9 The Dirac Equation

Dirac wanted an equation first order in time derivatives and Lorentz covariant, so it had to be first order in spatial derivatives too. His starting point was to assume a Hamiltonian of the form,

$$H_D = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \beta m \tag{9.1}$$

where P_i are the three components of the momentum operator $\mathbf{p}$, and α_i and β are some unknown quantities, which, as will be seen below, cannot simply be commuting numbers. If the requirement that $H_D^2 = \mathbf{p}^2 + m^2$ is imposed, this implies that α_i and β must be interpreted as 4×4 matrices, as we shall discuss. The first step is to write the momentum operators explicitly in terms of their differential operators, using equation (2.10). Then the Dirac equation (2.9) becomes, using the Dirac Hamiltonian in equation (9.1),

$$i\,\frac{\partial \psi}{\partial t} = (-i\,\boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + \beta m)\psi \tag{9.2}$$

which is the position space Dirac equation. Remember that in field theory, the Dirac equation is the equation of motion for the field operator describing spin 1/2 fermions. In order for this equation to be Lorentz covariant, it will turn out that ψ cannot be a scalar under Lorentz transformations. In fact this will be precisely how the equation turns out to describe spin 1/2 particles. We will return to this below.

If ψ is to describe a free particle it must satisfy the Klein-Gordon equation so that it has the correct energy-momentum relation. This requirement imposes relationships among $\alpha_1, \alpha_2, \alpha_3$ and β. To see these, apply the operator on each side of equation (9.2) twice, i.e. iterate the equation,

$$-\frac{\partial^2 \psi}{\partial t^2} = [-\alpha^i \alpha^j \nabla^i \nabla^j - i\,(\beta\alpha^i + \alpha^i\beta) m \nabla^i + \beta^2 m^2]\psi$$

with an implicit sum over i and j from 1 to 3. The Klein-Gordon equation by comparison is

$$-\frac{\partial^2 \psi}{\partial t^2} = [-\nabla^i \nabla^i + m^2]\psi \tag{9.3}$$

If we do not assume that the α^i and β commute then the KG will be satisfied if

$$\begin{aligned} \alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij} \\ \beta\alpha_i + \alpha_i\beta &= 0 \\ \beta^2 &= 1 \end{aligned} \tag{9.4}$$

for $i, j = 1, 2, 3$. The α_i and β cannot be ordinary numbers, but it is possible to give them a realisation as matrices. In this case, ψ must be a multi-component *spinor* on which these matrices act.

From eq.(9.5) we can see that

$$\alpha_i^2 = \beta^2 = 1. \tag{9.5}$$

Hence if we consider the action of the operators on an eigenstate

$$\begin{aligned}\alpha_i|e_i\rangle &= \lambda|e_i\rangle \\ \alpha_i^2|e_i\rangle &= \lambda^2|e_i\rangle = |e_i\rangle,\end{aligned} \tag{9.6}$$

and hence $\lambda = \pm 1$ and similarly for β. We can also see that

$$\begin{aligned}\operatorname{Tr}\alpha_i &= \operatorname{Tr}\beta^2\alpha_i = \operatorname{Tr}\beta\alpha_i\beta \\ &= -\operatorname{Tr}\beta^2\alpha_i = -\operatorname{Tr}\alpha_i,\end{aligned} \tag{9.7}$$

and hence $\operatorname{Tr}\alpha_i = 0$, and again a similar argument for β. But the trace is equal to the sum of the eigenvalues. Hence, we need an equal number of $+1$ eigenvalues as -1 eigenvalues and so the matrices representing α_i and β must be even dimensional.

In two dimensions a natural set of matrices for the $\boldsymbol{\alpha}$ would be the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.8}$$

However, there is no other independent 2×2 matrix with the right properties for β, so the smallest number of dimensions for which the Dirac matrices can be realized is four. One choice is the *Dirac representation*:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.9}$$

Note that each entry above denotes a two-by-two block and that the 1 denotes the 2×2 identity matrix.

There is a theorem due to Pauli that states that all sets of matrices obeying the relations in (9.5) are equivalent. Since the Hermitian conjugates $\boldsymbol{\alpha}^\dagger$ and $\beta^\dagger$ clearly obey the relations, you can, by a change of basis if necessary, assume that $\boldsymbol{\alpha}$ and β are Hermitian. All the common choices of basis already have this property. Furthermore, we would like α_i and β to be Hermitian so that the Dirac Hamiltonian (9.1) is Hermitian.

If we define

$$\rho = J^0 = \psi^\dagger\psi, \qquad \mathbf{J} = \psi^\dagger\boldsymbol{\alpha}\psi, \tag{9.10}$$

then it is a simple exercise using the Dirac equation to show that this satisfies the continuity equation $\partial_\mu J^\mu = 0$. We will see later that $(\rho, \mathbf{J})$ transforms as a four-vector, as it must. Note that ρ is now also positive definite.

9.1 Solutions to the Dirac Equation

We look for plane wave solutions of the form

$$\psi = \begin{pmatrix} \chi(\mathbf{p}) \\ \phi(\mathbf{p}) \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})} \tag{9.11}$$

where $\phi(\mathbf{p})$ and $\chi(\mathbf{p})$ are two-component spinors that depend on momentum $\mathbf{p}$ but are independent of $\mathbf{x}$. Using the Dirac representation of the matrices, and inserting the trial solution into the Dirac equation gives the pair of simultaneous equations

$$E\begin{pmatrix}\chi\\ \phi\end{pmatrix} = \begin{pmatrix} m & \boldsymbol{\sigma}\cdot\mathbf{p}\\ \boldsymbol{\sigma}\cdot\mathbf{p} & -m\end{pmatrix}\begin{pmatrix}\chi\\ \phi\end{pmatrix}. \tag{9.12}$$

There are two simple cases for which equation (9.12) can readily be solved, namely

1. $\mathbf{p} = 0$, $m \neq 0$, which might represent an electron in its rest frame.
2. $m = 0$, $\mathbf{p} \neq 0$, which might represent a massless neutrino.

We will consider case (2) later. For case (1), an electron in its rest frame, the equations (9.12) decouple and become simply,

$$E\chi = m\chi, \qquad E\phi = -m\phi. \tag{9.13}$$

So, in this case, we see that χ corresponds to solutions with $E = m$, while ϕ corresponds to solutions with $E = -m$. In light of our earlier discussions, we no longer need to recoil in horror at the appearance of these negative energy states.

The negative energy solutions persist for an electron with $\mathbf{p} \neq 0$ for which the solutions to equation (9.12) are

$$\phi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\,\chi, \qquad \chi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\,\phi. \tag{9.14}$$

Using the anti-commutation relations for the Pauli matrices it is easy to verify that $(\boldsymbol{\sigma}\cdot\mathbf{p})^2 = \mathbf{p}^2$. Using this we see that $E = \pm|\sqrt{\mathbf{p}^2 + m^2}|$. We write the positive energy solutions with $E = +|\sqrt{\mathbf{p}^2 + m^2}|$ as

$$\psi(x) = \begin{pmatrix}\chi\\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi\end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{9.15}$$

while the general negative energy solutions with $E = -|\sqrt{\mathbf{p}^2 + m^2}|$ are

$$\psi(x) = \begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\phi\\ \phi\end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{9.16}$$

for arbitrary constant ϕ and χ. Clearly when $\mathbf{p} = 0$ these solutions reduce to the positive and negative energy solutions discussed previously.

As an aside, it is interesting to see how Dirac coped with the negative energy states.

Dirac interpreted the negative energy solutions by postulating the existence of a "sea" of negative energy states. The vacuum or ground state has all the negative energy states full. An additional electron must now occupy a positive energy state since the Pauli exclusion principle forbids it from falling into one of the filled negative energy states. On promoting one of these negative energy states to a positive energy one, by supplying energy, an electron-hole pair is created, i.e. a positive energy electron and a hole in the negative energy sea. The hole is seen in nature as a positive energy positron. This was a radical new idea, and brought pair creation and antiparticles into physics.

The problem with Dirac's hole theory is that it does not work for bosons. Such particles have no exclusion principle to stop them falling into the negative energy states, releasing their energy.

We want to rewrite the solutions, eqs. (9.15) and (9.16), introducing the spinors $u_\alpha(s,\mathbf{p})$ and $v_\alpha(s,\mathbf{p})$. The label $\alpha \in \{1,2,3,4\}$ is a spinor index that often will be suppressed. Take the positive energy solution equation (9.15) and define

$$\sqrt{E+m}\begin{pmatrix} \chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s \end{pmatrix} e^{-ip\cdot x} \equiv u(s,p)e^{-ip\cdot x}. \tag{9.17}$$

For the negative energy solution of equation (9.16), change the sign of the energy, $E \to -E$, and the three-momentum, $\mathbf{p} \to -\mathbf{p}$, to obtain,

$$\sqrt{E+m}\begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s \\ \chi_s \end{pmatrix} e^{ip\cdot x} \equiv v(s,p)e^{ip\cdot x}. \tag{9.18}$$

In these two solutions E is now (and for the rest of the course) always positive and given by $E = (\mathbf{p}^2 + m^2)^{1/2}$. The argument s takes the values $1,2$, with

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{9.19}$$

For the simple case $\mathbf{p} = 0$ it seems natural to interpret χ_1 as the spin-up state and χ_2 as the spin-down state. Thus for $\mathbf{p} = 0$ the 4-component wave function has a very simple interpretation: the first two components describe electrons with spin-up and spin-down, while the second two components describe positrons with spin-up and spin-down. Thus we understand on physical grounds why the wave function had to have four components. The full discussion of spin, including the general case $\mathbf{p} \neq 0$, is slightly more involved and is considered in the next section.

The u-spinor solutions will correspond to particles and the v-spinor solutions to antiparticles. The role of the two χ's will become clear in the following section, where it will be shown that the two choices of s are spin labels. Note that each spinor solution depends on the three-momentum $\mathbf{p}$, and it is implicit that $p^0 = E$.

9.2 Spin

Now it is time to justify the statements we have been making that the Dirac equation describes spin-1/2 particles. The Dirac Hamiltonian in momentum space is given in equation (9.1) as

$$H_D = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m \tag{9.20}$$

and the orbital angular momentum operator is

$$\mathbf{L} = \mathbf{R} \times \mathbf{p}.$$

Evaluating the commutator of $\mathbf{L}$ with H_D,

$$\begin{aligned} [\mathbf{L}, H_D] &= [\mathbf{R}\times\mathbf{p}, \boldsymbol{\alpha}\cdot\mathbf{p}] \\ &= [\mathbf{R}, \boldsymbol{\alpha}\cdot\mathbf{p}]\times\mathbf{p} \\ &= i\boldsymbol{\alpha}\times\mathbf{p}, \end{aligned} \tag{9.21}$$

we see that the orbital angular momentum is not conserved (otherwise the commutator would be zero). We would like to find a *total* angular momentum $\mathbf{J}$ that *is* conserved, by adding an additional operator $\mathbf{S}$ to $\mathbf{L}$,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \quad [\mathbf{J}, H_D] = 0. \tag{9.22}$$

To this end, consider the three matrices,

$$\boldsymbol{\Sigma} \equiv \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} = -i\alpha_1\alpha_2\alpha_3\boldsymbol{\alpha}, \tag{9.23}$$

where the first equivalence is merely a definition of $\boldsymbol{\Sigma}$ and the last equality can be verified by an explicit calculation. The $\boldsymbol{\Sigma}/2$ have the correct commutation relations to represent angular momentum, since the Pauli matrices do, and their commutators with $\boldsymbol{\alpha}$ and β are,

$$[\boldsymbol{\Sigma}, \beta] = 0, \qquad [\Sigma_i, \alpha_j] = 2i\varepsilon_{ijk}\alpha_k. \tag{9.24}$$

From the relations in (9.24) we find that

$$[\boldsymbol{\Sigma}, H_D] = -2i\boldsymbol{\alpha} \times \mathbf{p}. \tag{9.25}$$

Comparing equation (9.25) with the commutator of $\mathbf{L}$ with H_D in equation (9.21), you see that

$$\left[\mathbf{L} + \frac{1}{2}\boldsymbol{\Sigma}, H_D\right] = 0,$$

and we can identify

$$\mathbf{S} = \frac{1}{2}\boldsymbol{\Sigma}$$

as the additional quantity that, when added to $\mathbf{L}$ in equation (9.22), yields a conserved total angular momentum $\mathbf{J}$. We interpret $\mathbf{S}$ as an angular momentum *intrinsic* to the particle. Now

$$\mathbf{S}^2 = \frac{1}{4}\begin{pmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} \end{pmatrix} = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and, recalling that the eigenvalue of $\mathbf{J}^2$ for spin j is $j(j+1)$, we conclude that $\mathbf{S}$ represents spin-1/2 and the solutions of the Dirac equation have spin-1/2 as promised. We worked in the Dirac representation of the matrices for convenience, but the result is necessarily independent of the representation.

Now consider the u-spinor solutions $u(s,p)$ of equation (9.17). Choose $\mathbf{p} = (0,0,p_z)$ and write

$$u_\uparrow \equiv u(1,p) = \begin{pmatrix} \sqrt{E+m} \\ 0 \\ \sqrt{E-m} \\ 0 \end{pmatrix}, \qquad u_\downarrow \equiv u(2,p) = \begin{pmatrix} 0 \\ \sqrt{E+m} \\ 0 \\ -\sqrt{E-m} \end{pmatrix}. \tag{9.26}$$

With these definitions, we get

$$S_z u_\uparrow = \frac{1}{2}u_\uparrow, \qquad S_z u_\downarrow = -\frac{1}{2}u_\downarrow.$$

So, these two spinors represent spin up and spin down along the z-axis respectively. For the v-spinors, with the same choice for $\mathbf{p}$, write,

$$v_\downarrow = v(1,p) = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix}, \qquad v_\uparrow = v(2,p) = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix}, \tag{9.27}$$

where now,

$$S_z v_\downarrow = \frac{1}{2} v_\downarrow, \qquad S_z v_\uparrow = -\frac{1}{2} v_\uparrow.$$

This apparently perverse choice of up and down for the v's is actually quite sensible when one realises that a negative energy electron carrying spin $+1/2$ backwards in time looks just like a positive energy positron carrying spin $-1/2$ forwards in time.

9.3 Normalization and Gamma Matrices

We have included a normalization factor $\sqrt{E+m}$ in our spinors. With this factor,

$$u^\dagger(r,p)u(s,p) = v^\dagger(r,p)v(s,p) = 2E\delta^{rs}, \tag{9.28}$$

as can easily be checked using eqs. (9.26) and (9.27) This corresponds to the standard relativistic normalization of $2E$ particles per unit volume (we shall justify this a bit later on). It also means that $u^\dagger u$ transforms like the time component of a 4-vector under Lorentz transformations as we will see in section 9.6.

There is a much more compact way of writing the Dirac equation, which requires that we get to grips with some more notation. Define the γ-*matrices*,

$$\gamma^0 = \beta, \qquad \boldsymbol{\gamma} = \beta\boldsymbol{\alpha}. \tag{9.29}$$

In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{9.30}$$

In terms of these, the relations between the $\boldsymbol{\alpha}$ and β in equation (9.5) can be written compactly as,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{9.31}$$

Combinations like $a_\mu\gamma^\mu$ occur frequently and are conventionally written as,

$$\not{a} = a_\mu\gamma^\mu = a^\mu\gamma_\mu,$$

pronounced "a slash." Note that γ^μ is not, despite appearances, a 4-vector. It just denotes a set of four matrices. However, the notation is deliberately suggestive, for when combined with

Dirac fields you can construct quantities that transform like vectors and other Lorentz tensors (see the next section).

Let us close this section by observing that using the γ-matrices the Dirac equation (9.2) becomes

$$(i\partial\!\!\!/ - m)\psi = 0, \tag{9.32}$$

or, in momentum space,

$$(p\!\!\!/ - m)\psi = 0. \tag{9.33}$$

The spinors u and v satisfy

$$(p\!\!\!/ - m)u(s,p) = 0, \tag{9.34}$$

$$(p\!\!\!/ + m)v(s,p) = 0, \tag{9.35}$$

since for $v(s,p)$, $E \to -E$ and $\mathbf{p} \to -\mathbf{p}$.

9.4 Lorentz Covariance

We want the Dirac equation (9.32) to preserve its form under Lorentz transformations eq. (1.1). We know that 4-vectors get their components mixed up by LT's, so we expect that the components of ψ might get mixed up too:

$$\psi(x) \to \psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}x') \tag{9.36}$$

where $S(\Lambda)$ is a 4×4 matrix acting on the spinor index of ψ. Note that the argument $\Lambda^{-1}x'$ is just a fancy way of writing x, i.e. each component of $\psi(x)$ is transformed into a linear combination of components of $\psi(x)$.

We now need to figure out what S is. The requirement is that the Dirac equation has the same form in any inertial frame. Thus, if we make a LT from our original frame into another ('primed') frame and write down the Dirac equation in this frame, it has to have the same form.

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0 \quad\longrightarrow\quad (i\gamma^\mu\partial'_\mu - m)\psi'(x') = 0, \tag{9.37}$$

where we used the fact that m is a scalar, i.e. $m' = m$. To determine S we rewrite the Dirac equation in the original frame in terms of the primed variables (just a mathematical substitution):

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = (i\gamma^\mu\Lambda^\sigma{}_\mu\partial'_\sigma - m)\psi(\Lambda^{-1}x') = 0. \tag{9.38}$$

where we used $\partial_\mu = \Lambda^\sigma{}_\mu\partial'_\sigma$. This last equality follows because

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial x'^\sigma}{\partial x^\mu}\frac{\partial}{\partial x'^\sigma} = \Lambda^\sigma{}_\mu\frac{\partial}{\partial x'^\sigma} \tag{9.39}$$

and equation (1.1) has been used in the last step. In the 'primed' frame, we have

$$(i\gamma^\mu\partial'_\mu - m)\psi'(x') = (i\gamma^\sigma\partial'_\sigma - m)S(\Lambda)\psi(\Lambda^{-1}x') = 0. \tag{9.40}$$

where we used the last identity in eq. (9.36). Multiplying eq. (9.38) by $S(\Lambda)$ and comparing the result with eq. (9.40) we see that we need $S(\Lambda)\Lambda^{\sigma}{}_{\mu}\gamma^{\mu} = \gamma^{\sigma}S(\Lambda)$. Thus, in order for the Dirac equation to be Lorentz invariant, $S(\Lambda)$ has to satisfy

$$\Lambda^{\sigma}{}_{\mu}\gamma^{\mu} = S^{-1}(\Lambda)\gamma^{\sigma}S(\Lambda) \tag{9.41}$$

We still haven't solved for S explicitly. We need to find an S that satisfies eq. (9.41). Since S depends on the LT, we first have to find a convenient parameterization of a LT and then express $S(\Lambda)$ in terms of these parameters. For an infinitesimal LT, it can be shown that,

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} - \varepsilon_{\rho\sigma}(g^{\rho\mu}\delta^{\sigma}{}_{\nu} - g^{\sigma\mu}\delta^{\rho}{}_{\nu}) \tag{9.42}$$

where $\varepsilon_{\rho\sigma}$ is an $\rho \leftrightarrow \sigma$ antisymmetric set of infinitesimal parameters and the pair (ρ, σ) label the six types of transformation, i.e. 3 boosts and 3 rotations. For example a boost along the z-axis corresponds to $\rho = 0$, $\sigma = 3$, since in this case,

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} - \varepsilon_{03}(g^{0\mu}\delta^{3}{}_{\nu} - g^{3\mu}\delta^{0}{}_{\nu}) = \begin{pmatrix} 1 & 0 & 0 & -\varepsilon_{03} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\varepsilon_{03} & 0 & 0 & 1 \end{pmatrix},$$

which is the usual matrix for an infinitesimal Lorentz boost, with $\beta = \varepsilon_{03}$ and $\gamma = 1$.

The combinations $(\rho, \sigma) = (0, 1)$ and $(0, 2)$ correspond to boosts along the x and y axes respectively. The remaining combinations, $(2, 3)$, $(3, 1)$ and $(1, 2)$, correspond to infinitesimal anti-clockwise rotations through an angle $\varepsilon_{\rho\sigma}$ about the x, y and z axes respectively. It's a nice exercise to check this out.

For an infinitesimal LT we are at liberty to write

$$S(\Lambda) = 1 + i\varepsilon_{\rho\sigma}s^{\rho\sigma}, \tag{9.43}$$

which is nothing but a definition of the set of matrices $s^{\rho\sigma}$. Our task is to determine these matrices. To do this, substitute the expression for S, eq. (9.43), into eq. (9.41) (and remember that $S^{-1}(\Lambda) = 1 - i\varepsilon_{\rho\sigma}s^{\rho\sigma}$. After some algebra, we can convince ourselves that the solution is

$$s^{\rho\sigma} = \frac{i}{4}\,[\gamma^{\rho}, \gamma^{\sigma}] \equiv \frac{1}{2}\,\sigma^{\rho\sigma}. \tag{9.44}$$

Here, I have taken the opportunity to define the matrix $\sigma^{\rho\sigma}$. Thus S is given explicitly in terms of γ-matrices for a general LT.

Now that we now how ψ transforms we can find quantities that are Lorentz invariant, or transform as vectors or tensors under LT's. To this end, we will find it useful to introduce the Dirac adjoint. The Dirac adjoint $\bar{\psi}$ of a spinor ψ is defined by

$$\bar{\psi} \equiv \psi^{\dagger}\gamma^{0} \tag{9.45}$$

With the help of

$$S^{\dagger}(\Lambda)\gamma^{0} = \gamma^{0}S^{-1}(\Lambda) \tag{9.46}$$

we see that $\bar{\psi}$ transforms as follows under LT's:

$$\bar{\psi} \to \bar{\psi}' = \bar{\psi} S^{-1}(\Lambda) \tag{9.47}$$

Combining the transformation properties of ψ and $\bar{\psi}$ in equations (9.36) and (9.47) we see that the bilinear $\bar{\psi}\psi$ is Lorentz invariant. In section 9.6 we will consider the transformation properties of general bilinears.

Let me close this section by recasting the spinor normalization equations (9.28) in terms of Dirac inner products. The conditions become

$$\begin{aligned} \bar{u}(r,p)u(s,p) &= 2m\delta^{rs} \\ \bar{u}(r,p)v(s,p) &= \bar{v}(r,p)u(s,p) = 0 \\ \bar{v}(r,p)v(s,p) &= -2m\delta^{rs} \end{aligned} \tag{9.48}$$

where, in analogy to eq (9.45), we defined $\bar{u} \equiv u^\dagger \gamma^0$ and $\bar{v} \equiv v^\dagger \gamma^0$.

9.5 Parity

In the next section we are going to construct quantities bilinear in ψ and $\bar{\psi}$, and classify them according to their transformation properties under LT's. We normally use LT's which can be obtained by a continuous deformation of the identity transformation (i.e. by lots of little transformations)[1]. These are the rotations and boosts, since the angle of the rotation or the boost parameter can be made arbitrarily small. Thereby, the LT approaches the identity transformation. This class of LT is often referred to as proper LT. However, the full Lorentz group consists not only of the proper transformations but also includes the discrete operations of parity (space inversion), P, and time reversal, T:

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

LT's satisfy $\Lambda^T g \Lambda = g$, so taking determinants shows that $\det \Lambda = \pm 1$. Proper LT's have determinant 1, since the identity does, but the P and T operations have determinant -1.

Let us now find the action of parity on the Dirac wave function and determine the wave function ψ_P in the parity-reversed system. According to the discussion of the previous section, we need to find a matrix S satisfying

$$S^{-1}\gamma^0 S = \gamma^0, \qquad S^{-1}\gamma^i S = -\gamma^i.$$

Using the properties of the γ-matrices we see that $S = S^{-1} = \gamma^0$ is an acceptable solution (Clearly one could multiply γ^0 by a phase and still have an acceptable definition for the parity transformation.), from which it follows that the wave function ψ_P is

$$\psi_P(t, -\mathbf{x})(\equiv \psi_P(x_P)) = \gamma^0 \psi(t, \mathbf{x})(\equiv S\psi(x)). \tag{9.49}$$

[1] Indeed in the last section we considered LT's very close to the identity in equation (9.42)

Since

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

the u-spinors and v-spinors at rest have opposite eigenvalues, corresponding to particle and antiparticle having opposite *intrinsic* parities.

9.6 Bilinear Covariants

Now, as promised, we will construct and classify the bilinears. These are useful for defining quantities with particular properties under Lorentz transformations, and appearing in Lagrangians for fermion field theories.

To begin, note that by forming products of the γ-matrices it is possible to construct 16 linearly independent 4×4 matrices. Any constant 4×4 matrix can then be decomposed into a sum over these basis matrices. In equation (9.44) we have defined

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu],$$

and now it is convenient to define

$$\gamma^5 \equiv \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{9.50}$$

where the last equality is valid in the Dirac representation. This new matrix satisfies

$$\gamma^{5\dagger} = \gamma^5, \qquad \{\gamma^5, \gamma^\mu\} = 0, \qquad (\gamma^5)^2 = 1.$$

It also changes sign under improper Lorentz transformations, e.g. $\gamma^5\gamma^0 = -\gamma^0\gamma^5$ so multiplying through by $(\gamma^0)^{-1} \equiv \gamma^0$ and using the fact that $\gamma^0 \equiv P$, where P is the parity operator, we have $P^{-1}\gamma^5 P = -\gamma^5$.

Now, the set of 16 matrices

$$\{1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5, \sigma^{\mu\nu}\}$$

form a basis for γ-matrix products. There are 16 matrices since there is 1 unit matrix, 1 γ^5 matrix, 4 γ^μ matrices and 4 $\gamma^\mu\gamma^5$ matrices, and 6 $\sigma^{\mu\nu}$ matrices (see equation (9.44) for the definition of $\sigma^{\mu\nu}$).

Using the transformations of ψ and $\bar{\psi}$ from eqs. (9.36) and (9.47), together with the transformation of γ^μ in eq. (9.41), the 16 fermion bilinears and their transformation properties can be written as follows:

$$\begin{array}{rcll}
\bar{\psi}\psi & \rightarrow & \bar{\psi}\psi & \text{S scalar} \\
\bar{\psi}\gamma^5\psi & \rightarrow & \det(\Lambda)\,\bar{\psi}\gamma^5\psi & \text{P pseudoscalar} \\
\bar{\psi}\gamma^\mu\psi & \rightarrow & \Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\psi & \text{V vector} \\
\bar{\psi}\gamma^\mu\gamma^5\psi & \rightarrow & \det(\Lambda)\,\Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\gamma^5\psi & \text{A axial vector} \\
\bar{\psi}\sigma^{\mu\nu}\psi & \rightarrow & \Lambda^\mu{}_\lambda\Lambda^\nu{}_\sigma\bar{\psi}\sigma^{\lambda\sigma}\psi & \text{T tensor}
\end{array} \tag{9.51}$$

In particular we note that

$$\bar{\psi}\gamma^{\mu}\psi = \psi^{\dagger}\gamma^{0}\gamma^{\mu}\psi = (\psi^{\dagger}\psi, \psi^{\dagger}\boldsymbol{\alpha}\psi) \tag{9.52}$$

which is our previous definition eq. (9.10) of the current 4–vector J^{μ}, i.e. we now see that it is really a 4–vector.

9.7 Charge Conjugation *Details of spinor transformation are Non-examinable*

There is one more discrete invariance of the Dirac equation in addition to parity. It is charge conjugation, which takes you from particle to antiparticle and vice versa. For scalar fields the symmetry is just complex conjugation, but in order for the charge conjugate Dirac field to remain a solution of the Dirac equation, you have to mix its components as well:

$$\psi \to \psi_C = C\bar{\psi}^T.$$

Here $\bar{\psi}^T = \gamma^{0T}\psi^*$ and C is a matrix satisfying the condition

$$C\gamma_{\mu}^{T}C^{-1} = -\gamma_{\mu} \quad (C^{-1} = C^{\dagger}).$$

One can show that this transformation leaves the Dirac Equation and $\bar{\psi}\psi$ invariant.

In the Dirac representation,

$$C = i\gamma^2\gamma^0 = \begin{pmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}.$$

When Dirac wrote down his equation everybody thought parity and charge conjugation were exact symmetries of nature, so invariance under these transformations was essential. Now we know that neither of them, nor the combination *CP*, is respected by the standard electroweak model.

9.8 Massless Fermions

Although there is now excellent evidence that neutrinos have masses, it is likely that these are very small. Let us look, therefore, at solutions of the Dirac equation with $m = 0$, on the basis that this will be an extremely good approximation for many situations. In fact it is also often used for electrons and up and down quarks, which have very small masses compared to the usual scale of the scattering processes we consider.

From equation (9.12) we have in this case

$$E\phi = \boldsymbol{\sigma}\cdot\mathbf{p}\,\chi, \qquad E\chi = \boldsymbol{\sigma}\cdot\mathbf{p}\,\phi. \tag{9.53}$$

These equations can easily be decoupled by taking the linear combinations and defining in a suggestive way the two component spinors ν_L and ν_R,

$$\nu_R \equiv \chi + \phi, \quad \nu_L \equiv \chi - \phi \tag{9.54}$$

which leads to

$$E\nu_R = \boldsymbol{\sigma}\cdot\mathbf{p}\,\nu_R, \qquad E\nu_L = -\boldsymbol{\sigma}\cdot\mathbf{p}\,\nu_L. \tag{9.55}$$

Since $E = |\mathbf{p}|$ for massless particles, these equations may be written

$$\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}\nu_L = -\nu_L, \qquad \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}\nu_R = \nu_R \tag{9.56}$$

Now, $\frac{1}{2}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}$ is known as the *helicity* operator (i.e. it is the spin operator projected in the direction of motion of the momentum of the particle). We see that the ν_L corresponds to solutions with negative helicity, while ν_R corresponds to solutions with positive helicity. In other words ν_L describes a left-handed neutrino while ν_R describes a right-handed neutrino, and each type of neutrino is described by a two-component spinor.

Under parity transformations $\boldsymbol{\sigma} \to \boldsymbol{\sigma}$ (like $\mathbf{R}\times\mathbf{p}$), $\mathbf{p} \to -\mathbf{p}$, therefore $\boldsymbol{\sigma}\cdot\mathbf{p} \to -\boldsymbol{\sigma}\cdot\mathbf{p}$, i.e. the spinors transform into each other:

$$\nu_L \leftrightarrow \nu_R. \tag{9.57}$$

So a theory which involves only ν_L without ν_R (such as the standard model) manifestly violates parity.

Although massless neutrinos can be described very simply using two component spinors as above, they may also be incorporated into the four-component formalism as follows. From equation (9.2) we have, in momentum space,

$$|\mathbf{p}|\psi = \boldsymbol{\alpha}\cdot\mathbf{p}\,\psi.$$

For such a solution,

$$\gamma^5\psi = \gamma^5\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{|\mathbf{p}|}\psi = 2\frac{\mathbf{S}\cdot\mathbf{p}}{|\mathbf{p}|}\psi,$$

using the spin operator $\mathbf{S} = \frac{1}{2}\boldsymbol{\Sigma} = \frac{1}{2}\gamma^5\boldsymbol{\alpha}$, with $\boldsymbol{\Sigma}$ defined in equation (9.23). But $\mathbf{S}\cdot\mathbf{p}/|\mathbf{p}|$ is the projection of spin onto the direction of motion, i.e. the helicity, and is equal to $\pm 1/2$. Thus $(1+\gamma^5)/2$ projects out the neutrino with helicity $1/2$ (right handed) and $(1-\gamma^5)/2$ projects out the neutrino with helicity $-1/2$ (left handed):

$$\frac{(1+\gamma^5)}{2}\psi \equiv \psi_R, \quad \frac{(1-\gamma^5)}{2}\psi \equiv \psi_L, \tag{9.58}$$

define the four-component spinors ψ_R and ψ_L. These are used extensively in practice in the Standard Model, as the weak force only couples to ψ_L, and not to ψ_R.