1.5 Free particle wave equations

The relativistic relation between energy, momentum and mass is given in (1.9):

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

If we replace the quantities E and p by the quantum mechanical operators

$$E_{\text{op}} = i\hbar\frac{\partial}{\partial t}, \quad p_{\text{op}} = -i\hbar\nabla = -i\hbar\frac{\partial}{\partial \mathbf{r}} \tag{1.12}$$

where $\mathbf{r}$ is the position vector, we get the *Klein–Gordon wave equation*

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \nabla^2\psi - \frac{m^2c^2}{\hbar^2}\psi \tag{1.13}$$

As described above, it is often more convenient to work in units such that $\hbar = c = 1$, in order to avoid writing these symbols repeatedly, so that the above equation becomes

$$\frac{\partial^2\psi}{\partial t^2} = (\nabla^2 - m^2)\psi \tag{1.14}$$

This wave equation is suitable for describing spinless (or scalar) bosons (since no spin variable has been introduced). In the non-relativistic case, if we define $E = p^2/(2m)$ as the kinetic energy rather than the total energy then substituting the above operators gives the *Schrödinger wave equation* for non-relativistic spinless particles:

$$\frac{\partial\psi}{\partial t} - \frac{i}{2m}\nabla^2\psi = 0 \tag{1.15}$$

Note that the Klein–Gordon equation is second order in the derivatives, while the Schrödinger equation is first order in time and second order in space. This is unsatisfactory when we are dealing with high energy particles, where the description of physical processes must be relativistically invariant, with space and time coordinates occurring to the same power.

Dirac set out to formulate a wave equation symmetric in space and time, which was *first order in both derivatives*. The simplest form that can be written down is that for massless particles, in the form of the *Weyl equations*

$$\frac{\partial\psi}{\partial t} = \pm\left(\sigma_1\frac{\partial\psi}{\partial x} + \sigma_2\frac{\partial\psi}{\partial y} + \sigma_3\frac{\partial\psi}{\partial z}\right) = \pm\boldsymbol{\sigma}\cdot\frac{\partial}{\partial\mathbf{r}}\psi \tag{1.16}$$

Here the σ's are unknown constants. In order to satisfy the Klein–Gordon equation (1.14), we square (1.16) and equate coefficients, whence we find

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$$

$$\sigma_1\sigma_2 + \sigma_2\sigma_1 = 0, \quad \text{etc.} \tag{1.17}$$

$$m = 0$$

These results hold for either sign on the right-hand side of (1.16), and both must be considered. The σ's cannot be numbers since they do not commute, but they can be represented by matrices, in fact the equations (1.17) define the 2×2 Pauli matrices, which we know from atomic physics to be associated with the description of the spin quantum number of the electron:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.18}$$

Using (1.12) we can also express (1.16) in the forms

$$E\chi = -\boldsymbol{\sigma}\cdot\mathbf{p}\chi \tag{1.19a}$$

$$E\phi = +\boldsymbol{\sigma}\cdot\mathbf{p}\phi \tag{1.19b}$$

where E and $\mathbf{p}$ are the energy and momentum operators. χ and ϕ are two-component wavefunctions, called spinors, and are separate solutions of the two Weyl equations, and $\boldsymbol{\sigma}$ denotes the Pauli spin vector, with Cartesian components $\sigma_1, \sigma_2, \sigma_3$ as above. As indicated below, the two Weyl equations have in total four solutions, corresponding to particle and antiparticle states with two spin substates of each.

If the fermion mass is now included, we need to enlarge (1.16) or (1.19) by including a mass term, giving the *Dirac equation*,

$$E\psi = (\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)\psi \tag{1.20a}$$

Here, the matrices $\boldsymbol{\alpha}$ and β are 4×4 matrices, operating on four-component

(spinor) wavefunctions (particle, antiparticle and two spin substates for each). The matrices $\boldsymbol{\alpha}$ and β are

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where each element denotes a 2×2 matrix and '1' denotes the unit 2×2 matrix. The matrix $\boldsymbol{\alpha}$ has three components, just as does $\boldsymbol{\sigma}$ in (1.18). Here, we have quoted the so-called Dirac–Pauli representation of these matrices, but other representations are possible.

Usually, the Dirac equation is quoted in a covariant form, using (1.12) in (1.20a), as

$$\left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m \right) \psi = 0 \tag{1.20b}$$

where the γ_μ (with $\mu = 1, 2, 3, 4$) are 4×4 matrices related to those above. In fact

$$\gamma_k = \beta\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}, \quad k = 1, 2, 3 \quad \text{and} \quad \gamma_4 = \beta \tag{1.20c}$$

The Dirac equation is fully discussed in books on relativistic quantum mechanics, and we have mentioned it here merely for completeness; we shall not discuss it in detail in this text. Occasionally we shall need to quote results from the Dirac equation without derivation. However, it turns out that, in most of the applications with which we shall be dealing in high energy physics, the fermions have extreme relativistic velocities so that the masses can be neglected and the Dirac equation breaks down into the two much simpler, decoupled, Weyl equations as described above.

1.6 Helicity states: helicity conservation

For a massless fermion of positive energy, $E = |\mathbf{p}|$ so that (1.19a) satisfies

$$\frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} \chi = -\chi \tag{1.21}$$

The quantity

$$H = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} = -1 \tag{1.22}$$

is called the *helicity* (or handedness). It measures the sign of the component of spin of the particle, $j_z = \pm\frac{1}{2}\hbar$, in the direction of motion (z-direction). The z-component of spin and the momentum vector $\mathbf{p}$ together define a screw sense, as in Figure 1.5. $H = +1$ corresponds to a right-handed (RH) screw, while particles with $H = -1$ are left-handed (LH).

The solution χ of (1.19a) represents a LH, positive energy, particle but it can also represent a particle with negative energy $-E$ and momentum $-\mathbf{p}$. Thus $-E\chi = -\boldsymbol{\sigma} \cdot (-\mathbf{p})\chi$, or $H = \boldsymbol{\sigma} \cdot (-\mathbf{p})/|\mathbf{p}| = +1$. This state is interpreted, as before, as that of the antiparticle. Thus, (1.19a) represents *either* a LH particle *or* a RH antiparticle, while the independent solution (1.19b) corresponds to a RH particle or a LH antiparticle state.

Helicity is a well-defined, Lorentz-invariant quantity for a massless particle, for the simple reason that such a particle travels at velocity c. In making a Lorentz transformation to another reference frame of relative velocity $v < c$, it is therefore impossible to reverse the helicity. As discussed below, neutrinos have very small, possibly even zero, masses, and are well described by one of the two Weyl equations. By contrast, it turns out that solutions of the Dirac equation (1.20), with its finite mass term, are *not* pure helicity eigenstates but some admixture of LH and RH functions. However, provided they are extreme relativistic, massive fermions (electrons for example) can also be described well enough by the Weyl equations.

For interactions involving *vector* or *axial vector fields*, i.e. those mediated by vector or axial vector bosons, *helicity is conserved in the relativistic limit.* The reason is that such interactions do not mix the separate LH and RH solutions of the Weyl equations. This means for example that a LH lepton, undergoing scattering in such an interaction, will emerge as a LH particle, irrespective of the angle of scatter, provided it is extreme relativistic. On the other hand, a *scalar* interaction does not preserve the helicity and does mix LH and RH states. In the Dirac equation, the mass term represents such a scalar-type interaction and because of its presence, massive leptons with v less than c are superpositions of LH and RH helicity states. In the successful theory of electroweak interactions discussed in Chapter 8, the elementary leptons and bosons start out as massless particles. Scalar field particles, called Higgs bosons, are associated with an all-pervading scalar field which is postulated to interact with, and give mass to, these hitherto massless objects.

Helicity conservation holds good in the relativistic limit for any interaction that has the Lorentz transformation properties of a vector or axial vector, and it therefore applies to strong, weak and electromagnetic interactions, which are all mediated by vector or axial vector bosons. Consequently, in a scattering process at high energy, e.g. of a quark by a quark or a lepton by a quark or lepton, a LH particle remains LH, and a RH particle remains RH. This fact, together with the conservation of angular momentum, determines angular distributions in many interactions, as described later in the text.

Dirac Equation

- **H of massive particle has an extra term βm which mixes two Weyl equations - $\Psi=\Psi_R+\Psi_L$ – 4-component spinors**

$$\hat{H}\psi = (\vec{\alpha}.\vec{p}+\beta m)\psi = i\frac{\partial\psi}{\partial t} \qquad \textbf{(D1)}$$

where $\hat{H}$ is the Hamiltonian operator and, as usual, $\vec{p} = -i\vec{\nabla}$

- **Writing (D1) in full:**

$$\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\psi = \left(i\frac{\partial}{\partial t}\right)\psi$$

"squaring" this equation gives

$$\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\psi = -\frac{\partial^2\psi}{\partial t^2}$$

Expanding

$$\begin{aligned}
-\frac{\partial^2\psi}{\partial t^2} \quad = \quad & -\alpha_x^2\frac{\partial^2\psi}{\partial x^2} - \alpha_y^2\frac{\partial^2\psi}{\partial y^2} - \alpha_z^2\frac{\partial^2\psi}{\partial z^2} + \beta^2 m^2\psi \\
& -(\alpha_x\alpha_y + \alpha_y\alpha_x)\frac{\partial^2\psi}{\partial x\partial y} - (\alpha_y\alpha_z + \alpha_z\alpha_y)\frac{\partial^2\psi}{\partial y\partial z} - (\alpha_z\alpha_x + \alpha_x\alpha_z)\frac{\partial^2\psi}{\partial z\partial x} \\
& -(\alpha_x\beta + \beta\alpha_x)m\frac{\partial\psi}{\partial x} - (\alpha_y\beta + \beta\alpha_y)m\frac{\partial\psi}{\partial y} - (\alpha_z\beta + \beta\alpha_z)m\frac{\partial\psi}{\partial z}
\end{aligned}$$

- **For this to be a reasonable formulation of relativistic QM, a free particle must also obey $E^2 = \vec{p}^2 + m^2$, i.e. it must satisfy the Klein-Gordon equation:**

$$-\frac{\partial^2\psi}{\partial t^2} = -\frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial y^2} - \frac{\partial^2\psi}{\partial z^2} + m^2\psi$$

- **Hence for the Dirac Equation to be consistent with the KG equation require:**

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1 \qquad \textbf{(D2)}$$

$$\alpha_j\beta + \beta\alpha_j = 0 \qquad \textbf{(D3)}$$

$$\alpha_j\alpha_k + \alpha_k\alpha_j = 0 \quad (j \neq k) \qquad \textbf{(D4)}$$

★ **Immediately we see that the α_j and β cannot be numbers. Require 4 mutually anti-commuting matrices**

★ **Must be (at least) 4x4 matrices**

Probability Density and Current

- Now consider probability density/current – this is where the perceived problems with the Klein-Gordon equation arose.
- Start with the Dirac equation

$$-i\alpha_x\frac{\partial\psi}{\partial x}-i\alpha_y\frac{\partial\psi}{\partial y}-i\alpha_z\frac{\partial\psi}{\partial z}+m\beta\psi=i\frac{\partial\psi}{\partial t} \qquad \textbf{(D6)}$$

and its Hermitian conjugate

$$+i\frac{\partial\psi^\dagger}{\partial x}\alpha_x^\dagger+i\frac{\partial\psi^\dagger}{\partial y}\alpha_y^\dagger+i\frac{\partial\psi^\dagger}{\partial z}\alpha_z^\dagger+m\psi^\dagger\beta^\dagger=-i\frac{\partial\psi^\dagger}{\partial t} \qquad \textbf{(D7)}$$

- Consider $\psi^\dagger\times(\mathbf{D6})-(\mathbf{D7})\times\psi$ remembering α,β are Hermitian ➡

$$\psi^\dagger\left(-i\alpha_x\frac{\partial\psi}{\partial x}-i\alpha_y\frac{\partial\psi}{\partial y}-i\alpha_z\frac{\partial\psi}{\partial z}+\beta m\psi\right)-\left(i\frac{\partial\psi^\dagger}{\partial x}\alpha_x+i\frac{\partial\psi^\dagger}{\partial y}\alpha_y+i\frac{\partial\psi^\dagger}{\partial z}\alpha_z+m\psi^\dagger\beta\right)\psi=i\psi^\dagger\frac{\partial\psi}{\partial t}+i\frac{\partial\psi^\dagger}{\partial t}\psi$$

$$\Rightarrow\ \underbrace{\psi^\dagger\left(\alpha_x\frac{\partial\psi}{\partial x}\right.}+\left.\alpha_y\frac{\partial\psi}{\partial y}+\alpha_z\frac{\partial\psi}{\partial z}\right)+\left(\underbrace{\frac{\partial\psi^\dagger}{\partial x}\alpha_x}+\frac{\partial\psi^\dagger}{\partial y}\alpha_y+\frac{\partial\psi^\dagger}{\partial z}\alpha_z\right)\psi+\frac{\partial(\psi^\dagger\psi)}{\partial t}=0$$

- Now using the identity:

$$\psi^\dagger\alpha_x\frac{\partial\psi}{\partial x}+\frac{\partial\psi^\dagger}{\partial x}\alpha_x\psi\equiv\frac{\partial(\psi^\dagger\alpha_x\psi)}{\partial x}$$

gives the continuity equation $\boxed{\vec{\nabla}.(\psi^\dagger\vec{\alpha}\psi)+\frac{\partial(\psi^\dagger\psi)}{\partial t}=0}$ **(D8)**

where $\psi^\dagger=(\psi_1^*,\psi_2^*,\psi_3^*,\psi_4^*)$

- The probability density and current can be identified as:

$$\boxed{\rho=\psi^\dagger\psi} \quad \text{and} \quad \boxed{\vec{j}=\psi^\dagger\vec{\alpha}\psi}$$

where $\rho=\psi^\dagger\psi=|\psi_1|^2+|\psi_2|^2+|\psi_3|^2+|\psi_4|^2>0$

- Unlike the KG equation, the Dirac equation has probability densities which are always positive.
- In addition, the solutions to the Dirac equation are the four component Dirac Spinors. A great success of the Dirac equation is that these components naturally give rise to the property of intrinsic spin.
- It can be shown that Dirac spinors represent spin-half particles (appendix II) with an intrinsic magnetic moment of

$$\vec{\mu}=\frac{q}{m}\vec{S}=2\frac{q}{2m}\vec{\sigma}\rightarrow-g\frac{q}{e}\mu_m\vec{\sigma} \qquad \mu_m \text{ – magneton}$$

g-factor = 2 for Dirac particle, g_e~2, g_p=2.79 – proton is not point-like

the Dirac γ Matrices

- The Dirac equation can be written more elegantly by introducing the four Dirac gamma matrices:

$$\gamma^0 \equiv \beta; \;\; \gamma^1 \equiv \beta\alpha_x; \;\; \gamma^2 \equiv \beta\alpha_y; \;\; \gamma^3 \equiv \beta\alpha_z$$

Premultiply the Dirac equation (D6) by β

$$i\beta\alpha_x \frac{\partial\psi}{\partial x} + i\beta\alpha_y \frac{\partial\psi}{\partial y} + i\beta\alpha_z \frac{\partial\psi}{\partial z} - \beta^2 m\psi = -i\beta\frac{\partial\psi}{\partial t}$$

$$\Rightarrow \quad i\gamma^1 \frac{\partial\psi}{\partial x} + i\gamma^2 \frac{\partial\psi}{\partial y} + i\gamma^3 \frac{\partial\psi}{\partial z} - m\psi = -i\gamma^0 \frac{\partial\psi}{\partial t}$$

using $\partial_\mu = (\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ this can be written compactly as:

$$\boxed{(i\gamma^\mu \partial_\mu - m)\psi = 0} \qquad \text{(D9)}$$

★ **NOTE:** it is important to realise that the Dirac gamma matrices are <u>not</u> four-vectors - they are constant matrices which remain invariant under a Lorentz transformation. However it can be shown that the Dirac equation is itself Lorentz covariant

- From the properties of the α and β matrices (D2)-(D4) immediately obtain:

$$(\gamma^0)^2 = \beta^2 = 1 \quad \text{and} \quad (\gamma^1)^2 = \beta\alpha_x\beta\alpha_x = -\alpha_x\beta\beta\alpha_x = -\alpha_x^2 = -1$$

- The full set of relations is

$$\begin{aligned} (\gamma^0)^2 &= 1 \\ (\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 &= -1 \\ \gamma^0\gamma^j + \gamma^j\gamma^0 &= 0 \\ \gamma^j\gamma^k + \gamma^k\gamma^j &= 0 \quad (j \neq k) \end{aligned}$$

which can be expressed as:

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \qquad \text{(defines the algebra)}$$

- Are the gamma matrices Hermitian?
  - β is Hermitian so γ^0 is Hermitian.
  - The α matrices are also Hermitian, giving

$$\gamma^{1\dagger} = (\beta\alpha_x)^\dagger = \alpha_x^\dagger\beta^\dagger = \alpha_x\beta = -\beta\alpha_x = -\gamma^1$$

  - Hence $\gamma^1, \gamma^2, \gamma^3$ are anti-Hermitian

$$\boxed{\gamma^{0\dagger} = \gamma^0, \;\; \gamma^{1\dagger} = -\gamma^1, \;\; \gamma^{2\dagger} = -\gamma^2, \;\; \gamma^{3\dagger} = -\gamma^3}$$

Pauli-Dirac Representation

- From now on we will use the Pauli-Dirac representation of the gamma matrices:

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}; \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}$$ which when written in full are

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \ \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}; \ \gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}; \ \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- Using the gamma matrices $\rho = \psi^\dagger \psi$ and $\vec{j} = \psi^\dagger \vec{\alpha} \psi$ can be written as:

$$j^\mu = (\rho, \vec{j}) = \psi^\dagger \gamma^0 \gamma^\mu \psi$$

where j^μ is the <u>four-vector</u> current.
(The proof that j^μ is indeed a four vector is given in Appendix V.)

- In terms of the four-vector current the continuity equation becomes

$$\partial_\mu j^\mu = 0$$

- Finally the expression for the four-vector current

$$j^\mu = \psi^\dagger \gamma^0 \gamma^\mu \psi$$

can be simplified by introducing the adjoint spinor

- The adjoint spinor is defined as

$$\boxed{\overline{\psi} = \psi^\dagger \gamma^0}$$

i.e. $$\overline{\psi} = \psi^\dagger \gamma^0 \quad = \quad (\psi^*)^T \gamma^0 = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\boxed{\overline{\psi} \quad = \quad (\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*)}$$

- In terms the adjoint spinor the four vector current can be written:

$$\boxed{j^\mu = \overline{\psi} \gamma^\mu \psi}$$

- **this 4-vector current is used for derivation of the Lorentz invariant matrix element for the fundamental interactions.**