Chapter 3

The Dirac Theory

In this chapter, we will study the Dirac equation which describes spin-1/2 particles such as quarks and leptons. We start from the observation that the apparent reason why the Klein-Gordon equation gave solutions with negative energies and negative probabilities was that the energy appeared as E^2 in the relation $E^2 = \vec{P}^2 + m^2$, allowing negative as well as positive value of E. Thus, we will proceed by trying to construct a wave equation which is linear in the time derivative, hoping that it will give us solutions with only positive energies and positive probabilities.

It is not clear exactly what kind of thought process was followed by Dirac when he set out to construct a relativistic theory of electron in late 1920's, even though it seems that he was particularly disturbed by the negative probability. As we will see, he actually succeeded in solving the problem of negative probability, but not that of negative energy. In fact, even if one constructs a wave equation that is linear in time derivative, there is no guarantee that the energy will always be positive. When one applies a differential equation linear in $i\partial_0$ to a plane-wave form, one obtains a relation between E and $\vec{P}$ in the form $E = f(\vec{P}, m)$ as opposed to $E^2 = f(\vec{P}, m)$. The function $f(\vec{P}, m)$, however, may in general be positive or negative, which is actually what happens in the case of the Dirac equation. Dirac partially 'solved' the negative-energy problem by the so-called hole theory in which the vacuum is assumed to be the state where all the negative energy-states are filled up thereby preventing positive energy states to fall into the negative-energy states by the exclusion principle. However, such a scenario cannot work for particles with integer spin for which the exclusion principle does not apply. As in the case of the Klein-Gordon theory, the problem of negative energy in the Dirac theory will be satisfactorily solved in the framework of the quantum field theory.

In retrospect, the importance of the Dirac equation does not have much to do with negative energy or negative probability. Its importance lies in the fact that, in trying to solve these 'problems', Dirac stumbled upon a quantity called a spinor which describes a particle with spin 1/2, and led to a correct description of the magnetic moment of the electron. Here, we will use the linearity in the time derivative simply as a convenient guide to introduce the Dirac equation. In this section and later, we will generically refer to the particle represented by the Dirac equation as an 'electron', but the discussions apply to any point-like spin-1/2 particle which is not antiparticle of itself. [1]

3.1 The Dirac equation

We now search for a wave equation that is linear in time derivative and consistent with the relativistic energy-momentum relation $E^2 = \vec{P}^2 + m^2$. Let's start from the following Schrödinger form of equation:

$$i\frac{\partial}{\partial t}\psi = H\psi\,, \tag{3.1}$$

where ψ is some wave function representing the state of the electron. The operator H is assumed to be some function of the momentum operator $-i\vec{\nabla}$ and the mass m, presumably representing the energy of the electron. Since we are now interested in a free electron, we assume that H does not depend on time. We then want this to give the relativistic relation $E^2 = \vec{P}^2 + m^2$ when acting upon a plane wave solution $\psi \propto e^{-ip\cdot x}$. This can be accomplished if (3.1) somehow leads to

$$\left(i\frac{\partial}{\partial t}\right)^2\psi = [(-i\vec{\nabla})^2 + m^2]\psi\,. \tag{3.2}$$

We first take the time derivative $i\partial/\partial t$ of (3.1) to get

$$\begin{aligned}\left(i\frac{\partial}{\partial t}\right)^2\psi &= H\,\underbrace{i\frac{\partial}{\partial t}\psi}_{H\psi} = H^2\psi \\ &\text{(which we would like to become)} \\ &= [(-i\vec{\nabla})^2 + m^2]\psi\,,\end{aligned} \tag{3.3}$$

or

$$H^2 = (-i\vec{\nabla})^2 + m^2\,. \tag{3.4}$$

Note that the equation (3.2) is nothing but the Klein-Gordon equation. Yes, if a wave function satisfies the Dirac equation, then it will satisfy the Klein-Gordon equation. In fact, any relativistic wave function would satisfy the Klein-Gordon equation as long as it is consistent with $E^2 = \vec{P}^2 + m^2$.

[1] A spin-1/2 particle that is antiparticle of itself is called a Majorana particle. So far, no such particles have been found in nature.

Since H^2 should become a quadratic function of $-i\vec{\nabla}$ and m, it seems reasonable (though not mandatory) to look for H which is linear in $-i\vec{\nabla}$ and m:

$$H = \vec{\alpha} \cdot (-i\vec{\nabla}) + \beta m \,, \tag{3.5}$$

where $\vec{\alpha} \equiv (\alpha_1, \alpha_2, \alpha_3)$ and β are some *constants* which turn out to be matrices. Keeping track of the ordering of the (matrix) products of α_i and β while noting that m is just a real number, the condition (3.4) then becomes

$$\begin{aligned}
H^2 &= [\alpha_i(-i\nabla_i) + \beta m]\,[\alpha_j(-i\nabla_j) + \beta m] \\
&= \sum_{i,j} \alpha_i\alpha_j(-i\nabla_i)(-i\nabla_j) \quad \to \left(\begin{array}{l} \sum_i \alpha_i^2(-i\nabla_i)^2 \\ +\sum_{i>j}(\alpha_i\alpha_j + \alpha_j\alpha_i)(-i\nabla_i)(-i\nabla_j) \end{array} \right. \\
&\quad \left. \begin{array}{l} +\sum_i \alpha_i\beta(-i\nabla_i)m \\ +\sum_j \beta\alpha_j(-i\nabla_j)m \end{array} \right\} \quad \to \sum_i (\alpha_i\beta + \beta\alpha_i)(-i\nabla_i)m \\
&\quad +\beta^2 m^2 \\
&\text{(which should be equal to)} \\
&= (-i\vec{\nabla})^2 + m^2 \,,
\end{aligned} \tag{3.6}$$

which is satisfied if

$$\alpha_i^2 = \beta^2 = 1, \qquad \begin{pmatrix} \alpha_i\alpha_j + \alpha_j\alpha_i = 0 \quad (i \neq j) \\ \alpha_i\beta + \beta\alpha_i = 0 \end{pmatrix}, \quad (i,j = 1,2,3)\,. \tag{3.7}$$

Using the anticommutator symbol,

$$\{A, B\} \stackrel{\text{def}}{\equiv} AB + BA \,, \tag{3.8}$$

this condition can be written as

$$\boxed{\begin{array}{ll} \{\alpha_i\,, \alpha_j\} = 0 \quad (i \neq j), & \\ \{\alpha_i\,, \beta\} = 0, & \end{array} \quad \alpha_i^2 = \beta^2 = 1, \quad (i,j = 1,2,3)}\,. \tag{3.9}$$

Note that this condition is symmetric among the four quantities $\alpha_1, \alpha_2, \alpha_3$ and β; namely, the square of each is unity and each anticommutes with another. They cannot be ordinary numbers since they do not commute. They can be matrices, however, and we will now show that they have to be at least 4×4:

1. *α_i and β are hermitian.* Since H represents the energy, we want it to have real eigenvalues; thus, H is hermitian. Since the operator $-i\vec{\nabla}$ acts as a hermitian operator, then α_i and β should also be hermitian:

$$\alpha_i^\dagger = \alpha_i, \qquad \beta^\dagger = \beta \,. \tag{3.10}$$

2. *The traces of α_i and β are zero.* From (3.9), we have

$$\alpha_i\beta = -\beta\alpha_i \quad \xrightarrow{\times\beta \text{ from right}} \quad \alpha_i \underbrace{\beta^2}_{1} = -\beta\alpha_i\beta\,. \tag{3.11}$$

   Taking trace of both sides,

$$\mathrm{Tr}\alpha_i = -\underbrace{\mathrm{Tr}(\beta\alpha_i\beta)}_{\mathrm{Tr}(\beta^2\alpha_i)} = -\mathrm{Tr}\alpha_i\,, \quad \rightarrow \quad \mathrm{Tr}\alpha_i = 0\,, \tag{3.12}$$

   where we have used

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)\,, \tag{3.13}$$

   where A and B are two arbitrary square matrices of same rank. Similarly, multiplying α_i instead of β in (3.11), we get $\mathrm{Tr}\beta = 0$.

3. *The eigenvalues of α_i and β are ± 1.* Suppose u is an eigenvector of β with an eigenvalue c; namely, $\beta u = cu$. Applying β from the left again, we get

$$\beta^2 u = c\,\beta u \quad \rightarrow \quad u = c^2 u \quad \rightarrow \quad c^2 = 1. \tag{3.14}$$

   Thus, the eigenvalues of β should be ± 1. Similarly, the eigenvalues of α_i should also be ± 1.

4. *The rank of α_i and β is even.* Since β is hermitian, it can be diagonalized by some matrix S:

$$S\beta S^{-1} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}, \tag{3.15}$$

   where c_i are the eigenvalues of β. Taking the trace, and using (3.13),

$$\mathrm{Tr}(S\beta S^{-1}) = \mathrm{Tr}(S^{-1}S\beta) = \mathrm{Tr}\beta = \sum_i c_i\,. \tag{3.16}$$

   Since $\mathrm{Tr}\beta = 0$ and $c_i = \pm 1$, n should be even. The proof is the same for α_i.

5. *α_i and β are linearly independent.* Suppose β can be written as a linear combination of α_i: $\beta = b_i\alpha_i$ where b_i are some c-numbers. Then, using $\{\beta\,,\alpha_j\} = 0$ (3.9),

$$0 = \{\beta\,,\alpha_j\} = \{b_i\alpha_i\,,\alpha_j\} = b_i\{\alpha_i\,,\alpha_j\} = 2b_j\,. \tag{3.17}$$

   where we have used the linearity of the anticommutator

$$\{cA, B\} = \{A, cB\} = c\{A, B\} \quad (c : c\text{-number};\ A, B : \text{matrices}). \tag{3.18}$$

   We thus have $b_j = 0$ ($j = 1, 2, 3$); namely, β cannot be written as a linear combination of α_i's. Similarly, α_i cannot be written as a linear combination of the rest.

An $n \times n$ complex matrix A has $2n^2$ degrees of freedom, and out of which n^2 are taken away by the hermitian condition $A^\dagger = A$.[2] Together with the traceless requirement which takes away one degree of freedom, we are left with $2n^2 - n^2 - 1 = n^2 - 1$ degrees of freedom. Thus, n should be larger than 2 to have at least four independent such matrices ($\vec{\alpha}$ and β). Since n should be even, α_i and β are then independent, traceless, hermitian anticommuting matrices of rank 4 or more.

Now we will explicitly construct the 4×4 matrices α_i and β. A well-known set of linearly independent anticommuting matrices is the Pauli matrices σ_i $(i = 1, 2, 3)$ given by

$$\boxed{\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},} \tag{3.19}$$

which are also hermitian and traceless

$$\sigma_i^\dagger = \sigma_i, \qquad \mathrm{Tr}\sigma_i = 0 \quad (i = 1, 2, 3) \tag{3.20}$$

and satisfy

$$\sigma_i^2 = 1 \quad (i = 1, 2, 3), \qquad \sigma_i \sigma_j = -\sigma_j \sigma_i = i\sigma_k \quad (i, j, k : \text{cyclic}), \tag{3.21}$$

or equivalently,

$$\boxed{\begin{array}{rcl} \{\sigma_i, \sigma_j\} & = & 2\,\delta_{ij} \\ [\sigma_i, \sigma_j] & = & 2\,i\,\epsilon_{ijk}\sigma_k \end{array} \qquad (i, j, k = 1, 2, 3)}\,. \tag{3.22}$$

In (3.21), '1' is actually a 2×2 identity matrix, and in (3.22) the identity matrix on the right hand side is omitted. Hereafter, the $n \times n$ identity matrix is often written as the number '1' for simplicity.

Exercise 3.1 *Explicitly verify the relations (3.22).*

Exercise 3.2 *Use the properties of the Pauli matrixes (3.22) to prove*

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} + i\vec{\sigma} \cdot (\vec{a} \times \vec{b}) \tag{3.23}$$

and

$$e^{i\vec{a}\cdot\vec{\sigma}} = \cos a + i\hat{a} \cdot \vec{\sigma} \sin a \tag{3.24}$$

where $\vec{a}$ and $\vec{b}$ are 3-component vectors, and

$$a = |a|, \quad \hat{a} = \vec{a}/a, \quad \vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)\,. \tag{3.25}$$

[2] The hermiticity condition of an $n \times n$ matrix A is given by $A^*_{ij} = A_{ji}$. It contains n equations for the diagonal elements $A^*_{ii} = A_{ii}$ each of which removes one degree of freedom, and $n(n-1)/2$ equations of the form $A^*_{ij} = A_{ji}$ $(i \neq j)$ for off diagonal elements each of which removes two degrees of freedom. Thus, the total degrees of freedom removed is $n + 2 \times n(n-1)/2 = n^2$.

One important feature of the Pauli matrices is that $\sigma_i/2$ act as the spin-1/2 angular momentum operators. Indeed, from (3.22), we see that $\sigma_i/2$ satisfy the commutation relation of angular momentum $[J_i\,,J_j] = i\,\epsilon_{ijk}J_k$:

$$\left[\frac{\sigma_i}{2},\frac{\sigma_j}{2}\right] = i\,\epsilon_{ijk}\frac{\sigma_k}{2}\,. \tag{3.26}$$

The Pauli matrices are independent, hermitian, traceless, anticommuting, and thier squares are unity. The problem of course is that there are only three of them while we need four. It is, however, a good place to start. So we try the following:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{3.27}$$

where I is the 2×2 identity matrix. These are clearly hermitian and traceless. It is also straightforward to see that they satisfy the relations (3.9). For example,

$$\begin{aligned}
\{\alpha_i\,,\alpha_j\} &= \alpha_i\alpha_j + \alpha_j\alpha_i \\
&= \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \\
&= \begin{pmatrix} \sigma_i\sigma_j & 0 \\ 0 & \sigma_i\sigma_j \end{pmatrix} + \begin{pmatrix} \sigma_j\sigma_i & 0 \\ 0 & \sigma_j\sigma_i \end{pmatrix} \\
&= \begin{pmatrix} \{\sigma_i,\sigma_j\} & 0 \\ 0 & \{\sigma_i,\sigma_j\} \end{pmatrix} = 2\delta_{ij}\,, \quad \text{etc.}
\end{aligned} \tag{3.28}$$

Thus, we now have an equation which is linear in time derivative and consistent with the relativistic relation $E^2 = \vec{P}^2 + m^2$. Namely, combining (3.1) and (3.5),

$$i\frac{\partial}{\partial t}\psi(x) = [\vec{\alpha}\cdot(-i\vec{\nabla}) + \beta m]\psi(x)\,, \tag{3.29}$$

where α_i and β are 4×4 matrices defined by the relations (3.9) and explicitly given by (3.27). Since α_i and β are 4×4 matrices, the wave function ψ has to be a 4-component quantity (called a *Dirac spinor*, a *4-component spinor*, or simply a *spinor*):

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}, \qquad \psi_n \; (n = 1,2,3,4) : \text{complex} \tag{3.30}$$

The indexes of the four components have *nothing to do* with the space-time. The multiple components were introduced to satisfy the required relations among α_i and β. We will see later that they represent the degrees of freedom corresponding to spin up and down as well as those corresponding to electron and its antiparticle positron. We will call this four-dimensional space the 'spinor space'.

The equation (3.29) can be written in a way that treats the time and space component more symmetrically. This can be done by multiplying β from the left:

$$i\beta\frac{\partial}{\partial t}\psi = [\beta\alpha_i(-i\nabla_i) + m]\psi\,, \tag{3.31}$$

and defining four 'gamma' matrices out of α_i and β:

$$\gamma^0 \equiv \beta, \quad \gamma^i \equiv \beta\alpha_i \quad (i = 1, 2, 3)\,, \tag{3.32}$$

where we will adopt the standard superscript/subscript rule for the index of the gamma matrices. Then, together with $\partial_0 = \partial/\partial t, \nabla_i = \partial_i$, it can be written as

$$\Big[i(\gamma^0\partial_0 + \gamma^i\partial_i) - m\Big]\psi = 0 \tag{3.33}$$

or

$$\boxed{(i\gamma^\mu\partial_\mu - m)\psi = 0}\,. \tag{3.34}$$

This equation is called the *Dirac equation.* It is often written as

$$(i\partial\!\!\!/ - m)\psi = 0\,. \tag{3.35}$$

where we have defined the notation

$$a\!\!\!/ \stackrel{\text{def}}{\equiv} \gamma^\mu a_\mu \tag{3.36}$$

with a_μ being any 4-component quantity.

From the properties (3.9) of α_i and β, it can be easily shown that the four matrices introduced in (3.32) satisfy the following anticommutation relations of utmost importance:

$$\boxed{\{\gamma^\mu, \gamma^\nu\} = 2\,g^{\mu\nu}\,,} \tag{3.37}$$

or equivalently, ${\gamma^0}^2 = 1$, ${\gamma^i}^2 = -1$, and all four matrices anticommute among themselves. This relation (3.37) is sometimes called the Clifford algebra, and is entirely equivalent to (3.9).

Exercise 3.3 *Gamma matrices.*
Use the definition of the γ matrices $\gamma^0 = \beta$, $\gamma^i = \beta\alpha_i \quad (i = 1, 2, 3)$, and the relations among β and α_i (3.9) to verify the anticommutation relations (3.37).

The matrix γ^0 is hermitian while γ^i are anti-hermitian (i.e. its hermitian conjugate is the negative of itself) which can be easily seen as follows: since α_i and β are hermitian,

$$\begin{aligned} {\gamma^0}^\dagger &= \gamma^0, \qquad \text{(since } \gamma^0 = \beta) \\ {\gamma^i}^\dagger &= (\beta\alpha_i)^\dagger = \alpha_i^\dagger\beta^\dagger = \alpha_i\beta = -\beta\alpha_i = -\gamma^i\,. \end{aligned} \tag{3.38}$$

The following, however, is true for all μ:

$$\gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu \,. \tag{3.39}$$

Using the explicit representation (3.27) of α_i and β, the gamma matrices can be written as

$$\boxed{\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}}\,. \tag{3.40}$$

This is not the only explicit expression of the 4×4 matrices that satisfy (3.37). This particular representation is called the Dirac representation, and is the standard one we will use in this book. The anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2\, g^{\mu\nu}$ and the hermiticity relation $\gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu$, however, are independent of representation since they are direct consequences of the anticommutation relation (3.9) and hermiticity of α_i and β.

Just to make sure that we know what is going on, let's completely expand the Dirac equation $(i\partial\!\!\!/ - m)\psi = 0$, or $i\gamma^\mu \partial_\mu \psi = m\psi$, using the Dirac representation:

$$i\left[\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}_{\gamma^0} \partial_0 + \underbrace{\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}}_{\gamma^1} \partial_1 + \underbrace{\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}}_{\gamma^2} \partial_2 \right.$$
$$\left. + \underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_{\gamma^3} \partial_3 \right] \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = m \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \tag{3.41}$$

which is a set of four equations given by

$$\begin{aligned} i\,[\partial_0\psi_1 + (\partial_1 - i\partial_2)\psi_4 + \partial_3\psi_3] &= m\psi_1\,, \\ i\,[\partial_0\psi_2 + (\partial_1 + i\partial_2)\psi_3 - \partial_3\psi_4] &= m\psi_2\,, \\ i\,[-\partial_0\psi_3 - (\partial_1 - i\partial_2)\psi_2 - \partial_3\psi_1] &= m\psi_3\,, \\ i\,[-\partial_0\psi_4 - (\partial_1 + i\partial_2)\psi_1 + \partial_3\psi_2] &= m\psi_4\,. \end{aligned} \tag{3.42}$$

This looks complicated, but what matters is the structure of the gamma matrices, and it is almost all contained in the relation $\{\gamma^\mu, \gamma^\nu\} = 2\, g^{\mu\nu}$. It is actually as simple as indicated by the concise expression $(i\partial\!\!\!/ - m)\psi = 0$.

Exercise 3.4 *Hermiticity of $H = \vec{\alpha} \cdot (-i\vec{\nabla}) + m\beta$.*

Suppose a 4-component spinor ψ is an eigenvector of H with an eigenvalue c :

$$H\psi(x) = c\psi(x), \qquad \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \tag{3.43}$$

with

$$H = \alpha_i(-i\nabla_i) + m\beta\,, \tag{3.44}$$

or more explicitly,

$$-i\alpha_i(\nabla_i\psi) + m\beta\psi = c\psi\,. \tag{3.45}$$

The parameter m is real, and α_i, β are constant hermitian matrixes:

$$\alpha_i^\dagger = \alpha_i, \qquad \beta^\dagger = \beta\,. \tag{3.46}$$

Show that the eigenvalue c is real. Assume that the wave function $\psi(x)$ vanishes at infinite distance from the origin. (*hint: Multiply the row vector $\psi^\dagger$ to (3.45) from the left and integrate over d^3x, Then take hermitian conjugate of (3.45), multiply ψ from the right, and integrate over d^3x. Compare the two using a partial integration.*)

3.2 Conserved current

Now, let's see if we can construct a conserved current out of solutions of the Dirac equation. If we can somehow find a conserved current, then it is naturally interpreted as the probability current. The basic procedure is similar to the cases of the Schrödinger equation or the Klein-Gordon equation. Multiplying the hermitian conjugate of a spinor ψ

$$\psi^\dagger \stackrel{\text{def}}{\equiv} (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)\,. \tag{3.47}$$

to the Dirac equation $i\gamma^\mu\partial_\mu\psi = m\psi$ on the left,

$$i\,\psi^\dagger\gamma^\mu\partial_\mu\psi = m\psi^\dagger\psi\,, \tag{3.48}$$

which has the form

$$\overbrace{(\cdot\quad\cdot\quad\cdot\quad\cdot)}^{i\psi^\dagger}\overbrace{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}^{\gamma^\mu}\overbrace{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}}^{\partial_\mu\psi} = \overbrace{(\cdot\quad\cdot\quad\cdot\quad\cdot)}^{m\psi^\dagger}\overbrace{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}}^{\psi}\,. \tag{3.49}$$

Next, we take the hermitian conjugate of the Dirac equation $i\gamma^\mu \partial_\mu \psi = m\psi$ to obtain

$$-i\,(\partial_\mu\psi)^\dagger \gamma^{\mu\dagger} = m\,\psi^\dagger\,. \tag{3.50}$$

Multiplying ψ from the right,

$$-i\,(\partial_\mu\psi)^\dagger \gamma^{\mu\dagger}\psi = m\,\psi^\dagger\psi\,. \tag{3.51}$$

Subtracting (3.51) from (3.48),

$$i\,[\psi^\dagger\gamma^\mu(\partial_\mu\psi) + (\partial_\mu\psi)^\dagger \underbrace{\gamma^{\mu\dagger}}_{\gamma^\mu ?}\psi] = 0 \tag{3.52}$$

Now, only if $\gamma^{\mu\dagger} = \gamma^\mu$ were true, then this would become $i\,\partial_\mu(\psi^\dagger\gamma^\mu\psi) = 0$, and thus we would identify $\psi^\dagger\gamma^\mu\psi$ as the conserved current. However, γ^i ($i = 1, 2, 3$) are not hermitian as we have seen in (3.38). Instead, what we have is $\gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu$ (3.39). As we will see below, we can take advantage of this relation if we define a new kind of inner product of two spinors ψ_1 and ψ_2 by

$$\psi_1^\dagger\gamma^0\psi_2 \overset{\text{def}}{\equiv} \bar\psi_1\psi_2\,, \tag{3.53}$$

where we have defined a new kind of adjoint by

$$\boxed{\bar\psi \overset{\text{def}}{\equiv} \psi^\dagger\gamma^0}\,, \tag{3.54}$$

which we will call 'spinor adjoint'. Similarly, we define the spinor adjoint of a 4×4 matrix by

$$\boxed{\bar M \overset{\text{def}}{\equiv} \gamma^0 M^\dagger \gamma^0}\,. \tag{3.55}$$

Using this definition, the relation $\gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu$ can be written as

$$\boxed{\overline{\gamma^\mu} = \gamma^\mu}\,, \tag{3.56}$$

which is equivalent to $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$. Before moving on, let's become familiar with operations of the spinor adjoints.

Spinor adjoints

For two spinors a, b, and 4×4 matrices M_i ($i = 1, \ldots, n$) the quantity $\bar b\,M_1 \ldots M_n a$ is a complex number, and its complex conjugate is

$$(\bar b\,M_1 \ldots M_n a)^* \;=\; (b^\dagger\gamma^0 M_1 \ldots M_n a)^*$$

$$= a^\dagger M_n^\dagger \dots M_1^\dagger \underbrace{\gamma^{0\dagger}}_{\gamma^0} b$$

(inserting ${\gamma^0}^2 = 1$ almost everywhere)

$$= \underbrace{a^\dagger\gamma^0}_{\bar a}\underbrace{\gamma^0 M_n^\dagger\gamma^0}_{\bar M_n}\gamma^0 \dots \gamma^0 \underbrace{\gamma^0 M_1^\dagger\gamma^0}_{\bar M_1} b \tag{3.57}$$

Thus, we have

$$\boxed{(\bar b\, M_1 \dots M_n a)^* = \bar a\, \bar M_n \dots \bar M_1 b}\,. \tag{3.58}$$

Similarly, the adjoint of a column vector $M_1 \dots M_n a$ is by the definition (3.54),

$$\begin{aligned}
\overline{M_1 \dots M_n a} &\equiv (M_1 \dots M_n a)^\dagger \gamma^0 \\
&= a^\dagger M_n^\dagger \dots M_1^\dagger \gamma^0 \\
&= a^\dagger {\gamma^0}^2 M_n^\dagger {\gamma^0}^2 \dots {\gamma^0}^2 M_1^\dagger \gamma^0 \\
&= \bar a\, \bar M_n \dots \bar M_1\,.
\end{aligned} \tag{3.59}$$

Namely,

$$\boxed{\overline{M_1 \dots M_n a} = \bar a \bar M_n \dots \bar M_1}\,. \tag{3.60}$$

The general rule is thus similar to the case of hermitian conjugate; namely, one simply takes spinor adjoint of each, or take it away if it already has the adjoint symbol (e.g. $\bar{\bar a} = a$, $\bar{\bar M} = M$, etc.), and reverse the order. This applies also to products of matrices. Furthermore, a complex number becomes its complex conjugate when the spinor adjoint is taken:

$$\boxed{\overline{\eta a} = \eta^* \bar a \quad (\eta : \text{complex number})} \tag{3.61}$$

▮

Now we will repeat the failed attempt to construct a conserved current, this time replacing the hermitian conjugation with the spinor adjoint. Multiplying $\bar\psi$ to the Dirac equation $i\gamma^\mu\partial_\mu\psi = m\psi$ from the left,

$$i\,\bar\psi\gamma^\mu\partial_\mu\psi = m\,\bar\psi\psi\,. \tag{3.62}$$

Noting that $\overline{\partial_\mu\psi} = \partial_\mu\bar\psi$, we take the spinor adjoint of the Dirac equation,

$$-i\,(\partial_\mu\bar\psi)\,\underbrace{\overline{\gamma^\mu}}_{\gamma^\mu} = m\,\bar\psi\,, \tag{3.63}$$

where we have used $\overline{\gamma^\mu} = \gamma^\mu$ which is the critical step, then multiply ψ from the right to get

$$-i\,(\partial_\mu\bar{\psi})\gamma^\mu\psi = m\,\bar{\psi}\psi\,. \tag{3.64}$$

Subtracting (3.64) from (3.62), we then have a conserved current:

$$i\,[\bar{\psi}\gamma^\mu(\partial_\mu\psi) + (\partial_\mu\bar{\psi})\gamma^\mu\psi] = i\,\partial_\mu(\bar{\psi}\gamma^\mu\psi) = 0 \tag{3.65}$$

or

$$\boxed{\partial_\mu j^\mu = 0, \quad \text{with} \quad j^\mu \equiv \bar{\psi}\gamma^\mu\psi}\,. \tag{3.66}$$

Is this a Lorentz-invariant equation? Actually, we do not know how j^μ transforms under Lorentz transformation since we do not know how the spinor ψ transforms; it is not any one of the quantities we know so far: scalar, vector, tensor, etc. The spinor ψ is a 4-component quantity, but it cannot be a Lorentz 4-vector since the indexes have nothing to do with the space-time. Since we do not know how ψ transforms under Lorentz transformation, we do not even know if the Dirac equation is Lorentz-invariant or not at this point. Shortly, we will find how a spinor transforms under a Lorentz transformation by requiring that the Dirac equation becomes Lorentz-invariant. For now, however, let's examine the time component of the conserved current we just found, which is supposed to be the probability density: using $\bar{\psi} \equiv \psi^\dagger\gamma^0$,

$$\begin{aligned} j^0 &= \bar{\psi}\gamma^0\psi = (\psi^\dagger\underbrace{\gamma^0)\gamma^0}_{1}\psi = \psi^\dagger\psi \\ &= |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \geq 0\,. \end{aligned} \tag{3.67}$$

Thus, the probability density is always positive, and Dirac seems to have fixed the problem of negative probability. As mentioned earlier, however, this quantity will be reinterpreted as the charge current in the framework of the quantum field theory, and will acquire both positive and negative values.

Incidentally, (3.63) gives the equation that has to be satisfied by the adjoint spinor $\bar{\psi}$ in order for the original wave function ψ to be a solution of the Dirac equation:

$$-i\,(\partial_\mu\bar{\psi})\gamma^\mu = m\,\bar{\psi} \tag{3.68}$$

which is often written as

$$\bar{\psi}\,(i\gamma^\mu\overleftarrow{\partial}_\mu + m) = 0\,, \tag{3.69}$$

or

$$\bar{\psi}\,(i\overleftarrow{\not\partial} + m) = 0\,, \tag{3.70}$$

where the symbol $\overleftarrow{\partial}_\mu$ is defined to operate on everything to its left:

$$a\overleftarrow{\partial}_\mu b \stackrel{\text{def}}{\equiv} (\partial_\mu a)b\,. \tag{3.71}$$

This funny notation is a result of a compromise between the desire to put γ^μ and ∂_μ next to each other so that it can be combined as $\overleftarrow{\not\partial}$ and the necessity to put the matrix γ^μ to the right of $\bar\psi$ which is a row vector.

3.3 Lorentz invariance of the Dirac equation

We have seen that the Klein-Gordon equation is Lorentz-invariant, which meant the following: if a function of space-time $\phi(x)$ satisfies the equation $(\partial^2 + m^2)\phi(x) = 0$, then a new function defined by

$$\phi'(x') = \phi(x), \quad x' = \Lambda x \tag{3.72}$$

satisfies $(\partial'^2 + m^2)\phi'(x') = 0$. Note that the function $\phi'(x')$ is uniquely defined once the original function $\phi(x)$ and the Lorentz transformation Λ are given. In particular, a critical condition in defining $\phi'(x')$ uniquely was the definition of scalar field $\phi'(x') = \phi(x)$, which means that the value at an event point x in the original frame is the same in the transformed frame if measured at the same event point (now given by the coordinate $x' = \Lambda x$ in that frame). The value of ϕ' at a given event point is completely defined by the value of ϕ at the same event point, and does not depend on the values of any other event points. Also, note that the functional shape of ϕ' is in general different from that of ϕ; namely, if we give the same argument to ϕ' and ϕ, then in general $\phi'(x) \neq \phi(x)$.

Similarly, the electromagnetic wave equation

$$\partial^2 A^\mu(x) = j^\mu(x) \quad (\mu = 0, 1, 2, 3)\,, \tag{3.73}$$

where the electromagnetic 4-potential A^μ and the charge current j^μ are both Lorentz 4-vectors, is Lorentz-invariant. [3]

Namely, if $A^\mu(x)$ and $j^\mu(x)$ satisfy the above equation, then a new set of fields defined by the vector field condition

$$A'^\mu(x') = \Lambda^\mu{}_\alpha A^\alpha(x), \quad j'^\mu(x') = \Lambda^\mu{}_\beta\, j^\beta(x), \quad x' = \Lambda x\,, \tag{3.74}$$

satisfy $\partial'^2 A'^\mu(x') = j'^\mu(x')$ as can be readily verified. The function $A^\mu(x)$ [or $A'^\mu(x')$] assigns a set of four numbers to each event point, and the four values in the transformed frame $A'^\mu(x')$ at an event point are completely determined by the four values in the original frame at the same event point, they are simply mixed up by the matrix Λ. Note that the new values $A'^\mu(x')$ do not depend on the values of $A^\mu(x)$ at other event points.

[3] If one uses a gauge which is not Lorentz-invariant, such as the Coulomb gauge, then the correct statement is that the equation is Lorentz-invariant up to gauge transformation. This complication comes about due to the masslessness of photon. We will discuss this point in detail later.

Then, how does a Dirac spinor field $\psi(x)$ transform? We do not know at this point, except that as in the case of scalar and vector fields, the four values $\psi'(x')$ in a transformed frame will depend only on the four values in the original frame $\psi(x)$ associated with the same event point, presumably mixed up by some 4×4 matrix S:

$$\boxed{\psi'(x') = S(\Lambda)\psi(x)\,, \quad x' = \Lambda x}\,. \tag{3.75}$$

Also, just like the mixing matrix (Λ) for the vector field depended only on the Lorentz transformation and did not depend on event point x, we expect that S also depends only on the Lorentz transformation Λ.

Our strategy is to derive the condition for S that makes the Dirac equation Lorentz-invariant, then explicitly construct such a matrix, out of which we will obtain explicit solutions of the Dirac equation. We will then find that the solutions contain the spin-1/2 structure and the particle-antiparticle degrees of freedom.

Thus, we require that if $\psi'(x')$ satisfy $(i\not\partial' - m)\psi'(x') = 0$ then it leads to $(i\not\partial - m)\psi(x) = 0$ when $\psi'(x')$ and $\psi(x)$ are related by (3.75). Specifically, we require that the numerical values of γ^μ will be the same in any frame; namely, for the Dirac representation, the form of the four equations is given by (3.42) and is the same in any frame. Using $\psi'(x') = S\psi(x)$, $(i\not\partial' - m)\psi'(x') = 0$ becomes

$$\begin{aligned} 0 &= (i\gamma^\mu \underbrace{\partial'_\mu}_{\Lambda_\mu{}^\alpha \partial_\alpha} - m)\, \underbrace{\psi'(x')}_{S\psi(x)} \\ &= i\,\gamma^\mu \underbrace{\Lambda_\mu{}^\alpha \partial_\alpha S\psi(x)}_{S\Lambda_\mu{}^\alpha \partial_\alpha \psi(x)} - mS\psi(x) \\ (\times S^{-1}) \qquad &= i\,S^{-1}\gamma^\mu S \Lambda_\mu{}^\alpha \partial_\alpha \psi(x) - m\psi(x)\,, \end{aligned} \tag{3.76}$$

where in the second line, S can come out of ∂_α since S is a constant matrix, and S can move past $\Lambda_\mu{}^\alpha$ because $\Lambda_\mu{}^\alpha$ is just a number for given α and μ. Note that we have consistently suppressed the spinor indexes while all space-time indexes have been explicitly written out. The necessary condition for this to become $i\,\gamma^\alpha \partial_\alpha \psi(x) - m\psi(x) = 0$ for any $\psi(x)$ is then

$$S^{-1}\gamma^\mu S \Lambda_\mu{}^\alpha = \gamma^\alpha \tag{3.77}$$

or multiplying $\Lambda^\nu{}_\alpha$ and summing over α,

$$S^{-1}\gamma^\mu S \underbrace{\Lambda_\mu{}^\alpha \Lambda^\nu{}_\alpha}_{g_\mu{}^\nu} = \Lambda^\nu{}_\alpha \gamma^\alpha\,; \tag{3.78}$$

namely,

$$\boxed{S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha \gamma^\alpha}\,. \tag{3.79}$$

By tracing back the derivation, one sees that this is also a sufficient condition; namely, if a spinor field transforms under a Lorentz transformation Λ by $\psi'(x') = S\psi(x)$ where S is a matrix that satisfies $S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha\gamma^\alpha$, then the Dirac equation becomes Lorentz-invariant. Note that we did not restrict ourselves to proper and orthochronous Lorentz transformations; thus, the condition (3.79) applies to any Lorentz transformations including T and P.

The left hand side of (3.79) indicates a certain transformation of γ^μ, and the right hand side is exactly like the transformation of a vector. Does that mean that the gamma matrices that appear in the Dirac equation transform as a vector? The answer is no. In fact, the condition (3.79) was obtained by requiring that the gamma matrices in the Dirac equation *do not change* under a Lorentz transformation. It will turn out, however, that one can use (3.79) to construct quantities that transform as a vector. One example is the conserved current $j^\mu = \bar\psi\gamma^\mu\psi$ as will be discussed in detail shortly.

Let's move on to actually constructing $S(\Lambda)$. Our starting point is that the mapping between S and Λ preserves the product rule, which can be seen as follows. Suppose $S(\Lambda_1)$, $S(\Lambda_2)$, and $S(\Lambda_1\Lambda_2)$ correspond to Λ_1, Λ_2, and $\Lambda_1\Lambda_2$, respectively:

$$\begin{aligned} S(\Lambda_1) &\leftrightarrow \Lambda_1\,, \\ S(\Lambda_2) &\leftrightarrow \Lambda_2\,, \\ S(\Lambda_1\Lambda_2) &\leftrightarrow \Lambda_1\Lambda_2\,. \end{aligned} \tag{3.80}$$

Under the Lorentz transformation Λ_1, a spinor ψ will be transformed to $S(\Lambda_1)\psi$. If we perform an additional transformation Λ_2, which makes the total transformation $\Lambda_2\Lambda_1$, then this spinor will transform to $S(\Lambda_2)[S(\Lambda_1)\psi] = [S(\Lambda_2)S(\Lambda_1)]\psi$. Since this should hold for any spinor ψ, we have $S(\Lambda_2\Lambda_1) = S(\Lambda_2)S(\Lambda_1)$, or

$$S(\Lambda_2)S(\Lambda_1) \leftrightarrow \Lambda_2\Lambda_1\,, \tag{3.81}$$

which means that S's and Λ's have exactly the same group structure [see (1.151) through (1.158)]; in fact, S is said to be the *representation of the Lorentz group* in the spinor-space. If we restrict ourselves to Lorentz transformations that are continuously connected to the identity, then S and Λ can be written in exponential forms using generators, and the generators for S and those for Λ should have the same structure constants. The problem then reduces to finding the set of generators for S that satisfy the same commutation relations as $M^{\mu\nu}$ (or K_i and L_i) *and* are consistent with $S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha\gamma^\alpha$.

Suppose $B^{\mu\nu}$ be the generators for S corresponding to the generators $M^{\mu\nu}$ for Λ. Namely, the mapping between S and Λ is given by the same real parameters $a_{\mu\nu}$:

$$S = e^{\frac{1}{2}a_{\mu\nu}B^{\mu\nu}} \leftrightarrow \Lambda = e^{\frac{1}{2}a_{\mu\nu}M^{\mu\nu}}\,. \tag{3.82}$$

Let's first find the condition for $B^{\mu\nu}$ that satisfy $S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha\gamma^\alpha$. Assume that $a_{\mu\nu}$ are small and expand both sides of $S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^\nu$: using $(e^A)^{-1} = e^{-A}$ (1.114),

$$\begin{aligned} S^{-1}\gamma^\mu S &= e^{-\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}}\gamma^\mu e^{\frac{1}{2}a_{\alpha'\beta'}B^{\alpha'\beta'}} \\ &= \Big(1-\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}\Big)\gamma^\mu\Big(1+\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}\Big)+\cdots \\ &= \gamma^\mu + \frac{1}{2}\underbrace{\Big(\gamma^\mu a_{\alpha\beta}B^{\alpha\beta} - a_{\alpha\beta}B^{\alpha\beta}\gamma^\mu\Big)}_{a_{\alpha\beta}[\gamma^\mu, B^{\alpha\beta}]}+\cdots\,, \end{aligned} \tag{3.83}$$

and the right hand side is

$$\begin{aligned} \Lambda^\mu{}_\nu\gamma^\nu &= \Big(e^{\frac{1}{2}a_{\alpha\beta}M^{\alpha\beta}}\Big)^\mu{}_\nu\gamma^\nu \\ &= \Big(1+\frac{1}{2}a_{\alpha\beta}M^{\alpha\beta}\Big)^\mu{}_\nu\gamma^\nu+\cdots \\ &= \Big[g^\mu{}_\nu + \frac{1}{2}a_{\alpha\beta}(M^{\alpha\beta})^\mu{}_\nu\Big]\gamma^\nu+\cdots \\ &= \gamma^\mu + \frac{1}{2}a_{\alpha\beta}(M^{\alpha\beta})^\mu{}_\nu\gamma^\nu+\cdots\,. \end{aligned} \tag{3.84}$$

Requiring that (3.83) is equal to (3.84) for any $a_{\alpha\beta}$, we obtain

$$[\gamma^\mu, B^{\alpha\beta}] = (M^{\alpha\beta})^\mu{}_\nu\gamma^\nu\,, \tag{3.85}$$

which is the condition that the generators of S ($B^{\mu\nu}$'s) have to satisfy in order for the Dirac equation to be Lorentz-invariant under the (infinitesimal) transformation $\psi'(x') = S\psi(x)$. Note that γ^μ and $B^{\mu\nu}$ are 4×4 matrices (for given μ,ν) that operate in the spinor space while $(M^{\alpha\beta})^\mu{}_\nu$ is just a number. Again, we are suppressing the spinor indexes while writing out space-time indexes.

We will now show that the solution to (3.85) is given by

$$B^{\alpha\beta} = \frac{1}{4}[\gamma^\alpha, \gamma^\beta] \tag{3.86}$$

or equivalently (using $\{\gamma^\alpha, \gamma^\beta\} = 2g^{\alpha\beta}$),

$$B^{\alpha\beta} = \begin{cases} \frac{1}{2}\gamma^\alpha\gamma^\beta & (\alpha\neq\beta) \\ 0 & (\alpha=\beta) \end{cases}\,. \tag{3.87}$$

For $\alpha=\beta$, the condition (3.85) is clearly satisfied since in that case $M^{\alpha\beta}=0$ and $B^{\alpha\beta}=0$. For $\alpha\neq\beta$, we have to work out the commutator $[\gamma^\mu, \gamma^\alpha\gamma^\beta]$. Using the identity

$$[A, BC] = \underbrace{ABC - BCA}_{+BAC-BAC} = \{A,B\}C - B\{A,C\}\,, \tag{3.88}$$

we have

$$
\begin{aligned}
[\gamma^\mu, \gamma^\alpha\gamma^\beta] &= \underbrace{\{\gamma^\mu, \gamma^\alpha\}}_{2g^{\mu\alpha}}\gamma^\beta - \gamma^\alpha \underbrace{\{\gamma^\mu, \gamma^\beta\}}_{2g^{\mu\beta}} \\
&= 2\, g^{\alpha\mu}\gamma^\beta - 2\, g^{\beta\mu}\gamma^\alpha \\
&= 2(\underbrace{g^{\alpha\mu}g^\beta{}_\nu - g^{\beta\mu}g^\alpha{}_\nu}_{(M^{\alpha\beta})^\mu{}_\nu})\gamma^\nu\,, \qquad (3.89)
\end{aligned}
$$

where we have used the explicit expression $(M^{\alpha\beta})_{\mu\nu} = g^\alpha{}_\mu\, g^\beta{}_\nu - g^\beta{}_\mu\, g^\alpha{}_\nu$ (1.91). Dividing by 2 on both sides, we have

$$
\left[\gamma^\mu, \frac{1}{2}\,\gamma^\alpha\gamma^\beta\right] = (M^{\alpha\beta})^\mu{}_\nu\gamma^\nu \qquad (\alpha \neq \beta)\,. \tag{3.90}
$$

Namely, the condition (3.85) is satisfied for $B^{\mu\nu}$ given by (3.87) [or (3.86)] for all α and β.

We have shown that for an *infinitesimal Lorentz transformation* given by parameters $a_{\alpha\beta}$, the matrix $S = 1 + (a_{\alpha\beta}/2)B^{\alpha\beta}$ satisfies $S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha\gamma^\alpha$ to the first order in $a_{\alpha\beta}$ if $B^{\alpha\beta}$ is given by $B^{\alpha\beta} = [\gamma^\alpha, \gamma^\beta]/4$. Does it hold for a finite Lorentz transformation? Namely, if Λ and S are finite transformations mapped as (3.82) with the same set of parameters $a_{\alpha\beta}$, then do they satisfy $S^{-1}\gamma^\nu S = \Lambda^\nu{}_\alpha\gamma^\alpha$? This can be easily proven by dividing Λ and S into n consecutive small transformations:

$$
\Lambda = \lambda^n = \left(1 + \frac{1}{2}\frac{a_{\alpha\beta}}{n}M^{\alpha\beta}\right)^n, \qquad S = s^n = \left(1 + \frac{1}{2}\frac{a_{\alpha\beta}}{n}B^{\alpha\beta}\right)^n. \tag{3.91}
$$

Since s and λ are infinitesimal transformations mapped by the same set of parameters, we have already proven that

$$
s^{-1}\gamma^\mu s = \lambda^\mu{}_\nu\gamma^\nu\,. \tag{3.92}
$$

Multiplying s^{-1} from the left and s from the right, we get

$$
(s^{-1})^2\gamma^\mu s^2 = \lambda^\mu{}_\nu \underbrace{s^{-1}\gamma^\nu s}_{\lambda^\nu{}_\alpha\gamma^\alpha} = (\lambda^\mu{}_\nu\lambda^\nu{}_\alpha)\gamma^\alpha = (\lambda^2)^\mu{}_\alpha\gamma^\alpha\,. \tag{3.93}
$$

Repeating the process n times, we obtain

$$
\begin{aligned}
\underbrace{(s^{-1})^n}_{(s^n)^{-1}}\gamma^\mu s^n &= \underbrace{(\lambda^n)^\mu{}_\nu}_{\Lambda^\mu{}_\nu}\gamma^\nu \\
\to \quad S^{-1}\gamma^\mu S &= \Lambda^\mu{}_\nu\gamma^\nu\,. \qquad (3.94)
\end{aligned}
$$

Thus, we have shown that for any proper and orthochronous Lorentz transformation Λ given by

$$
\Lambda = e^{\frac{1}{2}a_{\alpha\beta}M^{\alpha\beta}}\,, \tag{3.95}
$$

one can construct a transformation in the spinor space by

$$\boxed{S = e^{\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}} \quad \text{with} \quad B^{\alpha\beta} = \frac{1}{4}\left[\gamma^\alpha, \gamma^\beta\right]}, \tag{3.96}$$

such that $S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu$ holds, or equivalently, the Dirac equation becomes invariant under the transformation $\psi'(x') = S\psi(x)$. It is straightforward to show that the generators $B^{\alpha\beta}$ satisfy the same commutation relations as those of $M^{\alpha\beta}$, and the proof is left as an exercise.

Exercise 3.5 *Representation of Lorentz group in the spinor space.*
Define the generators of boost and rotation in the 4-component spinor space by

$$\text{boost}: \quad K_i \leftrightarrow B_i^b \equiv B^{0i}, \quad \text{rotation}: \quad L_i \leftrightarrow B_i^r \equiv B^{jk} \;\; (ijk : \text{cyclic}). \tag{3.97}$$

Show explicitly that the B^b's and B^r's satisfy the same commutation relations as satisfied by the K's and L's; namely

$$\begin{aligned} [B_i^b, B_j^b] &= -\epsilon_{ijk}B_k^r \\ [B_i^r, B_j^r] &= \epsilon_{ijk}B_k^r \\ [B_i^r, B_j^b] &= \epsilon_{ijk}B_k^b, \end{aligned} \tag{3.98}$$

(It is also possible to prove generally. In any case, all you need is $B^{\alpha\beta} = 1/2\gamma^\alpha\gamma^\beta$ ($\alpha \neq \beta$) and the relation $\{\gamma^\mu, \gamma^\nu\} = 2\,g^{\mu\nu}$; do not use any explicit representation of the γ matrixes).

An important property of S is that its spinor adjoint is the inverse:

$$\boxed{\bar{S} = S^{-1}, \quad \text{or} \quad \bar{S}S = 1}, \tag{3.99}$$

which can be seen by noting that for $\alpha \neq \beta$,

$$\overline{B^{\alpha\beta}} = \frac{1}{2}\overline{\gamma^\alpha\gamma^\beta} = \frac{1}{2}\overline{\gamma^\beta}\,\overline{\gamma^\alpha} = \frac{1}{2}\gamma^\beta\gamma^\alpha = -\frac{1}{2}\gamma^\alpha\gamma^\beta = -B^{\alpha\beta} \tag{3.100}$$

(for $\alpha = \beta$, $\overline{B^{\alpha\beta}} = -B^{\alpha\beta}$ holds trivially since $B^{\alpha\beta} = 0$) and thus,

$$\begin{aligned} \bar{S} &= \overline{e^{\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}}} = \sum_{k=0}^{\infty}\frac{1}{k!}\big(\tfrac{1}{2}a_{\alpha\beta}\underbrace{\overline{B^{\alpha\beta}}}_{-B^{\alpha\beta}}\big)^k \\ &= e^{-\frac{1}{2}a_{\alpha\beta}B^{\alpha\beta}} = S^{-1}, \end{aligned} \tag{3.101}$$

where we have used $\overline{M^k} = (\bar{M})^k$ and $e^{-A} = (e^A)^{-1}$.

space	Lorentz transformation	inner product	metric invariance
space-time	$A' = \Lambda A$	$A \cdot B \equiv A^T G B$	$\Lambda^T G \Lambda = G$
spinor	$a' = Sa$	$\bar{a}b \equiv a^\dagger \gamma^0 b$	$S^\dagger \gamma^0 S = \gamma^0$

Table 3.1: Correspondence between the space-time and the spinor space. A and B are 4-vectors and a and b are 4-component spinors.

Using (3.99), we can easily see that the inner product $\bar{a}b$ of two spinors a and b is invariant under Lorentz transformation in the spinor space:

$$a' = Sa\,, \quad b' = Sb\,, \tag{3.102}$$

$$\bar{a}'b' = \underbrace{\overline{Sa}}_{\bar{a}\bar{S}}\, Sb = \bar{a}\,\underbrace{\bar{S}S}_{1}\, b = \bar{a}b\,. \tag{3.103}$$

The relation $\bar{S} = S^{-1}$ or $\gamma^0 S^\dagger \gamma^0 = S^{-1}$ can also be written as (by multiplying γ^0 from the left and S from the right)

$$S^\dagger \gamma^0 S = \gamma^0\,. \tag{3.104}$$

Note the parallel between the space-time and the spinor space: in the space-time, the Lorentz transformation of a 4-vector A is given by $A' = \Lambda A$ which keeps the inner product $A \cdot B = A^T G B$ (1.23) invariant, and the metric G is unchanged under the transformation: $\Lambda^T G \Lambda = G$. In the spinor space, the Lorentz transformation of a spinor a is given by $a' = Sa$ which keeps the inner product $\bar{a}b = a^\dagger \gamma^0 b$ invariant, and the 'metric' γ^0 is invariant under the transformation: $S^\dagger \gamma^0 S = \gamma^0$. Table 3.1 summarizes the parallelism.

Now it is trivial to show that the conserved current we derived earlier $j^\mu(x)$ (3.66) is indeed a 4-vector: using $\psi'(x') = S\psi(x)$,

$$\begin{aligned}
j'^\mu(x') &= \bar{\psi}'(x')\gamma^\mu \psi'(x') \\
&= \bar{\psi}(x)\,\underbrace{\bar{S}\,\gamma^\mu\, S}_{S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu \text{ by (3.79)}}\,\psi(x) \\
&= \Lambda^\mu{}_\nu \bar{\psi}(x)\gamma^\nu \psi(x) \\
\to \; j'^\mu(x') &= \Lambda^\mu{}_\nu j^\nu(x)\,.
\end{aligned} \tag{3.105}$$

Note that the current j^μ is a 4-vector even if $\psi(x)$ does not satisfy the Dirac equation. If $\psi(x)$ does satisfy the Dirac equation, then j^μ is conserved: $\partial_\mu j^\mu = 0$, and this, we now know, is a Lorentz-invariant statement; namely, if the current is conserved in one frame, then it is conserved in any other frame. Similarly, a current consisting of two spinor fields $\psi_1(x)$ and $\psi_2(x)$

$$j^\mu_{12}(x) = \bar{\psi}_1(x)\gamma^\mu\psi_2(x) \tag{3.106}$$

is also a 4-vector; namely,

$$j'^\mu_{12}(x') = \Lambda^\mu{}_\nu j^\nu_{12}(x)\,, \tag{3.107}$$

which can be easily shown to be conserved if ψ_1 and ψ_2 are solutions of the Dirac equation (with the same mass).

Thus, we have constructed a scalar quantity and a vector quantity out of two spinors ψ_1 and ψ_2. These are not the only quantities one can form out of two spinors that transforms in a well-defined manner under Lorentz transformation. Such quantities are called the bilinear covariants which we will study systematically in the next section.

3.4 Bilinear covariants

Let's consider a complex number of the form

$$f = \bar{a}\Gamma b \tag{3.108}$$

where Γ is a constant 4×4 complex matrix, and a and b are spinors; namely they transform under a Lorentz transformation Λ as

$$a' = S(\Lambda)a\,, \quad b' = S(\Lambda)b\,. \tag{3.109}$$

Or a and b may be spinor *fields*:

$$a'(x') = S(\Lambda)a(x)\,, \quad b'(x') = S(\Lambda)b(x) \qquad (x' = \Lambda x)\,. \tag{3.110}$$

The following discussion applies to both cases.

A complex 4×4 matrix has 16 *complex* elements, and thus any such matrix can be written as a linear combination of 16 independent matrices with *complex* coefficients. As we will see below, the 16 independent matrices can be chosen such that the quantities

$$\bar{a}\Gamma_i b \quad (i = 1, \ldots, 16) \tag{3.111}$$

are grouped into five sets each of which transform in a well-defined manner under *proper and orthochronous transformations* (scalar, vector and tensor) and the *space inversion P* (how it changes sign). Such quantities are called the bilinear covariants. To study them, let's first find the spinor representation of the space inversion P.

Space inversion in the spinor space
Recall that the condition for the invariance of the Dirac equation, $S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^\nu$, holds for $\Lambda = P$ also, where P is the space inversion:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{3.112}$$

which is also called the *parity transformation*. Namely, $S_P \equiv S(P)$ should satisfy

$$S_P^{-1}\gamma^0 S_P = P^0{}_\nu\gamma^\nu = \gamma^0 \quad (\mu = 0) \tag{3.113}$$
$$S_P^{-1}\gamma^i S_P = P^i{}_\nu\gamma^\nu = -\gamma^i \quad (\mu = i)\,. \tag{3.114}$$

This is accomplished by taking

$$S_P = \eta\gamma^0\,, \tag{3.115}$$

where η is an arbitrary complex constant: using $\gamma^{0^{-1}} = \gamma^0$,

$$S_P^{-1}\gamma^0 S_P = (\eta^{-1}\gamma^0)\gamma^0\eta\gamma^0 = \gamma^0 \tag{3.116}$$
$$S_P^{-1}\gamma^i S_P = (\eta^{-1}\gamma^0)\gamma^i\eta\gamma^0 = -\gamma^i\,. \tag{3.117}$$

We then require that S_P also satisfies the property $\bar{S}S = 1$ (3.99):

$$\bar{S}_P S_P = \overline{\eta\gamma^0}\,\eta\gamma^0 = \eta^*\eta\,{\gamma^0}^2 = |\eta|^2 = 1\,. \tag{3.118}$$

We will now arbitrarily choose[4] η to be $+1$:

$$\boxed{S_P = \gamma^0 \qquad (\eta = +1)}\,. \tag{3.119}$$

Here we are requiring that the Dirac equation be invariant under space inversion, but should it be? Experiments tell us that electromagnetic and strong interactions are invariant under space inversion, but weak interaction is not. Since there are some interactions that are invariant under space inversion, the free field part should better be invariant also. In fact, the rule of the game is to build in as many symmetries as possible into the theory as long as they do not contradict with experimental facts; such theory will have more predictive power than otherwise. Later, we will deal with the question of discrete symmetries in more detail in the context of quantized field.

[4]As we will see in a later chapter, this choice of the phase factor for S_P corresponds to choosing the intrinsic parity of the particle represented by the spinor field to be $+1$ and that of the antiparticle to be -1, which is purely by convention. As it turns out, all that is physically significant is that the phase factor of a spin-1/2 particle be the negative of the complex conjugate of the phase factor of its antiparticle.

Now, in the Dirac representation the space inversion in the spinor space amounts to changing the sign of the bottom two components: noting that $x' = Px$,

$$\psi'(x') = S_P\psi(x) = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} = \begin{pmatrix} \psi_1(P^{-1}x') \\ \psi_2(P^{-1}x') \\ -\psi_3(P^{-1}x') \\ -\psi_4(P^{-1}x') \end{pmatrix} . \tag{3.120}$$

Namely, if $\psi(x)$ satisfies the Dirac equation $(i\partial\!\!\!/ - m)\psi(x) = 0$, then $\psi'(x')$ defined above satisfies the Dirac equation $(i\partial\!\!\!/' - m)\psi'(x') = 0$.

Five types of bilinear covariants

Now we will write down five types of bilinear covariants and check the transformation properties under proper and orthochronous transformations and under the parity transformation S_P. In this section, we will denote a proper and orthochronous transformation by S for simplicity.

All we use in the following is

$$\begin{gathered} \bar{S}\gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu , \qquad S_P = \gamma^0 , \\ \bar{S}S = 1 , \qquad \bar{S}_P S_P = 1 . \end{gathered} \tag{3.121}$$

Transformed quantities, either by a proper and orthochronous transformation S or by the parity transformation S_P, are denoted with a prime.

1. **Scalar**

$$f = \bar{a}b : \quad \text{transforms as} \quad f' = f \ (S) , \quad f' = f \ (S_P) . \tag{3.122}$$

We have already seen that $f' = f$ under a proper and orthochronous transformation S (3.103). Under the parity transformation, we have

$$f' = \overline{a'}b' = \overline{S_P a}\, S_P b = \bar{a}\, \underbrace{\bar{S}_P S_P}_{1}\, b = \bar{a}b = f \tag{3.123}$$

2. **Vector**

$$f^\mu = \bar{a}\gamma^\mu b : \quad \text{transforms as} \quad f'^\mu = \Lambda^\mu{}_\nu f^\nu \ (S) , \quad f'^\mu = f_\mu \ (S_P) \tag{3.124}$$

Again, we have already seen in (3.107) that the set of four quantities, $\bar{a}\gamma^\mu b$ ($\mu = 0, 1, 2, 3$), as a group transforms as a 4-vector (thereby the name 'bilinear *covariants*'). Under S_P, it transforms as

$$f'^\mu = \overline{a'}\gamma^\mu b' = \overline{S_P a}\gamma^\mu S_P b = \bar{a}\, \underbrace{\gamma^0\gamma^\mu\gamma^0}_{\gamma_\mu}\, b = f_\mu . \tag{3.125}$$

Note that the superscript μ changed to a subscript; namely, the sign of the time component is unchanged while that of space components flipped, which is the same as the transformation of the familiar energy-momentum 4-vector $P^\mu = (E, \vec{P})$ under space inversion. This use of the Lorentz index is a bit sloppy, since it is mixing up the metric $g_{\mu\nu}$ and the Lorentz transformation P which happen to have the same matrix value. The correct expression would be $f'^\mu = P^\mu{}_\nu f^\nu$. Still, it makes expressions compact that we will use it with this point in mind.

3. **Tensor**

$$f^{\mu\nu} = \bar{a}\sigma^{\mu\nu}b \; : \; \text{transforms as } f'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta f^{\alpha\beta} \; (S)\,, \; f'^{\mu\nu} = f_{\mu\nu} \; (S_P) \quad (3.126)$$

where

$$\boxed{\sigma^{\mu\nu} \overset{\text{def}}{\equiv} \frac{i}{2}[\gamma^\mu, \gamma^\nu] = \begin{cases} i\,\gamma^\mu\gamma^\nu & (\mu \neq \nu) \\ 0 & (\mu = \nu) \end{cases}} \quad (3.127)$$

Even though $\Gamma = \gamma^\mu\gamma^\nu$ would do just fine, defining $\sigma^{\mu\nu}$ this way makes it explicitly antisymmetric with respect to the μ and ν indexes, and also the addition of i makes it hermitian for $\mu, \nu = 1, 2, 3$. Note that we have $\sigma^{\mu\nu} = 2iB^{\mu\nu}$. Since $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$, $f^{\mu\nu}$ is an antisymmetric tensor, and there are six independent quantities that transform as a set.

The transformation properties can be verified straightforwardly: using the linearity of the commutator and

$$\bar{S}\,[\gamma^\mu, \gamma^\nu]S = \bar{S}(\gamma^\mu \underbrace{\,}_{S\bar{S}} \gamma^\nu - \gamma^\nu \underbrace{\,}_{S\bar{S}} \gamma^\mu)S = [\bar{S}\gamma^\mu S, \bar{S}\gamma^\nu S]\,, \quad (3.128)$$

the transformation under S is

$$\begin{aligned} f'^{\mu\nu} = \bar{a}'\sigma^{\mu\nu}b' &= \bar{a}\bar{S}\,\frac{i}{2}[\gamma^\mu, \gamma^\nu]Sb \\ &= \bar{a}\,\frac{i}{2}[\bar{S}\gamma^\mu S, \bar{S}\gamma^\nu S]b \\ &= \bar{a}\,\frac{i}{2}[\Lambda^\mu{}_\alpha\gamma^\alpha, \Lambda^\nu{}_\beta\gamma^\beta]b \\ &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\,\bar{a}\,\frac{i}{2}[\gamma^\alpha, \gamma^\beta]b = \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta f^{\alpha\beta} \end{aligned} \quad (3.129)$$

For $S_P (= \bar{S}_P = \gamma^0)$, it suffices to check only the non-zero components (namely, for $\mu \neq \nu$):

$$f'^{\mu\nu} = \bar{a}'\sigma^{\mu\nu}b' \;=\; \bar{a}\bar{S}_P\,\sigma^{\mu\nu}S_P\,b$$

$$
\begin{aligned}
&= \bar{a}\gamma^0(i\underbrace{\gamma^\mu\gamma^\nu)\gamma^0}_{\gamma^0\gamma_\mu\gamma_\nu} b \\
&= \bar{a}\,(i\gamma_\mu\gamma_\nu)b = f_{\mu\nu}\,, \qquad (3.130)
\end{aligned}
$$

where the indexes μ and ν became subscripts when γ^0 was moved over $\gamma^\mu\gamma^\nu$ because γ^0 commutes with γ^0 but anticommutes with γ^i $(i = 1, 2, 3)$.

4. **Pseudoscalar**

$$
f = \bar{a}\gamma_5 b\ : \quad \text{transforms as} \quad f' = f\ (S)\,, \quad f' = -f\ (S_P) \qquad (3.131)
$$

Namely, a pseudoscalar transforms under proper and orthochronous transformation just like a scalar (i.e. it does not change its value), but changes sign under space inversion. The γ_5 matrix in the expression above is defined as

$$
\boxed{\gamma_5 \stackrel{\text{def}}{\equiv} i\,\gamma^0\gamma^1\gamma^2\gamma^3}\,, \qquad (3.132)
$$

which, in the Dirac representation, is

$$
\gamma_5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\,. \qquad (3.133)
$$

The index '5' is not a Lorentz index and thus there is no distinction between superscript and subscript. It is designed to anticommute with any one of γ^μ $(\mu = 0, 1, 2, 3)$:

$$
\gamma^\mu\gamma_5 = \gamma^\mu(i\,\gamma^0\gamma^1\gamma^2\gamma^3) = -(i\,\gamma^0\gamma^1\gamma^2\gamma^3)\gamma^\mu = -\gamma_5\gamma^\mu \qquad (3.134)
$$

where the minus sign arises because, when γ^μ moves over the four gamma matrices $\gamma^0\gamma^1\gamma^2\gamma^3$, one is γ^μ itself and three others will anticommute with γ^μ. It is also easy to see that γ_5 is hermitian and the square is 1. Namely,

$$
\boxed{\begin{array}{c} \{\gamma^\mu, \gamma_5\} = 0 \quad (\mu = 0, 1, 2, 3) \\ \gamma_5^\dagger = \gamma_5\,, \quad {\gamma_5}^2 = 1 \end{array}}\,. \qquad (3.135)
$$

It follows that γ_5 commutes with $B^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$ and thus with any proper and orthochronous transformation S:

$$
\gamma_5 B^{\mu\nu} = \gamma_5\frac{1}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) = \frac{1}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\gamma_5 = B^{\mu\nu}\gamma_5 \qquad (3.136)
$$

$$
\gamma_5 S = \gamma_5\ e^{\frac{1}{2}a_{\mu\nu}B^{\mu\nu}} = \gamma_5\sum_k \frac{\left(\frac{1}{2}a_{\mu\nu}B^{\mu\nu}\right)^k}{k\,!} = S\gamma_5\,. \qquad (3.137)
$$

Namely,

$$\boxed{[S, \gamma_5] = 0}\,. \tag{3.138}$$

The transformation properties can now be verified easily. Under a proper and orthochronous transformation, we have

$$f' = (\bar{a}\bar{S})\underbrace{\gamma_5(S\,b)}_{S\gamma_5} = \bar{a}\,\underbrace{\bar{S}S}_{1}\,\gamma_5 b = f\,, \tag{3.139}$$

and under S_P, it transforms as

$$f' = (\bar{a}\bar{S}_P)\gamma^5(S_P b) = (\bar{a}\gamma^0)\underbrace{\gamma_5(\gamma^0}_{-\gamma^0\gamma_5} b) = -\bar{a}\gamma_5 b = -f\,. \tag{3.140}$$

5. **Pseudovector (or 'axial vector')**

$$f^\mu = \bar{a}\gamma^\mu\gamma_5 b\ : \quad \text{transforms as} \quad f'^\mu = \Lambda^\mu{}_\nu f^\nu\ (S)\,, \quad f'^\mu = -f_\mu\ (S_P) \tag{3.141}$$

Namely, a pseudovector transforms just like a vector under proper and orthochronous transformations, and under space inversion it transforms with signs opposite to those of a vector; i.e. the time component changes sign and the space components stay the same.

Under a proper and orthochronous transformation S, we have indeed

$$\begin{aligned} f'^\mu &= (\bar{a}\bar{S})\gamma^\mu\,\underbrace{\gamma_5(S\,b)}_{S\gamma_5} = \bar{a}\,\underbrace{\bar{S}\gamma^\mu S}_{\Lambda^\mu{}_\nu\gamma^\nu}\,\gamma_5 b \\ &= \Lambda^\mu{}_\nu(\bar{a}\gamma^\nu\gamma_5 b) = \Lambda^\mu{}_\nu f^\nu\,, \end{aligned} \tag{3.142}$$

as stated above, and under the space inversion S_P,

$$f' = (\bar{a}\bar{S}_P)\gamma^\mu\gamma_5(S_P b) = (\bar{a}\gamma^0)\,\underbrace{\gamma^\mu\gamma_5(\gamma^0}_{-\gamma^0\gamma_\mu\gamma_5}\,b) = -\bar{a}\gamma_\mu\gamma_5 b = -f_\mu\,, \tag{3.143}$$

where the index μ has changed from a superscript to a subscript when γ^0 is moved over γ^μ. The key points are that γ_5 commutes with a proper and orthochronous transformation S, thus the transformation under S is the same as a vector, and since γ_5 anticommutes with $S_P = \gamma^0$, the changes of sign under S_P are opposite to those of a vector.

And this is all. The five types of bilinear covariants are summarized in Table 3.2. For a given pair of spinors a and b, here are 16 such quantities formed by 16 matrices

Type	Γ_i	# of Γ_i	under S	under S_P
(S) Scalar	I	1	$f' = f$	$f' = f$
(V) Vector	γ^μ	4	$f'^\mu = \Lambda^\mu{}_\nu f^\nu$	$f'^\mu = f_\mu$
(T) Tensor	$\sigma^{\mu\nu}$	6	$f'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta f^{\alpha\beta}$	$f'^{\mu\nu} = f_{\mu\nu}$
(A) Axial vector	$\gamma^\mu\gamma_5$	4	$f'^\mu = \Lambda^\mu{}_\nu f^\nu$	$f'^\mu = -f_\mu$
(P) Pseudoscalar	γ_5	1	$f' = f$	$f' = -f$

Table 3.2: Bilinear covariants $\bar{a}\Gamma_i b$. S is a proper and orthochronous transformation, and S_P is the space inversion.

Γ_i ($i = 1, \ldots, 16$). It is straightforward to show that these matrices are independent and complete; namely, a bilinear quantity formed by an arbitrary complex 4×4 matrix $\bar{a}Mb$ can be written uniquely in terms of these bilinear covariants with a set of complex coefficients c_i:

$$\bar{a}Mb = \sum_{i=1}^{16} c_i \, \bar{a}\Gamma_i b \,, \tag{3.144}$$

and thus we know exactly how it transforms under Lorentz transformations (or more precisely, under proper and orthochronous transformations and space inversion).

Exercise 3.6 *Linear independence of bilinear covariants.*
Follow the steps below to prove that the 16 4×4 matrixes Γ_i of Table 3.2 are linearly independent. Use only representation-independent relations. All you need should be $\{\gamma^\mu, \gamma^\nu\} = 2\, g^{\mu\nu}$ and $\{\gamma^\mu, \gamma_5\} = 0$.
(a) Show that $(\Gamma_i)^2 = +1$ or -1 for $i = 1 \ldots 16$.
(b) Verify that for any Γ_i ($i \neq 1$) there is at least one Γ_k that anticommutes with it:

$$\Gamma_i \Gamma_k = -\Gamma_k \Gamma_i \,. \tag{3.145}$$

Using this, show that for $i \neq 1$, $\mathrm{Tr}\Gamma_i = 0$. (hint: express $\mathrm{Tr}(\Gamma_k \Gamma_i \Gamma_k)$ in two ways, one using $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, and the other using the above anticommutation relation.)
(c) Show that, for any pair $i \neq j$, the product of the two Γ's is just another Γ_k ($k \neq 1$) times a constant:

$$\Gamma_i \Gamma_j = c \Gamma_k \quad (i \neq j, k \neq 1) \tag{3.146}$$

(d) Suppose a set of constants a_i ($i = 1 \ldots 16$) exists to satisfy

$$\sum_{i=1}^{16} a_i \Gamma_i = 0 \,. \tag{3.147}$$

Take the trace of this equation to show $a_1 = 0$. Multiply by Γ_k $(k \neq 1)$ and then take the trace to show that $a_k = 0$. Thus all a's are zero. This completes the proof of independence.

You may be wondering what happened to the *pseudotensor* which would be $f^{\mu\nu} = \bar{a}\sigma^{\mu\nu}\gamma_5 b$. It actually has the 'desired' transformation properties: $f'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta f^{\alpha\beta}$ and $f'^{\mu\nu} = -f_{\mu\nu}$ as can be readily verified, but we have already exhausted all 16 independent $\Gamma' s$, and $\sigma^{\mu\nu}\gamma_5$ is already covered by the ones we have listed: in fact, it is easily shown that

$$\sigma^{0i}\gamma_5 = i\,\sigma^{jk}, \quad \sigma^{ij}\gamma_5 = i\,\sigma^{k0}, \quad (i,j,k : \text{cyclic})\,. \tag{3.148}$$

Namely, a pseudotensor quantity can be constructed out of a pair of spinors, but it is just a rearrangement of the components of the tensor bilinear covariants.

What is the importance of the bilinear covariants? We have already encountered the conserved current $j^\mu = \bar{\psi}\gamma^\mu\psi$ which is a bilinear covariant; the usefulness of the bilinear covariants, however, goes far beyond the probability current. In general, when we study an interaction of two spin-1/2 particles 1 and 2 creating another particle 3, we consider the *overlap* of the waves $\psi_1(x)$ and $\psi_2(x)$ of particles 1 and 2 acting as a *source* of the wave of particle 3. The stronger the overlap, the more intense the source is. But what do we mean by the overlap of ψ_1 and ψ_2 which have four components each? The classification of the bilinear covariants tells us that there are only five ways to define the overlap. And the transformation properties of the overlap should be consistent with the transformation properties of the particle created; for example, if a particle is created by a vector bilinear covariant, then the particle has to be represented by a vector field, etc.

Later on when we form Lagrangians of particles and interactions, we will see that spin-1/2 fields always appear in pairs as bilinear covariants. This is because in order to form a Lagrangian, which is a Lorentz scalar, 4-component spinors have to be first combined into bilinear covariants which have definite transformation properties which can then be combined to form a scalar quantity.

3.5 Representations of the γ matrices

We mentioned earlier that the Dirac representation of the γ matrices is not the only explicit expression that satisfies the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = 2\,g^{\mu\nu}$. An important theorem in this regard is the *Pauli's fundamental theorem* which states:

If two sets of 4×4 matrices γ^μ and γ'^μ satisfy

$$\{\gamma^\mu, \gamma^\nu\} = 2\, g^{\mu\nu}, \quad \{\gamma'^\mu, \gamma'^\nu\} = 2\, g^{\mu\nu}\,, \tag{3.149}$$

then there exists a matrix V such that

$$\gamma'^\mu = V \gamma^\mu V^{-1} \tag{3.150}$$

and V is unique up to a multiplicative constant.

Furthermore, if γ^μ and γ'^μ are to satisfy the property $\overline{\gamma^\mu} = \gamma^\mu$ and $\overline{\gamma'^\mu} = \gamma'^\mu$, then the matrix V (which exists by the abobe theorem) can be shown to be unitary. Noting that $\overline{\gamma^\mu} = \gamma^\mu$ means $\gamma^{\mu\dagger} = \gamma_\mu$ and using the unitarity condition $V^\dagger = V^{-1}$, the relation $\gamma^{\mu\dagger} = \gamma_\mu$ leads to

$$\gamma'^{\mu\dagger} = (V\gamma^\mu V^\dagger)^\dagger = V\gamma^{\mu\dagger} V^\dagger = V\gamma_\mu V^\dagger = \gamma'_\mu\,. \tag{3.151}$$

The converse can easily be proven using the completeness of the set Γ_i ($i = 1, \cdots, 16$) that we have seen in (3.2).

Suppose a spinor wave function $\psi(x)$ satisfies the Dirac equation $(i\,\partial\!\!\!/ - m)\psi = 0$, then

$$\begin{aligned}
& (i\overbrace{\gamma^\mu}^{V}\partial_\mu - m)\overbrace{\psi}^{V^{-1}V} \;=\; 0 \\
\to\quad & (i\underbrace{V\gamma^\mu V^{-1}}_{\gamma'^\mu}\partial_\mu - m)V\psi \;=\; 0 \\
\to\quad & (i\gamma'^\mu\partial_\mu - m)\psi' \;=\; 0 \quad \text{with} \quad \boxed{\psi' = V\psi}\,;
\end{aligned} \tag{3.152}$$

namely, the new wave function $V\psi(x)$ satisfies the Dirac equation with the new gamma matrices given by $\gamma'^\mu = V\gamma^\mu V^{-1}$. Note that the two wave functions $\psi(x)$ and $V\psi(x)$ represent exactly the same physical state (same $E, \vec{P}$, spin etc.), but simply given in different representation where the meaning of the spinor indexes are modified (which indexes correspond to spin-up, down etc.). This is in contrast to the Lorentz transformation where the physical parameters of the particle such as E, $\vec{P}$ and spin are transformed to different values.

Other representations often used are the Weyl representation or the 'chiral representation':

$$\gamma^0 = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{3.153}$$

which decouples two spin states (left and right handed spins) for massless particles, and the Majorana representation, which is entirely imaginary:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix},$$

$$\gamma^5 = \begin{pmatrix} \sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}. \tag{3.154}$$

The matrix γ_5 is always defined as $\gamma_5 \equiv i\,\gamma^0\gamma^1\gamma^2\gamma^3$.

Exercise 3.7 *Verify that γ matrixes in Weyl and Majorana representations indeed satisfy the anticommutation relations (3.37). Check for all combinations of (μ, ν), but do so systematically.*

3.6 Spin of the electron

Review of spin-1/2 formalism

Let's review the spin-1/2 formalism in non-relativistic quantum mechanics. We define $|\uparrow\rangle$ to be the state with spin in the $+z$ direction, and $|\downarrow\rangle$ to be the state with spin in the $-z$ direction. In terms of two-component column vectors, they can be written as

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.155}$$

In the space spanned by $|\uparrow\rangle$ and $|\downarrow\rangle$, the angular momentum operator $\vec{J}$ is represented by the Pauli matrices:

$$J_i = \frac{\sigma_i}{2} \quad (i = 1, 2, 3 \text{ or } x, y, z) \tag{3.156}$$

and they satisfy the commutation relations of angular momentum operators $[J_i, J_j] = i\,\epsilon_{ijk} J_k$ as shown earlier in (3.26). In the space spanned by states with a given angular momentum j, the square of the angular momentum operator $\vec{J}^2$ is given by

$$\vec{J}^2 \equiv J_1^2 + J_2^2 + J_3^2 = j(j+1). \tag{3.157}$$

Using $\sigma_i^2 = 1$ (3.21), we have

$$\left(\frac{\vec{\sigma}}{2}\right)^2 = \left(\frac{\sigma_1}{2}\right)^2 + \left(\frac{\sigma_2}{2}\right)^2 + \left(\frac{\sigma_3}{2}\right)^2 = \frac{3}{4} = \frac{1}{2}\left(\frac{1}{2} + 1\right) \tag{3.158}$$

which shows that $\vec{\sigma}/2$ is indeed a spin-1/2 representation of angular momentum.

The eigenstates with spin polarized along an arbitrary unit vector $\vec{s}$ can be obtained by applying a rotation to the states $|\uparrow\rangle$ and $|\downarrow\rangle$ of (3.155). In general, the rotation operator $R(\vec{\theta})$ by an angle $\theta \equiv |\vec{\theta}|$ around the axis $\hat{\theta} \equiv \vec{\theta}/\theta$ is generated by the angular momentum operators as

$$R(\vec{\theta}) = e^{-i\,\vec{\theta}\cdot\vec{J}}\,. \tag{3.159}$$

For rotations in the 3-dimensional space, the generators L_i $(i = 1,2,3)$ we have obtained in (1.100) can be redefined to be consistent with this form:

$$J_i \overset{\text{def}}{\equiv} iL_i\ (i = 1,2,3) \quad \rightarrow \quad e^{\theta_i L_i} = e^{-i\theta_i J_i}\,, \tag{3.160}$$

then the commutation relation (1.103) of L_i's indeed leads to that of angular momentum:

$$[L_i\,, L_j\,] = \epsilon_{ijk}L_k \quad \rightarrow \quad [J_i\,, J_j\,] = i\,\epsilon_{ijk}J_k\,. \tag{3.161}$$

In the spin-1/2 space, the corresponding rotation $u(\vec{\theta})$ is a 2×2 matrix obtained by the replacement $\vec{J} \rightarrow \vec{\sigma}/2$ in (3.159):

$$u(\vec{\theta}) = e^{-i\,\frac{\vec{\theta}}{2}\cdot\vec{\sigma}} = \cos\frac{\theta}{2} - i\,(\hat{\theta}\cdot\vec{\sigma})\sin\frac{\theta}{2} \tag{3.162}$$

where we have used the identity

$$e^{i\,\vec{a}\cdot\vec{\sigma}} = \cos a + i\,(\hat{a}\cdot\vec{\sigma})\sin a \qquad \left(a \equiv |\vec{a}|,\ \hat{a} \equiv \frac{\vec{a}}{a}\right). \tag{3.163}$$

The rotation matrix (3.162) tells us that a rotation around any axis by 2π changes sign of the state $(u = -1)$, and it takes two complete rotations to recover the original state, which is a general feature of half-integer spin states.

The z-axis can be rotated to the direction of $\vec{s}$ by first rotating around the y-axis by θ and then around the original z-axis (not the rotated one) by ϕ, where (θ,ϕ) are the polar angles of the direction $\vec{s}$ (Figure 3.1):

$$\vec{s} = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)\,. \tag{3.164}$$

The rotation matrix is then

$$R(\theta,\phi) = e^{-i\phi J_z}e^{-i\theta J_y} \quad \rightarrow \quad u(\theta,\phi) = e^{-i\frac{\phi}{2}\sigma_z}e^{-i\frac{\theta}{2}\sigma_y}\,, \tag{3.165}$$

and the resulting eigenvectors are (up to a common overall phase)

$$\begin{aligned}
\chi_+ &= u(\theta,\phi)\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2(1+s_z)}}\begin{pmatrix}1+s_z\\ s_+\end{pmatrix}\\
\chi_- &= u(\theta,\phi)\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{\sqrt{2(1-s_z)}}\begin{pmatrix}s_z-1\\ s_+\end{pmatrix}\,,
\end{aligned} \tag{3.166}$$

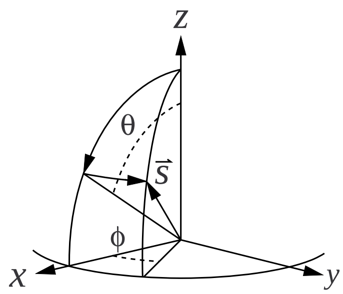

Figure 3.1: The rotation to take the z-axis to a general direction $\vec{s}$.

where

$$s_{\pm} \stackrel{\text{def}}{\equiv} s_x \pm i s_y \tag{3.167}$$

and the orthonormality is given by

$$\chi_+^\dagger \chi_+ = \chi_-^\dagger \chi_- = 1\,, \quad \chi_+^\dagger \chi_- = \chi_-^\dagger \chi_+ = 0\,. \tag{3.168}$$

On the other hand, the spin component in an arbitrary direction $\vec{s}$ is represented by the operator

$$\begin{aligned} \vec{s}\cdot\vec{\sigma} &= s_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + s_y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + s_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} s_z & s_x - i s_y \\ s_x + i s_y & -s_z \end{pmatrix} \\ &= \begin{pmatrix} s_z & s_- \\ s_+ & -s_z \end{pmatrix}\,. \end{aligned} \tag{3.169}$$

Using the formula

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = \vec{a}\cdot\vec{b} + i\,\vec{\sigma}\cdot(\vec{a}\times\vec{b}) \tag{3.170}$$

we have

$$(\vec{s}\cdot\vec{\sigma})^2 = \vec{s}^{\,2} + i\,\vec{\sigma}\cdot(\underbrace{\vec{s}\times\vec{s}}_{0}) = 1\,, \tag{3.171}$$

which means that the eigenvalues of $\vec{s}\cdot\vec{\sigma}$ is ± 1. In fact, it is easily verified that the states obtained in (3.166) are indeed eigenvectors of $\vec{s}\cdot\vec{\sigma}$:

$$(\vec{s}\cdot\vec{\sigma})\chi_{\pm} = \pm\chi_{\pm}\,. \tag{3.172}$$

Using the fact that $\vec{s}\cdot\vec{\sigma}$ has the eigenvalues ± 1, we can construct projection operators which project out $\chi_{\pm}$ from any vector. Define the operators $P_{\pm}$ by

$$P_{\pm} \stackrel{\text{def}}{\equiv} \frac{1 \pm \vec{s}\cdot\vec{\sigma}}{2}\,. \tag{3.173}$$

Then using $(\vec{s}\cdot\vec{\sigma})\chi_\pm = \pm\chi_\pm$

$$\begin{aligned} P_\pm\chi_\pm &= \frac{1\pm\vec{s}\cdot\vec{\sigma}}{2}\chi_\pm = \frac{1\pm(\pm1)}{2}\chi_\pm = \chi_\pm \\ P_\pm\chi_\mp &= \frac{1\pm\vec{s}\cdot\vec{\sigma}}{2}\chi_\mp = \frac{1\pm(\mp1)}{2}\chi_\mp = 0\,. \end{aligned} \tag{3.174}$$

Writing any vector v as a linear combination of $\chi_\pm$

$$v = c_+\chi_+ + c_-\chi_- \qquad (c_\pm : \text{complex constants}) \tag{3.175}$$

we see that $P_\pm$ indeed projects out $\chi_\pm$ out of v:

$$P_+v = c_+\chi_+, \quad P_-v = c_-\chi_-\,. \tag{3.176}$$

One can also easily verify that the operators $P_\pm$ satisfy the properties of projection operators:

$$P_\pm^2 = P_\pm, \quad P_+P_- = P_-P_+ = 0\,. \tag{3.177}$$

Exercise 3.8 *Non-relativistic spin-1/2 states.*
(a) Show that $u(\theta,\phi)$ can be written as

$$u(\theta,\phi) = e^{-i\frac{\phi}{2}}\begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} & \cos\frac{\theta}{2}e^{i\phi} \end{pmatrix}, \tag{3.178}$$

and verify (3.166). (hint: Use $e^{i\,\vec{a}\cdot\vec{\sigma}} = \cos a + i\,(\hat{a}\cdot\vec{\sigma})\sin a$.)
(b) For any 2-component vector v, $P_\pm(\vec{s})v$ are eigenvectors of $\vec{s}\cdot\vec{\sigma}$ with eigenvalues ±1 assuming that they are not zero. One can use this feature to construct the eigenvectors of $\vec{s}\cdot\vec{\sigma}$. Take any state, say the state with spin in $+z$ direction, as the vector v to obtain the same result for $\chi_\pm$ as above.

Electron spin

From the spinor-space representation of the proper and orthochronous Lorentz group (3.96), a pure rotation by $\vec{\theta}$, $U(\vec{\theta})$, is obtained by extracting the terms that correspond to the generators of rotation $M^{ij} = L_k$ (i,j,k: cyclic):

$$U(\vec{\theta}) = \exp\left(\frac{1}{2}\sum_{i,j}a_{ij}B^{ij}\right) = \exp\left(\sum_{i<j}a_{ij}B^{ij}\right) = e^{-i\theta_i\frac{\Sigma_i}{2}} \tag{3.179}$$

with

$$\frac{\Sigma_i}{2} \stackrel{\text{def}}{\equiv} iB^{jk}\,, \quad \theta_i = a_{jk} \quad (i,j,k : \text{cyclic})\,, \tag{3.180}$$

or using the definition of $\sigma^{\mu\nu}$ (3.127),

$$\boxed{\Sigma_i \stackrel{\text{def}}{\equiv} \sigma^{jk} = i\gamma^j\gamma^k \quad (i,j,k:\text{cyclic})}\,. \tag{3.181}$$

Comparing (3.179) with the general form of a rotation (3.159), we identify a set of operators that is acting as angular momentum operators:

$$J_i = \frac{\Sigma_i}{2}\,. \tag{3.182}$$

Since B^{jk} and M^{jk} satisfy the same commutation relations, $iL_i = iM^{jk}$ and $\Sigma_i/2 = iB^{jk}$ should satisfy the same commutation relations. On the other hand, we have seen in (3.161) that iL_i $(i = 1,2,3)$ satisfy the commutation relations of angular momentum, and thus so do $\frac{\Sigma_i}{2}$ $(i = 1,2,3)$. In fact, we will show below that $\vec{\Sigma}_i$ satisfies exactly the same commutation and anticommutation relations as the Pauli matrices $\vec{\sigma}$. First, the square of Σ_i is unity:

$$\begin{aligned}
(\Sigma_i)^2 &= (i\gamma^j\gamma^k)^2 = -\gamma^j \underbrace{\gamma^k\gamma^j}_{-\gamma^j\gamma^k \text{ since } j\neq k} \gamma^k \\
&= (\gamma^j)^2(\gamma^k)^2 = 1
\end{aligned} \tag{3.183}$$

where i,j,k are cyclic and no summation over repeated indexes is implied. Also, we have

$$\Sigma_1\Sigma_2 = (i\gamma^2 \underbrace{\gamma^3)(i\gamma^3}_{-i} \gamma^1) = -\gamma^1\gamma^2 = i\Sigma_3\,. \tag{3.184}$$

Similarly, we obtain $\Sigma_2\Sigma_3 = i\Sigma_1$ and $\Sigma_3\Sigma_1 = i\Sigma_2$, and thus

$$\Sigma_i\Sigma_j = i\Sigma_k \quad (i,j,k:\text{cyclic})\,. \tag{3.185}$$

Multiplying Σ_j on the left of $\Sigma_j\Sigma_k = i\Sigma_i$ (which is equivalent to the above),

$$\underbrace{(\Sigma_j)^2}_{1}\Sigma_k = i\Sigma_j\Sigma_i \quad \to \quad \Sigma_j\Sigma_i = -i\Sigma_k \quad (i,j,k:\text{cyclic})\,. \tag{3.186}$$

Comparing this with (3.185) shows that Σ_i's anticommute. These relations are summarized as

$$\boxed{\{\Sigma_i\,,\Sigma_j\} = 2\,\delta_{ij}\,, \quad [\Sigma_i\,,\Sigma_j] = 2i\epsilon_{ijk}\Sigma_k}\,. \tag{3.187}$$

In the Dirac representation Σ_i is explicitly written as

$$\begin{aligned}
\Sigma_i &= i\,\gamma^j\gamma^k = i\begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} = i\begin{pmatrix} -\sigma_j\sigma_k & 0 \\ 0 & -\sigma_j\sigma_k \end{pmatrix} \\
&= i\begin{pmatrix} -i\sigma_i & 0 \\ 0 & -i\sigma_i \end{pmatrix} = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}\,,
\end{aligned} \tag{3.188}$$

which explicitly shows that Σ_i's satisfy the same relations among themselves as do σ_i's. This is, however, independent of representation as we have seen.

To see the spin structure more clearly, apply the rotation operator U (3.179) to a plane wave

$$\psi(x) = \psi_0 e^{-i\,p\cdot x} \tag{3.189}$$

where ψ_0 is a constant spinor and $p^\mu = (E, \vec{P})$ represents the energy and momentum of the state. Using $x = \Lambda^{-1}x'$, or

$$x_\mu = (\Lambda^{-1})_\mu{}^\nu x'_\nu = \Lambda^\nu{}_\mu x'_\nu\,, \tag{3.190}$$

where Λ is the 4×4 Lorentz transformation matrix corresponding to the rotation, the transformed wave function $\psi'(x')$ can be written as

$$\begin{aligned}\psi'(x') &= U\psi_0 \exp(-i\,p^\mu x_\mu) = (U\psi_0)\exp(-i\,\overbrace{p^\mu \Lambda^\nu{}_\mu}^{p'^\nu}\, x'_\nu) \\ &= (U\psi_0)e^{-ip'\cdot x'}\,,\end{aligned} \tag{3.191}$$

where

$$p'^\nu = \Lambda^\nu{}_\mu p^\mu \tag{3.192}$$

is nothing but the rotated energy-momentum 4-vector. If it were a scalar field, then this rotation of the energy-momentum 4-vector (actually, the energy stays the same) would have been the only change. For a spinor field, however, we have additional 'rotation' $(U\psi_0)$ associated with some internal structure of the particle, and we know that this internal rotation is exactly like that of a spin-1/2 particle. Thus, we suspect that a particle represented by a spinor field carries an intrinsic spin 1/2.

To obtain a more physical understanding, however, we will now show that the angular momentum is conserved only when the spin is added to the orbital angular momentum. We start from the Dirac equation written as (3.29):

$$i\frac{\partial}{\partial t}\psi = H\psi\,, \quad H = \vec{\alpha}\cdot\vec{p} + \beta m \tag{3.193}$$

with

$$\beta = \gamma^0, \quad \alpha_i = \gamma^0\gamma^i \;(\leftarrow\; \gamma^i \equiv \beta\alpha_i), \quad \vec{p} \equiv -i\vec{\nabla}\,. \tag{3.194}$$

Recall that if an operator O commutes with the Hamiltonian H, then it is a constant of motion. Does it apply to the operator H we have here? Actually it does: suppose states $|a\rangle$ and $|b\rangle$ are solutions of a Schrödinger-form equation $i\frac{\partial}{\partial t}|\;\rangle = H|\;\rangle$. Then we have

$$i|\dot{a}\rangle = H|a\rangle \quad \rightarrow \quad -i\langle\dot{a}| = \langle a|H^\dagger = \langle a|H, \tag{3.195}$$

The time derivative of the matrix element $\langle a|O|b\rangle$ is then

$$\begin{aligned}\frac{\partial}{\partial t}\langle a|O|b\rangle &= \underbrace{\langle\dot a|}_{i\langle a|H}O|b\rangle + \langle a|O\underbrace{|\dot b\rangle}_{-iH|b\rangle} \\ &= i\langle a|HO|b\rangle - i\langle a|OH|b\rangle \\ &= i\langle a|[H,O]|b\rangle\,; \end{aligned}\tag{3.196}$$

namely, the matrix element $\langle a|O|b\rangle$ is a constant of motion if $[H,O]=0$. This holds as long as the equation is in the Schrödinger form and the 'Hamiltonian' is hermitian.

We first evaluate the commutator of the orbital angular momentum and the Hamiltonian, $[L_i,H]$, with

$$L_i = (\vec x\times\vec p)_i = \epsilon_{ijk}\,x^j p^k \tag{3.197}$$

$$H = \underbrace{\alpha_i}_{\gamma^0\gamma^i} p^i + \underbrace{\beta}_{\gamma^0} m = \gamma^0(\gamma^i p^i + m)\,, \tag{3.198}$$

where sum over repeated space indexes are implicit regardless of superscript or subscript, and we are consistently using the definition $x^\mu = (x^0,\vec x)$ and $p^\mu = (p^0,\vec p)$. Useful relations are

$$[AB,C] = A[B,C] + [A,C]B\,, \qquad [C,AB] = A[C,B] + [C,A]B\,. \tag{3.199}$$

It follows that, in either of the formulas, if A commutes with C, A can simply come out of the commutator to the left, and if B commutes with C, B can simply come out to the right.

Noting that γ^μ commutes with x^i and p^i, and that $[x^i\,,p^j] = i\delta_{ij}$, we have

$$\begin{aligned}[L_i,H] &= [\epsilon_{ijk}\,x^j p^k\,,\gamma^0(\gamma^l p^l + m)] \\ &= \epsilon_{ijk}\gamma^0[x^j p^k\,,\gamma^l p^l + \cancel{m}\,] \\ &= \epsilon_{ijk}\gamma^0\gamma^l\,\underbrace{[x^j p^k\,,p^l]}_{\underbrace{[x^j,p^l]}_{i\delta_{jl}}p^k} = i\,\epsilon_{ijk}\underbrace{\gamma^0\gamma^j}_{\alpha_j}p^k \\ &= i(\vec\alpha\times\vec p)_i \end{aligned}\tag{3.200}$$

This looks non-zero and cannot be simplified further, indicating that the orbital angular momentum is not a constant of motion by itself.

We now evaluate $[\Sigma_i\,,H]$ using $\Sigma_i = i\gamma^j\gamma^k$, $(i,j,k$: cyclic):

$$[\Sigma_i\,,H] = [i\gamma^j\gamma^k,\gamma^0(\gamma^l p^l + m)]$$

$$
\begin{aligned}
&= i\gamma^0 \underbrace{[\gamma^j\gamma^k, \gamma^l p^l + \cancel{m}\,]}_{\underbrace{[\gamma^j\gamma^k, \gamma^l]}_{(3.89) \to \; -2(g^{jl}g^k{}_\nu - g^{kl}g^j{}_\nu)\gamma^\nu = 2(\delta_{jl}\delta_{k\nu} - \delta_{kl}\delta_{j\nu})\gamma^\nu} p^l} \\
&= 2i\gamma^0(\gamma^k p^j - \gamma^j p^k) = -2i(\alpha_j p^k - \alpha_k p^j) \\
&= -2i\,(\vec{\alpha} \times \vec{p})_i
\end{aligned} \tag{3.201}
$$

Combining (3.200) and (3.201), we have a conserved quantity:

$$
[\vec{J}, H] = 0\,, \quad \text{with} \quad \vec{J} = \vec{L} + \frac{\vec{\Sigma}}{2}\,, \tag{3.202}
$$

which shows that when the spin operator $\vec{\Sigma}/2$ is added to the orbital angular momentum, the total angular momentum is conserved, indicating that electron does carry an intrinsic spin angular momentum whose absolute value is 1/2. Note also that the above equation clearly identifies Σ_i to be the i-th component of the spin which has to be added to the i-th component of $\vec{L}$; for example, $\Sigma_3 = i\gamma^1\gamma^2$ always represents the z-component of the spin regardless of the representation of the γ matrices.

Let's reflect upon how this descritpion of electron spin has come about. We set out to solve the problems of the Klein-Gordon equation, in particular its negative probability, and in making the equation linear in time and space derivatives in hope of solving it, the equation became 4-component, and it introduced the possibility that the particle has some internal structure. The spin of the electron, or more correctly the magnetic moment of electron, had been 'discovered' by G. Uhlenbeck and S. Goudsmit three years earlier in 1925 in order to explain the anomalous Zeeman effect (splitting of energy levels in alkali atoms in presence of electric and magnetic fields). The electron spin was now beautifully explained by a theory which was consistent with the special relativity and the quantum mechanics. Actually, we have not shown that electron has a magnetic moment yet, which involves interactions with the electromagnetic field $A^\mu(x)$. This will be done later, and it will be shown that the value of the electron magnetic moment is twice as large as the natural value classically expected from its charge and spin. Now, let's move on to finding explicit plane-wave solutions of the Dirac equation.

3.7 Plane-wave solutions of the Dirac equation

In this section, we will construct plane-wave solutions of the Dirac equation. Even though we will use the Dirac representation for explicit expressions, many of the essential relations are independent of the representation as will be shown. Let's start from the solutions for a particle at rest.

Electron at rest ($\vec{p} = 0$)

From the correspondence $\vec{p} \leftrightarrow -i\vec{\nabla}$, the solution for $\vec{p} = 0$ should have no space dependence, or $\partial_i \psi = 0$, $(i = 1, 2, 3)$. Then the Dirac equation $i\gamma^\mu \partial_\mu \psi = m\psi$ becomes

$$i\gamma^0 \partial_0 \psi = m\psi\,, \quad \text{or} \quad i \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} \dot{\psi}_1 \\ \dot{\psi}_2 \\ \dot{\psi}_3 \\ \dot{\psi}_4 \end{pmatrix} = m \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} . \tag{3.203}$$

The four equations are

$$\begin{cases} i\dot{\psi}_1 &= m\psi_1 \\ i\dot{\psi}_2 &= m\psi_2 \\ i\dot{\psi}_3 &= -m\psi_3 \\ i\dot{\psi}_4 &= -m\psi_4 \end{cases} \tag{3.204}$$

There are thus four independent solutions:

$$\begin{gathered} \psi^{(1)} = \omega^{(1)} e^{-imt},\ \psi^{(2)} = \omega^{(2)} e^{-imt},\ \psi^{(3)} = \omega^{(3)} e^{+imt},\ \psi^{(4)} = \omega^{(4)} e^{+imt}, \\ \omega^{(1)} \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^{(2)} \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^{(3)} \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \omega^{(4)} \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} . \end{gathered} \tag{3.205}$$

Since we define the eigenvalue of the operator $i\partial/\partial t$ to be the energy, $\psi^{(1)}$ and $\psi^{(2)}$ are positive energy solutions and $\psi^{(3)}$ and $\psi^{(4)}$ are negative energy solutions.

That this is a complete set of solutions for $\vec{p} = 0$ can be seen as follows. At a given time, say $t = 0$, the most general form of ψ is $\psi_i(x)|_{t=0} = a_i$ where a_i $(i = 1, 2, 3, 4)$ are some complex constants (there is no space dependence since $\vec{p} = 0$). Then, in order for this wave function to be a solution of the Dirac equation, the time dependence of each component in the past and future is uniquely determined by (3.204); namely, the first and second components has to vary as e^{-imt} and the third and fourth as e^{+imt}:

$$\psi(x)|_{t=0} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} \quad \to \quad \psi(x) = \begin{pmatrix} a_1 e^{-imt} \\ a_2 e^{-imt} \\ a_3 e^{+imt} \\ a_4 e^{+imt} \end{pmatrix} , \tag{3.206}$$

which is thus a completely general $\vec{p} = 0$ solution of the Dirac equation. This can of course be written as a linear combination of the four solutions (3.205) which are thus complete:

$$\psi(x) = a_i \psi^{(i)}(x)\,, \tag{3.207}$$

where the spinor index i is summed over 1 through 4.

Leaving the discussion of the negative energy states to a later time when we describe the hole theory, let's look at how $\psi^{(i)}$'s respond to the spin z-component operator Σ_3, which, in the Dirac representation, is written as [see (3.188)]

$$\Sigma_3 = \begin{pmatrix} \sigma_3 & \\ & \sigma_3 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}. \tag{3.208}$$

Clearly, $\psi^{(i)}$'s are eigenvectors of the spin z-component with eigenvalues given by

$$\psi^{(1)} : \Sigma_3 = +1, \quad \psi^{(2)} : \Sigma_3 = -1, \quad \psi^{(3)} : \Sigma_3 = +1, \quad \psi^{(4)} : \Sigma_3 = -1\,. \tag{3.209}$$

Eigenstates of spin component in any arbitrary direction $\vec{s}$ can be obtained by rotating these solutions, and we know exactly how to rotate spinors. The operator to rotate the z-axis to an arbitrary direction $\vec{s}$ can be taken as in (3.165): $R(\theta,\phi) = e^{-i\phi J_z} e^{-i\theta J_y}$ where (θ,ϕ) are the polar angles of the direction $\vec{s}$. The corresponding rotation in the spinor space is given by the replacement $\vec{J} \to \vec{\Sigma}/2$ (3.182):

$$U(\theta,\phi) = e^{-i\frac{\phi}{2}\Sigma_z} e^{-i\frac{\theta}{2}\Sigma_y}\,, \tag{3.210}$$

which can be rewritten as follows: First, due to the block-diagonal form of (3.188), we have

$$\Sigma_i = \begin{pmatrix} \sigma_i & \\ & \sigma_i \end{pmatrix} \quad \to \quad (c\Sigma_i)^k = \begin{pmatrix} (c\,\sigma_i)^k & \\ & (c\,\sigma_i)^k \end{pmatrix}, \tag{3.211}$$

where c is any constant number. Then, the exponential can go inside the block-diagonal subsections:

$$\begin{aligned} e^{c\Sigma_i} &= \sum_k \frac{1}{k!}(c\Sigma_i)^k = \sum_k \frac{1}{k!} \begin{pmatrix} (c\,\sigma_i)^k & \\ & (c\,\sigma_i)^k \end{pmatrix} \\ &= \begin{pmatrix} e^{c\,\sigma_i} & \\ & e^{c\,\sigma_i} \end{pmatrix} \end{aligned} \tag{3.212}$$

Thus, the spinor operator $U(\theta,\phi)$ can be written as

$$\begin{aligned} U(\theta,\phi) &= e^{-i\frac{\phi}{2}\Sigma_z} e^{-i\frac{\theta}{2}\Sigma_y} = \begin{pmatrix} e^{-i\frac{\phi}{2}\sigma_z} & \\ & e^{-i\frac{\phi}{2}\sigma_z} \end{pmatrix} \begin{pmatrix} e^{-i\frac{\theta}{2}\sigma_y} & \\ & e^{-i\frac{\theta}{2}\sigma_y} \end{pmatrix} \\ &= \begin{pmatrix} e^{-i\frac{\phi}{2}\sigma_z} e^{-i\frac{\theta}{2}\sigma_y} & \\ & e^{-i\frac{\phi}{2}\sigma_z} e^{-i\frac{\theta}{2}\sigma_y} \end{pmatrix} \\ &= \begin{pmatrix} u(\theta,\phi) & \\ & u(\theta,\phi) \end{pmatrix} \end{aligned} \tag{3.213}$$

where $u(\theta,\phi)$ is the 2×2 matrix defined in (3.165). Applying this rotation to $\psi^{(1)}$,

$$\psi'^{(1)} = U(\theta,\phi)\psi^{(1)} = \begin{pmatrix} u(\theta,\phi)\begin{pmatrix}1\\0\end{pmatrix} \\ \begin{pmatrix}0\\0\end{pmatrix}\end{pmatrix} e^{-imt} = \begin{pmatrix}\chi_+\\0\end{pmatrix} e^{-imt}\,, \tag{3.214}$$

where χ_+ is defined by (3.166). Similarly, applying the rotation $U(\theta,\phi)$ to $\psi^{(i)}$ ($i = 2,3,4$), we obtain a set of four states polarized along a general direction $\vec{s}$ (and $\vec{p}=0$):

$$\begin{aligned}
\psi^{(1)}_{\vec{s}} &= N\begin{pmatrix}\chi_+\\0\end{pmatrix} e^{-imt}\ (\text{spin } +\vec{s}), \qquad & \psi^{(2)}_{\vec{s}} &= N\begin{pmatrix}\chi_-\\0\end{pmatrix} e^{-imt}\ (\text{spin } -\vec{s}),\\
\psi^{(3)}_{\vec{s}} &= N\begin{pmatrix}0\\\chi_+\end{pmatrix} e^{+imt}\ (\text{spin } +\vec{s}), \qquad & \psi^{(4)}_{\vec{s}} &= N\begin{pmatrix}0\\\chi_-\end{pmatrix} e^{+imt}\ (\text{spin } -\vec{s})\,,
\end{aligned} \tag{3.215}$$

where we have omitted the primes on ψ's, and N is a normalization factor to be chosen later. It can be easily checked that these are indeed eigenvectors of the operator $\vec{s}\cdot\vec{\Sigma}$, which should represent the component of spin along the direction $\vec{s}$, and that they have the stated eigenvalues: Using (3.188), we have

$$\vec{s}\cdot\vec{\Sigma} = \begin{pmatrix}\vec{s}\cdot\vec{\sigma} & \\ & \vec{s}\cdot\vec{\sigma}\end{pmatrix}. \tag{3.216}$$

Applying this to $\psi^{(1)}_{\vec{s}}$ of (3.215), for example,

$$(\vec{s}\cdot\vec{\Sigma})\psi^{(1)}_{\vec{s}} = \begin{pmatrix}(\vec{s}\cdot\vec{\sigma})\chi_+\\0\end{pmatrix} e^{-imt} = \begin{pmatrix}\chi_+\\0\end{pmatrix} e^{-imt} = +\psi^{(1)}_{\vec{s}}\,, \tag{3.217}$$

where we have used (3.172). Similarly, other states can be shown to be eigenvectors of $\vec{s}\cdot\vec{\Sigma}$; thus, we have

$$(\vec{s}\cdot\vec{\Sigma})\psi^{(1,3)}_{\vec{s}} = +\psi^{(1,3)}_{\vec{s}}\,,\quad (\vec{s}\cdot\vec{\Sigma})\psi^{(2,4)}_{\vec{s}} = -\psi^{(2,4)}_{\vec{s}}\,. \tag{3.218}$$

Electron in motion

Now that we have a set of solutions for an electron at rest with its spin polarized in an arbitrary direction, the next step is to construct solutions for an electron in motion. One could solve the Dirac equation explicitly without dropping the space derivatives, but we are now well equipped to deal with it more systematically: we know solutions at rest and we know exactly how to boost them.

The matrix S in the spinor space corresponding to a Lorentz transformation in space-time Λ which is a boost in a direction $\hat{\xi}$ by velocity $\beta = \tanh|\xi|$ is

$$\Lambda = e^{\xi_i K_i} = e^{\xi_i M^{0i}} \quad\rightarrow\quad S = e^{\xi_i B^{0i}}\,, \tag{3.219}$$

with

$$B^{0i} = \frac{1}{2}\gamma^0\gamma^i = \frac{1}{2}\alpha_i\,. \tag{3.220}$$

Writing $\vec{\xi} = \xi\hat{\xi}$ ($\xi \equiv |\vec{\xi}|$, $\hat{\xi} = \vec{\xi}/\xi$),

$$\begin{aligned}
(\hat{\xi}\cdot\vec{\alpha})^2 &= \hat{\xi}_i\alpha_i\hat{\xi}_j\alpha_j = \hat{\xi}_i\hat{\xi}_j\alpha_i\alpha_j \\
&= \frac{1}{2}(\hat{\xi}_i\hat{\xi}_j\,\overbrace{\alpha_i\alpha_j + \hat{\xi}_j\hat{\xi}_i}^{(i\leftrightarrow j)}\,\alpha_j\alpha_i) \\
&= \frac{1}{2}\hat{\xi}_i\hat{\xi}_j(\underbrace{\alpha_i\alpha_j + \alpha_j\alpha_i}_{2\delta_{ij}}) = |\hat{\xi}|^2 = 1\,.
\end{aligned} \tag{3.221}$$

The matrix S can then be written as

$$\begin{aligned}
S &= e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}} = e^{\frac{\xi}{2}(\hat{\xi}\cdot\vec{\alpha})} \\
&= 1 + \left(\frac{\xi}{2}\right)(\hat{\xi}\cdot\vec{\alpha}) + \frac{1}{2!}\left(\frac{\xi}{2}\right)^2\underbrace{(\hat{\xi}\cdot\vec{\alpha})^2}_{1} + \frac{1}{3!}\left(\frac{\xi}{2}\right)^3\underbrace{(\hat{\xi}\cdot\vec{\alpha})^3}_{(\hat{\xi}\cdot\vec{\alpha})} + \cdots \\
&= \left[1 + \frac{1}{2!}\left(\frac{\xi}{2}\right)^2 + \cdots\right] + \left[\left(\frac{\xi}{2}\right) + \frac{1}{3!}\left(\frac{\xi}{2}\right)^3 + \cdots\right](\hat{\xi}\cdot\vec{\alpha}) \\
&= \cosh\frac{\xi}{2} + (\hat{\xi}\cdot\vec{\alpha})\sinh\frac{\xi}{2}\,.
\end{aligned} \tag{3.222}$$

Using

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \quad\to\quad \hat{\xi}\cdot\vec{\alpha} = \begin{pmatrix} 0 & \hat{\xi}\cdot\vec{\sigma} \\ \hat{\xi}\cdot\vec{\sigma} & 0 \end{pmatrix} \tag{3.223}$$

we have

$$\begin{aligned}
S = e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}} &= \begin{pmatrix} \cosh\frac{\xi}{2} & (\hat{\xi}\cdot\vec{\sigma})\sinh\frac{\xi}{2} \\ (\hat{\xi}\cdot\vec{\sigma})\sinh\frac{\xi}{2} & \cosh\frac{\xi}{2} \end{pmatrix} \\
&= \cosh\frac{\xi}{2}\begin{pmatrix} 1 & (\hat{\xi}\cdot\vec{\sigma})\tanh\frac{\xi}{2} \\ (\hat{\xi}\cdot\vec{\sigma})\tanh\frac{\xi}{2} & 1 \end{pmatrix}\,.
\end{aligned} \tag{3.224}$$

It is more convenient to express it in terms of the energy and momentum of the particle. Recall that the boost parameter ξ is related to γ and η of the boost by

$$\gamma = \cosh\xi\,, \quad \eta = \sinh\xi\,, \tag{3.225}$$

then using the hyperbolic half-angle formulas,

$$\cosh\frac{\xi}{2} = \sqrt{\frac{\cosh\xi + 1}{2}} = \sqrt{\frac{\gamma+1}{2}} \tag{3.226}$$

$$\tanh\frac{\xi}{2} = \frac{\sinh\xi}{\cosh\xi + 1} = \frac{\eta}{\gamma+1}\,. \tag{3.227}$$

The energy-momentum 4-vector $p^\mu \equiv (E, \vec{p})$ acquired by a rest mass m after the boost is

$$\boxed{p^\mu = m\eta^\mu \quad \text{with} \quad \eta^\mu \stackrel{\text{def}}{\equiv} (\gamma, \vec{\eta})}. \tag{3.228}$$

where $\vec{\eta}$ is the vector with length η pointing to the direction of the boost. We then have $E = m\gamma$ and $p \equiv |\vec{p}| = m\eta$, and thus we can write

$$\cosh\frac{\xi}{2} = \sqrt{\frac{\gamma+1}{2}} = \sqrt{\frac{E+m}{2m}}, \quad \tanh\frac{\xi}{2} = \frac{\eta}{\gamma+1} = \frac{p}{E+m}. \tag{3.229}$$

Since $\vec{\xi}$ is in the direction of the boost as discussed just below (1.144), we have $\vec{p} = \hat{\xi}p$, and thus

$$(\hat{\xi}\cdot\vec{\sigma})\tanh\frac{\xi}{2} = (\hat{\xi}\cdot\vec{\sigma})\frac{p}{E+m} = \frac{\vec{p}\cdot\vec{\sigma}}{E+m}, \tag{3.230}$$

The boost matrix in the spinor space (3.224) can now be written as

$$\boxed{S = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 & \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m} \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m} & 1 \end{pmatrix}}, \tag{3.231}$$

which is a boost under which a rest mass m (> 0) acquires an energy-momentum $p^\mu = (E, \vec{p})$. Note that $p^0 \equiv E$ as defined is *always positive*. Using [see (3.169)]

$$\vec{p}\cdot\vec{\sigma} = \begin{pmatrix} p_z & p_- \\ p_+ & -p_z \end{pmatrix} \qquad (p_\pm = p_x \pm ip_y), \tag{3.232}$$

it could be written in a fully 4×4 form:

$$S = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 & 0 & \frac{p_z}{E+m} & \frac{p_-}{E+m} \\ 0 & 1 & \frac{p_+}{E+m} & \frac{-p_z}{E+m} \\ \frac{p_z}{E+m} & \frac{p_-}{E+m} & 1 & 0 \\ \frac{p_+}{E+m} & \frac{-p_z}{E+m} & 0 & 1 \end{pmatrix}. \tag{3.233}$$

All we need now is to apply S to the solutions at rest $\psi^{(i)}_{\vec{s}}$, and strictly follow the definition

$$\psi'(x') = S\psi(x), \quad x' = \Lambda x, \tag{3.234}$$

to write it as a function of x'. In doing so, we need to write the exponent imt in terms of x'. This can easily done by noting that

$$mt = p \cdot x, \qquad p = (m, \vec{0}). \tag{3.235}$$

Then we have the Lorentz invariance relation

$$p \cdot x = p' \cdot x', \qquad (x' = \Lambda x, \quad p' = \Lambda x'). \tag{3.236}$$

Applying S to the solution $\psi(x) = \psi^{(1)}_{\vec{s}}(x)$ with $\vec{p} = 0$ and spin $+\vec{s}$ (3.215), the boosted wave function, which we will denote as $\psi^{(1)}_{\vec{s},\vec{p}}(x')$, is then

$$\begin{aligned} \psi'(x') = S\psi^{(1)}_{\vec{s}}(x) &= N\sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m} \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m} & 1 \end{pmatrix} \begin{pmatrix} \chi_+ \\ 0 \end{pmatrix} e^{-i\,m\,t} \\ &= N\sqrt{\frac{E+m}{2m}} \begin{pmatrix} \chi_+ \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \end{pmatrix} e^{-i\,p'\cdot x'} \stackrel{\text{def}}{\equiv} \psi^{(1)}_{\vec{s},\vec{p}}(x') \,. \end{aligned} \tag{3.237}$$

At this point, we choose the normalization by $N = \sqrt{2m}$, which will avoid the divergence in the limit $m \to 0$. Similarly applying the same boost S to the rest of the states, we obtain (omitting the primes)

$$\begin{aligned} &\psi^{(1)}_{\vec{s},\vec{p}}(x) = \omega_{p,\vec{s}}\, e^{-ip\cdot x}\,, && \psi^{(2)}_{\vec{s},\vec{p}}(x) = \omega_{p,-\vec{s}}\, e^{-ip\cdot x}\,, \\ &\omega_{p,\vec{s}} = \sqrt{E+m} \begin{pmatrix} \chi_+ \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \end{pmatrix}, && \omega_{p,-\vec{s}} = \sqrt{E+m} \begin{pmatrix} \chi_- \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_- \end{pmatrix}, \\ &\psi^{(3)}_{\vec{s},\vec{p}}(x) = \omega_{-p,\vec{s}}\, e^{ip\cdot x}\,, && \psi^{(4)}_{\vec{s},\vec{p}}(x) = \omega_{-p,-\vec{s}}\, e^{ip\cdot x}\,, \\ &\omega_{-p,\vec{s}} = \sqrt{E+m} \begin{pmatrix} \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \\ \chi_+ \end{pmatrix}, && \omega_{-p,-\vec{s}} = \sqrt{E+m} \begin{pmatrix} \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_- \\ \chi_- \end{pmatrix}, \end{aligned} \tag{3.238}$$

with

$$p^\mu \equiv (E,\vec{p}) \equiv m\eta^\mu\,, \qquad \text{in particular,} \qquad p^0 \equiv E \geq 0\,, \tag{3.239}$$

where we have labeled the constant spinors ω by the spin direction and the eigenvalue of the operator $i\partial^\mu$ ($p \equiv p^\mu$). Thus, we have here a set of four solutions corresponding to a moving electron. The rule is that the constant spinor $\omega_{p,\pm\vec{s}}$ has to be attached to the space-time dependence $e^{-ip\cdot x}$ in order for the wave function to become a solution of the Dirac equation, where the parameters p^μ appearing in ω and the exponent have to be the same, and $\omega_{-p,\pm\vec{s}}$ has to be attached to $e^{ip\cdot x}$ to be a solution of the Dirac equation. Again, we emphasize that the 4-vector p^μ above is simply m times the boost 4-vector η^μ, and $p^0 \equiv E$ is always positive by definition. Thus, as long as the energy is defined by $i\partial_0$, the solutions $\psi^{(1,2)}_{\vec{s},\vec{p}}$ have a positive energy and $\psi^{(3,4)}_{\vec{s},\vec{p}}$ have a negative energy.

Conserved current of the plane wave solutions

Let's evaluate the conserved current $j^\mu = \bar{\psi}\gamma^\mu\psi$ for the rest frame solutions (3.215) with $N = \sqrt{2m}$. The time component, which should be the probability density, for the solution $\psi^{(1)}_{\vec{s}}$ is

$$\begin{aligned} j^0 &= \bar{\psi}^{(1)}_{\vec{s}} \gamma^0 \psi^{(1)}_{\vec{s}} = \psi^{(1)\dagger}_{\vec{s}} \psi^{(1)}_{\vec{s}} \\ &= 2m\, \chi^\dagger_+ \chi_+ = 2m\,, \end{aligned} \tag{3.240}$$

and the result is the same for all other solutions. The space component vanishes as expected for a particle at rest:

$$\begin{aligned} j^k &= \overline{\psi^{(1)}_{\vec{s}}} \gamma^k \psi^{(1)}_{\vec{s}} = \psi^{(1)\dagger}_{\vec{s}} \underbrace{\gamma^0 \gamma^k}_{\alpha_k} \psi^{(1)}_{\vec{s}} \\ &= 2m \,(\, \chi^\dagger_+ \quad 0\,) \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \begin{pmatrix} \chi_+ \\ 0 \end{pmatrix} = 0\,, \end{aligned} \tag{3.241}$$

which also vanishes for other solutions. Thus, we have for all the four solutions with $\vec{p} = 0$,

$$j^\mu = (2m, \vec{0}) \quad \text{for} \quad \psi^{(i)}_{\vec{s}} \quad (i = 1, 2, 3, 4)\,. \tag{3.242}$$

Since we know that j^μ transforms as a vector, we should have for the boosted solutions,

$$j^\mu = 2m\eta^\mu = 2p^\mu \tag{3.243}$$

where η^μ is the boost 4-velocity as before, and thus p^0 is positive by definition. This can be explicitly verified from $j^\mu = \bar{\psi}\gamma^\mu\psi$ and (3.238).

Our normalization of the wave function is such that for the solutions at rest, there are $2m$ particles per unit volume. When the states are boosted, we see that the probability density becomes denser by the contraction factor γ to become $2m\gamma = 2E$. These are all things we already know from general analysis. Here, however, we have verified them using explicit plane-wave solutions.

Negative energy solutions - the hole theory

We have defined the energy as the eigenvalue of the time derivative operator $i\partial^0$, and according to this definition, two of our plane-wave solutions (3.238) have negative energy. One cannot simply ignore the negative energy solutions, since when interactions are included in the theory, it becomes unavoidable to have transitions to negative energy states. As it turns out, such transitions cannot be excluded without violating the conservation of probability, and the true solution lies in the Quantum Field Theory. For now we will follow Dirac's argument to wiggle out of the problem. The main aim of this exercise is to assign proper quantum numbers to the negative energy solutions which happen to represent the antiparticle of the electron - the positron.

Consider an atom in its ground state with many electrons. Due to the Pauli's exclusion principle, another electron cannot fill the low-energy states which are already occupied. If one of the electrons is kicked out of the atom ('ionization'), it will leave a positively charged atom. Now imagine the vacuum as something like a gigantic atom where all the negative energy states are filled up. Then a positive energy electron cannot drop into one of the negative energy states since it is already filled up.

If a negative energy electron is 'excited' to a positive energy state, it will leave a 'hole'. If the original negative energy electron had 4-momentum $-p^\mu$ defined as the eigenvalues of $i\partial^\mu$ (with $p^0 \geq 0$) and spin $-\vec{s}$, then the hole would look like it has a 4-momentum p^μ and spin $\vec{s}$ relative to the vacuum. Also, the charge of the hole would look like the opposite of that of the electron. The mass of the hole then should be the same as that of the electron since $p^2 = (-p)^2 = m^2$. Thus, the hole has the same mass, same absolute spin, and opposite charge to those of the electron. Such particle is called the *antiparticle* of the electron, or 'positron' denoted as e^+, while electron is denoted as e^-. The correspondence between the missing negative energy electron and the resulting positron is summarized as

$$\begin{array}{ccc} \text{When missing } e^- \text{ of} & & \text{it is equivalent to } e^+ \text{ of} \\ \left\{ \begin{array}{rl} \text{energy} & -E \\ \text{momentum} & -\vec{p} \\ \text{spin} & -\vec{s}\,, \end{array} \right. & \begin{array}{c} \Longleftrightarrow \\ \text{all signs flipped} \end{array} & \left\{ \begin{array}{rl} \text{energy} & E \\ \text{momentum} & \vec{p} \\ \text{spin} & \vec{s}\,. \end{array} \right. \end{array} \tag{3.244}$$

Since in reality we deal with positrons with positive energies rather than electrons with negative energy, it is convenient to label the spinors in (3.238) by physical quantities. Using u for electron spinors and v for positron spinors, we define

$$\begin{array}{rclrcl} u_{\vec{p},\vec{s}} & \equiv & \omega_{p,\vec{s}}\,, \qquad & u_{\vec{p},-\vec{s}} & \equiv & \omega_{p,-\vec{s}}\,, \\ v_{\vec{p},-\vec{s}} & \equiv & \omega_{-p,\vec{s}}\,, \qquad & v_{\vec{p},\vec{s}} & \equiv & \omega_{-p,-\vec{s}}\,. \end{array} \tag{3.245}$$

Now for both u and v spinors, the labeling corresponds to physical quantities of electron and positron, respectively. Of course, in order to be a solution of the Dirac equation, the $u_{\vec{p},\pm\vec{s}}$ spinor has to be attached to $e^{-ip\cdot x}$ and the $v_{\vec{p},\pm\vec{s}}$ spinor has to be attached to $e^{ip\cdot x}$, where p^0 is given by $p^0 = \sqrt{\vec{p}^2 + m^2}$ which is always positive. Since p^0 is uniquely defined for a given $\vec{p}$, we have chosen to label the u, v spinors by $\vec{p}$ rather than by p^μ. The four plane-wave solutions are now

$$\boxed{\begin{array}{cc} u_{\vec{p},\pm\vec{s}}\, e^{-ip\cdot x}\,, & v_{\vec{p},\pm\vec{s}}\, e^{ip\cdot x}\,, \\[1ex] \begin{pmatrix} e^- \text{ with spin } \pm\vec{s} \\ \text{4-momentum } p^\mu \end{pmatrix} & \begin{pmatrix} e^+ \text{ with spin } \pm\vec{s} \\ \text{4-momentum } p^\mu \end{pmatrix} \end{array}} \tag{3.246}$$

where the u, v spinors are defined by (3.245) and (3.238); namely,

$$u_{\vec{p},\vec{s}} = \sqrt{E+m}\begin{pmatrix} \chi_+ \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \end{pmatrix}, \quad u_{\vec{p},-\vec{s}} = \sqrt{E+m}\begin{pmatrix} \chi_- \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_- \end{pmatrix},$$
$$v_{\vec{p},-\vec{s}} = \sqrt{E+m}\begin{pmatrix} \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \\ \chi_+ \end{pmatrix}, \quad v_{\vec{p},\vec{s}} = \sqrt{E+m}\begin{pmatrix} \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_- \\ \chi_- \end{pmatrix}. \tag{3.247}$$

3.8 Energy and spin projection operators

In the previous section we have used the Dirac representation. Almost all discussions in this section and the next, at least all the boxed formulas, are independent of representation.

Dirac equation in momentum space

Since $u_{\vec{p},\pm\vec{s}}\,e^{-ip\cdot x}$ and $v_{\vec{p},\pm\vec{s}}\,e^{ip\cdot x}$ satisfy the Dirac equation, we can extract the matrix equations for the u and v spinors:

$$\begin{aligned} 0 &= (i\gamma^\mu\partial_\mu - m)u_{\vec{p},\pm\vec{s}}\,e^{-ip\cdot x} = (\gamma^\mu p_\mu - m)u_{\vec{p},\pm\vec{s}}\,e^{-ip\cdot x} \\ 0 &= (i\gamma^\mu\partial_\mu - m)v_{\vec{p},\pm\vec{s}}\,e^{ip\cdot x} = (-\gamma^\mu p_\mu - m)v_{\vec{p},\pm\vec{s}}\,e^{ip\cdot x} \end{aligned}$$

$$\rightarrow \boxed{(\not{p} - m)u_{\vec{p},\pm\vec{s}} = 0\,, \quad (\not{p} + m)v_{\vec{p},\pm\vec{s}} = 0\,.} \tag{3.248}$$

These equations hold for any $\vec{p}$ and $\vec{s}$ as long as the u, v spinors are constructed as defined by (3.247) and p^μ that appears in $\not{p}$ is given by $p^\mu = (p^0, \vec{p})$ where $\vec{p}$ is the same $\vec{p}$ that appears in (3.247) and $p^0 \overset{\text{def}}{\equiv} \sqrt{\vec{p}^2 + m^2} \geq 0$. These equations for the u, v spinors are sometimes referred to as the Dirac equations in momentum space. Corresponding equations for the adjoint spinors $\bar{u}$ and $\bar{v}$ can be obtained by simply taking the spinor adjoint of (3.248). Noting that for any *real* 4-vector a^μ, we have

$$\overline{\not{a}} = \overline{\gamma^\mu a_\mu} = \underbrace{\overline{\gamma^\mu}}_{\gamma^\mu} a_\mu = \not{a}\,, \tag{3.249}$$

the adjoint of (3.248) is

$$\overline{(\not{p} - m)u_{\vec{p},\pm\vec{s}}} = \overline{u_{\vec{p},\pm\vec{s}}}\,(\underbrace{\overline{\not{p}}}_{\not{p}} - m) = 0\,, \quad \overline{(\not{p} + m)v_{\vec{p},\pm\vec{s}}} = \overline{v_{\vec{p},\pm\vec{s}}}\,(\underbrace{\overline{\not{p}}}_{\not{p}} + m) = 0\,,$$

$$\rightarrow \boxed{\overline{u_{\vec{p},\pm\vec{s}}}\,(\not{p} - m) = 0\,, \quad \overline{v_{\vec{p},\pm\vec{s}}}\,(\not{p} + m) = 0\,.} \tag{3.250}$$

Now note that the equations (3.248) can be written as

$$\frac{\not p}{m}\, u_{\vec p,\pm\vec s} = u_{\vec p,\pm\vec s}\,, \quad \frac{\not p}{m}\, v_{\vec p,\pm\vec s} = -v_{\vec p,\pm\vec s}\,, \tag{3.251}$$

which means that the u and v spinors are eigenvectors of the matrix $\not p/m$ with eigenvalues $+1$ and -1, respectively. Namely, the operator $\not p/m$ represents the sign of the energy, or equivalently, whether the spinor is an electron solution or a positron solution. Using this, one can construct an *energy projection operator* that projects out a spinor of a given 4-momentum and a given energy sign:

$$\boxed{\Lambda_\pm(p) \stackrel{\text{def}}{\equiv} \frac{1}{2}\left(1 \pm \frac{\not p}{m}\right)}\,. \tag{3.252}$$

Then, it is easy to verify that (for a given $\vec p$)

$$\begin{aligned} \Lambda_+(p)\, u_{\vec p,\pm\vec s} &= u_{\vec p,\pm\vec s}\,, \qquad & \Lambda_-(p)\, u_{\vec p,\pm\vec s} &= 0\,, \\ \Lambda_-(p)\, v_{\vec p,\pm\vec s} &= v_{\vec p,\pm\vec s}\,, \qquad & \Lambda_+(p)\, v_{\vec p,\pm\vec s} &= 0\,, \end{aligned} \tag{3.253}$$

Since the four spinors $(u_{\vec p,\pm\vec s}, v_{\vec p,\pm\vec s})$ form a complete set, applying $\Lambda_+(p)$ to any 4-component quantity a projects out an electron solution with 4-momentum p^μ, and applying $\Lambda_-(p)$ projects out a positron solution with the same 4-momentum. It is also straightforward to show that they satisfy the property of projection operators:

$$\begin{gathered} \Lambda_\pm^2(p) = \Lambda_\pm(p)\,, \quad \Lambda_+(p) + \Lambda_-(p) = 1\,, \\ \Lambda_+(p)\Lambda_-(p) = \Lambda_-(p)\Lambda_+(p) = 0\,. \end{gathered} \tag{3.254}$$

In proving the above, it is useful to note

$$\not a \not b + \not b \not a = a_\mu \gamma^\mu b_\nu \gamma^\nu + b_\nu \gamma^\nu a_\mu \gamma^\mu = a_\mu b_\nu (\underbrace{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}_{2\, g^{\mu\nu}}) = 2a\cdot b\,. \tag{3.255}$$

where a and b are c-number 4-vectors. Applying this to $a = b = p$, we get

$$\not p^2 = p^2 = m^2\,. \tag{3.256}$$

Then, we have for example,

$$\Lambda_+(p)^2 = \frac{1}{4}\left(1 + \frac{\not p}{m}\right)^2 = \frac{1}{4}\Big(1 + 2\frac{\not p}{m} + \underbrace{\frac{\not p^2}{m^2}}_{1}\Big) = \frac{1}{2}\left(1 + \frac{\not p}{m}\right) = \Lambda_+(p)\,, \tag{3.257}$$

and other relations of (3.254) can be shown similarly.

Spin projection operators

We have just seen that the matrix $\not{p}/m$ acts as an energy sign operator for a given 4-momentum p^μ. We now look for an operator that represents the *physical spin* in a given direction $\vec{s}$, where by 'physical spin' we mean the spin we measure experimentally, in particular, for a positron we want a projection operator that projects out a given spin of the positron and not that of the missing electron in the sea of negative energy. In doing so, there are two problems we have to deal with: one is that for a positron, we want to make sure that it is the physical spin that is measured and not the eigenvalue of $\vec{s}\cdot\vec{\Sigma}$, and the other is that we want the spin measured in the rest frame of the particle even for a solution with a finite momentum.

For the first problem all we have to do is to somehow flip the sign of the spin for a positron. The second problem needs some care, since when an eigenspinor of $\vec{s}\cdot\vec{\Sigma}$ is boosted, it is in general no longer an eigenspinor of $\vec{s}\cdot\vec{\Sigma}$. Take for example the $\vec{p}=0$ solution $u_{\vec{0},+\vec{s}}$ which is an eigenspinor of $\vec{s}\cdot\vec{\Sigma}$ with eigenvalue $+1$. If we apply the operator $\vec{s}\cdot\vec{\Sigma}$ to the boosted spinor $u_{\vec{p},+\vec{s}}$, we get (dropping the normalization factor $\sqrt{E+m}$ for simplicity)

$$(\vec{s}\cdot\vec{\Sigma})\,u_{\vec{p},+\vec{s}} = \begin{pmatrix} \vec{s}\cdot\vec{\sigma} & 0 \\ 0 & \vec{s}\cdot\vec{\sigma} \end{pmatrix} \begin{pmatrix} \chi_+ \\ \dfrac{\vec{p}\cdot\vec{\sigma}}{E+m}\chi_+ \end{pmatrix} = \begin{pmatrix} (\vec{s}\cdot\vec{\sigma})\chi_+ \\ \dfrac{(\vec{s}\cdot\vec{\sigma})(\vec{p}\cdot\vec{\sigma})}{E+m}\chi_+ \end{pmatrix}. \tag{3.258}$$

Since $(\vec{s}\cdot\vec{\sigma})\chi_+ = \chi_+$, if $\vec{s}\cdot\vec{\sigma}$ and $\vec{p}\cdot\vec{\sigma}$ commute, then $u_{\vec{p},+\vec{s}}$ becomes an eigenvector of $\vec{s}\cdot\vec{\Sigma}$. However, we have

$$[\vec{s}\cdot\vec{\sigma},\vec{p}\cdot\vec{\sigma}] = [s^i\sigma_i\,,p^j\sigma_j] = s^ip^j\,\underbrace{[\sigma_i\,,\sigma_j]}_{2i\epsilon_{ijk}\sigma_k} = 2i(\vec{s}\times\vec{p})\cdot\vec{\sigma} \tag{3.259}$$

which is in general not zero; thus, $u_{\vec{p},+\vec{s}}$ is in general *not* an eigenspinor of $\vec{s}\cdot\vec{\Sigma}$ (unless $\vec{s}\times\vec{p}=0$). Our strategy then is to go to the rest frame of the particle, construct an operator that reflects the physical spin there, express it in a Lorentz-invariant form, and then carefully 'boost' it.

Let's start with constructing the eigenstates of $\vec{s}\cdot\vec{\Sigma}$ in the rest frame in a way that is independent of representation. In the Dirac representation, the $\vec{p}=0$ solutions $u_{\vec{0},\pm\vec{s}}$ and $v_{\vec{0},\pm\vec{s}}$ look like

$$u_{\vec{0},\pm\vec{s}} \propto \begin{pmatrix} \chi_\pm \\ 0 \end{pmatrix}, \quad v_{\vec{0},\pm\vec{s}} \propto \begin{pmatrix} 0 \\ \chi_\mp \end{pmatrix}, \tag{3.260}$$

and they are eigenstates of $\vec{s}\cdot\vec{\Sigma}$:

$$(\vec{s}\cdot\vec{\Sigma})\,u_{\vec{0},\pm\vec{s}} = \pm u_{\vec{0},\pm\vec{s}}\,, \quad (\vec{s}\cdot\vec{\Sigma})\,v_{\vec{0},\pm\vec{s}} = \mp v_{\vec{0},\pm\vec{s}}\,. \tag{3.261}$$

These relations are independent of representation. In fact, if we change the representation by a unitary matrix V, then by (3.150) and (3.152), the spinor and Σ_i in the new representation are given by

$$
\begin{aligned}
u'_{\vec{0},\pm\vec{s}} &= V u_{\vec{0},\pm\vec{s}}\,, \\
\Sigma'_i = i\gamma'^j\gamma'^k &= i(V\gamma^j \underbrace{V^\dagger)(V}_{1} \gamma^k V^\dagger) = V i\,\gamma^j\gamma^k V^\dagger = V\Sigma_i V^\dagger\,.
\end{aligned} \tag{3.262}
$$

If $(\vec{s}\cdot\vec{\Sigma})\,u_{\vec{0},\pm\vec{s}} = \pm u_{\vec{0},\pm\vec{s}}$ holds in one representation, then multiplying V from the left,

$$
\begin{aligned}
\overbrace{V}\,(\vec{s}\cdot\vec{\Sigma})\,\overbrace{V^\dagger V}\,u_{\vec{0},\pm\vec{s}} &= \pm\overbrace{V}\,u_{\vec{0},\pm\vec{s}} \\
\to (\vec{s}\cdot\vec{\Sigma}')\,u'_{\vec{0},\pm\vec{s}} &= \pm u'_{\vec{0},\pm\vec{s}}\,,
\end{aligned} \tag{3.263}
$$

demonstrating that the same form holds in the new representation.

Note that for the v spinors in (3.261), the sign on the subscript $\vec{s}$ is opposite to the sign of the eigenvalue. We thus want an operator whose eigenvalues for v spinors are sign-flipped with respect to $\vec{s}\cdot\vec{\Sigma}$. This can be accomplished by (3.251) which can be written in the rest frame as

$$
\gamma^0 u_{\vec{0},\pm\vec{s}} = u_{\vec{0},\pm\vec{s}}\,, \quad \gamma^0 v_{\vec{0},\pm\vec{s}} = -v_{\vec{0},\pm\vec{s}}\,, \tag{3.264}
$$

where we have used $\not{p} = m\gamma^0$ [since $p^\mu = (m,\vec{0})$ in the rest frame]. It will thus flip the sign for the v spinors only. Together with (3.261), we have

$$
(\vec{s}\cdot\vec{\Sigma}\,\underbrace{\gamma^0)\,u_{\vec{0},\pm\vec{s}}}_{u_{\vec{0},\pm\vec{s}}} = \pm u_{\vec{0},\pm\vec{s}}\,, \quad (\vec{s}\cdot\vec{\Sigma}\,\underbrace{\gamma^0)\,v_{\vec{0},\pm\vec{s}}}_{-v_{\vec{0},\pm\vec{s}}} = \pm v_{\vec{0},\pm\vec{s}}\,, \tag{3.265}
$$

which is exactly what we wanted. This, however, works only for the rest frame solutions. In order to extend it to the boosted states, we first show that the operator $\vec{s}\cdot\vec{\Sigma}\gamma^0$ can be written in the rest frame as

$$
\vec{s}\cdot\vec{\Sigma}\gamma^0 = \gamma_5\not{s} \quad \text{(rest frame)}, \tag{3.266}
$$

where the *real* quantity s^μ is defined in the rest frame by

$$
s^\mu \stackrel{\text{def}}{\equiv} (0,\vec{s}) \quad \text{(rest frame)}, \tag{3.267}
$$

and assumed to transform as a Lorentz 4-vector. Since $\not{s} = s^i\gamma_i$ in the rest frame, and

$$
\Sigma_i \equiv i\gamma^j\gamma^k = -i\gamma^1\gamma^2\gamma^3\gamma^i \quad (i,j,k:\text{cyclic}) \tag{3.268}
$$

which can be verified for $k = 1, 2, 3$ explicitly, we have

$$\begin{aligned}\gamma_5 \not{s} &= (i\boxed{\gamma^0}\overrightarrow{\gamma^1\gamma^2\gamma^3)(s^i\gamma_i)} = s^i \underbrace{i\gamma^1\gamma^2\gamma^3\gamma_i}_{-i\gamma^1\gamma^2\gamma^3\gamma^i = \Sigma_i}\gamma^0 \\ &= s^i\Sigma_i\gamma^0 = \vec{s}\cdot\vec{\Sigma}\gamma^0 \,. \end{aligned} \tag{3.269}$$

Thus, (3.265) can be written as

$$\gamma_5 \not{s}\, w_{\vec{0},\pm\vec{s}} = \pm w_{\vec{0},\pm\vec{s}} \quad (w : u \text{ or } v)\,. \tag{3.270}$$

The above equation is for spinors representing a particle at rest. A spinor with momentum $\vec{p}$ was defined by

$$w_{\vec{p},\pm\vec{s}} \stackrel{\text{def}}{\equiv} S(\Lambda) w_{\vec{0},\pm\vec{s}}\,, \tag{3.271}$$

where as before w represents u or v, and Λ is the boost that makes the rest mass m acquire a momentum $\vec{p}$. First, we set $\Lambda \to \Lambda^{-1}$ in $S^{-1}(\Lambda)\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\alpha\gamma^\alpha$ (3.79) and use $S(\Lambda^{-1}) = S^{-1}(\Lambda)$, and thus $S^{-1}(\Lambda^{-1}) = S(\Lambda)$, to get

$$\underbrace{S^{-1}(\Lambda^{-1})\gamma^\mu S(\Lambda^{-1})}_{S(\Lambda)\gamma^\mu S^{-1}(\Lambda)} = \underbrace{(\Lambda^{-1})^\mu{}_\alpha}_{\Lambda_\alpha{}^\mu}\gamma^\alpha\,. \tag{3.272}$$

Multiplying (3.270) by $S(\Lambda)$ on the left, we obtain

$$S(\Lambda) \times \overbrace{\ \gamma_5 \not{s}\ \ w_{\vec{0},\pm\vec{s}}}^{}{}_{S^{-1}(\Lambda)S(\Lambda)} = S(\Lambda) \times\ \pm w_{\vec{0},\pm\vec{s}} \tag{3.273}$$

Together with $[\gamma_5, S] = 0$ [see (3.138)], this becomes

$$\gamma_5\ \underbrace{s_\mu \underbrace{S(\Lambda)\gamma^\mu S^{-1}(\Lambda)}_{(3.272)\ \to\ \Lambda_\alpha{}^\mu\gamma^\alpha}}_{s'_\alpha\gamma^\alpha = \not{s}'} \underbrace{S(\Lambda)w_{\vec{0},\pm\vec{s}}}_{w_{\vec{p},\pm\vec{s}}} = \pm \underbrace{S(\Lambda)w_{\vec{0},\pm\vec{s}}}_{w_{\vec{p},\pm\vec{s}}}\,. \tag{3.274}$$

where we have defined the boosted spin 4-vector

$$s'^\mu \stackrel{\text{def}}{\equiv} \Lambda^\mu{}_\nu s^\nu\,. \tag{3.275}$$

Thus, we have

$$\boxed{\gamma_5 \not{s}\, w_{\vec{p},\pm\vec{s}} = \pm w_{\vec{p},\pm\vec{s}} \quad (w : u \text{ or } v)}\,, \tag{3.276}$$

where we have dropped the prime on s with the understanding that s^μ is always the transformed 4-vector of $(0, \vec{s})$ by the same boost under which the rest mass m

acquires the momentum $\vec{p}$. Thus, we see that the operator $\gamma_5 \not{s}$ has eigenvalues ± 1, and properly represents the component of the *physical spin* along the direction $\vec{s}$ in the rest frame (times two, to be precise), where it is understood that the rest frame is reached by a pure boost. And it works for a positron as well as for an electron. Also, note that our derivation was independent of the representation of the γ matrices.

Just like we constructed the energy projection operator from the energy-sign operator $\not{p}/m$, we can construct spin projection operators as

$$\boxed{\Sigma_\pm(s) \stackrel{\text{def}}{\equiv} \frac{1 \pm \gamma_5 \not{s}}{2}}\,, \tag{3.277}$$

which project out the eigenstates of the spin operator $\gamma_5 \not{s}$ for the given direction $\vec{s}$ and the given boost represented by $\vec{p}$. Again, it can be readily verified that the operators $\Sigma_\pm(s)$ satisfy the properties of projection operators:

$$\begin{gathered} \Sigma_\pm^2(s) = \Sigma_\pm(s)\,, \quad \Sigma_+(s) + \Sigma_-(s) = 1\,, \\ \Sigma_+(s)\Sigma_-(s) = \Sigma_-(s)\Sigma_+(s) = 0\,. \end{gathered} \tag{3.278}$$

Orthonormalities of the u, v spinors

In the rest frame, we have $s^\mu = (0, \vec{s})$ and $p^\mu = (m, \vec{0})$. Thus,

$$\boxed{s^2 = -1\,, \quad s \cdot p = 0}\,, \tag{3.279}$$

which are in Lorentz-invariant form and thus true in any frame. Then from (3.255), we see that $\not{s}$ and $\not{s}$ anticommute:

$$\not{s}\not{p} + \not{p}\not{s} = 2 s \cdot p = 0\,. \tag{3.280}$$

For any given set of $\vec{s}$ and $\vec{p}$, the energy-sign operator $\not{p}/m$ and the spin operator $\gamma_5 \not{s}$ then commute:

$$\begin{aligned} \underbrace{\not{p}\,(\gamma_5}_{-\gamma_5 \not{p}} \not{s}) = -\gamma_5 \underbrace{\not{p}\not{s}}_{-\not{s}\not{p}} \;&=\; (\gamma_5 \not{s})\not{p} \\ \to \quad \left[\frac{\not{p}}{m}, \gamma_5 \not{s}\right] \;&=\; 0\,. \end{aligned} \tag{3.281}$$

We have seen that the two commuting operators $\gamma_5 \not{s}$ and $\not{p}/m$ both have eigenvalues ± 1. Thus, there should be a set of four spinors that are simultaneous eigenspinors of the two operators. In fact, they are nothing but $u_{\vec{p},\pm\vec{s}}$ and $v_{\vec{p},\pm\vec{s}}$. Table 3.3 summarizes their eigenvalues. Now we will show that they are also orthonormal.

		$\not{p}/m$	
		$+1$	-1
$\gamma_5\not{s}$	$+1$	$u_{\vec{p},\vec{s}}$	$v_{\vec{p},\vec{s}}$
	-1	$u_{\vec{p},-\vec{s}}$	$v_{\vec{p},-\vec{s}}$

Table 3.3: The u and v spinors for a given momentum $\vec{p}$ and a spin quantization axis $\vec{s}$ shown with the eigenvalues of the operators $\not{p}/m$ and $\gamma_5\not{s}$.

First, the operators $\not{p}/m$ and $\gamma_5\not{s}$ are self-adjoint: using $\overline{\not{a}} = \not{a}$ (3.249),

$$\overline{\frac{\not{p}}{m}} = \frac{\not{p}}{m}\,, \qquad \overline{\gamma_5\not{s}} = \overline{\not{s}}\,\underbrace{\overline{\gamma_5}}_{-\gamma_5} = -\not{s}\gamma_5 = \gamma_5\not{s}\,, \tag{3.282}$$

where we have used

$$\overline{\gamma_5} \equiv \gamma^0 \underbrace{\gamma_5^\dagger}_{\gamma_5 \text{ by (3.135)}} \gamma^0 = \gamma^0\gamma_5\gamma^0 = -\gamma_5\,. \tag{3.283}$$

This indicates that the eigenspinors are orthogonal in terms of the inner product defined by the spinor adjoint. In fact, using (3.251),

$$\frac{\not{p}}{m}\,u_{\vec{p},\pm\vec{s}} = u_{\vec{p},\pm\vec{s}} \quad \xrightarrow{\overline{v_{\vec{p},\pm\vec{s}}}\,\times} \quad \overline{v_{\vec{p},\pm\vec{s}}}\,\frac{\not{p}}{m}\,u_{\vec{p},\pm\vec{s}} = \overline{v_{\vec{p},\pm\vec{s}}}\,u_{\vec{p},\pm\vec{s}}\,. \tag{3.284}$$

On the other hand, taking adjoint of $(\not{p}/m)v_{\vec{p},\pm\vec{s}} = -v_{\vec{p},\pm\vec{s}}$, and using $\overline{\not{p}/m} = \not{p}/m$,

$$\overline{v_{\vec{p},\pm\vec{s}}}\,\frac{\not{p}}{m} = -\overline{v_{\vec{p},\pm\vec{s}}} \quad \xrightarrow{\times\, u_{\vec{p},\pm\vec{s}}} \quad \overline{v_{\vec{p},\pm\vec{s}}}\,\frac{\not{p}}{m}\,u_{\vec{p},\pm\vec{s}} = -\overline{v_{\vec{p},\pm\vec{s}}}\,u_{\vec{p},\pm\vec{s}}\,. \tag{3.285}$$

Subtracting (3.285) from (3.284), we obtain

$$\overline{v_{\vec{p},\pm\vec{s}}}\,u_{\vec{p},\pm\vec{s}} = 0\,, \tag{3.286}$$

which holds for all spin combinations. Note that it was critical that $\not{p}/m$ was self-adjoint in order to make the left-hand sides of (3.285) and (3.284) identical.

Similarly, two spinors are orthogonal if they have different eigenvalues of $\gamma_5\not{s}$ as can be shown by simply replacing $\not{p}/m$ by $\gamma_5\not{s}$ in the above derivation (3.284) through (3.286). For example, we have

$$\gamma_5\not{s}\,u_{\vec{p},\vec{s}} = u_{\vec{p},\vec{s}}\,, \quad \gamma_5\not{s}\,u_{\vec{p},-\vec{s}} = -u_{\vec{p},-\vec{s}}\,. \tag{3.287}$$

Multiplying the first with $\overline{u_{\vec{p},-\vec{s}}}$ from the left, and multiplying the spinor adjoint of the second with $u_{\vec{p},\vec{s}}$ from the right, we obtain

$$\overline{u_{\vec{p},-\vec{s}}}\,\gamma_5\not{s}\,u_{\vec{p},\vec{s}} = \overline{u_{\vec{p},-\vec{s}}}\,u_{\vec{p},\vec{s}}\,, \quad \overline{u_{\vec{p},-\vec{s}}}\,\gamma_5\not{s}\,u_{\vec{p},\vec{s}} = -\overline{u_{\vec{p},-\vec{s}}}\,u_{\vec{p},\vec{s}}\,, \tag{3.288}$$

where we have used $\overline{\gamma_5 \not{s}} = \gamma_5 \not{s}$ (3.282). Taking the difference, we get

$$\overline{u_{\vec{p},-\vec{s}}}\, u_{\vec{p},\vec{s}} = 0 \,. \tag{3.289}$$

Similarly for v spinors, we obtain

$$\overline{v_{\vec{p},-\vec{s}}}\, v_{\vec{p},\vec{s}} = 0 \,. \tag{3.290}$$

Thus, all four spinors in Table 3.3 are orthogonal to each other.

There is another set of four spinors that are orthogonal to each other where the orthogonality is defined by the ordinary inner product $a^\dagger b$. First, replace $\vec{p}$ by $-\vec{p}$ in $(\not{p}/m) v_{\vec{p},\pm\vec{s}} = -v_{\vec{p},\pm\vec{s}}$ to get

$$\frac{\not{p}'}{m} v_{-\vec{p},\pm\vec{s}} = -v_{-\vec{p},\pm\vec{s}} \qquad \text{with} \quad p'^\mu \equiv (p^0, -\vec{p}) = p_\mu \,. \tag{3.291}$$

Note that the time component of p' did not change since $p'^0 \overset{\text{def}}{\equiv} \sqrt{(-\vec{p})^2 + m^2} = p^0$. We first note

$$(\not{p}')^\dagger = (\gamma_\mu p'^\mu)^\dagger = \underbrace{\gamma_\mu^\dagger}_{\gamma^\mu} \underbrace{p'^\mu}_{p_\mu} = \not{p} \,. \tag{3.292}$$

Then, taking the hermitian conjugate of (3.291), we have

$$v^\dagger_{-\vec{p},\pm\vec{s}} \frac{\overbrace{(\not{p}')^\dagger}^{\not{p}}}{m} = -v^\dagger_{-\vec{p},\pm\vec{s}} \,, \tag{3.293}$$

to which we multiply $u_{\vec{p},\pm\vec{s}}$ from the right to get

$$v^\dagger_{-\vec{p},\pm\vec{s}} \frac{\not{p}}{m} u_{\vec{p},\pm\vec{s}} = -v^\dagger_{-\vec{p},\pm\vec{s}} u_{\vec{p},\pm\vec{s}} \,. \tag{3.294}$$

On the other hand, multiplying $v^\dagger_{-\vec{p},\pm\vec{s}}$ to $(\not{p}/m) u_{\vec{p},\pm\vec{s}} = u_{\vec{p},\pm\vec{s}}$,

$$v^\dagger_{-\vec{p},\pm\vec{s}} \frac{\not{p}}{m} u_{\vec{p},\pm\vec{s}} = v^\dagger_{-\vec{p},\pm\vec{s}} u_{\vec{p},\pm\vec{s}} \,. \tag{3.295}$$

Taking the difference of (3.294) and (3.295), we see that

$$v^\dagger_{-\vec{p},\pm\vec{s}} u_{\vec{p},\pm\vec{s}} = 0 \,, \tag{3.296}$$

which holds for any spin combinations.

Now, let's look at the orthogonality of spinors with different spin eigenvalues. First, we can see that $u^\dagger_{\vec{p},+\vec{s}}\,u_{\vec{p},-\vec{s}} = 0$ explicitly in the Dirac representation as follows. Recalling $u_{\vec{p},\pm\vec{s}} = e^{\vec{\xi}\cdot\vec{\alpha}/2}\,u_{\vec{0},\pm\vec{s}}$ and, since α_i is hermitian,

$$(e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}})^\dagger = e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}^\dagger} = e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}}\,, \tag{3.297}$$

we have

$$\begin{aligned}
u^\dagger_{\vec{p},+\vec{s}}\,u_{\vec{p},-\vec{s}} &= u^\dagger_{\vec{0},+\vec{s}}\underbrace{(e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}})^\dagger}_{e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}}}e^{\frac{1}{2}\vec{\xi}\cdot\vec{\alpha}}u_{\vec{p},-\vec{s}} = u^\dagger_{\vec{0},+\vec{s}}\,e^{\vec{\xi}\cdot\vec{\alpha}}\,u_{\vec{p},-\vec{s}} \\
&= 2m\,(\,\chi^\dagger_+ \quad 0\,)\begin{pmatrix} \cosh\xi & (\hat{\xi}\cdot\vec{\sigma})\sinh\xi \\ (\hat{\xi}\cdot\vec{\sigma})\sinh\xi & \cosh\xi \end{pmatrix}\begin{pmatrix}\chi_- \\ 0\end{pmatrix} \\
&= 2m\cosh\xi\,(\chi^\dagger_+\chi_-) = 0\,,
\end{aligned} \tag{3.298}$$

where we have used the expression of $e^{\vec{\xi}\cdot\vec{\alpha}/2}$ given in (3.224) with $\vec{\xi}/2$ replaced by $\vec{\xi}$. This is representation-independent, since if we move to a different representation by a unitary matrix V, then the orthogonality remains valid:

$$u'^\dagger_{\vec{p},+\vec{s}}\,u'_{\vec{p},-\vec{s}} = (V u_{\vec{p},+\vec{s}})^\dagger(V u_{\vec{p},-\vec{s}}) = u^\dagger_{\vec{p},+\vec{s}}\underbrace{V^\dagger V}_{1}u_{\vec{p},-\vec{s}} = 0\,. \tag{3.299}$$

Similarly, we can show that $v^\dagger_{\vec{p},+\vec{s}}\,v_{\vec{p},-\vec{s}} = 0$ which is also representation-independent. Thus, for given $\vec{p}$ and $\vec{s}$, we have two sets of orthogonal spinors; one defined by the inner product $\bar{a}b$: $(u_{\vec{p},\pm\vec{s}}\,,\ v_{\vec{p},\pm\vec{s}})$, and the other by $a^\dagger b$: $(u_{\vec{p},\pm\vec{s}}\,,\ v_{-\vec{p},\pm\vec{s}})$.

How about the normalization? We have seen in (3.243) that the probability density is $j^0 = \psi^\dagger\psi = 2E$ which is always positive. Using the plane wave form (3.246) for ψ, the normalization of u, v spinors are then

$$u^\dagger_{\vec{p},\pm\vec{s}}\,u_{\vec{p},\pm\vec{s}} = v^\dagger_{\vec{p},\pm\vec{s}}\,v_{\vec{p},\pm\vec{s}} = 2E \tag{3.300}$$

Once this is given, the values of $\overline{u_{\vec{p},\pm\vec{s}}}\,u_{\vec{p},\pm\vec{s}}$ and $\overline{v_{\vec{p},\pm\vec{s}}}\,v_{\vec{p},\pm\vec{s}}$ are already fixed as follows: For the solution $\psi = u_{\vec{p}}e^{-ip\cdot x}$, the current is $j^\mu = \bar{u}_{\vec{p}}\gamma^\mu u_{\vec{p}}$ (we have dropped the spin indexes for simplicity). Using $(\not{p}/m)u_{\vec{p}} = u_{\vec{p}}$ (3.251) and its adjoint $\bar{u}_{\vec{p}}(\not{p}/m) = \bar{u}_{\vec{p}}$ this can be written as

$$\begin{aligned}
j^\mu = \bar{u}_{\vec{p}}\gamma^\mu u_{\vec{p}} &= \bar{u}_{\vec{p}}\gamma^\mu\Big(\frac{\not{p}}{m}u_{\vec{p}}\Big) = \Big(\bar{u}_{\vec{p}}\frac{\not{p}}{m}\Big)\gamma^\mu u_{\vec{p}} \\
&= \frac{1}{2}\bar{u}_{\vec{p}}\Big(\gamma^\mu\frac{\not{p}}{m} + \frac{\not{p}}{m}\gamma^\mu\Big)u_{\vec{p}} \\
&= \frac{1}{2m}\bar{u}_{\vec{p}}\,[\gamma^\mu(p_\nu\gamma^\nu) + (p_\nu\gamma^\nu)\gamma^\mu]u_{\vec{p}} \\
&= \frac{p_\nu}{2m}\bar{u}_{\vec{p}}\,(\underbrace{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}_{2g^{\mu\nu}})u_{\vec{p}} \\
&= \eta^\mu\bar{u}_{\vec{p}}u_{\vec{p}}\,.
\end{aligned} \tag{3.301}$$

It is instructive to compare this with the current $j^\mu = \rho^0 \eta^\mu$ we encountered in (1.80). We see that the scalar quantity $\bar{u}_{\vec{p}} u_{\vec{p}}$ is acting as ρ^0: the density in the rest frame of the flow. For the v spinors, the only difference is the minus sign in $(\not{p}/m)v_{\vec{p}} = -v_{\vec{p}}$ which leads to

$$\bar{v}_{\vec{p}} \gamma^\mu v_{\vec{p}} = -\eta^\mu \bar{v}_{\vec{p}} v_{\vec{p}} \,. \tag{3.302}$$

The time components of (3.301) and (3.302) give

$$u^\dagger_{\vec{p}} u_{\vec{p}} = \gamma \, \bar{u}_{\vec{p}} u_{\vec{p}} \,, \quad v^\dagger_{\vec{p}} v_{\vec{p}} = -\gamma \, \bar{v}_{\vec{p}} v_{\vec{p}} \,, \tag{3.303}$$

where $\gamma \equiv \eta^0 = E/m$. Together with (3.300) this leads to

$$\bar{u}_{\vec{p}} u_{\vec{p}} = 2m \,, \quad \bar{v}_{\vec{p}} v_{\vec{p}} = -2m \,. \tag{3.304}$$

Thus, for given $\vec{p}$ and $\vec{s}$, we have two orthonormality relations for u, v spinors :

$$\boxed{\begin{array}{c} \overline{u_{\vec{p},\vec{s}_1}} \, u_{\vec{p},\vec{s}_2} = 2m \, \delta_{\vec{s}_1,\vec{s}_2} \,, \quad \overline{v_{\vec{p},\vec{s}_1}} \, v_{\vec{p},\vec{s}_2} = -2m \, \delta_{\vec{s}_1,\vec{s}_2} \,, \\ \overline{v_{\vec{p},\vec{s}_1}} \, u_{\vec{p},\vec{s}_2} = 0 \qquad (\vec{s}_1, \vec{s}_2 : \pm\vec{s}) \end{array}} \tag{3.305}$$

as defined by the inner product $\bar{a}b$, and

$$\boxed{\begin{array}{c} u^\dagger_{\vec{p},\vec{s}_1} \, u_{\vec{p},\vec{s}_2} = v^\dagger_{-\vec{p},\vec{s}_1} \, v_{-\vec{p},\vec{s}_2} = 2E \, \delta_{\vec{s}_1,\vec{s}_2} \,, \\ v^\dagger_{-\vec{p},\vec{s}_1} \, u_{\vec{p},\vec{s}_2} = 0 \qquad (\vec{s}_1, \vec{s}_2 : \pm\vec{s}) \end{array}} \tag{3.306}$$

as defined by the inner product $a^\dagger b$. Any spinor, or any set of four complex numbers, can be written uniquely as a linear combination of either of the above orthonormal sets. Note that one can construct such orthonormal sets for any given $\vec{p}$ and $\vec{s}$; thus, there are an infinite number of orthonormal sets. The coefficients of the linear combinations can be readily obtained by using the orthonormality relations above. For example, for the second orthonormal basis, the coefficient of $v_{-\vec{p},\vec{s}}$ for an arbitrary spinor a is given by taking the inner product of a with $v_{-\vec{p},\vec{s}}$ and accounting for the normalization:

$$\frac{v^\dagger_{-\vec{p},\vec{s}} \, a}{2E} \,. \tag{3.307}$$

Exercise 3.9 *Orthonormality of u, v spinors (by the product rule $a^\dagger b$).*
Use the Dirac representation of the u, v spinors to explicitly verify the orthonormality relations (3.306).

Note that the normalization (3.306) tells us that the norm $\psi^\dagger \psi$ of a Dirac spinor is in general not invariant under a boost since E changes its value. In fact, as we have seen in (3.297), a boost $S = e^{\vec{\xi}\cdot\vec{\alpha}/2}$ is hermitian and *not* unitary, while a rotation in the spinor space $U = e^{-i\vec{\theta}\cdot\vec{\Sigma}/2}$ *is* unitary:

$$U^\dagger = \left(e^{-i\frac{\vec{\theta}}{2}\cdot\vec{\Sigma}}\right)^\dagger = e^{i\frac{\vec{\theta}}{2}\cdot\vec{\Sigma}^\dagger} = e^{i\frac{\vec{\theta}}{2}\cdot\vec{\Sigma}} = U^{-1} \tag{3.308}$$

and thus the norm $\psi^\dagger\psi$ is invariant under a rotation:

$$\psi^\dagger\psi \quad \xrightarrow{U} \quad (U\psi)^\dagger(U\psi) = \psi^\dagger \underbrace{U^\dagger U}_{1} \psi = \psi^\dagger\psi\,. \tag{3.309}$$

There exist useful relations between the u, v spinors and the energy and spin projection operators that will be used later in actual calculations (the proof is left as an exercise):

$$\begin{array}{rcrcl} u_{\vec{p},\pm\vec{s}}\,\overline{u_{\vec{p},\pm\vec{s}}} & = & 2m\,\Lambda_+(p)\Sigma_\pm(s) & = & (\not{p}+m)\dfrac{1\pm\gamma_5\not{s}}{2} \\ v_{\vec{p},\pm\vec{s}}\,\overline{v_{\vec{p},\pm\vec{s}}} & = & -2m\,\Lambda_-(p)\Sigma_\pm(s) & = & (\not{p}-m)\dfrac{1\pm\gamma_5\not{s}}{2} \end{array} \tag{3.310}$$

where for any spinors a and b, which are column vectors, a 4×4 matrix ab^T is defined by

$$(ab^T)_{ij} \stackrel{\text{def}}{\equiv} a_i\,b_j\,. \tag{3.311}$$

By this definition, the multiplications of matrices and vectors become 'associative'; for example, if a, b, and c are column vectors, we have

$$[(ab^T)c]_i = (ab^T)_{ij}c_j = a_ib_jc_j = a_i(b^Tc) \quad \rightarrow \quad (ab^T)c = a(b^Tc)\,, \tag{3.312}$$

which has the form

$$\overbrace{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}^{(ab^T)} \overbrace{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}}^{c} = \overbrace{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}}^{a} \overbrace{\begin{pmatrix} \cdot \end{pmatrix}}^{(b^Tc)} \,. \tag{3.313}$$

Exercise 3.10 *u, v spinors and energy-spin projection operators.*
Prove the identities (3.310). (hint: Since, for a given set of $\vec{p}$ and $\vec{s}$, the spinor space is spanned by $u_{\vec{p},\pm\vec{s}}$ and $v_{\vec{p},\pm\vec{s}}$, all that is needed is to show that the LHS and the RHS of the identities behave the same for these 4 spinors.)

Similarly,

$$c^T(ab^T) = (c^Ta)b^T\,. \tag{3.314}$$

In terms of the spinor adjoint, we have

$$(a\bar{b})c = a(\bar{b}c) \qquad \bar{c}(a\bar{b}) = (\bar{c}a)\bar{b}\,. \tag{3.315}$$

Summing over the sign of spins in (3.310), we obtain

$$\boxed{\begin{aligned} \sum_{\pm\vec{s}} u_{\vec{p},\vec{s}}\,\overline{u_{\vec{p},\vec{s}}} &= (\not{p}+m) \\ \sum_{\pm\vec{s}} v_{\vec{p},\vec{s}}\,\overline{v_{\vec{p},\vec{s}}} &= (\not{p}-m) \end{aligned}}\,, \tag{3.316}$$

which will be used extensively when we calculate transition rates averaged over spins. Taking the difference of these two equations, we obtain a relation which expresses the completeness of the first orthonormal set:

$$\frac{1}{2m}\sum_{\pm\vec{s}}\Big(u_{\vec{p},\vec{s}}\,\overline{u_{\vec{p},\vec{s}}} - v_{\vec{p},\vec{s}}\,\overline{v_{\vec{p},\vec{s}}}\Big) = 1\,, \tag{3.317}$$

where the right hand side is actually the 4×4 identity matrix. This can be used to write any spinor a as a linear combination of the orthonormal basis ($u_{\vec{p},\pm\vec{s}}\,,\ v_{\vec{p},\pm\vec{s}}$):

$$a = \frac{1}{2m}\sum_{\pm\vec{s}}\Big(u_{\vec{p},\vec{s}}\,\overline{u_{\vec{p},\vec{s}}} - v_{\vec{p},\vec{s}}\,\overline{v_{\vec{p},\vec{s}}}\Big)a = \frac{1}{2m}\sum_{\pm\vec{s}}\Big[u_{\vec{p},\vec{s}}\,(\overline{u_{\vec{p},\vec{s}}}\,a) - v_{\vec{p},\vec{s}}\,(\overline{v_{\vec{p},\vec{s}}}\,a)\Big]\,, \tag{3.318}$$

where we have used the associativity (3.312). Namely, the coefficient of $u_{\vec{p},\vec{s}}$ is $\overline{u_{\vec{p},\vec{s}}}\,a/2m$, and that of $v_{\vec{p},\vec{s}}$ is $-\overline{v_{\vec{p},\vec{s}}}\,a/2m$. The completeness relation for the second orthogonal basis ($u_{\vec{p},\pm\vec{s}}\,,\ v_{-\vec{p},\pm\vec{s}}$) is given by

$$\frac{1}{2E}\sum_{\pm\vec{s}}\Big(u_{\vec{p},\vec{s}}\,u^\dagger_{\vec{p},\vec{s}} + v_{-\vec{p},\vec{s}}\,v^\dagger_{-\vec{p},\vec{s}}\Big) = 1\,, \tag{3.319}$$

which is readily verified to give the correct coefficients for an arbitrary spinor a. For example, by applying the above identity to a, we get

$$\frac{1}{2E}\sum_{\pm\vec{s}}\Big(u_{\vec{p},\vec{s}}\,(u^\dagger_{\vec{p},\vec{s}}a) + v_{-\vec{p},\vec{s}}\,(v^\dagger_{-\vec{p},\vec{s}}a)\Big) = a\,, \tag{3.320}$$

where we see that the coefficient of $v_{-\vec{p},\vec{s}}$ is $(v^\dagger_{-\vec{p},\vec{s}}a)/2E$ which is what we obtained in (3.307).

3.9 Low energy limit - electron magnetic moment

In order to find out the value of the electron magnetic moment predicted by the Dirac equation, we have to somehow introduce the coupling to the electromagnetic field $A^\mu = (\Phi, \vec{A})$ and then look for a term in the potential which looks like μB where B is the magnetic field and μ is the magnetic moment. For now, we accept that the interaction is introduced by the so-called minimal substitution

$$\partial_\mu \quad \rightarrow \quad D_\mu \overset{\text{def}}{\equiv} \partial_\mu + ieA_\mu \qquad \text{with} \quad A^\mu(x) = (\Phi(x), \vec{A}(x))\,, \tag{3.321}$$

where e is the charge of electron ($e < 0$). Starting from the Schrödinger form of the Dirac equation (3.29): $i\partial_0\psi = (-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\psi$, the minimal substitution

$$\partial_\mu = (\partial_0, \vec{\nabla}) \quad \rightarrow \quad D_\mu = (\partial_0 + ie\Phi, \vec{D}) \tag{3.322}$$

with

$$D_i \overset{\text{def}}{\equiv} \nabla_i + ie\underbrace{A_i}_{-A^i} = \nabla_i - ieA^i\,, \quad \text{or} \quad \vec{D} \overset{\text{def}}{\equiv} \vec{\nabla} - ie\vec{A} \tag{3.323}$$

yields

$$\begin{aligned} (i\partial_0 - e\Phi)\psi &= \Big(-i\vec{\alpha}\cdot\vec{D} + \beta m\Big)\psi \\ \rightarrow \quad i\dot{\psi} &= \Big(-i\vec{\alpha}\cdot\vec{D} + \beta m + e\Phi\Big)\psi\,. \end{aligned} \tag{3.324}$$

We are interested in a low energy electron; namely, $|\vec{p}| \ll m$ and $E \sim m$, or equivalently, $i\partial_i\psi \ll m\psi$ and $i\partial_0\psi \sim m\psi$. Thus, if we write $\psi(x)$ as

$$\psi(x) = \psi_s(x)e^{-imt}\,, \quad \psi_s(x) = \begin{pmatrix} \phi(x) \\ \eta(x) \end{pmatrix} \tag{3.325}$$

where ϕ and η are 2-component functions, then the space-time derivatives of ϕ and η are small: [5]

$$i\partial_\mu\phi \ll m\phi\,, \qquad i\partial_\mu\eta \ll m\eta\,. \tag{3.326}$$

Also, since the potential energies are assumed to be small compared to the rest mass,

$$e\Phi \ll m\,, \qquad eA^i \ll m \quad (i = 1, 2, 3)\,. \tag{3.327}$$

Furthermore, the solution in the Dirac representation for an electron at rest, $u_{\vec{0},\pm\vec{s}}\,e^{-imt}$, has only upper two components being non-zero. Since we are considering small perturbations to this state, we should have

$$\eta \ll \phi\,. \tag{3.328}$$

[5] In the order-of-magnitude relations shown, absolute values are implicit.

Substituting (3.325) in (3.324), and using the Dirac representations of $\vec{\alpha}$ and β,

$$\Big(m\psi_s + i\dot{\psi}_s\Big)e^{-imt} = \left[-i\begin{pmatrix} & \vec{\sigma}\cdot\vec{D} \\ \vec{\sigma}\cdot\vec{D} & \end{pmatrix} + m\begin{pmatrix} I & \\ & -I \end{pmatrix} + e\Phi\right]\psi_s e^{-imt} \tag{3.329}$$

or in terms of ϕ and η,

$$\begin{cases} m\phi + i\,\dot{\phi} &= -i\,\vec{\sigma}\cdot\vec{D}\,\eta + (\quad m + e\Phi)\phi \\ m\eta + i\,\dot{\eta} &= -i\,\vec{\sigma}\cdot\vec{D}\,\phi + (-m + e\Phi)\eta \end{cases},$$

$$\rightarrow \quad \begin{cases} i\,\dot{\phi} &= -i\,\vec{\sigma}\cdot\vec{D}\,\eta + e\Phi\phi \\ \underbrace{i\,\dot{\eta}}_{\text{small}} &= -\underbrace{i\,\vec{\sigma}\cdot\vec{D}}_{\text{small}}\underbrace{\phi}_{\text{big}} - 2m\eta + \underbrace{e\Phi\eta}_{\text{small}} \end{cases}. \tag{3.330}$$

In the second equation, the dominant terms are $i\,\vec{\sigma}\cdot\vec{D}\,\phi$ and $2m\eta$; thus, we have

$$\eta = -\frac{i\,\vec{\sigma}\cdot\vec{D}}{2m}\phi\,. \tag{3.331}$$

Substituting this in the first equation of (3.330),

$$i\,\dot{\phi} = \left[-\frac{(\vec{\sigma}\cdot\vec{D})^2}{2m} + e\Phi\right]\phi\,. \tag{3.332}$$

Now, can we set $(\vec{\sigma}\cdot\vec{D})^2 = \vec{D}^2$? That would be correct if D_i and D_j commute; in this case, however, they do not commute since D_i contains ∂_i and $A^i(x)$ in it. Since D_i is not a matrix, it does commute with σ_i. Noting that

$$\{\sigma_i, \sigma_j\} = 2\,\delta_{ij}\,, \quad [\sigma_i, \sigma_j] = 2\,i\,\epsilon_{ijk}\sigma_k \quad \rightarrow \quad \sigma_i\sigma_j = \delta_{ij} + i\,\epsilon_{ijk}\sigma_k\,, \tag{3.333}$$

we can write $(\vec{\sigma}\cdot\vec{D})^2$ as

$$\begin{aligned} (\vec{\sigma}\cdot\vec{D})^2 &= (\sigma_i D_i)(\sigma_j D_j) = \sigma_i\sigma_j D_i D_i = (\delta_{ij} + i\,\epsilon_{ijk}\sigma_k)D_i D_j \\ &= \vec{D}^2 + i\,\sigma_k\,\epsilon_{ijk}D_i D_j\,. \end{aligned} \tag{3.334}$$

Since ∂_i applies to everything on its right including ϕ (unless otherwise indicated by angle brackets as in $\langle\partial_i A^j\rangle$), we have

$$\partial_i A^j\phi = A^j\partial_i\phi + \langle\partial_i A^j\rangle\phi \quad \rightarrow \quad \partial_i A^j = A^j\partial_i + \langle\partial_i A^j\rangle\,, \tag{3.335}$$

then,

$$\begin{aligned}
\epsilon_{ijk} D_i D_j &= \epsilon_{ijk} \quad \underbrace{(\partial_i - ieA^i)(\partial_j - ieA^j)} \\
&\qquad \underbrace{\partial_i\partial_j - e^2 A^i A^j}_{0\ (i \leftrightarrow j \text{ symmetric})} - ie(A^i\partial_j + \underbrace{\partial_i A^j}_{A^j\partial_i + \langle\partial_i A^j\rangle}) \\
&= -\,ie\,\epsilon_{ijk}[\underbrace{A^i\partial_j + A^j\partial_i}_{0\ (i \leftrightarrow j \text{ symmetric})} + \langle\partial_i A^j\rangle] \\
&= -ie\,\epsilon_{ijk}\langle\partial_i A^j\rangle = -ie\,\langle\vec{\nabla}\times\vec{A}\rangle_k = -ie\,B_k \qquad (3.336)
\end{aligned}$$

where we have used the ordinary definition of the magnetic field $\vec{B}$:

$$\vec{B} = \vec{\nabla}\times\vec{A}\,. \qquad (3.337)$$

Thus, we have

$$(\vec{\sigma}\cdot\vec{D})^2 = \vec{D}^2 + e\vec{\sigma}\cdot\vec{B}\,, \qquad (3.338)$$

and (3.332) becomes

$$i\,\dot{\phi} = \left[-\frac{(\vec{\nabla} - ie\vec{A})^2}{2m} - \frac{e}{2m}\vec{\sigma}\cdot\vec{B} + e\Phi\right]\phi\,. \qquad (3.339)$$

The first term is the coupling of orbital motion to the photon field, the last term is the potential energy due to the Φ field, and these are familiar terms from non-relativistic extension of the Schrödinger equation to include electromagnetic interaction. The second term, however, is a new term and it represents the potential energy of the electron magnetic moment in a magnetic field $\vec{B}$. This can be seen as follows: suppose that $\vec{B} = (0,0,B)$ and that ϕ represents a state with the spin in z-direction; namely, $\sigma_z\phi = \phi$. Then,

$$-\frac{e}{2m}\vec{\sigma}\cdot\vec{B} = -\frac{e}{2m}\sigma_z B \quad \rightarrow \quad -\frac{e}{2m}B \qquad (3.340)$$

Comparing this to the potential energy of a magnetic moment parallel to the magnetic field $-\mu B$, we identify the magnetic moment to be

$$\mu = \frac{e}{2m}\,. \qquad (3.341)$$

The gyro-magnetic ratio g is defined by

$$\mu = gs\mu_o\,, \qquad \mu_0 \equiv \frac{e}{2m}: \text{ Bohr magneton} \qquad (3.342)$$

where s is the absolute spin of the particle and the Bohr magneton μ_0 is the magnetic moment of a classical particle with charge e and one unit ($\hbar$) of orbital angular momentum for which the distributions of charge and mass are such that charge to mass ratio is the same everywhere. Since $s = 1/2$ for an electron, we have

$$g = 2\,. \tag{3.343}$$

The current experimental value is

$$g = 2 \times 1.001159652193\,(\underset{\substack{\uparrow\\ \text{uncertainty}}}{10})\,, \tag{3.344}$$

which is quite close to the value we have just obtained. The theoretical calculation including higher-order effects has a comparable accuracy and is consistent with the experiment. This is one of the most accurately measured and calculated quantities in physics, and the striking agreement demonstrates the correctness of quantum electrodynamics.

Does the derivation above mean that all spin-1/2 particles should have $g \sim 2.0$? Apparently not, since we know that the magnetic moment of proton is $g = 2.79$ and that of neutron is $g = -1.91$ (even though the above derivation indicates that a neutral particle should not have a magnetic moment since it does not couple to A_μ to begin with). The key step in the derivation was the minimal substitution (3.321) which does not in general work for particles which are not pointlike, and proton and neutron are known to be made of quarks and have size of order 1 fm (femto meter = 10^{-15} m, also called 'fermi').

3.10 High energy limit - massless case

We now examine the behavior of free Dirac fields at high energy, or equivalently, in the massless limit. We will see that the helicity + and the helicity − components decouple in such cases, satisfying separate equations of motion, where the *helicity* is defined to be the spin component in the direction of the momentum.

Define the 'right-handed' and 'left-handed' components by

$$\psi_R \overset{\text{def}}{\equiv} P_R\psi\,, \qquad \psi_L \overset{\text{def}}{\equiv} P_L\psi\,, \tag{3.345}$$

with

$$\boxed{P_R \overset{\text{def}}{\equiv} \frac{1+\gamma_5}{2}\,, \quad P_L \overset{\text{def}}{\equiv} \frac{1-\gamma_5}{2}\,.} \tag{3.346}$$

One can readily verify that they satisfy the properties of projection operators:

$$\begin{gathered} P_{R,L}^2 = P_{R,L}\,, \quad P_R + P_L = 1\,, \\ P_R P_L = P_L P_R = 0\,. \end{gathered} \tag{3.347}$$

In particular, $P_R + P_L = 1$, gives

$$\psi = \psi_R + \psi_L\,. \tag{3.348}$$

Also, note that $\{\gamma^\mu, \gamma_5\} = 0$ (3.135) and $\overline{\gamma_5} = -\gamma_5$ (3.283) lead to

$$P_R\gamma^\mu = \gamma^\mu P_L\,, \quad P_L\gamma^\mu = \gamma^\mu P_R\,, \tag{3.349}$$

$$\overline{P_R} = P_L\,, \quad \overline{P_L} = P_R\,. \tag{3.350}$$

Now, set the mass to zero in the Dirac equation to get

$$i\,\gamma^\mu\partial_\mu\psi = 0 \quad (m = 0)\,. \tag{3.351}$$

Then, we see that ψ_R also satisfies the massless Dirac equation:

$$i\gamma^\mu \underbrace{\partial_\mu\psi_R}_{P_R\partial_\mu\psi} = i\,\underbrace{\gamma^\mu P_R}_{P_L\gamma^\mu}\partial_\mu\psi = P_L\,\underbrace{i\gamma^\mu\partial_\mu\psi}_{0} = 0\,. \tag{3.352}$$

Simply exchanging L and R everywhere in the above, we also obtain $i\gamma^\mu\partial_\mu\psi_L = 0$. Thus, we see that ψ_R and ψ_L separately satisfy the massless Dirac equation:

$$i\partial\!\!\!/\psi_R = 0\,, \quad i\partial\!\!\!/\psi_L = 0 \quad (m = 0)\,. \tag{3.353}$$

If the mass is non-zero, then they cannot be separated:

$$(i\gamma^\mu\partial_\mu - m)\psi_R = \underbrace{i\gamma^\mu\partial_\mu\psi_R}_{P_L i\gamma^\mu\partial_\mu\psi} - m\,\underbrace{\psi_R}_{P_R\psi} \neq 0\,, \tag{3.354}$$

where if the mass term also acquired P_L, then it would have been factored out as $P_L(i\gamma^\mu\partial_\mu - m)\psi = 0$.

One important property of $\psi_{R,L}$ is that they do not mix under a proper and orthochronous transformation S. In fact, since $[\gamma_5, S] = 0$ (3.138), we have

$$[P_{R,L}, S] = \Big[\frac{1 \pm \gamma_5}{2}, S\Big] = 0\,. \tag{3.355}$$

Then, using $\psi'(x') = S\psi(x)$, we find

$$\psi'_R(x') \equiv P_R\psi'(x') = P_R S\psi(x) = SP_R\psi(x) = S\psi_R(x)\,, \tag{3.356}$$

and similarly for ψ_L. Thus, ψ_R and ψ_L transform separately under a proper and orthochronous transformation S:

$$\psi'_R(x') = S\psi_R(x)\,, \quad \psi'_L(x') = S\psi_L(x)\,. \tag{3.357}$$

This holds even if the mass is non-zero.

The above discussion is independent of representation. The situation, however, becomes simpler in the Weyl representation (3.153) in which $P_{R,L}$ are written as

$$\gamma_5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad \to \quad P_R = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad P_L = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \quad (\text{Weyl}), \tag{3.358}$$

which means that P_R filters out the top two components and P_L filters out the bottom two components. Namely, in the Weyl representation the top half and the bottom half of the spinor transform independently under S. Or equivalently, the matrix S becomes block diagonal:

$$S = \left(\begin{array}{c|c} \begin{matrix} \cdot & \cdot \\ \cdot & \cdot \end{matrix} & 0 \\ \hline 0 & \begin{matrix} \cdot & \cdot \\ \cdot & \cdot \end{matrix} \end{array}\right) \quad : \text{ proper and orthochronous (Weyl)}. \tag{3.359}$$

Therefore, in order to represent all *proper and orthochronous* transformations, that is, to reflect the product rule of the space-time Lorentz transformations, one needs only 2×2 matrices. In fact, one could use the 2×2 matrices that act on, say, the top half of the spinor as the representation. In this sense, the 4×4 representation of proper and orthochronous Lorentz transformations is said to be *reducible*. When the parity transformation $S_P = \gamma^0$ is included, however, the 2×2 representation is not enough. This can be easily seen by the expression of γ^0 in the Weyl representation (3.153) which mixes the top half and the bottom half of the spinor.

We called ψ_R and ψ_L as the right-handed and left-handed components because in the massless limit the fermion have helicity $+$ and $-$, respectively. Now, we will show that it is indeed the case. We take the spin quantization axis $\vec{s}$ to be the direction of the boost under which a rest mass m will acquire a momentum $\vec{p}$:

$$\vec{s} = \hat{p} \equiv \frac{\vec{p}}{|\vec{p}|}\,. \tag{3.360}$$

Noting that in the rest frame, we have $s^0 = 0$ and $s_\parallel = 1$, where $s_\parallel$ is the component of $\vec{s}$ along the direction of boost, the spin 4-vector s^μ in the boosted frame is given by

$$\begin{pmatrix} s^0 \\ s_\parallel \end{pmatrix} = \begin{pmatrix} \gamma & \eta \\ \eta & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \eta \\ \gamma \end{pmatrix} = \begin{pmatrix} |\vec{p}|/m \\ E/m \end{pmatrix}, \quad \vec{s}_\perp = 0\,, \tag{3.361}$$

where γ, η are related to $E, \vec{p}$ as before (3.228), and E and $\vec{p}$ are the energy and momentum of an electron or a positron (with $E \geq 0$).

In the massless limit (or high energy limit) we have $E = |\vec{p}|$, and s^μ becomes proportional to p^μ:

$$s^\mu = \Big(\frac{|\vec{p}|}{m}, \frac{E}{m}\hat{p}\Big) \quad \stackrel{E\to\infty}{\to} \quad s^\mu = \Big(\frac{E}{m}, \frac{|\vec{p}|}{m}\hat{p}\Big) = \frac{p^\mu}{m}\,. \tag{3.362}$$

Since $\not{p}/m$ is $+1$ or -1 depending on whether the state is an electron or a positron (3.251), γ_5 is then equal to the spin operator $\gamma_5\not{s}$ (with $\vec{s}=\hat{p}$) up to a sign:

$$\gamma_5\not{s} = \gamma_5 \frac{\not{p}}{m} = \begin{cases} \gamma_5 & \text{(electron)} \\ -\gamma_5 & \text{(positron)} \end{cases} \qquad (\vec{s}=\hat{p}\,,\ m\to 0)\,. \tag{3.363}$$

Since the eigenvalue of $\gamma_5\not{s}$ correctly represents the spin component along $\hat{p}$ for both electron and positron solutions, we have

$$\begin{cases} \gamma_5 : \pm 1 \ \leftrightarrow \ \text{helicity } \pm \text{ for } e^- \ (u\text{'s}) \\ \gamma_5 : \pm 1 \ \leftrightarrow \ \text{helicity } \mp \text{ for } e^+ \ (v\text{'s}) \end{cases} \qquad (m\to 0)\,, \tag{3.364}$$

or equivalently,

$$\boxed{P_{R,L} \equiv \frac{1\pm\gamma_5}{2} : \begin{cases} \text{helicity } \pm \text{ projection operator for } e^- \ (u\text{'s}) \\ \text{helicity } \mp \text{ projection operator for } e^+ \ (v\text{'s}) \end{cases} \qquad (m\to 0)}\,. \tag{3.365}$$

Since γ_5 and any proper and orthochronous transformation S commute, once a state is an eigenstate of γ_5, one cannot change the eigenvalue of γ_5 by boosting it or rotating it:

$$\gamma_5\psi = \pm\psi \quad \to \quad \gamma_5(S\psi) = S\gamma_5\psi = \pm(S\psi)\,. \tag{3.366}$$

Namely, in the massless limit, one cannot change the value of helicity by boosting or rotating. This can be understood intuitively. For a rotation, it is plausible since the spin and the direction of motion rotate together, and thus the component of spin along the motion would stay the same. For a boost, a classical picture also works just fine. The only way to reverse the spin component along the motion is for the observer to move faster than the particle and overtake it. Then, the direction of momentum viewed by the observer will flip while the spin will stay the same, and thus the helicity will change its sign. In the massless limit, the particle will be moving at the speed of light, and thus it is impossible to overtake it.

As we have seen, it does not require the mass to be zero in order for ψ_R and ψ_L to transform independently under proper and orthochronous transformations. If the mass is non-zero, however, ψ_R and ψ_L do not correspond to helicity $+1$ and -1, respectively. The distinction between ψ_R and ψ_L, or what we called 'handedness', is usually referred to as the *chirality* regardless of the mass, while helicity is defined as the spin component along the direction of motion. Helicity $+(-)$ and right-handedness (left-handedness), however, are often used synonymously.