclusively that the K_2^0 mesons do in fact decay into two pions, thereby violating CP conservation, was that of Christenson, Cronin, Fitch and Turlay.[29] This discovery led to a change in nomenclature for the neutral kaons. The short-lived state (predominantly $CP = +1$) was henceforth labelled K_S^0 and the long-lived state (predominantly $CP = -1$) was labelled K_L^0.

Figure 11.26 shows a plan view of the apparatus used in the experiment. It was performed at the Brookhaven alternating-gradient synchrotron (AGS), where a beam from an internal Be target was selected at an angle of 30° to the direction of the 30 GeV internal proton beam by a lead collimator. A 4 cm thick lead block, on the target side of the collimator, attenuated γ rays and a bending magnet after the collimator deflected

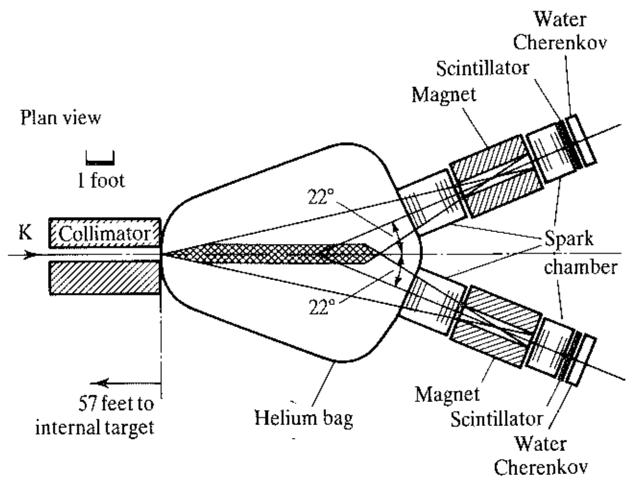


Figure 11.26
Schematic diagram of apparatus used in the discovery of *CP* violation in K^0 decays (Christenson J H *et al.* 1964 *Phys Rev Lett* **13** (138)).

charged particles from the beam. The detector was placed behind a further collimator some 17 m downstream of the target, at which point the beam consisted of the long-lived K^0 mesons (K_2^0), neutrons and some unattenuated γ rays. The detector was a double-armed spectrometer with each arm containing two spark chambers separated by a magnet which provided a momentum measurement. The spark chambers were triggered on a coincidence between the scintillators and water Cherenkov counters placed immediately behind the spectrometers. To minimize interactions, K_2^0 decays were observed in a volume, shown shaded in the figure, contained within a helium-filled bag.

The dominant decay modes of the K_2^0 meson are the three-pion decays $K_{\pi 3}$ and the three-body semi-leptonic decays K_{e3} and $K_{\mu 3}$. The identification of the two-pion decay mode $K_L^0 \to \pi^+\pi^-$ in this large background was performed as follows. For the two-pion decay mode the resultant vector momentum of the decay products should coincide with the direction of the tightly collimated K_2^0 beam and the effective mass of the two pions should be equal to the K^0 mass, within experimental errors. In general, neither of these conditions will be fulfilled for a three-body decay mode. The angular distribution in the very forward direction, for two oppositely charged particles, is shown for three different effective mass ranges in figure 11.27. The pronounced forward peak for the central mass range corresponds to the *CP*-violating decays $K_2^0 \to \pi^+\pi^-$. A three-body decay would not exhibit this strong correlation between K^0 mass and angle. Furthermore, the density of helium in the decay region was insufficient to provide a large enough regeneration of the short-lived K_S^0 to explain the number of events observed. Christenson, Cronin, Fitch and Turlay obtained

$$R = \frac{\Gamma K_L^0 \to \pi^+\pi^-)}{\Gamma(K_L^0 \to \text{all charged})} = (2.0 \pm 0.4) \times 10^{-3}$$

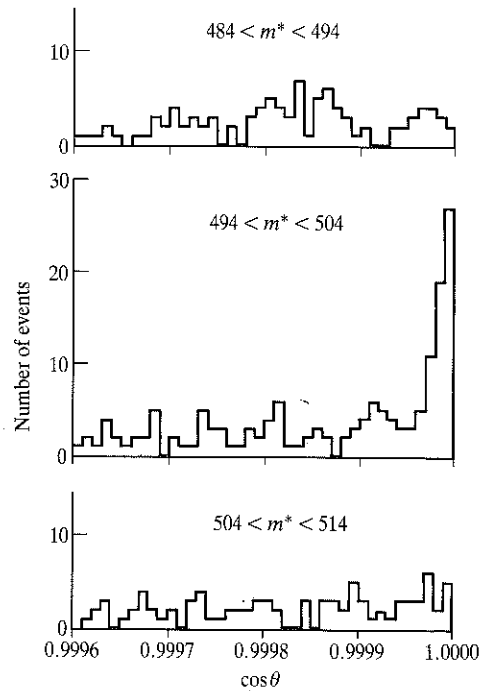


Figure 11.27 Angular distribution of two oppositely charged pions in three mass ranges for events with $\cos\theta > 0.9995$. The strong forward peak in the central mass interval is due to π^+ and π^- from the *CP*-violating decays $K_2^0 \to \pi^+\pi^-$.

and this result has been confirmed in a number of experiments with K_L^0 beams covering a wide range of momenta.

This important discovery means that we cannot identify $|K_S^0\rangle$ with $|K_1^0\rangle$ and $|K_L^0\rangle$ with $|K_2^0\rangle$. Instead, if *CP* is violated, but *CPT* conserved through a corresponding violation of *T*, we have

$$|K_S^0\rangle = \frac{|K_1^0\rangle + \epsilon|K_2^0\rangle}{\sqrt{(1+|\epsilon|^2)}}$$

and

$$|K_L^0\rangle = \frac{|K_2^0\rangle + \epsilon|K_1^0\rangle}{\sqrt{(1+|\epsilon|^2)}}$$

where ϵ is a small complex number which measures the degree of *CP* violation induced by kaon state mixing. *CP* violation in the 2π decay mode is commonly expressed in terms of the amplitude ratios

$$\eta_{+-} = \frac{\langle\pi^+\pi^-|T|K_L^0\rangle}{\langle\pi^+\pi^-|T|K_S^0\rangle} = |\eta_{+-}|\exp(\mathrm{i}\phi_{+-})$$

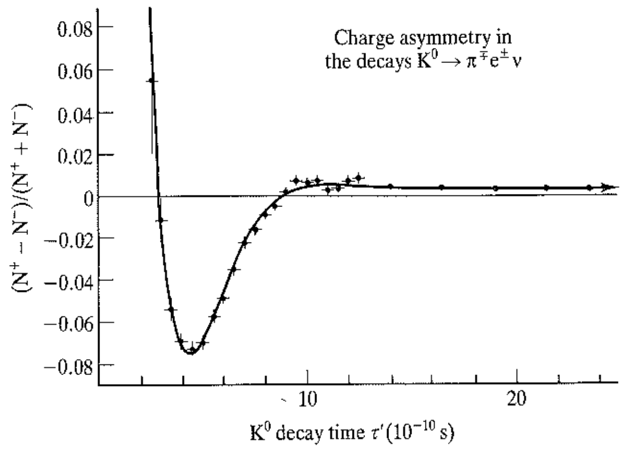


Figure 11.28
Time distribution of the charge asymmetry in the semi-leptonic decays $K^0 \to \pi^{\mp} e^{\pm} \nu$ (after Gjesdal S *et al.* 1974 *Phys Lett* **B52** (113)).

$$\eta_{00} = \frac{\langle \pi^0 \pi^0 | T | K_L^0 \rangle}{\langle \pi^0 \pi^0 | T | K_S^0 \rangle} = |\eta_{00}| \exp(i\phi_{00})$$

and all four quantities $|\eta_{+-}|$, $|\eta_{00}|$, ϕ_{+-} and ϕ_{00} have been measured.[30]

CP violation has also been observed in the semi-leptonic decays of neutral kaons (K_{l3} decays). In the decays

$$K^0 \to l^+ \nu_l \pi^- \qquad (\Delta S = \Delta Q)$$
$$K^0 \to l^- \bar{\nu}_l \pi^+ \qquad (\Delta S = -\Delta Q)$$

where the lepton l is either e or μ, the final states transform into each other under *CP*. Therefore *CP* violation will give rise to a small charge asymmetry, defined as $\delta = (\Gamma^+ - \Gamma^-)/(\Gamma^+ + \Gamma^-)$, where Γ^+ and Γ^- are the rates for the decays $K^0 \to l^+ \nu_l \pi^-$ and $K^0 \to l^- \bar{\nu}_l \pi^+$ respectively. Additionally, this asymmetry will vary with time and show interference effects between the K_L^0 and K_S^0 states. This is clearly demonstrated in figure 11.28 which shows the results of Gjesdal *et al.*[31] for the K_{e3} decays $K^0 \to e^- \bar{\nu}_e \pi^+$ and $K^0 \to e^+ \nu_e \pi^-$.

The current values[24] of the *CP* violation parameters are

$$|\eta_{00}| = (2.253 \pm 0.024) \times 10^{-3} \qquad |\eta_{+-}| = (2.268 \pm 0.023) \times 10^{-3}$$
$$\phi_{00} = (46.6 \pm 2.0)^0 \qquad \phi_{+-} = (46.6 \pm 1.2)^0$$
$$\delta = (0.327 \pm 0.012) \text{ per cent.}$$

The value quoted for δ is a weighted average of the asymmetry parameters for K_{e3} and $K_{\mu 3}$ decays.

11.13.5 Direct CP *violation*

Even in the absence of K_L^0–K_S^0 mixing ($\epsilon = 0$), a mechanism known as direct *CP* violation can result in non-zero values of η_{00} and η_{+-}.

Consider the decays of K_L^0 and K_S^0 into two pions. Bose symmetry restricts the two pions to symmetric *I* spin states; we designate these states $|0\rangle$ and $|2\rangle$ corresponding to $I = 0$ and $I = 2$ respectively. There are thus four amplitudes describing the 2π decays of K_S^0 and K_L^0:

$$\langle 0|H_w|K_S^0\rangle \qquad \langle 2|H_w|K_S^0\rangle$$
$$\langle 0|H_w|K_L^0\rangle \qquad \langle 2|H_w|K_L^0\rangle$$

where H_w is the weak-interaction Hamiltonian responsible for the decays. We can express these amplitudes in terms of the physical pion states $\pi^+\pi^-$ and $\pi^0\pi^0$ using Clebsch–Gordan coefficients and obtain

$$\langle \pi^+\pi^-| = \sqrt{\tfrac{1}{3}}\langle 2| + \sqrt{\tfrac{2}{3}}\langle 0|$$
$$\langle \pi^0\pi^0| = \sqrt{\tfrac{2}{3}}\langle 2| - \sqrt{\tfrac{1}{3}}\langle 0|$$

where by $\langle \pi^+\pi^-|$ we mean the properly symmetrized state

$$2^{-1/2}(\langle \pi_1^+\pi_2^-| + \langle \pi_1^-\pi_2^+|).$$

When the kaons decay the final state pions undergo phase shifts which depend on the *I* spin states of the two pions. Defining $\exp(i\delta_0)$ and $\exp(i\delta_2)$ as the π–π phase shifts at the kaon mass for the $I = 0$ and $I = 2$ states respectively we have

$$\langle \pi^+\pi^-| = \sqrt{\tfrac{1}{3}}\exp(i\delta_2)\langle 2| + \sqrt{\tfrac{2}{3}}\exp(i\delta_0)\langle 0|$$
$$\langle \pi^0\pi^0| = \sqrt{\tfrac{2}{3}}\exp(i\delta_2)\langle 2| - \sqrt{\tfrac{1}{3}}\exp(i\delta_0)\langle 0|.$$

We define the amplitudes

$$A_0 = \langle 0|H_w|K^0\rangle \qquad \text{and} \qquad A_2 = \langle 2|H_w|K^0\rangle$$

from which, assuming *CPT* invariance, we can determine amplitudes for the decays of $\bar{K}^0$. We recall that $CP|K^0\rangle = -|\bar{K}^0\rangle$, therefore $CPT|K^0\rangle = -\langle \bar{K}^0|$. Since any two-pion state has $CP = +1$ we have

$$CPT\langle 0| = |0\rangle$$
$$CPT\langle 2| = |2\rangle.$$

Therefore, if the weak Hamiltonian is *CPT* invariant

$$\langle 0|\mathrm{H_w}|\bar{K}^0\rangle \xrightarrow{CPT} -\langle K^0|\mathrm{H_w}|0\rangle = -A_0^*$$
$$\langle 2|\mathrm{H_w}|\bar{K}^0\rangle \xrightarrow{CPT} -\langle K^0|\mathrm{H_w}|2\rangle = -A_2^*$$

It is conventional to eliminate an overall phase by defining A_0 to be real. Using the expressions

$$|K_\mathrm{S}^0\rangle = \frac{(1+\epsilon)|K^0\rangle - (1-\epsilon)|\bar{K}^0\rangle}{\sqrt{[2(1+|\epsilon|^2)]}}$$
$$|K_\mathrm{L}^0\rangle = \frac{(1+\epsilon)|K^0\rangle + (1-\epsilon)|\bar{K}^0\rangle}{\sqrt{[2(1+|\epsilon|^2)]}}$$

we can express the observed transitions $\mathrm{K_L^0} \to \pi^+\pi^-$, $\pi^0\pi^0$ and $\mathrm{K_S^0} \to \pi^+\pi^-$, $\pi^0\pi^0$ in terms of A_0, A_2 and the *CP*-violating, *CPT*-conserving parameter ϵ, and hence obtain the ratios of η_{+-} and η_{00}. If we neglect second-order terms in the small parameters ϵ and A_2 we obtain (see example 11.15)

$$\eta_{+-} \approx \epsilon + \epsilon'$$

and

$$\eta_{00} \approx \epsilon - 2\epsilon'$$

where

$$\epsilon' = \frac{1}{\sqrt{2}}\frac{\mathrm{Im}\,A_2}{A_0}\exp[\mathrm{i}(\delta_2 - \delta_0)].$$

Observation of a non-zero value of ϵ' would imply the existence of *CP* violation without kaon state mixing, i.e. direct *CP* violation.

Such an observation would be important not only in its own right but also because it would help to distinguish between various models of *CP* violation. The most important of these is the superweak model proposed by Wolfenstein.[32] In this model *CP* violation is presumed to arise through a new $\Delta S = 2$ superweak interaction which transforms $\mathrm{K}^0 \leftrightarrow \bar{\mathrm{K}}^0$. The *CP*-violating $\mathrm{K_L^0} \to 2\pi$ decays are assumed to take place by a two-step process in which the $\mathrm{K_L^0}$ couples to $\mathrm{K_S^0}$ through the superweak interaction, which then decays to 2π through the normal weak interaction. Because of the very small mass difference between $\mathrm{K_L^0}$ and $\mathrm{K_S^0}$ a tiny perturbation can induce the transition $\mathrm{K_L^0} \to \mathrm{K_S^0}$. The superweak coupling is required to be only of the order of 10^{-10} of the normal weak coupling to explain the observed (state mixing) *CP* violation. The superweak model has the advantage that it makes definite predictions for the parameters used in the description of *CP* violation. These are:

(a) $\epsilon' = 0$ and consequently $\left|\frac{\eta_{00}}{\eta_{+-}}\right| = 1$

(b) $\phi_{+-} = \phi_{00} = \tan^{-1}\left(\frac{2\Delta m \tau_s}{\hbar}\right)$

(c) $\mathrm{Re}\,\epsilon = |\eta_{+-}|\left[1 + \left(\frac{2\Delta m \tau_s}{\hbar}\right)^2\right]^{-1/2} \approx \frac{\delta}{2}$.

Using the most recent values,[24] $\Delta m = (0.5351 \pm 0.0024) \times 10^{10}\,\hbar\,\mathrm{s}^{-1}$, $\tau_S = (0.8922 \pm 0.0020) \times 10^{-10}$ s and $|\eta_{+-}| = (2.268 \pm 0.023) \times 10^{-3}$, the superweak model predicts

$$\phi_{+-} = \phi_{00} = (43.68 \pm 0.14)^0 \qquad \text{and}$$
$$\mathrm{Re}\,\epsilon = (1.648 \pm 0.015) \times 10^{-3}.$$

These predictions compare favourably with the experimental values

$$\left|\frac{\eta_{00}}{\eta_{+-}}\right| = 0.9935 \pm 0.0032 \qquad \phi_{+-} = (46.6 \pm 1.2)^0$$
$$\phi_{00} = (46.6 \pm 2)^0 \qquad \mathrm{Re}\,\epsilon = (1.635 \pm 0.006) \times 10^{-3}.$$

Evidence for direct *CP* violation was first observed in an experiment[33] at CERN where the double ratio

$$R = \left|\frac{\eta_{00}}{\eta_{+-}}\right|^2 = \frac{\Gamma(\mathrm{K}_L^0 \to \pi^0\pi^0)/\Gamma(\mathrm{K}_L^0 \to \pi^+\pi^-)}{\Gamma(\mathrm{K}_S^0 \to \pi^0\pi^0)/\Gamma(\mathrm{K}_S^0 \to \pi^+\pi^-)}$$

was measured to be $0.980 \pm 0.004 \pm 0.005$; the uncertainties are statistical and systematic respectively. The deviation of *R* from unity implies that $\epsilon' \neq 0$ and corresponds to a ratio $\epsilon'/\epsilon = (3.3 \pm 1.1) \times 10^{-3}$, a three-standard deviation departure from the prediction of the superweak model. However, measurements of this ratio at Fermilab are consistent with zero.

In chapter 14 we shall see that direct *CP* violation as well as that induced by kaon state mixing can be accommodated in the standard model of electroweak interactions by transitions involving heavy-quark intermediate states.

REFERENCES 11

1 **Perl M L** *et al.* 1975 *Phys Rev Lett* **35** (1489)

2 **Perl M L** *et al.* 1976 *Phys Lett* **63B** (366)

3 **Lee T D and Yang C N** 1956 *Phys Rev* **104** (254)

4 **Wu C S, Ambler E, Haywood R W, Hoppes D D and Hudson R P** 1957 *Phys Rev* **105** (1413)

5 **Gorter C J 1948** *Physica* **14** (504)

6 **Rose M E** 1949 *Phys Rev* **75** (213)

7 **Goldhaber M, Grodzins L and Sunyar A W 1958** *Phys Rev* **109** (1015)

8 **Grodzins L** 1959 *Prog Nucl Phys* **7** (163)

9 **Burgy M T, Krohn V E, Novey T B, Ringo G R and Telegdi V A** 1958 *Phys Rev* **110** (1214)

10 **Garwin R L, Lederman L M and Weinrich M** 1957 *Phys Rev* **105** (1415)

11 **Dirac P A M** 1928 *Proc Roy Soc* (*London*) **A117** (610) and **A118** (351)

12 **Dirac P A M** 1930 *Proc. Roy Soc* (*London*) **A126** (360)

13 **Anderson C D** 1933 *Phys Rev* **43** (491)

14 **Stueckelberg E C G** 1941 *Helv Phys Acta* **14** (588)

15 **Feynman R P** 1949 *Phys Rev* **76** (749, 769)

16 **Halzen F and Martin A D** 1984 *Quarks and Leptons: An Introductory Course in Modern Particle Physics* Wiley

17 **Klein O** 1948 *Nature* **161** (897)

18 **Puppi G** 1948 *Nuovo Cimento* **5** (587)

19 **Lee T D, Rosenbluth M and Yang C N** 1949 *Phys Rev* **75** (905)

20 **Timono J and Wheeler J A** 1949 *Rev Mod Phys* **21** (144)

21 **Gerstein S S and Zel'dovitch Ya B** 1956 *Soviet Phys JETP* **2** (579)

22 **Feynman R P and Gell-Mann M** 1958 *Phys Rev* **109** (193)

23 **Cabibbo N** 1963 *Phys Rev Lett* **10** (531)

24 **Particle Data Group** 1992 'Review of particle properties' *Phys Rev* **D45**

25 **Glashow S L, Iliopoulos J and Maiani L** 1970 *Phys Rev* **D2** (1585)

26 **Kobayashi M and Maskawa K** 1973 *Prog Theor Phys* **49** (652)

27 **Gaillard M K, Lee B W and Rosner J L** 1975 *Rev Mod Phys* **47** (277)

28 **Pais A and Piccioni O** 1955 *Phys Rev* **100** (1487)

29 **Christenson J H, Cronin J, Fitch V and Turlay R** 1964 *Phys Rev Lett* **13** (138)

30 **Christenson J H** *et al.* 1979 *Phys Rev Lett* **43** (1209, 1212)

31 **Gjesdal S** *et al.* 1974 *Phys Lett* **B52** (113)

32 **Wolfenstein L** 1964 *Phys Rev Lett* **13** (562)

33 **Burkhardt H** *et al.* 1988 *Phys Lett* **B206** (169)

EXAMPLES 11

11.1 The helicity of a particle with spin $\boldsymbol{s}$ is defined as $\lambda = \boldsymbol{s}\cdot\hat{\boldsymbol{p}}$ where $\hat{\boldsymbol{p}}$ is a unit vector in the direction of the particle momentum. By considering how the helicity expectation value behaves under the parity transformation show that $\langle\lambda\rangle$ must be zero in a parity-conserving interaction.

11.2 ^{152}Eu can decay via K electron capture to an excited state of samarium (energy E_0) according to the reaction $e^- + {}^{152}\text{Eu} \rightarrow {}^{152}\text{Sm}^* + \nu_e$. If the neutrino has an energy E_ν show that the recoil velocity of the ^{152}Sm* nucleus is $v/c = E_\nu/Mc^2$, where M is the mass of the samarium nucleus. Suppose the recoiling ^{152}Sm* nucleus emits a γ ray at an angle θ with respect to its direction of motion. Determine the energy of this γ ray and hence show that the condition for resonant absorption to the level E_0 is $E_\nu \cos\theta = E_0$.

11.3 In the parity-violating weak decay $\Lambda \rightarrow p\pi^-$, both S and P waves are present. Show that the decay angular distribution relative to the Λ spin direction is

$$W(\theta) = \tfrac{1}{4}(|S|^2 + |P|^2)\left(1 - \frac{2\,\text{Re}\,S^*P}{|S|^2 + |P|^2}\cos\theta\right)$$

where S and P denote the (complex) amplitudes for S and P wave decay respectively.

11.4 Obtain equation (11.2) in the text by multiplying the Klein–Gordon equation by $-i\phi^*$ and its complex conjugate by $-i\phi$ and subtracting. Write down a plane wave solution to the Klein–Gordon equation for a free particle with energy E and momentum $\boldsymbol{p}$ and show that for this solution $\rho = 2|N|^2E$ and $\boldsymbol{j} = 2|N|^2\boldsymbol{p}$, where N is a normalization constant.

11.5 Use the definitions of α_i and β given in equation (11.17) to verify that

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}I \qquad \{\alpha_i, \beta\} = 0$$

where I is the unit matrix.

11.6 Use the results of the last example and the definitions $\gamma^0 = \beta$, $\gamma^k = \beta\alpha_k$ ($k = 1, 2, 3$), to verify the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I \qquad (\mu, \nu = 0, 1, 2, 3).$$