# 2 The Dirac Equation

In 1928, Dirac proposed a wave equation for the electron which was consistent with special relativity and successfully accounted for the intrinsic angular momentum (spin) and intrinsic magnetic moment of the electron postulated earlier by Goudsmit and Uhlenbeck. In addition, the Dirac equation predicted the existence of a particle with the same mass as the electron but with opposite electric charge, namely the positron. This prediction was confirmed by Anderson's observation in 1933 of positrons in cosmic rays. The existence of antiparticles thus emerges as an inevitable consequence of combining quantum mechanics and special relativity.

In this handout, we consider the Dirac equation and its predictions in detail, beginning with a brief review of the non-relativistic Schrödinger equation and its straightforward relativistic generalisation, the Klein-Gordon equation.

## 2.1 Non-Relativistic Quantum Mechanics

For a non-relativistic free particle, with $V(\boldsymbol{r}) = 0$, the operator replacements (with $\hbar = 1$)

$$\boldsymbol{p} \to -i\boldsymbol{\nabla}, \quad E \to i\frac{\partial}{\partial t} \tag{1}$$

applied to the equation

$$\frac{|\boldsymbol{p}|^2}{2m} = E$$

give the *time-dependent Schrödinger equation*

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

This is *not* a Lorentz invariant equation, being first order in $\partial/\partial t$ but second order in $\partial/\partial x$. Transforming the right hand side of Equation (2) to another frame, for example, would introduce terms containing the first order spatial derivatives $\partial\psi/\partial x'$, $\partial\psi/\partial y'$, $\partial\psi/\partial z'$, which have no equivalent in the original (untransformed) equation.

Taking the complex conjugate of Schrödinger's equation gives

$$-\frac{1}{2m}\nabla^2\psi^* = -i\frac{\partial\psi^*}{\partial t}. \tag{3}$$

Forming the linear combination $\psi^* \times$Eqn (2)$-\psi\times$Eqn (3) then gives

$$-\frac{1}{2m}\left[\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right] = i\left[\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right] = i\frac{\partial}{\partial t}\left(\psi^*\psi\right).$$

Using the identity

$$\psi^*\nabla^2\psi - \psi\nabla^2\psi^* = \boldsymbol{\nabla}\cdot\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right) \tag{4}$$

we obtain

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

A comparison with the *continuity equation*

$$\boldsymbol{\nabla}\cdot\boldsymbol{j} + \frac{\partial\rho}{\partial t} = 0 \tag{5}$$

then gives the following expressions for the probability density and current:

$$\rho = \psi^*\psi = |\psi|^2 \tag{6}$$

$$\boldsymbol{j} = \frac{1}{2im}\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right)\ . \tag{7}$$

Schrödinger's equation admits plane wave solutions of the form

$$\psi = Ne^{i(\boldsymbol{p}.\boldsymbol{r}-Et)} = Ne^{-ip.x},$$

as can be checked by direct substitution into Equation (2). The plane wave solutions are easily seen to be eigenstates of the energy and momentum operators of Equation (1):

$$\left(-i\boldsymbol{\nabla}\right)\psi = \boldsymbol{p}\psi, \qquad \left(i\frac{\partial}{\partial t}\right)\psi = E\psi\ ,$$

and hence correspond to free particles of definite energy $E$ and momentum $\boldsymbol{p}$. Substituting into Equations (6) and (7), the probability density $\rho$ and current $\boldsymbol{j}$ are found to be

$$\rho = |N|^2, \qquad \boldsymbol{j} = |N|^2\frac{\boldsymbol{p}}{m} = |N|^2\boldsymbol{v}.$$

Hence the number of particles per unit volume is $|N|^2$ and the current $\boldsymbol{j}$ is a vector pointing along the particle's direction of motion with magnitude equal to the *particle flux*, *i.e.* the number of particles per unit area per unit time passing any fixed point along the particle trajectory.

## 2.2 The Klein-Gordon Equation

Applying the replacements $\boldsymbol{p}\to -i\boldsymbol{\nabla}, E\to i\partial/\partial t$ to the relativistic equation

$$|\boldsymbol{p}|^2 + m^2 = E^2$$

(with $\hbar = c = 1$) in place of $|\boldsymbol{p}|^2/2m = E$ gives immediately the *Klein-Gordon equation*

$$\boxed{\nabla^2\psi - m^2\psi = \frac{\partial^2\psi}{\partial t^2}}\ . \tag{8}$$

Introducing the partial derivative 4-vector

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \left(\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right)$$

with the property

$$\Box^2 \equiv \partial^\mu \partial_\mu = \left(\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right) \cdot \left(\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right) = \frac{\partial^2}{\partial t^2} - \nabla^2$$

the Klein-Gordon equation can be written in the manifestly Lorentz invariant forms

$$\boxed{(\partial_\mu \partial^\mu + m^2)\psi = 0} \qquad \text{or} \qquad \boxed{(\Box^2 + m^2)\psi = 0}\,.$$

Plane wave solutions $\psi = Ne^{i(\boldsymbol{p}.\boldsymbol{r} - Et)}$ satisfy the Klein-Gordon equation provided that

$$|\boldsymbol{p}|^2 + m^2 = E^2; \qquad \textit{i.e.}\ \text{provided} \quad E = \pm\sqrt{|\boldsymbol{p}|^2 + m^2}.$$

Thus the Klein-Gordon equation suffers from the problem that there exist negative energy solutions for a free particle.

A further problem arises when the probability density and current are derived for the Klein-Gordon equation. Taking the complex conjugate of Equation (8) gives

$$\nabla^2 \psi^* - m^2 \psi^* = \frac{\partial^2 \psi^*}{\partial t^2}\,. \tag{9}$$

Taking $\psi^* \times$Eqn (8)$-\psi\times$Eqn (9), we then obtain

$$\psi^* \left(\nabla^2 \psi - m^2 \psi\right) - \psi\left(\nabla^2 \psi^* - m^2 \psi^*\right) = \psi^* \frac{\partial^2 \psi}{\partial t^2} - \psi \frac{\partial^2 \psi^*}{\partial t^2}\,.$$

Applying the identity in Equation (4) to the left-hand side of this equation then gives

$$\boldsymbol{\nabla} \cdot (\psi^* \boldsymbol{\nabla} \psi - \psi \boldsymbol{\nabla} \psi^*) = \frac{\partial}{\partial t}\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right)\,.$$

Thus we again arrive at the continuity equation, Equation (5), with

$$\rho = i\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right) \tag{10}$$

$$\boldsymbol{j} = i\left(\psi \boldsymbol{\nabla} \psi^* - \psi^* \boldsymbol{\nabla} \psi\right) \tag{11}$$

where the overall normalisation factors of $i$ are chosen to give a probability density $\rho$ which is real. Substituting the plane wave solution $\psi = Ne^{i(\boldsymbol{p}.\boldsymbol{r} - Et)}$ into Equations (10) and (11) gives

$$\rho = 2|N|^2 E, \qquad \boldsymbol{j} = 2|N|^2 \boldsymbol{p}\,. \tag{12}$$

The probability density $\rho$ therefore has the same sign as the energy $E$. In particular, if the energy $E$ is negative the probability density $\rho$ is also negative.

The Klein-Gordon equation therefore suffers from the presence of negative energy states and negative probability densities. These features, and the fact that the Klein-Gordon equation was unable to reproduce various features of atomic spectra, led Dirac (1928) to search for an alternative relativistic wave equation for the electron. The *Dirac equation* that he proposed turned out to describe spin half particles and antiparticles.

It was shown later by Pauli and Weisskopf (1934) that the problems with the Klein-Gordon equation could be resolved by treating $\psi(x)$ as a *field operator* rather than a single particle wavefunction. The resulting *quantum field theory* turns out to describe systems of spin 0 particles and antiparticles.

Before leaving the Klein-Gordon equation, we demonstrate that the vector $\boldsymbol{j}$ again represents the particle flux. The expression for $\rho$ in Equation (12) corresponds to a density of $2|N|^2E$ particles per unit volume. For a beam of particles of this density travelling with velocity $\boldsymbol{v}$, the particle flux is $2|N|^2E\boldsymbol{v}$. But the momentum and energy of a particle of velocity $\boldsymbol{v}$ are (with $c = 1$)

$$\boldsymbol{p} = \gamma m\boldsymbol{v}, \qquad E = \gamma m \; .$$

where $\gamma = 1/\sqrt{1 - v^2}$. Hence $E\boldsymbol{v} = \boldsymbol{p}$, and the particle flux is therefore $2|N|^2\boldsymbol{p} = \boldsymbol{j}$, as required.

Equations (10) and (11) can be combined into a single covariant equation by introducing the current 4-vector $j^\mu = (\rho, \boldsymbol{j})$:

$$j^\mu = i\left(\psi^*\partial^\mu\psi - \psi\partial^\mu\psi^*\right) .$$

(Setting $\mu = 0$ recovers Equation (10) while setting $\mu = 1, 2, 3$ recovers Equation (11)). Similarly, in covariant notation, Equation (12) is equivalent to

$$j^\mu = 2|N|^2 p^\mu \; ,$$

while the continuity equation, Equation (5), can be expressed in covariant form as

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

## 2.3 The Dirac Equation

Dirac attempted to resolve the problems associated with the Klein-Gordon equation by looking for an alternative equation *first order* in both $\partial/\partial t$ and $\partial/\partial x$, namely the *Dirac equation*

$$\boxed{H\psi = (\boldsymbol{\alpha}.\boldsymbol{p} + \beta m)\psi = i\frac{\partial\psi}{\partial t}}$$

where $H$ is the Hamiltonian operator, $\boldsymbol{p} = -i\boldsymbol{\nabla}$ is the momentum operator, and $\boldsymbol{\alpha}$ and $\beta$ are constants to be determined. Writing this explicitly in terms of its components, we have

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

"Squaring" this equation by operating on the left with $(-i\alpha_x\partial/\partial x - \ldots + \beta m)$ and on the right with $i\partial/\partial t$ gives

$$\left(-i\alpha_x\frac{\partial}{\partial x} - \ldots + \beta m\right)\left(-i\alpha_x\frac{\partial}{\partial x} - \ldots + \beta m\right)\psi = \left(i\frac{\partial}{\partial t}\right)\left(i\frac{\partial}{\partial t}\right)\psi \; .$$

This can be expanded as

$$\begin{aligned}
&-\alpha_x^2\frac{\partial^2\psi}{\partial x^2}-\alpha_y^2\frac{\partial^2\psi}{\partial y^2}-\alpha_z^2\frac{\partial^2\psi}{\partial z^2}\\
&\qquad-(\alpha_x\alpha_y+\alpha_y\alpha_x)\frac{\partial^2\psi}{\partial x\partial y}-(\alpha_x\alpha_z+\alpha_z\alpha_x)\frac{\partial^2\psi}{\partial x\partial z}-(\alpha_y\alpha_z+\alpha_z\alpha_y)\frac{\partial^2\psi}{\partial y\partial z}\\
&\qquad-i(\alpha_x\beta+\beta\alpha_x)m\frac{\partial\psi}{\partial x}-i(\alpha_y\beta+\beta\alpha_y)m\frac{\partial\psi}{\partial y}-i(\alpha_z\beta+\beta\alpha_z)m\frac{\partial\psi}{\partial z}\\
&\qquad\qquad+\beta^2m^2\psi=-\frac{\partial^2\psi}{\partial t^2}\ . \qquad (13)
\end{aligned}$$

But a free particle (with $E^2 = p^2 + m^2$) must also satisfy the Klein-Gordon equation:

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}-m^2\psi=\frac{\partial^2\psi}{\partial t^2}\ . \qquad (14)$$

For Equations (13) and (14) to be consistent, we must have:

$$\beta^2=\alpha_x^2=\alpha_y^2=\alpha_z^2=1 \qquad (15)$$

$$\beta\alpha_j+\alpha_j\beta=0 \qquad (16)$$

$$\alpha_j\alpha_k+\alpha_k\alpha_j=0 \quad (j\neq k) \qquad (17)$$

Hence the $\alpha_j$ and $\beta$ can not simply be numbers; they must be (at least) $4\times4$ matrices.

For $H$ to be Hermitian, $\boldsymbol{\alpha}$ and $\beta$ must be Hermitian:

$$\alpha_x^\dagger=\alpha_x;\qquad\alpha_y^\dagger=\alpha_y;\qquad\alpha_z^\dagger=\alpha_z;\qquad\beta^\dagger=\beta\ .$$

A convenient choice is:

$$\beta=\begin{pmatrix}I&0\\0&-I\end{pmatrix},\qquad\alpha_j=\begin{pmatrix}0&\sigma_j\\\sigma_j&0\end{pmatrix}$$

where $I$ is the $2\times2$ unit matrix and the $\sigma_j$ are the $2\times2$ Pauli spin matrices:

$$I=\begin{pmatrix}1&0\\0&1\end{pmatrix};\quad\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix};\quad\sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix};\quad\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\ . \qquad (18)$$

Then $\psi$ is represented by a 4-component object called a *spinor*:

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

(not to be confused with a 4-vector).

We shall see below that the four spinor degrees of freedom correspond to the two spin states of a spin $\frac{1}{2}$ *particle* plus the two spin states of the associated spin $\frac{1}{2}$ *antiparticle*.

## 2.4 Probability Density and Current

Writing out each component explicitly, the Dirac equation is

$$-i\alpha_x \frac{\partial\psi}{\partial x} - i\alpha_y \frac{\partial\psi}{\partial y} - i\alpha_z \frac{\partial\psi}{\partial z} + \beta m\psi = i\frac{\partial\psi}{\partial t} \ . \tag{19}$$

Taking the Hermitian (rather than just complex) conjugate of this equation, remembering that $A^\dagger = (A^*)^T$ so $i^\dagger = -i$, and that $(AB)^\dagger = B^\dagger A^\dagger$, we obtain

$$i\frac{\partial\psi^\dagger}{\partial x}\alpha_x^\dagger + i\frac{\partial\psi^\dagger}{\partial y}\alpha_y^\dagger + i\frac{\partial\psi^\dagger}{\partial z}\alpha_z^\dagger + \psi^\dagger\beta^\dagger m = -i\frac{\partial\psi^\dagger}{\partial t} \ . \tag{20}$$

Taking the combination $\psi^\dagger \times \text{Eqn}\,(19) - \text{Eqn}\,(20) \times \psi$ and using the fact that $\boldsymbol{\alpha}$ and $\beta$ are Hermitian then gives

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

Using the identity

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

(and similarly for $y$, $z$) we arrive at the continuity equation

$$\boxed{\boldsymbol{\nabla}\cdot(\psi^\dagger\boldsymbol{\alpha}\psi) + \frac{\partial}{\partial t}(\psi^\dagger\psi) = 0} \ . \tag{21}$$

The probability density $\rho$ and current $\boldsymbol{j}$ can therefore be identified as

$$\rho = \psi^\dagger\psi \equiv |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \tag{22}$$

$$\boldsymbol{j} = \psi^\dagger\boldsymbol{\alpha}\psi \tag{23}$$

where Equation (22) follows from the fact that $\psi^\dagger$ is the $1 \times 4$ matrix

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

The probability density $\rho$ is therefore guaranteed to be positive, solving one of the problems of the Klein-Gordon equation.

## 2.5 Orbital Angular Momentum

Consider the commutator $[H, \boldsymbol{L}]$, where $\boldsymbol{L} = \boldsymbol{r} \wedge \boldsymbol{p}$ is the orbital angular momentum operator:

$$[H, \boldsymbol{L}] = [\boldsymbol{\alpha}.\boldsymbol{p} + \beta m, \boldsymbol{r}\wedge\boldsymbol{p}] = [\boldsymbol{\alpha}.\boldsymbol{p}, \boldsymbol{r}\wedge\boldsymbol{p}] + [\beta m, \boldsymbol{r}\wedge\boldsymbol{p}] \ .$$

Since $\beta m$ is a constant matrix, independent of $\boldsymbol{r}$ and $\boldsymbol{p}$, we have $[\beta m, \boldsymbol{r} \wedge \boldsymbol{p}] = 0$. Hence

$$[H, \boldsymbol{L}] = [\boldsymbol{\alpha.p}, \boldsymbol{r} \wedge \boldsymbol{p}] \ .$$

Consider first the $x$ component of this equation:

$$\begin{aligned}[H, L_x] &= [\boldsymbol{\alpha.p}, (\boldsymbol{r} \wedge \boldsymbol{p})_x] \\ &= [\alpha_x p_x + \alpha_y p_y + \alpha_z p_z \ , \ y p_z - z p_y].\end{aligned}$$

The various operators in this expression all commute with each other except that $[x, p_x] = [y, p_y] = [z, p_z] = i$. Hence

$$\begin{aligned}[H, L_x] &= \alpha_y p_z [p_y, y] - \alpha_z p_y [p_z, z] \\ &= -i(\alpha_y p_z - \alpha_z p_y) \\ &= -i(\boldsymbol{\alpha} \wedge \boldsymbol{p})_x \ .\end{aligned}$$

Similar expressions hold for the $y$ and $z$ components, giving

$$[H, \boldsymbol{L}] = -i\boldsymbol{\alpha} \wedge \boldsymbol{p} \ . \tag{24}$$

Hence the angular momentum operator $\boldsymbol{L} = \boldsymbol{r} \wedge \boldsymbol{p}$ does not commute with the Hamiltonian $H$, and $\boldsymbol{L}$ is therefore not conserved.

## 2.6 Spin

Now consider the operator $\boldsymbol{\Sigma}$ defined as

$$\boldsymbol{\Sigma} \equiv \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$

where $\boldsymbol{\sigma}$ represents the three $2 \times 2$ Pauli spin matrices. Explicitly, we have

$$\Sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ; \quad \Sigma_y = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} ; \quad \Sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

The commutator $[H, \boldsymbol{\Sigma}]$ is given by

$$[H, \boldsymbol{\Sigma}] = [\boldsymbol{\alpha.p} + \beta m, \boldsymbol{\Sigma}] = [\boldsymbol{\alpha.p}, \boldsymbol{\Sigma}] + [\beta m, \boldsymbol{\Sigma}] \ .$$

But

$$[\beta, \boldsymbol{\Sigma}] = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = 0 \ ,$$

and hence

$$[H, \boldsymbol{\Sigma}] = [\boldsymbol{\alpha.p}, \boldsymbol{\Sigma}] \ .$$

The $x$ component of this equation is

$$\begin{aligned}[H, \Sigma_x] &= [\alpha_x p_x + \alpha_y p_y + \alpha_z p_z, \Sigma_x] \\ &= p_x[\alpha_x, \Sigma_x] + p_y[\alpha_y, \Sigma_x] + p_z[\alpha_z, \Sigma_x] \ .\end{aligned}$$

Evaluating each of the commutators on the right-hand side in turn gives

$$\begin{aligned}[\alpha_x, \Sigma_x] &= \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} = 0 \\ [\alpha_y, \Sigma_x] &= \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \sigma_y\sigma_x - \sigma_x\sigma_y \\ \sigma_y\sigma_x - \sigma_x\sigma_y & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2i\sigma_z \\ -2i\sigma_z & 0 \end{pmatrix} \\ &= -2i\alpha_z \\ [\alpha_z, \Sigma_x] &= 2i\alpha_y\end{aligned}$$

where the commutation properties $[\sigma_x, \sigma_y] = 2i\sigma_z$ etc. have been used. Hence

$$\begin{aligned}[H, \Sigma_x] &= -2ip_y\alpha_z + 2ip_z\alpha_y \\ &= 2i(\boldsymbol{\alpha} \wedge \boldsymbol{p})_x.\end{aligned}$$

A similar result holds also for the $y$ and $z$ components, giving altogether

$$[H, \boldsymbol{\Sigma}] = 2i\boldsymbol{\alpha} \wedge \boldsymbol{p}. \tag{25}$$

Hence, from Equations (24) and (25), the combination

$$\boldsymbol{J} = \boldsymbol{L} + \tfrac{1}{2}\boldsymbol{\Sigma}$$

commutes with the Hamiltonian and is therefore conserved:

$$\boxed{[H, \boldsymbol{L} + \tfrac{1}{2}\boldsymbol{\Sigma}] = 0} \ .$$

The operator $\frac{1}{2}\boldsymbol{\Sigma}$ can be identified as the intrinsic angular momentum or *spin* of the electron:

$$\boldsymbol{S} = \tfrac{1}{2}\boldsymbol{\Sigma} \ .$$

The operators $S^2$ and $S_z$ are

$$S^2 = \tfrac{1}{4}(\Sigma_x^2 + \Sigma_y^2 + \Sigma_z^2) = \tfrac{3}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ; \qquad S_z = \tfrac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

Thus $S^2$ and $S_z$ commute and have eigenvalues $S(S+1) = \frac{3}{4}$ and $S_z = \pm\frac{1}{2}$, respectively. The operator $\boldsymbol{S}$ therefore corresponds to a spin angular momentum $S = \frac{1}{2}$. Such an intrinsic electron angular momentum had been postulated by Goudsmit and Uhlenbeck in 1925 to account for various features of atomic spectra, and the Dirac equation (1928) finally provided a proper justification for its introduction.

## 2.7 Covariant Notation: the Dirac $\boldsymbol{\gamma}$ Matrices

The Dirac equation can be written more elegantly by introducing the four Dirac gamma matrices $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ defined by

$$\alpha_x = \gamma^0\gamma^1; \qquad \alpha_y = \gamma^0\gamma^2; \qquad \alpha_z = \gamma^0\gamma^3; \qquad \beta = \gamma^0 \ . \tag{26}$$

Equation (19) then becomes

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

Premultiplying by $-\gamma^0$, and using $(\gamma^0)^2 = \beta^2 = 1$, then gives

$$i\gamma^1\frac{\partial\psi}{\partial x} + i\gamma^2\frac{\partial\psi}{\partial y} + i\gamma^3\frac{\partial\psi}{\partial z} - m\psi = -i\gamma^0\frac{\partial\psi}{\partial t} \ . \tag{27}$$

Introducing the 4-vector

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

we have

$$\gamma^\mu\partial_\mu = \gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^2\partial_2 + \gamma^3\partial_3 = \gamma^0\frac{\partial}{\partial t} + \gamma^1\frac{\partial}{\partial x} + \gamma^2\frac{\partial}{\partial y} + \gamma^3\frac{\partial}{\partial z} \ .$$

Hence Equation (27), the Dirac equation, can be written compactly as

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

From Equations (15) and (26), we have $\beta^2 = 1$ and $\gamma^0 = \beta$, and hence $(\gamma^0)^2 = 1$. From Equation (26), we then have $\beta\alpha_x = \gamma^0.\gamma^0\gamma^1 = \gamma^1$, and, similarly, $\beta\alpha_y = \gamma^2$, $\beta\alpha_z = \gamma^3$. Thus the inverse relations to those of Equation (26) are

$$\gamma^0 = \beta; \qquad \gamma^1 = \beta\alpha_x; \qquad \gamma^2 = \beta\alpha_y; \qquad \gamma^3 = \beta\alpha_z \ . \tag{28}$$

Then, for example,

$$(\gamma^1)^2 = \beta\alpha_x \cdot \beta\alpha_x = -\alpha_x\beta \cdot \beta\alpha_x = -\alpha_x^2 = -1$$

where we have used $\alpha_x\beta = -\beta\alpha_x$ from Equation (16). Continuing in this vein, it is easily seen that the anticommutation relations (15), (16), (17) for $\boldsymbol{\alpha}$ and $\beta$ become

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

These relations are all contained within the single expression

$$\boxed{\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}} \ . \tag{29}$$

Since $\beta$ is Hermitian and $\gamma^0 = \beta$, then clearly $\gamma^0$ is Hermitian. Similarly, since the $\boldsymbol{\alpha}$ matrices are Hermitian, we have

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

and similarly $\gamma^{2\dagger} = -\gamma^2$, $\gamma^{3\dagger} = -\gamma^3$. In summary, $\gamma^0$ is Hermitian while $\gamma^1, \gamma^2, \gamma^3$ are anti-Hermitian:

$$\gamma^{0\dagger} = \gamma^0, \quad \gamma^{1\dagger} = -\gamma^1, \quad \gamma^{2\dagger} = -\gamma^2, \quad \gamma^{3\dagger} = -\gamma^3 \, .$$

It is often convenient to work with a particular representation of the $\gamma$ matrices, such as the Pauli-Dirac representation:

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

where $I$ is the $2 \times 2$ unit matrix and the $\sigma_k$ are the $2 \times 2$ Pauli spin matrices of Equation (18). In full, we have

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix};$$

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

Equations (22) and (23) for the probability density and current, $\rho = \psi^\dagger\psi$ and $\boldsymbol{j} = \psi^\dagger\boldsymbol{\alpha}\psi$, can be written in the form

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

where $j^\mu \equiv (\rho, \boldsymbol{j})$ is the 4-vector current, while the continuity equation, Equation (21), becomes

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

It is useful to define here the *adjoint spinor*

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

which, in the Pauli-Dirac representation, has components

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

In terms of the adjoint spinor, the current can be written

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

We shall encounter many constructions of this form when evaluating matrix elements.

Note that, despite the suggestive notation, the Dirac $\gamma$ matrices do not themselves constitute a 4-vector; they are simply constant $4 \times 4$ matrices and remain invariant under Lorentz transformations. However, it can be shown that, for any Lorentz transformation $x \to x'$, there exists a corresponding transformation $\psi'(x') = S\psi(x)$, with $S$ a suitable $4 \times 4$ matrix, which keeps the form of the Dirac equation invariant. It can also be shown that the current $\overline{\psi}\gamma^\mu\psi$ does indeed transform as a 4-vector.

## 2.8 Plane Wave Solutions

### 2.8.1 Derivation of Plane Wave Solutions

We now look for free particle solutions to the Dirac equation of the form

$$\psi = u(E, \boldsymbol{p})e^{i(\boldsymbol{p}.\boldsymbol{r} - Et)} .$$

The first order derivatives $\partial_\mu \psi$ are:

$$\partial_0\psi = \frac{\partial\psi}{\partial t} = -iE\psi; \quad \partial_1\psi = \frac{\partial\psi}{\partial x} = ip_x\psi; \quad \partial_2\psi = \frac{\partial\psi}{\partial y} = ip_y\psi; \quad \partial_3\psi = \frac{\partial\psi}{\partial z} = ip_z\psi .$$

Substituting into the Dirac equation, Equation (27), then gives

$$i\gamma^1.ip_x u + i\gamma^2.ip_y u + i\gamma^3.ip_z u - mu = -i\gamma^0. - iEu .$$

This can be rearranged as

$$(\gamma^0 E - \gamma^1 p_x - \gamma^2 p_y - \gamma^3 p_z - m)u = 0 ,$$

or, more compactly,

$$\boxed{(\gamma^\mu p_\mu - m)\, u = 0} .$$

But

$$\begin{aligned}\gamma^\mu p_\mu - m &= \gamma^0 E - \boldsymbol{\gamma}.\boldsymbol{p} - m \\ &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} E - \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix} \cdot \boldsymbol{p} \; - \; m \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} E - m & -\boldsymbol{\sigma}.\boldsymbol{p} \\ \boldsymbol{\sigma}.\boldsymbol{p} & -E - m \end{pmatrix}\end{aligned}$$

giving

$$\begin{pmatrix} E - m & -\boldsymbol{\sigma}.\boldsymbol{p} \\ \boldsymbol{\sigma}.\boldsymbol{p} & -E - m \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

where $u_A$ and $u_B$ are two-component spinors known as *Weyl spinors*. We thus obtain two coupled equations for $u_A$ and $u_B$:

$$(\boldsymbol{\sigma}.\boldsymbol{p})u_B = (E - m)u_A \tag{30}$$

$$(\boldsymbol{\sigma}.\boldsymbol{p})u_A = (E + m)u_B \tag{31}$$

which can be solved to yield four linearly independent solutions.

The operator $\boldsymbol{\sigma}.\boldsymbol{p}$ is given by

$$\begin{aligned}\boldsymbol{\sigma}.\boldsymbol{p} &= \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_x + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} p_y + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} p_z \\ &= \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}\end{aligned} \tag{32}$$

Hence

$$(\boldsymbol{\sigma}.\boldsymbol{p})^2 = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} = \begin{pmatrix} |\boldsymbol{p}|^2 & 0 \\ 0 & |\boldsymbol{p}|^2 \end{pmatrix} = |\boldsymbol{p}|^2 I$$

Operating on Equation (31) with $\boldsymbol{\sigma}.\boldsymbol{p}$ then gives

$$\begin{aligned} & (\boldsymbol{\sigma}.\boldsymbol{p})(\boldsymbol{\sigma}.\boldsymbol{p})u_A = (E+m)(\boldsymbol{\sigma}.\boldsymbol{p})u_B \\ \Rightarrow \quad & |\boldsymbol{p}|^2 u_A = (E+m)(E-m)u_A \\ \Rightarrow \quad & |\boldsymbol{p}|^2 = E^2 - m^2 \end{aligned}$$

Hence the Dirac equation still admits negative energy solutions with $E = -\sqrt{|\boldsymbol{p}|^2 + m^2}$.

### 2.8.2 Standard Form of Plane Wave Solutions

From Equation (31), we have

$$u_B = \frac{\boldsymbol{\sigma}.\boldsymbol{p}}{E+m} u_A = \frac{1}{E+m} \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} u_A$$

Two independent plane wave solutions can then be obtained with the simple choices

$$u_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{or} \qquad u_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

This gives

$$\boxed{u_1 = N \begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x + ip_y)/(E+m) \end{pmatrix}, \qquad u_2 = N \begin{pmatrix} 0 \\ 1 \\ (p_x - ip_y)/(E+m) \\ -p_z/(E+m) \end{pmatrix}} \qquad (33)$$

where $N$ is a normalisation factor, determined below. For these solutions, we *must* choose the positive energy solution, $E = +\sqrt{|\boldsymbol{p}|^2 + m^2}$, to prevent the denominator $E+m$ becoming zero as $E \to -m$.

The two remaining independent solutions can be found by choosing

$$u_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{or} \qquad u_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

In this case, $u_A$ is determined by Equation (30):

$$u_A = \frac{\boldsymbol{\sigma}.\boldsymbol{p}}{E-m} u_B = \frac{1}{E-m} \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} u_B$$

giving

$$\boxed{u_3 = N \begin{pmatrix} p_z/(E-m) \\ (p_x + ip_y)/(E-m) \\ 1 \\ 0 \end{pmatrix}, \qquad u_4 = N \begin{pmatrix} (p_x - ip_y)/(E-m) \\ -p_z/(E-m) \\ 0 \\ 1 \end{pmatrix}} \qquad (34)$$

where now we must choose the negative energy solution $E = -\sqrt{|\boldsymbol{p}|^2 + m^2}$ to prevent the denominator becoming zero when $E = +m$.

For a particle travelling along the $z$ axis, with $\boldsymbol{p} = (0, 0, p_z)$, the free particle spinors become

$$u_1 = N \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ 0 \end{pmatrix} \; ; \quad u_2 = N \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-p_z}{E+m} \end{pmatrix} \; ; \quad u_3 = N \begin{pmatrix} \frac{p_z}{E-m} \\ 0 \\ 1 \\ 0 \end{pmatrix} \; ; \quad u_4 = N \begin{pmatrix} 0 \\ \frac{-p_z}{E-m} \\ 0 \\ 1 \end{pmatrix}$$

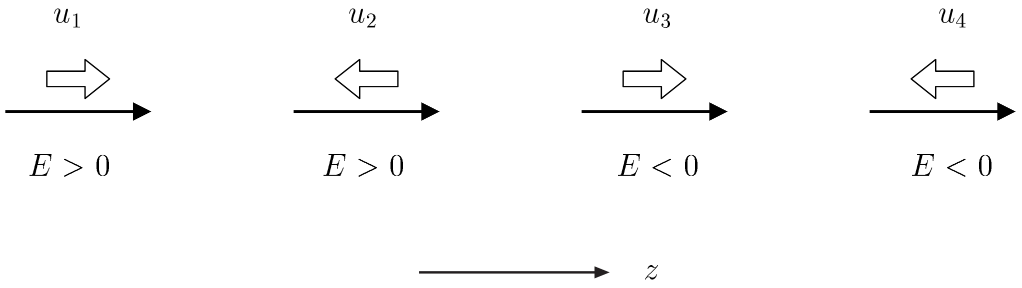
These spinors are easily seen to be eigenstates of the spin operator

$$S_z = \tfrac{1}{2}\Sigma_z = \tfrac{1}{2}\begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

with eigenvalues $\pm\frac{1}{2}$:

$$\begin{aligned} S_z u_1 &= +\tfrac{1}{2} \cdot u_1 \qquad & S_z u_3 &= +\tfrac{1}{2} \cdot u_3 \\ S_z u_2 &= -\tfrac{1}{2} \cdot u_2 \qquad & S_z u_4 &= -\tfrac{1}{2} \cdot u_4 \end{aligned}$$

*i.e.* for a particle travelling (forwards or backwards) along the $z$-axis, the spinors $u_1$ and $u_3$ have the spin vector pointing in the $+z$ direction ($S_z = +\frac{1}{2}$) while $u_2$ and $u_4$ have the spin vector pointing in the $-z$ direction ($S_z = -\frac{1}{2}$). For motion in the $+z$ direction for example, we have

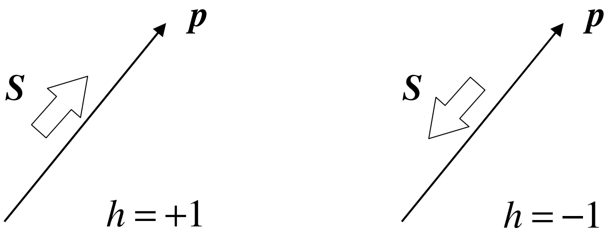


## 2.9 Helicity

The *helicity* operator $h$ measures the projection of the particle spin along the direction of motion. It is defined as

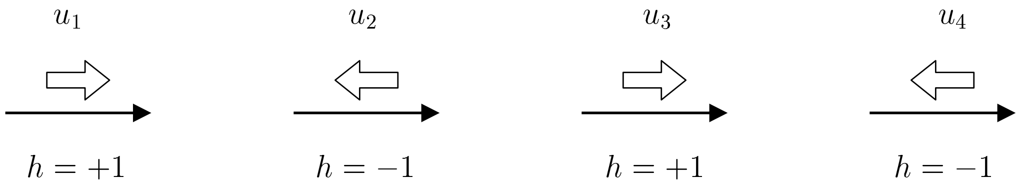
$$h \equiv 2\frac{\boldsymbol{S}.\boldsymbol{p}}{|\boldsymbol{p}|} = \frac{\boldsymbol{\Sigma}.\boldsymbol{p}}{|\boldsymbol{p}|} = \boldsymbol{\Sigma}.\hat{\boldsymbol{p}} = \begin{pmatrix} \boldsymbol{\sigma}.\hat{\boldsymbol{p}} & 0 \\ 0 & \boldsymbol{\sigma}.\hat{\boldsymbol{p}} \end{pmatrix}$$

where $\hat{\boldsymbol{p}}$ is the unit vector along $\boldsymbol{p}$, and $\boldsymbol{\Sigma} = 2\boldsymbol{S}$. For a free spin $\frac{1}{2}$ particle or antiparticle there are two possible helicity eigenstates: one with spin eigenvalue $S_{z'} = +\frac{1}{2}$ with respect to an axis $z'$ oriented along the direction of motion (helicity eigenvalue $h = +1$), and one with spin eigenvalue $S_{z'} = -\frac{1}{2}$ (helicity eigenvalue $h = -1$). Schematically, the helicity eigenstates can be pictured as:

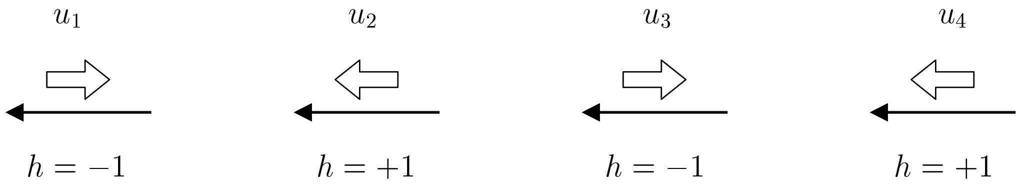


A particle or antiparticle in a helicity eigenstate with $h = -1$ is said to be *left-handed*, while one in an eigenstate with $h = +1$ is said to be *right-handed*.

For a particle or antiparticle travelling along the $z$-axis, with $p^\mu = (E, 0, 0, p_z)$, the helicity eigenstates are given by the spinors $u_i$. For motion in the $+z$ direction ($p_z > 0$), the helicity eigenvalues corresponding to the spin states $u_1$, $u_2$, $u_3$, $u_4$ are $+1$, $-1$, $+1$, $-1$, respectively:



For motion in the $-z$ direction ($p_z < 0$), the spin direction remains unchanged and the helicity eigenvalues are therefore reversed:



In general, *i.e.* for particle motion not directed along the $z$ axis, the spinors $u_i$ are *not* helicity eigenstates. In the next section, we construct the helicity eigenstates for free particles moving in a general direction $\theta$ with respect to the $z$ axis.

## 2.10 Construction of Helicity Eigenstates

The positive and negative helicity eigenstates $u_\uparrow$ and $u_\downarrow$ are the solutions of the equations

$$(\boldsymbol{\Sigma}.\hat{\boldsymbol{p}})u_\uparrow = +u_\uparrow$$
$$(\boldsymbol{\Sigma}.\hat{\boldsymbol{p}})u_\downarrow = -u_\downarrow \,.$$

In terms of two-component Weyl spinors $u_A$ and $u_B$, this is

$$\begin{pmatrix} \boldsymbol{\sigma}.\hat{\boldsymbol{p}} & 0 \\ 0 & \boldsymbol{\sigma}.\hat{\boldsymbol{p}} \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix} = \pm \begin{pmatrix} u_A \\ u_B \end{pmatrix} \,,$$

which gives two identical decoupled equations

$$(\boldsymbol{\sigma}.\hat{\boldsymbol{p}})u_A = \pm u_A \qquad (35)$$
$$(\boldsymbol{\sigma}.\hat{\boldsymbol{p}})u_B = \pm u_B \,. \qquad (36)$$

For a particle moving at an angle $\theta$ to the $z$ axis, we can take the 4-momentum to be $p^\mu = (E, p\sin\theta, 0, p\cos\theta)$, so that $\hat{\boldsymbol{p}} = (\sin\theta, 0, \cos\theta)$. Using Equation (32) and denoting the two components of the spinors $u_A$ or $u_B$ by $a$ and $b$, Equations (35) and (36) are both equivalent to

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \pm \begin{pmatrix} a \\ b \end{pmatrix} \,.$$

The first row of this equation gives

$$\frac{b}{a} = \frac{\pm 1 - \cos\theta}{\sin\theta} \,.$$

For the positive helicity solution $u_\uparrow$ with $h = +1$, we have

$$\frac{b}{a} = \frac{1-\cos\theta}{\sin\theta} = \frac{2\sin^2\theta/2}{2\sin\theta/2\cos\theta/2} = \tan\theta/2 \,,$$

and we can therefore take

$$u_A \text{ or } u_B = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix} \,.$$

For the negative helicity solution $u_\downarrow$ with $h = -1$, we have instead

$$\frac{b}{a} = \frac{-(1+\cos\theta)}{\sin\theta} = \frac{-2\cos^2\theta/2}{2\sin\theta/2\cos\theta/2} = -\cot\theta/2 \,,$$

and can take

$$u_A \text{ or } u_B = \begin{pmatrix} -\sin\theta/2 \\ \cos\theta/2 \end{pmatrix} \,.$$

For positive energy solutions with $E > 0$, $u_B$ is given in terms of $u_A$ by Equation (31):

$$u_B = \frac{\boldsymbol{\sigma}.\boldsymbol{p}}{E+m}u_A = \frac{p}{E+m}(\boldsymbol{\sigma}.\hat{\boldsymbol{p}})u_A = \pm\frac{p}{E+m}u_A \,.$$

Using the two possible choices of $u_A$ in turn then gives the positive energy helicity eigenstates as

$$u_\uparrow(p) = N\begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \\ \frac{p}{E+m}\cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \end{pmatrix}, \qquad u_\downarrow(p) = N\begin{pmatrix} -\sin\theta/2 \\ \cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \\ -\frac{p}{E+m}\cos\theta/2 \end{pmatrix} \,.$$

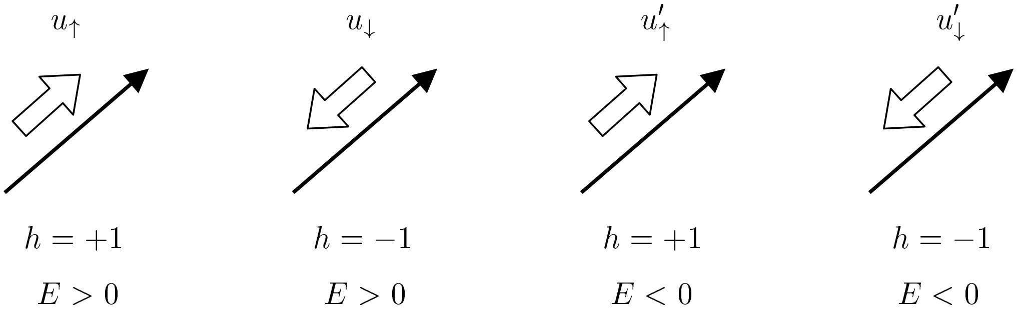
For negative energy solutions with $E < 0$, we must use Equation (30) to determine $u_A$ in terms of $u_B$:

$$u_A = \frac{\boldsymbol{\sigma.p}}{E-m}u_B = \frac{p}{E-m}(\boldsymbol{\sigma}.\hat{\boldsymbol{p}})u_B = \pm\frac{p}{E-m}u_B\ .$$

Taking the two possible choices of $u_B$ in turn then gives the negative energy helicity eigenstates $u'_\uparrow$ and $u'_\downarrow$:

$$u'_\uparrow(p) = N\begin{pmatrix}\frac{p}{E-m}\cos\theta/2\\ \frac{p}{E-m}\sin\theta/2\\ \cos\theta/2\\ \sin\theta/2\end{pmatrix}, \qquad u'_\downarrow(p) = N\begin{pmatrix}\frac{p}{E-m}\sin\theta/2\\ -\frac{p}{E-m}\cos\theta/2\\ -\sin\theta/2\\ \cos\theta/2\end{pmatrix}\ . \tag{37}$$

Schematically, we have



Note that when $\theta = 0$ the spinor $u_\uparrow$ becomes equal to $u_1$, $u_\downarrow$ becomes equal to $u_2$, $u'_\uparrow$ becomes equal to $u_3$, and $u'_\downarrow$ becomes equal to $u_4$, as expected. Similarly, when $\theta = \pi$, $u_\uparrow$ becomes equal to $u_2$, $u_\downarrow$ becomes equal to $-u_1$, $u'_\uparrow$ becomes equal to $u_4$, and $u'_\downarrow$ becomes equal to $-u_3$, where the arbitrary overall normalisation factors of $-1$ are of no physical consequence.

## 2.11 Normalisation of Plane Wave Solutions

The normalisation of a multi-component spinor $\psi$ is determined by the value of $\psi^\dagger\psi = (\psi^*)^T\psi$, in place of $\psi^*\psi$ for a one-dimensional wavefunction. For a plane wave of the form $\psi = u(p)e^{ip.x}$, we have $\psi^\dagger = u(p)^\dagger e^{-ip.x}$ and hence $\psi^\dagger\psi = u^\dagger u$.

For the basis spinor $u_1$ of Equation (33), we have

$$u_1^\dagger u_1 = |N|^2\left(1 + \frac{p_z^2}{(E+m)^2} + \frac{p_x^2+p_y^2}{(E+m)^2}\right) = |N|^2\frac{(E+m)^2+p^2}{(E+m)^2}$$

Using the relation $E^2 = p^2 + m^2$, this can be written

$$u_1^\dagger u_1 = |N|^2\frac{2E^2+2Em}{(E+m)^2} = |N|^2\frac{2E}{E+m}\ .$$

Similar calculations for $u_2^\dagger u_2$, $u_3^\dagger u_3$, $u_4^\dagger u_4$ give

$$u_1^\dagger u_1 = u_2^\dagger u_2 = |N|^2\frac{2E}{E+m}; \qquad u_3^\dagger u_3 = u_4^\dagger u_4 = |N|^2\frac{2E}{E-m}$$

For reasons which will become clear later, it is conventional to normalise to $2|E|$ particles per unit volume and choose

$$u_1^\dagger u_1 = u_2^\dagger u_2 = u_3^\dagger u_3 = u_4^\dagger u_4 = 2|E| \; .$$

This requires taking

$$\boxed{N = \sqrt{|E| + m}} \; .$$

The same normalisation factor $N$ applies equally to the helicity eigenstates $u_\uparrow$, $u_\downarrow$, $u'_\uparrow$, $u'_\downarrow$.

We note in passing that the spinors $u_i$ can easily be shown to be mutually orthogonal:

$$u_i^\dagger u_j = 0 \qquad (i \neq j) \; ,$$

and similarly for the helicity eigenstates $u_\uparrow$, $u_\downarrow$, $u'_\uparrow$, $u'_\downarrow$.

## 2.12 Interpretation of Negative Energy Solutions (Antiparticles)

The negative energy solutions $u_3$ and $u_4$, or equivalently $u'_\uparrow$ and $u'_\downarrow$, correspond to *particles* with momentum $\boldsymbol{p}$ and with *negative* energy $E = -\sqrt{|\boldsymbol{p}|^2 + m^2}$. Feynman suggested that the negative energy solutions be interpreted physically as *antiparticles* with momentum $-\boldsymbol{p}$ and *positive* energy $E = +\sqrt{|\boldsymbol{p}|^2 + m^2}$.

For antiparticles, it is convenient to replace the negative energy particle spinors $u$ by corresponding spinors $v$ expressed in terms of the *physical* $E, \boldsymbol{p}$ of the antiparticle. Schematically, we define

$$v(E, \boldsymbol{p}) \equiv u(-E, -\boldsymbol{p})$$

where $E > 0$ is the antiparticle energy and $\boldsymbol{p}$ is its three-momentum. Under the transformation $(E, \boldsymbol{p}) \to (-E, -\boldsymbol{p})$, the position vector $\boldsymbol{r}$ remains unchanged and hence the orbital angular momentum $\boldsymbol{L} = \boldsymbol{r} \wedge \boldsymbol{p}$ must change sign: $\boldsymbol{L} \to -\boldsymbol{L}$. Conservation of angular momentum, $[H, \boldsymbol{L} + \boldsymbol{S}] = 0$, can then only be maintained if $\boldsymbol{S} \to -\boldsymbol{S}$ also. Since $\boldsymbol{p}$ and $\boldsymbol{S}$ both change sign, the helicity $h \propto \boldsymbol{S}.\boldsymbol{p}$ must remain unchanged. In summary, we have

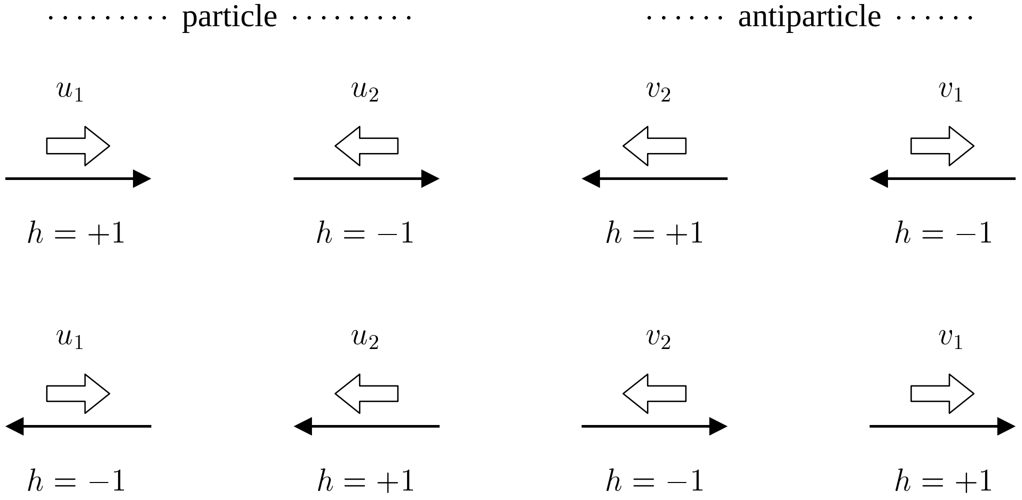
$$E \to -E, \quad \boldsymbol{p} \to -\boldsymbol{p}, \quad \boldsymbol{S} \to -\boldsymbol{S}, \quad h \to h \; .$$

In place of $u_3$ and $u_4$ of Equation (34), we introduce antiparticle spinors $v_1$ and $v_2$ defined as

$$v_1(E, \boldsymbol{p}) \equiv u_4(-E, -\boldsymbol{p}) = \sqrt{E + m} \begin{pmatrix} (p_x - ip_y)/(E + m) \\ -p_z/(E + m) \\ 0 \\ 1 \end{pmatrix} \tag{38}$$

$$v_2(E, \boldsymbol{p}) \equiv u_3(-E, -\boldsymbol{p}) = \sqrt{E + m} \begin{pmatrix} p_z/(E + m) \\ (p_x + ip_y)/(E + m) \\ 1 \\ 0 \end{pmatrix} \tag{39}$$

with $E > 0$ throughout. Note that $v_1$ is defined in terms of $u_4$ (rather than $u_3$) and $v_2$ is defined in terms of $u_3$ (rather than $u_4$). This choice ensures that $u_1$ and $v_1$ correspond to particles and antiparticles in the same physical spin state, and similarly for $u_2$ and $v_2$. For motion along the $z$ axis for example, the above transformations applied to the configurations shown towards the end of Section 2.9 result in



Thus the antiparticle spin associated with the spinor $v_1$ is $S_z = +\frac{1}{2}$, the same as the particle spin associated with the spinor $u_1$. Similarly, the antiparticle spinor $v_2$ has the same spin, $S_z = -\frac{1}{2}$, as the particle spinor $u_2$.

Under the transformation $E \to -E$, $\boldsymbol{p} \to -\boldsymbol{p}$, the plane wave factor $e^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)}$ becomes $e^{-i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)}$ and the physical antiparticle wavefunctions therefore have the form

$$\psi = v(E,\boldsymbol{p})e^{-i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)}$$

where $E > 0$.

Similar considerations apply to the negative energy helicity eigenstates $u'_\uparrow$ and $u'_\downarrow$ of Equation (37). We introduce antiparticle spinors $v_\uparrow$ and $v_\downarrow$ defined as

$$v_\uparrow(E,\boldsymbol{p}) \equiv u'_\uparrow(-E,-\boldsymbol{p}), \qquad v_\downarrow(E,\boldsymbol{p}) \equiv u'_\downarrow(-E,-\boldsymbol{p}) \ .$$

Under the transformation $\boldsymbol{p} \to -\boldsymbol{p}$, the angle $\theta$ with respect to the $z$ axis transforms as $\theta \to \theta + \pi$ and we must therefore make the replacements

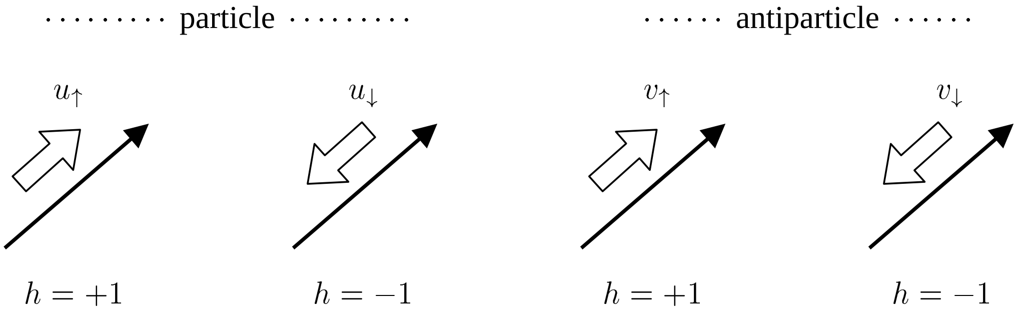
$$\begin{aligned} \cos\theta/2 \ &\longrightarrow \ \cos\left(\theta/2 + \pi/2\right) = -\sin\theta/2 \\ \sin\theta/2 \ &\longrightarrow \ \sin\left(\theta/2 + \pi/2\right) = \ \ \ \cos\theta/2 \end{aligned}$$

The negative energy helicity spinors of Equation (37) therefore become the antiparticle spinors

$$v_\uparrow(p) = N \begin{pmatrix} \frac{p}{E+m}\sin\theta/2 \\ -\frac{p}{E+m}\cos\theta/2 \\ -\sin\theta/2 \\ \cos\theta/2 \end{pmatrix}, \qquad v_\downarrow(p) = N \begin{pmatrix} \frac{p}{E+m}\cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \\ \cos\theta/2 \\ \sin\theta/2 \end{pmatrix}$$

where $E > 0$. For convenience, $v_\downarrow(p)$ has also been multiplied by an arbitrary overall normalisation factor of $-1$.

Schematically, we now have



it can be shown that an antiparticle has the same mass but opposite electric charge to its corresponding particle. This is done by deriving the equation of motion of a spin-half particle in an electromagnetic field and by introducing the *charge conjugation* operator, $C$. As its name suggests, this operator reverses the sign of all electric charges and hence transforms particles into antiparticles and *vice-versa*. The correspondence between $u_1$ and $v_1$, $u_2$ and $v_2$ *etc.* emerges formally from this analysis.

As emphasised by Feynman, an antiparticle travelling "forward in spacetime" can be considered equivalent to a particle travelling "backward in spacetime". A formal demonstration of this equivalence is done via the combined application of the parity operator $P$ to reverse the sign of the spatial coordinates $(x, y, z)$, the time reversal operator $T$ to change the sign of the time coordinate $t$, and the charge conjugation operator $C$ to change the sign of the electric charge. In Feynman diagrams, a common and useful convention is that the arrows drawn on lines representing spin-half particles or antiparticles indicate the *particle* direction. For antiparticles, the arrow therefore points in the opposite sense to the actual direction of the antiparticle.

## 2.13 Summary of Plane Wave Solutions

For a free *particle*, with 4-momentum $(E, \boldsymbol{p})$, the plane wave solutions are

$$\psi = u_i(E, \boldsymbol{p})e^{i(\boldsymbol{p}.\boldsymbol{r}-Et)}$$

$$\boxed{u_1 = \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x+ip_y)/(E+m) \end{pmatrix} \quad ; \qquad u_2 = \sqrt{E+m}\begin{pmatrix} 0 \\ 1 \\ (p_x-ip_y)/(E+m) \\ -p_z/(E+m) \end{pmatrix}}$$

For a free *antiparticle*, with physical 4-momentum $(E > 0, \boldsymbol{p})$, the solutions are

$$\psi = v_i(E, \boldsymbol{p})e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)}$$

$$\boxed{v_1 = \sqrt{E+m}\begin{pmatrix} (p_x-ip_y)/(E+m) \\ -p_z/(E+m) \\ 0 \\ 1 \end{pmatrix} \quad ; \qquad v_2 = \sqrt{E+m}\begin{pmatrix} p_z/(E+m) \\ (p_x+ip_y)/(E+m) \\ 1 \\ 0 \end{pmatrix}}$$

For a particle or antiparticle travelling (forwards or backwards) along the $z$-axis, the spinors $u_i$ and $v_i$ are eigenstates of the spin operator $S_z$, *i.e.* they are helicity eigenstates. In this case, the *physical* spin of the particle or antiparticle is $S_z = +\frac{1}{2}$ for $u_1$, $v_1$ and $S_z = -\frac{1}{2}$ for $u_2$, $v_2$.

In general, the spinors $u_1$, $u_2$, $v_1$, $v_2$ are *not* helicity eigenstates. For a particle moving at an angle $\theta$ to the $z$ axis with 4-momentum $p^\mu = (E, p\sin\theta, 0, p\cos\theta)$, the $h = +1$ (right-handed) and $h = -1$ (left-handed) helicity eigenstates are

$$\boxed{u_\uparrow(p) = \sqrt{E+m}\begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \\ \frac{p}{E+m}\cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \end{pmatrix}, \qquad u_\downarrow(p) = \sqrt{E+m}\begin{pmatrix} -\sin\theta/2 \\ \cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \\ -\frac{p}{E+m}\cos\theta/2 \end{pmatrix}} \tag{40}$$

For an antiparticle, the helicity eigenstates are (with $E > 0$)

$$\boxed{v_\uparrow(p) = \sqrt{E+m}\begin{pmatrix} \frac{p}{E+m}\sin\theta/2 \\ -\frac{p}{E+m}\cos\theta/2 \\ -\sin\theta/2 \\ \cos\theta/2 \end{pmatrix}, \qquad v_\downarrow(p) = \sqrt{E+m}\begin{pmatrix} \frac{p}{E+m}\cos\theta/2 \\ \frac{p}{E+m}\sin\theta/2 \\ \cos\theta/2 \\ \sin\theta/2 \end{pmatrix}} \tag{41}$$

Either of the two sets of basis states $(u_1, u_2, v_1, v_2)$ or $(u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow)$ can be used equally well in matrix element calculations. In what follows, we shall tend to prefer the helicity eigenstate spinors in order to bring out more clearly the underlying spin structure of the electromagnetic, weak and strong interactions.

The spinors above are all normalised to $2E$ particles per unit volume:

$$\rho = \psi^\dagger(p)\psi(p) = \overline{\psi}(p)\gamma^0\psi(p) = 2E\ .$$

More generally, it is shown on the examples sheet that the current $j^\mu = \overline{\psi}(p)\gamma^\mu\psi(p)$ for a free (anti)particle is proportional to the physical four-momentum of the (anti)particle:

$$\boxed{j^\mu = \overline{u}(p)\gamma^\mu u(p) = \overline{v}(p)\gamma^\mu v(p) = 2p^\mu}\ ,$$

and that the (anti)particle flux is given by the three-vector current $\boldsymbol{j} = 2\boldsymbol{p}$.

## 2.14 Chirality

In the extreme relativistic limit $E \gg m$, where particle masses can be neglected, the factor $p/(E+m)$ in Equations (40) and (41) tends to unity and the helicity eigenstate spinors become

$$u_\uparrow = \sqrt{E}\begin{pmatrix} c \\ s \\ c \\ s \end{pmatrix}\ ; \qquad u_\downarrow = \sqrt{E}\begin{pmatrix} -s \\ c \\ s \\ -c \end{pmatrix}\ ; \qquad v_\uparrow = \sqrt{E}\begin{pmatrix} s \\ -c \\ -s \\ c \end{pmatrix}\ ; \qquad v_\downarrow = \sqrt{E}\begin{pmatrix} c \\ s \\ c \\ s \end{pmatrix} \tag{42}$$

where $c \equiv \cos\theta/2$ and $s \equiv \sin\theta/2$. Introducing the matrix $\gamma^5$ defined by

$$\boxed{\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}}$$

it is straightforward to check that the helicity eigenstate spinors are eigenstates of $\gamma^5$:

$$\gamma^5 u_\uparrow = u_\uparrow, \qquad \gamma^5 u_\downarrow = -u_\downarrow, \qquad \gamma^5 v_\uparrow = -v_\uparrow, \qquad \gamma^5 v_\downarrow = v_\downarrow\ . \tag{43}$$

These equations do *not* hold in the more general case when the mass $m$ cannot be neglected, as is seen immediately by applying $\gamma^5$ to the spinors in Equations (40) and (41). Thus, in the extreme relativistic limit, and only in this limit, the helicity eigenstates $u_\uparrow$, $u_\downarrow$, $v_\uparrow$, $v_\downarrow$ become eigenstates of the $\gamma^5$ operator. For a *particle*, the $\gamma^5$ eigenvalue gives the helicity of the particle state while for an *antiparticle* $\gamma^5$ gives *minus* the helicity.

We now define the *left-handed* and *right-handed projection operators* $P_\mathrm{L}$ and $P_\mathrm{R}$ to be

$$P_\mathrm{L} \equiv \tfrac{1}{2}(1-\gamma^5) = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

$$P_\mathrm{R} \equiv \tfrac{1}{2}(1+\gamma^5) = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}\ .$$

Any spinor $\psi$ can trivially be decomposed as

$$\psi = \psi_{\rm L} + \psi_{\rm R} \tag{44}$$

where $\psi_{\rm L}$ and $\psi_{\rm R}$ are the *left-handed* and *right-handed chiral components* of $\psi$, defined as

$$\psi_{\rm L} \equiv \tfrac{1}{2}(1-\gamma^5)\psi = P_{\rm L}\psi \tag{45}$$
$$\psi_{\rm R} \equiv \tfrac{1}{2}(1+\gamma^5)\psi = P_{\rm R}\psi \tag{46}$$

Using Equation (43), we have

$$\begin{aligned} P_{\rm L}u_\uparrow &= \tfrac{1}{2}(1-\gamma^5)u_\uparrow = 0 & \qquad P_{\rm R}u_\uparrow &= \tfrac{1}{2}(1+\gamma^5)u_\uparrow = u_\uparrow \\ P_{\rm L}u_\downarrow &= \tfrac{1}{2}(1-\gamma^5)u_\downarrow = u_\downarrow & \qquad P_{\rm R}u_\downarrow &= \tfrac{1}{2}(1+\gamma^5)u_\downarrow = 0\ . \end{aligned}$$

A general particle spinor $u$ can always be expressed as a linear combination of the basis spinors $u_\uparrow$ and $u_\downarrow$: $u = \alpha_1 u_\uparrow + \alpha_2 u_\downarrow$ where $\alpha_1$ and $\alpha_2$ are constants with $|\alpha_1|^2 + |\alpha_2|^2 = 1$. The chiral components of this general spinor are:

$$\begin{aligned} P_{\rm L}u &= \tfrac{1}{2}(1-\gamma^5)u = \alpha_2 u_\downarrow & \qquad (h=-1) \\ P_{\rm R}u &= \tfrac{1}{2}(1+\gamma^5)u = \alpha_1 u_\uparrow & \qquad (h=+1)\ . \end{aligned}$$

Thus, the operators $P_{\rm L}$ and $P_{\rm R}$ project out the negative (left-handed) and positive (right-handed) helicity components, respectively, of a general particle spinor. Specifically, for any particle spinor $u$, the projection $P_{\rm L}u = \frac{1}{2}(1-\gamma^5)u$ is always a negative helicity (left-handed) eigenstate [1], while the projection $P_{\rm R}u = \frac{1}{2}(1+\gamma^5)u$ is always a positive helicity (right-handed) eigenstate.

For an *antiparticle*, we have

$$\begin{aligned} P_{\rm L}v_\uparrow &= \tfrac{1}{2}(1-\gamma^5)v_\uparrow = v_\uparrow & \qquad P_{\rm R}v_\uparrow &= \tfrac{1}{2}(1+\gamma^5)v_\uparrow = 0 \\ P_{\rm L}v_\downarrow &= \tfrac{1}{2}(1-\gamma^5)v_\downarrow = 0 & \qquad P_{\rm R}v_\downarrow &= \tfrac{1}{2}(1+\gamma^5)v_\downarrow = v_\downarrow\ . \end{aligned}$$

The operator $P_{\rm L}$ now projects out the *positive* (*right*-handed) helicity component while $P_{\rm R}$ projects out the *negative* (*left*-handed) component.

In summary, in the extreme relativistic limit, for any particle spinor $u$ or antiparticle spinor $v$ we have:

| | Left-handed<br>$(h=-1)$ | Right-handed<br>$(h=+1)$ |
|---|---|---|
| Particle | $\frac{1}{2}(1-\gamma^5)u\ \ (= P_{\rm L}u)$ | $\frac{1}{2}(1+\gamma^5)u\ \ (= P_{\rm R}u)$ |
| Antiparticle | $\frac{1}{2}(1+\gamma^5)v\ \ (= P_{\rm R}v)$ | $\frac{1}{2}(1-\gamma^5)v\ \ (= P_{\rm L}v)$ |

It should be emphasised that the definition of the left-handed and right-handed chiral components in Equations (45) and (46) applies quite generally to any spinor $\psi$. In the extreme relativistic limit $E \gg m$, and only in this limit, the chiral components become states with definite helicity, *i.e.* left-handed or right-handed helicity eigenstates. Note that, for antiparticles, the *left*-handed projection operator $P_{\rm L}$ projects out the *right*-handed helicity eigenstate, while the *right*-handed projection operator $P_{\rm R}$ projects out the *left*-handed helicity eigenstate.

---

[1] except for the special case that $u$ is a pure right-handed state, in which case $P_{\rm L}u = \frac{1}{2}(1-\gamma^5)u = 0$

## 2.17 Magnetic Moment of a Dirac Particle (non-examinable)

In the Part II course on Relativity and Electrodynamics, it is shown that the equation of motion for a particle of charge $q$ in an electromagnetic field $\boldsymbol{A}^{\mu} = (\phi, \boldsymbol{A})$ can be obtained by making the *minimal substitution*

$$\boldsymbol{p} \to \boldsymbol{p} - q\boldsymbol{A}; \qquad E \to E - q\phi.$$

Applying this prescription to Equations (30) and (31), the plane wave solutions for a spin $\frac{1}{2}$ particle in an electromagnetic field become

$$(\boldsymbol{\sigma}.\boldsymbol{p} - q\boldsymbol{\sigma}.\boldsymbol{A})u_B = (E - m - q\phi)u_A \tag{47}$$

$$(\boldsymbol{\sigma}.\boldsymbol{p} - q\boldsymbol{\sigma}.\boldsymbol{A})u_A = (E + m - q\phi)u_B. \tag{48}$$

Multiplying Equation (47) by $E + m - q\phi$ gives

$$(\boldsymbol{\sigma}.\boldsymbol{p} - q\boldsymbol{\sigma}.\boldsymbol{A})(E + m - q\phi)u_B = (T + 2m - q\phi)(T - q\phi)u_A$$

where $T \equiv E - m$ is the kinetic energy of the particle. Taking the non-relativistic limit $T \ll m$, and assuming that the electric potential energy is small compared to the rest mass energy, $|q\phi| \ll m$, we obtain

$$(\boldsymbol{\sigma}.\boldsymbol{p} - q\boldsymbol{\sigma}.\boldsymbol{A})(\boldsymbol{\sigma}.\boldsymbol{p} - q\boldsymbol{\sigma}.\boldsymbol{A})u_A \approx 2m(T - q\phi)u_A \ .$$

Multiplying this out gives

$$\left[(\boldsymbol{\sigma}.\boldsymbol{p})^2 - q(\boldsymbol{\sigma}.\boldsymbol{A})(\boldsymbol{\sigma}.\boldsymbol{p}) - q(\boldsymbol{\sigma}.\boldsymbol{p})(\boldsymbol{\sigma}.\boldsymbol{A}) + q^2(\boldsymbol{\sigma}.\boldsymbol{A})^2\right] u_A \approx 2m(T - q\phi)u_A \ . \tag{49}$$

Using Equation (32), for any two vector operators $\boldsymbol{A}$ and $\boldsymbol{B}$ we have

$$\boldsymbol{\sigma}.\boldsymbol{A} = \begin{pmatrix} A_z & A_x - iA_y \\ A_x + iA_y & -A_z \end{pmatrix}; \qquad \boldsymbol{\sigma}.\boldsymbol{B} = \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix} \ .$$

It is then straightforward to check that

$$(\boldsymbol{\sigma}.\boldsymbol{A})(\boldsymbol{\sigma}.\boldsymbol{B}) = \boldsymbol{A}.\boldsymbol{B} + i\boldsymbol{\sigma} \cdot \boldsymbol{A} \wedge \boldsymbol{B} \ .$$

This contains the special cases

$$(\boldsymbol{\sigma}.\boldsymbol{p})^2 = |\boldsymbol{p}|^2, \qquad (\boldsymbol{\sigma}.\boldsymbol{A})^2 = |\boldsymbol{A}|^2 \ .$$

Hence the operator on the left-hand side of Equation (49) becomes

$$\begin{aligned}
&\boldsymbol{p}^2 - q\left[\boldsymbol{A}.\boldsymbol{p} + i\boldsymbol{\sigma} \cdot \boldsymbol{A} \wedge \boldsymbol{p} + \boldsymbol{p}.\boldsymbol{A} + i\boldsymbol{\sigma} \cdot \boldsymbol{p} \wedge \boldsymbol{A}\right] + q^2\boldsymbol{A}^2 \\
&= (\boldsymbol{p} - q\boldsymbol{A})^2 - iq\boldsymbol{\sigma} \cdot \left[\boldsymbol{A} \wedge \boldsymbol{p} + \boldsymbol{p} \wedge \boldsymbol{A}\right] \\
&= (\boldsymbol{p} - q\boldsymbol{A})^2 - iq\boldsymbol{\sigma} \cdot -i\left[\boldsymbol{A} \wedge \boldsymbol{\nabla} + \boldsymbol{\nabla} \wedge \boldsymbol{A}\right] \\
&= (\boldsymbol{p} - q\boldsymbol{A})^2 - iq\boldsymbol{\sigma} \cdot -i\boldsymbol{\nabla} \wedge \boldsymbol{A} \\
&= (\boldsymbol{p} - q\boldsymbol{A})^2 - q\boldsymbol{\sigma}.\boldsymbol{B}
\end{aligned}$$

where we have set $\boldsymbol{p} = -i\boldsymbol{\nabla}$, $\boldsymbol{B} = \boldsymbol{\nabla} \wedge \boldsymbol{A}$, and made use of the identity

$$(\boldsymbol{\nabla} \wedge \boldsymbol{A})\psi = \boldsymbol{\nabla} \wedge (\boldsymbol{A}\psi) + \boldsymbol{A} \wedge (\boldsymbol{\nabla}\psi) \ .$$

Substituting into Equation (49) gives the *Schrodinger-Pauli equation* describing the motion of a non-relativistic spin $\frac{1}{2}$ particle in an electromagnetic field:

$$\boxed{\left[\frac{1}{2m}(\boldsymbol{p}-q\boldsymbol{A})^2 - \frac{q}{2m}\boldsymbol{\sigma}.\boldsymbol{B} + q\phi\right] u_A = T u_A} .$$

Since the energy of a magnetic moment $\boldsymbol{\mu}$ in a magnetic field $\boldsymbol{B}$ is $-\boldsymbol{\mu}.\boldsymbol{B}$, we can associate with the Dirac particle an intrinsic (or spin) magnetic moment

$$\boldsymbol{\mu} = \frac{q}{2m}\boldsymbol{\sigma} .$$

In terms of the spin, $\boldsymbol{S} = \frac{1}{2}\boldsymbol{\sigma}$, this can be expressed as

$$\boxed{\boldsymbol{\mu} = g\frac{q}{2m}\boldsymbol{S}}$$

where the *gyromagnetic ratio* is $g = 2$. This reflects the fact that the magnetic moment associated with the spin angular momentum is twice as large as would be expected classically.

## 2.18 Intrinsic Parity of a Dirac Particle (non-examinable)

The parity operation $\hat{P}$ is defined as spatial inversion through the origin:

$$x \to -x, \qquad y \to -y, \qquad z \to -z .$$

In terms of transformed coordinates

$$x' \equiv -x, \qquad y' \equiv -y, \qquad z' \equiv -z, \qquad t' \equiv t ,$$

the Dirac equation

$$i\gamma^1\frac{\partial\psi}{\partial x} + i\gamma^2\frac{\partial\psi}{\partial y} + i\gamma^3\frac{\partial\psi}{\partial z} - m\psi = -i\gamma^0\frac{\partial\psi}{\partial t} \qquad (27)'$$

becomes

$$-i\gamma^1\frac{\partial\psi}{\partial x'} - i\gamma^2\frac{\partial\psi}{\partial y'} - i\gamma^3\frac{\partial\psi}{\partial z'} - m\psi = -i\gamma^0\frac{\partial\psi'}{\partial t} .$$

Now define the spinor $\psi'$ as

$$\psi'(x',t') \equiv \gamma^0\psi(x,t) . \qquad (50)$$

Since $(\gamma^0)^2 = I_4$, this can be inverted to give

$$\psi(x,t) = \gamma^0\psi'(x',t') .$$

The Dirac equation then becomes

$$-i\gamma^1\gamma^0\frac{\partial\psi'}{\partial x'} - i\gamma^2\gamma^0\frac{\partial\psi'}{\partial y'} - i\gamma^3\gamma^0\frac{\partial\psi'}{\partial z'} - m\gamma^0\psi' = -i\left(\gamma^0\right)^2\frac{\partial\psi'}{\partial t'} .$$

Since $\gamma^0$ anticommutes with $\gamma^1$, $\gamma^2$, $\gamma^3$, we obtain

$$i\gamma^0\gamma^1\frac{\partial\psi'}{\partial x'} + i\gamma^0\gamma^2\frac{\partial\psi'}{\partial y'} + i\gamma^0\gamma^3\frac{\partial\psi'}{\partial z'} - m\gamma^0\psi' = -i\frac{\partial\psi'}{\partial t'}$$

Premultiplying both sides by $\gamma^0$ then gives

$$i\gamma^1\frac{\partial\psi'}{\partial x'} + i\gamma^2\frac{\partial\psi'}{\partial y'} + i\gamma^3\frac{\partial\psi'}{\partial z'} - m\psi' = -i\gamma^0\frac{\partial\psi'}{\partial t'} \ .$$

This is of the same form as the original equation, Equation (27), but now in terms of the primed coordinates.

In summary, under a parity transformation, the form of the Dirac equation remains unchanged provided that Dirac spinors are transformed as

$$\boxed{\psi \to \psi' = \hat{P}\psi = \gamma^0\psi} \ .$$

Equivalently, if $\psi$ is a solution of the Dirac equation in the original frame, then $\psi' = \gamma^0\psi$ is a solution of the Dirac equation in the space-inverted frame.

For a particle or antiparticle at rest, the free particle solutions to the Dirac equation in the original frame are $\psi = u_1 e^{-imt}$, $u_2 e^{-imt}$, $v_1 e^{imt}$ and $v_2 e^{imt}$ where

$$u_1 = N\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \ ; \quad u_2 = N\begin{pmatrix}0\\1\\0\\0\end{pmatrix} \ ; \quad v_1 = N\begin{pmatrix}0\\0\\0\\1\end{pmatrix} \ ; \quad v_2 = N\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$$

and $N = \sqrt{2m}$. For the spinor $u_1$, for example, from Equation (50), the solution $\psi = u_1 e^{-imt}$ transforms under parity to $\psi' \equiv \hat{P}\psi$ where

$$\psi' = \gamma^0\psi = \gamma^0 u_1 e^{-imt} = \begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & -1 & 0\\0 & 0 & 0 & -1\end{pmatrix} \cdot N\begin{pmatrix}1\\0\\0\\0\end{pmatrix} e^{-imt} = N\begin{pmatrix}1\\0\\0\\0\end{pmatrix} e^{-imt} = \psi \ .$$

Thus $\psi' \equiv \hat{P}\psi = \psi$, or equivalently $u_1' \equiv \hat{P}u_1 = +u_1$: a particle at rest in the state $u_1$ is an eigenstate of parity with eigenvalue $+1$.

A similar treatment of the remaining solutions $u_2 e^{-imt}$, $v_1 e^{imt}$, $v_2 e^{imt}$ gives altogether

$$u_1' = +u_1, \qquad u_2' = +u_2, \qquad v_1' = -v_1, \qquad v_2' = -v_2 \ .$$

Hence an *antiparticle* at rest ($v_1, v_2$) has *opposite* intrinsic parity to a *particle* at rest ($u_1, u_2$):

$$\begin{aligned}\hat{P}\psi &= +\psi \qquad \text{for } u_1, u_2\\ \hat{P}\psi &= -\psi \qquad \text{for } v_1, v_2.\end{aligned}$$

In fact, it is only the *relative* intrinsic parities of particles and antiparticles which can unambiguously be determined; the *absolute* parities are assigned by convention to be $+1$ for particles and $-1$ for antiparticles. For example, we could equally well have chosen $\psi' = -\gamma^0\psi$ in place of Equation (50), which would have given parity eigenvalues $-1$ for $u_1$,$u_2$ and $+1$ for $v_1$,$v_2$. [2]

---

[2] This is reminiscent of electric charge: particles and antiparticles have opposite electric charge but the assignment of negative charge to the electron is purely conventional.

## 2.19 Time Reversal Invariance (non-examinable)

The effect of reversing the time coordinate $t$ can be treated in a similar fashion to the analysis of the parity operation. The transformed coordinates are now defined as

$$x' \equiv x, \qquad y' \equiv y, \qquad z' \equiv z, \qquad t' \equiv -t\ ,$$

and the Dirac equation

$$i\gamma^1\frac{\partial\psi}{\partial x} + i\gamma^2\frac{\partial\psi}{\partial y} + i\gamma^3\frac{\partial\psi}{\partial z} - m\psi = -i\gamma^0\frac{\partial\psi}{\partial t} \tag{27$'$}$$

becomes

$$i\gamma^1\frac{\partial\psi}{\partial x'} + i\gamma^2\frac{\partial\psi}{\partial y'} + i\gamma^3\frac{\partial\psi}{\partial z'} - m\psi = i\gamma^0\frac{\partial\psi'}{\partial t'}\ . \tag{51}$$

Now define the spinor $\psi'$ as

$$\psi'(x',t') \equiv i\gamma^1\gamma^3\psi^*(x,t)\ . \tag{52}$$

Since $(\gamma^1)^2 = (\gamma^3)^2 = -I_4$, premultiplying by $\gamma^1$ gives

$$\gamma^1\psi' = -i\gamma^3\psi^*\ .$$

Premultiplying by $\gamma^3$ we obtain

$$\gamma^3\gamma^1\psi' = i\psi^*\ ,$$

and taking the complex conjugate of this equation then gives

$$\psi(x,t) = i\gamma^3\gamma^1\psi'^*(x',t')\ .$$

Substituting for $\psi$ in Equation (51) and cancelling a factor of $i$, the Dirac equation becomes

$$i\gamma^1\gamma^3\gamma^1\frac{\partial\psi'^*}{\partial x'} + i\gamma^2\gamma^3\gamma^1\frac{\partial\psi'^*}{\partial y'} + i\gamma^3\gamma^3\gamma^1\frac{\partial\psi'^*}{\partial z'} - m\gamma^3\gamma^1\psi'^* = i\gamma^0\gamma^3\gamma^1\frac{\partial\psi'^*}{\partial t'}\ .$$

(Anti)commuting the $\gamma$ matrices to put a factor $\gamma^3\gamma^1$ on the left, which then cancels gives

$$-i\gamma^1\frac{\partial\psi'^*}{\partial x'} + i\gamma^2\frac{\partial\psi'^*}{\partial y'} - i\gamma^3\frac{\partial\psi'^*}{\partial z'} - m\psi'^* = i\gamma^0\frac{\partial\psi'^*}{\partial t'}\ .$$

Taking the complex conjugate and using the fact that $(\gamma^2)^* = -\gamma^2$ finally gives

$$i\gamma^1\frac{\partial\psi'}{\partial x'} + i\gamma^2\frac{\partial\psi'}{\partial y'} + i\gamma^3\frac{\partial\psi'}{\partial z'} - m\psi' = -i\gamma^0\frac{\partial\psi'}{\partial t'}\ .$$

This is of the same form as the original equation, Equation (27), but now expressed in terms of the primed coordinates.

In summary, under a time-reversal transformation, the form of the Dirac equation remains unchanged provided that Dirac spinors $\psi$ are transformed as

$$\boxed{\psi \to \psi' = \hat{T}\psi = i\gamma^1\gamma^3\psi^*}\ .$$

## 2.20 Charge Conjugation (non-examinable)

Since $p^\mu = (E, \boldsymbol{p}) = i\partial^\mu$ and $A^\mu = (\phi, \boldsymbol{A})$, the minimal substitution

$$\boldsymbol{p} \to \boldsymbol{p} - e\boldsymbol{A}; \qquad E \to E - e\phi$$

of Section 2.17 can be written more compactly as

$$\partial_\mu \to \partial_\mu + ieA_\mu .$$

The equation of motion for a Dirac particle in an electromagnetic field is then [3]

$$\boxed{\gamma^\mu(\partial_\mu + ieA_\mu)\psi + im\psi = 0} . \tag{53}$$

The complex conjugate of this equation is

$$\gamma^{\mu *}(\partial_\mu - ieA_\mu)\psi^* - im\psi^* = 0.$$

Premultiplying by $-i\gamma^2$ then gives

$$-i\gamma^2\gamma^{\mu *}(\partial_\mu - ieA_\mu)\psi^* - m\gamma^2\psi^* = 0.$$

In the Pauli-Dirac representation of the gamma matrices, only $\gamma^2$ is imaginary:

$$\gamma^{0*} = \gamma^0; \qquad \gamma^{1*} = \gamma^1; \qquad \gamma^{2*} = -\gamma^2; \qquad \gamma^{3*} = \gamma^3$$

and it can easily be checked that

$$\gamma^2\gamma^{\mu *} = -\gamma^\mu\gamma^2.$$

Hence we obtain

$$i\gamma^{\mu *}\gamma^2(\partial_\mu - ieA_\mu)\psi^* - m\gamma^2\psi^* = 0.$$

Introducing the *charge conjugation* transformation

$$\boxed{\psi' \equiv \hat{C}\psi = i\gamma^2\psi^*} ,$$

the above equation can be expressed as

$$\boxed{\gamma^\mu(\partial_\mu - ieA_\mu)\psi' + im\psi' = 0} .$$

This is of the same form as the original (untransformed) Dirac equation, Equation (53), but with charge $-e$ in place of $e$. In other words, the spinor $\psi'$ describes a particle with the same mass $m$ as $\psi$ but with opposite electric charge, namely the *antiparticle*. The charge conjugation operation transforms particles into antiparticles (and *vice versa*).

We now consider how the free particle (plane wave) solutions of the Dirac equation transform under charge conjugation. For example the *particle* plane wave solution

$$\psi = u_1 e^{i(\boldsymbol{p}.\boldsymbol{r} - Et)}$$

---

[3] The term $ie\gamma^\mu A_\mu\psi$ in this equation contains both the electromagnetic field $A_\mu$ and the particle spinor $\psi$. It ultimately gives rise to the vertex factor $-ie\gamma^\mu$ in the QED Feynamn rules for the interaction between a spin-half particle and a photon.

becomes

$$\psi' = i\gamma^2\psi^* = i\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x+ip_y)/(E+m) \end{pmatrix}^* e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)}$$

$$= \sqrt{E+m}\begin{pmatrix} (p_x-ip_y)/(E+m) \\ -p_z/(E+m) \\ 0 \\ 1 \end{pmatrix} e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)} = v_1 e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)}$$

which represents a free *antiparticle*. Thus, under charge conjugation,

$$\psi = u_1 e^{i(\boldsymbol{p}.\boldsymbol{r}-Et)} \quad \xrightarrow{C} \quad \psi' = v_1 e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)} \ .$$

Similarly, it can be shown that

$$\psi = u_2 e^{i(\boldsymbol{p}.\boldsymbol{r}-Et)} \quad \xrightarrow{C} \quad \psi' = -v_2 e^{-i(\boldsymbol{p}.\boldsymbol{r}-Et)} \ .$$

Thus, under charge conjugation, the particle spinors $u_1$ and $u_2$ transform to the antiparticle spinors $v_1$ and $v_2$, respectively.

## 2.21 CPT and Antiparticles (non-examinable)

To summarise the results of the previous three sections, the Dirac equation remains invariant under the operations of parity, $P$, time-reversal, $T$, and charge conjugation, $C$, provided the wavefunction $\psi$ transforms in each case as:

$$\begin{aligned} P: &\qquad \psi' = \gamma^0\psi \\ T: &\qquad \psi' = i\gamma^1\gamma^3\psi^* \\ C: &\qquad \psi' = i\gamma^2\psi^* \end{aligned}$$

We now consider the combined effect of applying all three of these operators together. Applying $P$ followed by $T$ gives (since $\gamma^0$ is real)

$$\begin{aligned} \psi' &= i\gamma^1\gamma^3(\gamma^0\psi)^* \\ &= i\gamma^1\gamma^3\gamma^0\psi^* \ . \end{aligned}$$

Then applying $C$ gives

$$\begin{aligned} \psi' &= i\gamma^2(i\gamma^1\gamma^3\gamma^0\psi^*)^* \\ &= \gamma^2\gamma^1\gamma^3\gamma^0\psi \\ &= -\gamma^1\gamma^2\gamma^3\gamma^0\psi \\ &= \gamma^0\gamma^1\gamma^2\gamma^3\psi \\ &= -\gamma^5\psi \ . \end{aligned}$$

Thus the overall effect of the combined operator $CPT$ is that $\psi$ transforms as

$$\psi'(\boldsymbol{r'}, t') = -\gamma^5 \psi(\boldsymbol{r}, t) \ ,$$

where, due to the combined effect of $P$ and $T$, all four spacetime coordinates have now been reversed:

$$x' \equiv -x, \qquad y' \equiv -y, \qquad z' \equiv -z, \qquad t' \equiv -t \ .$$

In addition, the effect of $C$ is to reverse the sign of the electric charge. The same overall result would be obtained applying $C$, $P$ and $T$ in any order.

Consider for example the negative energy solution

$$\psi(\boldsymbol{r}, t) = u_3(E, \boldsymbol{p}) e^{i(\boldsymbol{p.r} - Et)} = N \begin{pmatrix} p_z/(E-m) \\ (p_x + ip_y)/(E-m) \\ 1 \\ 0 \end{pmatrix} e^{i(\boldsymbol{p.r} - Et)}$$

with energy $E < 0$. Applying a CPT transformation gives

$$\begin{aligned} \psi'(\boldsymbol{r'}, t') &= - \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} N \begin{pmatrix} p_z/(E-m) \\ (p_x + ip_y)/(E-m) \\ 1 \\ 0 \end{pmatrix} e^{i(\boldsymbol{p.r} - Et)} \\ &= -N \begin{pmatrix} 1 \\ 0 \\ p_z/(E-m) \\ (p_x + ip_y)/(E-m) \end{pmatrix} e^{-i(\boldsymbol{p.r'} - Et')} \ . \end{aligned}$$

Apart from an arbitrary overall factor of $-1$, this is of the same form as the standard positive energy particle solution $u_1$. The overall effect of a CPT transformation can therefore be summarised as

$$\psi = u_3(E, \boldsymbol{p}) e^{i(\boldsymbol{p.r} - Et)} \quad \overset{CPT}{\longrightarrow} \quad \psi' = -u_1(E', \boldsymbol{p'}) e^{i(\boldsymbol{p'.r'} - E't')} \ .$$

where $E' \equiv -E > 0$ and $\boldsymbol{p}' \equiv -\boldsymbol{p}$.

In summary, viewed from a frame where all space and time coordinates are inverted, the original negative energy state with energy $E < 0$, momentum $\boldsymbol{p}$ and charge $q$ appears as a positive energy particle with energy $-E > 0$, momentum $-\boldsymbol{p}$ and charge $-q$. The assumption that a particle travelling "backwards in spacetime" is manifested in the laboratory as an antiparticle travelling "forwards in spacetime" then completes the Feynman interpretation of the negative energy solutions.