## Chapter 4

## **DIRAC EQUATION**

The non-relativistic Schrödinger equation was obtained by noting that the Hamiltonian

$$H = \frac{P^2}{2m} \tag{1}$$

can be transformed into an operator form with the substitutions

 $H \to i \frac{\partial}{\partial t} (\hbar = 1, c = 1)$   $P \to -i \vec{\nabla}$ (2)

$$i\frac{\partial}{\partial t}\psi(x,t) = \frac{\nabla^2}{2m}\psi(x,t)$$
(3)

This equation is linear in the time-derivative and quadratic in the space derivative, and is manifestly non-covariant under Lorentz transformation. Clearly, the left-hand side and the right-hand side of Equation 3 transforms differently under Lorentz transformation.

We briefly review Lorentz transformation.

Space-time coordinates (t, x, y, z) are denoted by the contravariant 4-vector

$$x^{\mu} \equiv (x^{o}, x^{1}, x^{2}, x^{3}) \equiv (t, x, y, z)$$

The covariant 4-vector,  $x_{\mu}$ , is

$$x_{\mu} \equiv (x_{o}, x_{1}, x_{2}, x_{3}) \equiv (t, -x, -y, -z) = g_{\mu\nu} x^{\nu}$$

$$g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
(4)

 $g_{\mu\nu}$  is the metric tensor

The Lorentz transformation relates  $x^{\mu}$  to  $x^{\mu}$ :

$$\left(x^{\mu}\right)' = \Lambda^{\mu}_{\ \nu} x^{\nu} \tag{5}$$

For an inertial system S' moving along z with  $\frac{v}{c} = \beta$ , we have

$$\Lambda = \begin{bmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{bmatrix}$$
(6)

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

Any set of four quantities which transform like  $x^{\mu}$  under Lorentz transformation is called a four-vector. The total energy E and the momentum  $\vec{P}$  form a four-vector

$$P^{\mu} = (P^{o}, P^{1}, P^{2}, P^{3}) = (E, \vec{P})$$
 (7)

Scalar product of two 4-vectors are invariant under Lorentz transformation

$$A_{\mu}B^{\mu} = g_{\mu\nu}A^{\nu}B^{\mu} \tag{8}$$

Some examples:

$$x \bullet x = x_{\mu} x^{\mu} = t^{2} - x^{2}$$
$$P \bullet P = P_{\mu} P^{\mu} = E^{2} - P^{2}$$
$$x \bullet P = x_{\mu} P^{\mu} = t \ E - \vec{x} \bullet \vec{P}$$

The  $H \rightarrow i \frac{\partial}{\partial t}$ ,  $\vec{P} \rightarrow -i \vec{\nabla}$  transcription is Lorentz covariant

$$P^{\mu} \to i \frac{\partial}{\partial x_{\mu}} = i \partial^{\mu}$$

One can readily obtain the continuity equation from the Schrödinger equation:

P570

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \tag{9}$$

where  $\rho = \psi^* \psi \ge 0$ ,  $\vec{J} = -\frac{i}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*)$ 

 $\rho$  is the probability density, and  $\vec{J}$  is the current density.

Note that the continuity equation can be written in a covariant form

$$\partial_{\mu}j^{\mu} = 0, \quad j^{\mu} = \left(\rho, \vec{J}\right) \tag{10}$$

The early attempt for formulating a Lorentz covariant equation started from the relativistic energy-momentum relation

$$E^2 = P^2 + m^2$$

with the  $E \rightarrow i \frac{\partial}{\partial t}, \ \vec{P} \rightarrow -i \vec{\nabla}$  transcription, one obtains

$$-\frac{\partial^2}{\partial t^2}\psi = \left(-\nabla^2 + m^2\right)\psi \tag{11}$$

which is the relativistic Schrödinger equation, or the Klein-Gordon equation.

Equation 11 can be written as

$$\left(\Box^2 + m^2\right)\psi = 0 \tag{12}$$

$$\Box^{2} = \frac{\partial^{2}}{\partial t^{2}} - \nabla^{2} = \frac{\partial}{\partial x^{\mu}} \frac{\partial}{\partial x_{\mu}} = \partial_{\mu} \partial^{\mu}$$
(13)

is the d'Alembertian operator, a Lorentz invariant operator.

From the Klein-Gordon equation, one can readily derive the following continuity equation:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0$$

$$\rho = i \left( \phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t} \right) \bigg|_{j^{\mu}} = i \left( \phi^* \partial^{\mu} \phi - \phi \partial^{\mu} \phi^* \right)$$
(14)

$$\vec{J} = -i\left(\phi * \vec{\nabla}\phi - \phi \vec{\nabla}\phi *\right) \int^{J} (15)$$

For a plane-wave solution to Klein-Gordon equation

$$\phi = N e^{-i\left(Et - \vec{P} \cdot \vec{x}\right)} \tag{16}$$

We have

$$\rho = 2E|N|^2, \ \vec{J} = 2\vec{P}|N|^2$$
(17)

Since  $E = \pm (P^2 + m^2)^{\frac{1}{2}}$ ,  $\rho$  can be negative for the E < 0 solutions.

The origin of this problem is the second derivative in time for the term  $\partial^2/\partial t^2$ .

Dirac tried to find a relativistic covariant equation with a positive-definite probability density. He assumed an equation linear in  $\frac{\partial}{\partial t}$  and in  $\vec{\nabla}$ :

$$H\psi = \left(\vec{\alpha} \cdot \vec{P} + \beta m\right)\psi$$

$$\vec{\alpha} \cdot \vec{P} = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3$$
(18)

It is evident that  $\alpha$ 's and  $\beta$  cannot be just numbers, because the equation would not be invariant even under spatial rotation. Dirac proposed that the equation be considered as a matrix equation. It should satisfy  $H^2\psi = (P^2 + m^2)\psi$ .

Applying *H* to Equation 18 leads to

$$H^{2}\psi = \left[\sum_{i} \alpha_{i}^{2} P_{i}^{2} + \sum_{i \neq j} (\alpha_{i} \alpha_{j} + \alpha_{j} \alpha_{i}) P_{i} P_{j} + \sum_{i} (\alpha_{i} \beta + \beta \alpha_{i}) m P_{i} + \beta^{2} m^{2}\right]\psi$$

Therefore

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} ; \{\alpha_i, \beta\} = 0$$
$$\alpha_i^2 = \beta^2 = 1$$

What are the other properties of  $\alpha_i$ ,  $\beta$ ? First,  $\alpha_i$ ,  $\beta$  are Hermitian, since H is Hermitian (see Equation 18).

In addition,  $Tr(\alpha_i) = 0$ ,  $Tr(\beta) = 0$ . This can be shown by the anticommutator relations in Equation 19. Recall Tr(AB) = Tr(BA)

$$Tr(\alpha_{i}) = Tr(\alpha_{i}\beta^{2}) = Tr(\beta\alpha_{i}\beta) = Tr(-\alpha_{i}\beta\beta)$$
$$= -Tr(\alpha_{i}) \Rightarrow Tr(\alpha_{i}) = 0$$
(20)

Similarly,  $Tr(\beta) = 0$ . Therefore,  $\alpha_i$ ,  $\beta$  are traceless, Hermitian matrices.

The eigenvalues of  $\alpha_i$ ,  $\beta$  can only be +1 or -1 since  $\alpha_i^2 = \beta^2 = 1$ .

$$\alpha_i x = \lambda x \implies \alpha_i^2 x = \lambda^2 x = x \implies \lambda = \pm 1$$
(21)

The trace of a matrix is equal to the sum of all its eigenvalues. Therefore, for a traceless matrix having +1, -1 as possible eigenvalues, it must have an even number of dimension *n*. *n* cannot be 2, since only three traceless n = 2 independent matrices exist ( $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$ , for example), and we need four traceless matrices to accommodate  $\alpha_1, \alpha_2, \alpha_3, \beta$ .

n = 4 is the smallest possible dimension for  $\alpha_i$ ,  $\beta$ . The Pauli-Dirac representation is given as

$$\alpha_{i} = \begin{bmatrix} 0 & \sigma_{i} \\ \sigma_{i} & 0 \end{bmatrix} \qquad \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$
(22)

The Dirac equation  $\left[\vec{\alpha} \cdot \vec{P} + m\beta\right]\psi = H\psi$  can be expressed in a more covariant fashion by multiplying  $\beta$  from the left:

$$\left(\beta H - \left(\beta \vec{\alpha}\right) \cdot \vec{P} - m\right) \psi = 0 \tag{23}$$

Introducing  $\gamma^{\mu} = (\beta, \beta \vec{\alpha})$ , Equation 23 becomes

$$\left(\gamma^{\mu}P_{\mu}-m\right)\psi=0\tag{24}$$

 $P_{\mu} = i\partial_{\mu}$ , therefore

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0$$
 (recall  $\partial_{\mu} = \frac{\partial}{\partial x^{\mu}}$ ) (25)

Explicitly,

$$\gamma^{o} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}; \quad \gamma^{i} = \begin{pmatrix} 0 & \sigma_{i} \\ -\sigma_{i} & 0 \end{pmatrix}$$
(26)

 $\gamma^{\mu}$ 's satisfy the anticommutation relation

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

$$\left(\gamma^{o}\right)^{2} = I, \quad \left(\gamma^{i}\right)^{2} = -I$$
(27)

Although 
$$\gamma^{o}$$
 is Hermitian,  $(\gamma^{o})^{+} = \gamma^{o}$ ,  $\gamma^{i}$  is anti-Hermitian:  
 $(\gamma^{i})^{+} = -\gamma^{i}$ 
(28)

However,

$$(\gamma^{\mu})^{+} = \gamma^{o} \gamma^{\mu} \gamma^{o}$$
 and  $\gamma^{\mu} = \gamma^{o} (\gamma^{\mu})^{+} \gamma^{o}$  (29)

To obtain the adjoint Dirac equation, one takes Hermitian conjugate of Equation 25

$$-i\partial_{\mu}\psi^{+}(\gamma^{\mu})^{+} - m\psi^{+} = 0$$
(30)

Since  $\gamma^{\mu}$  is not Hermitian for  $\mu = 1, 2, 3$ , it is useful to multiply Equation 30 from the right by  $\gamma^{\circ}$  to yield (using Equation 29)

$$-i\partial_{\mu}\left(\psi^{+}\gamma^{o}\right)\gamma^{\mu}-m\left(\psi^{+}\gamma^{o}\right)=0$$
(31)

Define

$$\overline{\nu} = \psi^+ \gamma^o \tag{32}$$

Define  $\overline{\psi} = \psi^+ \gamma^o$ the adjoint Dirac equation is  $i\partial_\mu \overline{\psi} \gamma^\mu + m\overline{\psi} = 0$ (33) or

where

and *P* acts to its left in Equation 34.

One is now ready to check the continuity equation and the probability density and current density for the Dirac equation.

 $\overline{\psi} \left( \not\!\!\!P + m \right) = 0$  $\not\!\!\!P = \gamma^{\mu} P_{\mu}$ 

Multiplying Equation 25 by  $\overline{\psi}$  to its left, and Equation 33 by  $\psi$  to its right, we obtain

$$i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - \overline{\psi}m\psi = 0$$

$$i\partial_{\mu}\overline{\psi}\gamma^{\mu}\psi + m\overline{\psi}\psi = 0$$
(35)

Adding these two equations yields

$$\partial_{\mu} \left( \bar{\psi} \gamma^{\mu} \psi \right) = 0 \tag{36}$$

Define

$$J = \varphi T = \varphi$$

the continuity equation

 $\partial_{\mu} j^{\mu} = 0$  follows.

Note that 
$$\rho = j^o = \overline{\psi}\gamma^o\psi = \psi^+\psi = \sum_{i=1}^4 |\psi_i|^2 \ge 0$$

Hence the probability density is indeed positive definite for the Dirac equation.

 $i^{\mu} = \overline{W} \gamma^{\mu} W$ 

As will be shown later  $j^{\mu}$  transforms like a 4-vector, and  $\rho = j^{\circ}$  is not a Lorentz invariant. Rather, it is the time component of a 4-vector. It is not too surprising that  $\rho$  is not a Lorentz invariant, since a probability density satisfies the following:

$$\int \rho d^3 x = \text{const} \tag{38}$$

Since  $d^3x$  undergoes Lorentz contraction, it is natural that  $\rho$  must also change under the Lorentz transformation in order to preserve the integral.

(34)

(37)

Each component of the four Dirac wave function components satisfies the Klein-Gordon equation. Multiplying Equation 25 from the left by  $i\gamma^{\mu}\partial_{\mu}$ , we have

$$-\gamma^{\mu}\gamma^{\nu}\partial_{\mu}\partial_{\nu}\psi - m^{2}\psi = 0 \tag{39}$$

or

 $\left[\frac{1}{2}\left(\gamma^{\mu}\gamma^{\nu}+\gamma^{\nu}\gamma^{\mu}\right)\partial_{\mu}\partial_{\nu}+m^{2}\right]\psi=0$ 

 $\left(g^{\mu\nu}\partial_{\mu}\partial_{\nu}+m^{2}\right)\psi=0$ 

hence

or

$$\left(\partial_{\mu}\partial^{\mu} + m^{2}\right)\psi = 0 \tag{40}$$

Note that Equation 40 is understood as four separate uncoupled equations for each component of  $\psi$ .

We turn next to the solutions to the Dirac equation for a free particle.

Because of Equation 40, the Dirac equation admits a free-particle solution as follows

$$\psi = u(P)e^{-iP \cdot x} \tag{41}$$

where u(P) is independent of x and is a 4-component spinor. Substituting Equation 41 into Equation 25 gives

$$\left(\gamma^{\mu}P_{\mu}-m\right)u(P)=0, \quad \text{or} \quad \left(\not P-m\right)u(P)=0$$

$$(42)$$

Eq. 42 is the Dirca equation in the momentum space. Note that  $P_{\mu}$  in Eq. 42 is no longer a differential operator (just a multiplicative operator). We now simply write u(P) as u. Similarly, the adjoint equation is readily obtained:

We now try to find solutions to the Dirac Equation. For a Dirac particle at rest,  $\vec{p} = 0$ , and Equation 42 becomes

$$(\gamma^{o} \mathbf{E} - m) u = 0 \qquad (\psi = u e^{-i\mathbf{E}t})$$
 (44)

or

$$Eu = m\gamma^{o}u = \begin{pmatrix} m & & & \\ & m & & \\ & & -m & \\ & & & -m \end{pmatrix} u$$
(45)

The eigenvectors are

$$u_{1} = N \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \quad u_{2} = N \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; \quad u_{3} = N \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \quad u_{4} = N \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$
(46)

with eigenvalues m, m, -m, -m, respectively. N is the normalization factor.

To obtain the solutions for  $\vec{p} \neq 0$ , one can either solve the Dirac Equation directly, or by boosting the  $\psi(p=0)$  solution to another inertia frame moving with  $-\vec{p}$  with respect to the rest frame.

It is instructive to examine the method involving the Lorentz boost. To use this method, we need to first address the question – How does  $\psi$  transform under Lorentz transformation? Let us suppose that in reference frame *S*, the Dirac equation is written as

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x^{\mu}}-m\right)\psi(x)=0$$
(47)

In another frame S', where  $x^{\mu'} = \Lambda^{\mu}_{\nu} x^{\nu}$ , the Dirac equation becomes

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x^{\mu'}}-m\right)\psi'(x')=0$$
(48)

Now,

$$\psi'(x') = \psi'(\Lambda x) = S\psi(x) \tag{49}$$

and

$$\psi(x) = S^{-1}\psi'(x') \tag{50}$$

Substituting Equation 50 into Equation 47 and noting

$$\frac{\partial}{\partial x^{\mu}} = \frac{\partial}{\partial x^{\nu'}} \frac{\partial x^{\nu'}}{\partial x^{\mu}} = \Lambda^{\nu}{}_{\mu} \frac{\partial}{\partial x^{\nu'}}$$
(51)

Equation 47 becomes, upon multiplying S from the left,

$$\left(iS\gamma^{\mu}\Lambda^{\nu}{}_{\mu}\frac{\partial}{\partial x^{\nu'}}S^{-1}-m\right)\psi'(x')=0$$
(52)

Requiring Equation 52 to be identical to Equation 48 gives

$$iS\gamma^{\mu}\Lambda^{\nu}{}_{\mu}\frac{\partial}{\partial x^{\nu'}}S^{-1} = i\gamma^{\mu}\frac{\partial}{\partial x^{\mu'}}$$
$$S\gamma^{\mu}\Lambda^{\nu}{}_{\mu}S^{-1} = \gamma^{\nu}$$
$$\Lambda^{\nu}{}_{\mu}\gamma^{\mu} = S^{-1}\gamma^{\nu}S$$

or

Equivalently,

Our task is to find *S*, which is a 4 x 4 matrix, for any Lorentz transformation  $\Lambda$ , such that Equation 53 is satisfied.

We consider three cases of Lorentz transformation:

- 1) Lorentz boost
- 2) Spatial rotation
- 3) Space inversion (parity operation)

The cases 1) and 2) can be built up with infinitesimal Lorentz transformations.

Consider an infinitesimal Lorentz transformation

$$\Lambda^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \Delta \omega^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \Delta \omega (I_n)^{\nu}{}_{\mu}$$
(54)

P570

 $\Delta \omega$  is an infinitesimal quantity, and  $I_n$  is a 4 x 4 matrix corresponding to a specific Lorentz transformation.

For a transformation to S' moving along the x-axis with an infinitesimal velocity  $\Delta \omega = \Delta \beta$ , we have

Similarly, for a rotation through an angle  $\Delta \omega$  around the z-axis, the correspond  $I^{\nu}_{\mu}$  is

$$I^{\nu}{}_{\mu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad I^{\nu\mu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
(56)

In order to satisfy the relation for Lorentz transformation

$$\Lambda^{\nu}_{\ \mu}\Lambda^{\mu}_{\ \sigma} = \delta^{\nu}_{\ \sigma} \tag{57}$$

the infinitesimal transformation, Equation 54, satisfies

$$I^{\mu\nu} = -I^{\nu\mu} \tag{58}$$

as can be readily verified for Equations 55 and 56.

There are 6 independent  $I^{\mu\nu}$ , 3 for boosts and 3 for rotations.

Expand S in powers of  $\Delta \omega$ :

$$S = 1 - \frac{i}{4} \sigma_{\mu\nu} \Delta \omega^{\mu\nu} \tag{59}$$

 $\sigma_{\mu\nu}$  is the generator for the corresponding infinitesimal Lorentz transformation.

Inserting Equation 59 and Equation 54 into Equation 53, we have

$$2i\left[g^{\nu}_{\ \alpha} \ \gamma_{\beta} - g^{\nu}_{\ \beta} \ \gamma_{\alpha}\right] = \left[\gamma^{\nu}, \ \sigma_{\alpha\beta}\right]$$
(60)

which is satisfied by

$$\sigma_{\mu\nu} = \frac{i}{2} \Big[ \gamma_{\mu}, \gamma_{\nu} \Big] \tag{61}$$

The anti-symmetric tensor  $\sigma_{\mu\nu}$  has the following non-vanishing elements

$$\sigma_{oi} = -\sigma_{io} = -i\alpha_i \tag{62}$$

for boost and

$$\sigma_{ij} = -\sigma_{ji} = \begin{bmatrix} \sigma_k & 0\\ 0 & \sigma_k \end{bmatrix}$$
(63)

for rotation *ijk* cyclic.

We are ready to construct a finite spinor transformation S from successive infinitesimal transformation

$$S = \lim_{N \to \infty} \left( 1 - \frac{i}{4} \frac{\omega}{N} \sigma_{\mu\nu} I_n^{\mu\nu} \right)^N$$
  
=  $\exp\left( -\frac{i}{4} \omega \sigma_{\mu\nu} I_n^{\mu\nu} \right)$  (64)

To prove Equation 60

$$S = 1 - \frac{i}{4} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta}$$

$$S^{+} = 1 + \frac{i}{4} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta}$$

$$\Lambda^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \Delta \omega^{\nu}{}_{\mu}$$
Since
$$\Lambda^{\nu}{}_{\mu} \gamma^{\mu} = S^{-1} \gamma^{\nu} S$$
We have
$$\left(\delta^{\nu}{}_{\mu} + \Delta \omega^{\nu}{}_{\mu}\right) \gamma^{\mu} = \left(1 + \frac{i}{4} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta}\right) \gamma^{\nu} \left(1 - \frac{i}{4} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta}\right)$$

Sinc

or

$$\gamma^{\nu} + \Delta \omega^{\nu}{}_{\mu} \gamma^{\mu} = \gamma^{\nu} - \frac{i}{4} \gamma^{\nu} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta} + \frac{i}{4} \sigma_{\alpha\beta} \Delta \omega^{\alpha\beta} \gamma^{\nu}$$
  
(ignoring  $(\Delta \omega)^2$  term)

Therefore 
$$\Delta \omega^{\nu}{}_{\mu} \gamma^{\mu} = -\frac{i}{4} \Delta \omega^{\alpha\beta} [\gamma^{\nu}, \sigma_{\alpha\beta}]$$
 (a)

but

$$\Delta \omega^{\nu}{}_{\mu} \gamma^{\mu} = \Delta \omega^{\nu \mu} \gamma_{\mu} = \frac{1}{2} \Big[ \Delta \omega^{\nu \beta} \gamma_{\beta} + \Delta \omega^{\nu \alpha} \gamma_{\alpha} \Big]$$
$$= \frac{1}{2} \Big[ \Delta \omega^{\nu \beta} \gamma_{\beta} - \Delta \omega^{\alpha \nu} \gamma_{\alpha} \Big]$$
$$= \frac{1}{2} \Big[ g^{\nu}{}_{\alpha} \gamma_{\beta} - g^{\nu}{}_{\beta} \gamma_{\alpha} \Big] = \Big[ \gamma^{\nu}, \sigma_{\alpha\beta} \Big]$$
(b)

Inserting (b) into (a), we obtain

$$2i\left[g^{\nu}_{\ \alpha}\ \gamma_{\beta}-g^{\nu}_{\ \beta}\ \gamma_{\alpha}\right]=\left[\gamma^{\nu},\ \sigma_{\alpha\beta}\right]$$

For a Lorentz boost along the *x*-axis,

$$\psi'(x') = \exp\left[\left(-\frac{i}{2}\right)\omega \ \sigma_{01}\right]\psi(x)$$

$$= \exp\left(-\frac{\omega}{2}\alpha_{1}\right)\psi(x)$$
(65)

where  $\tanh \omega = \beta$ ,  $\cosh \omega = \frac{1}{\sqrt{1 - \beta^2}} = \gamma$ ,  $\sinh \omega = \beta \gamma$ 

Similarly, for a rotation around the *z*-axis of an angle  $\omega$ :

$$\psi'(x') = \exp\left[\frac{i}{2}\omega \ \sigma^{12}\right]\psi(x)$$

$$\sigma^{12} = \begin{bmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{bmatrix} = \begin{bmatrix}1 & 0\\ -1 & \\ 0 & -1\end{bmatrix}$$
(66)

where

Equation 66 is similar to the rotation of two-component Pauli spinor

$$\psi'(x') = e^{\binom{i}{2}\vec{\omega}\cdot\vec{\sigma}}\psi(x) \tag{67}$$

Are the spinor transformation *S* unitary?

For spatial rotation,  $S = S_{Rot}$ , S is unitary, since

$$S_{Rot}^{+} = e^{-\frac{i}{2}\omega(\sigma^{ij})^{+}} = e^{-\frac{i}{2}\omega(\sigma^{ij})} = S_{Rot}^{-1}$$
(68)

but for Lorentz boost

$$S_{Lor}^{+} = e^{\left(-\frac{\omega}{2}\right)\alpha_{i}^{+}} = e^{\left(-\frac{\omega}{2}\right)\alpha_{i}} = S_{Lor} \neq S_{Lor}^{-1}$$
(69)

Nevertheless, both  $S_{Rot}$  and  $S_{Lor}$  have the property

$$S^{-1} = \gamma_o S^+ \gamma_o \tag{70}$$

which can be verified by expanding  $S_{Rot}$ ,  $S_{Lor}$  in power series.

Here is proof for  $\beta$  = tanh  $\omega$ 

Consider a Lorentz boost along the *x*-axis

$$= \left(e^{\omega I}\right)_{\mu}^{\nu} x^{\mu}$$

$$= \left(\cosh \omega I + \sinh \omega I\right)_{\mu}^{\nu} x^{\mu}$$

$$= \left(1 - I^{2} + I^{2} \cosh \omega + I \sinh \omega\right)_{\mu}^{\nu} x^{\mu}$$

$$x' = \begin{bmatrix}\cosh \omega & -\sinh \omega & 0 & 0\\ -\sinh \omega & \cosh \omega & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} x$$

$$t' = \left(\cosh \omega\right) t - \left(\sinh \omega\right) x$$

$$x' = -\left(\sinh \omega\right) t + \left(\cosh \omega\right) x$$

$$t' = \gamma \left(t - \beta x\right)$$

$$x' = \gamma \left(-\beta t + x\right)$$

Therefore

We also have

$$\gamma = \cosh \omega$$
  $\gamma \beta = \sinh \omega$   $\beta = \tanh \omega$ 

The fact that  $S_{Lor}$  is not unitary (Equation 69) should not be too surprising. This simply reflects the situation that  $\psi^+\psi$  is not a conserved quantity in a Lorentz boost.

We now consider space inversion:

$$\Lambda^{\nu}{}_{\mu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

This discrete transformation cannot be constructed out of infinitesimal transformation. Rather, we rely on the definition of *S* (Equation 53) to find the parity transformation  $S_p$  for the Dirac spinor  $\psi$ .

Equation 53 and Equation 71 imply

$$S_{p}^{-1}\gamma^{o}S_{p} = \gamma^{o}$$

$$S_{p}^{-1}\gamma^{i}S_{p} = -\gamma^{i}$$
(72)

This can be satisfied with

$$S_{p} = \gamma^{o} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Note that  $S_p$  also satisfies Equation 70 since

$$S_p^{-1} = \gamma^o S_p^+ \gamma_o$$

For a Dirac particle at rest, Equation 73 shows that  $U_1$  and  $U_2$  have positive parity while  $U_3$  and  $U_4$  have negative parity (Equation 46)!

Having established the properties of *S*, we can show that  $\overline{\psi}\psi$  transforms like a scalar, while  $\overline{\psi}\gamma^{\mu}\psi$  transforms like a 4-vector under Lorentz transformation.

$$\overline{\psi}'(x') = \psi'(x')^{+} \gamma_{o} = \left(S\psi(x)\right)^{+} \gamma_{o} = \psi(x)^{+} S^{+} \gamma_{o}$$

$$= \psi(x)^{+} \gamma_{o} S^{-1} = \overline{\psi}(x) S^{-1}$$
(74)

Therefore

$$\overline{\psi}'(x')\psi'(x') = \overline{\psi}(x)S^{-1}S\psi(x) = \overline{\psi}(x)\psi(x)$$

$$\Rightarrow \overline{\psi}\psi \text{ is a scalar}$$

$$\overline{\psi}'(x')\gamma^{\mu}\psi'(x') = \overline{\psi}(x)S^{-1}\gamma^{\mu}S\psi(x)$$

$$= \overline{\psi}(x)\Lambda^{\mu}_{\nu}\gamma^{\nu}\psi(x)$$

$$= \Lambda^{\mu}_{\nu}(\overline{\psi}(x)\gamma^{\nu}\psi(x))$$

$$\Rightarrow \overline{\psi}\gamma^{\mu}\psi \text{ is a 4-vector}$$
(75)

P570

To obtain the plane-wave solution to the Dirac equation for a particle moving along the *x*-axis with velocity  $\beta$ , we boost along -x with a velocity of  $-\beta$ .

$$S = e^{-\binom{i}{2}\omega\sigma_{01}}; \quad \sigma_{01} = -i\alpha_{1}$$

$$= e^{-\frac{\omega}{2}\alpha_{1}} = \cosh\frac{\omega}{2} - \alpha_{1}\sinh\frac{\omega}{2}$$
where  $\tanh\omega = -\beta$ 
(76)

The spinors  $u^{\nu}(p)$  are

$$u^{\nu}(p) = \left[\cosh\left(\frac{\omega}{2}\right) - \alpha_{1}\sinh\left(\frac{\omega}{2}\right)\right]U^{\nu}(0)$$

$$= \cosh\left(\frac{\omega}{2}\right) \begin{bmatrix} 1 & 0 & 0 & -\tanh\frac{\omega}{2} \\ 0 & 1 & -\tanh\frac{\omega}{2} & 0 \\ 0 & -\tanh\frac{\omega}{2} & 1 & 0 \\ -\tanh\frac{\omega}{2} & 0 & 0 & 1 \end{bmatrix} u^{\nu}(0) \quad (77)$$

where

$$\tanh\left(\frac{\omega}{2}\right) = \frac{\tanh\omega}{1+\sqrt{1-\tanh^2\omega}} = \frac{-\beta}{1+\left(1-\beta^2\right)^{\frac{1}{2}}}$$
$$= \frac{-\beta}{1+\frac{1}{\gamma}} = \frac{-\beta\gamma}{1+\gamma} = \frac{-\beta\gamma m}{(1+\gamma)m} = \frac{-P}{E+m}$$
(78)

and

$$\cosh\left(\frac{\omega}{2}\right) = \frac{1}{\left(1 - \tanh^2\left(\frac{\omega}{2}\right)\right)^{\frac{1}{2}}} = \frac{1}{\left(\frac{2}{1+\gamma}\right)^{\frac{1}{2}}} = \sqrt{\frac{m+E}{2m}}$$

using Equation 78, Equation 77 becomes

P570

$$u^{\nu}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} 1 & 0 & 0 & \frac{Px}{E+m} \\ 0 & 1 & \frac{Px}{E+m} & 0 \\ 0 & \frac{Px}{E+m} & 1 & 0 \\ \frac{Px}{E+m} & 0 & 0 & 1 \end{bmatrix} u^{\nu}(0)$$
(79)

For a boost along an arbitrary direction

$$I^{\mu}_{\nu} = \begin{bmatrix} 0 & -Px & -Py & -Pz \\ |p| & |p| & |p| \\ -Px & 0 & 0 & 0 \\ |p| & & & \\ -Py & 0 & 0 & 0 \\ |p| & & & & \\ -Pz & 0 & 0 & 0 \\ |p| & & & & 0 \end{bmatrix}$$
(80)

and

$$S = e^{-\frac{\omega}{2}\tilde{\alpha} \cdot \hat{P}}$$
(81)

$$u^{\nu}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} 1 & 0 & \frac{Pz}{E+m} & \frac{Px-iPy}{E+m} \\ 0 & 1 & \frac{Px+iPy}{E+m} & \frac{-Pz}{E+m} \\ \frac{Pz}{E+m} & \frac{Px-iPy}{E+m} & 1 & 0 \\ \frac{Px+iPy}{E+m} & \frac{-Pz}{E+m} & 0 & 1 \end{bmatrix} u^{\nu}(0)$$
(82)

The four solutions of the spinors are

$$u^{1}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} 1\\0\\\frac{Pz}{E+m}\\\frac{Px+iPy}{E+m} \end{bmatrix}; \quad u^{2}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} 1\\0\\\frac{Px-iPy}{E+m}\\\frac{-Pz}{E+m} \end{bmatrix}$$

$$u^{3}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} \frac{Pz}{E+m}\\\frac{Px+iPy}{E+m}\\1\\0 \end{bmatrix}; \quad u^{4}(p) = \sqrt{\frac{m+E}{2m}} \begin{bmatrix} \frac{Px-iPy}{E+m}\\\frac{-Pz}{E+m}\\0\\1 \end{bmatrix}$$
(83)

Note that Equation 83 is obtained by using a normalization:

$$u^{1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad u^{2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \quad u^{3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; \quad u^{4} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$
(84)

We can also solve the Dirac equation directly:

$$Eu = \left(\vec{\alpha} \cdot \vec{P} + \beta m\right)u = \begin{pmatrix} m & \vec{\sigma} \cdot \vec{P} \\ \vec{\sigma} \cdot \vec{P} & -m \end{pmatrix}u$$
(85)

 $u = \begin{pmatrix} u_A \\ u_B \end{pmatrix}, \text{ where } u_A, u_B \text{ are two-component wave functions}$  $E \begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} m & \vec{\sigma} \cdot \vec{P} \\ \vec{\sigma} \cdot \vec{P} & -m \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix}$ 

We obtain the coupled equations

$$\vec{\sigma} \cdot \vec{P} u_B = (E - m) u_A$$
and
$$\vec{\sigma} \cdot \vec{P} u_A = (E + m) u_B$$
(86)

A possible solution to Equation 86 is

For E > 0, 
$$u^{(S)} = N \begin{pmatrix} x^{(S)} \\ \frac{\vec{\sigma} \cdot \vec{P}}{E+m} x^{(S)} \end{pmatrix}, \quad x^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
  
$$S = 1, 2$$
  
For E < 0,  $u^{(S+2)} = N \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{P}}{E-m} x^{(S)} \\ \frac{x^{(S)}}{x} \end{pmatrix}$ 

Explicitly:

$$u^{1}(p) = N \begin{pmatrix} 1 \\ 0 \\ \frac{Pz}{E+m} \\ \frac{Px+iPy}{E+m} \end{pmatrix}; \quad u^{2}(p) = N \begin{pmatrix} 0 \\ 1 \\ \frac{Px-iPy}{E+m} \\ \frac{-Pz}{E+m} \end{pmatrix}; \quad E > 0$$
(87)

$$u^{3}(p) = N \begin{pmatrix} \frac{Pz}{E-m} \\ \frac{Px+iPy}{E-m} \\ 1 \\ 0 \end{pmatrix}; \quad u^{4}(p) = N \begin{pmatrix} \frac{Px-iPy}{E-m} \\ \frac{-Pz}{E-m} \\ 0 \\ 1 \end{pmatrix}; \quad E < 0$$

Note that  $u^{3}(p)$ ,  $u^{4}(p)$  from Equation 87 are different from those obtained by the boost method (Equation 83). Apart from a difference in the normalization factor N,

$$u_{\text{boost}}^{(3,4)}(p) = u_{\text{direct}}^{(3,4)}(-p)$$
, where  $p = (E, \vec{p})$ 

Note that E > 0 for  $u_{\text{boost}}^{(3,4)}(p)$ , while E < 0 for  $u_{\text{direct}}^{(3,4)}(p)$ .

In order to interpret the E < 0 solutions for  $u^{(3,4)}(p)$ , we examine the chargeconjugation transformation on the Dirac equation.

The Dirac equation for an electron in an EM field can be obtained with the gauge substitution

$$\partial^{\mu} \to D^{\mu} \equiv \partial^{\mu} + iqA^{\mu} \tag{89}$$

For an electron with charge q = -e (e > 0), we have

$$i\partial_{\mu} \to i\partial_{\mu} + eA_{\mu} \tag{90}$$

The Dirac equation for an electron becomes

$$\left[\gamma^{\mu}\left(i\partial_{\mu}+eA_{\mu}\right)-m\right]\psi=0$$
(91)

The Dirac equation for a positron,  $\psi_c$ , would be

$$\left[\gamma^{\mu}\left(i\partial_{\mu}-eA_{\mu}\right)-m\right]\psi_{c}=0$$
(92)

Our task is to find the transformation linking  $\psi$  to  $\psi_c$ . To change the relative sign between the  $\partial_{\mu}$  and  $A_{\mu}$  terms in Equation 91, one can take a complex conjugate of Equation 91

$$\left[\left(\gamma^{\mu}\right)^{*}\left(-i\partial_{\mu}+eA_{\mu}\right)-m\right]\psi^{*}=0$$
(93)

To find a 4 x 4 matrix  $S_c$ , which satisfies

$$S_c \,\psi^* = \psi_c \tag{94}$$

We insert  $S_c^+ S_c$  term in front of  $\psi^*$  in Equation 93 and multiply  $S_c$  from the left:

$$\left[S_{c}\left(\gamma^{\mu}\right)^{*}S_{c}^{-1}\left(-i\partial_{n}+eA_{\mu}\right)-m\right]\left(S_{c}\psi^{*}\right)=0$$
(95)

From Equations 94 and 95, it follows that  $S_c$  satisfies the following equation:

$$S_c \left(\gamma^{\mu}\right)^* S_c^{-1} = -\gamma^{\mu} \tag{96}$$

Now, the Pauli-Dirac representation of  $\gamma^{\mu}$  gives

$$\left(\gamma^{\mu}\right)^{*} = \gamma^{\mu} \quad \text{for } \mu = 0, 1, 3$$

$$\left(\gamma^{\mu}\right)^{*} = -\gamma^{\mu} \quad \text{for } \mu = 2$$

$$(97)$$

Therefore, Equation 96 implies that  $S_c$  commutes with  $\gamma^2$  and anticommutes with  $\gamma^0$ ,  $\gamma^1$ ,  $\gamma^3$ . This can be satisfied if  $S_c$  is proportional to  $\gamma^2$ :

$$S_{c} = i\gamma^{2} = \begin{pmatrix} & & 1 \\ & -1 & \\ & -1 & \\ 1 & & \end{pmatrix}$$
(98)

The factor *i* is chosen by convention.

To obtain the positron wave function, one takes a complex conjugate of the electron wave function, followed by a multiplication of  $i\gamma^2$ :

$$\psi_c = i\gamma^2 \psi^* \tag{99}$$

We can apply Equation 99 to  $\psi^{(1)}(p)$ :

$$\psi^{(1)}(p) = N \begin{pmatrix} 1 \\ 0 \\ \frac{Pz}{E+m} \\ \frac{Px+iPy}{E+m} \end{pmatrix} e^{-iP \cdot x}, \quad P^{o} = E > 0$$
(100)

$$i\gamma^{2}\psi^{(1)*}(p) = \begin{pmatrix} & & 1 \\ & -1 & \\ 1 & & \end{pmatrix} N \begin{pmatrix} 1 & \\ 0 \\ \frac{Pz}{E+m} \\ \frac{Px-iPy}{E+m} \end{pmatrix} e^{iP\cdot x}$$
(101)

$$= N \begin{pmatrix} \frac{Px - iPy}{E + m} \\ \frac{-Pz}{E + m} \\ 0 \\ 1 \end{pmatrix} e^{iP \cdot x}; \quad P^o = E > 0$$
(101)

We can also write  $\psi^{(4)}(p')$  as

$$\psi^{(4)}(p') = N \begin{pmatrix} \frac{Px' - iPy'}{E - m} \\ \frac{-Pz'}{E - m} \\ 0 \\ 1 \end{pmatrix} e^{-iP' \cdot x} \quad \text{where} \quad (P')^0 = E' < 0 \quad (102)$$

Substituting p' = -p into Equation 102, we obtain

$$\psi^{(4)}(-p) = N \begin{pmatrix} \frac{-Px - iPy}{E - m} \\ \frac{Pz}{-E - m} \\ 0 \\ 1 \end{pmatrix} e^{iP \cdot x}; \quad P^{0} = E > 0$$
(103)

Equation 103 is identical to Equation 101.

Hence, 
$$\psi_c^{(1)} = i\gamma^2 \psi^{(1)*}(p) = \psi^{(4)}(-p)$$
  
Similarly,  $\psi_c^{(2)} = i\gamma^2 \psi^{(2)*}(p) = -\psi^{(3)}(-p)$  (104)

It also follows

$$u^{(4)}(-p) = v^{(1)}(p)$$
  
- $u^{(3)}(-p) = v^{(2)}(p)$  (105)

where  $v^{(1)}$ ,  $v^{(2)}$  are the spinors for positron wave functions.

P570

It is instructive to consider time-reversal operation on the Dirac equation. In a procedure analogous to that for the charge-conjugation operation, we can write down the Dirac equation for a time-reversed electron state,  $\psi_t$ , as

$$\left(-i\gamma^{0}\frac{\partial}{\partial t}+i\gamma^{1}\frac{\partial}{\partial x}+i\gamma^{2}\frac{\partial}{\partial y}+i\gamma^{3}\frac{\partial}{\partial z}-m\right)\psi_{t}=0$$
(106)

and the ordinary Dirac equation as

$$\left(i\gamma^{0}\frac{\partial}{\partial t}+i\gamma^{1}\frac{\partial}{\partial x}+i\gamma^{2}\frac{\partial}{\partial y}+i\gamma^{3}\frac{\partial}{\partial z}-m\right)\psi=0$$
(107)

Note that Equation 106 is obtained by  $t \rightarrow -t$  transformation applied to the Dirac equation.

Taking the complex conjugate of Equation 107 gives

$$\left(-i\left(\gamma^{0}\right)^{*}\frac{\partial}{\partial t}-i\left(\gamma^{1}\right)^{*}\frac{\partial}{\partial x}-i\left(\gamma^{2}\right)^{*}\frac{\partial}{\partial y}-i\left(\gamma^{3}\right)^{*}\frac{\partial}{\partial z}-m\right)\psi^{*}=0$$
(108)

To find a 4 x 4 matrix which transforms  $\psi$  to  $\psi_t$ 

$$S_t \, \psi^* = \psi_t \tag{109}$$

When we multiply  $S_t^{-1}S_t$  in front of  $\psi^*$  in Equation 108 followed by multiplying  $S_t$  from the left, we get

$$\left\{ \left[ S_t \left( -i\left(\gamma^0\right)^* \frac{\partial}{\partial t} - i\left(\gamma^1\right)^* \frac{\partial}{\partial x} - i\left(\gamma^2\right)^* \frac{\partial}{\partial y} - i\left(\gamma^3\right)^* \frac{\partial}{\partial z} \right] S_t^{-1} \right] - m \right\} S_t \psi^* = 0$$
(110)

Equations 109 and 110 imply that  $S_t$  should have the following properties:

$$S_t \left(\gamma^0\right)^* S_t^{-1} = \gamma^0$$

$$S_t \left(\gamma^i\right)^* S_t^{-1} = -\gamma^i \qquad (i = 1 \to 3)$$
(111)

Since  $(\gamma^0)^* = \gamma^0$ ,  $(\gamma^1)^* = \gamma^1$ ,  $(\gamma^2)^* = -\gamma^2$ ,  $(\gamma^3)^* = \gamma^3$ , Equation 111 implies

 $S_t$  commutes with  $\gamma^0$ ,  $\gamma^2$  and anticommutes with  $\gamma^1$ ,  $\gamma^3$ .  $S_t$  is therefore proportional to  $\gamma^1$ ,  $\gamma^3$  and, by convention, one chooses

$$S_{t} = i\gamma^{1}\gamma^{3}$$

$$= \begin{pmatrix} -\sigma_{2} & 0 \\ 0 & \sigma_{2} \end{pmatrix} = -i \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
(112)

Applying  $S_t \psi^*(p)$  for  $\psi^{(1)}$ , one obtains

$$S_{t}\psi^{(1)*}(p) = -i \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} N \begin{pmatrix} 1 \\ 0 \\ \frac{Pz}{E+m} \\ \frac{Px-iPy}{E+m} \end{pmatrix} e^{iP \cdot x}$$

$$= -iN \begin{pmatrix} 0 \\ 1 \\ \frac{-Px-iPy}{E+m} \\ \frac{Pz}{E+m} \\ \frac{Pz}{E+m} \end{pmatrix} e^{iP \cdot x}$$
(113)

Hence, the time-reversal operation transforms  $\psi^{(1)}(\vec{p},t)$  to  $\psi^{(2)}(-\vec{p},-t)$ , as one would expect (since time-reversal flips  $\vec{p}$ , *t* and spin, and  $\psi^{(1)}$  and  $\psi^{(2)}$  have opposite spin).

The combined operation of CPT can be obtained too.

Since

$$P\psi = \gamma^{0}\psi$$
$$C\psi = i\gamma^{2}\psi^{*}$$
$$T\psi = i\gamma^{1}\gamma^{3}\psi^{*}$$

$$\psi_{CPT} = CPT\psi = CP(i\gamma^{1}\gamma^{3}\psi^{*})$$
$$= C(\gamma^{0}i\gamma^{1}\gamma^{3}\psi^{*})$$
$$= -i\gamma^{2}\gamma^{0}i\gamma^{1}\gamma^{3}\psi$$
$$= \gamma^{0}\gamma^{1}\gamma^{2}\gamma^{3}\psi$$
$$= -i\gamma^{5}\psi$$

we have

We can also examine the operation of CPT on solutions for the Klein-Gordon equation

$$\left(\partial^{\mu}\partial_{\mu}+m^{2}\right)\phi(x)=0$$

In the presence of an EM field:

$$i\partial^{\mu} \rightarrow i\partial^{\mu} - eA^{\mu}$$

and the K-G equation becomes

$$(i\partial^{\mu}-eA^{\mu})(i\partial_{\mu}-eA_{\mu})\phi(x)=m^{2}\phi(x)$$

or

$$\left[\Box^{2}+m^{2}+ie\left(\partial^{\mu}A_{\mu}+A^{\mu}\partial_{\mu}\right)-e^{2}A^{2}\right]\phi(x)=0$$

It is straight forward to show that

P: 
$$\phi'_P(x') = \phi(x)$$
  
C:  $\phi'_C(x') = \phi^*(x)$   
T:  $\phi'_T(x') = \phi^*(x)$   
CPT:  $\phi'_{CPT}(x') = \phi(x)$ 

One can also work out the non-relativistic limit of the Dirac equation. From Equation 86, the following coupled equations can be obtained:

$$\begin{pmatrix} \vec{\sigma} \cdot \vec{P} \end{pmatrix} \psi_B = (E - m) \psi_A$$

$$\begin{pmatrix} \vec{\sigma} \cdot \vec{P} \end{pmatrix} \psi_A = (E + m) \psi_B$$

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}$$
(114)

where

Making the substitution  $P \rightarrow P + eA = \pi$ , Equation 114 becomes

$$(\vec{\sigma} \cdot \vec{\pi}) \psi_B = (E + eA^o - m) \psi_A$$

$$(\vec{\sigma} \cdot \vec{\pi}) \psi_A = (E + eA^o + m) \psi_B$$
(115)

In the non-relativistic limit  $E \simeq m$  and  $eA^{\circ} \ll m$ . Therefore

$$\psi_B \simeq \frac{\vec{\sigma} \cdot \vec{\pi}}{2m} \psi_A \tag{116}$$

$$\frac{\left(\vec{\sigma} \cdot \vec{\pi}\right)^2}{2m} \psi_A = \left(\mathbf{E}_{NR} + eA^o\right) \psi_A \tag{117}$$

where

$$E_{NR} > E - m$$

$$\left(\vec{\sigma} \cdot \vec{\pi}\right)^2 = \vec{\pi}^2 + i\vec{\sigma} \cdot \left(\vec{\pi} \times \vec{\pi}\right) = \left(\vec{P} + e\vec{A}\right)^2 + i\vec{\sigma} \cdot \left(\vec{P} + e\vec{A}\right) \times \left(\vec{P} + e\vec{A}\right)$$

$$= \left(\vec{P} + e\vec{A}\right)^2 + e\left(\vec{\nabla} \times \vec{A}\right) \cdot \vec{\sigma}$$

$$(118)$$

Note that

$$\vec{P} \times \left(\vec{A}\psi\right) + \vec{A} \times \left(\vec{P}\psi\right) = -i\left(\vec{\nabla} \times \vec{A}\right)\psi$$

Equations 117 and 118 imply

$$\left[\frac{1}{2m}\left(\vec{P}+e\vec{A}\right)^2 + \frac{e}{2m}\vec{\sigma}\cdot\vec{B} - eA^o\right]\psi_A = \mathbf{E}_{NR}\psi_A \tag{119}$$

which is the Pauli equation.

For a uniform  $\vec{B}$  field,  $\vec{A} = \frac{1}{2}\vec{B} \times \vec{\gamma}$ 

$$\left(\vec{P} + e\vec{A}\right)^2 \cong P^2 + 2e\vec{P} \cdot \frac{1}{2} \left(\vec{B} \times \vec{\gamma}\right) = P^2 + e\vec{B} \cdot \left(\vec{\gamma} \times \vec{P}\right) = P^2 + e\vec{B} \cdot \vec{L}$$
(120)

Using Equation 120, Equation 119 becomes

$$\left[\frac{P^2}{2m} + \frac{e}{2m}\left(\vec{L} + 2\vec{S}\right) \cdot \vec{B} - eA^o\right] \psi_A = \mathcal{E}_{NR} \psi_A$$
(121)

The Dirac equation leads to  $g_s = 2$  for spin- $\frac{1}{2}$  particle.

However,  $g_s = 2$  is not a unique feature of Dirac equation. To illustrate this point, one can start with a non-relativistic expression

$$H = \frac{P^{2}}{2m} = \frac{\left(\vec{\sigma} \cdot \vec{P}\right)^{2}}{2m}$$
(122)

Making the  $\vec{P} \rightarrow \vec{P} + e\vec{A}$  substitution, Equation 122 becomes

$$H = \left[\vec{\sigma} \cdot \left(\vec{P} + e\vec{A}\right)\right]^2 / 2m$$
  
$$= \frac{1}{2m} \left[\left(\vec{P} + e\vec{A}\right)^2 + e\vec{\sigma} \cdot \vec{B}\right]$$
  
$$= \frac{1}{2m} \left[\left(\vec{P} + e\vec{A}\right)^2\right] + \frac{e}{2m} \left(2\vec{S} \cdot \vec{B}\right)$$
(123)

Hence,  $g_s = 2$  also follows from a non-relativistic approach.

The spin operator is defined as

$$\sum_{k} = \sigma_{ij} = \begin{bmatrix} \sigma_{k} & 0 \\ 0 & \sigma_{k} \end{bmatrix} \quad (i, j, k = 1, 2, 3 \text{ cyclic})$$
(124)

It can be shown that the Dirac Hamiltonian  $H = \vec{\alpha} \cdot \vec{P} + \beta m$  does not commute with  $\Sigma$ , namely

$$\left[H,\vec{\Sigma}\right] = 2i\left(\vec{\alpha}\times\vec{P}\right) \neq 0 \tag{125}$$

It does not commute with the orbital angular momentum  $\vec{L}$  either

$$\left[H,\vec{L}\right] = -i\left(\vec{\alpha}\times\vec{P}\right) \neq 0 \tag{126}$$

However, Equations 125 and 126 show that *H* commutes with the total angular momentum  $\vec{J} = \vec{L} + \frac{1}{2}\vec{\Sigma}$ 

$$\begin{bmatrix} H, \vec{J} \end{bmatrix} = \begin{bmatrix} H, \vec{L} \end{bmatrix} + \begin{bmatrix} H, \frac{1}{2}\vec{\Sigma} \end{bmatrix} = 0$$
(127)

Although spin itself is not conserved, the spin projection along the momentum vector is conserved. This spin projection is also called helicity and is defined as

$$\frac{1}{2}\vec{\Sigma}\cdot\vec{P} = \frac{1}{2}\begin{bmatrix}\vec{\sigma}\cdot\hat{P} & 0\\0 & \vec{\sigma}\cdot\hat{P}\end{bmatrix}$$
(128)

One can readily show that  $\left[H, \frac{1}{2}\vec{\Sigma}\cdot\hat{P}\right] = 0$ . The reason why  $\frac{1}{2}\vec{\Sigma}\cdot\hat{P}$  is conserved

can be understood if one recognizes that there is no orbital angular momentum along the direction of momentum. Hence the particle's spin is the same as its total angular momentum in this case, and it is conserved.

Note that  $\Sigma^k = \gamma^5 \gamma^0 \gamma^k$  (k = 1, 2, 3)

The Dirac spinors given in Equation 87 are not eigenstates of the helicity operator:

$$\left(\frac{1}{2}\vec{\Sigma}\cdot\hat{P}\right)u^{(1)}(P) = \frac{1}{2|P|} \begin{pmatrix} Pz & Px-iPy & 0 & 0\\ Px+iPy & -Pz & 0 & 0\\ 0 & 0 & Pz & Px-iPy\\ 0 & 0 & Px+iPy & -Pz \end{pmatrix} N \begin{pmatrix} 1\\ 0\\ Pz/\\ E+m\\ Px+iPy/\\ E+m \end{pmatrix}$$

$$= \frac{N}{2|P|} \begin{pmatrix} Pz\\ Px+iPy\\ P^2/\\ E+m\\ 0 \end{pmatrix}$$
(129)

Equation 129 shows that in general,  $U^{(1)}(P)$  is not an eigenstate of  $\frac{1}{2}\vec{\Sigma}\cdot\hat{P}$ . In the special case when Px = Py = 0,  $U^{(1)}(P)$  is an eigenstate of  $\frac{1}{2}\vec{\Sigma}\cdot\hat{P}$  with an eigenvalue of  $\pm \frac{1}{2}$ .

It is possible to construct basis states which are eigenstates of the helicity operator. It can easily be shown that the following  $u^{(\pm)}$  are hilicity eigenstates with eigenvalues of  $+\frac{1}{2}$ ,  $-\frac{1}{2}$ , respectively.

$$u^{(\pm)} = N \begin{pmatrix} Pz \pm |P| \\ Px + iPy \\ (Pz \pm |P|) \frac{|P|}{E+m} \\ (Px + iPy) \frac{|P|}{E+m} \end{pmatrix}$$
(130)

We summarize here some of the pertinent properties of the Dirac spinors *u* and *v*: First, *u* and *v* satisfy the following equations:

$$(\not P - m)u = 0, \quad (\not P + m)v = 0$$

$$\overline{u}(\not P - m) = 0, \quad \overline{v}(\not P + m) = 0$$
(131)

The orthogonality relations are

$$u^{(r)^{+}}u^{(s)} = 2E \ \delta_{rs}, \quad v^{(r)^{+}}v^{(s)} = 2E \ \delta_{rs}$$
  
$$\overline{u}^{(s)}u^{(s)} = 2m, \quad \overline{v}^{(s)}v^{(s)} = -2m$$
(132)

Note that different normalizations have been adopted in the literature. For example

$$\overline{u}^{(s)}u^{(s)} = 1, \quad \overline{v}^{(s)}v^{(s)} = 1$$
 (133)

have been chosen by many textbooks

The completeness relations are

$$\sum_{s=1,2} u^{(s)}(P)\overline{u}^{(s)}(P) = \not P + m$$

$$\sum_{s=1,2} v^{(s)}(P)\overline{v}^{(s)}(P) = \not P - m$$
(134)

Why are these relations called completeness relations?

One can show easily that

$$\sum_{s=1,2} u^{(s)}(P) \overline{u}^{(s)}(P) - v^{(s)}(P) \overline{v}^{(s)}(P) = 2m$$
(135)

Using the normalization scheme of Equation 133, one obtains

$$\sum_{s=1,2} u^{(s)}(P)\overline{u}^{(s)}(P) - v^{(s)}(P)\overline{v}^{(s)}(P) = 1$$

Equations 134 are very important for evaluating cross sections later on.

Explicitly

P570

$$u^{(1)} = \sqrt{E} + m \begin{pmatrix} 1 \\ 0 \\ Pz/E + m \\ Px + iPy/E + m \end{pmatrix}; \quad u^{(2)} = \sqrt{E} + m \begin{pmatrix} 0 \\ 1 \\ Px - iPy/E + m \\ -Pz/E + m \end{pmatrix}$$
$$v^{(1)} = \sqrt{E} + m \begin{pmatrix} Px - iPy/E + m \\ -Pz/E + m \\ -Pz/E + m \\ 0 \\ 1 \end{pmatrix}; \quad v^{(2)} = \sqrt{E} + m \begin{pmatrix} Pz/E + m \\ Px + iPy/E + m \\ -Pz/E + m \\ 1 \\ 0 \end{pmatrix}$$

In addition to  $\overline{\psi}\psi$  and  $\overline{\psi}\gamma^{\mu}\psi$  discussed earlier, there are other bilinear covariants. One can show that the following 16 bilinear covariants are independent and complete:

$$\overline{\psi}\psi \qquad (1) \quad \text{scalar} \\
\overline{\psi}\gamma^{\mu}\psi \qquad (4) \quad \text{vector} \\
\overline{\psi}\sigma^{\mu\nu}\psi \qquad (6) \quad \text{tensor} \qquad (136) \\
\overline{\psi}\gamma^{5}\psi \qquad (1) \quad \text{pseudoscalar} \\
\overline{\psi}\gamma^{5}\gamma^{\mu}\psi \qquad (4) \quad \text{axial vector}$$

where 
$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3, \quad \gamma^5 = \gamma^{5^+}$$
 (137)  
 $\gamma^5$  anticommutes with all  $\gamma^{\mu}$ 

$$\left\{\gamma^{5},\gamma^{\mu}\right\}=0\tag{138}$$

$$\gamma^5$$
 commutes with proper Lorentz transformation *S*

$$\gamma^5 S = S \gamma^5 \tag{139}$$

but anticommutes with space inversion  $S_P$ 

$$\gamma^5 S_P = -S_P \gamma^5 \tag{140}$$

The transformation of the 16 bilinear covariants under proper Lorentz transformations and space inversion, as given by Equation 136, can be readily proved using Equations 53, 74, 139 and 140.

The  $\gamma^5$  matrix is often encountered in the context of chirality. Although  $[H, \gamma^5] \neq 0$ , it can be shown that at high energies E >> m

$$\gamma^5 u = \begin{bmatrix} \vec{\sigma} \cdot \vec{P} & 0\\ 0 & \vec{\sigma} \cdot \vec{P} \end{bmatrix} u$$

and  $\frac{1-\gamma^5}{2}$ ,  $\frac{1+\gamma^5}{2}$  projects the left and right-handed components of the Dirac spinor *u*.

We have based our discussion on the Dirac equation upon the Pauli-Dirac representation of the  $\gamma$  matrices. How would the wave function be modified if we use different representations for the  $\gamma$  matrices?

Suppose we use another representation  $\gamma^{\mu\prime}$  for the Dirac equation with wave function  $\psi'$ :

$$(i\gamma^{\mu'}\partial_{\mu} - m)\psi' = 0$$
 where  $[\gamma^{\mu'}, \gamma^{\nu'}] = 2g^{\mu\nu}$  (141)

How does  $\psi'$  differ from  $\psi$  in the 'regular' Dirac equation?

$$\left(i\gamma^{\mu}\partial_{\mu}-m\right)\psi=0\tag{142}$$

The 'Pauli's fundamental theorem' shows that for any  $\gamma^{\mu'}$  satisfying the relations  $\left[\gamma^{\mu'}, \gamma^{\nu'}\right] = 2g^{\mu\nu}$ , a transformation *S* can be found such that

$$S\gamma^{\mu}S^{-1} = \gamma^{\mu} \tag{143}$$

Equation 141 becomes

$$\left(i S \gamma^{\mu} S^{-1} \partial_{\mu} - m\right) S S^{-1} \psi' = 0 \tag{144}$$

multiplying  $S^{-1}$  from the left gives

$$\left(i \gamma^{\mu} \partial_{\mu} - m\right) S^{-1} \psi' = 0 \tag{145}$$

Equation 145 is equivalent to Equation 142 provided

$$\psi = S^{-1}\psi'$$
 or  $\psi' = S\psi$  (146)

Therefore, the wave functions obtained using different representations of the  $\gamma$ -matrices are related by the same transformation *S* used to relate  $\gamma^{\mu}$  and  $\gamma^{\mu\prime}$ .

We complete our discussion on the Dirac equation by presenting an alternative way to derive the Dirac equation.

We can start with the relativistic energy-momentum relation

$$E^2 - P^2 = m^2$$
 (147)

or equivalently

$$\mathbf{E}^2 - \left(\vec{\boldsymbol{\sigma}} \cdot \vec{\boldsymbol{P}}\right)^2 = m^2 \tag{148}$$

Making the substitution  $E \rightarrow i \frac{\partial}{\partial t}$ ,  $\vec{P} \rightarrow -i \vec{\nabla}$  and acting on wave function  $\phi$ , we have

$$\left(i\frac{\partial}{\partial t} + i\vec{\sigma}\cdot\vec{\nabla}\right)\left(i\frac{\partial}{\partial t} - i\vec{\sigma}\cdot\vec{\nabla}\right)\phi = m^2\phi \tag{149}$$

Now  $\phi$  is a two-component wave function. Define

$$\phi^{(R)} = \frac{1}{m} \left( i \frac{\partial}{\partial t} - i \vec{\sigma} \cdot \vec{\nabla} \right) \phi$$

$$\phi^{(L)} = \phi$$
(150)

then Equation 149 gives

$$\phi^{(R)} = \frac{1}{m} \left( i \frac{\partial}{\partial t} - i \vec{\sigma} \cdot \vec{\nabla} \right) \phi^{(L)}$$

$$\phi^{(L)} = \frac{1}{m} \left( i \frac{\partial}{\partial t} + i \vec{\sigma} \cdot \vec{\nabla} \right) \phi^{(R)}$$
(151)

We can now construct a 4-component wave function  $\boldsymbol{\psi}$ 

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = \begin{pmatrix} \phi^{(R)} + \phi^{(L)} \\ \phi^{(R)} - \phi^{(L)} \end{pmatrix}$$
(152)

From Equation 151, one can form the following equations:

$$\phi^{(R)} + \phi^{(L)} = \frac{1}{m} \left[ i \frac{\partial}{\partial t} \left( \phi^{(R)} + \phi^{(L)} \right) + i \vec{\sigma} \cdot \vec{\nabla} \left( \phi^{(R)} - \phi^{(L)} \right) \right]$$
(153)

$$\phi^{(R)} - \phi^{(L)} = \frac{1}{m} \left[ -i \frac{\partial}{\partial t} \left( \phi^{(R)} - \phi^{(L)} \right) + i \vec{\sigma} \cdot \vec{\nabla} \left( \phi^{(R)} + \phi^{(L)} \right) \right]$$
(154)

From Equation 152, Equations 153, 154 can be expressed

$$\begin{pmatrix} i\frac{\partial}{\partial t} & i\vec{\sigma}\cdot\vec{\nabla} \\ -i\vec{\sigma}\cdot\vec{\nabla} & -i\frac{\partial}{\partial t} \end{pmatrix} \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = m \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}$$
(155)

which is simply

$$\left(i\gamma^{\mu}\partial_{\mu}-m\right)\psi=0\tag{156}$$

the Dirac equation.