Copyright © 2021 by Robert G. Littlejohn

Physics 221B
Academic Year 2021–22
Notes 48
Covariance of the Dirac Equation†

1. Introduction

In accordance with the principle of relativity, physics must "look the same" in all Lorentz frames. This means that physical theories that are consistent with the principle of relativity must have the same form in all Lorentz frames, that is, they must be covariant. In this set of notes we examine the covariance of the Dirac equation.

In these notes we mainly deal with the Dirac wave function ψ, which is understood to be a four-component spinor. But occasional reference is made to the scalar Klein-Gordon wave function, which will be denoted by ψ_{KG}.

2. Covariant Form of the Dirac Equation

We begin with the free particle. The Dirac equation, as developed in Notes 46, is

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar c\,\boldsymbol{\alpha}\cdot\nabla\psi + mc^2\beta\,\psi. \tag{1}$$

The operator $\partial/\partial t$ on the left-hand side is not a Lorentz scalar, because the time t represents just one component of the 4-vector $x^\mu = (ct, \mathbf{x})$. The Dirac equation, as written, is not manifestly Lorentz covariant. Let us bring all the derivatives over to one side, and write the Dirac equation as

$$i\hbar c\Big(\frac{\partial\psi}{\partial(ct)} + \boldsymbol{\alpha}\cdot\nabla\psi\Big) = mc^2\beta\,\psi. \tag{2}$$

To put this into covariant form, we multiply through by β, using $\beta^2 = 1$, to obtain

$$i\hbar c\Big(\beta\frac{\partial\psi}{\partial(ct)} + \beta\boldsymbol{\alpha}\cdot\nabla\psi\Big) = mc^2\psi. \tag{3}$$

The constant operator on the right-hand side, mc^2, is a Lorentz scalar, while the operators

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \Big(\frac{\partial}{\partial(ct)}, \nabla\Big) \tag{4}$$

that appear on the left-hand side transform as a covariant vector (see Appendix E). Therefore we guess that the coefficients that multiply ∂_μ on the left-hand side must transform as a contravariant 4-vector, so that the entire operator on the left-hand side will be a Lorentz scalar.

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2122/221.html`.

To bring this out notationally, we define

$$\gamma^0 = \beta, \qquad \gamma^i = \beta\alpha_i, \tag{5}$$

for $i = 1, 2, 3$, so that the free particle Dirac equation can be written,

$$i\hbar\gamma^\mu \frac{\partial\psi}{\partial x^\mu} = mc\,\psi, \tag{6}$$

after cancelling a factor of c. We have written the four matrices defined in Eq. (5) as γ^μ, $\mu = 0, 1, 2, 3$, which we can think of as a 4-vector of Dirac matrices, much as the Pauli matrices $\boldsymbol{\sigma}$ constitute a 3-vector of matrices. If we introduce the covariant momentum operators,

$$p_\mu = i\hbar\frac{\partial}{\partial x^\mu} = i\hbar\Big(\frac{\partial}{\partial(ct)}, \nabla\Big) \tag{7}$$

[see Eq. (45.20)], then the free particle Dirac equation takes on the suggestive form,

$$(\gamma^\mu p_\mu - mc)\psi = 0. \tag{8}$$

The notation suggests that $\gamma^\mu p_\mu$ is a Lorentz scalar, but we will not have proven that until we see how and in what sense γ^μ constitutes a 4-vector. To do that, we will have to show that it transforms as a 4-vector under Lorentz transformations. Moreover, since γ^μ is a 4-vector of 4×4 matrices, not ordinary numbers, its transformation law will not be the same as that of the 4-vectors encountered in classical relativity theory, as discussed in Appendix E. Instead, as we shall see, it transforms by a four-dimensional generalization of the definition of a vector operator in quantum mechanics, which is discussed in Sec. 19.4.

To introduce the coupling with the electromagnetic field, we use the covariant version of the minimal coupling prescription (46.16),

$$p_\mu \to p_\mu - \frac{q}{c}A_\mu, \tag{9}$$

where A^μ is the 4-vector potential,

$$A^\mu = \begin{pmatrix} \Phi \\ \mathbf{A} \end{pmatrix}, \qquad A_\mu = \begin{pmatrix} \Phi \\ -\mathbf{A} \end{pmatrix}. \tag{10}$$

See also Sec. B.17. This puts the Dirac equation for a particle interacting with the electromagnetic field into the form,

$$\Big(\gamma^\mu p_\mu - \frac{q}{c}\gamma^\mu A_\mu - mc\Big)\psi = 0. \tag{11}$$

Contractions between γ^μ and ordinary 4-vectors such as A_μ, or 4-vectors of operators such as p_μ, are very common in the Dirac theory. Here is some notation for such contractions. Let X^μ be any 4-vector. Then we define

$$\not{X} = \gamma^\mu X_\mu, \tag{12}$$

which is called the *Feynman slash*. When you see the Feynman slash, you must recognize that it is a 4×4 Dirac matrix, with components that are numbers, possibly with a space-time dependence, as in $\not{A}$, or operators, as in $\not{p}$. In terms of this notation, the Dirac equation becomes

$$\Big(\not{p} - \frac{q}{c}\not{A} - mc\Big)\psi = 0. \tag{13}$$

This is regarded as the covariant version of the Dirac equation. It is equivalent to the Hamiltonian version, $i\hbar\partial\psi/\partial t = H\psi$, with H given by Eq. (46.17).

The covariant version of the Dirac equation (13) produces the Pauli equation (46.1) in the nonrelativistic limit with $g = 2$, as we showed in Sec. 46.9. And yet it is simpler in form than the Pauli equation, in spite of the extra notation. In fact, as we will see in a later set of notes, Eq. (13) contains even more physics than the Pauli equation, for if it is expanded to fourth order in v/c it produces all the fine structure corrections we saw in Notes 25.

3. The Matrices γ^μ

The matrices γ^μ, defined by Eq. (5), constitute an alternative version of the Dirac matrices $\boldsymbol{\alpha}$ and β, useful when we wish to reveal the covariant aspects of the Dirac equation. All the properties of the $\boldsymbol{\alpha}$ and β matrices can be converted into properties of the matrices γ^μ. Here we list some of them.

First, there are the values of the matrices. In the usual Dirac-Pauli representation, they are

$$\gamma^0 = \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \gamma^i = \beta\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \qquad \text{(Dirac-Pauli)} \tag{14}$$

while in the Weyl representation they are

$$\gamma^0 = \beta = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \qquad \gamma^i = \beta\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}. \qquad \text{(Weyl)} \tag{15}$$

Next there are the Hermiticity properties. Recall that $\boldsymbol{\alpha}$ and β are Hermitian. This implies that γ^0 is Hermitian, while γ^i, $i = 1, 2, 3$, are anti-Hermitian. This is easily proved using the properties of the $\boldsymbol{\alpha}$ and β matrices, for example,

$$\begin{aligned} (\gamma^0)^\dagger &= \beta^\dagger = \beta = \gamma^0, \\ (\gamma^i)^\dagger &= (\beta\alpha_i)^\dagger = \alpha_i\beta = -\beta\alpha_i = -\gamma^i, \end{aligned} \tag{16}$$

where we have used the anticommutation relation, $\{\alpha_i, \beta\} = 0$ and $\beta^2 = 1$.

Finally, there are the anticommutation properties of the γ^μ, which are easily derived from those of $\boldsymbol{\alpha}$ and β. Explicitly, we have

$$\begin{aligned} &\{\gamma^0, \gamma^0\} = \{\beta, \beta\} = 2, \\ &\{\gamma^0, \gamma^i\} = \{\beta, \beta\alpha_i\} = \beta^2\alpha_i + \beta\alpha_i\beta = \alpha_i - \alpha_i = 0, \\ &\{\gamma^i, \gamma^j\} = \{\beta\alpha_i, \beta\alpha_j\} = \beta\alpha_i\beta\alpha_j + \beta\alpha_j\beta\alpha_i = -\alpha_i\alpha_j - \alpha_j\alpha_i = -2\delta_{ij}, \end{aligned} \tag{17}$$

where again we use the anticommutation relations (46.8).

We can summarize the anticommutation relations of the matrices γ^μ by writing,

$$\boxed{\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}.} \tag{18}$$

Here it is understood that the right hand side multiplies the identity matrix, usually denoted by 1 in the Dirac theory. This equation looks covariant, and indeed we will see that it is, once we have worked out the transformation properties of the matrices γ^μ. Equation (18) is a compact, covariant form of the Dirac algebra first presented in Eq. (46.8), which as we recall arose from the demand that the relativisitc energy-momentum relations should be satisfied by free particle solutions of the Dirac equation. It is used frequently in the Dirac theory.

We make a remark on the positions of the spatial index i in γ^i and α_i. We use an upper (superscript) position on γ^i because this is seen as the spatial components of a contravariant vector γ^μ (it is $\gamma^{\mu\to i}$, in the notation discussed in Sec. E.20). On the other hand, $\boldsymbol{\alpha}$ is not the spatial components of any 4-vector (there is no α^0), rather $\boldsymbol{\alpha}$ is seen as an object that is intrinsically a 3-vector, so we just use lower indices for its components. The case of the velocity vector is similar; we write simply v_i for the usual components of the velocity vector $\mathbf{v}$, since this is never the spatial part of a 4-vector [see the comment after Eq. (E.9), and note that in the Dirac theory, $\mathbf{v} = c\boldsymbol{\alpha}$].

4. Transformation of the Dirac Wave Function

We turn to the transformation properties of the Dirac wave function ψ under Lorentz transformations. We begin by guessing the form of the transformation law, by analogy with the transformation laws for different types of fields under ordinary rotations, as well as certain types of fields under Lorentz transformations.

The transformation laws for scalar and vector fields in 3-dimensional space under ordinary rotations was reviewed in Sec. 47.11. If $S(\mathbf{x})$ is a scalar field and R specifies a rotation, then the rotated field S' is given in terms of the original field S by

$$S'(\mathbf{x}) = S\big(\mathsf{R}^{-1}\mathbf{x}\big). \tag{19}$$

See Eq. (47.85). This transformation law applies in particular to the wave function of a spin-0 particle, as explained in Sec. 15.2. See Eq. (15.13). Next, if $\mathbf{E}(\mathbf{x})$ is a vector field, then the rotated field is given by

$$\mathbf{E}'(\mathbf{x}) = \mathsf{R}\mathbf{E}\big(\mathsf{R}^{-1}\mathbf{x}\big). \tag{20}$$

See Eq. (47.87). Next, the case of the wave function of a particle of spin s was treated in Prob. 18.1(a). If $\psi(\mathbf{x})$ is the $2s+1$-component spinor of a particle of spin s and R is a rotation, then the rotated wave function is

$$\psi'(\mathbf{x}) = D^s(\mathsf{R})\psi\big(\mathsf{R}^{-1}\mathbf{x}\big) \tag{21}$$

(this is the solution to the problem). Here D^s is the $(2s+1)\times(2s+1)$ rotation matrix, as explained in Notes 13, and it is understood that this matrix multiplies the spinor ψ. The rotation matrices satisfy the representation property,

$$D^s(\mathsf{R}_1)D^s(\mathsf{R}_2) = \pm D^s(\mathsf{R}_1\mathsf{R}_2), \tag{22}$$

where the $\pm$ sign only applies in the case of half-integer s. See Eqs. (13.72)–(13.74) and the discussion of double-valued representations in Sec. 12.9.

We see that in all cases (scalar, vector, spinor), the rotated field at point $\mathbf{x}$ is expressed in terms of the original field at the inverse rotated point $\mathsf{R}^{-1}\mathbf{x}$, while the value of the field is transformed by the appropriate rotation matrix (no rotation at all for a scalar, the classical rotation matrix R for a vector field, and the D^s-matrix for a spinor field of spin s).

In special relativity we consider fields on space-time, and subject them to Lorentz transformations. Letting x stand for the four space-time coordinates x^μ or $(\mathbf{x},t)$, the transformation rule for a scalar field under a Lorentz transformation $\Lambda^\mu{}_\nu$ is

$$S'(x) = S\big(\Lambda^{-1}x\big). \tag{23}$$

See Eq. (47.88). The Klein-Gordon wave function ψ_{KG} is a scalar field that represents a spin-0 particle, and it transforms according to this rule,

$$\psi'_{\rm KG}(x) = \psi_{\rm KG}\big(\Lambda^{-1}x\big). \tag{24}$$

In the case of a contravariant vector field X^μ, the transformation law is

$$X'^\mu(x) = \Lambda^\mu{}_\nu\, X^\nu\big(\Lambda^{-1}x\big). \tag{25}$$

In these cases the Lorentz transformed field at space-time point x is expressed in terms of the original field at space-time point $\Lambda^{-1}x$, while the value of the field transforms by the appropriate matrix (none at all for a scalar, and $\Lambda^\mu{}_\nu$ for a contravariant vector field).

These examples suggest that the Dirac wave function ψ should transform under Lorentz transformations according to

$$\psi'(x) = D(\Lambda)\psi\big(\Lambda^{-1}x\big), \tag{26}$$

where $D(\Lambda)$ is some kind of 4×4 matrix that acts on the spin part of ψ. We know that the Dirac equation is giving us the relativistic quantum mechanics of a spin-$\frac{1}{2}$ particle, and we know that when the nonrelativistic wave function of a spin-$\frac{1}{2}$ particle is rotated the matrix $D^{1/2}$ appears as in Eq. (21). Since rotations are special cases of Lorentz transformations, we must expect some kind of spin matrix as in Eq. (26) when the Dirac wave function is subjected to a Lorentz transformation. Moreover, we expect that the matrices $D(\Lambda)$ should satisfy the representation property,

$$D(\Lambda_1)D(\Lambda_2) = \pm D(\Lambda_1\Lambda_2), \tag{27}$$

since this means that if we apply a Lorentz transformation Λ_2 to a Dirac wave function, then a second one Λ_1, the effect is the same as applying the single Lorentz transformation $\Lambda_1\Lambda_2$. That

is, the matrices $D(\Lambda)$ should form a representation of the Lorentz group. We include a $\pm$ sign in Eq. (27) because we know this sign is necessary in the case of ordinary rotations of a spin-$\frac{1}{2}$ particle, and because rotations are special cases of Lorentz transformations.

5. The Space-Time Part of the Transformation

Let us concentrate first on the space-time part of the transformation (26), to see if it makes sense. Consider an eigenfunction of energy and momentum, that is, an eigenfunction of the free-particle Dirac equation. We write this as

$$\psi(x) \sim e^{i(\mathbf{p}\cdot\mathbf{x}-Et)/\hbar}, \tag{28}$$

where $\sim$ means that we are concentrating on the space-time dependence of the wave function and ignoring the spin. (If we were talking about the Klein-Gordon wave function then we could replace $\sim$ by $=$.) We can put this into covariant form by using

$$x^\mu = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}, \qquad p^\mu = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \tag{29}$$

so that

$$p_\mu x^\mu = Et - \mathbf{p}\cdot\mathbf{x}, \tag{30}$$

and

$$\psi(x) \sim \exp\Big(-\frac{i}{\hbar}p_\mu x^\mu\Big) = \exp\Big[-\frac{i}{\hbar}(p\cdot x)\Big], \tag{31}$$

where we use the notation of Eq. (47.104) for the scalar product of two 4-vectors. Now subjecting this to a Lorentz transformation specified by Λ and using the transformation law (26), this becomes

$$\psi'(x) \sim \exp\Big[-\frac{i}{\hbar}p\cdot(\Lambda^{-1}x)\Big]. \tag{32}$$

But by Eq. (47.105) the scalar product in the exponent can also be written $p'\cdot x$ where $p' = \Lambda p$, or

$$p'^{\,\mu} = \Lambda^\mu{}_\nu\, p^\nu. \tag{33}$$

The momentum $p'^{\,\mu}$ is the Lorentz transformed version of the original momentum p^μ, in the active sense. That is, when we boost and/or rotate a free-particle Dirac eigenfunction according to Eq. (26), the energy-momentum 4-vector of the new wave function is the boosted and/or rotated version of the original energy-momentum 4-vector. This is just what we want.

6. The Spin Part of the Lorentz Transformation

The spin part of the Lorentz transformation is specified by the as-yet-unknown matrices $D(\Lambda)$. We obtain a condition on these by requiring that the Dirac equation be covariant. For this it suffices to work with the free particle. Suppose $\psi(x)$ is a free-particle solution, that is, it satisfies

$$i\hbar\gamma^\mu\,\frac{\partial\psi(x)}{\partial x^\mu} = mc\,\psi(x). \tag{34}$$

Let us demand that the Lorentz-transformed wave function $\psi'(x) = D(\Lambda)\psi(\Lambda^{-1}x)$ also satisfy the free-particle Dirac equation, that is, let us demand that Lorentz transformations map free particle solutions into other free particle solutions. Then we have

$$i\hbar\gamma^\mu D(\Lambda)\frac{\partial\psi(\Lambda^{-1}x)}{\partial x^\mu} = mc\, D(\Lambda)\psi(\Lambda^{-1}x). \tag{35}$$

In this formula, space-time indices are indicated explicitly, for example, $\gamma^\mu\partial/\partial x^\mu$, but spinor indices are not. For example, ψ is a 4-component column spinor and $D(\Lambda)$ is a 4×4 matrix, and matrix multiplication is implied. We have pulled the $\partial/\partial x^\mu$ past $D(\Lambda)$ since the latter depends on Λ but not on x.

Let us write

$$y^\mu = (\Lambda^{-1})^\mu{}_\nu\, x^\nu \tag{36}$$

and multiply Eq. (35) by $D(\Lambda)^{-1}$ to clear the $D(\Lambda)$ on the right hand side. Then we get

$$i\hbar\, D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\frac{\partial\psi(y)}{\partial x^\mu} = mc\,\psi(y). \tag{37}$$

But since ψ satisfies the Dirac equation, we have

$$i\hbar\gamma^\nu\,\frac{\partial\psi(y)}{\partial y^\nu} = mc\,\psi(y) \tag{38}$$

which is just Eq. (34) with $x\to y$ and $\mu\to\nu$. We also have

$$\frac{\partial\psi(y)}{\partial x^\mu} = \frac{\partial\psi(y)}{\partial y^\nu}\frac{\partial y^\nu}{\partial x^\mu} = \frac{\partial\psi(y)}{\partial y^\nu}(\Lambda^{-1})^\nu{}_\mu. \tag{39}$$

Now combining Eqs. (37) and (38) and cancelling $i\hbar$, we get

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\frac{\partial\psi(y)}{\partial y^\nu}(\Lambda^{-1})^\nu{}_\mu = \gamma^\nu\frac{\partial\psi(y)}{\partial y^\nu}. \tag{40}$$

But $\partial\psi(y)/\partial y^\nu$ is arbitrary (by choosing different free particle solutions we can make it anything we want), so

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\,(\Lambda^{-1})^\nu{}_\mu = \gamma^\nu, \tag{41}$$

or, by multiplying by Λ to clear the Λ^{-1} on the left hand side,

$$\boxed{D(\Lambda)^{-1}\gamma^\mu D(\Lambda) = \Lambda^\mu{}_\nu\,\gamma^\nu.} \tag{42}$$

This result is important because it is the fundamental relation that the matrices $D(\Lambda)$ must satisfy. In fact, it allows those matrices to be determined, as we shall see. It is also a statement that the 4-vector of matrices γ^μ actually transforms as a 4-vector. For comparison recall the adjoint formula for rotations of a spin-$\frac{1}{2}$ particle,

$$U(\mathsf{R})^\dagger\boldsymbol{\sigma}U(\mathsf{R}) = \mathsf{R}\boldsymbol{\sigma} \tag{43}$$

(see Eq. (12.29)). This is a special case of the transformation law for a vector operator, that is, the transformation law for a 3-vector under spatial rotations. See also Eq. (19.14) for the general case of a vector operator.

We remark that had we been using the passive point of view, then our requirement that Lorentz transformations map free particle solutions of the Dirac equation into other free particle solutions would become the requirement that the free-particle Dirac equation have the same form in all Lorentz frames.

7. The Matrices $D(\Lambda)$ for Infinitesimal Lorentz Transformations

The case of infinitesimal Lorentz transformations is especially important. To see why, note that if Eq. (42) is valid for two Lorentz transformations,

$$\begin{aligned} D(\Lambda_1)^{-1}\gamma^\mu D(\Lambda_1) &= (\Lambda_1)^\mu{}_\nu\,\gamma^\nu \\ D(\Lambda_2)^{-1}\gamma^\mu D(\Lambda_2) &= (\Lambda_2)^\mu{}_\nu\,\gamma^\nu, \end{aligned} \tag{44}$$

then it is true for their product $\Lambda_1\Lambda_2$. We see this by combining Eqs. (44) to get

$$\begin{aligned} D(\Lambda_2^{-1})D(\Lambda_1^{-1})\gamma^\mu D(\Lambda_1)D(\Lambda_2) &= D(\Lambda_2^{-1})(\Lambda_1)^\mu{}_\nu\,\gamma^\nu D(\Lambda_2) = (\Lambda_1)^\mu{}_\nu\,D(\Lambda_2^{-1})\gamma^\nu D(\Lambda_2) \\ &= (\Lambda_1)^\mu{}_\nu\,(\Lambda_2)^\nu{}_\sigma\gamma^\sigma = (\Lambda_1\Lambda_2)^\mu{}_\sigma\,\gamma^\sigma. \end{aligned} \tag{45}$$

But according to the representation law (27), the left-hand side of Eq. (45) can be written

$$D\big((\Lambda_1\Lambda_2)^{-1}\big)\gamma^\mu D(\Lambda_1\Lambda_2), \tag{46}$$

where the $\pm$ sign cancels out. This proves the assertion.

Thus, if we can find $D(\Lambda)$ satisfying Eq. (42) for infinitesimal Lorentz transformations, and if we use the representation property (27) to build up $D(\Lambda)$ for finite, proper Lorentz transformations as products of infinitesimal ones, then the finite ones will automatically satisfy Eq. (42). We recall that any proper Lorentz transformation can be built up a product of infinitesimal ones (see the end of Sec. 47.3). This simplifies the problem of finding the matrices $D(\Lambda)$ considerably.

The general form of an infinitesimal Lorentz transformation was presented in Eq. (47.63). It is

$$\Lambda = \mathsf{I} + \frac{1}{2}\theta_{\mu\nu}\mathsf{J}^{\mu\nu}, \tag{47}$$

where $\theta_{\mu\nu}$ is an antisymmetric tensor or matrix of small numbers, specifying the infinitesimal Lorentz transformation, and where $\mathsf{J}^{\mu\nu}$ is an antisymmetric "tensor" of 4×4 matrices. We recall that for fixed value of μ and ν, $\mathsf{J}^{\mu\nu}$ is a 4×4 matrix, that is, the μ and ν are labels of the matrix, not its component indices. The components are $(\mathsf{J}^{\mu\nu})^\alpha{}_\beta$, and are given by Eq. (47.64). Because $\theta_{\mu\nu} = -\theta_{\nu\mu}$, there are only 6 independent components of $\theta_{\mu\nu}$, which are obtained if we restrict the indices to $\mu < \nu$. Thus we can think of the infinitesimal Λ in Eq. (47) as functions of these six independent $\theta_{\mu\nu}$.

The D-matrix representing this infinitesimal Lorentz transformation, $D(\Lambda)$, must also therefore be a function of the six independent $\theta_{\mu\nu}$. Let us expand this $D(\Lambda)$ out to first order in the independent $\theta_{\mu\nu}$,

$$D(\Lambda) = D(\theta_{\mu\nu}) = 1 + \sum_{\mu<\nu} \theta_{\mu\nu} \frac{\partial D}{\partial \theta_{\mu\nu}}(0). \tag{48}$$

Now we define matrices $\sigma^{\mu\nu}$ for $\mu < \nu$ by

$$\frac{\partial D}{\partial \theta_{\mu\nu}}(0) = -\frac{i}{2}\,\sigma^{\mu\nu}, \tag{49}$$

and then define $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$ for $\mu \geq \nu$. Since the derivatives are evaluated at $\theta_{\mu\nu} = 0$, the matrices $\sigma^{\mu\nu}$ are independent of $\theta_{\mu\nu}$, that is, they are constants. The factor $-i/2$ is conventional, but it is intended to make the answers come out in familiar form for rotations of nonrelativistic spinors, which we know about already. The matrices $\sigma^{\mu\nu}$ form an antisymmetric "tensor" of 4×4 Dirac matrices (that is, matrices that act on spin space). This is a generalization of the 4-vector of Dirac matrices γ^μ. Later we will see that $\sigma^{\mu\nu}$ actually transforms as a tensor under Lorentz transformations, in a generalization of Eq. (42). We must find the explicit form of the matrices $\sigma^{\mu\nu}$ to determine $D(\Lambda)$ for infinitesimal Lorentz transformations.

We extend the sum in Eq. (48) to all μ, ν, so that

$$D(\Lambda) = 1 - \frac{i}{4}\,\theta_{\mu\nu}\,\sigma^{\mu\nu}. \tag{50}$$

To determine the matrices $\sigma^{\mu\nu}$, we substitute the infinitesimal Lorentz transformation (47) or its spinor representative (50) into the fundamental transformation law for γ^μ, Eq. (42), switching indices $\mu\nu \to \alpha\beta$ on σ to avoid collision of indices. This gives

$$\Big(1 + \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}\Big)\gamma^\mu\Big(1 - \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}\Big) = \Big[\mathsf{I} + \frac{1}{2}\theta_{\alpha\beta}\,\mathsf{J}^{\alpha\beta}\Big]^\mu{}_\nu\,\gamma^\nu, \tag{51}$$

or, on multiplying this out and keeping terms that are first order in $\theta_{\alpha\beta}$,

$$\frac{i}{4}\theta_{\alpha\beta}[\sigma^{\alpha\beta}, \gamma^\mu] = \frac{1}{2}\theta_{\alpha\beta}(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha), \tag{52}$$

where we have used Eq. (47.64) for the components of $\mathsf{J}^{\alpha\beta}$. The antisymmetric but otherwise arbitrary coefficients $\theta_{\alpha\beta}$ on both sides are contracted with objects antisymmetric in $(\alpha\beta)$, so we can cancel $\theta_{\alpha\beta}$ to obtain

$$\frac{i}{4}[\sigma^{\alpha\beta}, \gamma^\mu] = \frac{1}{2}(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha). \tag{53}$$

This equation must be solved for $\sigma^{\alpha\beta}$.

The notation suggests that $\sigma^{\alpha\beta}$ transforms as a tensor under Lorentz transformations. We guess that that is true. We already know that γ^μ transforms as a 4-vector under Lorentz transformations, in the sense of Eq. (42). This implies that $\gamma^\alpha\gamma^\beta$ (the product of two γ-matrices, another 4×4 spin matrix) transforms as a second rank tensor under Lorentz transformations, that is, that

$$D(\Lambda^{-1})\gamma^\alpha\gamma^\beta D(\Lambda) = \Lambda^\alpha{}_\sigma\,\Lambda^\beta{}_\tau\gamma^\sigma\gamma^\tau. \tag{54}$$

This fact will be left as a easy exercise. But $\gamma^\alpha\gamma^\beta$ is not antisymmetric, in fact, it has a nonvanishing symmetric part given by the fundamental anticommutation relations (18). However its antisymmetric part is an antisymmetric, second rank tensor of Dirac matrices, so it is a good guess that it must be the same as $\sigma^{\alpha\beta}$ to within a multiplicative constant.

That is, let us guess that $\sigma^{\alpha\beta} = k(\gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha) = k[\gamma^\alpha, \gamma^\beta]$ for some constant k. Substituting this into Eq. (53) and using the anticommutation relations (18), we find after some algebra that it works with $k = i/2$. Altogether, we find

$$\boxed{\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu],} \tag{55}$$

switching back to indices μ, ν. This is the explicit solution for the matrices that appear in $D(\Lambda)$ for an infinitesimal Lorentz transformation, as shown in Eq. (50).

The matrices $\sigma^{\mu\nu}$ are considered the *generators* of the Dirac representation $D(\Lambda)$ of Lorentz transformations. They are analogous to the Pauli matrices $\boldsymbol{\sigma}$ of the nonrelativistic theory, which play as similar role as the generators of spin rotations for a spin-$\frac{1}{2}$ particle. That is, an infinitesimal rotation matrix for a spin-$\frac{1}{2}$ particle is given by

$$U(\hat{\mathbf{n}}, \theta) = 1 - \frac{i}{2}\theta\hat{\mathbf{n}}\cdot\boldsymbol{\sigma} = 1 - \frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}, \tag{56}$$

where $\boldsymbol{\theta} = \theta\hat{\mathbf{n}}$ is a vector of small angles specifying the infinitesimal rotation. This is discussed in Notes 12, and Eq. (56) is a small-angle version of Eq. (12.27). Equation (56) in the nonrelativistic Pauli theory may be compared to Eq. (50) in the relativistic Dirac theory.

The matrices $\sigma^{\mu\nu}$ are a new set of 4×4 Dirac matrices, in addition to γ^μ. As the notation indicates, $\sigma^{\mu\nu}$ transforms as a second rank tensor, in the sense of

$$D(\Lambda^{-1})\sigma^{\mu\nu}D(\Lambda) = \Lambda^\mu{}_\alpha\,\Lambda^\nu{}_\beta\,\sigma^{\alpha\beta}, \tag{57}$$

as follows from Eqs. (54) and (55).

8. $D(\Lambda)$ for Pure Rotations

Let us specialize to the case of pure rotations, which is summarized by Eqs. (47.68), (47.70), (47.72), (47.74) and (47.75). In this case, of the six $\theta_{\mu\nu}$ we have $\theta_{0i} = 0$. As for the remaining three components θ_{ij}, we express these in terms of a 3-vector of angles θ_i by $\theta_i = (1/2)\epsilon_{ijk}\,\theta_{jk}$, or its inverse, $\theta_{ij} = \epsilon_{ijk}\,\theta_k$. The vector of small angles θ_i is related to the axis and angle of the infinitesimal rotation by $\boldsymbol{\theta} = \theta\hat{\mathbf{n}}$, that is, $\theta = |\boldsymbol{\theta}|$ and $\hat{\mathbf{n}} = \boldsymbol{\theta}/\theta$. Then the Dirac D-matrix for an infinitesimal rotation can be written,

$$D(\hat{\mathbf{n}}, \theta) = 1 - \frac{i}{4}\theta_{ij}\,\sigma^{ij} = 1 - \frac{i}{4}\epsilon_{ijk}\,\sigma^{ij}\,\theta_k = 1 - \frac{i}{2}\Sigma_k\,\theta_k, \tag{58}$$

where we define

$$\Sigma_i = \frac{1}{2}\epsilon_{ijk}\,\sigma^{jk}. \tag{59}$$

In this notation we can also write the infinitesimal rotation as

$$D(\hat{\mathbf{n}},\theta) = 1 - \frac{i}{2}\theta\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma} = 1 - \frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\Sigma}, \tag{60}$$

which should be compared to Eq. (56) in the nonrelativistic theory. The formulas look the same except for $\boldsymbol{\sigma}\to\boldsymbol{\Sigma}$; of course, $\boldsymbol{\sigma}$ is a vector of 2×2 matrices, while $\boldsymbol{\Sigma}$ is a vector of 4×4 matrices.

The vector of Dirac matrices $\boldsymbol{\Sigma}$ is a new set to be added to the collection we have so far. The menagerie of Dirac matrices can be divided into those that are useful in a $3+1$-description of the theory, and those that are useful in a covariant description. The former set includes $\boldsymbol{\alpha}$, β, and $\boldsymbol{\Sigma}$. We put a lower index on the components Σ_i of $\boldsymbol{\Sigma}$ for the same reasons we did on α_i. The set of Dirac matrices useful in a covariant description includes γ^μ, $\sigma^{\mu\nu}$ and one more to be added later.

As for the matrices $\boldsymbol{\Sigma}$, they can be worked out by evaluating the commutators in the definition (55) and using Eq. (14). Since the matrices γ^i are the same in both the Dirac-Pauli and Weyl representations, the answers are the same in both representations. We find

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \qquad \text{(Dirac-Pauli and Weyl)}. \tag{61}$$

We can now find the Dirac D-matrices for rotations of any finite angle, by composing rotations by an infinitesimal angle. When θ is not small we write

$$D(\hat{\mathbf{n}},\theta) = \big[D\big(\hat{\mathbf{n}},\frac{\theta}{N}\big)\big]^N = \lim_{N\to\infty}\Big[1 - \frac{i}{2}\frac{\theta}{N}\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma}\Big]^N = \exp\Big(-\frac{i}{2}\theta\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma}\Big), \tag{62}$$

where we use a matrix version of the limit (4.42). Another method of building up finite rotations out of infinitesimal ones is to use a differential equation, as in Secs. 11.9 and 12.5.

If we use the explicit form (61) for $\boldsymbol{\Sigma}$, we obtain the Dirac rotation matrices in the form

$$D(\hat{\mathbf{n}},\theta) = \begin{pmatrix} U(\hat{\mathbf{n}},\theta) & 0 \\ 0 & U(\hat{\mathbf{n}},\theta) \end{pmatrix} \qquad \text{(Dirac-Pauli and Weyl)}, \tag{63}$$

where $U(\hat{\mathbf{n}},\theta)$ is the 2×2 rotation matrix for spin-$\frac{1}{2}$ particles, that is,

$$U(\hat{\mathbf{n}},\theta) = \exp\Big(-i\frac{\theta}{2}\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\Big) = \cos\frac{\theta}{2} - i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2} \tag{64}$$

(this is Eq. (12.27)). We see that to subject a Dirac 4-component spinor to a purely spatial rotation, we rotate both the upper and lower 2-component spinors by the nonrelativistic rotation matrix for a spin-$\frac{1}{2}$ particle. This is true in both the Dirac-Pauli and Weyl representations.

9. $D(\Lambda)$ for Pure Boosts

The case of pure boosts is summarized by Eqs. (47.69), (47.71), (47.73), (47.76) and (47.77). In this case, of the six independent $\theta_{\mu\nu}$, the three parameters contained in θ_{ij} vanish leaving the three θ_{0i} to parameterize the boost. We write $\lambda_i = \theta_{0i}$ for these parameters, the components of a boost

vector $\boldsymbol{\lambda}$ with rapidity $\lambda = |\boldsymbol{\lambda}|$ and boost axis $\hat{\mathbf{b}} = \boldsymbol{\lambda}/\lambda$. The correction term in the infinitesimal D-matrix in Eq. (50) is

$$-\frac{i}{4}\theta_{\mu\nu}\sigma^{\mu\nu} = -\frac{i}{4}(\theta_{0i}\sigma^{0i} + \theta_{i0}\sigma^{i0}) = -\frac{i}{2}\theta_{0i}\sigma^{0i} = -\frac{i}{2}\lambda_i\sigma^{0i}. \tag{65}$$

But

$$\sigma^{0i} = \frac{i}{2}[\gamma^0, \gamma^i] = \frac{i}{2}[\beta, \beta\alpha_i] = \frac{i}{2}(\beta^2\alpha_i - \beta\alpha_i\beta) = i\alpha_i, \tag{66}$$

so for infinitesimal boosts we have

$$D(\hat{\mathbf{b}}, \lambda) = 1 + \frac{1}{2}\boldsymbol{\lambda}\cdot\boldsymbol{\alpha} = 1 + \frac{\lambda}{2}\hat{\mathbf{b}}\cdot\boldsymbol{\alpha}. \tag{67}$$

We see that the velocity matrices $\boldsymbol{\alpha}$ are the generators of boosts, perhaps not a surprise.

To obtain finite boosts we follow the same procedure shown in Eq. (62) for rotations, which gives

$$D(\hat{\mathbf{b}}, \lambda) = \exp\Big(\frac{\lambda}{2}\hat{\mathbf{b}}\cdot\boldsymbol{\alpha}\Big). \tag{68}$$

The exponential is easiest to carry out in the Weyl representation (46.13), in which the matrices $\boldsymbol{\alpha}$ are block-diagonal,

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix} \qquad \text{(Weyl)}. \tag{69}$$

The result can be expressed as

$$D(\hat{\mathbf{b}}, \lambda) = \begin{pmatrix} V(\hat{\mathbf{b}}, \lambda) & 0 \\ 0 & V(\hat{\mathbf{b}}, \lambda)^{-1} \end{pmatrix} \qquad \text{(Weyl)}, \tag{70}$$

where

$$V(\hat{\mathbf{b}}, \lambda) = \exp\Big(\frac{\lambda}{2}\hat{\mathbf{b}}\cdot\boldsymbol{\sigma}\Big) = \cosh\frac{\lambda}{2} + (\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})\sinh\frac{\lambda}{2} \tag{71}$$

and

$$V(\hat{\mathbf{b}}, \lambda)^{-1} = V(\hat{\mathbf{b}}, -\lambda) = \cosh\frac{\lambda}{2} - (\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})\sinh\frac{\lambda}{2}. \tag{72}$$

Although the rotation matrices $U(\hat{\mathbf{n}}, \theta)$ (see Eq. (64)) are unitary, the matrices $V(\hat{\mathbf{b}}, \lambda)$ which appear in the boosts are not (in fact, they are Hermitian). One can see that the V-matrices are like the U-matrices but with an imaginary angle.

To obtain the boosts in the Dirac-Pauli representation, we can either exponentiate the 4-dimensional $\boldsymbol{\alpha}$ matrices in that representation, or else change the basis according to

$$X_{\rm DP} = W X_{\rm W} W^\dagger, \tag{73}$$

where X is any Dirac matrix and the subscripts DP and W refer to the Dirac-Pauli and Weyl representations, respectively, and where

$$W = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{74}$$

See Eqs. (46.14)–(46.15). This gives

$$D(\hat{\mathbf{b}}, \lambda) = \begin{pmatrix} \cosh(\lambda/2) & (\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})\sinh(\lambda/2) \\ (\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})\sinh(\lambda/2) & \cosh(\lambda/2) \end{pmatrix} \qquad \text{(Dirac-Pauli)}. \tag{75}$$

10. Properties of the matrices $D(\Lambda)$

We recall the theorem quoted in Sec. 47.6, that every proper Lorentz transformation can be represented uniquely as a product of a rotation times a boost, $\Lambda = \mathsf{RB}$. Since we now have the D-matrices for both pure rotations and pure boosts, we can use this theorem to find the D-matrix for an arbitrary Lorentz transformation. It is the product of a rotation matrix of the form (63), times a boost matrix of the form (70) or (75). As mentioned, the D-matrices for pure rotations are unitary, while those for pure boosts are Hermitian, and not generally unitary. The product of a rotation times a boost is a matrix with no particular symmetry, in general.

We usually say that unitary transformations are needed to implement symmetry operations, in order to preserve probabilities. Why then are the Dirac D-matrices not unitary? The answer is that the $D(\Lambda)$ only implement the spin part of a Lorentz transformation, but there is a spatial part, too. Since the probability density is $\psi^\dagger\psi$ in the Dirac theory, a normalized wave function satisfies

$$\int d^3\mathbf{x}\,\psi^\dagger(\mathbf{x},t)\psi(\mathbf{x},t) = 1, \tag{76}$$

for all t. Under a Lorentz transformation the volume element $d^3\mathbf{x}$ is not invariant, but rather scales by the relativistic factor $\gamma = 1/\sqrt{1-(v/c)^2}$ of length contraction and time dilation (obviously not to be confused with a Dirac matrix). Similarly, the spin part of a Lorentz transformation guarantees that $\psi^\dagger\psi$ is not an invariant, either (it is the time component of a 4-vector), rather it also acquires a factor of γ upon being subjected to a Lorentz transformation, which cancels the factor of γ coming from the volume element. Thus, the normalization integral is invariant, but $\psi^\dagger\psi$ is not. That is, the overall transformation is unitary, even if the spin part is not. The exception is a pure rotation, under which both $\psi^\dagger\psi$ and $d^3\mathbf{x}$ are invariant, and D is unitary. This discussion has been rather imprecise and lacking in details because a proper understanding of probability conservation in a relativistic theory must take into account the relativity of simultaneity, and is best expressed in terms of currents and 3-forms in space-time. Nevertheless, we will see the factor of γ appear when we consider the transformation of spinors for free particles.

Notice that the Dirac D-matrices for pure rotations in Eq. (63) have the same double-valuedness seen in nonrelativistic rotations. That is, rotations about an axis $\hat{\mathbf{n}}$ by angles θ and $\theta + 2\pi$ give the same classical rotation, but the Dirac rotations differ by a sign, $D(\hat{\mathbf{n}}, \theta + 2\pi) = -D(\hat{\mathbf{n}}, \theta)$. On the other hand, there is no double-valuedness in the boosts, as seen in Eq. (70) or (75); a classical boost with a given axis $\hat{\mathbf{b}}$ and rapidity λ corresponds to a unique Dirac D matrix, which boosts a spinor. Overall, the Dirac D-matrices have the same double-valuedness seen in spinor rotations, with nothing added on account of boosts. Thus, a sign ambiguity appears if we attempt to parameterize a Dirac D-matrix by a classical Lorentz transformation, as in the notation $D(\Lambda)$, but not if we parameterize rotations and boosts by $(\hat{\mathbf{n}}, \theta)$ and $(\hat{\mathbf{b}}, \lambda)$.

11. Conjugation of Dirac Matrices by γ^0

The generators of unitary transformations are Hermitian, but since the Dirac D-matrices are in general not unitary, the generators $\sigma^{\mu\nu}$ cannot be in general Hermitian. In fact, the spatial parts, σ^{ij}, or, equivalently, $\boldsymbol{\Sigma}$, are Hermitian, as can be seen in Eq. (59) and (61). But the components $\sigma^{0i} = i\alpha_i$, which generate boosts, are anti-Hermitian. See Eq. (66).

But there is a simple relation between $\sigma^{\mu\nu}$ and $(\sigma^{\mu\nu})^\dagger$, given by

$$\boxed{\gamma^0\sigma^{\mu\nu}\gamma^0 = \big(\sigma^{\mu\nu}\big)^\dagger.} \tag{77}$$

This in turn implies

$$\boxed{\gamma^0 D(\Lambda)^{-1}\gamma^0 = D(\Lambda)^\dagger.} \tag{78}$$

This replaces the usual property for unitary matrices, $U^{-1} = U^\dagger$, when working with Dirac D-matrices.

To prove Eq. (77), we begin with

$$\boxed{\gamma^0\gamma^\mu\gamma^0 = \big(\gamma^\mu\big)^\dagger.} \tag{79}$$

We recall that $\gamma^0 = \beta$ is Hermitian, and $\gamma^i = \beta\alpha_i$ is anti-Hermitian (see Eq. (16)). So

$$\gamma^0\gamma^0\gamma^0 = \beta^3 = \beta = \gamma^0 = \big(\gamma^0\big)^\dagger, \tag{80}$$

and

$$\gamma^0\gamma^i\gamma^0 = \beta^2\alpha_i\beta = \alpha_i\beta = -\beta\alpha_i = -\gamma^i = \big(\gamma^i\big)^\dagger. \tag{81}$$

This proves Eq. (79).

To prove Eq. (77) we use the definition (55), and write

$$\begin{aligned}\gamma^0\sigma^{\mu\nu}\gamma^0 &= \frac{i}{2}\gamma^0\big(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\big)\gamma^0 = \frac{i}{2}\big[\big(\gamma^0\gamma^\mu\gamma^0\big)\big(\gamma^0\gamma^\nu\gamma^0\big) - (\mu\leftrightarrow\nu)\big] \\ &= \frac{i}{2}\big[\big(\gamma^\mu\big)^\dagger\big(\gamma^\nu\big)^\dagger - (\mu\leftrightarrow\nu)\big] = \frac{i}{2}\big(\gamma^\nu\gamma^\mu - \gamma^\mu\gamma^\nu\big)^\dagger = \big(\sigma^{\mu\nu}\big)^\dagger.\end{aligned} \tag{82}$$

Now from Eq. (50), we have, in the case of an infinitesimal Lorentz transformation,

$$D(\Lambda)^{-1} = 1 + \frac{i}{4}\theta_{\mu\nu}\,\sigma^{\mu\nu}, \tag{83}$$

and

$$D(\Lambda)^\dagger = 1 + \frac{i}{4}\theta_{\mu\nu}\big(\sigma^{\mu\nu}\big)^\dagger. \tag{84}$$

Equation (78) easily follows from this in the case of an infinitesimal Lorentz transformation. But it is easy to show that if Eq. (78) is true for two proper Lorentz transformations, then it is true

for their product. Since an arbitrary proper Lorentz transformation can be built up as a product of infinitesimal Lorentz transformations, the result must be true for an arbitrary proper Lorentz transformation.

The case of parity, an improper Lorentz transformation, is taken up below.

12. The Adjoint Spinor and Probability Current

We are now prepared to show that the Dirac probability current, defined by Eq. (46.23), transforms as a 4-vector under Lorentz transformations. To begin we write the components of the current in the following way:

$$\begin{aligned} J^0 &= c\rho = c\psi^\dagger\psi = c\psi^\dagger\gamma^0\gamma^0\psi, \\ J^i &= c\psi^\dagger\alpha_i\psi = c\psi^\dagger\gamma^0\gamma^0\alpha_i\psi = c\psi^\dagger\gamma^0\gamma^i\psi. \end{aligned} \tag{85}$$

The combination $\psi^\dagger\gamma^0$ is of frequent occurrence in the Dirac theory, so we give it a special notation,

$$\boxed{\bar\psi = \psi^\dagger\gamma^0,} \tag{86}$$

and we call it the *adjoint spinor*. Notice that the adjoint spinor contains a complex conjugated version of ψ, and is a row spinor. In terms of the adjoint spinor, we can write

$$\boxed{J^\mu = c\,\bar\psi\gamma^\mu\psi.} \tag{87}$$

This is regarded as the covariant form of the probability current.

Let us see how $J^\mu(x)$ transforms under a Lorentz transformation. The Dirac spinor ψ itself transforms according to Eq. (26), which we write in the form

$$\psi(x) \xrightarrow{\Lambda} D(\Lambda)\psi\big(\Lambda^{-1}x\big). \tag{88}$$

Thus the adjoint spinor transforms according to

$$\bar\psi(x) = \psi^\dagger(x)\gamma^0 \xrightarrow{\Lambda} \psi^\dagger\big(\Lambda^{-1}x\big)D(\Lambda)^\dagger\gamma^0 = \psi^\dagger\big(\Lambda^{-1}x\big)\gamma^0\gamma^0D(\Lambda)^\dagger\gamma^0 = \bar\psi\big(\Lambda^{-1}x\big)D(\Lambda)^{-1}, \tag{89}$$

where we have used Eq. (78). Therefore J^μ itself transforms according to

$$\begin{aligned} J^\mu(x) = c\bar\psi(x)\gamma^\mu\psi(x) &\xrightarrow{\Lambda} c\bar\psi\big(\Lambda^{-1}x\big)D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\psi\big(\Lambda^{-1}x\big) \\ &= \Lambda^\mu{}_\nu\, c\bar\psi\big(\Lambda^{-1}x\big)\gamma^\nu\psi\big(\Lambda^{-1}x\big) = \Lambda^\mu{}_\nu\, J^\nu\big(\Lambda^{-1}x\big), \end{aligned} \tag{90}$$

where we have used Eq. (42). The result is that $J^\mu(x)$ transforms as a vector field on space-time should under an active Lorentz transformation, as shown in Eq. (25). Thus, the continuity equation (46.28) is covariant, as we require.

13. How Fields Transform

We are collecting several examples of different kinds of fields and how they transform under Lorentz transformations. Table 1 summarizes the ones we have encountered so far, and several new ones as well. The transformation properties of Dirac spinors and adjoint spinors has been discussed in these notes. The case of scalars (that is, scalars under Lorentz transformations) has been discussed in Notes 47, in which it was pointed out that $E^2 - B^2$ is a Lorentz scalar that can be constructed out of the electromagnetic field [see Eqs. (47.88)–(47.89]. It has also been mentioned that the Klein-Gordon wave function $\psi_{\rm KG}$ is a Lorentz scalar. Can we construct a Lorentz scalar out of the Dirac wave function? Yes, it turns out that the quantity $\bar\psi(x)\psi(x)$ is a Lorentz scalar. The proof of this fact will be left as an exercise. Notice that $\psi^\dagger(x)\psi(x)$ is not a Lorentz scalar, in fact it is the time-component of a 4-vector (essentially the current J^μ).

Spinor	$\psi(x) \xrightarrow{\Lambda} D(\Lambda)\psi\big(\Lambda^{-1}x\big)$
Adjoint Spinor	$\bar\psi(x) \xrightarrow{\Lambda} \bar\psi\big(\Lambda^{-1}x\big) D(\Lambda)^{-1}$
Scalar	$S(x) \xrightarrow{\Lambda} S\big(\Lambda^{-1}x\big)$
Vector	$V^\mu(x) \xrightarrow{\Lambda} \Lambda^\mu{}_\nu\, V^\nu\big(\Lambda^{-1}x\big)$
Tensor	$T^{\mu\nu}(x) \xrightarrow{\Lambda} \Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\, T^{\alpha\beta}\big(\Lambda^{-1}x\big)$
Pseudoscalar	$K(x) \xrightarrow{\Lambda} (\det\Lambda) K\big(\Lambda^{-1}x\big)$
Pseudovector	$W^\mu(x) \xrightarrow{\Lambda} (\det\Lambda)\, \Lambda^\mu{}_\nu\, W^\nu\big(\Lambda^{-1}x\big)$

Table 1. Transformation properties of various types of fields under an active Lorentz transformation.

The table shows the transformation law for a generic 4-vector field $V^\mu(x)$; examples include the 4-vector potential A^μ in electromagnetism (see Eq. (10)), the Klein-Gordon probability current (see Eq. (45.27)) and the Dirac probability current (which we have gone to some trouble in these notes to show is a genuine 4-vector, see Sec. 12).

The table also refers to a generic second-rank, relativistic tensor $T^{\mu\nu}$ and its transformation law. Such tensors occur in electromagnetic theory, for example, the field tensor $F^{\mu\nu}$ and the stress-energy tensor. A second-rank, antisymmetric tensor occurs in the Dirac theory; it is $\bar\psi(x)\sigma^{\mu\nu}\psi(x)$. This is proportional to a relativistic generalization of the magnetization $\mathbf{M}$ (the dipole moment per unit volume), which in the nonrelativistic Pauli theory of the electron can be written as $\psi^\dagger\boldsymbol{\mu}\psi$, where $\boldsymbol{\mu}$ is the magnetic moment operator for the electron. See Sec. 18.6. We will see this tensor appear in the Gordon decomposition of the current, which is discussed in Sec. 50.10.

14. Improper Lorentz Transformations and Parity

Recall that an improper Lorentz transformation is one that cannot be built up from near-identity transformations, rather it is the product of a proper Lorentz transformation times time-reversal or parity or both. See Sec. 47.3.

Recall also that in the nonrelativistic theory we classified vectors as either true vectors or pseudo-vectors, depending on how they transform under parity (odd and even, respectively). See Sec. 21.6. We also spoke of scalars and pseudoscalars.

In the relativistic theory a pseudoscalar (for example, K in the table, or $\mathbf{E}\cdot\mathbf{B}$ in electromagnetism) transforms as a scalar under proper Lorentz transformations but under improper ones it acquires a sign given by $\det\Lambda$. Similarly, a pseudovector (for example, W^μ in the table) transforms as a vector under proper Lorentz transformation but under improper ones it acquires the same sign. In the following the only improper Lorentz transformation we will consider is parity, or products of parity times proper Lorentz transformations. For simplicity we will leave time-reversal out of the picture. With this understanding, $\det\Lambda = +1$ for proper Lorentz transformations and $\det\Lambda = -1$ for improper ones.

The classical Lorentz transformation corresponding to parity is

$$\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{91}$$

in other words, it is the spatial inversion operation, which leaves time alone. Notice that $\det\mathsf{P} = -1$.

Let us look for the operator π that acts on Dirac wave functions, corresponding to the spatial inversion operation P. In the nonrelativistic theory, parity is a purely spatial operation with no effect on the spin [see Eq. (21.33)]. Therefore we might guess that the same is true in the Dirac theory, that is, that the transformation law for the Dirac wave function should be $\psi(\mathbf{x},t)\xrightarrow{\pi}\psi(-\mathbf{x},t)$. It turns out, however, that this does not work. Instead we must write

$$\psi(\mathbf{x},t) \xrightarrow{\pi} D(\mathsf{P})\psi(-\mathbf{x},t), \tag{92}$$

where $D(\mathsf{P})$ is a spin matrix to be determined. Notice that the transformation of the space-time dependence in Eq. (92) can also be written,

$$\psi(x) \xrightarrow{\pi} D(\mathsf{P})\psi\big(\mathsf{P}^{-1}x\big), \tag{93}$$

which makes the transformation law under parity a generalization of Eq. (26), which was originally intended to apply to proper Lorentz transformations only.

We determine $D(\mathsf{P})$ by requiring that parity map free particle solutions of the Dirac equation into other free particle solutions. The analysis of this question proceeds exactly as in Sec. 6, and it leads to a conclusion of the same form as Eq. (42), that is, $D(\mathsf{P})$ must satisfy

$$D(\mathsf{P})^{-1}\gamma^\mu D(\mathsf{P}) = P^\mu{}_\nu\,\gamma^\nu, \tag{94}$$

where $P^\mu{}_\nu$ are the components of P. This implies

$$\begin{aligned} \mu = 0: &\qquad D(\mathsf{P})^{-1}\gamma^0 D(\mathsf{P}) = \gamma^0 = \left(\gamma^0\right)^\dagger, \\ \mu = i: &\qquad D(\mathsf{P})^{-1}\gamma^i D(\mathsf{P}) = -\gamma^i = \left(\gamma^i\right)^\dagger. \end{aligned} \tag{95}$$

But in view of Eq. (79), we see that Eqs. (95) are satisfied if we take

$$D(\mathsf{P}) = e^{i\alpha}\,\gamma^0, \tag{96}$$

where α is any phase. And we see that $D(\mathsf{P}) = 1$ will not work; in the relativistic theory, parity *must* involve the spin. This is another indication of the more intimate coupling between spatial and spin degrees of freedom in the relativistic theory.

The classical Lorentz transformation P obeys certain properties, including

$$\mathsf{P}\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{R}(\hat{\mathbf{n}},\theta)\mathsf{P}, \tag{97a}$$

$$\mathsf{P}\mathsf{B}(\hat{\mathbf{b}},\lambda) = \mathsf{B}(\hat{\mathbf{b}},-\lambda)\mathsf{P}, \tag{97b}$$

$$\mathsf{P}^2 = \mathsf{I}, \tag{97c}$$

where R and B refer to pure rotations and pure boosts, respectively. (Spatial inversion commutes with rotations but inverts the direction of a boost.) If we demand that $D(\mathsf{P})$, taken along with the $D(\Lambda)$ that we have worked out, form a representation of the extended Lorentz group including parity (but still excluding time-reversal), then we should have

$$D(\mathsf{P})D(\hat{\mathbf{n}},\theta) = D(\hat{\mathbf{n}},\theta)D(\mathsf{P}), \tag{98a}$$

$$D(\mathsf{P})D(\hat{\mathbf{b}},\lambda) = D(\hat{\mathbf{b}},-\lambda)D(\mathsf{P}), \tag{98b}$$

$$D(\mathsf{P})^2 = 1, \tag{98c}$$

where $D(\hat{\mathbf{n}},\theta)$ and $D(\hat{\mathbf{b}},\lambda)$ are pure rotations and pure boosts as in Eqs. (63) and (70) or (75), respectively.

Equation (98c) implies that $e^{i\alpha}$ in Eq. (96) must be ± 1, so that

$$D(\mathsf{P}) = \pm\gamma^0. \tag{99}$$

The final choice of sign is purely a convention; if we take it to be $+1$, then it means that the *intrinsic parity* of the electron (see Sec. 21.5) is $+1$. This in turn means that the intrinsic parity of the positron is -1, although we are jumping the gun to be talking about positrons at this point. But no physics would change if we made the opposite convention. For any fermion, the intrinsic parities of the particle and antiparticle are opposite, but it is a matter of convention which is which. In the following we will settle the matter by taking

$$D(\mathsf{P}) = +\gamma^0. \tag{100}$$

Does this choice of $D(\mathsf{P})$ satisfy Eqs. (98a) and (98b)? In the case of Eq. (98a) for an infinitesimal rotation, the question boils down to

$$\gamma^0\sigma^{ij}\gamma^0 = \sigma^{ij}, \tag{101}$$

which is true in view of Eq. (77) and the fact that σ^{ij} consists of Hermitian matrices. In the case of Eq. (98b) for an infinitesimal boost, we must check

$$\gamma^0\alpha_i\gamma^0 = -\alpha_i, \tag{102}$$

which is also true. But if these relations are true for infinitesimal Lorentz transformations then they are true for any proper Lorentz transformation, by building up finite transformations as the products of infinitesimal ones (by now you must appreciate the power of this argument). We conclude that our definition of $D(\mathsf{P})$ (100) is satisfactory.

With this definition of $D(\mathsf{P})$ it is easy to check that the Dirac current, proportional to $\bar{\psi}\gamma^\mu\psi$, transforms as a true vector (not a pseudovector), and that $\bar{\psi}\psi$ transforms as a true scalar (not a pseudoscalar).

15. Pseudoscalars and Pseudovectors

Pseudoscalars and pseudovectors are important in the weak interactions, notably in the theory of the neutrino, which do not conserve parity. To construct these types of fields we must introduce a new (and final) Dirac matrix,

$$\gamma_5 = \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{103}$$

The 5 index is not a space-time index (obviously), rather it is a way of defining a new Dirac matrix without using up another letter of the Greek alphabet. We do not use a 4 because γ_4 is a Dirac matrix used in (mostly older) versions of the theory which use $x_4 = ict$ as an imaginary coordinate on space-time. See Sec. E.21. The upper or lower position of the index 5 is of no significance. The complete set of Dirac matrices useful in a covariant description includes γ^μ, $\sigma^{\mu\nu}$ and γ_5 (one can also include the identity matrix 1).

The explicit form of the matrix γ_5 is

$$\gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{(Dirac-Pauli)}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{(Weyl)}. \tag{104}$$

Other properties of γ_5 include the following:

$$\big(\gamma_5\big)^\dagger = \gamma_5, \tag{105a}$$

$$\big(\gamma_5\big)^2 = 1, \tag{105b}$$

$$\{\gamma_5, \gamma^\mu\} = 0, \tag{105c}$$

$$[\gamma_5, \sigma^{\mu\nu}] = 0, \tag{105d}$$

Notice that Eq. (105c) implies

$$\{\gamma_5, D(\mathsf{P})\} = 0, \tag{106}$$

and Eq. (105d) implies

$$[\gamma_5, D(\Lambda)] = 0, \tag{107}$$

for all proper Lorentz transformations Λ. We will just prove property (105c). Take the case $\mu = 2$. We have

$$\gamma_5\gamma^2 = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^2 = -i\gamma^0\gamma^1\gamma^2\gamma^2\gamma^3 = -i\gamma^2\gamma^0\gamma^1\gamma^2\gamma^3 = -\gamma^2\gamma_5, \tag{108}$$

where in the first step we move the γ^2 on the right past γ^3, incurring one minus sign, and in the second step we move the γ^2 on the left past γ^1 and γ^0, incurring two minus signs. We can summarize Eqs. (106) and (107) by writing

$$\gamma_5 D(\Lambda) = (\det\Lambda) D(\Lambda)\gamma_5, \tag{109}$$

valid for both proper and improper Lorentz transformations Λ (still excluding time-reversal).

It is now easy to construct the pseudoscalar and pseudovector in the Dirac theory. The pseudoscalar is $\bar\psi(x)\gamma_5\psi(x)$, and the pseudovector is $\bar\psi(x)\gamma_5\gamma^\mu\psi(x)$.

16. The Dirac Algebra and Bilinear Covariants

The various fields with the various transformation properties under Lorentz transformations that are summarized in Table 1 are needed to construct Lorentz invariant Lagrangians that model experimental data. That data shows that nature either does or does not respect various symmetries, which in turn dictates the kinds of fields that may appear in the Lagrangian. It is believed that all interactions are invariant under proper Lorentz transformations, at least at scales at which gravitational effects are unimportant, but it is known that some interactions do not respect parity or time-reversal (more precisely, CP-invariance). For example, at low energies the Lagrangian for the weak interactions involves the difference between a vector and a pseudovector (the "$V - A$" theory), which is responsible for parity violation.

The types of fields listed in Table 1 are *bilinear covariants*, which we now describe. These are associated with the *algebra* of the Dirac matrices γ^μ. The algebra is defined as the set of all linear combinations, with complex coefficients, of all matrices that can be formed by multiplying the γ^μ. It is the space of all complex polynomials that can be constructed out of the γ^μ.

The algebra generated by the 2×2 Pauli matrices is simpler, so let us look at it first. There are three Pauli matrices σ_i, $i = 1, 2, 3$, but $(\sigma_i)^2 = 1$ so the algebra includes the identity matrix. A general quadratic monomial in the Pauli matrices can be reduced to a polynomial of first degree by the formula,

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\,\sigma_k, \tag{110}$$

so the algebra generated by the Pauli matrices consists of all first degree polynomials, that is, all matrices of the form $a + \mathbf{b}\cdot\boldsymbol{\sigma}$, where a and $\mathbf{b}$ are complex coefficients. But the set of matrices, $\{1, \boldsymbol{\sigma}\}$ spans the space of all 2×2 matrices, so the algebra generated by the Pauli matrices consists of all 2×2 matrices.

The Dirac algebra is generated by the four 4×4 Dirac matrices γ^μ. Since $\left(\gamma^\mu\right)^2 = \pm 1$, the Dirac algebra contains the identity matrix. The quadratic monomial $\gamma^\mu\gamma^\nu$ looks like 16 matrices, but only six of these are independent, because of the anticommutator $\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$

(times the identity matrix). The antisymmetric part is captured by $\sigma^{\mu\nu} = (i/2)[\gamma^\mu, \gamma^\nu]$, which has 6 independent components.

As for cubic monomials, say, $\gamma^\mu\gamma^\nu\gamma^\sigma$, these can be reduced to a first degree polynomial if any of the indices μ, ν, and σ are equal. For example, $\gamma^2\gamma^3\gamma^2 = -\gamma^2\gamma^2\gamma^3 = \gamma^3$. But if all three indices are distinct, then one index must be omitted, so there are 4 independent cubic monomials that cannot be reduced to lower degree. In fact, these are given by $\gamma_5\gamma^\mu$, where μ indicates the index that is omitted. For example, if $\mu = 2$ we have

$$\gamma_5\gamma^2 = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^2 = -i\gamma^0\gamma^1\big(\gamma^2\big)^2\gamma^3 = +i\gamma^0\gamma^1\gamma^3, \tag{111}$$

where the final result is a cubic monomial with $\mu = 2$ omitted.

Finally, a quartic monomial $\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau$ can be reduced to a lower degree unless all four indices are distinct, in which case the matrix is proportional to γ_5. Thus, there is one independent quartic monomial. All higher degree monomials must contain repetitions of indices, and so can be reduced to lower degree.

Matrices	Name	Count
1	Scalar	1
γ^μ	Vector	4
$\sigma^{\mu\nu}$	Tensor	6
$\gamma_5\gamma^\mu$	Pseudovector	4
γ_5	Pseudoscalar	1
Total		16

Table 2. Bilinear covariants and the Dirac algebra.

The list of Dirac matrices that span the Dirac algebra is summarized in Table 2. By sandwiching these matrices between $\bar\psi$ and ψ we obtain one of the bilinear covariants given in Table 1. The matrices listed are linearly independent, so they span the space of all 4×4 Dirac matrices. Thus the Dirac algebra consists of all 4×4 matrices, and an arbitrary 4×4 matrix can be expressed uniquely as a linear combination of the the 16 basis matrices listed in Table 2.

17. Angular Momentum of the Dirac Particle

In Notes 12 we defined the angular momentum of a system as the generator of rotations. See Eq. (12.13). Since we now know how to apply rotations to the Dirac particle, we can work out the angular momentum operator by considering infinitesimal rotations and examininng the correction term.

The Dirac wave function $\psi(x)$ transforms under a general Lorentz transformation according to Eq. (26). The transformation has two parts, a space-time Lorentz transformaiton specified by Λ,

which in the case of pure rotations only affects the spatial coordinates $\mathbf{x}$ and leaves the time alone; and a spin part, specified by a Dirac matrix $D(\Lambda)$, which in the case of pure rotations is given by Eq. (63). Making the Lorentz transformation a pure rotation with an infinitesimal angle θ, we can write Eq. (26) as

$$\psi'(\mathbf{x},t) = \Big(1 - \frac{i}{2}\theta\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma}\Big)\psi\big(\mathbf{x} - \theta(\hat{\mathbf{n}}\cdot\mathbf{J})\mathbf{x}, t\big) = \psi(\mathbf{x},t) - \frac{i}{2}\theta(\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma})\psi - \theta[(\hat{\mathbf{n}}\cdot\mathbf{J})\mathbf{x}]\cdot\nabla\psi, \tag{112}$$

where we use Eqs. (47.74) and (47.70) for the infinitesimal rotation Λ and Eq. (62) for the corresponding infinitesimal D-matrix. But by Eq. (11.26) the second correction term in Eq. (112) can be written,

$$-\theta(\hat{\mathbf{n}}\times\mathbf{x})\cdot\nabla\psi = -\frac{i}{\hbar}(\hat{\mathbf{n}}\times\mathbf{x})\cdot\mathbf{p}\psi = -\frac{i}{\hbar}\hat{\mathbf{n}}\cdot(\mathbf{x}\times\mathbf{p})\psi = -\frac{i}{\hbar}\hat{\mathbf{n}}\cdot\mathbf{L}\psi, \tag{113}$$

where $\mathbf{p} = -i\hbar\nabla$ and $\mathbf{L} = \mathbf{x}\times\mathbf{p}$. This is essentially the same derivation of the orbital angular momentum as given in Sec. 15.3, but here we also have a spin part, the first correction term in Eq. (112). Writing $-(i/\hbar)\hat{\mathbf{n}}\cdot\mathbf{J}$ for the entire correction term, we obtain

$$\mathbf{J} = \mathbf{L} + \frac{\hbar}{2}\boldsymbol{\Sigma} \tag{114}$$

for the entire angular momentum of the Dirac particle. This is an obvious generalization of the angular momentum $\mathbf{J} = \mathbf{L} + (\hbar/2)\boldsymbol{\sigma}$ of a spin-$\frac{1}{2}$ particle in the Pauli theory.

Problems

1. The generators $\sigma^{\mu\nu}$ of the Dirac representation of Lorentz transformations must satisfy Eq. (53), which is a version of Eq. (42) when the Lorentz transformation is infinitesimal. We guess that $\sigma^{\mu\nu} = k[\gamma^\mu, \gamma^\nu]$ for some constant k, since both sides are antisymmetric tensors of Dirac matrices. Substitute this guess into Eq. (53) and verify that for an appropriate choice of k the guess is correct. This will give you some practice with working with Dirac matrices; you must pay attention to what is a matrix and what is a number.

2. The transformation properties of fields constructed out of the Dirac wave function $\psi(x)$ under Lorentz transformations. All examples in this problem consist of complete contractions over spin indices, that is, the quantities are scalars as far as the spin indices are concerned. However, they still have a space-time dependence (in this problem x means $(ct, \mathbf{x})$), and they may have space-time indices such as μ, ν etc.

(a) Show that $\bar{\psi}(x)\psi(x)$ transforms as a scalar under proper Lorentz transformations.

(b) Show that $\bar{\psi}(x)\sigma^{\mu\nu}\psi(x)$ transforms as a second rank tensor under proper Lorentz transformations.

(c) Show that $\bar{\psi}(x)\psi(x)$ transforms as a scalar (not a pseudoscalar) under parity. Show that the Dirac current transforms transforms as a vector (not a pseudovector) under parity.

(d) Show that $\bar{\psi}(x)\gamma_5\psi(x)$ transforms as a pseudoscalar, and that $\bar{\psi}(x)\gamma_5\gamma^\mu\psi(x)$ transforms as a pseudovector.

3. A continuation of Problem 46.1. A fact not mentioned in that earlier problem is that the representation of the Dirac algebra in $2+1$ dimensions has two *inequivalent* two-dimensional representations. Recall that in $3+1$ dimensions, the four-dimensional representation found by Dirac is the only one at that dimensionality (all others are equivalent). To do Problem 46.1 or this one, it does not matter which of the two-dimensional representations you use.

(a) Assume that the 2-component Dirac wave function transforms under proper Lorentz transformations Λ in $2+1$ dimensions according to

$$\psi'(x) = D(\Lambda)\psi(\Lambda^{-1}x), \tag{115}$$

where $D(\Lambda)$ is some (as yet unknown) 2×2 representation of the proper Lorentz transformations in $2+1$ dimensions and $x = (ct, x_1, x_2)$ (1,2 mean x, y). Assuming that $\psi(x)$ satisfies the free particle Dirac equation, and that $\psi'(x)$ is given by Eq. (115), demand that $\psi'(x)$ also satisfy the free particle Dirac equation and thereby derive a condition that the representation $D(\Lambda)$ must satisfy.

(b) Write out explicitly the matrices $D(\hat{\mathbf{z}}, \theta)$ for the case of pure rotations and $D(\hat{\mathbf{b}}, \lambda)$ for the case of pure boosts, where $\hat{\mathbf{b}}$ lies in the x-y plane. Do this in the Dirac-Pauli representation. Show that if you work in the Maiorana representation, the D-matrices are purely real.

(c) Show that in $2+1$ dimensions, the spatial inversion operation is a proper Lorentz transformation, that is, it can be continuously connected with the identity. Thus there is no problem of determining $D(\mathsf{P})$ as in $3+1$ dimensions; it is already taken care of by the proper Lorentz transformations worked out in part (b).

4. This problem is borrowed from Bjorken and Drell, *Relativistic Quantum Mechanics*, chapter 4.

We have seen that the Dirac equation with minimal coupling to the electromagnetic field gives a g-factor of 2, very close to the experimental value for the electron. What do we do with spin-$\frac{1}{2}$ particles such as the proton and neutron, which have "anomalous" g-factors?

In this problem we use natural units, $\hbar = c = 1$. Next, we modify the Dirac equation to include another coupling to the electromagnetic field, in addition to the minimal coupling,

$$\Big(\not{p} - q\not{A} - \frac{\kappa e}{4m}\sigma_{\mu\nu}F^{\mu\nu} - m\Big)\psi = 0, \tag{116}$$

where m is the mass, q the charge, and κ the strength of the anomalous magnetic moment term. For the electron, $q = -e$ and $\kappa = 0$; for the proton, $q = e$ and $\kappa = 1.79$; and for the neutron, $q = 0$ and $\kappa = -1.91$. Here $F_{\mu\nu}$ is defined in terms of the vector potential A_μ by

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}, \tag{117}$$

which agrees with Jackson. See also Sec. B.13, and Eqs. (B.56) and (B.57).

(a) Write out the modified Dirac Hamiltonian, and show that it is Hermitian.

(b) Show that probability is conserved, i.e.,

$$\frac{\partial J^\mu}{\partial x^\mu} = 0, \tag{118}$$

where J^μ is defined exactly as for the unmodified Dirac equation, $J^\mu = \bar{\psi}\gamma^\mu\psi$.

(c) Covariance. Suppose $\psi(x)$ satisfies the modified Dirac equation (116), and let

$$\begin{aligned} \psi'(x) &= D(\Lambda)\psi(\Lambda^{-1}x), \\ A'^\mu(x) &= \Lambda^\mu{}_\nu\, A^\nu(\Lambda^{-1}x), \\ F'^{\mu\nu}(x) &= \Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\, F^{\alpha\beta}(\Lambda^{-1}x). \end{aligned} \tag{119}$$

Then show that $\psi'(x)$ satisfies the modified Dirac equation (116), but with Lorentz transformed fields $A'^\mu(x)$ and $F'^{\mu\nu}(x)$ instead of the original fields.

(d) Assume $\mathbf{E} = 0$, $\mathbf{B} \neq 0$ (in order to see what the effective magnetic moment of the particle is). Perform a simple nonrelativistic approximation as in Sec. 46.9, and show that you get the right g-factors for the proton and neutron.